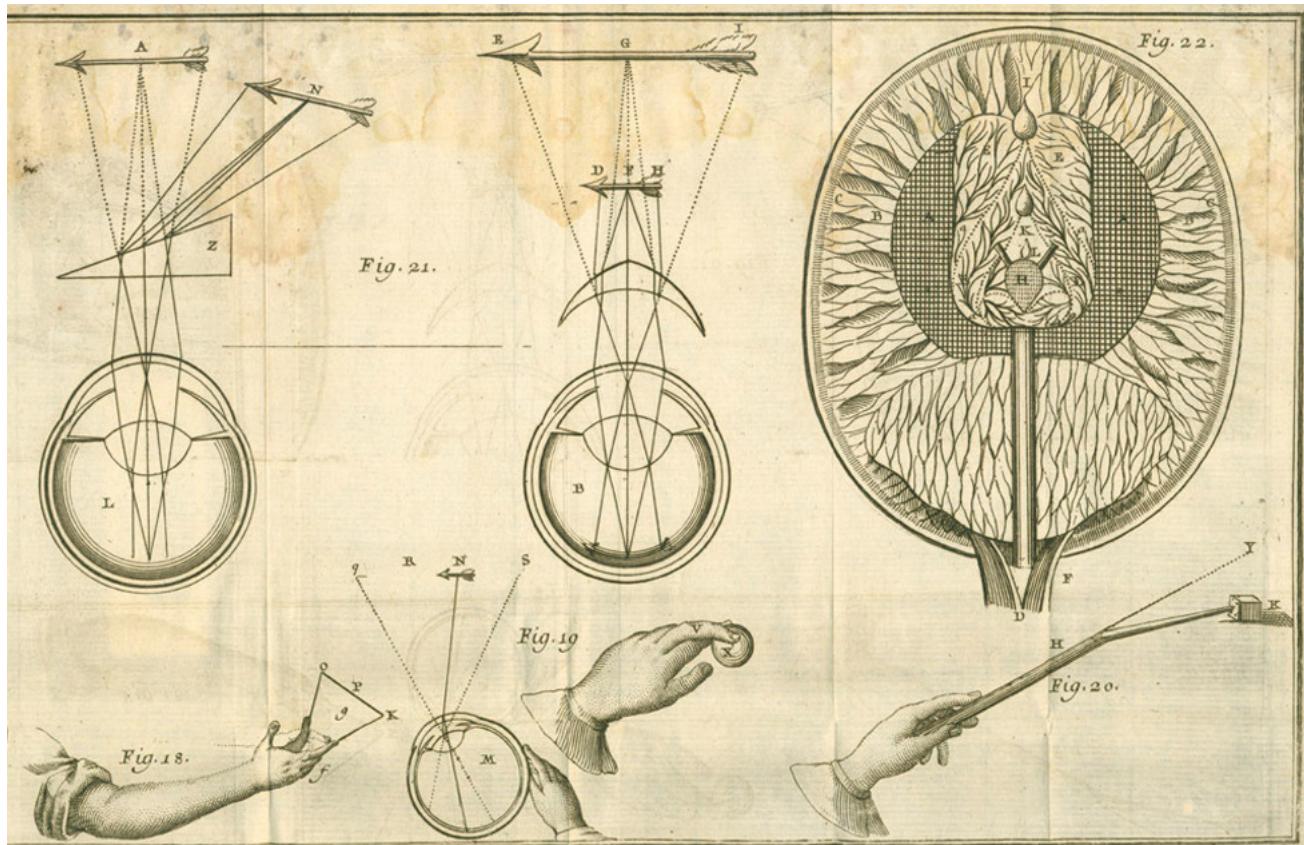


GEOMETRIA Z ALGEBRAŁ LINIOWĄ I

(notatki do wykładu i opowiadania o matematyce)

ARKADIUSZ MĘCEL



Kartezjusz, *La Geometrie*, 1637, źródło: <https://www.gutenberg.org/ebooks/26400>

Spis treści

Notatki nie są skryptem. Czym są?	6
Algebra liniowa z geometrią?	7
1 Układy równań liniowych	9
1.1 Wykład 1	9
1.2 Wybór przykładowych pytań	14
1.3 Zadania do samodzielnej pracy	15
1.4 Uzupełnienie. Wstępne uwagi o geometrii układów równań	16
1.5 Dodatek. Układy nierówności liniowych	19
1.6 Przypomnienie. Algebra zbiorów i rachunek zdań	21
1.7 Przypomnienie. Dowodzenie i metoda indukcji	23
2 Macierze. Operacje elementarne. Rozwiązywanie układów równań	27
2.1 Wykład 2	27
2.2 Wybór przykładowych pytań	32
2.3 Zadania do samodzielnej pracy	33
2.4 Uzupełnienie. Wierszowa równoważność macierzy	34
2.5 Dodatek. Nieliniowe układy równań	36
2.6 Trivia. Kwadraty magiczne	38
2.7 Coda. Eliminacja i podstawienie, czyli historia upraszczania	40
3 Działania i ich własności. Ciała	45
3.1 Wykład 3	45
3.2 Wybór przykładowych pytań	50
3.3 Zadania do samodzielnej pracy	51
3.4 Uzupełnienie. Charakterystyka ciała. Ciała skończone	52
3.5 Przypomnienie. Relacje równoważności. Funkcje.	54
3.6 Dodatek. Ciało ułamków	58
3.7 Trivia. Ciało na paraboli	61
3.8 Trivia. Równania językowe	62
4 Ciało liczb zespolonych	63
4.1 Wykład 4	63
4.2 Wybór przykładowych pytań	68
4.3 Zadania do samodzielnej pracy	69
4.4 Uzupełnienie. Geometria płaszczyzny zespolonej	70
4.5 Dodatek. Ciało liczb p -adycznych	72
4.6 Trivia. Wykres funkcji zespolonej?	76
4.7 Coda. O kształtowaniu się pojęcia liczby	78
5 Wielomiany i ich pierwiastki. Ciała algebraicznie domknięte	83
5.1 Wykład 5	83
5.2 Wybór przykładowych pytań	89
5.3 Zadania do samodzielnej pracy	90
5.4 Uzupełnienie. Elementy teorii podzielności wielomianów	91
5.5 Dodatek. Rozkładanie na czynniki i jego jednoznaczność	94
5.6 Dodatek. Kwaterniony	97

5.7 Coda. Wokół rozkładu na czynniki wielomianów i ich funkcji	99
6 Przestrzenie liniowe	102
6.1 Wykład 6	102
6.2 Wybór przykładowych pytań	108
6.3 Zadania do samodzielnej pracy	109
6.4 Trivia. Kody samokorekcyjne	110
6.5 Coda. O kształtowaniu się pojęcia wektora	112
7 Kombinacje liniowe. Podprzestrzeń rozpięta na układzie wektorów	116
7.1 Wykład 7	116
7.2 Wybór przykładowych pytań	121
7.3 Zadania do samodzielnej pracy	122
7.4 Uzupełnienie. Kombinacje liniowe i układy równań	123
7.5 Dodatek. Ciało jako przestrzeń liniowa nad podciąłem	125
8 Liniowo niezależne układy wektorów	128
8.1 Wykład 8	128
8.2 Wybór przykładowych pytań	133
8.3 Zadania do samodzielnej pracy	134
8.4 Dodatek. Nieprzeliczalne układy. Algebraiczna niezależność	135
8.5 Trivia. Wektory przynależności do klubów	137
8.6 Coda. Kombinacje, czyli o przestrzeni barw	139
9 Baza przestrzeni liniowej	142
9.1 Wykład 9	142
9.2 Wybór przykładowych pytań	147
9.3 Zadania do samodzielnej pracy	148
9.4 Uzupełnienie. Rekurencje liniowe. Wzór Bineta	149
9.5 Dodatek. Wielomiany ograniczonego stopnia i ich bazy	151
9.6 Trivia. Cykle i rozcięcia w grafach	154
10 Wymiar przestrzeni liniowej.	
Rząd macierzy	158
10.1 Wykład 10	158
10.2 Wybór przykładowych pytań	163
10.3 Zadania do samodzielnej pracy	164
10.4 Uzupełnienie. Zadanie o macierzach półmagicznych	165
10.5 Dodatek. Stopień rozszerzenia ciała	167
10.6 Coda. O kształtowaniu się pojęcia wymiaru	171
11 Twierdzenie Kroneckera-Capellego	176
11.1 Wykład 11	176
11.2 Wybór przykładowych pytań	182
11.3 Zadania do samodzielnej pracy	183
11.4 Uzupełnienie. Kilka uwag o prostopadłości	184
11.5 Dodatek. Odpowiedniość Galois i Nullstellensatz Hilberta	185
11.6 Trivia. Zadanie o wymiarze podprzestrzeni macierzy	188
11.7 Trivia. Lights Out	189
11.8 Coda. Bardzo wstępnie o twierdzeniach klasyfikacyjnych	190
12 Operacje na podprzestrzeniach	192
12.1 Wykład 12	192
12.2 Wybór przykładowych pytań	197
12.3 Zadania do samodzielnej pracy	198
12.4 Dodatek. Jednoznaczność daje wyniki o nieistnieniu	199
12.5 Trivia. Krata podprzestrzeni przestrzeni liniowej	202

13 Wstęp do przestrzeni ilorazowych	205
13.1 Wykład 13*	205
13.2 Wybór przykładowych pytań	209
13.3 Zadania do samodzielnej pracy	210
14 Każda przestrzeń liniowa ma bazę	211
14.1 Wykład 14*	211
14.2 Dodatek. Kilka zadań ilustrujących użycie Lematu K-Z	215
14.3 Trivia. Podział prostokąta na kwadraty	217
15 Przekształcenia liniowe	218
15.1 Wykład 15	218
15.2 Wybór przykładowych pytań	223
15.3 Zadania do samodzielnej pracy	224
15.4 Uzupełnienie. Liniowa zamiana współrzędnych	225
15.5 Trivia. Geometria przekształceń liniowych	226
16 Jądro i obraz. Monomorfizm, epimorfizm, izomorfizm	228
16.1 Wykład 16	228
16.2 Wybór przykładowych pytań	233
16.3 Zadania do samodzielnej pracy	234
16.4 Uzupełnienie. Kojądro. Twierdzenie o homomorfizmie.	235
16.5 Dodatek. Różniczkowanie i pierwiastki wielokrotne	236
16.6 Trivia. Układy współrzędnych i układy równań	237
17 Działania na przekształceniach liniowych. Diagramy przekształceń	238
17.1 Wykład 17	238
17.2 Wybór przykładowych pytań	244
17.3 Zadania do samodzielnej pracy	245
17.4 Uzupełnienie. Faktoryzacje i przekształcenia ilorazowe	246
17.5 Dodatek. Ciągi dokładne	248
17.6 Trivia. Być epimorfizmem, być monomorfizmem	251
18 Mnożenie macierzy	
Macierz odwrotna	252
18.1 Wykład 18	252
18.2 Wybór przykładowych pytań	258
18.3 Zadania do samodzielnej pracy	259
19 Macierz przekształcenia liniowego. Zmiana bazy	260
19.1 Wykład 19	260
19.2 Wybór przykładowych pytań	268
19.3 Zadania do samodzielnej pracy	269
20 Macierze odwracalne i izomorfizmy	
Macierze operacji elementarnych	270
20.1 Wykład 20	270
20.2 Wybór przykładowych pytań	276
20.3 Zadania do samodzielnej pracy	277
21 Macierze blokowe	
Równoważność macierzy	278
21.1 Wykład 21	278
21.2 Wybór przykładowych pytań	284
21.3 Zadania do samodzielnej pracy	285
22 Przestrzeń sprzężona	286
22.1 Wykład 22	286
22.2 Wybór przykładowych pytań	291
22.3 Zadania do samodzielnej pracy	292

23 Przekształcenie sprzężone	293
23.1 Wykład 23	293
23.2 Wybór przykładowych pytań	297
23.3 Zadania do samodzielnej pracy	298
24 Wyznacznik — rozwinięcie Laplace'a	299
24.1 Wykład 24	299
24.2 Wybór przykładowych pytań	305
24.3 Zadania do samodzielnej pracy	306
25 Wyznacznik — funkcja objętości	307
25.1 Wykład 25	307
25.2 Wybór przykładowych pytań	315
25.3 Zadania do samodzielnej pracy	316
26 Wyznacznik i układy równań	317
26.1 Wykład 26	317
26.2 Wybór przykładowych pytań	323
26.3 Zadania do samodzielnej pracy	324
27 Wzór permutacyjny. Minory	325
27.1 Wykład 27	325
27.2 Wybór przykładowych pytań	331
27.3 Zadania do samodzielnej pracy	332

Notatki nie są skryptem. Czym są?

Poniższe notatki pochodzą z mojego wykładu z Geometrii z Algebra Liniową na kierunku matematyka prowadzonym przez Wydział Matematyki, Informatyki i Mechaniki Uniwersytetu Warszawskiego. Kurs jest semestralny, złożony z trzydziestu wykładów w trakcie pierwszego semestru, i ma charakter wstępny – omawiane są układy równań liniowych o współczynnikach w ciele, liczby zespolone, przestrzenie i przekształcenia liniowe, funkcjonały i wyznacznik. Wykład kontynuowany jest w drugim semestrze.

Notatki podzielone są na części odpowiadające 30 wykładom. Każda część podzielona jest na moduły, ale nie każda z części złożona jest ze wszystkich modułów.

- **Wykład** – treść zasadnicza „czerpiąca bezpośrednio” (miejscami całe fragmenty) ze skryptu wydziałowego dra T. Koźniewskiego¹, rozszerzająca te treści o intuicje, przykłady i rozmaite zadania.
- **Wybór przykładowych pytań** – nietrudne pytania do autodiagnozy, sprawdzające zrozumienie podstawowych treści wykładu, tworzone wspólnie z dr Olgą Ziemiańską w roku 2024.
- **Zadania do samodzielnej pracy** – zestaw zadań ilustrujących umiejętności wymagane zarówno do zaliczenia przedmiotu na ocenę dostateczną (oznaczonych ♠), jak i do uzyskania biegłości w prowadzeniu bardziej złożonych rozumowań, koniecznej do uzyskania oceny bardzo dobrzej.
- **Uzupełnienie** – zawiera treści ściśle związane z tematem – takie, które chętnie dodałbym do wykładu, gdybym tylko miał czas. Poziomem trudności treści te nie odbiegają od wykładu. Często omawiane są przykłady, na które w bardzo skondensowanym harmonogramie nie ma miejsca.
- **Dodatek** – poruszam tu tematy niekiedy bardziej zaawansowane, o charakterze „gwiazdkowym”, odnoszące się do ważnych motywów pojawiających się na konkretnym wykładzie. Zagadnienia te mają charakter algebraiczny, ale nie zawsze czysto algebraliniowy (a na przykład: kombinatoryczny).
- **Przypomnienie** – uzupełnienie treści pojęciowych, czy to szkolnych, czy korzystających z innych kursów (zwłaszcza ze wstępu do matematyki).
- **Trivia** – tematy luźno związane z wykładem, ale moim zdaniem ładne i mające charakter popularyzatorski. Czasem mogą nawet dotyczyć zastosowań. Oglądanie takich rzeczy pomaga niektórym podtrzymać wiarę w to, że studiowanie „czystej” matematyki ma głęboki estetyczny sens.
- **Coda** – fragment ten wskazываć będzie szerszy kontekst pojęciowy i historyczny pojęć pojawiających się na wykładzie. Na wykładzie sygnalizujemy wiele fundamentalnych koncepcji, które poznawać będą Państwo przez całe studia. Będę się starał krótko wskazywać na osoby i wydarzenia związane z rozwojem danego pojęcia.

Ważne zastrzeżenie – **notatki (zwłaszcza te notatki) to nie jest skrypt**, czyli tekst realizujący obowiązujący sylabus. Skrypt zawiera to, co kluczowe dla uczestnika kursu – treści, z których będzie egzaminowany. Poniższe notatki zawierają także to, co mi w uczeniu tego kursu pomaga i co mi się w nim podoba. Obok ścisłych rozumowań ważne są dla mnie intuicje i motywacje. Ten tekst próbuje je zebrać i wciąż się zmienia – gdy coś mi nie wychodzi, gdy próbuję coś poprawić, gdy dostaję uwagi, gdy czytam innych autorów, gdy coś nowego mi się spodoba. Chciałbym, aby to był żywy, „mówiony” tekst. Zapraszam Czytelnika do krytycznej lektury oraz do wskazywania mi znalezionych błędów. Dziękuję za już wskazane i przepraszać, że pozwalam sobie nie wymieniać wszystkich osób, które je znajdowały.

Arkadiusz Męcel

¹T. Koźniewski, *Wykłady z algebry liniowej I*, Uniwersytet Warszawski, Warszawa 2008.

Algebra liniowa z geometrią?

Wykład GAL I poświęcony jest rozwiązywaniu układów równań liniowych nad ciałem. Omówiona będzie definicja ciała i jego podstawowe własności. Skupimy się na badaniu ciała liczb rzeczywistych i zespolonych. Pokażemy jak opisywać zbiory rozwiązań wprowadzając na nich strukturę przestrzeni liniowej. Podstawowym narzędziem będą macierze opisujące zarówno układy równań jak i przekształcenia liniowe.

Powyższe streszczenie wykładu zaczerpnięte z wydziałowego sylabusa należy uzupełnić bardziej poglądowym wyjaśnieniem nawiązującym do nazwy przedmiotu, który zamierzamy studiować. Co rozumieć będziemy przez algebrę, a co przez geometrię? Co oznacza pojęcie algebry „liniowej”? Czy odnosić się będziemy do treści szkolnych? Na czym polegać będzie nowość ujęcia, które zaprezentujemy? W jaki sposób przedmiot ów wprowadzi nas w studiowanie matematyki? Pełne odpowiedzi na te pytania przekraczają rama tego wprowadzenia. Możemy jednak poczynić kilka pozytycznych, mniejmy nadzieję, uwag.

Z perspektywy absolwenta szkoły przedmiot nasz wydawać się może uogólnieniem geometrii analitycznej (pod podobną nazwą wykładowany był zresztą – przy nieco innym rozkładzie akcentów – w latach powojennych). Będziemy mówić o obiektach geometrycznych opisywanych przy pomocy warunków algebraicznych. Będą to początkowo rozwiązania układów równań liniowych. Weźmy, nawiązując do wiedzy szkolnej, problem określenia wzajemnego położenia dwóch prostych leżących na płaszczyźnie kartezjańskiej, na podstawie opisujących ich równań (np. w postaci ogólnej). Trzy możliwe konfiguracje to:

- para prostych przecinających się w dokładnie jednym punkcie,
- para prostych równoległych,
- para prostych tożsamych (ta sama prosta może być opisana przez różne równania!).

Sytujom tym odpowiada interpretacja algebraiczna:

- układ równań ułożony z równań opisujących poszczególne proste ma dokładnie jedno rozwiązanie,
- układ równań ułożony z równań opisujących poszczególne proste nie ma rozwiązań,
- układ równań ułożony z równań opisujących poszczególne proste ma nieskończenie wiele rozwiązań.

Podczas zajęć szybko zrozumiemy, że przedstawiona wyżej odpowiedniość jest w istocie swego rodzaju utożsamieniem rzeczywistości algebraicznej i geometrycznej. Wspólną płaszczyzną pojęciową dla tego utożsamienia będzie pojęcie przestrzeni liniowej. Samo zaś utożsamienie dokonane zostanie przy pomocy aparatu pojęciowego związanego z operacjami na przestrzeniach liniowych. Co więcej, przestrzenie liniowe są pojęciem tak ogólnym i wygodnym w użyciu, że pozwolą na dostrzeganie innych, mniej oczywistych powiązań pomiędzy różnymi działami matematyki. Ciekawe co Czytelnik powiedziałby słysząc, że w odpowiednio dobranym kontekście funkcja sinus może być prostopadła do funkcji cosinus (a nawet i do samej siebie!) albo, że liniowo niezależne – cokolwiek to znaczy – są przy odpowiednich definicjach dwa podzbiory jego znajomych. Takie algebraiczne konteksty pojawią się np. w analizie i kombinatoryce.

Innymi słowy, chodzić nam będzie nie tylko o rozwiązywanie układów równań, ale o umiejętność wy-abstrahowania „struktury” tych rozwiązań i dostrzegania jej w rozmaitych matematycznych sytuacjach. Tym, co charakteryzuje naukowe podejście do uprawiania matematyki jest w tym kontekście nie tylko dążenie do znalezienia ogólnej struktury, ale też próba opisu „wszystkich możliwych struktur” z dokładnością do pewnych warunków. Przykład tego rodzaju opisu widzimy wyżej: układ dwóch prostych na płaszczyźnie prowadzić może tylko do trzech na swój sposób istotnie różnych konfiguracji. W języku algebry liniowej niezmiennikiem takich „istotnie różnych” konfiguracji będzie m.in. pojęcie wymiaru.

Podejście, o którym tu piszemy nazwać można podejściem klasycznym. Badamy grupę obiektów i próbujemy znaleźć cechy charakterystyczne, które by je odróżniały (na przykład odróżniliśmy rodziny par prostych na płaszczyźnie poprzez opis możliwej liczby ich punktów przecięcia). W kontekście geometrii analitycznej ojcem tego podejścia był R. Descartes (Kartezjusz), stąd też często (w szkole?) mówimy o kartezjańskim układzie współrzędnych. Podejście to bardzo rozwinęło badania geometryczne (do tej pory rozumiane głównie w stylu starożytnych „Elementów” Euklidesa – który to znany jest absolwentowi szkoły i nazywany jest mianem geometrii elementarnej) najpierw poprzez odkrycie rachunku różniczkowego i całkowego przez Newtona i Leibniza, które doprowadziło do powstania geometrii różniczkowej, potem przez prace dotyczące współrzędnych jednorodnych (tzw. współrzędne barycentryczne) poczynione na początku XIX wieku przez Möbiusa i Plückera, następnie przez rozwój geometrii rzutowej dzięki pracom Monge'a i Ponceleta, i wreszcie przez prace Łobaczewskiego i Bolyai, którzy „odkryli” geometrie nieeuklidesowe. Narodziny dziedziny, którą będziemy się zajmować przypisuje się różnym uczonym, zwłaszcza jednak nauczycielowi gimnazjalnemu H. Grassmannowi (prace od roku 1844). Był to czas rozwoju podejścia aksjomatycznego do geometrii, które w tej i innych dziedzinach zrewolucjonizowało w XX wieku Królową Nauk. Jest ono podstawą wykładu uniwersyteckiego matematyki i współczesnych badań.

Warto pamiętać, że przy całej abstrakcji jaka towarzyszy wykładowi geometrii z algebrą liniową, mowa jest wciąż o geometrii, a więc w najszerzym sensie w jakim rozumiemy to pojęcie – nauce o przestrzeni i jej własnościach. Zarówno wysiłki Kartezjusza, jak i Ponceleta, Łobaczewskiego, Riemanna, Kleina, Hilberta, Minkowskiego czy wielkich geometrów XX wieku miały w sobie zawsze silne pragnienie zrozumienia świata, w którym żyjemy i budowania przekonujących modeli matematycznych dla opisu jego działania. Zetkniamiemy się z tym trudnym zjawiskiem w drugim semestrze mówiąc o przestrzeniach afiniczych.

Prawdziwą rewolucją w myśleniu (bardzo charakterystyczną dla współczesnej matematyki) jest próba opisu obiektów nie w języku ich wewnętrznej struktury, ale w języku operacji, które na nich wykonujemy. Dla przykładu: kwadrat i okrąg odróżnić można w sposób „klasyczny” mówiąc dość nieprecyzyjnie (jakkolwiek byśmy nie lubili postulatów Euklidesa) o kątach czy wierzchołkach, lub bardziej formalnie: podając równania je opisujące (i tłumacząc kiedy równania opisują obiekty różne – i co to znaczy), ale można je także odróżnić bardzo elegancko poprzez obserwację, że tylko skończenie wiele izometrii płaszczyzny (na przykład obrotów) w siebie pozostawia kwadrat w miejscu, podczas gdy okrąg „jest niezmienny” (znowu ważne pojęcie) przy działaniu nieskończego wielu różnych izometrii. To nowe podejście do matematyki polegające na badaniu obiektów przez „niezmienniki operacji” pochodzi, w kontekście geometrii układów równań i ich przekształceń od matematyków angielskich A. Cayleya, J.J. Sylvestera i G. Salmona. Do ustalenia głębszych związków pomiędzy teorią przekształceń a geometrią przyczynił się tzw. program z Erlangen słynnego matematyka niemieckiego Felixa Kleina sformułowany w 1872 roku.

Nasze podejście pójdzie jednak jeszcze dalej. W powojennej matematyce rewolucji dokonały prace warszawskiego matematyka Samuela Eilenberga i cały ruch nazwany później „New Math” reprezentowany przez środowisko francuskich badaczy, zwanych bourbakistami, którzy zaproponowali jeszcze ogólniejsze pojmowanie matematyki w języku diagramów. Teoria ta, zwana teorią kategorii, jest obecnie językiem urzędowym całej algebraiczno–geometrycznej strony matematyki. Jednym z zaskakujących przejawów tego podejścia, które zobaczymy na naszym przedmiocie jedynie w załączku, jest istnienie naturalnych odpowiedniości, które zamazują rozróżnienie pomiędzy obiektami, a działającymi na nich przekształceniami. W języku matematycznym odpowiedniości takie nazywamy dualnościami. Dla przykładu: obiektem geometrycznym przypiszemy ich przekształcenia, i odwrotnie. Na jednych i drugich wprowadzimy te same operacje. Okaże się, że matematycznie nie ma sensu ich rozróżnianie! W dostrzeżeniu tych dualności podstawowym narzędziem będą niepozorne na pierwszy rzut oka tablice liczbowe zwane macierzami.

Podsumowując ten przydługi wstęp: droga od matematyki szkolnej, zajmującej się głównie rozwiązywaniem konkretnych zadań dotyczących struktury danych obiektów (np. wzajemne położenie dwóch konkretnych prostych na płaszczyźnie) do współczesnej geometrii algebraicznej czy innych dziedzin matematyki prowadzi przez trzy progi: klasyfikowanie „wszystkich konfiguracji” obiektów w języku opartym na współrzędnych, klasyfikowanie obiektów będących „niezmiennikami operacji” bez użycia współrzędnych, ustalanie „dualności” pomiędzy obiektami i operacjami na nich, także w obrębie różnych działów matematyki (ważne hasło na przyszłość dla aspirującego matematyka: odpowiedniość Galois). Życzę Czytelnikowi udanego wejścia w piękną przygodę z matematyczną abstrakcją dzięki algebrze liniowej z geometrią.

Rozdział 1

Układy równań liniowych

1.1 Wykład 1

Wykład z geometrii z algebrą liniową zaczynamy omówieniem teorii układów równań liniowych o współczynnikach w ciele. Czym jest ciało dowiemy się na wykładzie 3. Na początku spróbujemy natomiast uporządkować i usystematyzować język dotyczący samych układów równań i zbiorów ich rozwiązań. Zbiór liczb rzeczywistych oznaczamy przez \mathbb{R} , zaś zbiór liczb wymiernych – przez \mathbb{Q} .

Definicja 1.1.1: Równanie liniowe

RÓWNANIE LINIOWE o zmiennych x_1, \dots, x_n i o WSPÓŁCZYNNIKACH w zbiorze K , to wyrażenie postaci

$$a_1x_1 + a_2x_2 + \dots + a_nx_n = b, \text{ gdzie } a_1, a_2, \dots, a_n, b \in K.$$

ROZWIĄZANIE powyższego równania to ciąg (s_1, s_2, \dots, s_n) , gdzie $s_1, s_2, \dots, s_n \in K$, taki że

$$a_1 \cdot s_1 + a_2 \cdot s_2 + \dots + a_n \cdot s_n = b.$$

Przykłady

- Dla $K = \mathbb{R}$, równaniem liniowym zmiennej x jest $\sqrt{2}x = \pi$. Natomiast definicji powyższej nie spełniają równania $2x + 1 = 3$ lub $2x = x$ (choć można je przekształcić równoważnie do równań liniowych).
- Dla $K = \mathbb{Q}$, równaniem liniowym o zmiennych x_1, x_2, x_3 jest: $\frac{1}{2}x_1 + x_2 + 0 \cdot x_3 = 0$, ale nie jest nim równanie $\sqrt{2}x_1 = 2$.

Aby powyższa definicja miała sens, przyjmujemy, że K to \mathbb{R} lub \mathbb{Q} . Przyjmujemy także **konwencję**: zapisując równanie liniowe pomijamy składniki, w których zmienna przemnożona jest przez 0, np. równanie $2x_1 + 0 \cdot x_2 + x_3 = 0$ zapisujemy w postaci $2x_1 + x_3 = 0$, a równanie $0x_1 + 0x_2 = 1$ w postaci: $0 = 1$.

Definicja 1.1.2: Układ równań liniowych

UKŁAD m RÓWNAŃ LINIOWYCH o zmiennych x_1, \dots, x_n i o współczynnikach rzeczywistych, to ciąg m równań liniowych o współczynnikach rzeczywistych postaci:

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + \dots + a_{2n}x_n = b_2 \\ \dots \\ a_{m1}x_1 + a_{m2}x_2 + \dots + a_{mn}x_n = b_m. \end{cases} \quad (\heartsuit)$$

Rozwiązanie powyższego układu równań liniowych to ciąg elementów (s_1, \dots, s_n) ze zbioru \mathbb{R} , który jest rozwiązaniem każdego z m równań liniowych tego układu.

Ważne jest zwrócenie uwagi na duży stopień dowolności, jaki daje ta definicja. Nic nie stoi na przeszkodzie by układ składał się z pojedynczego równania, albo zawierał identyczne równania. Oto kilka przykładów:

$$\begin{cases} 2x_1 + 3x_2 + x_3 = 1 \\ -\frac{1}{2}x_2 + x_3 = 0 \end{cases}, \quad \begin{cases} x_1 + x_2 = 1 \\ x_1 + x_2 = 1 \\ 0x_1 + 0x_2 = 3 \end{cases}, \quad \begin{cases} 0x_1 + 0x_2 + 0x_3 = 0 \\ 0x_1 + 0x_2 + x_3 = 0 \\ x_1 + x_2 + x_3 = 0 \end{cases}.$$

Przykładem rozwiązania pierwszego z powyższych układów jest trójką $(-3, 2, 1)$. Drugi układ nie ma rozwiązań, ponieważ trzecie jego równanie nie ma rozwiązań. Trzeci układ ma choćby rozwiązanie $(0, 0, 0)$. Można sprawdzić, że dowolna trójką $(-s, s, 0)$, gdzie $s \in \mathbb{R}$, jest rozwiązaniem trzeciego układu.

Warto w tym miejscu poczynić uwagę terminologiczną o charakterze formalnym.

Definicja 1.1.3: Układ

Niech T, X będą dowolnymi zbiorami. Wówczas UKŁADEM ELEMENTÓW zbioru X o wskaźnikach przebiegających zbiór T będziemy nazywali dowolną funkcję $x : T \rightarrow X$ określona na zbiorze T o wartościach w zbiorze X . Wartość tej funkcji w punkcie $t \in T$ nazwiemy ELEMENTEM tego układu o WSKAŹNIKU t , ozn. x_t .

Definicja powyższa jest formalnym wyrażeniem myśli zawartej wyżej. Gdy rozważamy układ m równań, wówczas zbiór T równy jest $\{1, 2, \dots, m\}$, a X jest zbiorem m równań liniowych o n zmiennych. Wprowadzenie funkcji $x : T \rightarrow M$ daje nam to, że możemy mówić o pierwszym, drugim, trzecim, ..., m -tym równaniu tego układu. A to, że poszczególne równania mogą być zerowe, identyczne, bez rozwiązań — to jest inna kwestia. Na etapie konstruowania układu równań kwestie te nie mają znaczenia.

Definicja 1.1.4: Ważne typy układów równań liniowych

Układ równań liniowych postaci (\heartsuit) nazywamy:

- JEDNORODNYM, jeśli $b_i = 0$, dla każdego $i = 1, 2, \dots, m$,
- NIESPRZECZNYM, jeśli zbiór rozwiązań układu (\heartsuit) jest niepusty,
- SPRZECZNYM, jeśli zbiór rozwiązań układu (\heartsuit) jest pusty.

Układ jednorodny powstaje z układu (\heartsuit) przez przyjęcie $b_i = 0$, dla każdego $i = 1, 2, \dots, m$, nazywamy UKŁADEM JEDNORODNYM ODPOWIADAJĄCYM układowi (\heartsuit) .

Wśród powyższych trzech przykładów, jednorodny jest trzeci układ. Niesprzeczne są — pierwszy i trzeci, a sprzecznym jest drugi z wymienionych układów. Układem jednorodnym odpowiadającym pierwszemu z powyższych trzech układów jest

$$\begin{cases} 2x_1 + 3x_2 + x_3 = 0 \\ -\frac{1}{2}x_2 + x_3 = 0 \end{cases}.$$

Definicja 1.1.5: Układy równoważne

Układy równań liniowych o tym samym zbiorze zmiennych i o tym samym zbiorze współczynników nazywamy RÓWNOWAŻNYMI, jeśli mają one identyczne zbiory rozwiązań.

Przykład. Poniższe układy równań liniowych o zmiennych x_1, x_2 o współczynnikach w \mathbb{R} są równoważne:

$$\begin{cases} x_1 + x_2 = 0 \\ x_1 - x_2 = 2 \end{cases}, \quad \begin{cases} x_1 = 1 \\ x_2 = -1 \end{cases}.$$

Naszym celem jest opis rozwiązań układów równań liniowych o współczynnikach rzeczywistych. Metoda polega na zastępowaniu układu innym — równoważnym, i z jakiegoś powodu prostszym do rozwiązania.

Definicja 1.1.6: Operacje na równaniach liniowych

Dane są równania liniowe U, U' o zmiennych x_1, \dots, x_n i o współczynnikach w \mathbb{R} :

$$U : a_1x_1 + a_2x_2 + \dots + a_nx_n = b, \quad U' : a'_1x_1 + a'_2x_2 + \dots + a'_nx_n = b'.$$

Określamy równanie liniowe $U + U'$ oraz dla każdego $\lambda \in \mathbb{R}$ określamy równanie liniowe λU :

$$U + U' : (a_1 + a'_1)x_1 + (a_2 + a'_2)x_2 + \dots + (a_n + a'_n)x_n = b + b',$$
$$\lambda U : \lambda \cdot a_1x_1 + \lambda \cdot a_2x_2 + \dots + \lambda \cdot a_nx_n = \lambda \cdot b.$$

Przykład. Jeśli U jest równaniem $x_1 + x_2 - x_3 = 2$ oraz U' jest równaniem $3x_1 - x_2 + x_3 = 0$, to równanie $U + U'$ ma postać $4x_1 = 2$, a równanie $2U$ ma postać $2x_1 + 2x_2 - 2x_3 = 4$.

Uwaga 1.1.7

Jeśli (s_1, \dots, s_n) jest rozwiązaniem obydwu równań liniowych U_1 oraz U_2 , gdzie $s_1, \dots, s_n \in \mathbb{R}$, to jest też rozwiązaniem każdego równania liniowego postaci $\lambda_1U_1 + \lambda_2U_2$, gdzie $\lambda_1, \lambda_2 \in \mathbb{R}$.

Dowód. Niech $U_1 : a_1x_1 + \dots + a_nx_n = b$ oraz $U_2 : a'_1x_1 + \dots + a'_nx_n = b'$. Mamy wówczas

$$(\lambda_1a_1 + \lambda_2a'_1)s_1 + \dots + (\lambda_1a_n + \lambda_2a'_n)s_n = \lambda_1b + \lambda_2b'.$$

W szczególności (s_1, \dots, s_n) jest rozwiązaniem równania $\lambda_1U_1 + \lambda_2U_2$. \square

Twierdzenie 1.1.8

Poniższe operacje zamieniają układ równań liniowych U o współczynnikach rzeczywistych w układ równoważny:

- (1) dodanie do równania innego równania pomnożonego przez liczbę rzeczywistą,
- (2) zamiana dwóch równań miejscami,
- (3) pomnożenie równania przez liczbę rzeczywistą różną od zera.

Dowód. Założymy, że układ U ma m równań. Przez U_i oznaczamy i -te równanie układu U , dla $1 \leq i \leq m$.

Zacznijmy od dowodu dla operacji (3). Niech U' powstaje z U przez przemnożenie równania U_i przez $\lambda \neq 0$. Twierdzimy, że każde rozwiązanie (s_1, \dots, s_n) układu U jest też rozwiązaniem układu U' . Rzeczywiście, j -te równanie U' jest dla $j \neq i$, j -tym równaniem układu U , więc (s_1, \dots, s_n) jest jego rozwiązaniem. Natomiast i -te równanie U' to równanie λU_i . Zgodnie z Uwagą 1.1.7, skoro (s_1, \dots, s_n) jest rozwiązaniem U_i , to jest też rozwiązaniem λU_i . A zatem (s_1, \dots, s_n) jest rozwiązaniem U' .

Z drugiej strony, każde rozwiązanie (r_1, \dots, r_n) układu U' jest rozwiązaniem układu U . Rzeczywiście, za wyjątkiem i -tego równania, wszystkie równania U to równania U' , więc (r_1, \dots, r_n) jest ich rozwiązaniem. Natomiast i -te równanie układu U powstaje z i -tego równania układu U' przez przemnożenie go przez λ^{-1} . A zatem zgodnie z Uwagą 1.1.7 (r_1, \dots, r_n) jest rozwiązaniem tego równania. Ostatecznie więc ciąg (r_1, \dots, r_n) jest rozwiązaniem każdego równania układu U , czyli jest rozwiązaniem układu U . Pokazaliśmy, że każde rozwiązanie U jest rozwiązaniem U' i odwrotnie – każde rozwiązanie U' jest rozwiązaniem U . A zatem układy te są równoważne.

Wykazaliśmy, że operacja (3) przeprowadza układ równań liniowych w układ równoważny. Teza dla operacji (2) jest oczywista. Pozostała analiza operacji (1). Niech układ U' powstaje z U przez dodanie do i -tego równania U_i równania U_j przemnożonego przez $a \in \mathbb{R}$. Wtedy U powstaje z U' przez dodanie do i -tego równania $U_i + aU_j$ równania U_j przemnożonego przez $-a \in \mathbb{R}$. Na mocy Uwagi 1.1.7 zbiory rozwiązań układów U oraz U' są zatem identyczne. \square

Definicja 1.1.9: Operacje elementarne na układzie równań liniowych

Operacje (1)-(3) nazywać będziemy OPERACJAMI ELEMENTARNYMI NA UKŁADZIE U .

Do jakich postaci chcemy dochodzić w ramach takiego „upraszczania” przy pomocy operacji elementarnych? Do takich, gdzie jedne zmienne można wyrazić za pomocą innych. Tak jest w przypadku układu po prawej stronie — równoważnego, jak się okazuje, układowi po lewej.

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

Oto operacje elementarne przeprowadzające pierwszy układ w drugi: dodanie drugiego równania do pierwszego równania, odjęcie od trzeciego równania drugiego równania, przemnożenie pierwszego równania przez $\frac{1}{2}$, odjęcie od drugiego równania pierwszego równania, przemnożenie drugiego równania przez -1 :

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

Zbiór rozwiązań powyższych układów składa się z ciągów postaci $(\frac{1}{2}, -s_3 - s_4 + \frac{1}{2}, s_3, s_4)$, gdzie s_3, s_4 są dowolnymi liczbami rzeczywistymi. Bez trudu odczytaliśmy to rozwiązanie z układu uzyskanego poprzez operacje elementarne. Wyrziliśmy po prostu zmienne x_1, x_2 za pomocą stałych oraz zmiennych x_3, x_4 .

Definicja 1.1.10: Rozwiążanie ogólne układu równań liniowych

Niech U oraz U' będą równoważnymi układami równań liniowych o n zmiennych (niewiadomych). Przypuśćmy, że układ U' MOŻNA PRZEPISAĆ w postaci:

$$\begin{cases} \textcolor{red}{x_{j_1}} = c_{11}x_1 + c_{12}x_2 + \dots + c_{1n}x_n + d_1 \\ \vdots \\ \textcolor{red}{x_{j_k}} = c_{k1}x_1 + c_{k2}x_2 + \dots + c_{kn}x_n + d_k \end{cases}$$

przy czym $1 \leq j_1 < \dots < j_k \leq n$ oraz zmienne $\textcolor{red}{x_{j_1}}, \dots, \textcolor{red}{x_{j_k}}$ nie występują po prawej stronie (to znaczy: stoją przy nich współczynniki zerowe, czyli $c_{ij} = 0$, dla $j = j_1, \dots, j_k$). Mówimy wtedy, że:

- układ U' JEST ROZWIĄZANIEM OGÓLNYM (zadaje rozwiązanie ogólne) układu U ,
- w rozwiązaniu tym zmienne $\textcolor{red}{x_{j_1}}, \dots, \textcolor{red}{x_{j_k}}$ nazywamy ZMIENNYMI ZALEŻNYMI,
- pozostałe zmienne x_i , dla $i \neq j_s$, nazywamy ZMIENNYMI NIEZALEŻNYMI, albo PARAMETRAMI.

Potrzebna jest chwila uwagi, by zrozumieć powyższą definicję. Zdefiniowaliśmy te postaci układów równań, z których łatwo jest odczytywać rozwiązania, a które uzyskiwać chcemy z dowolnych układów poprzez stosowanie operacji elementarnych. Oto przykłady układów, będących rozwiązaniami ogólnymi.

$$\begin{cases} \textcolor{red}{x_1} &= 2 \\ \textcolor{red}{x_2} &= 1 \\ \textcolor{red}{x_3} &= 4 \end{cases}, \quad \begin{cases} \textcolor{red}{x_1} &= \frac{1}{2} \\ \textcolor{red}{x_2} + x_3 + x_4 &= \frac{1}{2} \end{cases}, \quad \begin{cases} \textcolor{red}{x_1} &+ x_3 &- 2x_5 &= 0 \\ \textcolor{red}{x_2} &&+ x_5 &= 0 \\ \textcolor{red}{x_4} &+ x_5 & &= 0 \end{cases}.$$

Przykład. Rozważmy układ równań liniowych:

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

To nie przypadek, że dwa razy wpisane jest to samo równanie. Przyzwyczaimy się do tego po wprowadzeniu pojęcia macierzy i stanie się to naturalne. Układ ten ma rozwiązanie ogólne. Jest nim na przykład:

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

A zatem jak przedstawić zbiór wszystkich rozwiązań? Jest to zbiór ciągów postaci

$$\left\{ \left(\frac{1}{2}, -x_3 - x_4 + \frac{1}{2}, x_3, x_4 \right), \text{ gdzie } x_3, x_4 \in \mathbb{R} \right\}.$$

Na marginesie istnieje także inny sposób zapisania rozwiązania ogólnego:

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

Ma ono zatem postać $\left\{ \left(\frac{1}{2}, x_2, x_3, -x_2 - x_3 + \frac{1}{2} \right), \text{ gdzie } x_2, x_3 \in \mathbb{R} \right\}$.

Mamy zatem ilustrację definicji. W pierwszym rozwiązaniu ogólnym zmiennymi zależnymi są x_1, x_2 (w zasadzie x_1 jest tutaj akurat stałą), parametrami zaś x_3, x_4 . W drugim rozwiązaniu ogólnym jest nieco inaczej. Parametrami są x_2, x_3 . A zatem formalnie rzecz biorąc jeden układ równań posiadać może kilka ogólnych postaci rozwiązań, co nie wydaje się do końca wygodne. Intuicja podpowiada jednak, że druga z powyższych konfiguracji nie jest „elegancka”, ponieważ zmienna zależna ma większy indeks niż zmienna niezależna. W dalszej części poradzimy sobie z tą drobną delikatnością. Wykorzystamy w tym celu język macierzy, pozwalający na zalgorytmizowanie procesu rozwiązywania układów równań metodą doprowadzania do układu równoważnego, będącego w postaci ogólnej.

Celem kolejnego wykładu będzie między innymi dowód następującego rezultatu.

Twierdzenie 1.1.11

Niech U będzie układem m równań liniowych o n zmiennych i współczynnikach w \mathbb{R} . Wówczas zachodzi następująca alternatywa wyłączająca:

- albo istnieje układ równań liniowych U' , który można uzyskać z U za pomocą skończonego ciągu operacji elementarnych na (kolejno otrzymywanych) układach równań, i który jest rozwiązaniem ogólnym układu U ,
- albo istnieje układ równań liniowych U'' , który można uzyskać z U za pomocą skończonego ciągu operacji elementarnych na (kolejno otrzymywanych) układach równań, i który zawiera równanie sprzeczne $0 = 1$.

1.2 Wybór przykładowych pytań

Proszę się nie uczyć na pamięć odpowiedzi czy wyjaśnienie! Takich pytań można ułożyć znacznie więcej i służą one jedynie sprawdzeniu podstawowego zrozumienia treści zajęć. Jeśli odpowiedź na dowolne z poniższych pytań sprawia Państwu problem, proszę pomyśleć który fragment notatek dotyczy tego pytania, wrócić do ich lektury, a w razie dalszych trudności poprosić prowadzącego o wskazówkę.

1. Sformułuj jednorodny układ równań liniowych odpowiadający układowi

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

2. Przypuśćmy, że ciąg $(1, 1, 1)$ jest rozwiązaniem jednorodnego układu równań liniowych. Czy dla dowolnego $s \in \mathbb{R}$ ciąg (s, s, s) jest rozwiązaniem tego układu?

3. Czy jednorodny układ równań liniowych może być układem sprzecznym?

4. Czy jednorodny układ równań liniowych może być równoważny z układem równań liniowych, który nie jest jednorodny?

5. Czy każde dwa sprzeczne układy równań liniowych są równoważne?

6. Układ U składa się z jednego równania. Czy układ ten może być sprzeczny?

7. Niech $(x_1, x_2, \dots, x_n), (x'_1, x'_2, \dots, x'_n)$ będą rozwiązaniami układu równań liniowych U . Czy ciąg

$$(x_1 + x'_1, x_2 + x'_2, \dots, x_n + x'_n)$$

jest również rozwiązaniem układu U ?

8. Niech (x_1, x_2, \dots, x_n) będzie rozwiązaniem układu równań liniowych U , zaś $(x'_1, x'_2, \dots, x'_n)$ — niech będzie rozwiązaniem układu równań liniowych U' . Czy ciąg

$$(x_1 + x'_1, x_2 + x'_2, \dots, x_n + x'_n)$$

jest rozwiązaniem układu składającego się z równań U i U' ?

9. Niech X będzie zbiorem rozwiązań układu m równań liniowych o współczynnikach rzeczywistych, gdzie $m \geq 1$. Niech Y będzie zbiorem rozwiązań układu, który powstaje z powyższego przez usunięcie pierwszego równania. Czy jest możliwe, aby $X \supseteq Y$?

10. Czy istnieje układ równań liniowych o trzech zmiennych x_1, x_2, x_3 , którego rozwiązaniem jest każdy ciąg (s_1, s_2, s_3) , gdzie $s_1, s_2, s_3 \in \mathbb{R}$?

11. Na układzie równań liniowych U_1 wykonano dwie operacje elementarne: najpierw do wiersza trzeciego dodano trzykrotność wiersza pierwszego, a następnie pomnożono pierwszy wiersz przez 3. Uzyskano układ równań układ liniowych U_2 . Jakie operacje elementarne należy wykonać na układzie U_2 , aby uzyskać z powrotem układ U_1 ?

12. Ile jest parametrów potrzebnych do opisu zbioru rozwiązań układu równań liniowych o 3 zmiennych złożonego z równania $x_1 + 2x_2 + 3x_3 = 0$?

13. Układ U złożony jest z dwóch równań liniowych z trzema zmiennymi o współczynnikach rzeczywistych. Czy układ ten może mieć dokładnie jedno rozwiązanie?

14. Czy rozwiązania ogólne dwóch niesprzecznych równoważnych układów równań liniowych mogą mieć różną liczbę parametrów?

15. Ile parametrów ma rozwiązanie ogólne układu m równań liniowych o n zmiennych, który ma dokładnie jedno rozwiązanie?

16. Założymy, że niejednorodny układ U składający się z 3 równań liniowych o współczynnikach rzeczywistych od 4 zmiennych ma dwa różne rozwiązania (a, b, c, d) oraz (e, f, g, h) . Rozstrzygnij, którzy z poniższych ciągów jest również rozwiązaniem układu U

- a) $(a - e, b - f, c - g, d - h)$,
- b) $(a + e, b + f, c + g, d + h)$,
- c) $(2a - e, 2b - f, 2c - g, 2d - h)$.

1.3 Zadania do samodzielnej pracy

Zadania ilustrujące umiejętności niezbędne do zaliczenia przedmiotu oznaczone są symbolem ♠.

Przy tych zadaniach dodany jest opis umiejętności, które sprawdzają.

1. (♠) Rozpoznawanie rozwiązań układu równań liniowych)

Które z poniższych ciągów $(-1, 1, 1, -1), (2, 3, 1, 4), (4, -3, 2, 1), (4, 0, -3, \frac{1}{2})$ są rozwiązaniami poniższego układu równań liniowych o współczynnikach rzeczywistych?

$$\begin{cases} 3x_1 + 2x_2 + 4x_3 + 2x_4 = 1 \\ 7x_1 + 5x_2 + 9x_3 + 4x_4 = 3 \\ 5x_1 - 3x_2 + 7x_3 + 4x_4 = 1 \end{cases}$$

2. (♠) Wyznaczanie rozwiązania ogólnego. Operacje elementarne na układzie równań)

Znajdź rozwiązanie ogólne poniższego układu równań, wskazując ciąg równoważnych układów równań, uzyskiwanych za pomocą elementarnych operacji. Opisz dokładnie te operacje.

$$\begin{cases} 2x_1 + 3x_2 + x_3 + 2x_4 = 1 \\ 4x_1 + 6x_2 + 3x_3 + 2x_4 = 3 \\ 6x_1 + 9x_2 + 5x_3 + 2x_4 = 5 \end{cases}$$

3. (♠) Zapisywanie zbioru rozwiązań układu równań liniowych w zależności od parametrów)

Rozpatrzmy układ równań liniowych o współczynnikach rzeczywistych postaci

$$\begin{cases} x_1 + 3x_2 - x_3 + 3x_4 = 1 \\ 2x_1 + 7x_2 + x_3 + 6x_4 = 4 \end{cases}$$

Znajdź rozwiązanie ogólne tego układu. Wskaż zmienne zależne i niezależne (parametry).

Wypisz wszystkie czwórki (s_1, s_2, s_3, s_4) będące rozwiązaniami tego układu.

4. • (♠) Niech ciąg (x_1, \dots, x_n) jest rozwiązaniem równania liniowego o współczynnikach rzeczywistych $a_1x_1 + \dots + a_nx_n = 0$. Uzasadnij, że dla dowolnego $\lambda \in \mathbb{R}$ ciąg $(\lambda x_1, \dots, \lambda x_n)$ jest również rozwiązaniem tego równania.
 • Niech ciągi $(1, 2, 3, -1)$ i $(3, 6, 9, -3)$ będą rozwiązaniami układu równań liniowych U . Uzasadnij, że ciąg $(0, 0, 0, 0)$ jest rozwiązaniem tego układu.
5. • (♠) Niech ciągi (x_1, \dots, x_n) oraz (y_1, \dots, y_n) będą rozwiązaniami równania $a_1x_1 + \dots + a_nx_n = a$. Udowodnij, że ciąg $(x_1 - y_1, \dots, x_n - y_n)$ jest rozwiązaniem równania $a_1x_1 + \dots + a_nx_n = 0$.
 • Niech ciągi $(1, 2, 3, 4)$ i $(2, 0, 0, 1)$ będą rozowaniami układu U . Uzasadnij, że układ ten ma nieskończenie wiele rozwiązań.

6. Niech $n \geq 3$. Rozwiąż, w zależności od n , układ równań:

$$\begin{cases} x_1 + x_2 = 0 \\ x_i + x_{i+1} + x_{i+2} = 0 \quad (1 \leq i \leq n-2) \\ x_{n-1} + x_n = 0. \end{cases}$$

7. Do niesprzecznego układu n równań liniowych U o współczynnikach rzeczywistych dołączono równanie liniowe, otrzymując układ sprzeczny U' złożony z $n+1$ równań. Rozstrzygnij, czy układ U'' powstałý z U' przez wykreślenie pewnego równania z U może być niesprzeczny?

8. Niech U będzie układem trzech równań liniowych o czterech zmiennych i współczynnikach rzeczywistych. Założymy, że układ równań V powstaje z U przez zastąpienie każdego równania w U sumą dwóch pozostałych. Rozstrzygnij, czy układy równań U oraz V są zawsze równoważne?

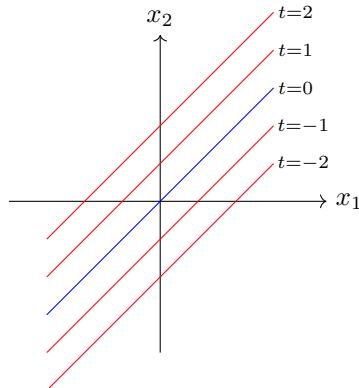
9. Założymy, że $n \geq 1$ oraz $m \geq 2$. Rozważmy układ równań

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \dots + a_{1m}x_m = b_1 \\ \vdots \\ a_{n1}x_1 + a_{n2}x_2 + \dots + a_{nm}x_m = b_n, \end{cases}$$

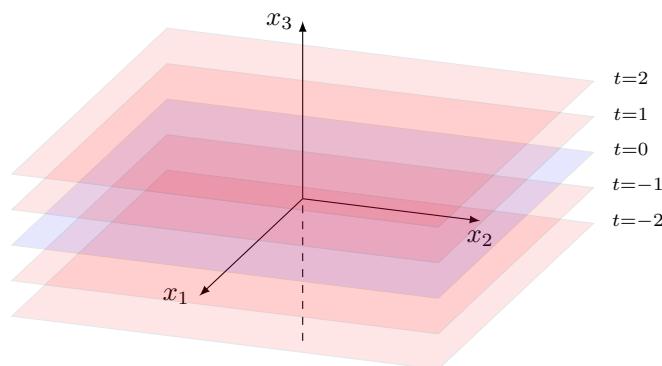
gdzie $a_{ij} \in \mathbb{R}$ oraz $b_i \in \mathbb{R}$ dla $1 \leq i \leq n$ oraz $1 \leq j \leq m$. Przypuśćmy, że dla pewnych $1 \leq r < s \leq m$ zachodzi $a_{ir} = a_{is}$, dla dowolnego $1 \leq i \leq n$. Wykaż, że układ ten jest albo sprzeczny, albo ma nieskończenie wiele rozwiązań.

1.4 Uzupełnienie. Wstępne uwagi o geometrii układów równań

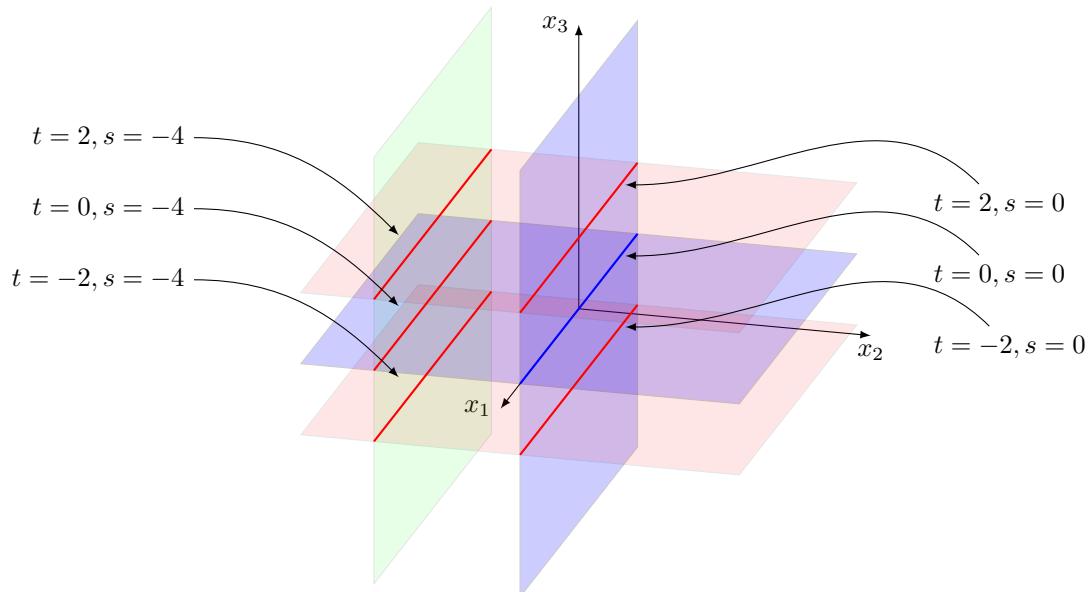
Jednym z celów naszego wykładu jest nadanie struktury geometrycznej obiektom, które (przynajmniej na poziomie intuicji) nie mają wiele wspólnego z geometrią. Rozważmy na przykład rodzinę równań U_t o zmiennych x_1, x_2 i współczynnikach w \mathbb{R} postaci: $-x_1 + x_2 = t$, gdzie $t \in \mathbb{R}$. Zgodnie ze szkolną geometrią analityczną rozwiązań (s_1, s_2) równania U_t można interpretować na płaszczyźnie kartezjańskiej jako współrzędne punktów stanowiących prostą przechodzącą przez punkty $(-t, 0)$ oraz $(0, t)$. Co istotne, dla różnych t uzyskujemy różne proste – wszystkie jednak równoległe do prostej $-x_1 + x_2 = 0$ (stąd np. układ równań $-x_1 + x_2 = 0, -x_1 + x_2 = 1$ jest sprzeczny).



Oto inna rodzina równań liniowych o współczynnikach w \mathbb{R} , o zmiennych x_1, x_2, x_3 , postaci $x_3 = t$, dla $t \in \mathbb{R}$. W trójwymiarowej przestrzeni zbiory rozwiązań tego układu złożone są z trójkę postaci (r, s, t) , gdzie r, s to dowolne liczby rzeczywiste. Są to zbiory stanowiące układ równoległy płaszczyzn.



Rozważmy wreszcie układ równań o zmiennych x_1, x_2, x_3 postaci $x_3 = t, x_2 = s$, gdzie $t, s \in \mathbb{R}$. Zbiory rozwiązań reprezentowane są przez przecięcia nieidentycznych i nierównoległych płaszczyzn, czyli proste



Przedstawimy teraz dwa proste fakty dotyczące zbiorów rozwiązań układów równań o n zmiennych o współczynnikach rzeczywistych, wiążące rozwiązania układu równań liniowych z odpowiadającym mu układem jednorodnym. Fakty te będą w przyszłości elementem dowodu tw. Kroneckera-Capellego.

Uwaga 1.4.1

Jeśli ciągi (s_1, s_2, \dots, s_n) oraz $(s'_1, s'_2, \dots, s'_n)$ są rozwiązaniami układu równań

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + \dots + a_{2n}x_n = b_2 \\ \dots \\ a_{m1}x_1 + a_{m2}x_2 + \dots + a_{mn}x_n = b_m. \end{cases}, \quad (1.1)$$

to ciąg

$$(s_1 - s'_1, s_2 - s'_2, \dots, s_n - s'_n)$$

jest rozwiązaniem układu równań:

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n = \mathbf{0} \\ a_{21}x_1 + a_{22}x_2 + \dots + a_{2n}x_n = \mathbf{0} \\ \dots \\ a_{m1}x_1 + a_{m2}x_2 + \dots + a_{mn}x_n = \mathbf{0}. \end{cases} \quad (1.2)$$

Dowód. Ciągi (s_1, s_2, \dots, s_n) oraz $(s'_1, s'_2, \dots, s'_n)$ są rozwiązaniami każdego z równań

$$a_{i1}x_1 + a_{i2}x_2 + \dots + a_{in}x_n = b_i,$$

gdzie $i = 1, 2, \dots, m$. Pisząc wprost mamy:

$$\begin{aligned} a_{i1}s_1 + a_{i2}s_2 + \dots + a_{in}s_n &= b_i \\ a_{i1}s'_1 + a_{i2}s'_2 + \dots + a_{in}s'_n &= b_i, \end{aligned}$$

czyli po odjęciu stronami widzimy, że $(s_1 - s'_1, s_2 - s'_2, \dots, s_n - s'_n)$ spełnia każde z m równań układu (1.2):

$$a_{i1}(s_1 - s'_1) + a_{i2}(s_2 - s'_2) + \dots + a_{in}(s_n - s'_n) = 0,$$

co oznacza, że jest to rozwiązanie całego układu (1.2). \square

Uwaga 1.4.2

Załóżmy, że $(\mathbf{s}_1, \dots, \mathbf{s}_n)$ jest rozwiązaniem układu równań (1.1). Wówczas każde rozwiązanie układu (1.2) jest postaci:

$$(\mathbf{s}_1 + \mathbf{u}_1, \dots, \mathbf{s}_n + \mathbf{u}_n) \quad (\diamond),$$

gdzie $(\mathbf{u}_1, \dots, \mathbf{u}_n)$ jest rozwiązaniem układu (1.2).

Używając (na razie poglądowo) interpretacji geometrycznej można powiedzieć, że zbiór rozwiązań układu niejednorodnego \mathbf{U} reprezentowany jest przez zbiór równoległy do zbioru rozwiązań układu jednorodnego \mathbf{U}' , odpowiadającego układowi U – zbiór zawierający dowolne rozwiązanie układu U .

Dowód. Weźmy rozwiązanie $(\mathbf{s}_1, \dots, \mathbf{s}_n)$ układu (1.1) i rozwiązanie $(\mathbf{u}_1, \dots, \mathbf{u}_n)$ układu (1.2). Wówczas ciąg $(\mathbf{s}_1 + \mathbf{u}_1, \dots, \mathbf{s}_n + \mathbf{u}_n)$ jest rozwiązaniem układu (1.1), bo spełnia dowolne z jego równań:

$$a_{i1}(\mathbf{s}_1 + \mathbf{u}_1) + \dots + a_{in}(\mathbf{s}_n + \mathbf{u}_n) = (a_{i1}\mathbf{s}_1 + \dots + a_{in}\mathbf{s}_n) + (a_{i1}\mathbf{u}_1 + \dots + a_{in}\mathbf{u}_n) = b_i + 0 = b_i.$$

Pozostaje wykazać, że dowolne rozwiązanie (s'_1, \dots, s'_n) układu (1.1) można przedstawić w postaci (\diamond) , dla pewnego rozwiązania $(\mathbf{u}_1, \dots, \mathbf{u}_n)$ układu (1.2). Zgodnie z poprzednią Uwagą wystarczy wziąć:

$$\mathbf{u}_1 = s'_1 - \mathbf{s}_1, \quad \dots, \quad \mathbf{u}_n = s'_n - \mathbf{s}_n$$

i dostajemy $(s'_1, \dots, s'_n) = (s_1 + (s'_1 - s_1), \dots, s_n + (s'_n - s_n))$. \square

Dla zilustrowania powyższych faktów rozważmy układ równań liniowych o współczynnikach w \mathbb{R} postaci:

$$\begin{cases} x + 2y + 3z = 1 \\ 4x + 5y + 6z = 1 \\ 7x + 8y + 9z = 1 \end{cases} \quad (\dagger)$$

Zgodnie z procedurą opisaną wyżej do opisania rozwiązań tego układu potrzebujemy dowolne jego rozwiązanie, na przykład $(s_1, s_2, s_3) = (-1, 1, 0)$ oraz wszystkie rozwiązania układu postaci:

$$\begin{cases} x + 2y + 3z = 0 \\ 4x + 5y + 6z = 0 \\ 7x + 8y + 9z = 0 \end{cases} \quad (\dagger)$$

Można sprawdzić, że rozwiązania tego układu są postaci $(z, -2z, z)$, gdzie $z \in \mathbb{R}$. W związku z tym każde rozwiązanie układu (\dagger) ma postać $(-1 + z, 1 - 2z, z)$, gdzie $z \in \mathbb{R}$.

Jaka jest zatem geometryczna natura naszego problemu? Chodziło o wyznaczenie miejsca geometrycznego przecięcia trzech płaszczyzn opisanych równaniami

$$x + 2y + 3z = 1, \quad 4x + 5y + 6z = 1, \quad 7x + 8y + 9z = 1.$$

Najpierw, wyznaczyliśmy miejsce geometryczne przecięcia równoległych do tych płaszczyzn postaci:

$$x + 2y + 3z = 0, \quad 4x + 5y + 6z = 0, \quad 7x + 8y + 9z = 0.$$

(to jest prosta $(z, -2z, z)$, gdzie $z \in \mathbb{R}$, którą oznaczaliśmy na niebiesko na wcześniejszych rysunkach). Później musielibyśmy sprawdzić, czy każdą z tych płaszczyzn da się przesunąć równolegle tak, by przecięcie spełniało trzy wyjściowe równania. Nie zawsze musi się to udać.

Interpretacja geometryczna układu równań zaprezentowana wyżej stanowi swego rodzaju **wierszowy obraz** tego układu – patrzmy na każde równanie osobno, interpretujemy jego rozwiązania jako podzbiory przestrzeni trójwymiarowej i szukamy części wspólnej tych zbiorów. Przejedźmy teraz do odrobinę mniej intuicyjnego, **kolumnowego obrazu** opisującego układ (\dagger) . Przepiszmy ten te układ do postaci:

$$x \begin{bmatrix} 1 \\ 4 \\ 7 \end{bmatrix} + y \begin{bmatrix} 2 \\ 5 \\ 8 \end{bmatrix} + z \begin{bmatrix} 3 \\ 6 \\ 9 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}. \quad (\clubsuit)$$

Widzimy, że naszym celem jest teraz znalezienie takich liczb rzeczywistych x, y, z , aby suma pewnych „skalarnej wielokrotności” wektorów $(1, 4, 7), (2, 5, 8), (3, 6, 9)$ stała się wektorem $(1, 1, 1)$. Ta nowa interpretacja geometryczna jest, co może być uznane za zaskoczenie, niezwykle istotna. Dlaczego? Przede wszystkim dlatego, że pozwala nam dużo powiedzieć o strukturze rozwiązań układu (\dagger) , o czym przekonamy się dalej. Po drugie, interpretacja ta jest wartościowa, bo pozwala nam zobaczyć na ile istotny jest wybór tego, a nie innego układu współczynników. Zmiana wektora $(1, 1, 1)$ na inny może całkowicie zmienić odpowiedź na pytanie czy dany układ ma rozwiązanie, czy nie. Na przykład układ postaci:

$$x \begin{bmatrix} 1 \\ 4 \\ 7 \end{bmatrix} + y \begin{bmatrix} 2 \\ 5 \\ 8 \end{bmatrix} + z \begin{bmatrix} 3 \\ 6 \\ 9 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}.$$

nie może mieć rozwiązań, bo wektor $(0, 0, 1)$ „nie leży” na tej samej płaszczyźnie w przestrzeni, co wektory $(1, 4, 7), (2, 5, 8), (3, 6, 9)$. To nie jest oczywiste, bo nie widać wcale, że te trzy wektory muszą leżeć na płaszczyźnie i to takiej, w której nie ma wektora $(0, 0, 1)$. Jeśli jednak przepiszemy w języku wektorowym operacje wykonane na wyjściowym układzie, wówczas okaże się, że jest to jasne. Zauważmy, że układ

$$\begin{cases} x = 1 - 2y - 3z \\ y + 2z = 1 \\ 0 = 0 \end{cases},$$

równoważny układowi (\dagger) przepisuje się w języku „wektorów kolumnowych” jako:

$$x \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} + y \begin{bmatrix} 2 \\ 1 \\ 0 \end{bmatrix} + z \begin{bmatrix} 3 \\ 2 \\ 0 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}. \quad (\spadesuit)$$

To, co jest kluczowe to fakt, że lewe strony warunków (\clubsuit) oraz (\spadesuit) opisują ten sam zbiór wektorów (tego nie widać „na wierszach”). Teraz jest już jasne, że $(0, 0, 1)$ nie jest jego elementem.

1.5 Dodatek. Układy nierówności liniowych

Szereg zagadnień praktycznych sprowadza się do rozwiązywania układów nierówności liniowych lub ogólniej – tzw. zadań programowania liniowego, o których powiemy w kolejnych rozdziałach. Zagadnienia te mają podstawowe znaczenie m.in. w ekonomii.

Definicja 1.5.1: Układ nierówności liniowych

UKŁADEM NIERÓWNOŚCI LINIOWYCH o n zmiennych x_1, \dots, x_n i współczynnikach rzeczywistych nazywamy układ postaci:

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n \leq b_1 \\ a_{21}x_1 + a_{22}x_2 + \dots + a_{2n}x_n \leq b_2 \\ \vdots \quad \vdots \quad \vdots \quad \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \dots + a_{mn}x_n \leq b_m \end{cases},$$

gdzie $a_{ij}, b_i \in \mathbb{R}$, dla $1 \leq i \leq m$ oraz $1 \leq j \leq n$.

ROZWIĄZANIEM UKŁADU nierówności liniowych nazywamy każdy ciąg liczb s_1, \dots, s_n , które po podstawieniu za zmienne x_1, \dots, x_n spełniają wszystkie nierówności tego układu, czyli dla każdego $1 \leq i \leq m$ spełniona jest nierówność

$$a_{i1}s_1 + \dots + a_{in}s_n \leq b_i.$$

Układ nierówności, który nie ma rozwiązania nazywamy SPRZECZNYM.

Podstawowa metoda rozwiązywania układów nierówności liniowych przypisywana jest Fourierowi. Kluczowe jest w niej zauważenie, że następujące warunki są równoważne dla liczb rzeczywistych a, b :

- (1) Istnieje liczba rzeczywista x spełniająca układ warunków $a \leq x$ oraz $x \leq b$.
- (2) Zachodzi nierówność $a \leq b$.

Przykład. Rozważmy układ nierówności liniowych postaci

$$\begin{cases} -x_1 \leq 0 \\ x_1 + 3x_2 \leq 9 \\ -x_1 - x_2 \leq -4 \\ x_1 - x_2 \leq 3 \\ -x_2 \leq 0 \end{cases}$$

Przepiszmy powyższy układ tak, aby wyizolować zmienną x_1 .

$$\begin{cases} 0 \leq x_1 \\ x_1 \leq 9 - 3x_2 \\ 4 - x_2 \leq x_1 \\ x_1 \leq 3 + x_2 \\ x_2 \geq 0 \end{cases}$$

Uzyskany układ możemy zapisać w równoważnej postaci

$$\begin{cases} x_1 \leq \min\{9 - 3x_2, 3 + x_2\} \\ \max\{0, 4 - x_2\} \leq x_1 \\ x_2 \geq 0 \end{cases}$$

Powyższy układ nierówności ma rozwiązanie (s_1, s_2) , wtedy i tylko wtedy, gdy s_2 spełnia

$$\begin{cases} 0 \leq 9 - 3s_2 \\ 0 \leq 3 + s_2 \\ 4 - s_2 \leq 9 - 3s_2 \\ 4 - s_2 \leq 3 + s_2 \\ 0 \leq s_2 \end{cases}$$

W ten sposób sprowadziliśmy problem istnienia rozwiązania układu nierówności o dwóch zmiennych do istnienia rozwiązania układu nierówności o jednej zmiennej. Udowodnimy, że podejście to działa w ogólnej sytuacji, dla dowolnego układu skończenie wielu nierówności liniowych.

Twierdzenie 1.5.2: Fourier-Motzkin

Załóżmy, że H jest zbiorem rozwiązań układu nierówności liniowych od n zmiennych x_1, \dots, x_n . Niech $p(H)$ będzie zbiorem złożonym ze wszystkich takich ciągów postaci s_1, \dots, s_{n-1} , że ciąg s_1, \dots, s_n jest elementem zbioru H . Wówczas $p(H)$ jest zbiorem rozwiązań układu nierówności liniowych od $n - 1$ zmiennych x_1, \dots, x_{n-1} .

Oczywiście zbiór H jest niepusty wtedy i tylko wtedy, gdy zbiór $p(H)$ jest niepusty. Stąd przez eliminację kolejnych zmiennych możemy sprowadzić problem rozwiązywania układu nierówności od n zmiennych do problemu rozwiązywania układu nierówności jednej zmiennej.

Dowód. Założymy, że zbiór H zadany jest układem m nierówności postaci

$$a_{i1}x_1 + a_{i2}x_2 + \dots + a_{in}x_n \leq b_i. \quad (\dagger)$$

Rozważmy trzy zbiory postaci

$$P = \{i : a_{in} > 0\}, \quad N = \{j : a_{jn} < 0\}, \quad Z = \{k : a_{kn} = 0\}.$$

Dla indeksów i, j należących do zbioru $P \cup N$ przyjmujemy, po podzieleniu odpowiednich nierówności przez a_{in} :

$$\begin{aligned} x_n &\leq a'_{i1}x_1 + \dots + a'_{i,n-1}x_{n-1} + b'_i, \text{ dla } i \in P \\ a'_{j1}x_1 + \dots + a'_{j,n-1}x_{n-1} + b'_j &\leq x_n, \text{ dla } j \in N. \end{aligned}$$

W konsekwencji, warunek mówiący, że ciąg s_1, \dots, s_{n-1} należy do zbioru $p(H)$ można zapisać za pomocą układu nierówności

$$a'_{j1}x_1 + \dots + a'_{j,n-1}x_{n-1} + b'_j \leq x_n \leq a'_{i1}x_1 + \dots + a'_{i,n-1}x_{n-1} + b'_i,$$

gdzie $i \in P$ oraz $j \in N$. Zauważmy przy tym, że jeśli $P = \emptyset$ lub $N = \emptyset$, to $p(H)$ opisany jest nierównosciami (\dagger) , dla $i \in Z$. Jeśli również Z jest zbiorem pustym, to do $p(H)$ należy dowolny ciąg $n - 1$ liczb.

Oznacza to, że zbiór $p(H)$ jest zadany za pomocą $|P| \cdot |N|$ nierówności postaci:

$$(a'_{j1} - a'_{i1})x_1 + \dots + (a'_{j,n-1} - a'_{i,n-1})x_{n-1} \leq (b'_i - b'_j)$$

oraz nierówności typu (\dagger) , dla których $a_{in} = 0$, i dla których indeksy i są spoza zbioru $P \cup N$. Stąd zbiór $p(H)$ jest zbiorem rozwiązań układu nierówności liniowych od zmiennych x_1, \dots, x_{n-1} . Ostatnia część tezy jest oczywistym wnioskiem z opisu $p(H)$. \square

W przypadku problemu rozwiązywania układu równań liniowych wykażemy, że zachodzi jedna z dwóch alternatywnych możliwości: układ U ma rozwiązanie lub istnieje taki układ równoważny do U , który zawiera równanie typu $0 = 1$, stanowiące swego rodzaju bezpośredni "certyfikat" tego, że U jest sprzeczny.

Okazuje się, że podobny rezultat, zwany Lematem Farkasa, można udowodnić dla układu nierówności liniowych. Mówi on, że dla układu m nierówności liniowych albo można wskazać rozwiązanie, albo istnieje układ m nieujemnych współczynników $\lambda_1, \dots, \lambda_n$, takich że przemnożenie odpowiednich nierówności przez te współczynniki i dodanie ich stronami sprawi, że uzyskamy nierówność postaci $0 \leq -1$, stanowiącą potwierdzenie, że wyjściowy układ jest sprzeczny.

Ogólny dowód tego lematu przedstawimy w dalszych rozdziałach. Odnotujmy, że każde równanie liniowe postaci $a_1x_1 + \dots + a_nx_n = b$ ma ten sam zbiór rozwiązań, co układ dwóch nierówności $a_1x_1 + \dots + a_nx_n \leq b$ oraz $-a_1x_1 - \dots - a_nx_n \leq -b$. Dowód Lematu Farkasa stanović będzie więc istotne uogólnienie rozważań, które przeprowadzimy w kolejnym rozdziale dla układów równań.

1.6 Przypomnienie. Algebra zbiorów i rachunek zdań

Przypomnijmy podstawową notację teorii zbiorów (teorii mnogości), których szczegółowe omówienie znaleźć można w literaturze, choćby w podręczniku Rasiowej. Czymy to, aby Czytelnik nie musiał w różnych miejscach odwoływać się do innych źródeł, chcąc uzyskać jedynie przypomnienie podstawowych pojęć.

Teorie matematyczne zbudowane są w sposób pozwalający na dowodzenie prawdziwości pewnych zdań, nazywanych tezami, poprzez ciąg wynikań (implikacji) rozpoczynający się od zdań, nazywanych założeniami. Chcemy przy tym uniknąć sytuacji, gdy w ramach tej samej teorii uzyskujemy zdania sprzeczne ze sobą, tzw. antynomie. Dopóki odwołujemy się jedynie do intuicji dotyczących określonych obiektów, występowania antynomii trudno uniknąć, o czym przekonali się twórcy teorii zbiorów w początkach XX wieku, formułując chociażby problem istnienia „zbioru wszystkich zbiorów”.

W celu uniknięcia występowania zdań sprzecznych, formułuje się tak zwane TEORIE AKSJOMATYCZNE. W ramach teorii aksjomatycznej wybierane są pojęcia, które w niej występują, nazywane POJĘCIAMI PIERWOTNYMI, oraz układ zdań charakteryzujących te pojęcia, zwanych AKSJOMATAMI. Za TWIERDZENIA teorii uznajemy zdania, które można uzyskać z aksjomatów poprzez poprawne rozumowanie. Samych aksjomatów nie dowodzimy. Wszelkie jednak własności pojęć danej teorii, które nie są ujęte w aksjomatach, wymagają dowodu. Zwykle przedstawiamy jedynie wybór istotnych konstrukcji i rezultatów.

Za pojęcia pierwotne teorii zbiorów przyjmujemy pojęcie zbioru oraz pojęcie przynależności elementu do zbioru. Przynależenie elementu x do zbioru X zapisujemy symbolicznie jako $x \in X$, a brak przynależenia – jako $x \notin X$. Przy tym przyjmujemy następujące nazewnictwo i oznaczenia.

- (1) Jeśli zbiory A i B mają te same elementy, to mówimy, że zbiory A i B są RÓWNE, co oznaczamy $A = B$.
- (2) Dla dowolnych zbiorów A i B zbiór, którego wszystkimi elementami są wszystkie elementy zbioru A i wszystkie elementy zbioru B , i który nie zawiera innych elementów, nazywamy SUMĄ ZBIORÓW A i B , co oznaczamy $A \cup B$.
- (3) Dla dowolnych zbiorów A i B zbiór zawierający te i tylko te elementy, które należą jednocześnie do obu tych zbiorów, nazywamy CZĘŚCIĄ WSPÓŁNĄ lub PRZECIĘCIEM zbiorów A i B , co oznaczamy $A \cap B$.
- (4) Dla dowolnych zbiorów A i B zbiór, którego elementami są wszystkie elementy zbioru A , które nie są elementami zbioru B , i który nie zawiera innych elementów, nazywany różnicą zbioru A i B , co oznaczamy $A \setminus B$.
- (5) Jeśli A i B są zbiorami oraz każdy element zbioru A jest elementem zbioru B , to zbiór A nazywamy PODZBIOREM zbioru B , co oznaczamy $A \subseteq B$. Mówimy też, że ma miejsce ZAWIERANIE lub INKLUSJA zbioru A w zbiorze B .

Zbiory A i B są równe wtedy i tylko wtedy, gdy $A \subseteq B$ oraz $B \subseteq A$. Jeśli zbiory A i B nie są równe, piszemy $A \neq B$. Jeśli natomiast $A \neq B$ oraz $A \subseteq B$, to piszemy $A \subset B$ lub $A \subsetneq B$ i mówimy, że zawieranie zbioru A w zbiorze B jest ścisłe lub, że A jest PODZBIOREM WŁAŚCIWYM zbioru B . Wprowadzamy też zbiór, który nie ma elementów, zwany ZBIOREM PUSTYM, oznaczanym przez \emptyset .

- (6) Dla każdego zbioru X zbiór, którego elementami są wszystkie podzbiory zbioru X i tylko one, oznaczany przez $P(X)$.

Zbiór mający n elementów nazywamy n -elementowym. Zbiór, do którego należą elementy a, b, c, d, e , i tylko one, oznaczamy jako

$$\{a, b, c, d, e\},$$

a zbiór, do którego należą elementy a, b, c, \dots, m, n , i tylko one, oznaczamy przez

$$\{a, b, c, \dots, m, n\}.$$

Zbiór liczb naturalnych oznaczamy przez \mathbb{N} i przyjmujemy, że 0 nie jest liczbą naturalną. Przez $\mathbb{Z}, \mathbb{Q}, \mathbb{R}$ oznaczamy odpowiednio zbiory liczb całkowitych, wymiernych i rzeczywistych.

Twierdzenia matematyczne są przykładami ZDAŃ, przy czym zajmujemy się jedynie zdaniem, które są prawdziwe lub fałszywe. Takie zdania nazywamy ZDANIAMI LOGICZNYMI. Oznaczamy je poniżej małymi literami p, q, r, \dots

Z dwóch dowolnych zdań logicznych można tworzyć nowe zdania, łącząc je za pomocą spójników „i”, „lub”, „jeśli..., to ...” lub „wtedy i tylko wtedy, gdy”. Dla dowolnego zdania logicznego p można sformułować również ZAPRZECZENIE lub NEGACJĘ tego zdania, czyli zdanie o przeciwniej wartości logicznej.

Zdanie „ p i q ” oznaczamy $p \wedge q$ i nazywamy KONIUNKcją ZDAŃ p i q . Koniunkcja $p \wedge q$ jest prawdziwa, gdy obydwa zdania p, q są prawdziwe. W każdym innym przypadku zdanie $p \wedge q$ jest fałszem.

Zdanie „ p lub q ” zapisujemy jako $p \vee q$ i nazywamy ALTERNATYWĄ ZDAŃ p i q . Alternatywa $p \vee q$ jest prawdziwa, gdy choć jedno ze zdań p, q jest prawdziwe. Jeśli obydwa te zdania są fałszywe, to również ich alternatywa jest fałszywa.

Zdanie „jeśli p , to q ” nazywamy IMPLIKACJĄ O POPRZEDNIKU p I NASTĘPNIKU q , którą oznaczamy $p \Rightarrow q$. Implikację uznajemy za fałszywą, gdy jej następnik jest fałszywy, a poprzednik – prawdziwy. We wszystkich innych przypadkach implikację uznajemy za prawdziwą. Jeśli implikacja $p \Rightarrow q$ jest prawdziwa, to mówimy, że „ q wynika z p ”. Najczęściej interesuje nas sytuacja, gdy zdanie p jest prawdziwe. Wówczas prawdziwość prawdziwość implikacji $p \Rightarrow q$ oznacza, że zdanie q jest prawdziwe.

Zdanie „ p wtedy i tylko wtedy, gdy q ” nazywamy RÓWNOWAŻNOŚCIĄ zdań p i q . Zapisujemy ją w postaci $p \Leftrightarrow q$. Zdanie to uznajemy za prawdziwe, jeśli zdania p, q mają te same wartości logiczne, czyli są oba prawdziwe lub oba fałszywe.

Na podstawie wprowadzonych wyżej konstrukcji formułować można złożone zdania logiczne. Szczególnie istotne jest poprawne stosowanie zaprzeczeń koniunkcji i alternatywy. Zaprzeczeniem alternatywy jest koniunkcja zaprzeczeń, a zaprzeczeniem koniunkcji jest alternatywa zaprzeczeń (prawa de Morgana).

FORMĄ ZDANIOWĄ określoną w zbiorze A nazywamy każde zdanie zawierające zmienną, które po podstawieniu w miejsce zmiennej dowolnego elementu zbioru A staje się zdaniem logicznym. Przykładem jest forma: „ $x > 5$ ”. Dla każdej formy zdaniowej $\varphi(x)$, gdzie $x \in X$, podzbiór zbioru X zawierający wszystkie elementy $x \in X$, dla których zdanie $\varphi(x)$ jest prawdziwe oznaczmy jako $\{x \in X \mid p(x)\}$. Dla formy „ $x > 5$ ” jest to zbiór $\{x \in \mathbb{R} \mid x > 5\}$.

Szczególnie istotnymi formami zdaniowymi są KWANTYFIKATORY. Umawiamy się, że jeśli dla formy zdaniowej $p(x)$ na zbiorze X wszystkie elementy zbioru X mają własność wyrażoną przez tę formę (czyli zdanie $p(x)$ jest prawdziwe, dla każdego $x \in X$), to piszemy symbolicznie: $\forall_{x \in X} p(x)$. W przypadku, gdy co najmniej jeden element zbioru X spełnia warunek $p(x)$, piszemy symbolicznie $\exists_{x \in X} p(x)$. Kwantyfikator \forall nazywamy OGÓLNYM lub UNIWERSALNYM, a kwantyfikator \exists nazywamy EGZYSTENCJALNYM.

Często formułujemy zaprzeczenia zdań, w których występują kwantyfikatory. Na przykład, zaprzeczeniem zdania: „dla każdego $x \in X$ prawdą jest $p(x)$ ” jest zdanie: „istnieje $x \in X$, że zdanie $p(x)$ nie jest prawdą”. Podobnie budujemy zaprzeczenia bardziej skomplikowanych zdań. Na przykład, zaprzeczeniem zdania: „ $\exists_{n \in \mathbb{Z}} \forall_{x \in \mathbb{R}} x \leq n$ ” jest zdanie „ $\forall_{n \in \mathbb{Z}} \exists_{x \in \mathbb{R}} x > n$.”

Gdy $p(x)$ oraz $q(x)$ są formami zdaniowymi na zbiorze X , to może się okazać, że każdy element $x \in X$ mający własność $p(x)$ ma też własność $q(x)$, czyli:

$$\{x \in X \mid p(x)\} \subseteq \{x \in X \mid q(x)\}.$$

Mówimy wtedy, że forma $q(x)$ WYNIKA z formy $p(x)$ lub, że forma $p(x)$ IMPLIKUJE formę $q(x)$, co piszemy symbolicznie w postaci: $p(x) \Rightarrow q(x)$. Zapis ten możemy odczytywać także jako: „z $p(x)$ wynika $q(x)$ ” lub „jeżeli $p(x)$, to $q(x)$ ”. Na przykład w zbiorze \mathbb{R} prawdziwa jest implikacja: $x > 0 \Rightarrow x^2 > 0$. Implikacja $x^2 > 0 \Rightarrow x > 0$ nie jest prawdziwa.

Dla implikacji $p(x) \Rightarrow q(x)$ implikację $q(x) \Rightarrow p(x)$ nazywamy IMPLIKACJĄ ODWROTNĄ. Formy zdaniowe $p(x)$ oraz $q(x)$, dla których zachodzą jednocześnie implikacje $p(x) \Rightarrow q(x)$ oraz $q(x) \Rightarrow p(x)$ nazywamy RÓWNOWAŻNYMI, co oznaczamy $p(x) \Leftrightarrow q(x)$. Dla przykładu, w zbiorze \mathbb{R} mamy

$$(x - 1)(x - 2) = 0 \Leftrightarrow (x = 1 \vee x = 2).$$

1.7 Przypomnienie. Dowodzenie i metoda indukcji

Umiejętność dowodzenia twierdzeń jest jedną z kluczowych kompetencji matematycznych. Twierdzenie matematyczne formułowane jest za pomocą zdania logicznego, którego prawdziwość należy uzasadnić. Podzielone jest ono na dwie części:

- ZAŁOŻENIA, czyli pewne prawdziwe zdania,
- TEZY – zdania, których prawdziwość należy rozstrzygnąć.

Istnieje szereg metod umożliwiających dowodzenie twierdzeń i sprawdzania poprawności dowodów, które opisuje dział matematyki, zwany logiką. Poniżej wyróżniamy trzy podstawowe metody dowodowe: dowód wprost, dowód nie wprost oraz indukcja.

* * *

1. DOWÓD WPROST (DEDUKCJA) polega na argumentowaniu przez wskazywanie ciągu prawdziwych implikacji prowadzących od założeń do tezy.

Zadanie 1.1. Jeśli liczba całkowita $n \neq 0$ jest dzielnikiem zarówno liczby całkowitej a , jak i liczby całkowitej b , to n jest dzielnikiem dowolnej liczby postaci $xa + yb$, gdzie $x, y \in \mathbb{Z}$.

Rozwiązanie. Oto założenia: liczba n jest całkowita i niezerowa, liczby a, b są całkowite, zaś liczba n jest dzielnikiem a i jest dzielnikiem b . Teza natomiast brzmi: dla dowolnych liczb całkowitych x, y liczba n jest dzielnikiem liczby $xa + yb$.

Z założeń wynikają następujące dwa zdania: istnieje liczba całkowita k , taka że $a = kn$ oraz: istnieje liczba całkowita l , taka że $b = ln$. Z powyższych dwóch zdań wynika zdanie: dla dowolnych liczb całkowitych x, y liczba $xa + yb$ jest postaci $xkn + yln = (xk + yl)n$. Z ostatniego zdania wynika teza zadania. ■

* * *

2. DOWÓD NIE WPROST (APAGOGICZNY) polega na tym, aby z zaprzeczenia tezy wydedukować zaprzeczenie jednego z założeń, lub zaprzeczenie jakiegoś prawdziwego zdania.

Zadanie 1.2. Niech n będzie liczbą całkowitą, taką że liczba n^2 jest nieparzysta. Wykaż, że liczba n jest nieparzysta.

Rozwiązanie. Teza zadania brzmi: „liczba n jest nieparzysta”. Przypuśćmy zatem, że nie jest to zdanie prawdziwe. Niech $n = 2k$, dla pewnej liczby całkowitej k . Ze zdania tego wynika, że $n^2 = (2k)^2 = 4k^2$. Stąd wnioskujemy, że liczba n^2 jest parzysta, co jest sprzeczne z założeniem, że liczba n^2 jest nieparzysta. Uzyskana sprzeczność oznacza, że zaprzeczenie tezy nie jest prawdą. Stąd teza zadania jest prawdziwa. ■

Zauważmy, że istnieje inne rozwiązanie powyższego zadania (spotkamy się wielokrotnie z tym, że istnieje wiele możliwych podejść do jednego problemu), które oprzeć można o rozumowanie wprost.

Rozwiązanie. Z założenia mówiącego, że liczba n jest nieparzysta możemy wnioskować, że istnieje liczba całkowita k , że $n = 2k + 1$. Stąd $n^2 = (2k + 1)^2 = 4k^2 + 4k + 1 = 2(2k^2 + 2k) + 1$. Uzyskana liczba jest postaci $2l + 1$, dla $l = 2k^2 + 2k$. Zatem n^2 jest liczbą nieparzystą. ■

Metoda dowodu nie wprost związana jest z wieloma podstawowymi technikami dowodowymi, zwłaszcza z tzw. zasadą ekstremum oraz zasadą szufladkową Dirichleta.

ZASADA SZUFLADKOWA mówi, że jeśli zbiór n -elementowy rozbijemy na mniej niż n rozłącznych podzbiorów, to co najmniej jeden z tych podzbiorów zawiera więcej niż jeden element. Inaczej mówiąc – funkcja ze zbioru mającego więcej niż n elementów do zbioru n elementowego nie jest różnowartościowa.

Zadanie 1.3. Ze zbioru $\{1, 2, \dots, 100\}$ wybieramy takie 51 liczb, że suma żadnych dwóch z nich nie jest równa 100. Pokazać, że zbiór ten zawiera kwadrat liczby całkowitej.

Rozwiązanie. Oznaczmy zbiór tych 50 liczb przez A . Przypuśćmy nie wprost, że zbiór A nie zawiera kwadratu. Rozważmy pary

$$\{1, 99\}, \{2, 98\}, \dots, \{49, 51\}.$$

Jest ich 49. Co więcej, zbiór A nie zawiera ani 36, ani 64. Zatem zbiór A zawiera przynajmniej 49 elementów z pozostałych par. Zawiera więc pewną parę w całości – na mocy zasady szufladkowej. Ale zakładaliśmy, że w A nie ma elementów, których suma równa jest 100. Uzyskujemy sprzeczność z przyjętym założeniem, że w zbiorze A nie ma kwadratu. ■

ZASADA EKSTREMUM mówi (w podstawowej wersji), że w skończonym zbiorze liczb rzeczywistych istnieje element największy i element najmniejszy.

Zadanie 1.4. Czy istnieje wielościan wypukły, w którym każde dwie ściany mają inną liczbę boków?

Rozwiązańe. Taki wielościan nie istnieje. Przypuśćmy przeciwnie. Weźmy ścianę F o największej liczbie m krawędzi. Wówczas ściana F oraz m sąsiadujących z nią ścian to wielokąty, które mogą mieć $3, 4, 5, \dots, m$ boków. Mamy zatem $m+1$ ścian, a tylko $m-2$ możliwości. Stąd przynajmniej jedna liczba boków musi wystąpić dwukrotnie. Doszliśmy do sprzeczności. ■

Zadanie 1.5. W pewnej grupie co najmniej jedna para osób to znajomi oraz żadne dwie osoby o tej samej liczbie znajomych nie mają wspólnych znajomych. Wykaż, że pewna osoba w tej grupie ma dokładnie jednego znajomego.

Rozwiązańe. Rozważmy osobę o największej liczbie znajomych, równej n . Niech ci znajomi to x_1, \dots, x_n . Z założenia zadania wynika, że każde dwie z tych osób mają inne liczby znajomych. Co więcej, każda z tych osób ma nie więcej niż n , a więc jedna z nich musi mieć jednego znajomego. ■

Zadanie 1.6. Na płaszczyźnie pokolorowano skończenie wiele punktów: część na biało, a część na czarno, przy czym wiadomo, że każdy odcinek łączący dwa punkty tego samego koloru zawiera punkt innego koloru. Pokazać, że wszystkie pokolorowane punkty leżą na jednej prostej.

Rozwiązańe. Przypuśćmy, wbrew tezie, że nie wszystkie pokolorowane punkty leżą na jednej prostej. Rozważmy trójkąt o najmniejszym (dodatnim) polu złożony z tych punktów. Dwa z jego wierzchołków są jednego koloru. Oznacza to, że jeden z boków tego trójkąta zawiera w swoim wnętrzu punkt innego koloru. Oznacza to jednak, że istnieje trójkąt o mniejszym polu o pokolorowanych wierzchołkach. Uzyskana sprzeczność kończy dowód. ■

* * *

3. ZASADA INDUKCJI jest w istocie jednym z aksjomatów liczb naturalnych (pochodzących od Peano), choć w praktyce dowodzenia korzystamy z następującego jej ujęcia.

Twierdzenie 1.7.1

Niech $k \in \mathbb{N}$. Niech $T(k), T(k+1), T(k+2) \dots$ będzie ciągiem zdań logicznych. Jeśli zachodzą jednocześnie dwa następujące warunki:

- (i) $T(k)$ jest zdaniem prawdziwym,
- (ii) dla każdej liczby naturalnej $n \geq k$ mamy $T(n) \implies T(n+1)$,

to wszystkie zdania $T(k), T(k+1), T(k+2) \dots$ w ciągu $\{T(n) \mid n \geq k\}$ są prawdziwe.

Przypomnijmy pewien słownik pojęć związany z metodą indukcji.

- Sprawdzenie warunków (i) i (ii) dla ciągu zdań $T(n)$ nazywamy *indukcją* ze względu na n lub *rozumowaniem indukcyjnym*.
- Warunek (i), czyli zdanie $T(k)$, nazywamy *bazą indukcji*.
- Rozumowanie sprawdzające prawdziwość zdania (i) nazywamy *krokiem bazowym* indukcji.
- Przy sprawdzaniu warunku (ii), założenie prawdziwości zdania $T(n)$ nazywać będziemy *założeniem indukcyjnym*, a zdanie $T(n+1)$ nazwiemy *tezą indukcyjną*.
- Samo rozumowanie sprawdzające implikację (ii) nazywamy *krokiem indukcyjnym*. Należy sprawdzać, że obydwa warunki (i), (ii) są spełnione.

Zadanie 1.7. Niech $T(n)$ będzie, dla $n \geq 1$ zdaniem postaci:

$$1 + 2 + \dots + n = \frac{n(n+1)}{2}.$$

Wykaż, że zdanie $T(n)$ jest prawdziwe, dla każdego $n \geq 1$.

Innymi słowy,

- zdanie $T(1)$ jest postaci $1 = \frac{1 \cdot 2}{2}$,
- zdanie $T(2)$ jest postaci $1 + 2 = \frac{2 \cdot 3}{2}$,
- zdanie $T(3)$ jest postaci $1 + 2 + 3 = \frac{3 \cdot 4}{2}$, i tak dalej.

Rozwiązanie. Przeprowadzamy rozumowanie indukcyjne, mające na celu uzasadnienie, że zdanie $T(n)$ jest prawdziwe dla każdej liczby naturalnej.

Baza indukcji, czyli zdanie $T(1)$, jest prawdziwe. Przechodzimy do kroku indukcyjnego. Zakładamy, że zdanie $T(n)$ jest prawdziwe (dla pewnej liczby $n > 1$). Wykażemy prawdziwość zdania $T(n+1)$. Na mocy założenia indukcyjnego, czyli zdania $T(n)$, mamy:

$$1 + 2 + \dots + n = \frac{n(n+1)}{2}.$$

Dodając do stron powyższej równości $n+1$ otrzymujemy

$$1 + 2 + \dots + n + (n+1) = \frac{n(n+1)}{2} + (n+1) = \frac{(n+1)(n+2)}{2}.$$

Stąd zdanie $T(n+1)$ jest prawdziwe. W rezultacie krok indukcyjny jest zakończony. Wykonaliśmy krok bazowy i krok indukcyjny, więc na mocy zasady indukcji, zdanie $T(n)$ jest prawdziwe, dla $n \geq 1$. ■

Zadanie 1.8. Niech $T(n)$ będzie, dla $n \geq 1$, zdaniem: zachodzi nierówność:

$$2^n \geq n+1.$$

Wykaż, że zdanie $T(n)$ jest prawdziwe, dla każdego $n \geq 1$.

Rozwiązanie. Baza indukcji jest oczywiście prawdziwa, gdyż zdanie $T(1)$ ma postać $2 \geq 2$. Założymy, że prawdziwe jest zdanie $T(n)$ postaci $2^n > n+1$. Mamy:

$$2^{n+1} = 2 \cdot 2^n > 2 \cdot (n+1) = 2n+2 > n+2.$$

W rezultacie uzyskaliśmy prawdziwość tezy indukcyjnej. Zatem na mocy zasady indukcji, zdanie $T(n)$ jest prawdziwe, dla każdego $n \geq 1$. ■

Jeśli z kontekstu wynika jasno jak sformułować zdanie $T(n)$, to nie będziemy tego czynić ani w treści zadania, ani w rozwiązaniu. Zaznaczać będziemy jedynie jak nazywa się zmienna, ze względu na którą prowadzona jest indukcja, a następnie przejdźmy do poszczególnych kroków indukcji.

Zadanie 1.9. Udowodnimy, że dla dowolnej liczby naturalnej liczba postaci

$$11^{n+2} + 12^{2n+1}$$

jest podzielna przez 133.

Rozwiązanie. Rozumowanie jest indukcją ze względu na n , dla $n \geq 0$. Krok bazowy jest jasny — dla $n = 0$ liczba $11^2 + 12 = 133$ jest podzielna przez 133.

Przechodzimy do kroku indukcyjnego. Rozważmy liczbę

$$11^{(n+1)+2} + 12^{2(n+1)+1} = 11^{n+3} + 12^{2n+3}.$$

Zauważmy, że liczba ta jest równa:

$$\begin{aligned} 11^{n+2} \cdot 11 + 12^{2n+1} \cdot 12^2 &= 11^{n+2} \cdot (144 - 133) + 12^{2n+1} \cdot 144 \\ &= 144(11^{n+2} + 12^{2n+1}) - 133 \cdot 12^{2n+1}. \end{aligned}$$

Zgodnie z założeniem indukcyjnym, liczba $144(11^{n+2} + 12^{2n+1})$ jest podzielna przez 133. Liczba $133 \cdot 12^{2n+1}$ jest oczywiście wielokrotnością liczby 133. Zatem także różnica tych liczb, czyli $11^{n+3} + 12^{2n+3}$, jest podzielna przez 133. Krok indukcyjny jest zatem zakończony. ■

Zadanie 1.10. Wykaż, że liczba przekątnych n -kąta wypukłego równa jest

$$\frac{n(n-3)}{2},$$

dla każdej liczby naturalnej $n \geq 3$.

Rozwiązanie. Rozumowanie jest indukcją ze względu na n . Skoro liczba przekątnych trójkąta równa jest 0, to baza indukcji zachodzi, dla $n = 3$.

Przechodzimy do kroku indukcyjnego. Założmy, że dowolny n -kąt wypukły ma $\frac{n(n-3)}{2}$ przekątnych. Rozważmy dowolny $n+1$ -kąt wypukły W i rozważmy jego kolejne wierzchołki A_1, \dots, A_{n+1} (idąc przeciwnie z ruchem wskazówek zegara).

Rozważmy wierzchołek A_{n+1} . Liczba wszystkich przekątnych wielokąta W równa jest sumie dwóch liczb: liczby przekątnych W o końcu w A_{n+1} i liczby przekątnych W , które nie mają końca w A_{n+1} . Pierwsza liczba jest oczywiście równa $n-2$. Druga liczba opisuje, ile jest przekątnych n -kąta wypukłego o wierzchołkach A_1, \dots, A_n oraz ile jest boków tego samego n -kąta, które nie są zarazem bokami $n+1$ -kąta W (jest jeden taki bok: $A_n A_1$), czyli – zgodnie z założeniem indukcyjnym – jest to liczba $\frac{n(n-3)}{2} + 1$. Stąd liczba przekątnych wielokąta W równa jest

$$n-2 + \frac{n(n-3)}{2} + 1 = \frac{2n-4+2+n^2-3n}{2} = \frac{(n+1)(n-2)}{2}.$$

■

* * *

Zasadę minimum można również formułować nieco ogólniej.

Twierdzenie 1.7.2: Zasada minimum

W każdym niepustym zbiorze liczb naturalnych istnieje liczba najmniejsza, czyli mniejsza lub równa od każdej liczby należącej do tego zbioru.

Dowód. Założmy, że A jest niepustym zbiorem liczb naturalnych i że w A nie ma liczby najmniejszej. Niech B będzie zbiorem liczb naturalnych zdefiniowanym w następujący sposób: liczba naturalna n należy do B wtedy i tylko wtedy, gdy dla każdej liczby naturalnej m , jeśli $m \leq n$, to $m \notin A$. Można powiedzieć, że jest to zbiór ograniczeń dolnych zbioru A .

Łatwo zauważyć, że $0 \in B$. W przeciwnym razie 0 byłaby najmniejszą liczbą w A , wbrew założeniu. Założmy, że $n \in B$. Z definicji zbioru B wynika, że dla każdej liczby naturalnej m , jeśli $m \leq n$, to $m \notin A$. Stąd również $n+1 \notin A$. W przeciwnym razie $n+1$ byłoby najmniejszą liczbą w zbiorze A , wbrew założeniu. W konsekwencji $n+1 \in B$. Wykazaliśmy, że dla zbioru B spełnione są założenia indukcji zupełnej. Stąd $B = \mathbb{N}$. Biorąc pod uwagę definicję zbioru B wnioskujemy, że A jest zbiorem pustym, co przeczy założeniu. □

Korzystając z powyższej zasady wykażemy, rozumując nie wprost, że liczba $\sqrt{2}$ jest niewymierna (dowód będzie nieco inny niż te, podawane w szkole).

Przypuśćmy, rozumując nie wprost, że $\sqrt{2}$ jest liczbą wymierną. Stąd następujący podzbiór zbioru liczb naturalnych jest niepusty:

$$\{k \in \mathbb{N} \mid k\sqrt{2} \in \mathbb{N}\}.$$

Weźmy, zgodnie z zasadą minimum, najmniejszy element tego zbioru równy n . Wówczas liczba $(\sqrt{2}-1)n$ jest naturalna i mniejsza od n . Pomóżmy ją przez $\sqrt{2}$ uzyskując

$$(\sqrt{2}-1)n\sqrt{2} = 2n - n\sqrt{2}.$$

Po raz kolejny otrzymaliśmy różnicę liczb naturalnych. Wskazaliśmy jednocześnie liczbę naturalną, czyli $(\sqrt{2}-1)n$ – mniejszą od n , która przemnożona przez $\sqrt{2}$ daje liczbę całkowitą, co oznacza sprzeczność.

Rozdział 2

Macierze. Operacje elementarne. Rozwiązywanie układów równań

2.1 Wykład 2

Przejdziemy teraz do systematycznego opisu rozwiązywania układów równań liniowych o współczynnikach rzeczywistych. W tym celu wprowadzamy fundamentalne dla całego wykładu pojęcie macierzy.

Definicja 2.1.1: Macierz

MACIERZĄ ROZMIARU $m \times n$ lub inaczej – macierzą o m wierszach i n kolumnach o WYRAZACH ze zbioru X nazywamy tablicę:

$$D = \begin{bmatrix} d_{11} & d_{12} & \dots & d_{1n} \\ d_{21} & d_{22} & \dots & d_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ d_{m1} & d_{m2} & \dots & d_{mn} \end{bmatrix}$$

gdzie $d_{ij} \in X$ dla $1 \leq i \leq m$ oraz $1 \leq j \leq n$.

Przy tym wprowadzamy następujące nazewnictwo.

- Rzędy poziome macierzy D nazywamy WIERSZAMI (uporządkowanymi od góry do dołu), a dokładniej – wyrazy $d_{i1}, d_{i2}, \dots, d_{in}$ tworzą i -ty wiersz macierzy D .
- Rzędy pionowe macierzy D nazywamy KOLUMNAMI (uporządkowanymi od lewej do prawej), a dokładniej – wyrazy $d_{1j}, d_{2j}, \dots, d_{mj}$ tworzą j -tą kolumnę macierzy D .
- Macierz D o wyrazach d_{ij} , gdzie $1 \leq i \leq m, 1 \leq j \leq n$ oznaczamy w skrócie $D = [d_{ij}]$.
- Zbiór wszystkich macierzy $m \times n$ o wyrazach ze zbioru X oznaczamy jako $M_{m \times n}(X)$.

O macierzy można myśleć też jako o funkcji ze zbioru par $\{1, \dots, m\} \times \{1, \dots, n\}$ do zbioru X przypisującej każdej parze (i, j) (dla odpowiednich indeksów) element $d_{ij} \in X$, który nazywamy będącym WYRAZEM i, j macierzy D .

Oto przykład macierzy D o 3 wierszach i 5 kolumnach o wyrazach w \mathbb{R} , czyli elementu zbioru $M_{3 \times 5}(\mathbb{R})$. W drugim wierszu i czwartej kolumnie tej macierzy znajduje się wyraz 1, czyli jeśli $D = [d_{ij}]$, to $d_{24} = 1$.

$$\begin{bmatrix} 1 & 2 & 3 & 0 & -1 \\ \sqrt{2} & 0 & 0 & \textcolor{red}{1} & -\frac{1}{2} \\ 0 & 1 & 3 & 7 & 9 \end{bmatrix}$$

Macierz jest fundamentalnym pojęciem nie tylko w algebrze liniowej, którego nowe interpretacje będziemy poznawać przez cały kurs. W tym miejscu macierze posłużą nam jako narzędzia do przechodzenia z jednego układu równań do układu równoważnego.

Definicja 2.1.2: Macierz układu równań liniowych

Każdemu układowi m równań liniowych o n zmiennych i współczynnikach w K postaci:

$$U : \begin{cases} a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + \dots + a_{2n}x_n = b_2 \\ \dots \\ a_{m1}x_1 + a_{m2}x_2 + \dots + a_{mn}x_n = b_m. \end{cases}$$

możemy przypisać macierz rozmiaru $m \times (n+1)$ o wyrazach w K postaci:

$$\left[\begin{array}{cccc|c} a_{11} & a_{12} & \dots & a_{1n} & b_1 \\ a_{21} & a_{22} & \dots & a_{2n} & b_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} & b_m \end{array} \right].$$

Nazywamy ją MACIERZĄ (ALBO MACIERZĄ ROZSZERZONĄ) UKŁADU U . Macierz powstała przez pojęcie ostatniej kolumny w macierzy układu U będziemy nazywać MACIERZĄ WSPÓŁCZYNNIKÓW UKŁADU U .

Rozważmy kilka przykładów macierzy i odpowiadających im układów równań liniowych:

układ	macierz rozszerzona	macierz współczynników
$\begin{cases} 2x_1 + 3x_2 + x_3 = 1 \\ -\frac{1}{2}x_2 + x_3 = 0 \end{cases}$	$\left[\begin{array}{ccc c} 2 & 3 & 1 & 1 \\ 0 & -\frac{1}{2} & 1 & 0 \end{array} \right]$	$\left[\begin{array}{ccc} 2 & 3 & 1 \\ 0 & -\frac{1}{2} & 1 \end{array} \right]$
$\begin{cases} x_1 + x_2 = 1 \\ x_1 + x_2 = 1 \\ x_1 + x_2 = 1 \\ 0x_1 + 0x_2 = 3 \end{cases}$	$\left[\begin{array}{cc c} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \\ 0 & 0 & 3 \end{array} \right]$	$\left[\begin{array}{cc} 1 & 1 \\ 1 & 1 \\ 1 & 1 \\ 0 & 0 \end{array} \right]$
$\begin{cases} 0x_1 + 0x_2 + 0x_3 = 0 \\ 0x_1 + 0x_2 + x_3 = 0 \\ x_1 + x_2 + x_3 = 0 \end{cases}$	$\left[\begin{array}{ccc c} 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 1 & 1 & 1 & 0 \end{array} \right]$	$\left[\begin{array}{ccc} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 1 & 1 & 1 \end{array} \right]$

Przyjrzyjmy się teraz macierjom układów będących rozwiązaniami ogólnymi układów równań liniowych, opisanymi w Definicji 1.7. Oto przykłady.

układ	macierz rozszerzona
$\begin{cases} \textcolor{red}{x}_1 = 2 \\ \textcolor{red}{x}_2 = 1 \\ \textcolor{red}{x}_3 = 4 \end{cases}$	$\left[\begin{array}{ccc c} 1 & 0 & 0 & 2 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 4 \end{array} \right]$
$\begin{cases} \textcolor{red}{x}_1 = \frac{1}{2} \\ \textcolor{red}{x}_2 + x_3 + x_4 = \frac{1}{2} \end{cases}$	$\left[\begin{array}{cccc c} 1 & 0 & 0 & 0 & \frac{1}{2} \\ 0 & 1 & 1 & 1 & \frac{1}{2} \end{array} \right]$
$\begin{cases} \textcolor{red}{x}_1 + x_3 - 2x_5 = 0 \\ \textcolor{red}{x}_2 + x_5 = 0 \\ \textcolor{red}{x}_4 + x_5 = 0 \end{cases}$	$\left[\begin{array}{ccccc c} 1 & 0 & 1 & 0 & -2 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 1 & 0 \end{array} \right]$

W powyższych przykładach pokreślono kolorem współczynniki odpowiadające zmiennym zależnym. Widać, że w każdym kolejnym wierszu współczynniki te znajdują się w kolumnach o większym indeksie.

Definicja 2.1.3: Operacje elementarne na wierszach

Niech $A \in M_{m \times n}(\mathbb{R})$. Następujące transformacje macierzy A nazywamy OPERACJAMI ELEMENTARNYMI NA WIERSZACH:

- (1) Dodanie do wiersza innego wiersza przemnożonego przez liczbę rzeczywistą.
- (2) Zamiana dwóch wierszy miejscami.
- (3) Pomnożenie wiersza przez liczbę różną od zera.

Analogicznie definiuje się OPERACJE ELEMENTARNE NA KOLUMNACH macierzy A .

Widzimy zatem, że operacjom elementarnym na układzie U odpowiadają operacje elementarne na wierszach macierzy tego układu. Proces znajdowania rozwiązania ogólnego układu U będzie polegał na kolejnym upraszczaniu macierzy tego układu, przy zastosowaniu operacji elementarnych na wierszach.

Przykłady operacji elementarnych na wierszach macierzy i notacja, którą będziemy stosować:

$$\begin{array}{c} \left[\begin{array}{ccccc} 1 & 2 & -1 & -1 & 0 \\ 2 & -1 & -2 & 1 & 1 \\ -1 & 0 & 1 & 4 & 1 \end{array} \right] \xrightarrow{w_2 - 2 \cdot w_1} \left[\begin{array}{ccccc} 1 & 2 & -1 & -1 & 0 \\ 0 & -5 & 0 & 3 & 1 \\ -1 & 0 & 1 & 4 & 1 \end{array} \right] \\ \xrightarrow{w_3 + w_1} \left[\begin{array}{ccccc} 1 & 2 & -1 & -1 & 0 \\ 0 & -5 & 0 & 3 & 1 \\ 0 & 2 & 0 & 3 & 1 \end{array} \right] \\ \xrightarrow{w_2 \leftrightarrow w_3} \left[\begin{array}{ccccc} 1 & 2 & -1 & -1 & 0 \\ 0 & 2 & 0 & 3 & 1 \\ 0 & -5 & 0 & 3 & 1 \end{array} \right] \\ \xrightarrow{2 \cdot w_2} \left[\begin{array}{ccccc} 1 & 2 & -1 & -1 & 0 \\ 0 & 4 & 0 & 6 & 2 \\ 0 & -5 & 0 & 3 & 1 \end{array} \right] \end{array}$$

Definicja 2.1.4: Macierz w postaci schodkowej

Mówimy, że macierz A jest w POSTACI SCHODKOWEJ, jeśli A spełnia następujące warunki:

- każdy wiersz zerowy (tzn. złożony z samych zer) znajduje się poniżej każdego wiersza niezerowego (czyli jeśli i -ty wiersz tej macierzy jest niezerowy, to j -ty wiersz jest zerowy tylko gdy $j > i$),
- dla każdego $i > 1$ pierwszy licząc od lewej niezerowy wyraz w i -tym wierszu znajduje się w kolumnie stojącej na prawo od pierwszego niezerowego wyrazu ($i - 1$)-szego wiersza.

Rozważmy, za skryptem, kilka prostych przykładów ilustrujących nową definicję: dwie pierwsze macierze są w postaci schodkowej, dwie kolejne nie. Czy Czytelnik widzi owe „schodki”?

$$\left[\begin{array}{cccc} 0 & 4 & 7 & 2 \\ 0 & 0 & 1 & 12 \\ 0 & 0 & 0 & 5 \end{array} \right], \quad \left[\begin{array}{cccc} 2 & 5 & 1 & 1 \\ 0 & 3 & 0 & 3 \\ 0 & 0 & 0 & 0 \end{array} \right], \quad \left[\begin{array}{cccc} 0 & 0 & 3 & 4 \\ 0 & 1 & 0 & -1 \\ 0 & 0 & 0 & 5 \end{array} \right], \quad \left[\begin{array}{cccc} 2 & 8 & 3 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 2 & 0 & 12 \end{array} \right].$$

Definicja 2.1.5: Macierz w postaci zredukowanej

Mówimy, że macierz jest w ZREDUKOWANEJ POSTACI SCHODKOWEJ (krócej: w postaci zredukowanej), jeśli jest w postaci schodkowej oraz w każdym niezerowym wierszu pierwszy niezerowy wyraz wynosi 1 i jest jedynym niezerowym wyrazem w swojej kolumnie.

Przykłady:

$$\left[\begin{array}{ccc|c} 1 & 0 & 0 & 2 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 4 \end{array} \right], \quad \left[\begin{array}{ccc|c} 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 4 \end{array} \right], \quad \left[\begin{array}{ccccc|c} 1 & 0 & 1 & 0 & -2 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 1 & 0 \end{array} \right], \quad \left[\begin{array}{cc|c} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{array} \right].$$

Jeśli układ ma macierz w postaci schodkowej zredukowanej, to są takie kolumny gdzie powstaje „schodek”, powiedzmy kolumny (od lewej) $j_1 < j_2 < \dots < j_k$ i zmiana z kolumny j_i raz jeden występuje w całym układzie właśnie w i -tym równaniu. Co więcej stoi przy niej współczynnik 1. A zatem jest to macierz, z której łatwo uzyskać rozwiązanie ogólne układu równań, o którym mówiliśmy wcześniej. Chyba, że... jedna z tych kolumn to ostatnia kolumna. Wtedy sytuacja jest nieco inna. Czy Czytelnik widzi jaka? Pokażemy teraz kluczową obserwację: do uzyskania postaci schodkowej i schodkowej zredukowanej wystarczy stosować na wierszach macierzy odpowiednie operacje elementarne.

Twierdzenie 2.1.6

Każda macierz $A \in M_{m \times n}(\mathbb{R})$ można:

- (i) za pomocą operacji elementarnych typu (1) i (2) sprowadzić do postaci schodkowej,
- (ii) za pomocą operacji elementarnych typu (1), (2) i (3) sprowadzić do postaci zredukowanej.

Dowód. Pokażemy jedynie (i). Dowód (ii) będzie ćwiczeniem.

Stosujemy zasadę indukcji matematycznej względem liczby wierszy macierzy¹. Dla $m = 1$ twierdzenie jest oczywiste. Każda macierz o jednym wierszu jest w postaci schodkowej.

Załóżmy, że twierdzenie jest udowodnione dla macierzy o co najwyżej $m - 1$ wierszach. Niech A będzie macierzą rozmiaru $m \times n$ o wyrazach rzeczywistych. Jeśli A jest macierzą zerową (to znaczy każdy jej wyraz jest zerem), to oczywiście jest schodkowa. Założymy więc, że A nie jest macierzą zerową. Niech s będzie indeksem pierwszej niezerowej kolumny macierzy A . Wybierzmy takie r , że $a_{rs} \neq 0$.

$$A = \begin{bmatrix} 0 & \dots & 0 & a_{1s} & * & * & * \\ 0 & \dots & 0 & a_{2s} & * & * & * \\ \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & \dots & 0 & \color{red}{a_{rs}} & * & * & * \\ \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & \dots & 0 & a_{ms} & * & * & * \end{bmatrix}$$

Zamieniamy miejscami wiersze: pierwszy i r -ty (operacja typu (2)). Za pomocą operacji (1) zerujemy² wszystkie, poza pierwszym, wyrazy w s -tej kolumnie (od i -tego wiersza **odejmujemy pierwszy przemnożony przez $\frac{a_{is}}{a_{rs}}$**). Otrzymujemy w ten sposób macierz $A' = [a'_{ij}]$, w której kolumny o indeksach $1, \dots, s-1$ są zerowe oraz $a'_{1s} \neq 0$ i $a'_{is} = 0$, dla $i > 1$.

$$A' = \begin{bmatrix} 0 & \dots & 0 & a_{rs} & * & * & * \\ 0 & \dots & 0 & \color{red}{a_{2s} - \frac{a_{2s}}{a_{rs}} \cdot a_{rs}} & * & * & * \\ \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & \dots & 0 & \color{red}{a_{1s} - \frac{a_{1s}}{a_{rs}} \cdot a_{rs}} & * & * & * \\ \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & \dots & 0 & \color{red}{a_{ms} - \frac{a_{ms}}{a_{rs}} \cdot a_{rs}} & * & * & * \end{bmatrix} = \begin{bmatrix} 0 & \dots & 0 & a_{rs} & * & * & * \\ 0 & \dots & 0 & 0 & * & * & * \\ \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & \dots & 0 & 0 & * & * & * \\ \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & \dots & 0 & 0 & * & * & * \end{bmatrix}$$

Niech $A'' \in M_{(m-1) \times n}(\mathbb{R})$ będzie macierzą otrzymaną z A' przez usunięcie pierwszego wiersza. Stosujemy do A'' założenie indukcyjne. Otrzymujemy macierz, która wraz z pierwszym wierszem macierzy A' tworzy macierz A''' w postaci schodkowej. Oczywiście A''' jest uzyskana z A przez ciąg operacji typu (1) i (2). \square

Metodę rozwiązywania układów równań liniowych opisaną w powyższym wniosku nazywamy METODĄ ELIMINACJI GAUSSA rozwiązywania układów równań. Zobaczmy jak działa na konkretnym przykładzie.

¹ Jeśli Czytelnik nie spotkał się dotąd z zasadą indukcji, można rozumować nieco inaczej: zakładamy, że m jest najmniejszą liczbą naturalną taką, że macierz o m wierszach nie można za pomocą operacji elementarnych sprowadzić do postaci schodkowej. Rozumowanie przedstawione w dowodzie wyżej pokazuje, że dojdziemy do sprzeczności z tym założeniem.

² Krok ten nazywany jest często ELIMINACJĄ i pochodzi od niego nazwa algorytmu opisującego uzyskiwanie postaci schodkowej (a także zredukowanej), jak również wielu równoważnych mu rozkładów macierzy.

$$\begin{cases} x_1 + 2x_2 - x_3 - x_4 = 0 \\ 2x_1 - x_2 - 2x_3 + x_4 + x_5 = 0 \\ -x_1 + x_3 + 4x_4 + x_5 = 0 \end{cases}$$

Za pomocą operacji wierszowych dokonujemy eliminacji w pierwszej kolumnie.

$$\left[\begin{array}{cccc|c} 1 & 2 & -1 & -1 & 0 \\ 2 & -1 & -2 & 1 & 1 \\ -1 & 0 & 1 & 4 & 1 \end{array} \right] \xrightarrow{w_2 - 2w_1} \left[\begin{array}{cccc|c} 1 & 2 & -1 & -1 & 0 \\ 0 & -5 & 0 & 3 & 1 \\ -1 & 0 & 1 & 4 & 1 \end{array} \right] \xrightarrow{w_3 + w_1} \left[\begin{array}{cccc|c} 1 & 2 & -1 & -1 & 0 \\ 0 & -5 & 0 & 3 & 1 \\ 0 & 2 & 0 & 3 & 1 \end{array} \right]$$

Teraz zajmujemy się drugą kolumną.

$$\left[\begin{array}{cccc|c} 1 & 2 & -1 & -1 & 0 \\ 0 & -5 & 0 & 3 & 1 \\ 0 & 2 & 0 & 3 & 1 \end{array} \right] \xrightarrow{w_3 + \frac{5}{2}w_2} \left[\begin{array}{cccc|c} 1 & 2 & -1 & -1 & 0 \\ 0 & -5 & 0 & 3 & 1 \\ 0 & 0 & 0 & \frac{21}{2} & \frac{7}{2} \end{array} \right]$$

Uzyskaliśmy macierz układu (równoważnego wyjściowemu) w postaci schodkowej. Teraz doprowadzamy ją do postaci zredukowanej. W tym celu wykonujemy operacje służące eliminacji w kolumnach, w których znajdują się wyrazy wiodące poszczególnych wierszy.

$$\left[\begin{array}{cccc|c} 1 & 2 & -1 & -1 & 0 \\ 0 & -5 & 0 & 3 & 1 \\ 0 & 0 & 0 & \frac{21}{2} & \frac{7}{2} \end{array} \right] \xrightarrow{\frac{2}{21}w_3} \left[\begin{array}{cccc|c} 1 & 2 & -1 & -1 & 0 \\ 0 & -5 & 0 & 3 & 1 \\ 0 & 0 & 0 & 1 & \frac{1}{3} \end{array} \right] \xrightarrow{w_2 - 3w_3} \left[\begin{array}{cccc|c} 1 & 2 & -1 & 0 & \frac{1}{3} \\ 0 & -5 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & \frac{1}{3} \end{array} \right]$$

Wreszcie, operacje eliminujące wyrazy w drugiej kolumnie prowadzą do uzyskania postaci zredukowanej.

$$\left[\begin{array}{cccc|c} 1 & 2 & -1 & 0 & \frac{1}{3} \\ 0 & -5 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & \frac{1}{3} \end{array} \right] \xrightarrow{-\frac{1}{5}w_2} \left[\begin{array}{cccc|c} 1 & 2 & -1 & 0 & \frac{1}{3} \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & \frac{1}{3} \end{array} \right] \xrightarrow{w_1 - 2w_2} \left[\begin{array}{cccc|c} 1 & 0 & -1 & 0 & \frac{1}{3} \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & \frac{1}{3} \end{array} \right]$$

Wniosek 2.1.7

Każdy niesprzeczny układ równań liniowych ma rozwiązanie ogólne, przy czym:

- aby znaleźć rozwiązanie układu U wystarczy macierz tego układu sprowadzić do zredukowanej postaci schodkowej elementarnymi operacjami na wierszach,
- jeśli otrzymana macierz nie zawiera wiersza postaci $0 \dots 0 1$, to można z niej odczytać rozwiązanie ogólne układu U ,
- jeśli otrzymana macierz zawiera wiersz postaci $0 \dots 0 1$, to układ U jest sprzeczny.

Dowód. Sprowadzamy macierz układu U do zredukowanej postaci schodkowej. Jeśli otrzymana macierz ma wiersz $0 \dots 0 1$, to odpowiadający jej układ równań zawiera równanie $0x_1 + 0x_2 + \dots + 0x_n = 1$, więc układ ten (i równoważny z nim układ U) jest sprzeczny. Jeśli otrzymana macierz w postaci schodkowej zredukowanej nie zawiera wiersza $0 \dots 0 1$, to odpowiadający jej układ równań ma (po pominięciu ewentualnych równań postaci $0x_1 + 0x_2 + \dots + 0x_n = 0$) postać

$$\begin{cases} x_{j_1} + a_{1(j_1+1)}x_{1(j_1+1)} + \dots + a_{1n}x_n = b_1 \\ x_{j_2} + a_{2(j_2+1)}x_{1(j_2+1)} + \dots + a_{2n}x_n = b_2 \\ \vdots \\ x_{j_k} + a_{k(j_k+1)}x_{1(j_k+1)} + \dots + a_{kn}x_n = b_k \end{cases}$$

przy czym $j_1 < j_2 < \dots < j_k$ oraz dla każdego $1 \leq s \leq k$ mamy $a_{ij_s} = 0$, dla wszystkich i . Stąd układ ten można przepisać do postaci

$$\begin{cases} x_{j_1} = -a_{1(j_1+1)}x_{1(j_1+1)} - \dots - a_{1n}x_n + b_1 \\ x_{j_2} = -a_{2(j_2+1)}x_{1(j_2+1)} - \dots - a_{2n}x_n + b_2 \\ \vdots \\ x_{j_k} = -a_{k(j_k+1)}x_{1(j_k+1)} - \dots - a_{kn}x_n + b_k \end{cases}$$

stanowiącej rozwiązanie ogólne układu U .

□

2.2 Wybór przykładowych pytań

Proszę się nie uczyć na pamięć odpowiedzi czy wyjaśnienie! Takich pytań można ułożyć znacznie więcej i służą one jedynie sprawdzeniu podstawowego zrozumienia treści zajęć. Jeśli odpowiedź na dowolne z poniższych pytań sprawia Państwu problem, proszę pomyśleć który fragment notatek dotyczy tego pytania, wrócić do ich lektury, a w razie dalszych trudności poprosić prowadzącego o wskazówkę.

1. Dana jest macierz $M \in M_{2 \times 6}(\mathbb{R})$

$$\begin{bmatrix} 1 & 0 & 1 & 2 & 2 & 3 \\ 3 & 4 & 2 & 3 & 2 & 3 \end{bmatrix}.$$

- a) Wymień wyrazy trzeciej kolumny tej macierzy.
- b) Wymień wyrazy drugiego wiersza tej macierzy.
- c) Niech $M = [m_{ij}]$, dla $1 \leq i \leq 2$ oraz $1 \leq j \leq 6$. Wypisz wyraz m_{25} .

2. Macierz A należy do $M_{3 \times 4}(\mathbb{R})$. Ile wierszy ma macierz A ? Ile ma wyrazów?

3. Jaki układ równań liniowych o wyrazach rzeczywistych ma macierz

$$\begin{bmatrix} 2 & 1 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \end{bmatrix}?$$

4. Macierz układu równań liniowych ma dwa wiersze i osiem kolumn. Ile zmiennych ma ten układ?

5. Czy macierz układu równań może się składać z jednej kolumny? A macierz współczynników?

6. Jakie wyrazy zawiera ostatnia kolumna macierzy jednorodnego układu m równań?

7. Czy macierz układu sprzecznego może mieć w ostatniej kolumnie wyłącznie wyrazy zerowe?

8. Opisz rozmiar i wyrazy macierzy układu równań liniowych postaci

$$x_1 = x_2 = \dots = x_n.$$

9. Wskaż przykładowy ciąg operacji elementarnych na wierszach potrzebny, aby poniższą macierz w postaci schodkowej sprowadzić do postaci zredukowanej

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

10. Kiedy poniższe dwie macierze A, B opisują równoważne układy równań liniowych?

$$A = \begin{bmatrix} a^2 & a & 2a & a \\ 1 - a^2 & 0 & 2 - 2a & 1 - a \\ 1 - a^2 & 1 - a & 0 & 1 - a \end{bmatrix}, \quad B = \begin{bmatrix} a & 1 & 2 & 1 \\ 1 + a & 0 & 2 & 1 \\ 1 + a & 1 & 0 & 1 \end{bmatrix}$$

11. Macierz A jest w postaci schodkowej. Czy każda operacja elementarna na wierszach tej macierzy przeprowadza ją w macierz w postaci schodkowej?

12. Rozstrzygnij, czy poniższa macierz jest w postaci schodkowej zredukowanej?

$$\begin{bmatrix} 1 & 4 & 0 & 7 & 2 \\ 0 & 0 & 1 & 3 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

13. Rozstrzygnij, czy istnieją takie $a, b, c \in \mathbb{R}$, dla których poniższa macierz o wyrazach rzeczywistych

$$\begin{bmatrix} a & b & 1 & 0 & 2 \\ 0 & c & 1 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 \end{bmatrix}$$

- a) nie jest w postaci schodkowej,
- b) jest w postaci schodkowej zredukowanej,
- c) jest macierzą układu sprzecznego.

2.3 Zadania do samodzielnej pracy

Zadania ilustrujące umiejętności niezbędne do zaliczenia przedmiotu oznaczone są symbolem ♠. Przy tych zadaniach dodany jest opis umiejętności, które sprawdzają.

1. (♠) Wyznaczanie postaci schodkowej i zredukowanej macierzy)

Wyznacz zredukowaną postać schodkową macierzy rzeczywistej:

$$\begin{bmatrix} 3 & 1 & 1 & 1 \\ 2 & -1 & -1 & 2 \\ -2 & 2 & 4 & 1 \end{bmatrix}, \quad \begin{bmatrix} 1 & 2 & 3 \\ -2 & 3 & 4 \\ 5 & 7 & 1 \\ 3 & 4 & -1 \end{bmatrix}, \quad \begin{bmatrix} 3 & 2 & 1 & -1 \\ 5 & -1 & 1 & 2 \\ 7 & 8 & 1 & -7 \\ 1 & -1 & 1 & 2 \end{bmatrix}.$$

2. (♠) Wyznacz zredukowaną postać schodkową macierzy rzeczywistej:

$$\begin{bmatrix} a & -b \\ b & a \end{bmatrix}, \quad \begin{bmatrix} 1 & 1 & 1 \\ 2 & 1 & t \\ -1 & 1 & -3 \end{bmatrix}$$

w zależności od parametrów $a, b \in \mathbb{R}$ oraz $t \in \mathbb{R}$.

3. (♠) Macierz układu równań liniowych, kryterium rozwiązywalności układu równań liniowych)

Wyznacz wartości parametru $a \in \mathbb{R}$, dla których poniższy układ równań jest niesprzeczny.

$$\begin{cases} ax_1 + x_2 - 2x_3 + x_4 = a \\ x_1 + ax_2 + x_3 + ax_4 = 3 \\ x_1 + 2x_2 + x_3 + 2x_4 = 2 \end{cases}$$

4. Czy dwa układy równań liniowych U, W o poniższych macierzach A, B są równoważne?

$$A = \begin{bmatrix} 1 & 3 & -2 & 2 \\ -1 & -2 & -1 & -1 \\ -1 & -5 & 8 & -3 \end{bmatrix}, \quad B = \begin{bmatrix} 1 & 2 & 1 & 2 \\ 1 & 1 & 4 & 0 \\ -1 & -1 & -4 & 1 \end{bmatrix}.$$

5. Wyznacz (dowolną) postać schodkową macierzy rozmiaru $n \times n$:

$$\begin{bmatrix} 1 & n & n & \dots & n \\ n & 2 & n & \dots & n \\ n & n & 3 & \dots & n \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ n & n & n & \dots & n \end{bmatrix}, \quad \begin{bmatrix} 1 & 1 & \dots & 1 & -n \\ 1 & 1 & \dots & -n & 1 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 1 & -n & \dots & 1 & 1 \\ -n & 1 & \dots & 1 & 1 \end{bmatrix}, \quad \begin{bmatrix} 1 & 1 & \dots & 1 & 1 \\ 1 & 2 & \dots & 1 & 2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 1 & 1 & \dots & n-1 & n-1 \\ 1 & 2 & \dots & n-1 & n \end{bmatrix}.$$

6. Wyrazy kwadratowej macierzy współczynników $A = [a_{ij}] \in M_n(\mathbb{R})$ pewnego jednorodnego układu równań liniowych spełniają warunek

$$|a_{ii}| > \sum_{j=1, j \neq i}^n |a_{ij}|, \text{ dla wszystkich } i = 1, 2, \dots, n.$$

Wykaż, że układ ten ma dokładnie jedno rozwiązanie.

7. Niech $A \in M_{m \times n}(\mathbb{R})$ i niech macierz A' powstaje z A przez zamianę dwóch wierszy, a następnie pomnożenie jednego z nich przez -1 . Czy A' można otrzymać z A ciągiem operacji elementarnych typu (1), czyli operacji polegających na dodaniu do wiersza innego wiersza pomnożonego przez liczbę.

8. Założmy, że $A \in M_{n \times n}(\mathbb{R})$ jest macierzą współczynników pewnego układu równań liniowych, który ma dokładnie jedno rozwiązanie. Uzasadnij, że macierz A można sprowadzić do postaci

$$\begin{bmatrix} 1 & 0 & \dots & 0 & 0 \\ 0 & 1 & \dots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \dots & 1 & 0 \\ 0 & 0 & \dots & 0 & d \end{bmatrix},$$

gdzie $d \neq 0$ jest pewną liczbą rzeczywistą, wyłącznie za pomocą operacji elementarnych polegających na dodaniu do wiersza innego wiersza pomnożonego przez liczbę.

2.4 Uzupełnienie. Wierszowa równoważność macierzy

Metoda eliminacji Gaussa-Jordana rozwiązywania układu równań liniowych zakłada, że macierz rozszerzoną układu równań sprowadzamy do zredukowanej postaci schodkowej i albo odczytujemy rozwiązanie ogólne wyjściowego układu, albo stwierdzamy, że jest to układ sprzeczny. Prowadzi to do naturalnego pytania o jednoznaczność tej procedury. Sformułujemy je w języku poniższej definicji.

Definicja 2.4.1: Macierze wierszowo równoważne

Powiemy, że macierze $A, B \in M_{m \times n}(\mathbb{R})$ są WIERSZOWO RÓWNOWAŻNE, jeśli jedną można otrzymać z drugiej poprzez ciąg operacji elementarnych na wierszach.

Jest jasne, że jeśli macierze są wierszowo równoważne, to są macierzami równoważnych układów równań liniowych. A odwrotnie? Problem postawiony na początku można sprowadzić zatem do dowodu następującego faktu.

Twierdzenie 2.4.2

Niech A, B będą macierzami w zredukowanej postaci schodkowej, które są wierszowo równoważne. Wówczas $A = B$.

Uzasadnienie poprzedzimy kilkoma uwagami wstępymi. Odnotujmy na początek następującą oczywistą obserwację.

Uwaga 2.4.3

Niech macierze $A', B' \in M_{m \times (n-1)}(\mathbb{R})$ powstają przez usunięcie i -tych kolumn odpowiednio macierzy $A, B \in M_{m \times n}(\mathbb{R})$. Jeśli macierze A, B są wierszowo równoważne, to również macierze A', B' są wierszowo równoważne.

Przykład. Poniższe macierze A, B są wierszowo równoważne i pozostają takimi po usunięciu pierwszych kolumn.

$$A = \begin{bmatrix} 1 & 2 & 3 \\ 1 & 1 & 0 \end{bmatrix}, \quad B = \begin{bmatrix} 0 & 1 & 3 \\ 2 & 2 & 0 \end{bmatrix}, \quad A' = \begin{bmatrix} 2 & 3 \\ 1 & 0 \end{bmatrix}, \quad B' = \begin{bmatrix} 1 & 3 \\ 2 & 0 \end{bmatrix}.$$

Wprowadźmy dodatkową definicję dotyczącą macierzy będącej w zredukowanej postaci schodkowej. Przypomnijmy, że macierz ta jest w postaci schodkowej oraz w każdym niezerowym wierszu pierwszy niezerowy wyraz wynosi 1 i jest jedynym niezerowym wyrazem w swojej kolumnie. Ów wyraz nazywać będziemy *wyrazem wiodącym*.

Dla przykładu, w macierzy $A = [a_{ij}] \in M_{3 \times 5}(\mathbb{R})$, gdzie

$$\begin{bmatrix} 0 & 1 & 5 & 0 & 3 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix},$$

wyrazy wiodące to a_{12} oraz a_{24} i zawierają je kolumna druga i czwarta.

* * *

Dowód TWIERDZENIA 2.4.2. Przypuśćmy nie wprost, że $A \neq B$. Niech i będzie najmniejszą liczbą taką, że i -te kolumny macierzy A oraz B są różne. Niech A' oraz B' będą macierzami powstającymi z A, B przez:

- usunięcie kolumn na prawo od i -tej, czyli kolumn o numerach $j > i$,
- usunięcie tych kolumn na lewo od i -tej, które nie zawierają wyrazów wiodących.

Na przykład, opisana procedura zastosowana do macierzy

$$A = \begin{bmatrix} 1 & 3 & 0 & 3 & 2 \\ 0 & 0 & 1 & 5 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}, \quad B = \begin{bmatrix} 1 & 3 & 0 & 4 & 7 \\ 0 & 0 & 1 & 6 & 9 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

daje macierze

$$A' = \begin{bmatrix} 1 & 0 & 3 \\ 0 & 1 & 5 \\ 0 & 0 & 0 \end{bmatrix}, \quad B' = \begin{bmatrix} 1 & 0 & 4 \\ 0 & 1 & 6 \\ 0 & 0 & 0 \end{bmatrix}$$

Zauważmy że zgodnie z Uwagą 2.4.3 macierze A', B' są wierszowo równoważne. Wykażemy, że $A' = B'$, co doprowadzi do sprzeczności.

Załóżmy, że macierze A', B' mają po $k+1$ kolumn. Każda z macierzy A', B' jest w zredukowanej postaci schodkowej i macierze te różnią się jedynie ostatnimi kolumnami, które są równe i -tym kolumnom macierzy A, B . Pierwsze k kolumn tych macierzy ma następującą postać: i -ty wyraz i -tej kolumny jest równy 1, a pozostałe są równe 0. Ostatnia kolumna natomiast albo zawiera wyraz wiodący na miejscu $k+1$, wtedy wszystkie inne są zerowe, albo ma co najwyżej k niezerowych wyrazów w pierwszych k wierszach.

Podsumowując, każda z macierzy A', B' jest w jednej z dwóch postaci

$$\begin{bmatrix} 1 & 0 & 0 & \dots & 0 & 0 \\ 0 & 1 & 0 & \dots & 0 & 0 \\ 0 & 0 & 1 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & 1 & 0 \\ 0 & 0 & 0 & \dots & 0 & 1 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \end{bmatrix} \quad \text{lub} \quad \begin{bmatrix} 1 & 0 & 0 & \dots & 0 & c_{1k} \\ 0 & 1 & 0 & \dots & 0 & c_{2k} \\ 0 & 0 & 1 & \dots & 0 & c_{3k} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & 1 & c_{kk} \\ 0 & 0 & 0 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \end{bmatrix}$$

przy czym wiersze poniżej wiersza $k+1$ są zerowe (o ile istnieją).

Skoro, jak ustaliliśmy, macierze A', B' są wierszowo równoważne, to muszą być obydwie w tej samej z wymienionych wyżej dwóch postaci. Macierz po lewej jest bowiem macierzą układu sprzecznego, a macierz po prawej – układu równań liniowych mającego jednoznaczne rozwiązanie postaci $x_1 = c_{1k}, \dots, x_k = c_{kk}$. W przypadku, gdy A', B' są obydwie równe macierzy po lewej, wtedy $A' = B'$. Natomiast gdy A', B' są równe macierzom, jak po prawej, wtedy

$$A' = \begin{bmatrix} 1 & 0 & 0 & \dots & 0 & a_{1k} \\ 0 & 1 & 0 & \dots & 0 & a_{2k} \\ 0 & 0 & 1 & \dots & 0 & a_{3k} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & 1 & a_{kk} \\ 0 & 0 & 0 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \end{bmatrix}, \quad B' = \begin{bmatrix} 1 & 0 & 0 & \dots & 0 & b_{1k} \\ 0 & 1 & 0 & \dots & 0 & b_{2k} \\ 0 & 0 & 1 & \dots & 0 & b_{3k} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & 1 & b_{kk} \\ 0 & 0 & 0 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \end{bmatrix}.$$

Skoro jednak A', B' są wierszowo równoważne, to są macierzami układów równań o identycznych rozwiązaniach. Stąd $a_{ik} = b_{ik}$, dla $i \leq k$. Zatem $A' = B'$, co przeczy założeniu, że ostatnie kolumny tych macierzy są różne. Otrzymana sprzeczność oznacza, że $A = B$.

* * *

Zaprezentowane wyżej rozumowanie opiera się o analizę wyrazów macierzy w zredukowanej postaci schodkowej i może wydawać się stosunkowo pomysłowe i zawiłe. W kolejnych rozdziałach wprowadzone zostaną narzędzia, pozwalające na uzasadnienie Twierdzenia 2.4.2 w zdecydowanie bardziej elementarny i naturalny sposób.

2.5 Dodatek. Nieliniowe układy równań

Pokazaliśmy, że rozwiązywanie układów równań liniowych polega na zastępowaniu ich przez inne układy o tym samym zbiorze rozwiązań. Zachowanie zbioru rozwiązań gwarantuje stosowanie operacji elementarnych. Czy podobna metoda da się stosować dla innych typów układów równań? Rozważmy układ:

$$\begin{cases} x^2 + yz + x = 0 \\ y^2 + xz + y = 0 \\ z^2 + xy + z = 0 \end{cases}$$

Układ ten nie jest układem równań liniowych, ale wydaje się, że nic nie stoi na przeszkodzie by wykonywać na nim wszystkie opisane wcześniej operacje elementarne. Czy ich stosowanie zmienia zbiór rozwiązań? Po odjęciu drugiego równania od pierwszego oraz trzeciego równania od drugiego dostajemy (co chyba jasne) układ równoważny postaci:

$$\begin{cases} (x-y)(x+y-z+1) = 0 \\ (y-z)(y+z-x+1) = 0 \\ z^2 + xy + z = 0 \end{cases}$$

Chwila skrupulatnego rozważania tych warunków (nic trudnego) prowadzi do trójkę (x, y, z) rozwiązań:

$$(0, 0, 0), (-1, 0, 0), (0, -1, 0), (0, 0, -1), (-1/2, -1/2, -1/2).$$

Czy rozwiążemy w ten sposób każdy układ równań wielomianowych (na razie rozumianych intuicyjnie)? Czy ma sens mówienie o macierzy współczynników tego układu? Jak ją określić?

Czy jest jakiś analog postaci schodkowej macierzy związanych z takimi układami?

Czy są jakieś specyficzne dla wielomianów „operacje elementarne”, które nie zmieniają zbioru rozwiązań? Czym są „zmienne niezależne” w przypadku istnienia nieskończenie wielu rozwiązań układu równań?

Konkretnie odpowiedzi na powyższe pytania związane są z działem algebra zwany teorią pierścieni oraz odkryciem Buchbergera z 1965 roku dotyczącym wyznaczania tak zwanych baz Gröbnera. Kto by chciał poczytać więcej o tym algorytmie i jego związku z algorytmem Gaussa, może zatrzymać się pod poniższymi adresami:

<https://www.math.usm.edu/perry/Research/GaussToGroebner.pdf>

<https://www.theoremoftheday.org/MathsStudyGroup/Buchberger.pdf>

W tym miejscu damy Czytelnikowi jedną tylko intuicję, która pozwala zobaczyć w nieco innym świetle sam algorytm Gaussa i sens wyznaczania postaci schodkowej (stąd też obecność tego tematu w tym miejscu). Założymy, że ograniczamy się do rozważania układów równań postaci:

$$\begin{cases} F_1(x, y) = 0 \\ \dots \\ F_s(x, y) = 0 \end{cases}$$

gdzie $F_i(x, y)$ są, dla $1 \leq i \leq s$, wielomianami zmiennych x, y , czyli formalnymi sumami jednomianów postaci $c \cdot x^m y^n$, $m, n \geq 0$, $c \in \mathbb{R}$, zaś $m, n \in \mathbb{Z}$ (liczby całkowite!). Jeśli każdy z takich jednomianów spełnia warunek $m + n = 1$, otrzymamy po prostu układ równań liniowych (o dwóch zmiennych).

W przypadku układu równań liniowych zmiennie oznaczaliśmy symbolami x_1, x_2, \dots . Oznacza to – choć na wykładzie nie zwracaliśmy na to uwagi – że zmienne tworzą **zbior uporządkowany**. Macierz współczynników układu, postać schodkowa, opis rozwiązań – wszystko to zależy od **kolejności zmiennych**, jaką przyjmujemy (choć nie zmienia się natura geometryczna zbioru rozwiązań). Idea jest następująca: w przypadku wielomianów uporządkować będziemy nie tylko same zmiennie, ale cały zbiór jednomianów.

Zamiast x pisać będziemy x_1 oraz zamiast y pisać będziemy x_2 . Chcemy przez to podkreślić, że x jest w porządku zmiennych „wcześniejszego”. Wprowadzamy następnie zasadę uporządkowania jednomianów. Przez stopień $\deg(x_1^m x_2^n)$ jednomianu $x_1^m x_2^n$ rozumiemy sumę $m + n$. Ustalamy zasadę, że zawsze „wcześniejszy” (czyli większy) jest jednomian wyższego stopnia. A zatem $x_1^4 x_2^2 > x_1^5$, $x_1 x_2^6 > x_1^2 x_2$ itd. Jeśli natomiast stopnie dwóch jednomianów są identyczne, wcześniejszy jest jednomian z wyższą potegą przy wcześniejszej zmienniej. A więc, na przykład, $x_1^4 x_2^3 > x_1^3 x_2^4$, chociaż $\deg(x_1^4 x_2^3) = \deg(x_1^3 x_2^4) = 7$.

Układ równań napisany na początku naszego dodatku jest, przy założeniu porządku $x > y > z$, napisany tak, że kolejne składniki są coraz mniejsze w „porządku jednomianowym”. Algorytm Buchbergera proponuje – do pewnego stopnia – wykonywanie podobnej procedury, co algorytm Gaussa. Chodzi o zastępowanie jednych wielomianów innymi i dążenie do prostszego układu. Możemy sobie już wyobrazić macierz układu równań wielomianowych. Co z operacjami elementarnymi? Wprowadza się (pozornie) dodatkową „operację elementarną” na dwóch wielomianach – tak zwany S-wielomian. Jak go określić?

Dla dwóch jednomianów $x_1^{m_1}x_2^{n_1}$ oraz $x_1^{m_2}x_2^{n_2}$ rozważa się tzw. najmniejszą wspólną wielokrotność (zbieżność użytego nazewnictwa z tym występującym w teorii podzielności liczb całkowitych nie jest przypadkowa). Nazwiemy ją $NWW(x_1^{m_1}x_2^{n_1}, x_1^{m_2}x_2^{n_2})$ i jest to po prostu jednomian $x_1^{\max\{m_1, m_2\}}x_2^{\max\{n_1, n_2\}}$.

Dla każdego wielomianu f przez $LT(f)$ oznaczamy największy jednomian – w określonym wyżej porządku – który występuje w f . Na przykład $LT(2x^3y^4 + y^6) = 2x^3y^4$ (uwaga – nie samo x^3y^4 , ale $2x^3y^4$).

Dla wielomianów f, g zmiennych x, y , przez $S(f, g)$ rozumiemy wielomian postaci:

$$S(f, g) = \frac{NWW(LT(f), LT(g))}{LT(f_1)} \cdot f_1 - \frac{NWW(LT(f), LT(g))}{LT(f_2)} \cdot f_2.$$

Przykład. Rozważmy przecięcie dwóch elips (proszę uwierzyć mi na słowo, że są to elipsy) postaci:

$$\begin{cases} 2x^2 + y^2 - 4x - 4y + 3 = 0 \\ x^2 + 3y^2 - 2x - 12y + 9 = 0 \end{cases}$$

Niech $f = 2x^2 + y^2 - 4x - 4y + 3$ oraz $g = x^2 + 3y^2 - 2x - 12y + 9$. Wówczas przyjmując $x > y$ i respektując opisany wyżej porządek jednomianowy dostajemy $LT(f) = 2x^2, LT(g) = x^2$, a zatem widzimy, że:

$$S(f, g) = \frac{-5}{2}y^2 + 10y - \frac{15}{2}.$$

Czy Czytelnik widzi co się stało? Jednomian wiodący S-wielomianu $S(f, g)$ jest mniejszy – w porządku jednomianowym – niż jednomian wiodący każdego z wielomianów f, g . Czy widać tu analogię do użyskiwania postaci schodkowej przez algorytm Gaussa, zwany często algorymem **eliminacji** Gaussa? Czy widać, że wykonanie S-wielomianu na dwóch równaniach liniowego układu równań i zastąpienie jednego z równań owym S-wielomianem (dwóch zmiennych – biorąc pod uwagę nasze definicje – ale po odpowiednim uogólnieniu – dowolnym) jest równoważne ze zwykłą operacją elementarną rozważaną na wykładzie?

Czy Czytelnik widzi w jaki sposób stosowanie S-wielomianu prowadzi do zastąpienia wyjściowego układu równań wielomianowych układem prostszym? Jak należy kontynuować ten algorytm i do jakich rozwiązań może on prowadzić? Poszukiwanie odpowiedzi polecam już Państwu. Polecam wskazane wyżej źródła.

Ps. Rozważane elipsy przecinają się w czterech punktach. Dowód wymaga jedynie matematyki szkolnej.

* * *

Zapomniałem wspomnieć: baz Gröbnera używa się do naprawdę ciekawych rzeczy (choćby w robotyce). Polecam dwa anglojęzyczne źródła (im szybciej Państwo zaczną czytać matematykę w tym języku, tym lepiej). Na razie mogą być one dla Państwa trudne w lekturze, ale za jakiś czas będzie można wrócić do tej lektury. A może warto przejść się na konsultacje i poprosić o wyjaśnienie trudniejszych fragmentów?

- Antoine Nectoux, *Map colouring and Gröbner Bases*
- Elizabeth Arnold, Stephen Lucas, and Laura Taalman, *Gröbner Basis Representations of Sudoku*

Aby znaleźć te teksty wystarczy wpisać tytuły w Google.

Polecam również dwugodzinny wykład prof. Przemysława Koprowskiego (Uniwersytet Śląski) dotyczący baz Gröbnera i ich zastosowań (<https://youtu.be/vdmyrbNqRlY>). Jest to znakomite, ścisłe wprowadzenie do tej tematyki, z licznymi przykładami zastosowań.

2.6 Trivia. Kwadraty magiczne

Pierwsza trivia dotyczy pięknego tematu, mianowicie kwadratów magicznych. Jest wiele poważnych tekstuów, zarówno popularyzatorskich, jak i naukowych dotyczących tych obiektów i problemów z nimi związanych. Znane były one od starożytności, a fascynowały między innymi samego Eulera, który poświęcił im kilka swoich artykułów. Obecnie można znaleźć wiele ich tłumaczeń w tym na język angielski. np.

- E530 (<http://eulerarchive.maa.org/docs/translations/E530.pdf>),
- E795³ (<https://arxiv.org/pdf/math/0408230.pdf>).

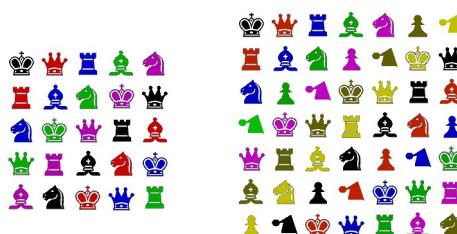
Czym więc są owe kwadraty? Może zamiast definicji podam problem 36 oficerów, badany przez Eulera. Mamy mianowicie 36 oficerów o sześciu różnych stopniach, wziętych z sześciu różnych oddziałów i próbujemy ustawić ich w kwadracie tak, by w każdym wierszu i każdej kolumnie tego kwadratu stało sześciu oficerów z innego oddziału i różnych stopni.

Oto stosowny obrazek pokazujący oficerów, ustawionych na razie zgodnie z przynależnością do oddziału.



Źródło: <http://www.ams.org/publicoutreach/feature-column/fcarch-latinii1>

Zachęcam do eksperymentu i próby dokonania żądanego ustawnienia. Okaże się, że nie bardzo chce się to udało. Eulerowi też to nie wychodziło. A problem był o tyle frustrujący, że analogiczne problemy 25 i 49 oficerów daje się gładko rozwiązać:



Źródło: <http://www.ams.org/publicoutreach/feature-column/fcarch-latinii1>

Euler rozpoznaje w badanym zagadnieniu problem algebraiczny, znany zresztą wcześniej. Założymy, że każdemu oficerowi nadamy plakietkę a^b , gdzie a oznaczać będzie stopień, a b – numer oddziału, a potem zapomnimy o numerach oddziału i popatrzymy tylko na stopnie, to rozkład wyżej ma postać:

7	6	5	4	3	2	1
5	4	3	2	1	7	6
3	2	1	7	6	5	4
1	7	6	5	4	3	2
6	5	4	3	2	1	7
4	3	2	1	7	6	5
2	1	7	6	5	4	3

Dla Eulera kwadratem magicznym jest tablica liczb rozmiaru $n \times n$ taka, że w każdym wierszu, każdej kolumnie i na obydwa przekątnych sumy wpisanych liczb są równe. A jakie liczby wpisujemy? W przypadku wyżej: siedem zestawów od 1 do 7. Czasem rozważa się kwadraty, w które wpisuje się kolejne liczby naturalne (tzw. normalne kwadraty magiczne), a czasem półmagiczne (bez warunku na przekątne) itd. Problem oficerów rozwiązały zostało dopiero w 1901 roku przez matematyka-amatora Gastona Tarry'ego.

³Pewnie ciekawi Państwa co znaczą te liczby? W latach 1910-1913 szwedzki matematyk Gustav Eneström dokonał gruntownych badań nad dziełami Eulera i z nieprzebranych archiwów publikacji oraz zbiorów notatek wyłonił 866 pozycji: książek, artykułów i istotnych listów, którym przydzielił symbole od E1 do E866.

Aby przybliżyć się nieco do materiału z wykładu zajmijmy się magicznymi kwadratami rozmiaru 3×3 o wyrazach rzeczywistych. Możemy je oczywiście utożsamiać z macierzami. Oto przykłady:

$$\begin{bmatrix} 2 & 9 & 4 \\ 7 & 5 & 3 \\ 6 & 1 & 8 \end{bmatrix}, \quad \begin{bmatrix} -1 & 5/2 & 0 \\ 3/2 & 1/2 & -1/2 \\ 1 & -3/2 & 2 \end{bmatrix}, \quad \begin{bmatrix} 1 & 0 & -1 \\ -2 & 0 & 2 \\ 1 & 0 & -1 \end{bmatrix}.$$

Ile jest macierzy magicznych? Powyższe przykłady sugerują, że jest ich nieskończonie wiele. Zauważmy, że wyznaczenie macierzy magicznej równoważne jest problemowi rozwiązyania układu równań liniowych. Aby rozstrzygnąć czy macierz:

$$\begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}$$

reprezentuje kwadrat magiczny należy nałożyć następujące warunki na jej wyrazy:

$$\begin{cases} a_{21} + a_{22} + a_{23} - a_{11} - a_{12} - a_{13} = 0 \\ a_{31} + a_{32} + a_{33} - a_{11} - a_{12} - a_{13} = 0 \\ a_{11} + a_{21} + a_{31} - a_{11} - a_{12} - a_{13} = 0 \\ a_{12} + a_{22} + a_{32} - a_{11} - a_{12} - a_{13} = 0 \\ a_{13} + a_{23} + a_{33} - a_{11} - a_{12} - a_{13} = 0 \\ a_{11} + a_{22} + a_{33} - a_{11} - a_{12} - a_{13} = 0 \\ a_{31} + a_{22} + a_{13} - a_{11} - a_{12} - a_{13} = 0 \end{cases}.$$

czyli rozwiązać układ równań, którego macierz ma aż 7 wierszy i 9 kolumn odpowiadających zmiennym

$$a_{11}, a_{12}, a_{13}, a_{21}, a_{22}, a_{23}, a_{31}, a_{32}, a_{33}$$

postaci

$$\left[\begin{array}{ccccccc|c} -1 & -1 & -1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 \\ -1 & -1 & -1 & 0 & 0 & 0 & 1 & 1 & 1 & 0 \\ 0 & -1 & -1 & 1 & 0 & 0 & 1 & 0 & 0 & 0 \\ -1 & 0 & -1 & 0 & 1 & 0 & 0 & 1 & 0 & 0 \\ -1 & -1 & 0 & 0 & 0 & 1 & 0 & 0 & 1 & 0 \\ 0 & -1 & -1 & 0 & 1 & 0 & 0 & 0 & 1 & 0 \\ -1 & -1 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 \end{array} \right].$$

Idę o zakład, że „schodkowanie” tej macierzy nie wydałoby się Państwu przyjemnością. Od ilu parametrów zależy rozwiązanie tego układu? Nie bardzo widać rozwiązanie. Jest jednak spryniejsza droga: skorzystamy z twierdzenia o rozwiązaniach układów jednorodnych i niejednorodnych. Proszę zauważać, że jeśli mamy macierz magiczną $[a_{ij}]$ taką, że suma wyrazów w każdym wierszu, kolumnie na przekątnych to $3S$, to macierz o wyrazach $a_{ij} - S$ również jest magiczna, a nawet 0-magiczna, bo suma wyrazów w każdym jej wierszu, kolumnie i na przekątnych to 0. Intuicja podpowiada zatem, że liczba parametrów potrzebna do opisania wszystkich macierzy magicznych jest o 1 większa niż liczba parametrów służących do opisu macierzy 0-magicznych. Zobaczmy, że opis macierzy t -magicznych wygląda przejrzyściej. Można go dokonać przez rozwiązywanie układu o macierzy:

$$\left[\begin{array}{ccccccc|c} 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & t \\ 0 & 0 & 0 & 1 & 1 & 1 & 0 & 0 & t \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & t \\ 1 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & t \\ 0 & 1 & 0 & 0 & 1 & 0 & 0 & 1 & t \\ 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & t \\ 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & t \\ 0 & 0 & 1 & 0 & 1 & 0 & 1 & 0 & t \end{array} \right].$$

Znacznie łatwiej analizować powyższą macierz 8×9 , czy ogólnie macierz rozmiaru $(2n+2) \times n^2$, uzyskiwaną dla układu opisującego macierze t -magiczne rozmiaru n . Tu już coś widać. Suma pierwszych trzech wierszy minus suma dwóch kolejnych daje wiersz szósty, więc po wyschodkowaniu jest co najmniej jeden wiersz zerowy... i jak się okazuje nie ma innych. Rozwiązywanie zależy od dwóch parametrów. A zatem wszystkie macierze magiczne rozmiaru 3×3 opisać można za pomocą trzech parametrów⁴.

⁴Czytelnik zechce zauważać, że (posługując się geometryczną intuicją z Uzupełnienia) twierdzimy, że zbiory macierzy t -magicznych rozmiaru 3×3 stanowią rodzinę równoległych płaszczyzn w trójwymiarowej przestrzeni macierzy magicznych.

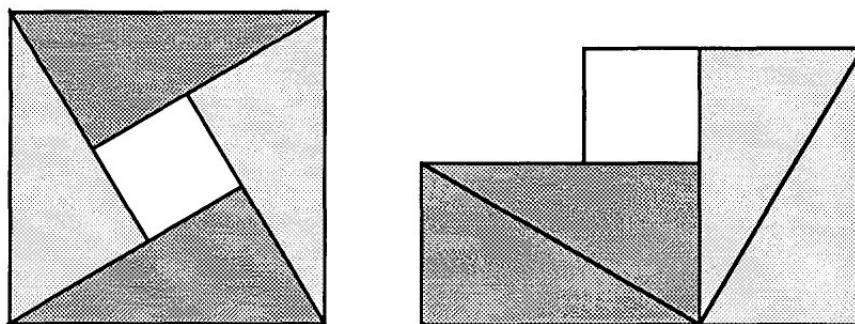
2.7 Coda. Eliminacja i podstawienie, czyli historia upraszczania

Gdyby chcieć wskazać jedną podstawową koncepcję przewijającą się przez całą metodologię nauki, to jest nią zasada mówiąca, że „Nie należy mnożyć bytów ponad potrzebę”. Często przypisuje się ją Ockhamowi, ale podobne zasady „ekonomii myślenia” formułowali już Arystoteles, Platon, a także filozofowie średniowieczni. W wyjaśnianiu zjawisk należy dążyć do prostoty, opierać się na najmniejszej możliwej liczbie pojęć i założeń. Wnioski nalezy wyciągać z możliwie malej liczby racjonalnych przesłanek.

W matematyce zasada dążenia do prostoty wiąże się z wieloma zagadnieniami, a na pierwszym wykładzie tego kursu widzieliśmy ją przede wszystkim w opisie podstawowej metody rozwiązywania układów równań, polegającej na zastępowaniu jednych układów innymi — równoważnymi, o prostszej postaci. Drogą w tym kierunku jest eliminacja zmiennych z równań, którą osiągamy poprzez sprowadzanie postaci macierzy tego układu do postaci schodkowej. Warto powiedzieć, że tego rodzaju podejście obecne jest w historii algebry — nie tylko zresztą akademickiej, ale także tej, którą poznawaliście Państwo przez lata.

W szkole mieli Państwo okazję rozwiązywać wiele typów równań algebraicznych, począwszy od liniowych, przez kwadratowe, wielomianowe, wykładnicze, logarytmiczne czy trygonometryczne. Poznawali Państwo również podstawowe techniki rozwiązywania układów równań, bez ogólnej metody. Czy na studiach poszerzać będziemy arsenał umiejętności w tym zakresie? Do pewnego stopnia tak będzie, choć podstawowe techniki będą dalej istotne: grupowanie, wyciąganie wspólnego czynnika przed nawias, zwijanie do iloczynu, korzystanie z wzorów skróconego mnożenia, wzorów dwumianowych, później także metody rachunku różniczkowego itd. Kluczowe jest dla nas zrozumienie podstawowej koncepcji rozwiązywania równań — PODSTAWIANIA. Algorytm Gaussa przedstawiony na wykładzie jest przypadkiem ogólniejszego sposobu myślenia, którego początki sięgają starożytnych problemów „algebry geometrycznej”.

Pierwsze formuły algebraiczne jakie poznajemy w szkole mają wszystkie, bez wyjątku pochodzenie geometryczne. Za pomocą wzorów zapisujemy obwody i pola podstawowych figur, oraz poprzez rozmaite techniki podziału figur na mniejsze, uzyskujemy nowe wzory — choćby na pole równoległoboku, czy w znacznie głębszym sensie — na pole koła. Za pomocą pól ilustrujemy wzory skróconego mnożenia na kwadrat sumy czy różnicy, a nawet być może zdarzyło się Państwu widzieć „obrazkowy dowód” twierdzenia Pitagorasa. Oto jedna z wielu możliwych wersji pochodzącej od hinduskiego uczonego Bhaskary.



Rys. 1. Dowód twierdzenia Pitagorasa. Źródło: R. Nielsen: *Proofs without words*.

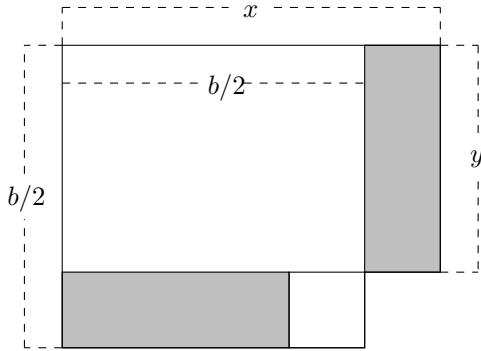
Problemy sprowadzające się do rozwiązywania równań znane były już w starożytnej Mezopotamii. Matematyka Babilończyków miała dwa źródła. Pierwszym było prowadzenie ksiąg rachunkowych, od początku istotnych dla funkcjonowania systemu biurokratycznego wczesnych dynastii rządzących tym obszarem 4000 lat temu. Termin *księgi* jest oczywiście umowny, bowiem teksty zapisywane były na tabliczkach glinianych. Drugim źródłem były problemy geometryczne związane głównie z zagadnieniami podziału terenu. Wiele starych tabliczek glinianych pochodzących z okresu 2000-1700 p.n.e. zawierają rozległe listy tego, co dziś nazwalibyśmy równaniami kwadratowymi, których celem było znalezienie takich wielkości jak długość czy szerokość prostokąta. Przykład takiego problemu pochodzi⁵ z tzw. tabliczki YBC 4663.

Dane są: suma długości i szerokości prostokąta: $6\frac{1}{2}$ oraz pole prostokąta: $7\frac{1}{2}$. Wyznaczyć długość i szerokość tego prostokąta. Skryba opisuje detalicznie kroki w celu uzyskania rozwiązania. Oto one.

⁵Źródło: Victor J. Katz, *Stages in the history of algebra with implications for teaching*.

- Przepołowić $6\frac{1}{2}$ otrzymując $3\frac{1}{4}$.
- Podnieść uzyskaną liczbę do kwadratu uzyskując $10\frac{9}{16}$.
- Od uzyskanego pola odjąć dane pole prostokąta $7\frac{1}{2}$, uzyskując $3\frac{1}{16}$.
- Z uzyskanej liczby wyciągnąć pierwiastek: $1\frac{3}{4}$.
- Stwierdzić, że długość prostokąta wynosi $3\frac{1}{4} + 1\frac{3}{4} = 5$, podczas gdy szerokość to $3\frac{1}{4} - 1\frac{3}{4} = 1\frac{1}{2}$.

Nie bez powodu opisujemy rozwiązywanie językiem książki kucharskiej, ponieważ tak w istocie uczyono się matematyki w starożytnej Mezopotamii — przez akumulację rozwiązywanych zadań, a nie przez abstrahowanie i wyciąganie ogólnych zależności. O co chodzi w tym rozwiązyaniu? Skryba miał pewnie przed oczami następujący obrazek, przy założeniu, że szukane wielkości to x, y , a znane są $x+y = b$ oraz $xy = c$.



Rys. 2. Problem z tabliczki YBC 4663.

Czy Czytelnik widzi, że kluczem tej metody jest porównanie pól kwadratu o boku $b/2$ i wyjściowego prostokąta, przez znalezienie w nich (odpowiednich miejscach) wspólnego szarego prostokąta tak, że różnicą pól jest również pole kwadratu, i to liczby $(x-y)/2$? Obrazek powyższy ilustruje więc wzory:

$$x = \frac{b}{2} + \sqrt{\left(\frac{b}{2}\right)^2 - c}, \quad y = \frac{b}{2} - \sqrt{\left(\frac{b}{2}\right)^2 - c}.$$

Wzory te nie były oczywiście znane Babilończykom, jak również żadna ogólna metoda rozwiązywania takich zadań. Znana była niebanalna procedura postępowania przy konkretnych danych liczbowych.

Zupełnie inaczej wyglądała sytuacja w starożytnej Grecji, w której manipulacje algebraiczne wciąż wykonywane były w oparciu o obiekty geometryczne, w oparciu jednak o wyraźnie sformułowane aksjomaty. Oto przykład z *Elementów Euklidesa* — Twierdzenie II-5, znane nam jako wzór na kwadrat sumy.

Jeśli podzielimy na dwie równe części, a także na dwie nierówne części, to prostokąt zawarty w nierównych częściach wraz z kwadratem linii łączącej punkty przecięcia jest równy kwadratowi połowy tej prostej.

Jeśli pomyślimy o „nierównych segmentach” jako o x oraz y , a o długości początkowego odcinka jako b , wówczas twierdzenie zdaje się twierdzić, że:

$$xy + \left(\frac{x-y}{2}\right)^2 = \left(\frac{x+y}{2}\right)^2$$

i w ten sposób rozwiązać można układ równań $x+y = b$ oraz $xy = c$. Dokładniej, jeśli PODSTAWIMY c zamiast xy oraz b zamiast $x+y$, orzymamy

$$\left(\frac{x-y}{2}\right)^2 = \left(\frac{b}{2}\right)^2 - c \iff \frac{x-y}{2} = \sqrt{\left(\frac{b}{2}\right)^2 - c} \quad \text{skąd} \quad x = \frac{x+y}{2} + \frac{x-y}{2} = \frac{b}{2} + \sqrt{\left(\frac{b}{2}\right)^2 - c}.$$

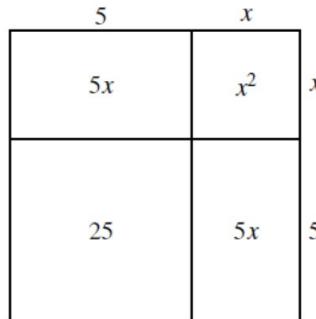
Stulecia później, arabscy matematycy cytowały będą powyższy wynik uzasadniając własne algorytmiczne rozwiązywanie równań kwadratowych. Sam Euklides ogólnych rozwiązań równań nie formułuje. Dowód jest czysto geometryczny. Idea zamiany zmiennych pochodząca z przesuwania pól pojawi się później.

O co więc chodzi z tym upraszczaniem i podejściem historycznym? Termin *algebra* pochodzi z tytułu arabskiego traktatu z początków IX wieku autorstwa perskiego uczonego Muhammada ibn Musa al-Chwarizmiego, bibliotekarza Bagdadzkiego *Domu Nauki*. Tytuł ten – *Zasady redukcji i przenoszenia* — ukrywa w sobie słowo *al-ğabr* oznaczające w języku arabskim *uzupełnianie, przenoszenie*, ale też *bilansowanie* (co miało praktyczne znaczenie). Dosłownie chodzi o operację przenoszenia wielkości z jednej strony na drugą, np. o zamianę równania $x^2 = 40x - 4x^2$ poprzez „al-ğabr” w równanie $5x^2 = 40x$. Przez redukcję (arab. *al-Muqabala*) autor rozumiał natomiast odjęcie tej samej dodatniej wielkości od obydwu stron równania, np. $x^2 + 5 = 40x + 4x^2$ zamieniamy na $5 = 40x + 3x^2$.

Ogólna metoda rozwiązywania równań kwadratowych, opracowana przez al-Khwarizmiego, prezentowana jest podobnie jak w matematyce greckiej w sposób geometryczny. Aby rozwiązać równanie

$$x^2 + 10x = 39$$

metodą arabską, przedstawiamy x^2 jako pole kwadratu o boku x , a $10x$ jako sumę pól dwóch prostokątów rozmiaru $5 \times x$. Dodatkowy kwadrat o polu 25 „uzupełnia” całą konfigurację do kwadratu o boku równym $x + 5$ o polu $25 + 39$, ponieważ 39 jest wartością wyrażenia $x^2 + 10x$. Stąd pole dużego kwadratu równe jest 64, czyli długość boku o długości $x + 5$ równa jest 8. Otrzymujemy zatem rozwiązanie $x = 3$.



Oczywiście matematyka arabska (ani grecka) nie uznawała liczb ujemnych, więc nie widziała też dodatkowego rozwiązania $x = -13$ tego równania. Ta konieczność unikania współczynników ujemnych w równaniach komplikowania rozważania algebraiczne. Nie było bowiem jednego ogólnego równania kwadratowego, ale aż trzy jego typy odpowiadające odpowiedniemu rozłożeniu dodatnich współczynników:

$$x^2 + ax = b, \quad x^2 = ax + b, \quad x^2 + b = ax.$$

Kluczowe jest jednak to, że owo zwijanie do kwadratu z algebraicznego punktu widzenia służy uproszczeniu równania przez podstawienie $t = x + 5$. Tego typu zamiana zmiennych przekształca równanie $x^2 + 10x = 39$ w $t^2 = 64$, a zatem w równanie prostsze. Podobną funkcję wzory skróconego mnożenia odgrywają w równaniach wyższego stopnia.

Przez kolejne stulecia także równania wyższych stopnia rozwiązywano przez podstawienie i eliminację. Cardano (powiemy o nim więcej za jakiś czas) wychodząc od podobnego geometrycznego ujęcia, od równania

$$x^3 + ax^2 + bx + c = 0$$

przechodzi za pomocą podstawienia

$$x = y - \frac{a}{3}$$

do równania typu

$$y^3 = py + q.$$

Dokonując kolejnej liniowej zamiany zmiennych postaci

$$y = u + v,$$

uzyskuje po lewej stronie:

$$(u^3 + v^3) + 3uv(u + v) = 3uvy + (u^3 + v^3),$$

które to wyrażenie równe jest prawej stronie wcześniejszego równania wtedy i tylko wtedy, gdy:

$$\begin{aligned} 3uv &= p, \\ u^3 + v^3 &= q. \end{aligned}$$

Eliminując v uzyskujemy równanie kwadratowe zmiennej u^3 postaci

$$u^3 + \left(\frac{p}{3u}\right)^3 = q,$$

o rozwiązaniach

$$\frac{q}{2} \pm \sqrt{\left(\frac{q}{2}\right)^2 - \left(\frac{p}{3}\right)^3}.$$

Rozumując symetrycznie, uzyskujemy te same wartości v^3 . Skoro $u^3 + v^3 = q$, to jeden z pierwiastków równy jest u^3 , a drugi — v^3 . Bez straty ogólności można przyjąć

$$u^3 = \frac{q}{2} + \sqrt{\left(\frac{q}{2}\right)^2 - \left(\frac{p}{3}\right)^3}, \quad v^3 = \frac{q}{2} - \sqrt{\left(\frac{q}{2}\right)^2 - \left(\frac{p}{3}\right)^3},$$

uzyskując

$$y = u + v = \sqrt[3]{\frac{q}{2} + \sqrt{\left(\frac{q}{2}\right)^2 - \left(\frac{p}{3}\right)^3}} + \sqrt[3]{\frac{q}{2} - \sqrt{\left(\frac{q}{2}\right)^2 - \left(\frac{p}{3}\right)^3}}.$$

Jeżeli wyobrażmy sobie teraz, że Cardano dokonywał tych manipulacji w języku geometrycznym, rozumiemy jak bardzo konieczne było przejście od geometrii do języka wyrażeń algebraicznych (zrobił to François Viète). A zatem ponownie — wzór skróconego mnożenia miał na celu dokonanie sensownego podstawienia i eliminacji zmiennych. Podobne przykłady można wskazać dla równań wyższych stopni.

Zanim przejdziemy do układów równań, zajrzyjmy jeszcze do XVIII wieku i do wielkiego Eulera⁶. Wkrótce poznacie Państwo twierdzenie udowodnione przez Gaussa na przełomie XVIII i XIX wieku mówiące, że każdy wielomian o współczynnikach rzeczywistych można rozłożyć na czynniki liniowe lub kwadratowe. Kilkadziesiąt lat wcześniej wieku ten był otwarty, a niektórzy wątpili w jego prawdziwość, wśród nich sam Leibniz. Co gorsze, Nicolas Bernoulli twierdził, że znalazł kontrprzykład, czyli wielomian

$$x^4 - 4x^3 + 2x^2 + 4x + 4.$$

Eulerowi udało się jednak znaleźć odpowiedni rozkład. W liście do Goldbacha z 1742 roku znalazł następujące czynniki kwadratowe wspomnianego wyżej wielomianu:

$$x^2 - \left(2 + \sqrt{4 + 2\sqrt{7}}\right)x + \left(1 + \sqrt{4 + 2\sqrt{7}} + \sqrt{7}\right),$$

$$x^2 - \left(2 - \sqrt{4 + 2\sqrt{7}}\right)x + \left(1 - \sqrt{4 + 2\sqrt{7}} + \sqrt{7}\right).$$

Fakt znalezienia tego rozkładu można chyba umieścić gdzieś pomiędzy cudem, a kuglarstwem. Euler był wielkim rachmistrzem. Właściwą intuicję daje wzór dwumianowy. Wspomniany wielomian przypomina bowiem rozwinięcie dwumianowe wyrażenia $(x - 1)^4$. Dokładniej:

$$(x - 1)^4 - (x^4 - 4x^3 + 2x^2 + 4x + 4) = 4x^2 - 8x - 3.$$

Pierwiastki uzyskanego wielomianu kwadratowego to $1 \pm \frac{\sqrt{7}}{2}$. Mamy więc:

$$x^4 - 4x^3 + 2x^2 + 4x + 4 = (x - 1)^4 - 4 \left((x - 1) - \frac{\sqrt{7}}{2} \right) \left((x - 1) + \frac{\sqrt{7}}{2} \right).$$

Podstawiając $u = (x - 1)^2$, uzyskujemy

$$x^4 - 4x^3 + 2x^2 + 4x + 4 = u^2 - 4u + 7.$$

Tak dochodzimy do rezultatu i rozumiemy już, że chodziło w istocie o wymyślenie zwykłego podstawienia typu $x = t + 1$ sprowadzającego problem do kwestii rozłożenia na czynniki wielomianu

$$t^4 - 4t^2 + 7.$$

⁶O historii tej przeczytać można także w szerszym kontekście w książce *Euler. The Master of Us All* Williama Dunhamia (Dolciani Mathematical Expositions Volume: 22; MAA1999).

Historia tak szerokiego tematu jak rozwiązywanie układów równań liniowych jest szalenie skomplikowana i w zasadzie nie jest możliwe choćby krótkie jej nakreślenie. Osobom zainteresowanym stosunkowo przystępnym tekstem polecam publicznie dostępny⁷ artykuł *How ordinary elimination became Gaussian elimination* autorstwa Josepha Greara (Historia Mathematica 38 (2011), 163–218). Pozwolę sobie przytoczyć tu krótki wstęp, odwołując się również do wspomnianej we wstępie książki Stillwella, wybierając fragmenty przydatne dla ogólnego spojrzenia na matematykę. Niecierpliwym polecam tekst: <https://www.fuw.edu.pl/kostecki/histmat.pdf>.

W istocie, ogólna teoria eliminacji dostała podwaliny w starożytnych Chinach w czasach dynastii Han (206 p.n.e. – 220 n.e.). Jest ona przedstawiona w słynnych Dziewięciu rozdziałach sztuki matematycznej, a jej ostateczna redakcja została dokonana przez Liu Hui w III wieku. W przeciwieństwie do greckiej matematyki, która budowała teorie wychodząc ze zbiorów aksjomatów, matematyka chińska szukała najbardziej ogólnych możliwych metod rozwiązywania zadań, niekoniecznie ubierając je w ogólne teorie.

Oczywiście nie pisano o układzie m równań z n niewiadomymi. Pokazano jednak na dostatecznej liczbie reprezentatywnych równań ogólną metodę odejmowania odpowiedniej wielokrotności niezerowego równania od pozostałych równań tak, by otrzymać układ w postaci trójkątnej (schodkowej). Ten typ rachunku bardzo odpowiadał urządzeniu upracowanemu do wykonywania kolejnych manipulacji na współczynnikach — analogicznych do tych, które wykonujemy na macierzach. Około XII wieku matematycy Chińscy opracowali podstawy metody wzmiankowanej w rozdziale o bazach Gröbnera, pozwalającej na eliminację zmiennej y z układu równań wielomianowych:

$$a_0(x)y^m + a_1(x)y^{m-1} + \dots + a_m(x) = 0, \quad b_0(x)y^m + b_1(x)y^{m-1} + \dots + b_m(x) = 0,$$

gdzie $a_i(x)$ oraz $b_j(x)$ są wielomianami zmiennej x . Oczywiście chodzi o przemnożenie pierwszego równania przez $b_0(x)$ i odjęcie drugiego przemnożonego przez $a_0(x)$. Uzyskujemy w ten sposób równanie stopnia $m - 1$. Nietrudno w ten sposób zamienić wyjściowy układ na prostszy — zamiast wielomianów stopnia m o zmiennej y (o współczynnikach wielomianowych) mamy dwa równania stopnia $m - 1$. W ten sposób możemy redukować (indukcyjnie) aż do uzyskania wielomianów zależnych tylko od x .

Jak się okazuje, opisana wyżej metoda odkryta została ponownie w XVII wieku na Zachodzie, przy okazji rozważania problemu znajdowania przecięć krzywych. Doprowadziło to do sformułowania metody eliminacji, którą niemal wiek później przeniesiono na grunt równań liniowych.

Jak to często bywa w historii matematyki, metodę eliminacji prowadzącą do rozwiązywania układów równań, którą tradycyjnie przypisuje się Gaussowi⁸, sam Książę Matematyków uważały za dobrze znany folklor (podobnie myślał o wielu innych tematach, na przykład o kwaterionach) — wiemy już czemu (była ogólniejsza teoria). Idea eliminacji pojawiała się w opublikowanych wbrew woli Newtona notatkach do prowadzonego w latach 1673–1687 wykładu z algebry w Cambridge, wydanych uroczyście w 1707 roku jako *Arithmetica Universalis*. Newton uważał, że publikowanie takich tekstów po 20 latach (a dalej tłumaczenie ich na różne języki) może prowadzić kogoś do wniosku, że są to jego najnowsze wyniki! Tymczasem w Anglii wydano w latach 1650–1750 kilkadziesiąt podręczników do algebry. Klasyczna złożliwość historii sprawiła, że ów podręcznik do algebry stał się jednym z najszerzej znanych i wpływowych dzieł matematycznych Newtona. Wielki Euler krytykował metodę eliminacji jako „niepolecaną”, jego następca Legendre nazywał ją „zwyczajną”. Gauss potraktował eliminację jak znane wszystkim narzędzie.

Pojęcie macierzy zawdzięczamy Brytyjczykom: Arthurowi Cayleyowi i Jamesowi Josephowi Sylvesterowi. Drugi z nich użył tego pojęcia po raz pierwszy w 1850 roku. Pierwszy natomiast, siedem lat później, napisał pierwszą rozprawę o macierzach (*Treatise on the Theory of Matrices*). Jak się nietrudno domyśleć oznacza to, że eliminacja przypisywana wcześniejszym autorom odbyła się bez użycia pojęcia macierzy.

Kto zatem, i dlaczego, przypisał Gaussowi autorstwo metody eliminacji? Stało się to głównie za sprawą rozwoju maszyn liczących. Kluczowe było to, że algorytmy opisane w pracy Gaussa były po prostu niezwykle użyteczne dla wykorzystania praprzodków komputerów. Stosowano je, z nielicznymi modyfikacjami jeszcze do II Wojny Światowej. Notacja „klamrowa” wprowadzona przez Gaussa podkreślała kolejność zmiennych i była bardzo wygodna. Pojęcie „algorytmu eliminacji Gaussa” zastosował po raz pierwszy, jak się wydaje, Alan Turing, który przez dwa tygodnie rozwiązywał, z pomocą „komputera biurkowego” układ 18 równań w roku 1946. Po II WŚ, pojęcie to stopniowo wchodziło do programów nauczania.

⁷Zarówno na ArXiV: <https://arxiv.org/pdf/0907.2397.pdf>, jak i na stronach Elseviera.

⁸Gauss, C. F., *Theoria Motus Corporum Coelestium in Sectionibus Conicis Solum Ambientium*, 1809.

Rozdział 3

Działania i ich własności. Ciała

3.1 Wykład 3

Na ostatnim wykładzie rozważaliśmy układy równań liniowych o współczynnikach rzeczywistych. Są to ciągi równań postaci $a_1x_1 + a_2x_2 + \dots + a_nx_n = b$, gdzie $a_1, a_2, \dots, a_n, b \in \mathbb{R}$. Sprawdzenie, że (s_1, \dots, s_n) jest rozwiązaniem równania wyżej dokonywaliśmy przez wykonanie **działań dodawania** i **mnożenia** w \mathbb{R} postaci $a_1 \cdot s_1 + a_2 \cdot s_2 + \dots + a_n \cdot s_n$, żądając, by wynikiem było b . Na tym wykładzie spróbujemy odpowiedzieć na pytanie: czy zamiast zbioru \mathbb{R} z działaniami $+$ oraz \cdot można rozważyć układy równań liniowych, gdzie współczynnikami są **inne zbiory X z innymi działaniami** dodawania i mnożenia \boxplus, \boxtimes tak, by metoda eliminacji Gaussa dalej działała? Na pewno nie mamy tu pełnej dowolności. Przecież choćby równanie liniowe $4x = 2$ nie ma rozwiązania w zbiorze liczb całkowitych.

Definicja 3.1.1: Działanie

Niech X będzie zbiorem niepustym. Przez X^n rozumieć będziemy zbiór ciągów postaci

$$(x_1, x_2, \dots, x_n), \quad \text{gdzie } x_i \in X, \text{ dla } 1 \leq i \leq n.$$

DZIAŁANIEM n -ARGUMENTOWYM na zbiorze X nazywamy każdą funkcję $\omega : X^n \rightarrow X$.

Najczęściej rozważanymi działaniami są działania dwuargumentowe. Oto ich przykłady.

zbiór X	działanie ω
liczby rzeczywiste/wymierne/całkowite/naturalne	dodawanie/mnożenie
liczby rzeczywiste	$a \boxplus b = a + b + ab$
zbiór podzbiorów danego zbioru	suma/część wspólna
zbiór funkcji ze zbioru X na zbiór X	złożenie

Kluczowym elementem definicji działania jest żądanie, by nie wyprowadzało ono poza zbiór, na którym jest określone. A więc na przykład odejmowanie nie jest działaniem w zbiorze dodatnich liczb całkowitych, bo $1 - 3 \notin \mathbb{Z}_+$.

Definicja 3.1.2: Łączność i przemienność

Niech $*$ będzie działaniem dwuargumentowym na zbiorze X . Mówimy, że $*$ jest:

- **ŁĄCZNE**, jeśli dla każdych $a, b, c \in X$ mamy $(a * b) * c = a * (b * c)$,
- **PRZEMIENNE**, jeśli dla każdych $a, b \in X$ mamy $a * b = b * a$.

Przyjrzyjmy się ponownie kilku przykładom.

- działanie $a \boxplus b = a^2 + b^2$ na zbiorze \mathbb{R} nie jest łączne, bo $(1 \boxplus 2) \boxplus 3 = 34$, zaś $1 \boxplus (2 \boxplus 3) = 170$.
- złożenie w zbiorze bijekcji (czyli odwzorowań 1-1 i „na”) zbioru \mathbb{R} nie jest przemienne.

Definicja 3.1.3: Ciało

CIAŁEM nazywamy piątkę $(K, \boxplus, \boxtimes, 0, 1)$, gdzie K jest zbiorem przynajmniej dwuelementowym z wyróżnionymi elementami $0 \neq 1$, zwanymi ZEREM i JEDYNKĄ, zaś \boxplus, \boxtimes są dwuargumentowymi działaniami zwanyimi **dodawaniem** i **mnożeniem**, spełniającymi następujące AKSJOMATY CIAŁA:

1)	$(a \boxplus b) \boxplus c = a \boxplus (b \boxplus c)$	$\forall_{a,b,c \in K}$	łączność dodawania
2)	$a \boxplus b = b \boxplus a$	$\forall_{a,b \in K}$	przemienność dodawania
3)	$a \boxplus 0 = a = 0 \boxplus a$	$\forall_{a \in K}$	własność elementu 0
4)	$a \boxplus b = 0 = b \boxplus a$	$\forall_{a \in K} \exists_{b \in K}$	element przeciwny
5)	$(a \boxtimes b) \boxtimes c = a \boxtimes (b \boxtimes c)$	$\forall_{a,b,c \in K}$	łączność mnożenia
6)	$a \boxtimes b = b \boxtimes a$	$\forall_{a,b \in K}$	przemienność mnożenia
7)	$a \boxtimes 1 = 1 \boxtimes a = a$	$\forall_{a \in K}$	własność elementu 1
8)	$a \boxtimes b = b \boxtimes a = 1$	$\forall_{a \in K \setminus \{0\}} \exists_{b \in K}$	odwrotność dla $a \neq 0$
9)	$a \boxtimes (b \boxplus c) = (a \boxtimes b) \boxplus (a \boxtimes c)$	$\forall_{a,b,c \in K}$	rozdzielnosc \boxtimes wzgl. \boxplus

Teoria ciał jest bardzo szeroką dziedziną algebry abstrakcyjnej i w trakcie studiów będziecie Państwo poznawać różne nowe jej aspekty. W ramach naszego wykładu skupimy się na podstawowych przykładach.

- Piątka $(\mathbb{R}, +, \cdot, 0, 1)$, jest ciałem, gdzie $+, \cdot$ – standardowe dodawanie i mnożenie liczb rzeczywistych.
- Piątka $(\mathbb{Q}, +, \cdot, 0, 1)$ jest ciałem, gdzie $+, \cdot$ – standardowe dodawanie i mnożenie liczb wymiernych.
- Piątka $(\mathbb{Z}, +, \cdot, 0, 1)$ nie jest ciałem, bo żaden niezerowy element poza $-1, 1$ nie ma elementu odwrotnego. Warto odnotować, że wszystkie inne aksjomaty poza (8) są przez $(\mathbb{Z}, +, \cdot, 0, 1)$ spełnione.

Uwaga 3.1.4

Niech K będzie ciałem. Dla każdego elementu $a \in K$ istnieje dokładnie jeden element przeciwny do a . Dla każdego niezerowego elementu $b \in K$ istnieje dokładnie jeden element odwrotny do b .

Dowód. Wykażemy jedynie jednoznaczność elementu przeciwnego. Drugą część dowodzi się analogicznie. Założmy, że dla pewnych elementów x, x' ciała K dla dowolnego $a \in K$ mamy $a \boxplus x = 0 = a \boxplus x'$. Korzystamy dalej z aksjomatów ciała i powyższej równości:

$$\begin{aligned} x &= x \boxplus 0 && \text{(aksjomat 3 –wł. elementu 0)} \\ &= x \boxplus (a \boxplus x') && \text{(równość wyżej)} \\ &= (x \boxplus a) \boxplus x' && \text{(aksjomat 1 –łączność } \boxplus\text{)} \\ &= x' \boxplus (a \boxplus x) && \text{(aksjomat 2 –przemienność } \boxplus\text{)} \\ &= x' \boxplus 0 && \text{(równość wyżej)} \\ &= x'. && \text{(aksjomat 3 –wł. elementu 0)} \end{aligned}$$

□

Uwaga 3.1.5

Niech K będzie ciałem. Wówczas jeśli $a, b \in K$ oraz $ab = 0$, to $a = 0$ lub $b = 0$.

Dowód. Niech $x, y \in K$. Wyprowadzimy kolejne wnioski z aksjomatów ciała.

- Jeśli $x + y = x$, to $-x + (x + y) = (-x + x) + y = 0 + y = y = -x + x = 0$.
- Mamy $0 \cdot x = (0 + 0) \cdot x = 0 \cdot x + 0 \cdot x$. Zatem na mocy (i) mamy $0 \cdot x = 0$.

Jeśli $a \neq 0$, to na mocy (ii) mamy $0 = a^{-1} \cdot 0 = a^{-1} \cdot (a \cdot b) = (a^{-1} \cdot a) \cdot b = b$.

□

W dalszym ciągu, o ile nie prowadzi to do nieporozumień, wprowadzamy następujące umowy: dodawanie i mnożenie w ciele K oznaczamy odpowiednio jako $+$ oraz \cdot , przy czym znak mnożenia może być pomijany.

Przez a^n rozumiemy wynik n -krotnego przemnożenia przez siebie elementu a . Przyjmujemy też:

oznaczenie	definicja
$-a$	element odwrotny do a ze względu na $+$
a^{-1}	element odwrotny do a ze względu na \cdot
$a - b$	element postaci $a + (-b)$
$\frac{a}{b}$	element postaci $a \cdot (b^{-1})$

Wszystkie pojęcia zdefiniowane na poprzednim wykładzie dla układów równań o współczynnikach w \mathbb{R} (równanie liniowe, układy równoważne, rozwiązania ogólne, macierz układu, operacje elementarne na wierszach, macierze w postaci schodkowej i zredukowanej) przenoszą się na układy o współczynnikach w dowolnym ciele K . Podobnie, opisana na poprzednim wykładzie metoda eliminacji Gaussa stosuje się do układów równań liniowych o współczynnikach w dowolnym ciele K . Mianowicie zachodzi twierdzenie:

Twierdzenie 3.1.6

Niech K będzie ciałem. Każdą macierz $A = M_{m \times n}(K)$ można za pomocą operacji elementarnych (1)-(2) na wierszach doprowadzić do postaci schodkowej. Każdą macierz $A = M_{m \times n}(K)$ można za pomocą operacji elementarnych (1)-(3) na wierszach doprowadzić do postaci zredukowanej. Każdy niesprzeczny układ równań liniowych o współczynnikach w K ma rozwiązanie ogólne. Aby je znaleźć wystarczy sprowadzić macierz tego układu do postaci schodkowej zredukowanej elementarnymi operacjami na wierszach, a następnie z otrzymanej macierzy otrzymać rozwiązanie ogólne.

Warto prześledzić dowody z poprzedniego wykładu, by zobaczyć w jaki sposób własności liczb rzeczywistych można w nich zastąpić przez kolejne aksjomaty ciała. Kluczowym elementem, który się wyłoni jest istnienie elementu przeciwnego oraz odwrotnego, niezbędnych m.in. do wykonania eliminacji Gaussa.

Celem tego wykładu nie jest systematyczny wykład teorii ciał (dotknimy tego zagadnienia w dodatkach) ale omówienie najważniejszych przykładów i prostych konstrukcji ciał.

Definicja 3.1.7: Podciało

Podzbiór $L \subset K$ nazywamy *podciągiem* ciała K , jeśli $0, 1 \in L$ oraz dla każdych $a, b \in L$ zachodzi $a + b \in L$, $ab \in L$, a także $-a \in L$ oraz, dla $a \neq 0$, gdy $a^{-1} \in L$.

Przykład. Niech \mathbb{Q} oznacza zbiór wszystkich liczb wymiernych i niech

$$\mathbb{Q}(\sqrt{2}) = \{a + b\sqrt{2} \mid a, b \in \mathbb{Q}\}.$$

Wówczas \mathbb{Q} oraz $\mathbb{Q}(\sqrt{2})$ są podciągami ciała \mathbb{R} .

Istotnie, zauważmy, że $a + b\sqrt{2} = c + d\sqrt{2}$ wtedy i tylko wtedy, gdy $a = c$ oraz $b = d$ (wynika to z niewymierności liczby $\sqrt{2}$). Co więcej, dla dowolnych $a + b\sqrt{2}$ oraz $c + d\sqrt{2}$ należących do $\mathbb{Q}(\sqrt{2})$ liczby

$$(a + b\sqrt{2}) + (c + d\sqrt{2}) = (a + c) + (b + d)\sqrt{2}, \quad (a + b\sqrt{2})(c + d\sqrt{2}) = (ac + 2bd) + (ad + bc)\sqrt{2}$$

są elementami $\mathbb{Q}(\sqrt{2})$. Prosty rachunek, między innymi zauważenie równości:

$$\frac{1}{a + b\sqrt{2}} = \frac{a - b\sqrt{2}}{a^2 - 2b^2}$$

pokazuje, że zbiór $\mathbb{Q}(\sqrt{2})$ jest ciałem.

Uwaga 3.1.8

Podciąg ciała K (z działaniami takimi jak w K) jest ciałem.

Pokażemy teraz, że dla dowolnej liczby pierwszej p istnieje ciało mające p elementów. Dla każdego $k \in \mathbb{Z}$ i każdego $n \in \mathbb{N}$ niech $k \pmod{n}$ oznacza resztę z dzielenia k przez n , to znaczy taką liczbę $l \in \mathbb{Z}$, że $0 \leq l < n$ oraz $k = an + l$, dla pewnego $a \in \mathbb{Z}$.

Definicja 3.1.9: Ciało reszt z dzielenia modulo p

Niech p będzie liczbą pierwszą. Ciało \mathbb{Z}_p jest to zbiór $\{0, 1, \dots, p-1\}$ z działaniami:

- dodawanie modulo p dane wzorem: $(a, b) \mapsto a + b \pmod{p}$,
- mnożenie modulo p dane wzorem: $(a, b) \mapsto a \cdot b \pmod{p}$

Należy sprawdzić, że tak określone działania spełniają wszystkie aksjomaty ciała. W tym celu przypominiemy pojęcia i fakty związane z działaniami na resztach.

Definicja 3.1.10: Relacja przystawania liczb całkowitych modulo k

Mówimy, że liczby całkowite a, b PRZYSTAJĄ MODULO k , co oznaczać będziemy przez $a \equiv b \pmod{k}$ lub krótko $a \equiv_k b$, jeśli liczba k jest dzielnikiem liczby $a - b$.

Przykład. Mamy $14 \equiv_3 2$, ponieważ 3 jest dzielnikiem liczby $14 - 2 = 12$.

Uwaga 3.1.11

Jeśli dla liczb całkowitych a, b, c, d oraz niezerowej liczby całkowitej k zachodzą warunki $a \equiv_k b$ oraz $c \equiv_k d$, to $a + c \equiv_k b + d$ oraz $ac \equiv_k bd$.

Dowód. Skoro liczba k jest dzielnikiem zarówno $b - a$, jak i $c - d$, to jest też dzielnikiem liczby $b + d - (a + c)$. Stąd $a + c \equiv_k b + d$. Liczba k jest dzielnikiem $c(b - a) + b(d - c) = bd - ac$, skąd $ac \equiv_k bd$. \square

Zgodnie z uzasadnioną wyżej własnością uzasadnić można bez trudu, że \mathbb{Z}_p spełnia wszystkie aksjomaty ciała, poza aksjomatem istnienia odwrotności dla dowolnego niezerowego elementu.

Uwaga 3.1.12

Niech p będzie liczbą pierwszą oraz niech k będzie dowolną liczbą całkowitą niepodzielną przez p . Wówczas istnieje taka liczba całkowita s , że $ks \equiv_p 1$.

Dowód. Niech r będzie niezerową resztą z dzielenia liczby k przez p . Rozważmy zbiór wszystkich niezerowych reszt z dzielenia przez p , czyli $1, \dots, p-1$. Przemuńźmy każdy z elementów tego zbioru przez r , otrzymując $r \cdot 1, \dots, r(p-1)$. Twierdzimy, że uzyskane elementy dają parami różne reszty z dzielenia przez p . Gdyby powiem, dla pewnych $1 \leq k < l \leq p-1$ zachodziła kongruencja $r \cdot k \equiv_p r \cdot l$, to by oznaczało, że liczba $rk - rl = r(k - l)$ jest podzielna przez p . Skoro jednak r jest liczbą niepodzielną przez p , zaś p liczbą pierwszą, to liczba $k - l$ jest podzielna przez p . To jest jednak niemożliwe, gdyż $|l - k| < p$ oraz $k \neq l$.

Uzasadniliśmy zatem, że liczby $r \cdot 1, \dots, r(p-1)$ dają parami różne reszty z dzielenia przez p i żadna z tych reszt nie jest równa 0. Skoro tych liczb jest $p-1$, to dają one wszystkie możliwe niezerowe reszty z dzielenia przez p . Istnieje więc liczba s taka, że $rs \equiv_p 1$. \square

Wniosek 3.1.13

Każdy niezerowy element ciała \mathbb{Z}_p posiada element odwrotny.

Przykłady

- W ciele $\mathbb{Z}_2 = \{0, 1\}$ działania dodawania i mnożenia wyglądają następująco: $0 + 0 = 1 + 1 = 0$, $0 + 1 = 1 + 0 = 1$, $0 \cdot 0 = 0 \cdot 1 = 1 \cdot 0 = 0$ oraz $1 \cdot 1 = 1$.
- W ciele \mathbb{Z}_3 mamy $1 + 2 = 0$, więc 1 i 2 są elementami wzajemnie przeciwnymi.
- W ciele \mathbb{Z}_5 mamy $2 \cdot 3 = 1$, a zatem $2^{-1} = 3$ oraz $3^{-1} = 2$.

Osobnym problemem jest kwestia znajdowania elementu odwrotnego modulo p . Proste rozwiązanie tego zagadnienia dostarcza procedura rozszerzająca działanie algorytmu Euklidesa, służącego do wyznaczania największego wspólnego dzielnika pary liczb całkowitych.

Twierdzenie 3.1.14: Algorytm Euklidesa

Niech a, b będą niezerowymi liczbami całkowitymi. Rozważmy ciągi liczb całkowitych $q_0, q_1, q_2, q_3, \dots$ oraz r_1, r_2, r_3, \dots spełniający warunki:

$$\begin{aligned} a &= b \cdot q_0 + r_1, \quad 0 \leq r_1 < b \\ b &= r_1 \cdot q_1 + r_2, \quad 0 \leq r_2 < r_1 \\ r_1 &= r_2 \cdot q_2 + r_3, \quad 0 \leq r_3 < r_2 \\ r_2 &= r_3 \cdot q_3 + r_4, \quad 0 \leq r_4 < r_3, \\ &\dots \end{aligned}$$

Niech n będzie taką liczbą, że $r_n \neq 0$ oraz $r_{n+1} = 0$. Wówczas $NWD(a, b) = r_n$.

Algorytm Euklidesa pozwala na znalezienie liczb całkowitych x, y , spełniających równość

$$ax + by = NWD(a, b).$$

Zauważmy, że gdy $NWD(a, b) = 1$, to liczby x, y spełniają równości

$$ax \equiv_b 1 \quad \text{oraz} \quad by \equiv_a 1.$$

Stąd, jeśli $a \in Z_p$, to reszta z dzielenia liczby x przez p jest elementem odwrotnym do a w tym ciele.

Przykład. Wyznaczmy odwrotność modulo k , dla $p = 13$ oraz $a = 9$. Z algorytmu Euklidesa wiemy, że

$$13 = 1 \cdot 9 + 4, \quad 9 = 2 \cdot 4 + 1,$$

skąd

$$1 = 9 - 2 \cdot 4 = 9 - 2 \cdot (13 - 9) = 3 \cdot 9 - 2 \cdot 13.$$

Stąd $9 \cdot 3 \equiv_{13} 1$, czyli liczba 3 jest odwrotnością liczby 9 modulo 13. W konsekwencji, w ciele Z_{13} elementy 3 oraz 9 są swoimi wzajemnymi odwrotnościami.

* * *

Spróbujmy rozwiązać układ równań liniowych o współczynnikach w ciele Z_3 :

$$\begin{cases} x_1 + x_2 = 2 \\ 2x_1 + x_2 = 1 \end{cases} \quad \text{o macierzy} \quad \left[\begin{array}{cc|c} 1 & 1 & 2 \\ 2 & 1 & 1 \end{array} \right] \in M_{3 \times 2}(Z_3).$$

Odejmujemy pierwszy wiersz przemnożony przez 2. Jaką on ma postać? Otóż jest to wiersz postaci $2 \cdot 1 | 1$, bo działania wykonujemy modulo 3. A zatem po tej operacji mamy (to samo, co po dodaniu wierszy):

$$\left[\begin{array}{cc|c} 1 & 1 & 2 \\ 0 & 2 & 0 \end{array} \right].$$

Cóż więc pozostaje? Przemożyc drugi wiersz przez... odwrotność 2, czyli 2 (bo $2 \cdot 3 \equiv 1$). A zatem mnożymy drugi wiersz przez 2 i odejmujemy od pierwszego. Dostajemy:

$$\left[\begin{array}{cc|c} 1 & 0 & 2 \\ 0 & 1 & 0 \end{array} \right].$$

A zatem rozwiązaniem tego układu jest para $(x_1, x_2) = (2, 0)$. Ten układ był tak prosty, że to rozwiązanie byłoby widoczne i bez sprowadzania macierzy do postaci schodkowej. **Układ równań liniowych o współczynnikach w ciele skończonym ma zawsze skończenie wiele rozwiązań!** Rozwiązania ogólne mogą być jednak nadal reprezentowane przez zmienne zależne i niezależne. Np. zbiorem rozwiązań równania $x_1 + x_2 = 0$ o współczynnikach w ciele skończonym Z_p są wszystkie pary $\{(-t, t) \mid t \in Z_p\}$, a więc równanie to ma p^2 rozwiązań. Równanie $x_1 + x_2 + x_3 = 0$ ma p^3 rozwiązań w ciele Z_p .

3.2 Wybór przykładowych pytań

Proszę się nie uczyć na pamięć odpowiedzi czy wyjaśnienie! Takich pytań można ułożyć znacznie więcej i służą one jedynie sprawdzeniu podstawowego zrozumienia treści zajęć. Jeśli odpowiedź na dowolne z poniższych pytań sprawia Państwu problem, proszę pomyśleć który fragment notatek dotyczy tego pytania, wrócić do ich lektury, a w razie dalszych trudności poprosić prowadzącego o wskazówkę.

1. Podaj przykłady zbiorów z działaniami dodawania i mnożenia, które nie są ciałami.
2. Czy dzielenie jest działaniem dwuargumentowym w zbiorze liczb wymiernych?
3. Podaj przykład działania 3-argumentowego.
4. Czy działanie przypisujące parze (a, b) liczb naturalnych liczbę a^b jest łączne?
5. Podaj przykłady dwóch funkcji $f, g : \mathbb{R} \rightarrow \mathbb{R}$, które nie spełniają równości $f \circ g = g \circ f$.
6. Element a ciała K spełnia warunek $a + b = b + a = b$, dla każdego $b \in K$. Czy $a = 0$?
7. Ile jest równy element $(1 + \sqrt{2})^{-1}$, a ile jest równy element $(1 + \sqrt[3]{2})^{-1}$ w ciele \mathbb{R} ?
8. Niech n będzie liczbą naturalną. Ile może być równa reszta z dzielenia liczby 7^n przez 3?
9. Oblicz iloczyn $3 \cdot 5$ w ciele \mathbb{Z}_7 .
10. Oblicz iloczyn 4×4 w ciele \mathbb{Z}_5 .
11. Oblicz iloczyn 2×7 w ciele \mathbb{Z}_{11} .
12. Wskaż element 2^{-1} w ciele \mathbb{Z}_5 .
13. Wskaż element 4^{-1} w ciele \mathbb{Z}_7 .
14. Wskaż element 9^{-1} w ciele \mathbb{Z}_{11} .
15. Wskaż element $-4 \cdot 3^{-1}$ w ciele \mathbb{Z}_{11} .
16. Oblicz sumę wszystkich elementów ciała \mathbb{Z}_{13} .
17. Wskaż wszystkie rozwiązania równania $x_1 + x_2 = 1$ w ciele \mathbb{Z}_5 .
18. Oblicz iloczyn 2^{100} w ciele \mathbb{Z}_3 .
19. Rozważmy równanie liniowe o współczynnikach w \mathbb{Z}_2 postaci

$$x_1 + \dots + x_5 = 0.$$

Ile rozwiązań ma to równanie?

20. Wykonaj operację elementarną dodawania do drugiego wiersza pierwszego wiersza pomnożonego przez 3, dla układu równań o współczynnikach w ciele \mathbb{Z}_5

$$\begin{cases} x_1 + 2x_2 + 3x_3 + 4x_4 &= 3 \\ x_1 + x_2 + 3x_3 &= 1 \end{cases}$$

21. Ile jest macierzy 2×3 o wyrazach w ciele \mathbb{Z}_3 ?

3.3 Zadania do samodzielnej pracy

Zadania ilustrujące umiejętności niezbędne do zaliczenia przedmiotu oznaczone są symbolem ♠. Przy tych zadaniach dodany jest opis umiejętności, które sprawdzają.

1. (♠ Wykonywanie działań w ciałach reszt modulo p)

Rozwiąż trzykrotnie równanie liniowe

$$4x + 3 = 0$$

zakładając, że współczynniki tego równania pochodzą kolejno z ciała: \mathbb{Z}_5 , \mathbb{Z}_7 oraz \mathbb{Z}_{13} .

2. (♠ Rozwiązywanie układów równań liniowych o współczynnikach w \mathbb{Z}_p)

Znajdź rozwiązanie ogólne następującego układu równań liniowych o współczynnikach w \mathbb{Z}_5 :

$$\begin{cases} 2x_1 + 3x_2 + x_3 + 4x_4 = 1 \\ 3x_1 + x_2 + 2x_3 + 4x_4 = 2 \\ 3x_1 + 3x_2 + x_3 + 3x_4 = 1. \end{cases}$$

3. Niech p będzie liczbą pierwszą. Rozważmy układ równań o współczynnikach w ciele \mathbb{Z}_p postaci

$$\begin{cases} 5x + 3y = 4 \\ 3x + 6y = 1 \end{cases}$$

Rozstrzygnij, dla jakich p układ ten nie ma rozwiązań/ma dokładnie jedno rozwiązanie/ma więcej niż jedno rozwiązanie?

4. Wyznacz element odwrotny do elementu 2 w ciele \mathbb{Z}_p , dla $p > 2$.

5. Uzasadnij, że dla dowolnych elementów $a, b \in \mathbb{Z}_p$ mamy $(a + b)^p = a^p + b^p$.

6. Niech $\mathbb{Z}_p^* = \{1, 2, \dots, p - 1\}$ będzie zbiorem niezerowych elementów ciała \mathbb{Z}_p . Wykaż, że

- Dla $p > 2$ suma wszystkich elementów zbioru \mathbb{Z}_p^* wynosi 0.
- Iloczyn wszystkich kwadratów elementów zbioru \mathbb{Z}_p^* wynosi 1.
- Iloczyn wszystkich elementów zbioru \mathbb{Z}_p^* wynosi -1 .
- Dla każdego elementu $a \in \mathbb{Z}_p^*$ mamy $a^{p-1} = 1$.
- Niech $a \in \mathbb{Z}_p^*$ oraz niech r będzie najmniejszą liczbą naturalną spełniającą warunek $a^r = 1$. Wykaż, że jeśli $a^n = 1$, to liczba r jest dzielnikiem liczby n .

7. Wykonaj następujące działania w ciele \mathbb{Z}_{13}

- $1 + 3^{3001} + 9^{3001}$

- $3^{2002} + 4^{4001}$

- $7 \cdot 11^{106} + 4$

8. Wykaż, że dla każdego $c \in \mathbb{Z}_p$ istnieją takie $a, b \in \mathbb{Z}_p$, że $a^2 + b^2 = c$.

9. Niech K będzie ciałem i niech $a, b, c \in K$. Wykaż, że:

- Jeśli $a + c = b + c$, to $a = b$.
- Jeśli $ac = bc$ i $c \neq 0$, to $a = b$.
- $0 \cdot a = 0$ oraz $(-1) \cdot a = -a$.
- $(-a) \cdot b = a \cdot (-b) = -ab$ oraz $(-a) \cdot (-b) = ab$.

10. Wykaż, że jeśli K jest podciałem ciała \mathbb{Q} , to $K = \mathbb{Q}$.

11. Uzasadnij, że jeśli wprowadzimy na zbiorze par $\{(p, q) \mid p, q \in \mathbb{Q}\}$ działania

$$(a, b) + (c, d) = (a + c, b + d), \quad (a, b) \cdot (c, d) = (ac + 5bd, ad + bc),$$

to uzyskamy ciało. Które elementy tego zbioru są zerem i jedynką?

12. W zbiorze liczb rzeczywistych określono działania \boxplus i \boxtimes postaci

$$a \boxplus b = a + b + 1, \quad a \boxtimes b = a + b + ab,$$

gdzie po prawych stronach wykonywane jest zwykłe dodawanie i mnożenie liczb rzeczywistych. Wykaż, że \mathbb{R} z działaniami \boxplus i \boxtimes jest ciałem. Jaki jest element zerowy oraz jedynka tego ciała?

3.4 Uzupełnienie. Charakterystyka ciała. Ciąła skończone

Przypomnijmy, że każde ciało K zawiera element 1, czyli element neutralny mnożenia, oraz wszystkie elementy postaci:

$$\underbrace{1 + 1 + \dots + 1}_m,$$

będące sumami m jedynek, gdzie $m \in \mathbb{N}$. Przyjrzymy się sytuacjom, gdy suma ta jest równa 0.

Uwaga 3.4.1

Niech K będzie ciałem i założymy, że m jest najmniejszą liczbą naturalną, taką że $\underbrace{1 + 1 + \dots + 1}_m = 0$ (zakładamy, że m istnieje). Wówczas m jest liczbą pierwszą.

Dowód. Oczywiście $1 \neq 0$, więc mamy $m \geq 2$. Przypuśćmy nie wprost, że $m = k \cdot l$, gdzie $k > 1$ oraz $l > 1$ są liczbami całkowitymi. Wówczas:

$$\underbrace{1 + 1 + \dots + 1}_m = (\underbrace{1 + 1 + \dots + 1}_k) \cdot (\underbrace{1 + 1 + \dots + 1}_l).$$

Skoro lewa strona powyższej równości to 0, to zgodnie z Uwagą 3.1.5, jeden z czynników po prawej stronie jest równy 0, co przeczy minimalności m . \square

Definicja 3.4.2: Charakterystyka ciała

Jeśli K jest ciałem i istnieje liczba pierwsza p , taka że

$$\underbrace{1 + 1 + \dots + 1}_p = 0,$$

to mówimy, że ciało K ma charakterystykę p (lub jest charakterystyki p). W przeciwnym przypadku mówimy, że ciało K ma charakterystykę 0.

Przykłady

- Ciąła \mathbb{Q} , \mathbb{R} oraz \mathbb{C} są charakterystyki 0. Ciało \mathbb{Z}_p jest charakterystyki p .
- Każde ciało o skończeniu wielu elementach ma dodatnią charakterystykę. Skoro bowiem liczba elementów postaci $\underbrace{1 + 1 + \dots + 1}_m$ jest skończona, to dla pewnych $k > l$ mamy

$$\underbrace{1 + 1 + \dots + 1}_k = \underbrace{1 + 1 + \dots + 1}_l.$$

Mamy stąd $\underbrace{1 + 1 + \dots + 1}_{k-l} = 0$. Teza wynika zatem z Uwagi 3.4.1.

Wniosek 3.4.3

Niech K będzie ciałem skończonym. Istnieje wówczas taka liczba pierwsza p oraz taka dodatnia liczba naturalna k , że ciało K ma p^k elementów.

Dowód. Skoro ciało K jest skończone, to jest charakterystyki dodatniej p , gdzie p jest liczbą pierwszą. Dla dowolnego elementu $x \in K$, niech $nx = \underbrace{x + x + \dots + x}_n$, który to element nazywać będziemy wielokrotnością x . Możemy przy tym przyjąć, że interesują nas jedynie liczby $n \in \{0, \dots, p-1\}$.

Dla ustalonych $x_1, \dots, x_s \in K$ rozważmy zbiór elementów

$$L(\{x_1, \dots, x_s\}) = \{n_1 x_1 + \dots + n_s x_s \mid n_i \in \{0, 1, \dots, p-1\}\} \subseteq K.$$

Skoro ciało K jest skończone, to możemy wybrać taki układ elementów x_1, \dots, x_k , aby powyższy zbiór był równy caemu ciału K . Rozważmy minimalny względem inkluzji podzbiór $S = \{x_1, \dots, x_k\}$, taki że $L(S) = K$. Innymi słowy — założymy, że nie istnieje podzbiór właściwy $S' \subsetneq S$ zbioru S , że $L(S') = K$. W szczególności zakładamy, że zbiór S nie zawiera elementu zerowego ciała K . Twierdzimy, że zbiór $L(S)$ ma p^k elementów.

Istotnie, żadne dwa różne elementy postaci $n_1 x_1 + \dots + n_k x_k$ nie mogą być równe jako elementy ciała K . W przeciwnym razie, jeśli $n_1 x_1 + \dots + n_k x_k = m_1 x_1 + \dots + m_k x_k$, dla pewnych $n_i, m_j \in \{0, \dots, p-1\}$, to biorąc najmniejsze i , że $m_i \neq n_i$ wiemy, że jedna z tych liczb jest większa, powiedzmy $m_i > n_i$. Wykażemy, że przeczy to wyborowi zbioru S .

Rozważmy dwa przypadki. Po pierwsze, jeśli $k = 1$ lub wszystkie inne współczynniki m_j, n_j są zerowe, dla $j \neq i$, to równość elementów zbioru $L(S)$ ma postać $m_i x_i = n_i x_i$, czyli mamy $(m_i - n_i)x_i = 0$, a stąd $x_i = 0$ (Uwaga 2.1.3), wbrew przypuszczeniu o minimalności B .

Jeśli choć jeden ze współczynników m_j, n_j jest różny od 0, dla $j \neq i$, to mamy $k \geq 2$ oraz stwierdzamy, że element $(m_i - n_i)x_i$ jest sumą wielokrotności pewnych elementów x_j , dla $j \neq i$. To jednak oznacza, że każdy element ciała K można przedstawić w postaci sumy wielokrotności elementów zbioru $S \setminus \{x_i\}$, co ponownie przeczy minimalności S . \square

Uzasadnimy teraz, że istnieje dokładnie jedno ciało czteroelementowe — to znaczy: jednoznacznie określmy działania na czterech elementach $\{0, 1, a, b\}$, spełniające aksjomaty ciała. Skorzystamy z tez Zadań 5a) i 5b) w podrozdziale 3.3.

Zauważmy, że w tabelkach opisujących działania w ciele skończonym K w każdym wierszu i w każdej kolumnie występuwać muszą wszystkie elementy z ciała — każdy dokładnie raz. W przypadku postulowanego ciała tabelki te miałyby zatem postać:

$+$	0	1	a	b	\cdot	0	1	a	b
0	0	1	a	b	0	0	0	0	0
1	1				1	0	1	a	b
a	a				a	0	a		
b	b				b	0	b		

Tabelkę mnożenia wypełnić można jeszcze dokładniej. Zauważmy bowiem, że wobec istnienia odwrotności dowolnego elementu niezerowego w ciele oraz łączności mnożenia, dla dowolnych $a \neq 1$ oraz $b \neq 0$ mamy $a \cdot b \neq b$. W przeciwnym bowiem razie mielibyśmy $a = abb^{-1} = bb^{-1} = 1$. A zatem $a \cdot b = 1$, czyli nie ma innej możliwości uzupełnienia trzeciego wiersza tabelki. W szczególności mamy:

$+$	0	1	a	b	\cdot	0	1	a	b
0	0	1	a	b	0	0	0	0	0
1	1				1	0	1	a	b
a	a				a	0	a	b	1
b	b				b	0	b	1	a

Teraz skorzystamy z rezultatów uzyskanych wyżej. Ciało czteroelementowe musi mieć charakterystykę równą 2. Stąd $1 + 1 = a + a = b + b = 0$. Mamy też $a + 1 \neq a$, czyli $a + 1 = b$. Analogicznie $b + 1 = a$, co pozwala do końca uzupełnić powyższe tabelki:

$+$	0	1	a	b	\cdot	0	1	a	b
0	0	1	a	b	0	0	0	0	0
1	1	0	b	a	1	0	1	a	b
a	a	b	0	1	a	0	a	b	1
b	b	a	1	0	b	0	b	1	a

Zauważmy, że sposób wypełniania powyższych tabelek był całkowicie jednoznaczny, przy założeniu symetrii różnych (od siebie i od 0, 1) elementów a, b rozważanego ciała. Nietrudno sprawdzić, że uzyskane tabelki działań zadają na zbiorze $\{0, 1, a, b\}$ strukturę ciała. W dodatku do wykładzie o przestrzeniach liniowych, przeprowadzimy ogólniejszą konstrukcję ciał p^k elementowych.

3.5 Przypomnienie. Relacje równoważności. Funkcje.

W tym nieco dłuższym podrozdziale, przypominamy podstawowe informacje związane z pojęciami relacji i funkcji, wybierając jedynie te, które są przydatne w kontekście definicji struktur algebraicznych występujących w niniejszym podręczniku.

Mając dwa elementy a, b zbioru X możemy utworzyć PARĘ UPORZĄDKOWANĄ (a, b) , w której wyróżniamy pierwszy element a — POPRZEDNIK PARY i drugi element b — NASTĘPNIK PARY, odróżniając tę parę od (b, a) , o ile tylko $a \neq b$. Pary uporządkowane (a, b) oraz (c, d) uznamy za równe wtedy i tylko wtedy, gdy $a = c$ oraz $b = d$.

Definicja 3.5.1: Iloczyn kartezjański, relacja dwuargumentowa

ILOCZYNEM KARTEZJAŃSKIM zbiorów X i Y nazywamy zbiór wszystkich par uporządkowanych (x, y) , gdzie $x \in X$ oraz $y \in Y$.

Przez RELACJĘ DWUARGUMENTOWĄ (BINARNĄ) R w zbiorze $X \times Y$ rozumiemy dowolny podzbiór $R \subseteq X \times Y$. Jeśli $X = Y$, mówimy że R jest relacją w zbiorze X .

Jeśli $(x, y) \in R$ jest elementem relacji na zbiorze $X \times Y$, to piszemy najczęściej xRy , czytając: element x jest w relacji R z elementem y .

Przykłady

- Dla dowolnego zbioru X relację R złożoną wyłącznie z par $(x, x) \in X \times X$ nazywamy *relacją identyczności* w zbiorze X .
- Dla dowolnego zbioru X w zbiorze $P(X)$ wprowadzić można *relację inkluzji* R złożoną wyłącznie z takich par podzbiorów $(A, B) \in P(X) \times P(X)$, że $A \subseteq B$.
- W zbiorze \mathbb{N} wprowadzić można relację podzielności R złożoną wyłącznie z takich par $(x, y) \in \mathbb{N} \times \mathbb{N}$, że liczba x jest dzielnikiem liczby y .
- W zbiorze \mathbb{R} wprowadzić można relację porządku R złożoną wyłącznie z takich par $(x, y) \in \mathbb{R} \times \mathbb{R}$, że $x \leq y$.

Definicja 3.5.2: Zwrotność, symetryczność i przechodniość relacji

Niech $R \subseteq X \times X$ będzie relacją dwuargumentową w zbiorze X . Relację R nazywamy

- (i) ZWROTNA, jeśli dla każdego $x \in X$ zachodzi: xRx .
- (ii) SYMETRYCZNA, jeśli dla każdych $x, y \in X$ zachodzi: $xRy \implies yRx$.
- (iii) PRZECHODNIA, jeśli dla dowolnych $x, y, z \in X$ zachodzi: xRy oraz $yRz \implies xRz$.

Przykłady

- Relacja identyczności na zbiorze X jest zwrotna, symetryczna i przechodnia.
- W zbiorze liczb całkowitych \mathbb{Z} określić można relację \equiv_n przystawania modulo n : liczby $x, y \in \mathbb{Z}$ spełniają $x \equiv_n y$ wtedy i tylko wtedy, gdy liczba $x - y$ jest podzielna przez n . Relacja \equiv_n jest zwrotna, symetryczna i przechodnia.
- Relacja przystawania pary trójkątów w zbiorze wszystkich trójkątów na płaszczyźnie jest zwrotna, symetryczna i przechodnia.
- Zdefiniowane wcześniej relacje: \leq w zbiorze \mathbb{R} , \subseteq w zbiorze $P(X)$ oraz relacja podzielności w zbiorze \mathbb{N} są zwrotne, przechodnie, ale nie są symetryczne. Relacje te są *antysymetryczne*, co oznacza dla relacji R na zbiorze X , że dla każdych $x, y \in X$ z warunków xRy oraz yRx wynika $x = y$.

Definicja 3.5.3: Relacja równoważności

Relację \equiv w zbiorze X , która jest jednocześnie zwrotna, symetryczna i przechodnia nazywamy RELACJĄ RÓWNOWAŻNOŚCI w zbiorze X .

Relacja równoważności jest swego rodzaju uogólnieniem relacji równości elementów zbioru — chodzić będzie o równość z jakiegoś precyzyjnie określonego punktu widzenia, czy pod jakimś wybranym względem. Relacje w pierwszych trzech przykładach wyżej są relacjami równoważności.

Wykażemy, że jeśli w zbiorze X jest relacja równoważności \equiv i dla elementów $x, y \in X$ wyróżnimy podzbiory $[x], [y]$ złożone z wszystkich elementów X , z którymi dany element jest on w relacji, to podzbiory te są równe lub rozłączne.

Definicja 3.5.4: Klasa abstrakcji (równoważności) elementu

Niech dana będzie relacja równoważności \equiv w zbiorze X . Dla dowolnego elementu $x \in X$ zbiór

$$\{y \in X \mid x \equiv y\}$$

nazywamy KLASĄ ABSTRAKCJI elementu x i oznaczamy przez $[x]$. Dowolny element należący do $[x]$ nazywamy REPREZENTANTEM klasy $[x]$.

Przykłady

- W zbiorze liczb całkowitych z relacją \equiv_2 klasą elementu 0 jest zbiór wszystkich liczb parzystych, a klasą elementu 1 jest zbiór wszystkich liczb nieparzystych.
- W zbiorze liczb rzeczywistych z relacją równoważności $x \equiv y$ wtedy i tylko wtedy, gdy $|x| = |y|$ klasą abstrakcji elementu 0 jest zbiór $\{0\}$, zaś klasą abstrakcji dowolnego elementu $x \neq 0$ jest podzbiór $\{-x, x\}$.
- W zbiorze liczb rzeczywistych z relacją równoważności $x \equiv y$ wtedy i tylko wtedy, gdy $x - y \in \mathbb{Z}$ klasą abstrakcji elementu 0 jest zbiór wszystkich liczb całkowitych, a klasą abstrakcji elementu $\sqrt{2}$ jest podzbiór

$$\{\sqrt{2} + c, \mid c \in \mathbb{Z}\}.$$

Twierdzenie 3.5.5: Zasada abstrakcji

Relacja równoważności \equiv w zbiorze X dzieli ten zbiór na rozłączne i niepuste klasy abstrakcji, to znaczy: każdy element należy do jednej klasy abstrakcji, a dwa elementy x, y zbioru X należą do tej samej klasy abstrakcji wtedy i tylko wtedy, gdy $x \equiv y$.

Dowód. Skoro relacja \equiv jest zwrotna, to mamy $x \in [x]$, dla każdego $x \in X$. Każdy element zbioru X należy więc do pewnej klasy abstrakcji relacji równoważności \equiv .

Załóżmy, że elementy $x, y \in X$ należą do klasy równoważności $[z]$ elementu $z \in X$. Z definicji klasy abstrakcji mamy więc $z \equiv x$ oraz $z \equiv y$. Z symetryczności relacji \equiv mamy natomiast $x \equiv z$, a stąd z przechodniości relacji \equiv wynika, że $x \equiv y$.

Z drugiej strony, jeśli $x \equiv y$, to $y \in [x]$. Oczywiście $x \in [x]$, więc $[x] \cap [y] \neq \emptyset$. Przypuśćmy, że istnieje element $z \in X$, że $z \in [x]$, ale $z \notin [y]$. Stąd $z \equiv x$. Oczywiście mamy też $x \equiv y$, więc z przechodniości relacji \equiv mamy $z \equiv y$, co przeczy założeniu, że $z \notin [y]$. Zatem $[x] \setminus [y] = \emptyset$. Analogicznie wykazujemy, że $[y] \setminus [x] = \emptyset$. Stąd $[x] = [y]$. \square

W ostatniej części przypomnijmy definicję funkcji, pozwalającą m.in. na wprowadzenie ogólnej definicji iloczynu kartezjańskiego oraz pojęcia działania n -argumentowego.

Definicja 3.5.6: Funkcja

Niech X i Y będą zbiorami. Relację dwuargumentową $f \subset X \times Y$ nazywamy FUNKCJĄ (lub PRZEKSZTAŁCENIEM, lub ODWZOROWANIEM), jeśli dla każdego $x \in X$ istnieje dokładnie jeden element $y \in Y$, taki że $x f y$.

Zbiór X nazywamy DZIEDZINĄ funkcji f , a jego elementy — ARGUMENTAMI. Dla każdego elementu $x \in X$ jednoznacznie wyznaczony element y , że $x f y$ oznaczamy przez $f(x)$ i nazywamy WARTOŚCIĄ funkcji f przy argumencie f . Zbiór Y nazywamy PRZECIWZDZIEDZINĄ funkcji f . Funkcję f o dziedzinie w zbiorze X i przeciwdziedzinie w zbiorze Y oznaczamy jako $f : X \rightarrow Y$.

Niech A będzie podzbiorem zbioru X . Zbiór

$$\{f(x) \mid x \in A\}$$

nazywamy OBRAZEM zbioru A przy funkcji f i oznaczamy przez $f(A)$.

Niech B będzie podzbiorem Y . Zbiór

$$\{x \in X \mid f(x) \in B\}$$

nazywamy PRZECIWOBRAZEM zbioru B przy funkcji f i oznaczamy przez $f^{-1}(B)$.

Przypomnijmy podstawowe definicje związane z pojęciem funkcji.

Definicja 3.5.7: Bijekcja, równoliczność, permutacja

Niech X, Y będą zbiorami. Mówimy, że funkcja $f : X \rightarrow Y$:

- (i) przekształca zbiór X na zbiór Y (jest *surjekcją*) wtedy i tylko wtedy, gdy zbiór $f(X)$ wartości funkcji f równy jest jej przeciwdziedzinie,
- (ii) jest *różnowartościowa* (jest *iniekcją*), jeśli dla dowolnych $x_1, x_2 \in X$ zachodzi $f(x_1) = f(x_2) \implies x_1 = x_2$,
- (iii) jest *bijekcją* (inaczej: *funkcją wzajemnie jednoznaczna*), jeśli jest różnowartościowa i przekształca X na Y .

O zbiorach X, Y , dla których istnieje bijekcja $f : X \rightarrow Y$ mówimy, że są RÓWNOLICZNE lub, że są TEJ SAMEJ MOCY.

Jeśli $X = \{1, 2, \dots, n\}$ jest zbiorem n -elementowym, to bijekcję $\sigma : X \rightarrow X$ nazywamy *permutacją*. Zbiór wszystkich permutacji zbioru X oznaczamy S_n .

Z pojęciem funkcji wzajemnie jednoznacznej związane jest w sposób naturalny pojęcie funkcji odwrotnej. Przypomnijmy podstawowe definicje i wyniki.

Definicja 3.5.8: Złożenie funkcji

Jeśli dane są dwie funkcje $f : X \rightarrow Y$ oraz $g : Y \rightarrow Z$ to funkcja $h : X \rightarrow Z$ zdefiniowana wzorem $h(x) = g(f(x))$ nazywana jest *złożeniem* (lub *superpozycją*) funkcji g z funkcją f . Oznaczamy ją symbolem $g \circ f$.

Dowolne funkcje $f, g : X \rightarrow X$ można złożyć, przy czym kolejność składania często ma znaczenie. Jeśli jednak dana jest funkcja $f : X \rightarrow X$, to oczywiście $(f \circ f) \circ f = f \circ (f \circ f)$, skąd wynika, że rozważać można n -tą iterację funkcji f będącą n -krotnym złożeniem funkcji f , czyli funkcję $f^n : X \rightarrow X$ określoną jako $f^n = \underbrace{f \circ f \circ \dots \circ f}_n$.

Definicja 3.5.9: Funkcja odwrotna

Funkcję $g : Y \rightarrow X$ nazywamy *odwrotną* do funkcji $f : X \rightarrow Y$ wtedy i tylko wtedy, gdy obie funkcje $g \circ f$ oraz $f \circ g$ są identycznościami (odpowiednio na zbiorach X, Y), to znaczy dla dowolnych $x \in X$ oraz $y \in Y$ mamy

$$g(f(x)) = x \quad \text{oraz} \quad f(g(y)) = y.$$

Jeśli oznaczymy funkcję identycznostwoową na zbiorze X przez id_X , to nietrudno pokazać, że dla funkcji $f : X \rightarrow Y$ warunek istnienia funkcji $g : Y \rightarrow X$, takiej że $g \circ f = \text{id}_X$ jest równoważny temu, że funkcja f jest różnowartościowa. Natomiast spełnianie przez g warunku $f \circ g = \text{id}_Y$ jest równoważne z tym, że funkcja f jest „na”.

Twierdzenie 3.5.10: O istnieniu funkcji odwrotnej

Funkcja $f : X \rightarrow Y$ ma funkcję odwrotną wtedy i tylko wtedy, gdy jest różnowartościowa i przekształca zbiór X na zbiór Y (to znaczy: jest injekcją i surjekcją).

Dowód. Twierdzenie jest sformułowane w postaci równoważności. Założymy najpierw, że funkcja f posiada funkcję odwrotną $g : Y \rightarrow X$. Wykażemy, że f jest bijekcją.

Niech x_1, x_2 będą elementami dziedziny X . Jeśli $f(x_1) = f(x_2)$, to

$$x_1 = g(f(x_1)) = g(f(x_2)) = x_2.$$

Funkcja f jest więc różnowartościowa. Wobec istnienia funkcji odwrotnej do f mamy $y = f(g(y))$, dla każdego $y \in Y$. Stąd każdy element y zbioru Y jest wartością funkcji f . A zatem f jest na, a w konsekwencji — f jest bijekcją.

Założymy teraz, że różnowartościowa funkcja $f : X \rightarrow Y$ przeprowadza zbiór X na zbiór Y . Określamy funkcję $g : Y \rightarrow X$ warunkiem:

$$g(y) = x \iff y = f(x).$$

Ponieważ dla każdego elementu $y \in Y$ istnieje dokładnie jeden element $x \in X$ spełniający powyższy warunek (definicja funkcji), więc wzór $g(y) = x$ rzeczywiście określa funkcję. Z tego określenia wynika także, że $f(g(y)) = f(x) = y$ oraz $g(f(x)) = g(y) = x$, co oznacza, że g jest funkcją odwrotną do f . \square

Wniosek 3.5.11

Jeśli funkcja f ma funkcję odwrotną, to tylko jedną.

Jako ćwiczenie pozostawiamy dwie inne nietrudne własności operacji składania funkcji, które będą szczególnie przydatne w dalszych rozdziałach.

Uwaga 3.5.12

Dla dowolnych funkcji $f : X \rightarrow Y$, $g : Y \rightarrow Z$ oraz $h : Z \rightarrow T$ mamy $h \circ (g \circ f) = (h \circ g) \circ f$. Dla dowolnych funkcji odwracalnych $f : X \rightarrow Y$ oraz $g : Y \rightarrow Z$ złożenie $g \circ f$ jest funkcją odwracalną, oraz $(g \circ f)^{-1} = f^{-1} \circ g^{-1}$.

Na koniec przywołamy ogólną definicję iloczynu kartezjańskiego.

Definicja 3.5.13: Iloczyn kartezjański rodzin podzbiorów

Niech $\{X_t\}_{t \in T}$ będzie daną rodziną indeksowaną podzbiorów pewnego zbioru X . Przez ILOCZYN KARTEZJAŃSKI zbiorów rodzin $\{X_t\}_{t \in T}$ rozumiemy zbiór wszystkich funkcji $f : T \rightarrow \bigcup_{t \in T} A_t$, spełniających warunek $f(t) \in A_t$, dla każdego $t \in T$.

3.6 Dodatek. Ciało ułamków

Pojęcie relacji równoważności wprowadzone wyżej jest podstawowym narzędziem wykorzystywanym niemal w każdym dziale matematyki. W tym dodatku przyjrzymy się pewnej konstrukcji algebraicznej, pozwalającej na uzyskiwanie ważnych przykładów ciał (i nie tylko, ale o tym nie powiemy). W tym celu potrzebne będzie najpierw wprowadzenie pojęcia dziedziny całkowitości — struktury ogólniejszej od ciała, obejmującej między innymi liczby całkowite, czy wielomiany (o których mowa będzie w rozdziale 5).

Definicja 3.6.1: Pierścień przemienny z 1

Pierścieniem przemiennym z 1 nazywamy piątkę $(R, \boxplus, \boxtimes, 0, 1)$, gdzie R jest zbiorem przynajmniej dwuelementowym z wyróżnionymi elementami $0 \neq 1$, zwanymi ZEREM i JEDYNKĄ, zaś \boxplus, \boxtimes są dwuargumentowymi działaniami zwany dodawaniem i mnożeniem, spełniającymi aksjomaty:

- | | | |
|----------------------------------------------------------------------------|---------------------------------------|-------------------------------------------|
| 1) $(a \boxplus b) \boxplus c = a \boxplus (b \boxplus c)$ | $\forall_{a,b,c \in K}$ | łączność dodawania |
| 2) $a \boxplus b = b \boxplus a$ | $\forall_{a,b \in K}$ | przemienność dodawania |
| 3) $a \boxplus 0 = a = 0 \boxplus a$ | $\forall_{a \in K}$ | własność elementu 0 |
| 4) $a \boxplus b = 0 = b \boxplus a$ | $\forall_{a \in K} \exists_{b \in K}$ | element przeciwny |
| 5) $(a \boxtimes b) \boxtimes c = a \boxtimes (b \boxtimes c)$ | $\forall_{a,b,c \in K}$ | łączność mnożenia |
| 6) $a \boxtimes b = b \boxtimes a$ | $\forall_{a,b \in K}$ | przemienność mnożenia |
| 7) $a \boxtimes 1 = 1 \boxtimes a = a$ | $\forall_{a \in K}$ | własność elementu 1 |
| 8) $a \boxtimes (b \boxplus c) = (a \boxtimes b) \boxplus (a \boxtimes c)$ | $\forall_{a,b,c \in K}$ | rozdzielność \boxtimes wzgl. \boxplus |

Widzimy zatem, że w definicji pierścienia przemiennego z 1 wymagamy wszystkich aksjomatów, które spełnia ciało, poza aksjomatem mówiącym, że każdy niezerowy element ma element odwrotny.

Przykłady pierścieni przemiennych z jedynką, które nie są ciałami

- zbiór liczb całkowitych,
- pierścień \mathbb{Z}_n reszt z dzielenia liczb całkowitych przez liczbę złożoną n z działaniami dodawania i mnożenia modulo n (gdy n jest liczbą pierwszą, pierścień ten jest ciałem).
- podzbiór zbioru liczb rzeczywistych złożony z liczb postaci $a + b\sqrt{2}$, gdzie $a, b \in \mathbb{Z}$,
- pierścień funkcji $f : \mathbb{R} \rightarrow \mathbb{R}$ z operacjami dodawania i składania.

Jedną z kluczowych własności wyprowadzonych na wykładzie było stwierdzenie, że jeśli w ciele K iloczyn dwóch elementów jest zerowy, to jeden z czynników jest zerem. Oczywiście własności tej nie musi mieć każdy pierścień przemienny z 1. W pierścieniu \mathbb{Z}_4 mamy $2 \cdot 2 = 0$.

Definicja 3.6.2: Dziedzina całkowitości

Pierścień przemienny z jedynką R nazywamy DZIEDZINĄ CAŁKOWITOŚCI, jeśli dla dowolnych elementów $a, b \in R$ warunek $ab = 0$ implikuje $a = 0$ lub $b = 0$.

Ponownie stwierdzamy, że każde ciało jest dziedziną całkowitości, ale nie na odwrót! Zauważymy, że skończony pierścień \mathbb{Z}_n nie jest ciałem, jeśli n nie jest liczbą pierwszą. Zachodzi ogólniejszy wniosek.

Uwaga 3.6.3

Skończona dziedzina całkowitości R jest ciałem.

Dowód. Niech $R^* = \{r_1, \dots, r_n\}$ będzie zbiorem niezerowych elementów R . Wykażemy, że każdy element D^* ma element odwrotny. Dla każdego $r \in R^*$ rozważmy zbiór elementów

$$\{rr_1, \dots, rr_n\}.$$

Twierdzimy, że powyższy zbiór równy jest R^* . Istotnie, wszystkie elementy tego zbioru są niezerowe, gdyż R jest dziedziną całkowitości, a gdyby dla pewnych i, j zachodziła równość $rr_i = rr_j$, to zgodnie z prawej rozdzielności mamy

$$r(r_i - r_j) = 0.$$

Skoro $r \neq 0$, to ponownie wnioskujemy $r_i = r_j$. Zatem w zbiorze wyżej znajduje się jedynka pierścienia R . \square

W jaki sposób powyższe rozważania o strukturach, które nie są ciałami mają nam pomóc poznać nowe przykłady ciał? Oto tytułowa konstrukcja, która to umożliwia.

Definicja 3.6.4: Ułamki dziedziny całkowitości

Niech R będzie dziedziną całkowitości i niech $R^* = R \setminus \{0\}$. Na zbiorze par uporządkowanych $R \times R^*$ wprowadzamy relację równoważności \sim określając, że dla dowolnych $r, s \in R$ oraz $r', s' \in R^*$ mamy

$$(r, r') \sim (s, s') \iff rs' = sr'.$$

Klasy abstrakcji tej relacji nazywamy **UŁAMKAMI** dziedziny R . Klasę $[(s, s')]$ oznaczamy przez $\frac{s}{s'}$. Zbiór wszystkich ułamków R określamy przez $Q(R)$.

Sprawdzenie, że zdefiniowana wyżej relacja jest relacją równoważności jest natychmiastowe. Oczywiście mamy $(r, r') \sim (r, r')$, podobnie warunek $(r, r') \sim (s, s')$ równoważny jest warunkowi $(s, s') \sim (r, r')$. Relacja \sim jest więc zwrotna i symetryczna. Aby wykazać, że jest to też relacja przechodnia, założymy, że $(r, r') \sim (s, s')$ oraz $(s, s') \sim (t, t')$. Mamy stąd

$$rs' = sr' \quad \text{oraz} \quad st' = ts'.$$

Mnożąc obustronnie pierwszą równość przez t' , dostajemy

$$rs't' = sr't'$$

Prawą stronę możemy zapisać, korzystając ponownie z założeń i przemienności w R :

$$sr't' = st'r' = ts'r' = r'ts'.$$

Lewą stronę możemy natomiast zapisać w postaci

$$rs't' = rt's'$$

W rezultacie

$$rt's' = r'ts' \iff (rt' - r't)s' = 0.$$

Stąd $rt' = r't$, czyli $(r, r') \sim (t, t')$. Relacja \sim jest więc przechodnia.

Oczywiście wzorcowa dla powyższych rozważań jest konstrukcja ciała liczb wymiernych jako ułamków pierścienia liczb całkowitych. Ułamki traktujemy jako identyczne, mówiąc nieprezencyjnie, jeśli każdy z nich można rozszerzyć w ten sposób, aby otrzymać iloraz identycznych liczb. Na przykład, pary liczb całkowitych $(1, 2), (2, 4)$ reprezentują ten sam ułamek, co oznaczamy po prostu

$$\frac{1}{2} = \frac{2}{4}.$$

Na zbiorze $Q(R)$ wprowadzić można działania dodawania i mnożenia, podobnie jak dla liczb wymiernych.

$$\frac{x}{y} + \frac{p}{q} = \frac{xq + py}{yq} \quad \text{oraz} \quad \frac{x}{y} \cdot \frac{p}{q} = \frac{xp}{yq}.$$

Definicje powyższe wymagają uzasadnienia poprawności. Należały w tym celu uzasadnić, czego tu nie zrobimy, że wyniki tych działań są niezależne od wyboru reprezentantów ułamków. Innymi słowy, jeśli

$$(r, r') \sim (s, s') \quad \text{oraz} \quad (t, t') \sim (u, u'), \quad \text{to} \quad (rs' + r's, r's') \sim (tu' + t'u, t'u'), \quad (rs, r's') \sim (tu, t'u').$$

Twierdzenie 3.6.5

Dla każdej dziedziny całkowitości R zbiór $Q(R)$ wraz z elementami $\frac{0}{1}, \frac{1}{1}$ oraz działaniami dodawania i mnożenia ułamków jest ciałem.

Nie będziemy dokonywać żmudnego sprawdzenia prawdziwości powyższego twierdzenia. Czytelnik zechce sprawdzić, że argumentacja przypomina uzasadnienie tego, że relacja \sim jest relacją równoważności.

Klasycznym przykładem jest, jak wspomnieliśmy, konstrukcja liczb wymiernych jako ułamków pierścienia liczb całkowitych. Innym ważnym przykładem jest ciało funkcji wymiernych, stanowiące ciało ułamków pierścienia wielomianów o współczynnikach w ciele K . Dla $K = \mathbb{Z}_2$ uzyskujemy w ten sposób przykład ciała $Q(\mathbb{Z}_2[x])$, które jest nieskończone i charakterystyki 2.

Konstrukcję powyższą zilustrujemy bardziej szczegółowo — ciałem liczb p -adycznych jest ciało ułamków dziedziny całkowitości, którą definiujemy na zbiorze tzw. koherentnych ciągów o wyrazach całkowitych. Więcej miejsca liczbom p -adycznym poświęcimy w osobnym dodatku, przedstawując ich alternatywną (topologiczną) konstrukcję jako analog konstrukcji Cantora liczb rzeczywistych (powstających jako tzw. uzupełnienie zbioru liczb wymiernych). W tym drugim ujęciu język podzielności przez odpowiednią potęgę liczby pierwszej przełożymy na język odległości, za pomocą tzw. normy p -adycznej.

Definicja 3.6.6: Ciąg koherentny

Niech p będzie liczbą pierwszą. Ciąg (x_n) o wyrazach całkowitych nazwiemy KOHERENTNYM, jeśli dla każdego $n \geq 1$ spełniony jest warunek $x_n \equiv x_{n-1} \pmod{p^n}$.

Na zbiorze ciągów koherentnych wprowadzamy relację równoważności \equiv postaci

$$(x_n) \equiv (y_n) \iff \forall_{n \geq 0} x_n \equiv y_n \pmod{p^{n+1}}.$$

Przykłady:

- Każdy ciąg stały (x_n) o wyrazach równych x spełnia pierwszy z warunków, ponieważ dla każdego n mamy $x \equiv x \pmod{p^n}$. Dwa ciągi stałe (x_n) o wszystkich wyrazach równych x i (y_n) o wszystkich wyrazach równych y są równoważne w powyższej relacji, wtedy i tylko wtedy, gdy $x \equiv y \pmod{p}$.
- Ciąg stany złożony z liczb -1 jest równoważny z ciągiem (x_n) o wyrazach $p^{n+1} - 1$.

Definicja 3.6.7: Całkowite liczby p -adyczne

Zbiór klas abstrakcji relacji \equiv oznaczamy nazywamy zbiorem CAŁKOWITYCH LICZB p -ADYCZNYCH.

Zauważmy, że biorąc ciąg koherentny (x_n) oraz liczby A_i będące resztami z dzielenia x_n przez p^{n+1} , dostajemy $[(x_n)] = [(A_n)]$. Ciąg (A_n) nazywamy często ZREDUKOWANĄ POSTACIĄ liczby p -adycznej $[(x_n)]$. Zauważmy, że dla każdego ciągu reszt z dzielenia przez p^{n+1} mamy reprezentanta w postaci całkowitej liczby p -adycznej. Liczby A_i związane są z tzw. *rozwinięciem p -adycznym* liczby $[(x_n)]$. Dla całkowitych liczb p -adycznych $[(x_n)]$ jak wyżej, jest to przedstawienie postaci

$$[(x_n)] = d_0 + d_1 p + d_2 p^2 + \dots,$$

gdzie $0 \leq d_i \leq p$, oraz gdzie sumy częściowe tego szeregu równe są A_i . Ogólną postać tego szeregu poznamy w kolejnym rozdziale. Zauważmy, że jeśli A jest liczbą całkowitą, to ciąg reszt z dzielenia A przez p^{n+1} jest skończony i liczbę A przypisać można jej rozwinięcie p -adyczne, będące po prostu przedstawieniem A w systemie pozycyjnym o podstawie p . Stąd liczby całkowite traktować można jako p -adyczne.

Twierdzenie 3.6.8: Dodawanie i mnożenie całkowitych liczb p -adycznych

Na zbiorze całkowitych liczb p -adycznych wprowadzamy działania:

$$[(x_n)] + [(y_n)] = [(x_n + y_n)], \quad [(x_n)] \cdot [(y_n)] = [(x_n y_n)].$$

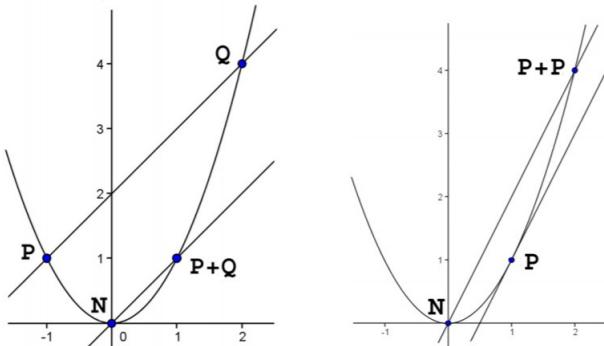
Działania te są dobrze określone i wprowadzają na tym zbiorze strukturę dziedziny całkowitości.

Zachęcamy Czytelnika do uzasadnienia tego twierdzenia oraz do wywnioskowania, że możliwa jest konstrukcja ciała ułamków całkowitych liczb p -adycznych — zwanych liczbami p -adycznymi. W kolejnym rozdziale przedstawimy inną konstrukcję oraz szkic uzasadnienia, że prowadzi ono do identycznych (izomorficznych) obiektów. Uzasadnienie oparte jest o dowód tego, że rozwinięcie p -adyczne jest jednoznaczne (a można je określić przy obydwu konstrukcjach).

3.7 Trivia. Ciało na paraboli

Dziwna może się wydawać definicja ciała liczb zespolonych, gdy patrzymy na nią z punktu widzenia działania na parach liczb. Przypomnij jednak uwagę z wykładu: zbiór \mathbb{R}^2 z działaniami dodawania i mnożenia „po współrzędnych” (tzn. z mnożeniem $(a, b) \otimes (c, d) = (ac, bd)$) nie jest ciałem! Ale nie o tym jest ten dodatek. Chciałbym w nim wspomnieć o mało znanym przykładzie ciała, w którym działania dodawania i mnożenia wyglądają bardziej intuicyjnie niż w \mathbb{C} , ale zdefiniowane są na dość nietypowym obiekcie.

Rozważamy mianowicie parabolę \mathcal{P} o równaniu $y = x^2$. Pokażemy najpierw, że można wprowadzić na niej strukturę przemienneego dodawania. Niech $N = (0, 0)$. Określamy sumę $P \oplus Q$ dwóch punktów paraboli \mathcal{P} jako drugi punkt przecięcia paraboli oraz prostej równoległej do prostej PQ przechodzącej przez punkt N . Jeśli $P = Q$, to zastępujemy prostą PQ przez prostą styczną do paraboli w punkcie P . Poniższe (zapożyczone) rysunki pokazują odpowiednio działania postaci $(-1, 1) \oplus (2, 4) = (1, 1)$ oraz $(1, 1) \oplus (1, 1) = (2, 4)$.



Źródło: Franz Lemmermeyer. *Pell Conics. An Alternative Approach to Elementary Number Theory*. <http://www.rzuser.uni-heidelberg.de/~hb3/pell.html>

Weryfikacja przypuszczenia, że wprowadzone działanie dwuargumentowe \oplus zadaje na \mathcal{P} działanie łączne z elementem neutralnym N wymaga odrobiny wiedzy szkolnej i wytrwałości. Wyznaczmy wzory na dodawanie dwóch punktów $P_1 = (x_1, x_1^2)$ oraz $P_2 = (x_2, x_2^2)$. Jeśli $P_1 \neq P_2$, to prosta przechodząca przez P_1 oraz P_2 ma współczynnik kierunkowy postaci: $m = (x_2^2 - x_1^2)/(x_2 - x_1) = x_1 + x_2$. Prosta równoległa do prostej P_1P_2 przechodząca przez punkt $N = (0, 0)$ ma równanie postaci $y = mx$. Aby znaleźć jej drugi punkt przecięcia z parabolą potrzebujemy rozwiązać równanie $mx = x^2$. To zaś daje nam dwa punkty przecięcia: $N = (0, 0)$ oraz $R = (m, m^2)$. A zatem na mocy naszej definicji działania \oplus mamy:

$$(x_1, x_1^2) \oplus (x_2, x_2^2) = (x_1 + x_2, (x_1 + x_2)^2).$$

Powyższa formuła pozostaje prawdziwa także gdy $P_1 = P_2$.

Ktoś powie: *to w zasadzie nic ciekawego*. Matematyk od razu widzi, że zdefiniowane działanie to „właściwie” (nie znamy pojęcia izomorfizmu) dodawanie liczb rzeczywistych. W podobny sposób na hiperboli $xy = 1$ wprowadzić można działanie mnożenia, które jest „izomorficzną kopią” mnożenia liczb rzeczywistych. Czy umelibyście Państwo zaproponować geometryczną konstrukcję mnożenia punktów na hiperboli? Aby dowiedzieć się więcej o podobnych konstrukcjach na stożkowych (i nie tylko) zachęcam do lektury książki, do której odsyłam pod powyższym obrazkiem. Znajdzicie tam Państwo łagodny wstęp do teorii punktów wymiernych na krzywych i ich zastosowań. W latach 90’ zaawansowane metody teorii krzywych eliptycznych zaowocowały dowodem Wielkiego Twierdzenia Fermata przez Andrew Wilesa. Polecam poglądowy wykład Wilesa i następujący po nim inspirujący wywiad mówiący o życiu zawodowym matematyka. Adres: <https://www.youtube.com/watch?v=uQgcpzKA5jk>.

To jednak nie koniec opowieści. Parabola ma tę szczególną cechę, że można na niej wprowadzić nie tylko strukturę tzw. grupy algebraicznej, ale i strukturę ciała. Spróbujmy więc określić działanie mnożenia. Niech $I = (1, 1)$. Dla punktów $P, Q \in \mathcal{P}$ definiujemy mnożenie w następujący sposób: prosta PQ przecina oś OY w punkcie A oraz prosta IA przecina parabolę \mathcal{P} w punkcie $B := P \star Q$ (proszę to dopracować).

Czy potraficie Państwo napisać algebraiczną formułę opisującą działanie \star ? Czy potraficie Państwo pokazać, za pomocą tych wzorów, że $(\mathcal{P}, \oplus, \star, N, I)$ jest ciałem? Okazuje się, że wcale nie potrzeba tu algebry. Zarówno łączność mnożenia, jak i rozdzielność mnożenia względem dodawania można udowodnić geometrycznie. Czy ktoś z Państwa potrafiłby to zrobić? Fakt ten, co zaskakujące, pochodzi z 2003 roku.

3.8 Trivia. Równania językowe

Wspomnieliśmy już o tym, że równania liniowe, których współczynniki nie są w ciele, ale np. w zbiorze liczb całkowitych, mogą nie mieć rozwiązania. Warto pociągnąć ten wątek, przyglądając się jak dalece ogólna może być definicja działania i równania liniowego oraz jak mało oczywistych dla nas własności może ona zachowywać. Rozważmy następujący, egzotyczny przykład struktury algebraicznej z działaniami dodawania i mnożenia. Będą to... zbiory słów, czyli w skrócie: słowniki.

Aby mieć słownik, trzeba mieć najpierw słowa. Niech $\Sigma_{a,b}$ będzie zbiorem złożonym ze wszystkich słów złożonych z liter a, b , np. $a, aaa, abaa, bbba$ oraz słowa pustego ϵ . W $\Sigma_{a,b}$ wprowadzamy działanie dwuargumentowe, tzw. **konkatenację**, · postaci $w_1 \cdot w_2 = w_1 w_2$, np.

$$\textcolor{red}{aba} \cdot \textcolor{blue}{bb} = \textcolor{red}{ababb}, \quad \epsilon \cdot abb = abb.$$

Rozważamy zbiór $X = P(\Sigma_{a,b})$ złożony z podzbiorów (także nieskończonych!) zbioru $\Sigma_{a,b}$, np.

$$\{\epsilon, a, ab\}, \quad \{a, aa, aaa, aaaa, \dots\}, \quad \{ab, abab, ababab, \dots\}.$$

Elementy zbioru X nazywać będziemy słownikami. W $P(\Sigma_{a,b})$ wprowadzamy działania dwuargumentowe:

- + oznaczające sumę mnogościową zbiorów, np. $\{aba, bb, ab\} + \{a, aa, bb\} = \{aba, bb, ab, a, aa\}$,
- · oznaczające zbiór powstający przez konkatenację wszystkich wyrazów z pierwszego zbioru ze wszystkimi elementami z drugiego. Innymi słowy, dla dowolnych słowników A, B , zbiór $A \cdot B$ złożony jest ze słów postaci $a \cdot b$, gdzie $a \in A, b \in B$. Np.

$$\{\textcolor{red}{aba}, \textcolor{blue}{bb}, \textcolor{red}{ab}\} \cdot \{\textcolor{blue}{a}, \textcolor{red}{aa}, \textcolor{blue}{bb}\} = \{\textcolor{red}{abaa}, \textcolor{blue}{abaaa}, \textcolor{red}{ababb}, \textcolor{blue}{bba}, \textcolor{red}{bbaa}, \textcolor{blue}{bbb}, \textcolor{red}{aba}, \textcolor{blue}{abb}\}$$

Rozważamy równania liniowe o współczynnikach w $P(\Sigma_{a,b})$, np. równanie o zmiennych x_1, x_2 :

$$\{a, aa\} \cdot x_1 + \{bb\} \cdot x_2 = \{ab, aab, bbb, aaab\},$$

którego **rozwiązaniem są pary elementów $P(\Sigma_{a,b})$** . W tym przypadku: $x_1 = \{b, ab\}$, $x_2 = \{b\}$.

Widzimy, że od strony formalnej cała opisana konstrukcja mieści się w definicji równania liniowego i jego rozwiązania. Równania liniowe, między innymi takie, jak wyżej, nazywa się **równaniami językowymi**. Rozwiązywanie tych równań w niczym nie przypomina znanych nam metod. Dlaczego?

- Pierwszy problem to konieczność określenia strony, z której piszemy współczynniki. Równania:

$$\{a\} \cdot x_1 = \{abaa\}, \quad x_1 \cdot \{a\} = \{abaa\}$$

mają różne rozwiązania! Przyczyna – nieprzemienność działania ·.

- Zauważmy, że w zbiorze słowników $P(\Sigma_{a,b})$ nie ma elementów *przeciwnych* i *odwrotnych*. Mając układ:

$$\begin{cases} \{a\} \cdot x_1 = \{abaa\} \\ \{a\} \cdot x_1 = \{aba\}. \end{cases}$$

nie sprowadzimy jego *macierzy* do postaci schodkowej lub zredukowanej, co utrudnia sprawdzanie kiedy jest on sprzeczny!

- Układy jednorodne nie mają sensu, bo $\{\epsilon\} \cup \{w\} = \{w\}$, ale $\{\epsilon\} \cdot \{w\} \neq \{\epsilon\}$.
- Trudno kontrolować zbiory rozwiązań. Rozwiązywanie równania wyżej było proste, bo współczynnikami były skończone słowniki. Proszę jednak pomyśleć na przykład o równaniu postaci:

$$X = \{a^n b^n \mid n \geq 1\} \cdot X \cdot \{b^n a^n \mid n \geq 1\} \cup \{\epsilon\},$$

gdzie ϵ jest słowem pustym.

Na potrzeby ilustracji problemów z nieprzemiennością i brakiem odwrotności operacji mnożenia (a to nie jedyne problemy, jak widać wyżej) dokonuję tu i tak dużego uproszczenia tego tematu. Czytelnik zainteresowany szczegółami może zapoznać się z prezentacją Michala Kunca pt. Language Equations (dostępna online) lub z monografią *Language Equations* autorstwa Ernsta Leissa (biblioteka IMPAN). Wcześniej jednak warto przejść/przeczytać wykład kursowy z Języków, automatów i obliczeń.

Rozdział 4

Ciało liczb zespolonych

4.1 Wykład 4

Przejdziemy teraz do drugiego fundamentalnego przykładu ciała, liczb zespolonych.

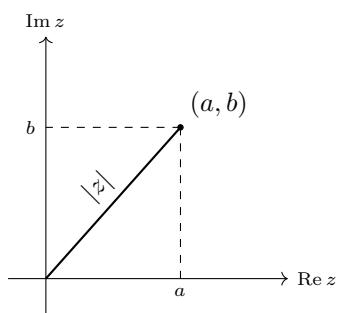
Definicja 4.1.1: Ciało liczb zespolonych

CIAŁO LICZB ZESPOLONYCH to piątka $(\mathbb{R}^2, \oplus, \otimes, (0, 0), (1, 0))$, oznaczana przez \mathbb{C} , którego elementami są wszystkie uporządkowane pary liczb rzeczywistych, i w którym działania \oplus, \otimes określone są za pomocą działań $+$ oraz \cdot w \mathbb{R} wzorami:

$$(a, b) \oplus (c, d) = (a + c, b + d), \quad (a, b) \otimes (c, d) = (ac - bd, ad + bc).$$

Dla dowolnej liczby zespolonej $z = (a, b)$ wprowadzamy oznaczenia:

- a nazywamy CZEŚCIĄ RZECZYWISTĄ liczby z i oznaczamy ją przez $\operatorname{Re} z$,
- b nazywamy CZEŚCIĄ UROJONĄ liczby z i oznaczamy $\operatorname{Im} z$,
- liczbę $\sqrt{a^2 + b^2}$ nazywamy MODULEM LICZBY z i oznaczamy jako $|z|$.



Rys. 1. Płaszczyzna zespolona. Część rzeczywista, urojona oraz moduł liczby zespolonej.

Odwołamy się teraz do niezwykle ważnej, geometrycznej interpretacji. Liczby zespolone to pary punktów i możliwe jest reprezentowanie liczb zespolonych na tzw. płaszczyźnie zespolonej. Oś odpowiadającą części rzeczywistej liczby zespolonej oznaczamy $\operatorname{Re} z$, a oś odpowiadającą części urojonej oznaczamy jako $\operatorname{Im} z$. Moduł $|z|$ liczby zespolonej z interpretujemy jako odległość (euklidesową) punktu z od punktu $(0, 0)$.

Zobaczmy kilka przykładowych działań w ciele \mathbb{C} .

- $(0, 1) \oplus (1, 0) = (1, 1)$,
- $(2, 1) \otimes (2, -1) = (2 \cdot 2 - 1 \cdot (-1), 2 \cdot (-1) + 1 \cdot 2) = (5, 0)$.

Przyporządkowanie $a \mapsto (a, 0)$ zadaje utożsamienie zbioru liczb rzeczywistych z podzbiorem zbioru liczb zespolonych złożonym ze wszystkich liczb postaci $(r, 0)$. Przy tym utożsamieniu działania na liczbach rzeczywistych odpowiadają działaniom na ich odpowiednikach w zbiorze liczb zespolonych. Mamy bowiem:

$$(a, 0) \oplus (a', 0) = (a + a', 0), \quad (a, 0) \otimes (a', 0) = (aa' - 0, 0 + 0) = (aa', 0).$$

W tym przyporządkowaniu:

- liczbę postaci $(a, 0)$ będziemy zapisywać jako a , dla każdego $a \in \mathbb{R}$,
- liczbę $(0, 1)$ oznaczać będziemy jako i .

Używając tych oznaczeń mamy na przykład:

$$(a, b) = (a, 0) \oplus (0, b) = (a, 0) \oplus (b, 0) \otimes (0, 1) = a + bi, \quad i^2 = (0, 1) \otimes (0, 1) = (-1, 0) = -1.$$

Widzimy więc, że liczbę $z = (a, b)$ zapisywać możemy w POSTACI OGÓLNEJ (algebraicznej)

$$z = a + bi,$$

przyjmując umowę, że jeśli $a + bi = c + di$, to $a = c$ oraz $b = d$. W przyjętej konwencji dodawanie i mnożenie liczb zespolonych staje się bardziej zrozumiałe i pozwala na opuszczenie oznaczeń \oplus, \otimes . Dodawanie i mnożenie liczb zespolonych w postaci ogólnej sprowadzają się do wykonywania operacji algebraicznych, uwzględniających zasadę: $i^2 = -1$, czyli np.

$$(a + bi)(c + di) = ac + adi + bci + bdi^2 = (ac - bd) + (ad + bc)i.$$

Przykład. Wyznaczmy wszystkie liczby $z = a + bi$ spełniające równanie:

$$(2 - i)z = 10 - z.$$

- krok pierwszy – podstawiamy $z = a + bi$:

$$(2 - i)(a + bi) = 10 - (a + bi)$$

- krok drugi – porządkujemy:

$$\begin{aligned} (2 - i)(a + bi) &= 10 - a - bi \iff \\ 2a + 2bi - ai - bi^2 &= 10 - a - bi \iff \\ (3a + b - 10) + (3b - a)i &= 0. \end{aligned}$$

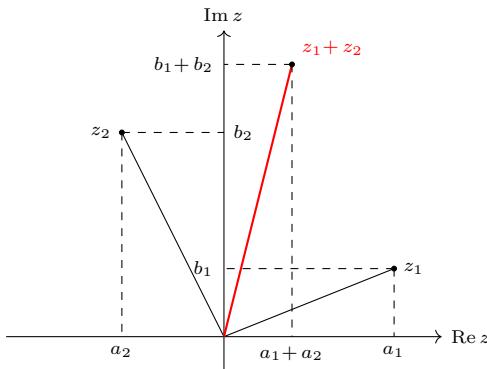
- krok trzeci – rozwiązujeśmy układ równań

$$\begin{cases} 3a + b - 10 = 0 \\ 3b - a = 0. \end{cases}$$

Rozwiązanie to: $a = 3, b = 1$, a więc jedynie rozwiązanie wyjściowego równania to $z = 3 + i$.

Zauważmy, że dodawanie liczb zespolonych $z_1 = a_1 + b_1i$ oraz $z_2 = a_2 + b_2i$ przypomina dodawanie wektorów. Mamy

$$(a_1 + b_1i) + (a_2 + b_2i) = (a_1 + a_2) + (b_1 + b_2)i.$$



Rys. 2. Interpretacja geometryczna dodawania liczb zespolonych.

Definicja 4.1.2: Postać trygonometryczna liczby zespolonej

Niech $z = a + bi \neq 0$ będzie liczbą zespoloną, zaś $\theta \in \mathbb{R}$ — miarą łukową kąta między półprostą o początku $(0, 0)$ przechodzącą przez $(1, 0)$, a półprostą o początku $(0, 0)$, przechodzącą przez z .

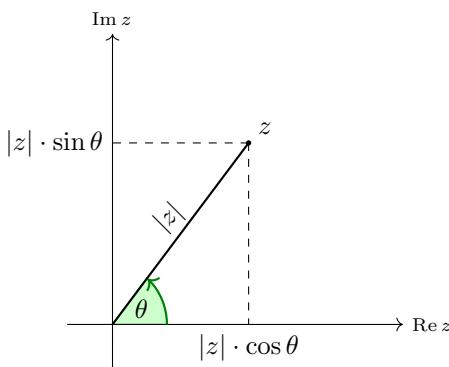
- Liczbę θ nazywamy ARGUMENTEM liczby z i oznaczamy przez $\arg z$. Argument liczby zespolonej $z \neq 0$ jest więc wyznaczony z dokładnością do całkowitej wielokrotności 2π . Liczbie 0 przypisujemy moduł 0 i dowolny argument $\theta \in \mathbb{R}$.
- Korzystając ze szkolnej definicji funkcji trygonometrycznych i z twierdzenia Pitagorasa mamy:

$$\cos \theta = \frac{a}{\sqrt{a^2 + b^2}} = \frac{a}{|z|}, \quad \sin \theta = \frac{b}{\sqrt{a^2 + b^2}} = \frac{b}{|z|}.$$

Stąd $a = |z| \cos \theta$ oraz $b = |z| \sin \theta$, a więc

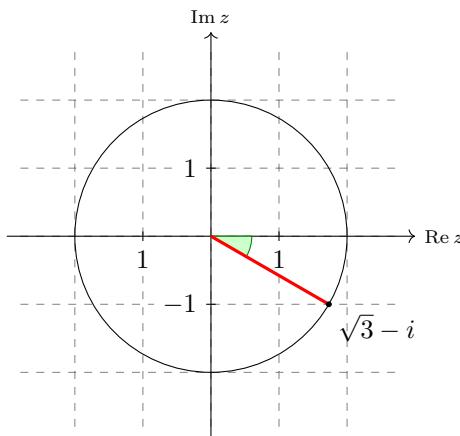
$$z = |z|(\cos \theta + i \cdot \sin \theta).$$

Jest to POSTAĆ TRYGONOMETRYCZNA LICZBY ZESPOLONEJ $z \neq 0$.



Rys. 3. Argument liczby zespolonej.

Przykład. Znajdziemy postać trygonometryczną liczby $z = \sqrt{3} - i$.



Rys. 4. Moduł i argument liczby $\sqrt{3} - i$.

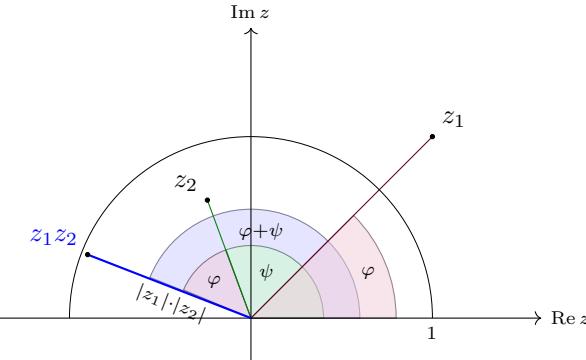
Mamy $|z| = \sqrt{\sqrt{3}^2 + (-1)^2} = 2$ oraz:

$$z = 2 \left(\frac{\sqrt{3}}{2} - \frac{1}{2} \cdot i \right) = 2 \left(\cos \frac{11\pi}{6} + i \cdot \sin \frac{11\pi}{6} \right) = 2 \left(\cos \frac{-\pi}{6} + i \cdot \sin \frac{-\pi}{6} \right).$$

Wniosek 4.1.3

Dla niezerowych liczb zespolonych w, z zachodzą równości

$$\arg(z \cdot w) = \arg(z) + \arg(w), \quad |zw| = |z| \cdot |w|.$$



Rys. 5. Mnożenie liczb zespolonych z_1 i z_2 w interpretacji geometrycznej.

Dowód. Następujące obliczenie pokazuje jak zachowuje się argument przy mnożeniu liczb zespolonych z, w danych w postaciach trygonometrycznych:

$$\begin{aligned} z \cdot w &= |z|(\cos \varphi + i \sin \varphi) \cdot |w|(\cos \psi + i \sin \psi) = \\ &= |z||w|(\cos \varphi \cos \psi - \sin \varphi \sin \psi + i(\sin \varphi \cos \psi + \cos \varphi \sin \psi)) = \\ &= |z||w|(\cos(\varphi + \psi) + i \sin(\varphi + \psi)). \end{aligned}$$

Pokażmy jeszcze, że $|z| \cdot |w| = |zw|$. Niech $z = a + bi, w = c + di \in \mathbb{C}$. Wówczas

$$\begin{aligned} |z \cdot w| &= |(a + bi)(c + di)| = |(ac - bd) + (bc + ad)i| = \\ &= \sqrt{(ac - bd)^2 + (bc + ad)^2} = \sqrt{a^2c^2 + b^2d^2 + b^2c^2 + a^2d^2} = \\ &= \sqrt{(a^2 + b^2)(c^2 + d^2)} = |z| \cdot |w|. \end{aligned}$$

□

Wniosek 4.1.4: Wzór de Moivre'a, 1730

Niech $z = |z|(\cos \varphi + i \cdot \sin \varphi)$. Wówczas dla każdej dodatniej liczby całkowitej n mamy:

$$z^n = |z|^n(\cos(n \cdot \varphi) + i \cdot \sin(n \cdot \varphi)).$$

Nie uzasadniliśmy tymczasem, że \mathbb{C} jest ciałem. To, że aksjomaty ciała są istotnie spełnione w zasadzie sprowadza się do manipulacji algebraicznych oraz własności liczb rzeczywistych (dowód łączności mnożenia czy rozdzielności). Odnotujmy tylko, że dla każdego $(a, b) \neq (0, 0)$ mamy:

$$(a, b)^{-1} = \left(\frac{a}{a^2 + b^2}, \frac{-b}{a^2 + b^2} \right).$$

W ramach naszych zajęć ważna będzie jeszcze jedna definicja, związana z liczbami zespolonymi.

Definicja 4.1.5: Sprzężenie liczby zespolonej

iech $z = a + bi$ będzie liczbą zespoloną. SPRZĘŻENIEM liczby zespolonej z nazywamy liczbę $a - bi$, oznaczaną przez \bar{z} . Na płaszczyźnie zespolonej punkt \bar{z} jest obrazem z w symetrii względem osi $\text{Re}(z)$. W szczególności $|z| = |\bar{z}|$ oraz $\arg(z) = -\arg(\bar{z})$.

Przykłady: $\overline{1+i} = 1-i$, $\overline{-i} = i$, $\overline{\sqrt{2}} = \sqrt{2}$.

Istotnym wnioskiem z wzoru Moivre'a są następujące formuły.

Wniosek 4.1.6: Wzory na pierwiastki zespolone stopnia n

Jeśli $0 \neq w = |w|(\cos \theta + i \sin \theta) \in \mathbb{C}$, to pierwiastkami stopnia n z liczby w są liczby postaci:

$$\sqrt[n]{|w|} \left(\cos \frac{\theta + 2\pi k}{n} + i \sin \frac{\theta + 2\pi k}{n} \right), \quad \text{dla } k = 0, 1, \dots, n-1.$$

W szczególności pierwiastki stopnia n z 1 to liczby:

$$\zeta_n^k = \cos \frac{2\pi k}{n} + i \sin \frac{2\pi k}{n}, \quad \text{dla } k = 0, 1, \dots, n-1.$$

Nietrudno pokazać, że powyższe listy pierwiastków zespolonych stopnia n z liczby $0 \neq w \in \mathbb{C}$ są pełne. Jedyna liczba zespolona, która w n -tej potędze daje liczbę w to taka, której moduł równy jest $\sqrt[n]{|w|}$, a której argument po przemnożeniu przez n równy jest, z dokładnością do 2π , liczbie $\arg w$.

- Są dokładnie 3 liczby zespolone, które podniesione do potęgi 3 dają 2:

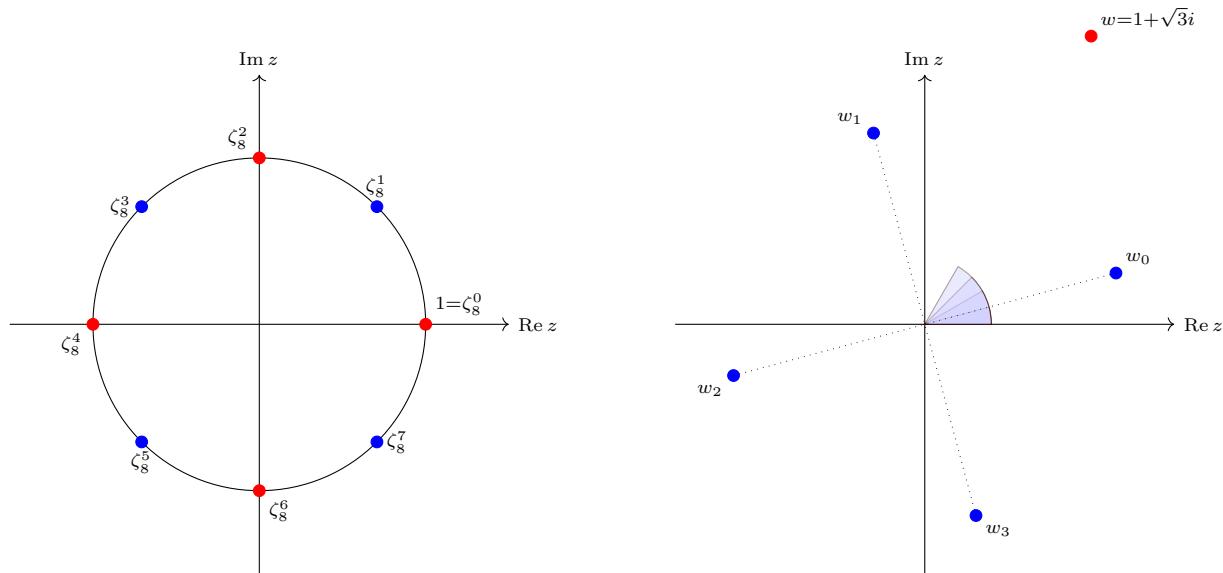
$$\sqrt[3]{2}(\cos 0 + i \sin 0), \quad \sqrt[3]{2}\left(\cos \frac{2\pi}{3} + i \sin \frac{2\pi}{3}\right), \quad \sqrt[3]{2}\left(\cos \frac{4\pi}{3} + i \sin \frac{4\pi}{3}\right).$$

Są dokładnie 4 liczby zespolone, które podniesione do potęgi 4 dają i:

$$\cos \frac{\pi}{8} + i \sin \frac{\pi}{8}, \quad \cos \frac{5\pi}{8} + i \sin \frac{5\pi}{8}, \quad \cos \frac{9\pi}{8} + i \sin \frac{9\pi}{8}, \quad \cos \frac{13\pi}{8} + i \sin \frac{13\pi}{8}.$$

Definicja 4.1.7: Pierwiastek pierwotny

Mówimy, że liczba $z \in \mathbb{C}$ jest PIERWIASTKIEM PIERWOTNYM stopnia n z 1, jeśli z jest pierwiastkiem stopnia n z 1, ale nie jest pierwiastkiem z 1 stopnia m , gdzie $m < n$.



Rys. 6. Po lewej — pierwiastki stopnia 8 z 1. Na niebiesko zaznaczono pierwiastki pierwotne stopnia 8 z 1. Po prawej — pierwiastki stopnia 4 z liczby $1 + \sqrt{3}i$ mają moduły równe $\sqrt[4]{2}$ oraz argumenty równe odpowiednio $\frac{\pi}{12}, \frac{4\pi}{12}, \frac{7\pi}{12}, \frac{10\pi}{12}$.

Nietrudno widzieć, że pierwiastki stopnia n z liczby zespolonej $w \neq 0$ o argumentie θ interpretować można jako wierzchołki n -kąta foremnego.

4.2 Wybór przykładowych pytań

Proszę się nie uczyć na pamięć odpowiedzi czy wyjaśnienie! Takich pytań można ułożyć znacznie więcej i służą one jedynie sprawdzeniu podstawowego zrozumienia treści zajęć. Jeśli odpowiedź na dowolne z poniższych pytań sprawia Państwu problem, proszę pomyśleć który fragment notatek dotyczy tego pytania, wrócić do ich lektury, a w razie dalszych trudności poprosić prowadzącego o wskazówkę.

1. Czy istnieje liczba zespolona, której część urojona równa jest i ?
2. Wyznacz moduł liczby $1 + i$ oraz moduł liczby $1 - i$.
3. Czy istnieje liczba zespolona, której moduł równy jest -1 ?
4. Czy istnieje liczba zespolona, której moduł równy jest i ?
5. Czy istnieją takie liczby rzeczywiste a, b , że $a + i = b - i$?
6. Czy istnieją takie liczby zespolone a, b , że $a + i = b - i$?
7. Czy istnieją takie liczby zespolone z , że $z^2 = -1$?
8. Czy istnieją takie liczby zespolone z , że $z^2 = i$?
9. Jaka jest część urojona liczby $\frac{1}{i}$?
10. Jaka jest część urojona liczby $\frac{1}{i+1}$?
11. Jaka jest część urojona liczby $\frac{i}{i+1}$?
12. Czy $\operatorname{Re} z = \operatorname{Re} \bar{z}$, dla dowolnej liczby zespolonej z ?
13. Czy warunek $\operatorname{Im} z = \operatorname{Im} \bar{z}$ implikuje $z = 0$?
14. Czy $z + \bar{z} \in \mathbb{R}$?
15. Czy $z - \bar{z} \in \mathbb{R}$?
16. Niech $a, b \in \mathbb{R}$. Rozłóż liczbę $a^2 + b^2$ na iloczyn liczb zespolonych.
17. Czy $z \cdot \bar{z} \in \mathbb{R}$?
18. Niech $|z_1| = |z_2|$. Czy $z_1 = \pm z_2$?
19. Wyznacz argument liczby $1 - i$.
20. Wyznacz argument liczby $1 - \sqrt{3}i$
21. Wyznacz argument liczby $2i$.
22. Wyznacz argument liczby $-\sqrt{2} + \sqrt{2}i$.
23. Argument liczby z równy jest θ . Jaki jest argument liczby $-z$?
24. Argument liczby z równy jest θ . Jaki jest argument liczby \bar{z} ?
25. Argument liczby z równy jest θ . Jaki jest argument liczby z^3 ?
26. Argument liczby z^2 równy jest π . Ile może być równy argument z ?
27. Argument liczby z^3 równy jest $\pi/2$. Ile może być równy argument z ?
28. Jakim warunkiem opisane są liczby zespolone z leżące na okręgu o promieniu 1 i środku i ?
29. Jakim warunkiem opisane są liczby zespolone z należące do koła o promieniu 2 i środku $1 + i$?
30. Jaka jest geometryczna interpretacja funkcji $f : \mathbb{C} \rightarrow \mathbb{C}$ danej wzorem $f(z) = z + 1$?
31. Jaka jest geometryczna interpretacja funkcji $f : \mathbb{C} \rightarrow \mathbb{C}$ danej wzorem $f(z) = z + i$?
32. Jaka jest geometryczna interpretacja funkcji $f : \mathbb{C} \rightarrow \mathbb{C}$ danej wzorem $f(z) = iz$?
33. Jaka jest geometryczna interpretacja funkcji $f : \mathbb{C} \rightarrow \mathbb{C}$ danej wzorem $f(z) = iz + 1$?

4.3 Zadania do samodzielnej pracy

Zadania ilustrujące umiejętności niezbędne do zaliczenia przedmiotu oznaczone są symbolem ♠.

Przy tych zadaniach dodany jest opis umiejętności, które sprawdzają.

1. (♠ Wykonywanie działań w ciele liczb zespolonych; część rzeczywista i urojona)

Wyznacz część rzeczywistą i część urojoną liczb zespolonych:

$$a) (2+i)(2-i) + (2+3i)(3+4i), \quad b) \frac{(5+i)(3+5i)}{2i}, \quad c) \frac{(1+3i)(8-i)}{(2+i)^2}, \quad d) \frac{(1-i)^4 - i}{(1+i)^4 + i}.$$

2. (♠) Czy następujący zbiór $X \subset \mathbb{C}$ jest skończony? Jeśli tak, to wypisz wszystkie jego elementy.

$$a) X = \{i^n \mid n \in \mathbb{N}\}, \quad b) X = \{(1-i)^n \mid n \in \mathbb{N}\}, \quad c) X = \left\{ \frac{(1+i)^{n+3}}{(1-i)^n} \mid n \in \mathbb{N} \right\}.$$

3. (♠ Wykorzystanie jednoznaczności części rzeczywistej i urojonej)

Znajdź wszystkie liczby zespolone z , które są rozwiązaniami równań:

$$\begin{aligned} a) \quad & (1+i)z^2 + (3-5i)z - 6 = 0, \\ b) \quad & 2z + 3\bar{z} - \operatorname{Re}(z) + 2\operatorname{Im}(z) = 8 - 3i, \\ c) \quad & |z| + 3\bar{z} = 2 + 6i. \end{aligned}$$

4. (♠ Rozwiązywanie układów równań liniowych o współczynnikach zespolonych)

Znajdź rozwiązanie ogólne układu równań liniowych

$$\begin{cases} (1-i)x_1 + ix_2 + 2x_3 - ix_4 = 1+i \\ (1+i)x_1 + x_2 + 2ix_3 + (1+2i)x_4 = 1-i \\ ix_1 + (-1+i)x_3 + ix_4 = 0 \end{cases}$$

5. (♠ Wyznaczanie postaci trygonometrycznej i stosowanie wzoru Moivre'a)

Poniższą liczbę $z \in \mathbb{C}$ przedstaw w postaci $a + bi$, gdzie $a, b \in \mathbb{R}$

$$a) z = (1-i)^{100}, \quad b) z = \left(\frac{i+\sqrt{3}}{i+1}\right)^{55}, \quad c) z = \frac{(\sqrt{3}+3i)^{40}}{(\sqrt{3}+i)^{20}}.$$

6. W zależności od $n \in \mathbb{N}$ oraz dla tych $x \in \mathbb{R}$, dla których to jest możliwe, wyznacz postać trygonometryczną liczby

$$z = \frac{(1+i \cos(x) + \sin(x))^n}{(1-i \cos(x) + \sin(x))^n}.$$

7. Wyznacz taki kąt θ , aby liczba $3+i$ była obrazem liczby $\sqrt{2} + 2\sqrt{2}i$ przy obrocie o kąt θ wokół punktu 0 (w kierunku przeciwnym do ruchu wskazówek zegara).

8. Niech $w = \frac{3+4i}{5}$. Znajdź taką liczbę zespoloną z , że $w = \frac{z}{\bar{z}}$. Wykaż, że każda liczba zespolona o module 1 jest ilorazem dwóch liczb zespolonych sprzężonych.

9. (♠ Szkicowanie prostych podzbiorów płaszczyzny zespolonej)

Naszkicuj następujące podzbiory zawarte w \mathbb{C} :

- $\{z \in \mathbb{C} \mid \operatorname{Im}(iz) < 0\}$,
- $\{z \in \mathbb{C} \mid \operatorname{Re}(1+i)z \geq 1\}$,
- $\{z \in \mathbb{C} \mid \operatorname{Im}(z^2) < 0\}$,
- $\{z \in \mathbb{C} \mid \operatorname{Im}(z^3) < \operatorname{Re}(z^3)\}$,
- $\{z \in \mathbb{C} \mid \operatorname{Re}(1+i)z \geq 1\}$,
- $\{z \in \mathbb{C} \mid \operatorname{Im}(1+i)z^2 < 0\}$,

10. Wykaż, że dla dowolnych liczb zespolonych z_1, z_2 mamy $|z_1 + z_2| \leq |z_1| + |z_2|$. Kiedy zachodzi równość? Wykaż też, że $\|z_1 - z_2\| \leq |z_1 - z_2|$.

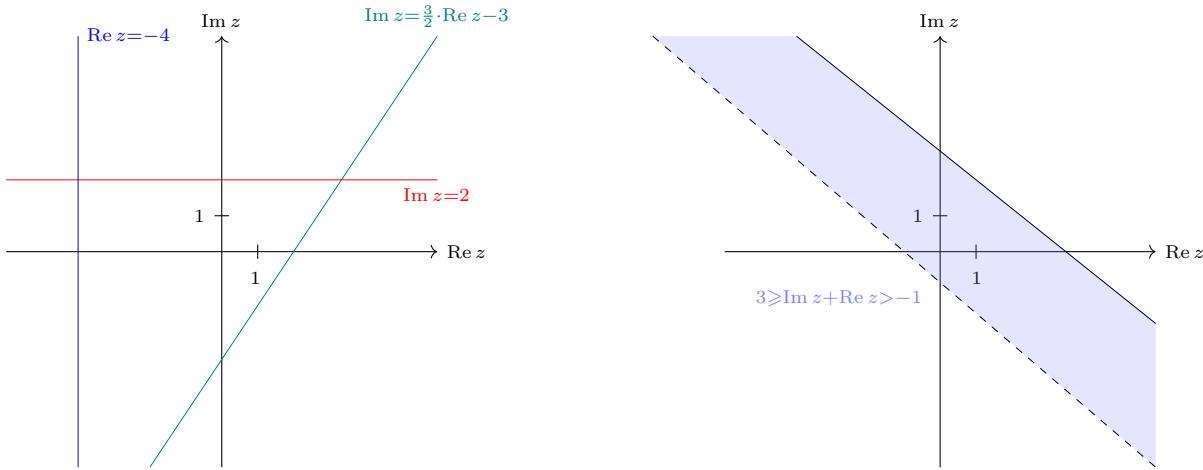
11. Liczba zespolona z spełnia warunek $|z| < 1$. Wykaż, że $|z^2 - z + i| < 3$.

12. Liczby zespolone z_1, z_2 spełniają warunek $|z_1| = |z_2| = 1$. Wykaż, że

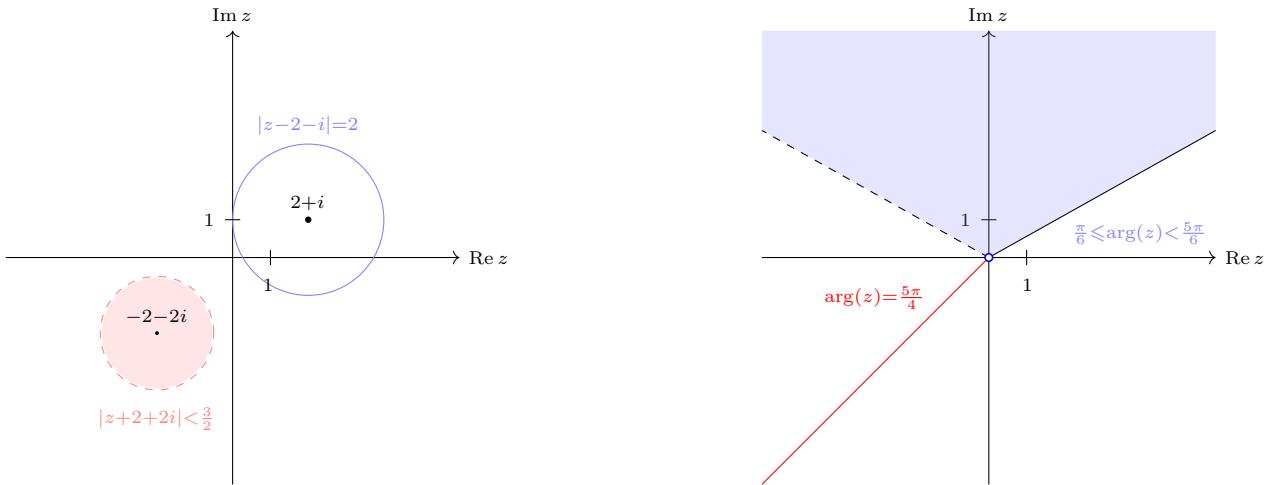
$$\frac{z_1 + z_2}{1 + z_1 z_2} \in \mathbb{R}.$$

4.4 Uzupełnienie. Geometria płaszczyzny zespolonej

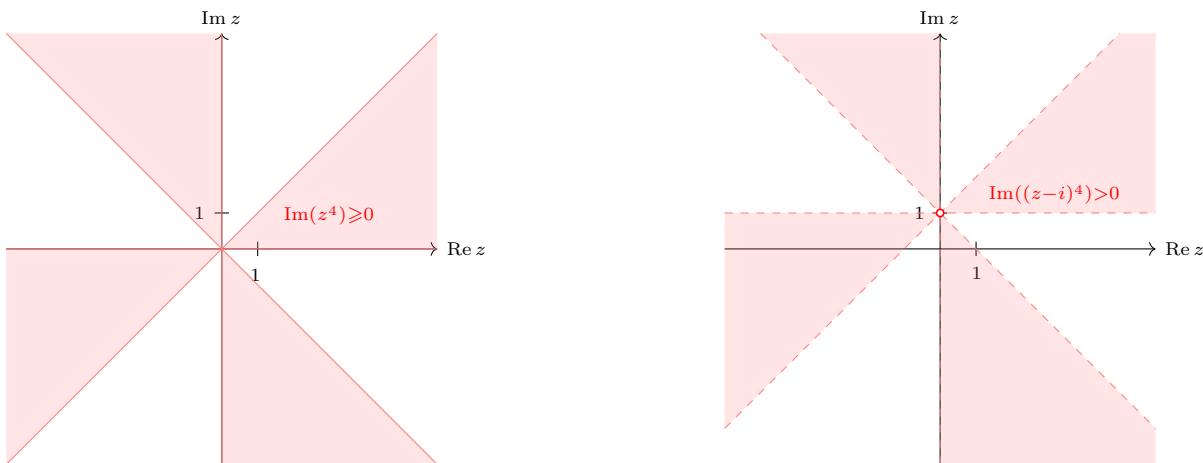
Ciało liczb zespolonych można rozumieć jako płaszczyznę z dodatkową strukturą algebraiczną. Struktura ta pozwala na wyrażenie w języku algebraicznym głębokich zależności geometrycznych. Zobaczmy kilka prostych przykładów.



Rys. 6. Każdy punkt z płaszczyzny zespolonej ma współrzędne $(\operatorname{Re} z, \operatorname{Im} z)$. Przechodząc do zapisu $z = x + yi$, współrzędne te wynoszą (x, y) . Wyznaczanie prostych i zbiorów ograniczonych przez półplaszczyzny wygląda więc podobnie jak w szkole.

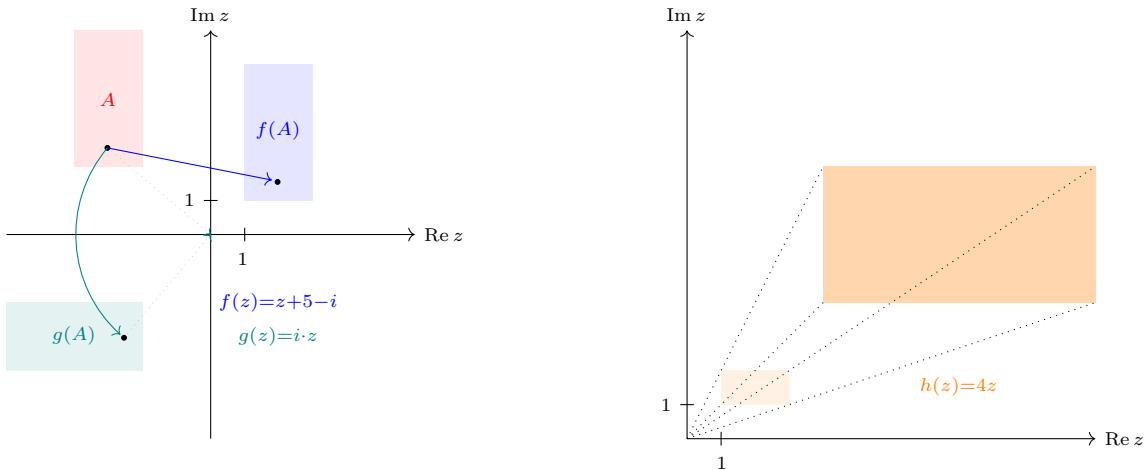


Rys. 7. Okrąg o środku w punkcie $z = a + bi$ i promieniu r opisany jest równaniem $|z - a - bi| = r$. Półprosta bez początku w punkcie $z = 0$ i o nachyleniu θ złożona jest z punktów z spełniających warunek $\arg z = \theta$.



Rys. 8. Zbiory opisane wyżej wyznaczamy zgodnie z postacią trygonometryczną. Jeśli $z = |z|(\cos \theta + i \sin \theta)$ to $z^4 = |z|^4(\cos 4\theta + i \sin 4\theta)$. Skoro więc $\arg(z^4) \in [0, \pi]$, to $\arg(z) \in [0, \frac{\pi}{4}] \cup [\frac{\pi}{2}, \frac{3\pi}{4}] \cup [\pi, \frac{5\pi}{4}] \cup [\frac{3\pi}{2}, \frac{7\pi}{4}]$.

W prostych funkcjach zespolonych rozpoznać można izometrie i podobieństwa płaszczyzny.



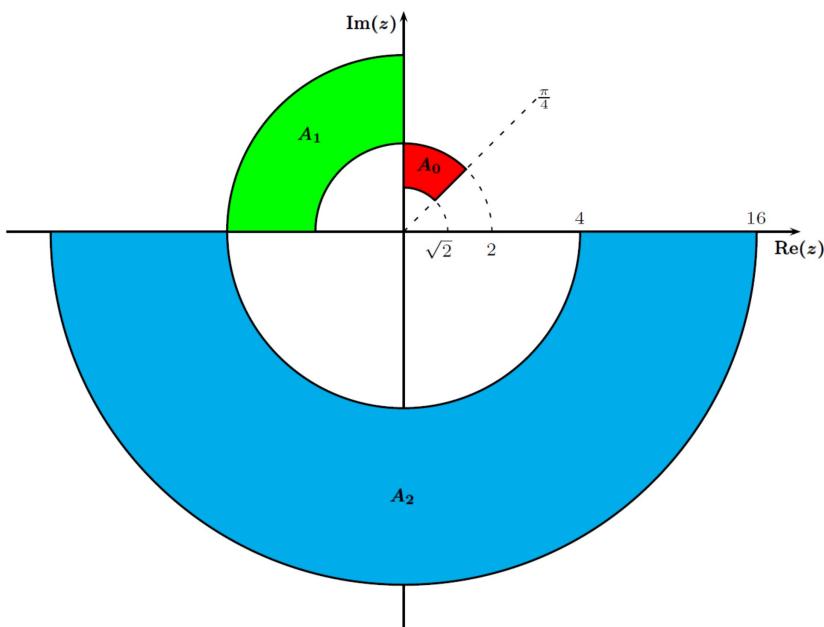
Rys. 9. Przesunięcie równoległe o wektor (a, b) opisane jest funkcją $z \mapsto z + a + bi$. Obrót wokół zera o kąt θ opisany jest wzorem $z \mapsto (\cos\theta + i \cdot \sin\theta) \cdot z$. Jednokładność o środku 0 i skali $t > 0$ można otrzymać poprzez przekształcenie $z \mapsto t \cdot z$.

* * *

Jeszcze inną sytuację opisuje zadanie z kolokwium z roku 2019. Definiujmy podzbiory \mathbb{C} postaci:

$$A_0 = \{z \in \mathbb{C} : \sqrt{2} \leq |z| \leq 2, \frac{\pi}{4} \leq \arg(z) \leq \frac{\pi}{2}\},$$

$$A_i = \{zz' : z, z' \in A_{i-1}\} \quad \text{dla } i > 0.$$



Rys. 10. Zbiory A_0, A_1, A_2 , przy czym dla czytelności zastosowano skalę logarytmiczną.

Dla $i > 1$ jeśli zbiór A_{i-1} jest zbiorem takich $z \in \mathbb{C}$, że $r_1 \leq |z| \leq r_2$ oraz $\phi_1 \leq \arg(z) \leq \phi_2$, to A_i jest zbiorem takich $z \in \mathbb{C}$, że $r_1^2 \leq |z| \leq r_2^2$ oraz $2\phi_1 \leq \arg(z) \leq 2\phi_2$. W szczególności

$$A_1 = \{z \in \mathbb{C} : 2 \leq |z| \leq 4, \pi/2 \leq \arg(z) \leq \pi\} \quad \text{oraz} \quad A_2 = \{z \in \mathbb{C} : 4 \leq |z| \leq 16, \pi \leq \arg(z) \leq 2\pi\}.$$

* * *

Liczby zespolone można z powodzeniem wykorzystywać w rozwiązywaniu nietrywialnych zadań z geometrii elementarnej, w tym zadań olimpijskich. Zainteresowanego Czytelnika odsyłam choćby do

- J. Jaszuńska, *Liczby zespolone w geometrii*, Delta 11/2010, <https://www.deltami.edu.pl/temat-matematyka/geometria/planimetria/2010/11/29/0905k25.pdf>
- E. Chen, *Bashing Geometry with Complex Numbers*, <https://web.evanchen.cc/handouts/cmplx/en-cmplx.pdf>.

4.5 Dodatek. Ciało liczb p -adycznych

W poprzednim rozdziale przedstawiliśmy ciało liczb p -adycznych jako zbiór „ułamków” tzw. całkowitych liczb p -adycznych. Struktura ta może być zdefiniowana w alternatywny sposób, wykorzystujący tzw. normę p -adyczną w ciele \mathbb{Q} . Będzie to uogólnienie klasycznej konstrukcji ciała \mathbb{R} , pochodzącej od Cantora, zakładającej rozważanie w ciele liczb wymiernych ciągów Cauchy'ego. Jak się okazuje, istnieje wiele sposobów ich definiowania. Aby mówić o zbieżności tych ciągów, wprowadzimy teraz pojęcie przestrzeni metrycznej i odległości, które należą do elementarnych pojęć działu matematyki zwanego topologią.

Definicja 4.5.1: Przestrzeń metryczna

Niech będzie X to niepustym zbiorem, zaś $d : X \times X \rightarrow [0, \infty)$ funkcją. Powiemy, że d jest METRYKĄ lub ODLEGŁOŚCIĄ na zbiorze X , jeśli spełnione są następujące założenia:

- (1) $d(x, y) = 0 \Leftrightarrow x = y$, dla każdych $x, y \in X$,
- (2) $d(x, y) = d(y, x)$, dla każdych $x, y \in X$,
- (3) $d(x, z) \leq d(x, y) + d(y, z)$, dla każdych $x, y, z \in X$.

Przykłady

- Funkcja $d(x, y) = |y - x|$ jest metryką w \mathbb{Q} , \mathbb{C} , a nawet w \mathbb{H} .
- Funkcja $d((x_1, y_1), (x_2, y_2)) = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2}$ w zbiorze \mathbb{R}^2 .
- Funkcja $d((x_1, y_1), (x_2, y_2)) = |x_1 - x_2| + |y_1 - y_2|$ w zbiorze \mathbb{R}^2 .

Wprowadzenie odległości pozwala na zdefiniowanie zbieżności

Definicja 4.5.2: Ciąg zbieżny

Niech X będzie niepustym zbiorem oraz d – metryką na X . Ciąg (x_n) elementów ze zbioru X nazywamy ZBIEŻNYM do punktu $x \in X$, ozn. $\lim_{n \rightarrow \infty} x_n = x$, jeśli dla każdego $\epsilon > 0$ istnieje $N \in \mathbb{N}$ takie, że dla każdego $n \geq N$ zachodzi $d(x_n, x) < \epsilon$.

Nietrudno widzieć, że w podobny sposób zdefiniować można pojęcie Ciągu Cauchy'ego w zbiorze X z metryką d . Przejdzmy do sytuacji, gdy X jest ciałem.

Definicja 4.5.3: Norma na ciele K

NORMĄ na ciele K nazywamy funkcję $\|\cdot\| : K \rightarrow [0, \infty)$ spełniającą warunki

- (1) $\|x\| = 0 \Leftrightarrow x = 0$, dla każdego $x \in K$,
- (2) $\|xy\| = \|x\| \cdot \|y\|$, dla każdych $x, y \in K$
- (3) $\|x + y\| \leq \|x\| + \|y\|$, dla każdych $x, y \in K$.

Przykłady

- Norma trywialna $\|\cdot\|$ na ciele K , określona warunkami: $\|0\| = 0$ oraz $\|x\| = 1$, dla $x \neq 0$.
- W ciałach \mathbb{Q} oraz \mathbb{R} mamy normy $\|x\| = |x|$, gdzie $|x|$ jest wartością bezwzględną.

Obserwacja 4.5.4: Ćwiczenie

Niech K będzie ciałem z normą $\|\cdot\|$. Funkcja $d : K \times K \rightarrow [0, \infty)$ dana wzorem $d(x, y) = \|x - y\|$ jest metryką na K . Mówimy, że jest to METRYKA INDUKOWANA przez normę $\|\cdot\|$.

Określimy teraz na zbiorze liczb wymiernych normy różne od wartości bezwzględnej.

Definicja 4.5.5: Norma p -adyczna

Niech p będzie dowolną liczbą pierwszą, zaś z – liczbą całkowitą.

- (1) Przez $v_p(z)$ oznaczamy największą liczbę całkowitą n taką że liczba p^n jest dzielnikiem liczby z , zwana WYKŁADNIKIEM p -ADYCZNYM LICZBY z .
- (2) Jeśli $x = \frac{a}{b}$, gdzie $a, b \in \mathbb{Z}$, $b \neq 0$, to określamy $v_p(x) = v_p(a) - v_p(b)$.
- (3) Normą p adyczną nazywamy funkcję $|.|_p : \mathbb{Q} \rightarrow [0, \infty)$ określoną wzorem:

$$|x|_p = \begin{cases} p^{-v_p(x)} & , x \neq 0 \\ 0 & , x = 0. \end{cases}.$$

Przykłady

- $|2|_2 = 2^{-v_2(2)} = \frac{1}{2}$,
- $|3|_2 = 2^{-v_2(3)} = 1$,
- $|4|_2 = 2^{-v_2(4)} = \frac{1}{4}$,
- $|- \frac{128}{7}|_2 = |\frac{2^7}{-7}|_2 = 2^{-v_2(2^7) + v_2(-7)} = 2^{-7} = 1/128$.
- $|13, 23|_3 = 1/27$, bo $13 + \frac{23}{100} = \frac{1323}{100} = \frac{3^3 \cdot 49}{100}$.

Wykładnik jest przydatnym elementarnym narzędziem tłumaczenia problemów podzielności na język algebraiczny. Poniższy rezultat jest natychmiastowym wnioskiem z twierdzenia o jednoznacznym rozkładzie liczb całkowitych na czynniki pierwsze.

Uwaga 4.5.6

Niech p będzie liczbą pierwszą, zaś a, b niech będą liczbami całkowitymi. Wówczas:

- (a) $v_p(ab) = v_p(a) + v_p(b)$ oraz $v_p(a/b) = v_p(a) - v_p(b)$,
- (b) $v_p(a^n) = nv_p(a)$, gdzie $n \in \mathbb{N}$
- (c) $v_p(NWD(a, b)) = \min\{v_p(a), v_p(b)\}$ oraz $v_p(NWW(a, b)) = \max\{v_p(a), v_p(b)\}$,
- (d) $v_p(a \pm b) \geq \min\{v_p(a), v_p(b)\}$. Gdy $v_p(a) \neq v_p(b)$, wówczas mamy równość.

Przykład. Niech p będzie liczbą pierwszą. Równanie $x^2 - py^2 = 0$ nie ma rozwiązań w liczbach całkowitych. Aby bowiem zachodziła równość $x^2 = py^2$ konieczne jest, aby $v_p(x^2) = v_p(py^2)$. Jedna strona tej równości jest liczbą parzystą $2v_p(x)$, druga zaś – liczbą nieparzystą $2v_p(y) + 1$, co jest niemożliwe. Zatem liczba \sqrt{p} jest niewymierna.

Przykład. Uzasadnijmy elementarną tożsamość $NWD(a, b) \cdot NWW(a, b) = a \cdot b$, zachodzącą dla dowolnych dodatnich liczb całkowitych. Dla dowolnej liczby pierwszej p mamy

$$v_p(NWD(a, b) \cdot NWW(a, b)) - v_p(ab) = \min\{v_p(a), v_p(b)\} + \max\{v_p(a), v_p(b)\} - v_p(a) - v_p(b).$$

Prawa strona jest równa 0, stąd iloraz liczby $NWD(a, b) \cdot NWW(a, b)$ przez ab jest równy 1, gdyż jest to jedyna dodatnia liczba całkowita n spełniająca warunek $v_p(n) = 0$, dla każdej liczby pierwszej p .

Uzasadnijmy, że wprowadzone normy p -adyczne są normamim, czyli indukują metrykę na \mathbb{Q} .

Uwaga 4.5.7

Norma p -adyczna jest normą na \mathbb{Q} .

Dowód. Jest jasne, że $|x|_p \geq 0$ przy czym równość zachodzi tylko dla $x = 0$. Niech:

$$x = \frac{p^a m}{n}, \quad y = \frac{p^b r}{s},$$

gdzie p jest liczbą pierwszą, która nie dzieli $r, s, m, n \in \mathbb{Z}$. Wówczas

$$xy = \frac{p^{a+b} mr}{ns}$$

oraz p nie jest dzielnikiem żadnej z liczb mr, ns , czyli $|xy|_p = |x|_p |y|_p$. Aby sprawdzić, że spełniony jest trzeci punkt definicji normy, bez straty ogólności możemy założyć, że $a \leq b$. Wtedy:

$$|x + y|_p = \left| \frac{p^a(sm + p^{b-a}nr)}{ns} \right|_p \leq p^{-a} = |x|_p.$$

Innymi słowy otrzymujemy, że: $|x + y|_p \leq \max\{|x|_p, |y|_p\} \leq |x|_p + |y|_p$. □

Definicja 4.5.8: Równoważność metryk

Powiemy, że metryki d_1, d_2 są *równoważne* na zbiorze X , jeśli ciąg (x_n) elementów z X jest Cauchy'ego względem d_1 wtedy i tylko wtedy, gdy jest ciągiem Cauchy'ego względem d_2 . Mówimy, że normy $\|\cdot\|_1$ i $\|\cdot\|_2$ są równoważne na w ciele K , jeśli indukują równoważne metryki.

Przykład. Na ciele \mathbb{Q} żadna z norm $|\cdot|_p$ jest równoważna formie $|\cdot|$, bo ciąg $x_n = p^n$ jest ciągiem Cauchy'ego względem $|\cdot|_p$, ale nie względem $|\cdot|$. Także jeśli $p_1 \neq p_2$ są liczbami pierwszymi, to ciąg $x_n = (p_1/p_2)^n$ spełnia: $|x_n|_{p_1} \rightarrow 0$ oraz $|x_n|_{p_2} \rightarrow \infty$.

Cytujemy bez dowodu istotne twierdzenie, motywujące nasze rozważania. Dowód można znaleźć na przykład w podręczniku Maurina.

Twierdzenie 4.5.9: Ostrowski

Każda nietrywialna norma na \mathbb{Q} jest równoważna jednej z norm $|\cdot|_p$ lub $|\cdot|$.

Zasadniczym twierdzeniem mówiącym o konstrukcji liczb p -adycznych jest następujący rezultat.

Twierdzenie 4.5.10

Niech K będzie ciałem z normą $\|\cdot\|$. Przez \widehat{K} oznaczamy zbiór klas abstrakcji relacji określonej na zbiorze ciągów Cauchy'ego na K postaci $[(x_n)]$, gdzie

$$[(a_n)] = [(b_n)] \Leftrightarrow (a_n) \sim (b_n) \Leftrightarrow \lim_{n \rightarrow \infty} \|a_n - b_n\| = 0.$$

Dla $k \in K$ niech \widehat{k} będzie ciągiem stale równym k . Wówczas na \widehat{K} określona jest struktura ciała z działaniami:

$$[(a_n)] + [(b_n)] = [(a_n + b_n)], \quad [(a_n)] \cdot [(b_n)] = [(a_n b_n)],$$

gdzie elementem zerowym jest $[\widehat{0}]$, zaś jedynką jest $[\widehat{1}]$.

Dowód. Prostym (choć nieco żmudnym) ćwiczeniem jest sprawdzenie, że działania $+, \cdot$ są dobrze określone na \widehat{K} , to znaczy: jeśli $(a_n) \sim (c_n)$, $(b_n) \sim (d_n)$, to:

$$[(a_n + c_n)] = [(b_n + d_n)], \quad [(a_n c_n)] = [(b_n d_n)].$$

Jedyną nietrywialną kwestią jest więc pokazanie, że jeśli ciąg Cauchy'ego $a_n \not\rightarrow 0$, to istnieje ciąg Cauchy'ego (b_n) taki, że $a_n b_n \rightarrow 1$.

Weźmy taki ciąg Cauchy'ego (a_n), że $\|a_n\| \neq 0$. Wówczas istnieje liczba dodatnia c oraz liczba całkowita dodatnia N , że dla $n > N$ mamy $\|a_n\| > c$. Określamy ciąg

$$b_n = \begin{cases} 0, & 1 \leq n \leq N-1 \\ a_n^{-1}, & n \geq N \end{cases}.$$

Ciąg (a_n) jest Cauchy'ego oraz dla $n, m \geq N$ mamy:

$$0 \leq \|b_m - b_n\| = \|a_m^{-1} - a_n^{-1}\| = \frac{\|a_m - a_n\|}{\|a_m\| \cdot \|a_n\|} \leq \frac{\|a_m - a_n\|}{c^2},$$

czyli ciąg (b_n) też jest Cauchy'ego. Oczywiście mamy

$$[(a_n)] \cdot [(c_n)] = [(a_n c_n)] = [(\underbrace{0, \dots, 0}_{N-1}, 1, 1, \dots)],$$

co znaczy, że $[(a_n)] \cdot [(c_n)] = [\widehat{1}]$. □

Powyższa konstrukcja pozwala na wspólne wprowadzenie zarówno liczb rzeczywistych, jak i ciała liczb p -adycznych.

Definicja 4.5.11: Liczby p -adyczne

Niech \mathbb{Q} będzie ciałem z metryką wyznaczoną przez normę p -adyczną. Wówczas ciało $\widehat{\mathbb{Q}}$ nazywamy CIAŁEM LICZB p -ADYCZNYCH, ozn. \mathbb{Q}_p .

Powyższa definicja nie jest oczywiście specjalnie użyteczna bez możliwości stosowania rozwinięć, podobnie jak ma to miejsce dla liczb rzeczywistych. Podobnie jak w wykładzie z analizy matematycznej, uzasadnić można następujący rezultat.

Twierdzenie 4.5.12: Reprezentacja liczb p -adycznych

Dla $0 < m \in \mathbb{Z}$ niech $d_{-m}, \dots, d_0, d_1, \dots$ będą nieujemnymi liczbami całkowitymi mniejszymi niż p , przy czym $d_{-m} > 0$. Rozważmy szereg:

$$d_{-m}p^{-m} + d_{-m+1}p^{-m+1} + \dots + d_0 + d_1p + d_2p^2 \dots \quad (*).$$

Wówczas sumy częściowe szeregu (*) tworzą ciąg Cauchy'ego w \mathbb{Q} (względem $|\cdot|_p$) i dla każdego elementu $A \in \mathbb{Q}_p$ istnieje dokładnie jeden reprezentujący go ciąg Cauchy'ego (A_i), którego wyrazami są sumy częściowe szeregu (*), spełniające warunki

- (i) $0 \leq A_i < p^i$, dla $i = 1, 2, \dots$
- (ii) $A_i \equiv A_{i+1} \pmod{p^i}$, dla $i = 1, 2, \dots$

Przykłady

- Liczba 320 ma w ciele \mathbb{Q}_7 rozwinięcie $5 + 3 \cdot 7 + 6 \cdot 7^2 = 635$.
- Liczba -1 ma w ciele \mathbb{Q}_7 przedstawienie $-1 = 6 + 6 \cdot 7 + 6 \cdot 7^2 + 6 \cdot 7^3 = \dots 6666$.
- Liczba $\frac{1}{2}$ ma w ciele \mathbb{Q}_5 przedstawienie $3 + 2 \cdot 5 + 2 \cdot 5^2 + 2 \cdot 5^3 + \dots = \dots 2223$.

W zbiorze liczb \mathbb{Q}_p wprowadzić można normę pochodząą od normy p -adycznej, podobnie jak definiuje się w zbiorze \mathbb{R} wartość bezwzględną $|\cdot|_p$, jako przedłużenie odpowiedniej normy p -adycznej $|\cdot|_p$ na zbiorze liczb wymiernych. Pominiemy formalną konstrukcję, pozostawiając jednak pewien komentarz do twierdzenia wyżej.

Liczby x oraz y mają te same i cyfr w rozwinięciu p -adycznym (od prawej), jeśli $|x - y|_p \leq p^{-i}$. Mówiąc intuicyjnie, liczby p -adyczne są tym bliżej siebie, im wyższą potegę liczby p wspólnie dzielą. Na przykład w ciele \mathbb{Q}_3 mamy

$$|81 - 1|_3 = |80|_3 = 1, \quad |81 - 27|_3 = |54|_3 = \frac{1}{27}, \quad |81 - 80|_3 = |1|_3 = 1.$$

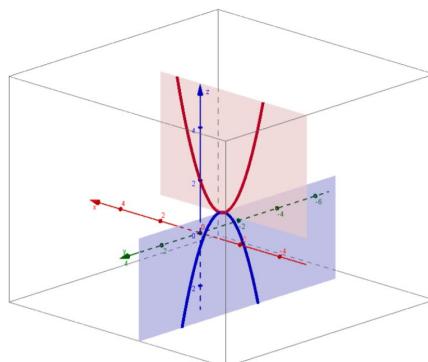
4.6 Trivia. Wykres funkcji zespolonej?

W przeciwnieństwie do funkcji zmiennej rzeczywistej, której przebieg zmienności prezentujemy często graficznie na płaszczyźnie kartezjańskiej, prezentowanie „wykresów” funkcji zmiennej zespolonej nie jest czytelne z uwagi na to, że wymagałoby operowania w przestrzeni czterowymiarowej (zbiór argumentów ma dwie współrzędne i zbiór wartości ma dwie współrzędne). Z uwagi jednak na to, że często interesuje nas rozwiązanie równania $f(z) = 0$, wprowadza się różne ciekawe metody wizualizacji tego problemu. Jedną z nich są tzw. krzywe bliźniacze, wprowadzone w jednym z podręczników licealnych (!) w USA w latach 50' przez Howarda Fehra (to były poczatki „New Math” w nauczaniu – kto by chciał przeczytać więcej polecam artykuł: *New Thinking in School Mathematics* słynnego matematyka J. Dieudonné'a).

Rozważmy funkcję $f(z) = z^2 + 2z + 2$. Nietrudno sprawdzić, że rozwiązaniami równania $f(z) = 0$ są liczby zespolone $-1 \pm i$. Jak to zobaczyć na „wykresie”? Niech $z = x + iy$. Wówczas:

$$f(z) = f(x + iy) = (x^2 - y^2 + 2x + 2) + (2y(x + 1))i.$$

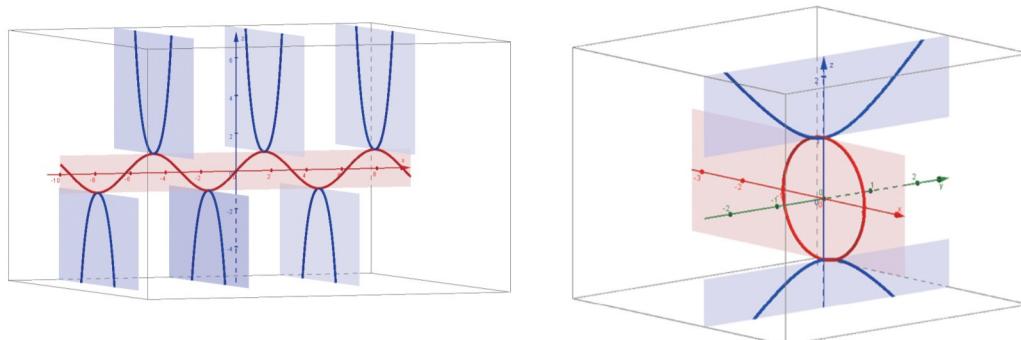
Skoro interesuje nas jedynie prezentacja warunku dotyczącego rzeczywistej wartości funkcji f (chodzi nam o wartość zero), to pomyślmy dla jakich x, y powyższa funkcja przyjmuje jedynie wartości rzeczywiste? Oczywiście dla $y = 0$ lub $x = -1$. Na płaszczyźnie $y = 0$ wartości $f(z)$ dane są przez $f(x) = x^2 + 2x + 2$, $x \in \mathbb{R}$, co reprezentowane jest przez dobrze znaną parabolę. W płaszczyźnie $x = -1$, prostopadlej do płaszczyzny $y = 0$, funkcja nasza ma postać $f(-1 + yi) = -y^2 + 1$, $y \in \mathbb{R}$. Oto stosowny obrazek:



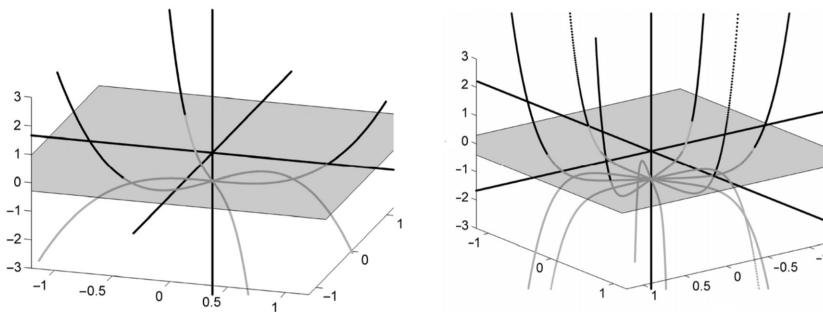
Rys. 1. Krzywe bliźniacze funkcji $f(z) = z^2 + 2z + 2$ w przestrzeni \mathbb{R}^3 to $x^2 + 2x + 2$ w płaszczyźnie $y = 0$ oraz $-y^2 + 1$ w płaszczyźnie $x = -1$. Źródło: Wiggins H., Harding A., Engelbrecht J.: *Visualising Complex Polynomials: A Parabola Is but a Drop in the Ocean of Quadratics*. J. Math. Research 10 (2018)

Co ten obrazek nam w zasadzie mówi? Otóż pokazuje nam on fragment czterowymiarowego wykresu funkcji $f(z)$ – ten mianowicie, na którym wartości funkcji są jedynie liczbami rzeczywistymi. Te wartości rzeczywiste reprezentowane są na osi OZ. Inaczej mówiąc: każda z powyższych dwóch krzywych ma punkty o trzech współrzędnych: (x, y, z) . Pierwsze dwie współrzędne „kodują” punkt $x + iy$ z dziedziny funkcji f , zaś współrzędna z zawiera wartość rzeczywistą funkcji $f(z)$. A zatem zgodnie z intuicją: czerwona parabola nie ma punktu o współrzędnej $z = 0$, natomiast niebieska parabola – owszem: przecina płaszczyznę $z = 0$ w punktach $(-1, -1)$ oraz $(-1, 1)$. Te punkty reprezentują oczywiście liczby zespolone $-1 \pm i$.

Podobnego typu obrazki generować można dla innych funkcji, korzystając niekiedy z postaci trygonometrycznej lub wykładniczej liczb zespolonych. Ciekaw jestem czy Czytelnik potrafiłby powiedzieć jakimi równaniami opisane są krzywe bliźniacze dla funkcji $f(z) = \sin(z)$ oraz dla krzywej postaci $y^2 + z^2 = 1$?



Pouczająco wyglądają także obrazki prezentujące rozwiązania równań $z^3 = 1$ lub $z^6 = 1$.

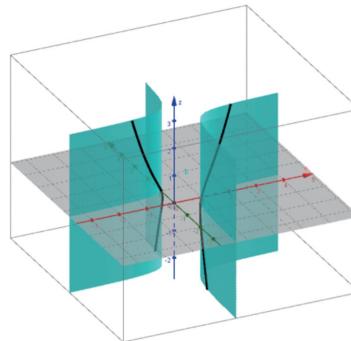


Rys. 3. Krzywe bliźniacze funkcji $f(z) = z^3 - 1$ oraz $f(z) = z^6 - 1$. Źródło: Harding A., Engelbrecht J.: *Sibling curves and complex roots 2: Looking ahead*. International Journal of Mathematical Education in Science and Technology, 38 (2017), 975-985.

Dla „funkcji kwadratowych” (zmiennej zespolonej) postaci $f(z) = az^2 + bz + c$, gdzie $a, b, c \in \mathbb{C}$, $a \neq 0$, pokazuje się, że zachodzić musi jedna z wykluczających się dwóch sytuacji:

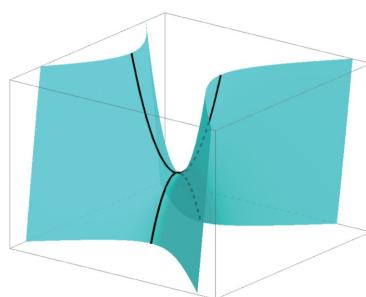
- dwie krzywe bliźniacze przecinają się – i wtedy są dwiema parabolami,
- dwie krzywe bliźniacze nie przecinają się – i stanowią gałęzie hiperboli.

Druga sytuacja zachodzi np. dla $f(z) = (z - 1)(z - i)$, gdzie krzywe bliźniacze dane są wzorem $y = \frac{x-1}{2x-1}$.



Rys. 4. Krzywe bliźniacze funkcji $f(z) = (z - 1)(z - i)$. Źródło jw.

Nie ma w tym rozróżnieniu, jak się okazuje, nic dziwnego. Jak pokazać, że taka klasyfikacja ma miejsce? Zasadniczo chodzi o sprowadzenie funkcji do postaci kanonicznej i rozważanie jedynie jej – jako geometrycznie „istotnej” dla kształtu krzywych bliźniaczych. Przez użycie przesunięcia, skalowania i obrotu można, bez straty ogólności, rozważyć jedynie krzywe bliźniacze dla równania $f(z) = z^2 + c$, dla pewnej liczby zespolonej c . Przyjmując $z = x + iy$ dostajemy, że krzywe bliźniacze zawsze leżą, z dokładnością do skalowania, przesunięcia czy obrotu, na tak zwanej **hiperboloidzie parabolicznej** $z = x^2 - y^2$.



Rys. 5. Krzywe bliźniacze funkcji $f(z) = z^2 - 1$ na paraboloidzie hiperbowlicznej $x^2 - y^2 = z$.

Należałyby, rzecz jasna, uściślić co znaczy stwierdzenie, że przesunięcie, skalowanie i obrót nie zmieniają „istoty geometrycznej” rozważanego problemu? Dlaczego wystarczyło rozważyć jedynie krzywe bliźniacze równania $f(z) = z^2 + c$? Znacznie ogólniejsze i bardziej szczegółowe wyjaśnienie otrzymacie Państwo w drugim semestrze, gdy rozważyć będziemy tak zwaną aficzną klasyfikację hipertpowierzchni stopnia 2.

Animację ukazującą zmianę położenia krzywych bliźniaczych na paraboloidzie hiperbowlicznej dla rodzin funkci o równaniach $f(z)z^2 + 2z + (1 + ki)$ w zakresie $-2 \leq k \leq 2$ znajdziecie Państwo pod adresem:
<https://cardanogroup.files.wordpress.com/2014/08/sibling-animation.gif>.

4.7 Coda. O kształtowaniu się pojęcia liczby

Liczby zespolone, a zwłaszcza ich trudny do uchwycenia przez stulecia aspekt bycia pierwiastkami z liczb ujemnych (których przecież *nie powinno* być!), każde przyjrzeć się nieco szerzej całej gamie przykładów historycznych sytuacji, w której jakimś liczbom odmawiano prawa do istnienia, do czasu gdy stojące za nimi koncepcje zostały nie tylko ugruntowane teoretycznie, ale i gdy ich użyteczność stała się w zasadzie ewidentna. Wielkie i historyczne kontrowersje dotyczyły nie tylko liczb zespolonych, o których historii powiemy w dalszej części, ale również innych rodzin liczb — niewymiernych, ujemnych, a nawet zera. Pogłębienie rozumienia koncepcji liczby znamionowało zawsze istotny przełom w matematyce.

Zacznijmy od liczb naturalnych. Czy kiedykolwiek ich nie akceptowano? Z pewnością małe liczby — owszem, ale czy również duże? Największa liczba mająca samodzielna nazwę u starożytnych Greków to *myrias*, czyli 10 000. Stąd największa wówczas — wciąż obecna w kulturze nazwa: miriady miriad — odnosi się do liczby stu milionów. Badacze starych dialektów oraz języków ludów pierwotnych zauważają, że budowa liczebników ma swoiste addytywną strukturę — w odróżnieniu od stosowanej przez nas notacji pozycyjnej, opartej na potęgowaniu. I tak w papuaskim plemieniu Wedau, liczba 2 ma nazwę *ruag'a*, a liczba 4 — *ruag'a ma-ruag'a*, czyli 2 + 2. Liczba 5 to *ura-i-ga*, a liczba 9 — *ura-g'ela-ruag'a-mu-ruga'a*, czyli 5 i 2 + 2. Nawet zapis łaciński cyfr rzymskich nie kojarzy nam się z wielkimi liczbami. Imperium Cezara interesujące się głównie praktycznymi zastosowaniami, wyrażało milion jako *dieces centena milia*, czyli dziesięć setek tysięcy. Samo słowo „milion” pojawiło się dopiero w trzynastowiecznej Francji.

Czy duże liczby nie istniały więc w starożytności? Istniały, choć nie znalazły wykorzystania. Były raczej ucieleśnieniem koncepcji — liczby nie są ograniczone. Archimedes oraz Diofantos dali początki notacji potęgowej, mówiąc o potęgach *miriady miriad* (osiągając liczbę $10^{8 \cdot 10^{16}}$, czyli miriadę miriad podniesioną do miriady miriad, która to jest podniesiona do potęgi o wykładniku miriady miriad). Inaczej sytuacja wyglądała na dalekim Wschodzie. Już starożytny Sanskryt zawiera nazwy potęg dziesiątki aż do 10^{12} . Późniejsze teksty hinduistyczne i buddyjskie rozszerzały zakres nazwanych liczb najpierw do 10^{421} , a później nawet do $10^{10 \cdot 2^{122}}$, nazwanej liczbą niewypowiadalną. Teksty tych kultur jeszcze przed początkiem naszej ery dzieliły liczby na trzy kategorie: policzalne (najmniejsze, pośrednie i najwyższe), niepoliczalne (prawie, istotnie i niepoliczalnie niepoliczalne) oraz nieskończoności (prawie, istotnie i nieskończonie).

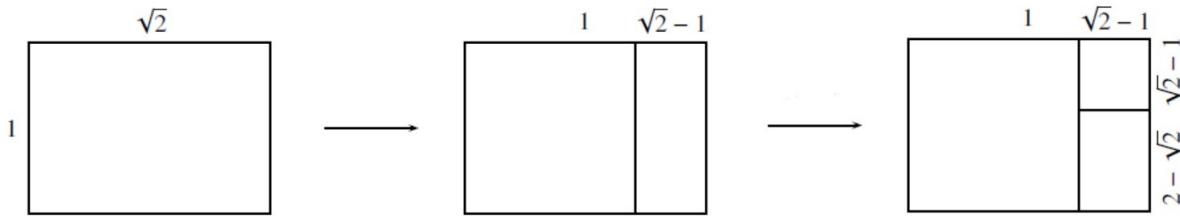
A co z samą nieskończonością? Temu zagadnieniu należałoby poświęcić osobny dodatek. Arystoteles wyróżniał nieskończoność potencjalną oraz aktualną. Ta pierwsza oznacza, że jakąkolwiek podamy wielkość geometryczną czy liczbową, zawsze znajdzie się wielkość (liczba) od niej większa. Tego rodzaju nieskończoność akceptowana była od starożytności. Nie zakładała ona istnienia nieskończonego bytu, a jedynie nieustającą możliwość powiększania lub pomniejszania, i tak stosowały ją choćby *Elementy* Euklidesa. Również nowożytne przejścia graniczne pozostają w istocie w obszarze nieskończoności potencjalnej.

Nieskończoność aktualna, to nieskończoność dokonana, to jest taka, w której nieskończonie wiele kroków jest wykonanych w jednym działaniu, np. poprzez wzięcie nieskończonego zbioru liczb naturalnych. Tego typu nieskończoność odrzucono zarówno w starożytności, jak i w epokach nowożytnych w zasadzie aż do XIX wieku, do badań Cantora. Wiązała się ona z lękiem przed paradoksami, począwszy od słynnego paradoksu Zenona, do współczesnych paradoksalnych rozkładów czy innych „pułapek” teorii mnogości. Stąd nawet XX-wieczny wielki matematyk Henri Poincaré powie, że „Następne pokolenia potraktują teorię zbiorów jako chorobę, z której udało im się wyleczyć”. Znany nam symbol nieskończoności ∞ pochodzi od Wallisa (1655). Symbole \aleph_0 oraz c związane z liczbami kardynalnymi pochodzą od Cantora.

Można by przypuszczać, że po dodatnich liczbach naturalnych służących do zliczania, naturalnym etapem rozwoju będą liczby ujemne, ale stało się inaczej. W istocie, wcześniej rozważane były ułamki (podobnie jest zresztą w szkole). Jak czytaliśmy w dodatku do wykładu pierwszego, już Babilończycy posługiwali się połówkami, ćwiartkami czy trzecimi częściami całości. Również zapis hieroglificzny stosowany w starożytnym Egipcie dopuszczał odwrotności liczb całkowitych, wraz z przekonaniem, że każdy ułamek można zapisać jako sumę takich odwrotności (i odpowiedniej liczby całości), czego oczywiście nikt wówczas nie dowodził — zrobił to dopiero Fibonacci w 1202 roku. Idea egipskiego zapisu ułamków przeniknęła do świata greckiego, gdzie przestrzegał go jak się wydaje sam Euklides. Egipcjanie mieli zresztą specjalne tablice rozkładu liczb $2/n$ na ułamki, a dla ułamka $2/3$ istniał osobny hieroglif. To, co jest interesujące z punktu widzenia historii rozwoju pojęcia liczby, to wstrząsające odkrycie, że nie wszystkie wielkości są proporcjonalne, które zachwiało filozoficznymi podstawami szkoły Pitagorejczyków.

Dlaczego dla greckiego świata odkrycie wielkości nieproporcjonalnych było aż tak wielkim szokiem? Pitagorejczycy wierzyli w spajającą świat *harmonię*, która ma być dla ludzi — jak pisze prof. Kordos w swoich *Wykładach z historii matematyki* — motywacją życia i pełnią człowieczeństwa. Pitagorejczycy wyróżnili dziedziny, w których harmonia jest najbardziej widoczna: muzykę, astronomię, arytmetykę i geometrię (późniejsze *quadrivium* stanowiące wyższy stopień kształcenia średniowiecznego, obok niższego *trivium* — gramatyki, retoryki i dialektyki). Oni to odkryli, że jeśli długości dwu napiętych jednakową siłą strun mają się jak 1 do 2, to wydają one harmonijne brzmienie. Podobnie z innymi stosunkami długości: 2:3, czy 3:4, które dały podstawy do wyróżnienia interwałów: oktawy, kwarty i kwinty. Tego typu głębokie związki budziły niezwykłe zainteresowanie. Koncepcja powiązania wielkości proporcjonalnych ze zjawiskami natury była tak atrakcyjna i narzucająca się, że powstała hipoteza, że w liczbach należy szukać istoty harmonii. Również z matematycznego, czy raczej — geometrycznego punktu widzenia, proporcje były w centrum rozważań pitagorejskich. Stąd rozbudowana teoria figur podobnych, skumulowana wokół twierdzenia Talesa. Z punktu widzenia arytmetyki (wciąż uprawianej geometrycznie) — kluczem był algorytm Euklidesa. Co ciekawe nie dotyczył on w istocie liczb całkowitych, ale *współmiernych*.

Wiemy dobrze, że długość boku kwadratu i jego przekątnej nie jest wspólna — znamy różne dowody tego faktu. Szczególnie interesujący polega na następującej obserwacji. Rozważmy prostokąt o bokach długości $a = 1$ oraz $b = \sqrt{2}$ (czy inaczej — jeden z boków to bok kwadratu, a drugi — ma długość jego przekątnej). Reprezentujemy odejmowanie mniejszej liczby od większej poprzez obcięcie z wyjściowego prostokąta kwadratu o krótszym boku. Zatem w dwóch krokach otrzymamy prostokąt o bokach długości $\sqrt{2} - 1$ oraz $\sqrt{2} - 2 = \sqrt{2}(\sqrt{2} - 1)$, mający taki sam kształt, jak wyjściowy prostokąt, przy czym dłuższy bok jest teraz pionowy, a krótszy poziomy. Z tego wynika, że opisany proces nigdy się nie skończy, a przecież dla liczb wymiernych jest inaczej! Oczywiście liczby niewymierne $\sqrt{2}$ i $3\sqrt{2}$ są wspólna.



Przejdzmy teraz do liczb ujemnych. Rozumiejąc już, że w ujęciu starożytnych Greków liczby stanowiły głównie reprezentację obiektów geometrycznych, nietrudno rozumieć, że nie stosowano liczb ujemnych. Długości, pola, objętości musiały być dodatnie. Około trzeciego wieku aleksandryjski matematyk Diofantos napisał wielkie i ważne dzieło *Arytmetyka*, w którym przedstawił zbiór problemów wraz z zaczątkami symboliki służącej do ich rozwiązywania. W jednym z rozwiązań Diofantos zapisze równanie, które dziś czytalibyśmy jako $4 = 4x + 20$, a które nazwie *absurdałnym*.

W istocie prawa działań na liczbach ujemnych zostały sformułowane już w VII wieku przez hinduskiego uczonego Brahmaguptę. Sięgając jeszcze głębiej w przeszłość, ślady liczb ujemnych znaleźć można w starożytnych chińskich dziełach, zawierających ciekawą koncepcję (dydaktyczną) oznaczania liczb dodatnich kolorem czerwonym, a ujemnych — kolorem czarnym. Wszystko to wprowadzano w kontekście obliczeń finansowych i podatkowych, w których liczby czarne bilansowały czerwone. Kwota sprzedaży była czerwona (otrzymujemy pieniądze), a która wydana na zakup była czarna (trzeba wydać). Bilans był dodatni, a deficyt — ujemny. Koncepcja ta pochodziła również z astronomii, gdzie przybliżano liczby z góry i z dołu, Przybliżenia z góry traktowano jako silne, a z dołu — jako słabe. Również Brahmagupta używał nie tylko liczb ujemnych, ale i specjalnego oznaczenia dla liczb ujemnych, a także liczby zero, wraz z odpowiednimi prawami działań. Co ciekawe, matematyka arabska, choć świadoma była osiągnięć hinduskich, odrzucała liczby ujemne, podobnie jak matematyka europejskiego Średniowiecza.

Nawet szesnastowieczni uczeni rozwiązujący równania wielomianowe metodą geometryczną, nie akceptowali do końca używania współczynników ujemnych, choć stosowali odpowiednią symbolikę. John Wallis (1616-1703) osowił nieco liczby ujemne poprzez wprowadzenie osi liczbowej. Jednak współczesny mu Kartezjusz — autor układu współrzędnych, zaniedbywał współrzędne ujemne. Dopiero rozwój rachunku wektorowego, począwszy od Galileusza, da liczbom ujemnym solidną pozycję w matematyce. Rozmaite wątpliwości pozostały jednak na dłużi czas. Nie rozumiano choćby znaczenia iloczynu $(-1) \times (-1)$.

Jeszcze w 1770 roku Euler „dowodzi” w swojej *Algebrze*, że $\sqrt{-2} \cdot \sqrt{-3} = \sqrt{6}$. Między Leibnizem, Eulerem, Bernoullim i d'Alembertem zaistniał poważny spór o to czy $\log(-x)$ jest tym samym, co $\log(x)$. W ten sposób doszło do pewnego rozdrożenia: liczb ujemnych (i zespolonych) używano jako narzędzi formalnych do uzyskiwania rozwiązań, choć niekoniecznie przypisywano im niezależny byt. Jeszcze w 1758 roku brytyjski matematyk Maseres napisze, że liczby ujemne zacieśniają całą naukę o równaniach i zacieśniają rzeczy, które w swej naturze są zbyt oczywiste i proste.

Zanim przejdziemy do liczb zespolonych, dodajmy jeszcze kilka zdań o zerze, które przybyło do Europy wraz z Fibonaccim (który jako syn kupca podróżował po basenie Morza Śródziemnego poznając matematykę hinduską oraz arabską) i popularyzowanym przezeń zapisem dziesiętnym. Pierwotnie zero funkcjonowało jako symbol pozwalający odróżnić 1 od 10 lub 100. Mimo rozbieżności zdań historyków co do tego kto pierwszy używał go w takim kontekście, wiadomo na pewno, że symbol ten wystąpił w hinduskiej notacji w roku 876 (już w formie zera, a nie kropki, czy innego symbolu). Jeśli chodzi o zero jako liczbę, rozważane było ono już przez Brahmaguptę, który starał się jednocześnie określić prawa działań na liczbach, i jako jeden z pierwszych natknął się na problem dzielenia przez zero, określając przy tym absurdalne prawa w rodzinie $0/0 = 0$. Również później uczeni hinduscy zmagali się z tym wyzwaniem, aż do żyjącego w XII wieku Bhaskary, który próbował przypisać ilorazowi $n/0$ wartość nieskończoną, oczywiście łącząc to z językiem religijnym. Sformułuje jednak poprawnie prawa $0^2 = 0$ oraz $\sqrt{0} = 0$.

Wielkim osiągnięciem matematyki arabskiej, a zwłaszcza Khwarizmiego było opisanie indyjskiego systemu pozycyjnego, oraz praw działań pisemnych, przejętych później przez Fibonacciego i Europejczyków (nie bez kontrowersji — zwyciężył pragmatyzm i szybkość, z jaką dokonywano rachunków). Również w Europie nie od razu akceptowano zero. Ufundowana na grecko-rzymskiej filozofii, i na dziełach Arystotelesa, nauka europejska była raczej przeciwna idei nicości. Poza tym zero łatwo był pomylić z cyfrą 6 lub 9, co powodowało nawet czasowe wyłączanie go z użycia (na przykład we Florencji na początku XIV wieku). Ważnym osiągnięciem, obecnym już u Claviusa (1608) było jednakże rozważanie zera jako współczynnika w równaniu wielomianowym i świadomość, że jego rozwiązywanie może być uzyskane przez rozkład na czynniki liniowe. Twierdzenie Bezout w istocie zdaje się pochodzić od Kartezjusza, a sformułowane jest w jego dziele *Geometria* z roku 1637. W roku 1657 wspomniany już Wallis zadeklaruje, że zero nie jest liczbą, a wspomniany już symbol nieskończoności wprowadzi właśnie dla oznaczenia wyniku $1/0$.

Słynny filozof i matematyki Alfred North Whitehead napisze w roku 1911 w swoim *Wstępie do Matematyki*, że „nikt nie idzie do sklepu, aby kupić zero ryb”. Doda jednak zaraz, że użycie zera jest na nas wymuszone potrzebami utartych już schematów myślenia. Być może nie utarły się one jednak całkowicie. W 2000 roku ludzkość witała „nowe millenium”, choć w istocie tak trzecie tysiąclecie, jak i XXI wiek rozpoczęły się 1 stycznia 2001 roku, z uwagi na brak roku zerowego.

* * *

Co z liczbami zespolonymi? Definicja, którą oglądaliśmy na wykładzie pochodzi z roku 1833 i jest zasługą Hamiltona — odkrywcy kwaternionów. Była ona swego rodzaju ukoronowaniem 300 lat wysiłków matematyków, którzy od 1545 roku mierzyli się z konsekwencjami wyników opisanych w dziele *Ars Magna* (Wielka Sztuka) autorstwa (jak mówią historycy — wybitnego uczonego, ale oszusta) Girolamo Cardano. Wyjawiona w nim była metoda rozwiązywania równań wielomianowych stopnia trzeciego i czwartego zakładająca konieczność wyciągania pierwiastków z liczb ujemnych, a która opracowana została w połowie XVI wieku przez del Ferro, Targaglię oraz Cardano. Ich rozwiązywanie równania trzeciego stopnia postaci

$$y^3 = py + q.$$

Historia ta warta jest opowiedzenia. Pierwszą osobą, która znalazła rozwiązywanie powyższego równania był boloński profesor Scipione del Ferro (1465-1526). Jego ojciec Floriano pracował w przemyśle drukarskim i już w wieku młodzieńczym Scipione miał dostęp do wielu klasycznych prac, w tym oczywiście do „Liber Quadratorum” (Księga Kwadratów) Fibonacciego. Czytelnika zaskoczy być może fakt, że nie zachowały się żadne prace del Ferro. Wydaje się, że obawiał się wyzwania i ewentualnej utraty pozycji na uniwersytecie w Bolonii. Trzymał więc swoją pracę w ukryciu, dzieląc się jedynie z najbliższymi uczniami. Zachował przy tym notatnik, w którym zapisał wszystkie największe osiągnięcia. Po śmierci, jego zięć Annibale della Nove — sam matematyk i były uczeń del Ferro, odziedziczył notatki stryja, wraz z pozycją na uniwersytecie. Same notatki pozostały jednak w ukryciu aż do roku 1543, gdy della Nove odwiedzili dwóch bardzo znani matematycy: Gerolamo Cardano i Lodovico Ferrari, poszukujący metody del Ferro. Skąd o niej wiedzieli?

Kilka lat wcześniej w Wenecji głośno było o matematyku samouku Nicolo Targaglii, chętnie uczestniczącym w pojedynkach matematycznych. Do jednego z tych pojedynków stanął z Tartaglia niejako Fior — jeden z uczniów del Ferro, niezbyt pojętny, jak podają relacje. Każda ze stron podała drugiej 30 równań do rozwiązyania. Wszystkie dotyczyły równania wielomianowego stopnia 3. Co o nich wówczas wiedziano?

Matematycy wiedzieli wówczas, że rozwiązywanie ogólne równania trzeciego stopnia może być zredukowane do rozwiązywania jednego z dwóch typów równań:

$$x^3 + mx = n, \quad x^3 = mx + n, \quad \text{gdzie } m, n > 0.$$

Dlaczego dwa typy? Nie uznawano wówczas liczb ujemnych i przekształceń równoważnych z ich użyciem. Fior wiedział jak rozwiązywać tylko pierwszy z wymienionych wyżej typów. Zadania Fiora dotyczyły zatem jedynie tej klasy równań, podczas gdy zadania Tartaglia były bardzo różnorodne. W trakcie pojedynku, 13 lutego nad ranem, Tartaglia odkrył metodę rozwiązywania równań pierwszego typu i mecz wygrał, co dało mu wielką sławę i zainteresowanie słynnego lombardzkiego matematyka Cardano.

Cardano poprosił Tartaglię w 1539 roku o wyjawienie metody rozwiązywania tych równań i obiecał dochowania tajemnicy i nieujawniania metody. W 1540 roku asystent Cardano Lodovico Ferrari opracował metodę redukcji równań czwartego stopnia do równań sześciennych, co motywowało dodatkowo do złamania obietnicy. Jak to zrobić? Rozwiązaaniem okazała się wizyta u zięcia del Ferro — wspomnianego już della Nove w 1543 roku. Cardano uznał, że to właśnie del Ferro odkrył jako pierwszy metodę rozwiązywania równań stopnia 3 i poczuł się zwolniony z tajemnicy danej Cardano.

W 1545 roku Cardano opublikował *Artis Magnae, Sive de Regulis Algebraicis Liber Unus* (Księga Pierwsza o Wielkiej Sztuce, lub o Zasadach Algebry, stanowiące obok słynnego "De revolutionibus" Kopernika i "De human corporis fabrica" Vasaliusa, jedno z trzech najważniejszych traktatów naukowych wczesnego renesansu. Pierwsze wydania tych trzech dzieł miały miejsce w latach 1543-1545.

Wielkość dzieła Cardano oparta była na kilku czynnikach. Kojarzymy przede wszystkim pierwsze bezpośrednie wprowadzenie do języka matematyki pierwiastków z liczb ujemnych, zwanych później liczbami zespolonymi. Istotą rewolucji nie były wówczas jednak same liczby zespolone (których występowanie batatyzował sam autor, nie potrafiąc im przypisać żadnego fizycznego znaczenia), ponieważ liczby te wcale nie posłużyły Cardano do rozwiązywania równań stopnia 3.

Z uwagi na to, że ówcześnie nie stosowano w przekształceniach algebraicznych liczb ujemnych, Cardano zmuszony był rozpatrywać aż trzynaście rozmaitych klas równań stopnia 3. Rozwiązywanie żadnej z tych klas nie wymagało użycia liczb zespolonych. Liczby zespolone wspomniane są przy rozwiązywaniu klasycznego problemu poszukiwania liczb, których suma równa jest 10, a iloczyn równy jest 40. Cardano wprowadził również koncepcję pierwiastka wielokrotnego wielomianu, między innymi rozpatrując liczbę -2 jako dwukrotny pierwiastek równania $x^3 = 12x + 16$.

Jak rozwiązywano równanie $y^3 = py + q$? Dokonując kolejnej liniowej zamiany zmiennej postaci $y = u + v$, uzyskuje się po lewej stronie:

$$(u^3 + v^3) + 3uv(u + v) = 3uvy + (u^3 + v^3),$$

które to wyrażenie równe jest prawej stronie wcześniejszego równania wtedy i tylko wtedy, gdy:

$$\begin{aligned} 3uv &= p, \\ u^3 + v^3 &= q. \end{aligned}$$

Eliminując v uzyskujemy równanie kwadratowe zmiennej u^3 postaci

$$u^3 + \left(\frac{p}{3u}\right)^3 = q,$$

o rozwiązaniach

$$\frac{q}{2} \pm \sqrt{\left(\frac{q}{2}\right)^2 - \left(\frac{p}{3}\right)^3}.$$

Rozumując symetrycznie, uzyskujemy te same wartości v^3 . Skoro $u^3 + v^3 = q$, to jeden z pierwiastków równy jest u^3 , a drugi — v^3 . Bez straty ogólności można przyjąć

$$u^3 = \frac{q}{2} + \sqrt{\left(\frac{q}{2}\right)^2 - \left(\frac{p}{3}\right)^3}, \quad v^3 = \frac{q}{2} - \sqrt{\left(\frac{q}{2}\right)^2 - \left(\frac{p}{3}\right)^3},$$

uzyskując

$$y = u + v = \sqrt[3]{\frac{q}{2} + \sqrt{\left(\frac{q}{2}\right)^2 - \left(\frac{p}{3}\right)^3}} + \sqrt[3]{\frac{q}{2} - \sqrt{\left(\frac{q}{2}\right)^2 - \left(\frac{p}{3}\right)^3}}.$$

Jeżeli wyobrażmy sobie teraz, że Cardano dokonywał tych manipulacji w języku geometrycznym, rozumiemy jak bardzo konieczne było przejście od geometrii do języka wyrażeń algebraicznych, dokonane przez Vite'a ponad 100 lat później. Co jest jednak kluczowe? Uzyskane rozwiązanie wymaga użycia liczb zespolonych w przypadku, gdy $(q/2)^2 - (p/3)^3 < 0$. Nie jest możliwe pominięcie tego rozwiązania, gdyż wielomian stopnia 3 zawsze posiada choćby jeden pierwiastek rzeczywisty. Stąd formula Cardano stawia problem wyciągnięcia liczby rzeczywistej z wyrażenia postaci:

$$\sqrt[3]{a + b\sqrt{-1}} + \sqrt[3]{a - b\sqrt{-1}}.$$

Cardano nie zmierzył się z tym problemem w swojej *Ars Magna* z 1545 roku. Wspomniał wprawdzie o liczbach zespolonych w kontekście równania kwadratowego, ale uznał je za „tak subtelne, jak bezużyteczne”. Problemem poważnym poważnie zajął się dopiero Rafael Bombelli w 1572 roku, w podręczniku, którego popularność trwała aż do czasów Leibniza i Eulera. To jemu zawdzięczamy symbol $\sqrt{-1}$ oraz redukcję wyrażenia postaci $\sqrt[3]{a + b\sqrt{-1}}$ do postaci algebraicznej $c + d\sqrt{-1}$. I rzeczywiście, rozpatrując równanie

$$x^3 = 15x + 4$$

uzyskujemy z wzoru Cardano rozwiązanie

$$x = \sqrt[3]{2 + 11\sqrt{-1}} + \sqrt[3]{2 - 11\sqrt{-1}},$$

a czysto bezpośrednie sprawdzenie pozwala znalezienie pierwiastka $x = 4$. Bombelli przypuszczał, że dwa składniki rozwiązania x są postaci $2 + n\sqrt{-1}$ oraz $2 - n\sqrt{-1}$. Podnosząc tę równość do potęgi 3 i używając formalizmu $(\sqrt{-1})^2 = -1$, uzyskał

$$\sqrt[3]{2 + 11\sqrt{-1}} = 2 + \sqrt{-1}, \quad \sqrt[3]{2 - 11\sqrt{-1}} = 2 - \sqrt{-1}.$$

Termin *liczby urojone*, przypisywany liczbom zespolonym, pochodzi od twórcy geometrii analitycznej — Kartezjusza i ma już pierwsze konotacje geometryczne. Reprezentację geometryczną sugerował zresztą już John Wallis w połowie wieku XVII. Oznaczenie $i = \sqrt{-1}$ pochodzi od Leonharda Eulera. On również wykorzystywał postać trygonometryczną, a także przedstawiał pierwiastki równania $z^n = 1$ jako wierzchołki wielokąta foremnego. Jednocześnie, przywołując znów książkę prof. Kordosa, sam Euler w swojej *Algebrze* odnosił się z dystansem do liczb zespolonych, pisząc:

Pierwiastki kwadratowe z liczb ujemnych nie są zerem, ani nie są ujemne, ani dodatnie.
Stąd wynika, że pierwiastki te nie mogą się znajdować wśród możliwych liczb. W konsekwencji są to niemożliwe liczby. I tak dochodzimy do pojęcia liczb na ogół zwanych urojonymi lub wyobrażalnymi dlatego, że istnieją one tylko w wyobraźni.

Reprezentacja geometryczna liczb zespolonych pochodzi od Caspara Wessela (1797) i jest zbliżona do rachunku na wektorach. Pojęcie płaszczyzny zespolonej, zwanej inaczej płaszczyzną Arganda, związane jest z geometryczną interpretacją Jean-Roberta Arganda z 1806 roku. Algebraiczną definicję, którą posłużyliśmy się na wykładzie podał w 1831 roku Rowan Hamilton. Pojęcie *liczb zespolonych* wprowadził w tym samym roku Gauss, zdecydowanie oponując przeciwko terminowi *liczb urojonych*.

Teoria liczb zespolonych znalazła wcześnie ważne zastosowania praktyczne, między innymi w równaniach hydrodynamicznych, dzięki pracom d'Alemberta (1752). Równania ta zostały dogłębnie zrozumiane i opracowane przez Cauchy'ego oraz Riemanna w połowie XIX wieku, którzy to zajmowali się już funkcjami zespolonymi w perspektywie szeregow potęgowych i rachunku różniczkowego. O tych zagadnieniach dowiecie się Państwo najpierw na Analizie, a potem na dedykowanym przedmiocie — Funkcjach Analitycznych. Wraz z rozwojem fizyki, liczby zespolone zagościły m.in. w teorii fal elektromagnetycznych czy w mechanice kwantowej, między innymi w słynnym równaniu Schrödingera. A co dalej? Kształtowanie się rozumienia pojęcia liczby rozwijało się i przed, i obok, i po zaakceptowaniu liczb zespolonych, gdy budowano teorie kwaternionów, teorię liczb hiperzespolonych, liczb algebraicznych, porządną teorię liczb rzeczywistych i dwudzieste skomplikowane teorie, choćby leżące u podstaw analizy niestandardowej.

Rozdział 5

Wielomiany i ich pierwiastki. Ciała algebraicznie domknięte

5.1 Wykład 5

Na poprzednim wykładzie wprowadziliśmy pojęcie ciała i skupiliśmy się omówieniu dwóch istotnych przykładów: ciała reszt z dzielenia przez p oraz ciała liczb zespolonych. Celem tego wykładu jest przedstawienie kilku uwag dotyczących funkcji o wartościach w tych ciałach. Zagadnienie to jest w ogólności niezwykle szerokie, natomiast mając na względzie program kolejnych wykładów, ograniczymy się jedynie do tzw. funkcji wielomianowych. Zaczniemy od pojęcia wielomianu o współczynnikach w ciele.

Definicja 5.1.1: Wielomian

WIELOMIANEM zmiennej x o współczynnikach w ciele K nazywamy wyrażenie:

$$a_0 + a_1x + a_2x^2 + \dots + a_nx^n,$$

gdzie n jest nieujemną liczbą całkowitą oraz $a_0, a_1, \dots, a_n \in K$. Utożsamiamy przy tym takie napisy, jeśli różnią się o składniki postaci $0 \cdot x^i$ oraz jeśli różnią się kolejnością składników.

Elementy a_i nazywamy WSPÓŁCZYNNIKAMI wielomianu. Zbiór wielomianów o współczynnikach z ciała K oznaczamy przez $K[x]$. Jeśli wszystkie współczynniki wielomianu w są równe zero, to piszemy $w = 0$, a wielomian w nazywamy wówczas WIELOMIANEM ZEROWYM.

Innymi słowy, wielomiany zmiennej x o współczynnikach w ciele K utożsamiać można z ciągami nieskończonymi (a_i) , dla $i \in \mathbb{N}$, w których $a_i \neq 0$ tylko dla skończenie wielu i . Jest to definicja nieco inna niż ta znana ze szkoły, w której utożsamialiśmy wielomiany (jako wyrażenia algebraiczne) i funkcje wielomianowe. Wprowadziłmy od razu to rozróżnienie.

Definicja 5.1.2: Funkcja wielomianowa

Niech $F = a_0 + a_1x + a_2x^2 + \dots + a_nx^n \in K[x]$. Funkcję $f : K \rightarrow K$ daną wzorem

$$f(s) = a_0 + a_1s + a_2s^2 + \dots + a_ns^n$$

nazwiemy FUNKcją WIELOMIANOWĄ odpowiadającą wielomianowi F .

Zauważmy, że z punktu widzenia podejścia funkcyjnego, dwie funkcje wielomianowe f, g są równe, jeśli istnieją takie wielomiany

$$F = a_0 + a_1x + a_2x^2 + \dots + a_nx^n, \quad G = b_0 + b_1x + b_2x^2 + \dots + b_mx^m \in K[x],$$

że dla każdego $s \in K$ mamy

$$f(s) = a_0 + a_1s + a_2s^2 + \dots + a_ns^n = b_0 + b_1s + b_2s^2 + \dots + b_ms^m = g(s).$$

To kryterium równości funkcji wielomianowych nie daje podstawy do łatwego rozstrzygnięcia, czy wielomiany F, G są równe. Wymaga to uzasadnienia. Wielomian $x^2 + x \in \mathbb{Z}_2[x]$ jest niezerowy, ale dla każdego $s \in \mathbb{Z}_2$ wyrażenie $s^2 + s$ równe jest 0. A zatem funkcja wielomianowa $f : \mathbb{Z}_2 \rightarrow \mathbb{Z}_2$ zadana wzorem $f(x) = x^2 + x$ jest funkcją zerową. A zatem znając jedynie zbiór wartości funkcji wielomianowej $w(s)$ nie zawsze rozpoznamy wielomian $w \in K[x]$, przynajmniej gdy K jest nad ciałem skończonym.

Wniosek 5.1.3: Równość wielomianów

Jeśli

$$w = a_0 + a_1x + a_2x^2 + \dots + a_nx^n \quad \text{oraz} \quad v = b_0 + b_1x + b_2x^2 + \dots + b_mx^m,$$

dla pewnych nieujemnych liczb całkowitych n, m , to następujące warunki są równoważne:

- $w = v$ jako elementy $K[x]$.
- $m = n$ oraz $a_i = b_i$, dla każdego $1 \leq i \leq n$.

Powyższy wniosek łatwo przełożyć na język ciągów nieskończonych o skończenie wielu niezerowych wyrazach ze zbioru K . Dwa wielomiany, widziane jako ciągi, są równe, gdy są równe jako ciągi.

Definicja 5.1.4: Stopień wielomianu

Niech $f = a_0 + a_1x + a_2x^2 + \dots + a_nx^n \in K[x]$, gdzie K jest ciałem. STOPNIEM WIELOMIANU f , ozn. $\deg(f)$, nazywamy:

- największe takie i , że $a_i \neq 0$, o ile f nie jest wielomianem zerowym.
- $-\infty$, jeśli f jest wielomianem zerowym.

Przykłady. Mamy $\deg f = 4, \deg p = 7, \deg h = 2$, przy czym

$$f = 1 - 2x + 7x^3 + 5x^4 \in \mathbb{R}[x], \quad p = 2t^7 - \sqrt{2}t^3 - 99 \in \mathbb{R}[t], \quad h = 9i + (5-i)z + (2+7i)z^2 \in \mathbb{C}[z].$$

W zbiorze $K[x]$ określamy działania 2-argumentowe dodawania i mnożenia, pochodzące od działań w K .

Definicja 5.1.5: Suma i iloczyn wielomianów

Dla wielomianów

$$f = a_0 + a_1x + \dots + a_nx^n, \quad g = b_0 + b_1x + \dots + b_mx^m$$

ze zbioru $K[x]$ określamy:

- sumę $f + g$ wielomianów f, g daną wzorem

$$f + g = (a_0 + b_0) + (a_1 + b_1)x + (a_2 + b_2)x^2 + \dots$$

Innymi słowy współczynnik wielomianu $f + g$ stojący przy x^i równy jest $a_i + b_i$.

- iloczyn $f \cdot g$ wielomianów f, g dany wzorem

$$f \cdot g = a_0b_0 + (a_0b_1 + a_1b_0)x + (a_0b_2 + a_1b_1 + a_2b_0)x^2 + \dots + a_nb_mx^{n+m}.$$

Innymi słowy współczynnik wielomianu $f \cdot g$ stojący przy x^i równy jest $\sum_{j=0}^i a_j b_{i-j}$.

Jako ćwiczenie pozostawiamy następujące własności stopnia, związane z wprowadzonymi operacjami:

$$\deg(f + g) \leq \max(\deg(f), \deg(g)), \quad \deg(fg) = \deg(f) + \deg(g). \quad (\heartsuit)$$

Definicja 5.1.6: Pierwiastek wielomianu, równanie wielomianowe

PIERWIASTKAMI WIELOMIAŃU $f \in K[x]$ (inaczej: miejscami zerowymi) nazywamy takie $s \in K$, że funkcja wielomianowa odpowiadająca wielomianowi s przyjmuje w s wartość 0, tzn. $f(s) = 0$. Jeśli $f \in K[x]$ jest wielomianem stopnia n , to równanie $f = 0$ nazywamy RÓWNANIEM WIELOMIAŃOWYM STOPNIA n o współczynnikach w K .

Przykłady.

- Równanie $x^2 - 2 = 0$ nie ma rozwiązań w ciele \mathbb{Z}_3 , ponieważ kwadrat żadnej liczby całkowitej nie daje reszty 2 z dzielenia przez 3. A zatem w ciele \mathbb{Z}_3 nie ma pierwiastka stopnia 2 z liczby 3.
- Równanie $x^2 - 1 = 0$ ma dwa pierwiastki w ciele \mathbb{Z}_3 , czyli $x_1 = 1$ oraz $x_2 = 2$. Oznacza to, że pierwiastek kwadratowy nie jest funkcją jednoznaczna w ciele \mathbb{Z}_3 .

Twierdzenie 5.1.7: Bezout

Niech $f \in K[x]$. Następujące warunki są równoważne.

- (1) Element $s \in K$ jest pierwiastkiem wielomianu f .
- (2) Istnieje $g \in K[x]$ taki, że $f = (x - s)g$.

Dowód. Korzystamy z następującej obserwacji. Dla każdego $s \in K$ mamy

$$x^n - s^n = (x - s)(x^{n-1} + x^{n-2}s + \dots + xs^{n-2} + s^{n-1}).$$

W szczególności, jeśli $f \in K[x]$ jest dowolnym wielomianem stopnia $n > 0$ postaci $a_nx^n + \dots + a_2x^2 + a_1x + a_0$, to możemy dla każdego $1 \leq i \leq n$ zapisać

$$a_i x^i = a_i x^i - a_i s^i + a_i s^i = a_i(x - s)(x^{i-1} + x^{i-2}s + \dots + xs^{i-2} + s^{i-1}) + a_i s^i$$

Dodając powyższe równości stronami wnioskujemy, że istnieje wielomian $q \in K[x]$ stopnia $n - 1$ taki, że

$$f = (x - s)q + f(s).$$

Jeśli s jest pierwiastkiem wielomianu f , to mamy $f(s) = 0$ i rozkład powyżej przybiera postać $f = (x - s)q$.

Implikacja z (2) do (1) jest jasna. Jeśli $f = (x - s)g$, dla pewnego wielomianu $g = a_k x^k + \dots + a_2 x^2 + a_1 x + a_0 \in K[x]$, to po wymnożeniu $(x - s)g$ mamy

$$f = a_k x^{k+1} + \dots + a_1 x^2 + a_0 - a_k x^k s - \dots - a_1 x s - a_0 s.$$

Po podstawieniu do powyższej równości $x = s$ dostajemy $f(s) = 0$.

□

Wniosek 5.1.8

Wielomian stopnia n o współczynnikach w ciele K ma nie więcej niż n parami różnych pierwiastków.

Dowód powyższego wniosku jest łatwą indukcją ze względu na stopień, korzystającą z twierdzenia Bezout. Udowodnimy natomiast ogólniejszy rezultat, związany z pojęciem krotności pierwiastka.

Definicja 5.1.9: Krotność pierwiastka wielomianu

Dla dowolnego pierwiastka $s \in K$ wielomianu $f \in K[x]$ o współczynnikach w ciele K , największą liczbę naturalną k , taką że istnieje wielomian $g \in K[x]$, że $f = (x - s)^k \cdot g$ nazywamy *krotnością pierwiastka s* wielomianu f .

Wielokrotnie będziemy korzystali z następującego faktu.

Twierdzenie 5.1.10

Niech f będzie niezerowym wielomianem o współczynnikach w ciele K i niech s_1, \dots, s_m będą różnymi pierwiastkami wielomianu f o krotnościami odpowiednio k_1, \dots, k_m . Wówczas istnieje wielomian $g \in K[x]$, że

$$f = (x - s_1)^{k_1} (x - s_2)^{k_2} \cdots \cdot (x - s_m)^{k_m} \cdot g.$$

Suma krotności pierwiastków wielomianu f równa jest co najwyżej $\deg f$.

Dowód. Dowód przeprowadzimy przez indukcję ze względu na liczbę m różnych pierwiastków wielomianu f . Jeśli $m = 1$, to teza wynika bezpośrednio z definicji krotności.

Załóżmy, że teza jest prawdziwa dla pewnego $m > 1$. Założmy, że f ma parami różne pierwiastki s_1, \dots, s_m, s_{m+1} . Zgodnie z założeniem f można przedstawić w postaci

$$f = (x - s_2)^{k_2} \cdots \cdot (x - s_{m+1})^{k_{m+1}} \cdot g,$$

dla pewnego wielomianu $g \in K[x]$. Niech s_1 będzie k_1 -krotnym pierwiastkiem wielomianu f różnym od s_2, \dots, s_{m+1} . Zgodnie z definicją krotności wielomianu możemy zapisać $f = (x - s_1)^{k_1} \cdot h$, gdzie $h(s_1) \neq 0$. Możemy też przyjąć, że s_1 jest k -krotnym pierwiastkiem wielomianu g (uwzględniamy możliwość, że $k = 0$, jeśli s_1 nie jest pierwiastkiem g), czyli $g = (x - s_1)^k \cdot g'$, dla pewnego $g' \in K[x]$, spełniającego $g'(s_1) \neq 0$. Stąd:

$$f = (x - s_1)^{k_1} \cdot h = (x - s_2)^{k_2} \cdots \cdot (x - s_{m+1})^{k_{m+1}} \cdot (x - s_1)^k \cdot g'.$$

Nietrudno widzieć¹, że dla wielomianów $w_1, w_2 \in K[x]$ oraz elementu $s \in K$ mamy

$$(x - s) \cdot w_1 = (x - s) \cdot w_2 \implies w_1 = w_2.$$

Stąd, gdyby zachodziła nierówność $k_1 > k$, to (korzystając k -krotnie z powyższej obserwacji o możliwości skracania przez wspólny czynnik $(x - s)$) skracając przez $(x - s_1)^k$ otrzymalibyśmy

$$(x - s_1)^{k_1 - k} \cdot h = (x - s_2)^{k_2} \cdots \cdot (x - s_{m+1})^{k_{m+1}} \cdot g',$$

co nie może mieć miejsca, gdyż wartość funkcji wielomianowej wielomianu po lewej stronie jest dla $s = s_1$ jest równa 0, a wartość funkcji wielomianowej dla wielomianu po prawej jest różna od 0. Analogicznie uzasadniamy, że nie jest możliwa nierówność $k_1 < k$. Stąd otrzymujemy $k_1 = k$ i mamy

$$f = (x - s_2)^{k_2} \cdots \cdot (x - s_{m+1})^{k_{m+1}} \cdot (x - s_1)^{k_1} \cdot g'.$$

Dowód indukcyjny jest zakończony. Ostatnie zdanie tezy jest natomiast oczywiste. □

W uzupełnieniu do tego rozdziału zarysujemy ogólną teorię podzielności wielomianów o współczynnikach w ciele, z której powyższe twierdzenie będzie bezpośrednio wynikało.

Przejdziemy teraz do omówienia ważnej klasy ciał. Wielomiany stopnia $n > 0$ o współczynnikach w tych ciałach mają zawsze n pierwiastków.

Definicja 5.1.11: Ciało algebraicznie domknięte

Jeśli każdy wielomian stopnia większego od 0 o współczynnikach z ciała K ma w ciele K pierwiastek, to K nazywamy CIAŁEM ALGEBRAICZNIE DOMKNIĘTYM.

Oczywiście \mathbb{R} , ani tym bardziej \mathbb{Q} nie jest algebraicznie domknięte, ponieważ wielomian $x^2 + 1$ nie ma pierwiastków w tych ciałach. Z twierdzenia Bezout wynika także łatwo, że żadne ciało skończone nie jest algebraicznie domknięte.

¹Istotnie, jeśli dla pewnych wielomianów $w_1 = a_n x^n + \dots + a_1 x + a_0$ stopnia n oraz $w_2 = b_m x^m + \dots + b_1 x + b_0$ stopnia m mamy $(x - a)w_1 = (x - b)w_2$, to oczywiście $m = n$ i korzystając z tego, że wielomiany są równe wtedy i tylko wtedy, gdy ich współczynniki są równe stwierdzamy, że z równości $(x - a)(a_n x^n + \dots + a_1 x + a_0) = (x - a)(b_m x^m + \dots + b_1 x + b_0)$ wynika równość $a_i = b_i$, dla $1 \leq i \leq n$.

Twierdzenie 5.1.12

Niech K będzie ciałem. Następujące warunki są równoważne.

- (1) Ciało K jest algebraicznie domknięte.
- (2) Każdy wielomian stopnia > 0 o współczynnikach z K rozkłada się nad K na czynniki stopnia 1 (to znaczy: jest iloczynem wielomianów stopnia 1 o współczynnikach z K).

Dowód. Założymy, że K jest ciałem algebraicznie domkniętym. Przy pomocy dowodu indukcyjnego pokażemy, że każdy wielomian stopnia > 0 rozkłada się nad K na współczynniki liniowe. Krok bazowy indukcji jest jasny – każdy wielomian stopnia 1 da się rozłożyć na iloczyn czynników liniowych. Założymy prawdziwość naszego założenia dla wielomianów stopnia $n - 1$. Niech f będzie wielomianem stopnia n . Ciało K jest algebraicznie domknięte, więc f ma pierwiastek $c \in K$. Stąd

$$f(x) = (x - c) \cdot g(x),$$

dla pewnego wielomianu $g \in K[x]$, na mocy twierdzenia Bezout. Wielomian g jest zatem stopnia $n - 1$, więc z założenia indukcyjnego g jest iloczynem czynników stopnia 1 o współczynnikach z K . Stąd f jest iloczynem czynników stopnia 1 o współczynnikach z K .

Na odwrót: jeśli f jest iloczynem wielomianów stopnia 1 o współczynnikach w K , to każdy taki czynnik stopnia 1 ma pierwiastek w K , będący też pierwiastkiem wielomianu f . A zatem (2) implikuje (1). \square

Wskazywanie ciał algebraicznie domkniętych zwykle jest skomplikowane. Jedną z kluczowych motywacji rozważania ciała liczb zespolonych jest następujący wynik.

Twierdzenie 5.1.13: Zasadnicze Twierdzenie Algebry Gaussa, 1799

Ciało \mathbb{C} jest algebraicznie domknięte.

W tym momencie nie mamy narzędzi do przedstawienia dowodu tego twierdzenia. Stosunkowo elementarny dowód będą Państwo (teoretycznie) w stanie przeprowadzić po pierwszym semestrze zajęć z Analizy. Na wyższych latach studiów poznacie Państwo krótkie (m.in. algebraiczne) dowody tego rezultatu. Przedstawimy rezultat korzystający z ZTA dotyczący rozkładów wielomianów rzeczywistych na czynniki.

Uwaga 5.1.14

Niech w będzie wielomianem stopnia większego od 0 o współczynnikach rzeczywistych, traktowanych jako liczby zespolone. Wówczas jeśli liczba zespolona z jest pierwiastkiem wielomianu w , to również liczba \bar{z} , sprzężona do z , jest pierwiastkiem wielomianu w .

Dowód. Istotnie, niech s będzie pierwiastkiem wielomianu $w = a_n x^n + a_{n-1} x^{n-1} + \dots + a_1 x + a_0$, gdzie $a_n, \dots, a_0 \in \mathbb{R}$. Korzystamy z tego, że sprzężenie liczby rzeczywistej jest tą liczbą oraz z formuł

$$\overline{z_1 + z_2} = \overline{z_1} + \overline{z_2}, \quad \overline{z_1 z_2} = \overline{z_1} \cdot \overline{z_2},$$

prawdziwych dla dowolnych $z_1, z_2 \in \mathbb{C}$. Mamy:

$$0 = \overline{0} = \overline{a_n s^n + \dots + a_1 s + a_0} = \overline{a_n} \overline{s^n} + \dots + \overline{a_1} \overline{s} + \overline{a_0} = a_n (\bar{s})^n + \dots + a_1 \bar{s} + a_0.$$

\square

Dla przykładu, można sprawdzić, że wielomian $z^2 - 2z + 5$ o współczynnikach rzeczywistych ma pierwiastek zespolony $1 + 2i$. Jak się okazuje, pierwiastkiem tego wielomianu jest również $1 - 2i$ i mamy:

$$z^2 - 2z + 5 = (z - 1 - 2i)(z - 1 + 2i).$$

Z drugiej strony wielomian z $\mathbb{C}[x]$, którego współczynniki nie są zespolone, nie musi (i nie spełnia) właściwości wyżej. Wielomian $x - i$ ma pierwiastek w \mathbb{C} równy i , natomiast $-i$ nie jest jego pierwiastkiem.

Wniosek 5.1.15

Każdy wielomian w stopnia większego od 0 o współczynnikach rzeczywistych rozkłada się na iloczyn wielomianów stopnia pierwszego i stopnia drugiego o współczynnikach rzeczywistych.

Dowód. Jeśli wielomian w jest stopnia pierwszego lub drugiego, to nie ma czego dowodzić. Dalsze rozumowanie jest indukcją ze względu na stopień n wielomianu. Założymy, że $\deg(w) > 2$. Jeśli $r_0 \in \mathbb{R}$ jest pierwiastkiem w , to z lematu Bezout

$$w = (x - r_0)g,$$

gdzie $g \in \mathbb{R}[x]$ i teza wynika z założenia indukcyjnego zastosowanego do wielomianu g . Jeśli w nie ma pierwiastków rzeczywistych, to postępowanie jest następujące. Bierzemy pierwiastek $z_0 \in \mathbb{C} \setminus \mathbb{R}$ wielomianu w , który musi istnieć na mocy ZTA. Wówczas na mocy poprzedniej obserwacji $\overline{z_0}$ też jest także pierwiastkiem w . Skoro z_0 i $\overline{z_0}$ to różne pierwiastki w dostajemy, że iloczyn

$$(x - z_0)(x - \overline{z_0}) \in \mathbb{R}[x]$$

jest dzielniakiem stopnia 2 wielomianu w . \square

Powyższe twierdzenie nie ma zastosowania do ciała \mathbb{Q} . Wielomian $x^4 + 2 \in \mathbb{Q}[x]$ nie ma żadnego rozkładu na czynniki niższego stopnia niż 4 w $\mathbb{Q}[x]$. Warto w kontekście ciała \mathbb{Q} pamiętać szkolne twierdzenie o wymiernych pierwiastkach wielomianu (i umieć je udowodnić, co jest raczej nietrudnym ćwiczeniem). Problem rozkładalności wielomianu na czynniki (nierożkładowalne) jest ważnym zagadnieniem, nie tylko na naszym przedmiocie czy w różnych działach matematyki, ale także w jej (praktycznych) zastosowaniach.

Przykład. Wyznaczmy wszystkie pierwiastki zespolone wielomianu

$$w = x^4 - 3x^3 + 6x^2 + 2x - 60 \in \mathbb{R}[x],$$

jeśli wiadomo, że jednym z tych pierwiastków jest $1 + 3i$.

Rozwiążanie. Skoro wielomian w ma współczynniki rzeczywiste, to na mocy Uwagi 5.1.12 kolejnym pierwiastkiem w , obok $1 + 3i$, jest $\overline{1 + 3i} = 1 - 3i$. Na mocy twierdzenia Bezout istnieje więc wielomian g stopnia 2 taki, że:

$$w = (x - 1 - 3i)(x - 1 + 3i) \cdot g = (x^2 - 2x + 10) \cdot (x^2 - x - 6).$$

Iloraz g wielomianów $x^4 - 3x^3 + 6x^2 + 2x - 60$ oraz $x^2 - 2x + 10$ uzyskaliśmy poprzez algorytmu dzielenia wielomianów znanego ze szkoły. Alternatywnie, na mocy warunku $w = (x^2 - 2x + 10) \cdot g$ można wyznaczyć takie a, b, c , porównując współczynniki wielomianów $w = (x^2 - 2x + 10) \cdot (ax^2 + bx + c)$.

Pozostaje zatem sprawdzić, że $x^2 - x - 6 = (x - 3)(x + 2)$, co oznacza, że pierwiastkami wielomianu w są liczby $-2, 3, 1 + 3i, 1 - 3i$. Wielomian w ma też następujący rozkład na iloczyn czynników liniowych i kwadratowych należących do $\mathbb{R}[x]$ (opisany we Wniosku 5.1.13) $w = (x - 3)(x + 2)(x^2 - 2x + 10)$. ■

Oto inny przykład.

Zadanie. Wielomian $w(x) = x^4 + ax^3 + bx^2 + cx + d$ ma współczynniki rzeczywiste oraz pierwiastki nierzeczywiste z_1, z_2, z_3, z_4 . Wiadomo, że $z_1 z_2 = 13 + i$ oraz $z_3 + z_4 = 3 + 4i$. Wyznacz $z_1 + z_2$ oraz $z_3 z_4$

Rozwiążanie. Wielomian w jest stopnia 4, a zatem z_1, z_2, z_3, z_4 są wszystkimi jego pierwiastkami.

Skoro $z_1 \notin \mathbb{R}$, to jedna z liczb z_2, z_3, z_4 musi być sprzężonym do z_1 pierwiastkiem w . Jednak dla każdego $z \in \mathbb{C}$ mamy $z\bar{z} = |z|^2 \in \mathbb{R}$, a zatem $\overline{z_1} \neq z_2$, bo $z_1 z_2 \notin \mathbb{R}$. Analogicznie $\overline{z_3} \neq z_4$. Stąd $\{\overline{z_1}, \overline{z_2}\} = \{z_3, z_4\}$ (nieco dokładniej: po podzieleniu w przez $(z - z_1)(z - \overline{z_1})$ mamy wielomian, którego pierwiastkami są $z_2, \overline{z_2}$). A zatem mamy:

$$z_3 z_4 = \overline{z_1 z_2} = 13 - i \quad \text{oraz} \quad z_1 + z_2 = \overline{z_3 + z_4} = 3 - 4i.$$

■

5.2 Wybór przykładowych pytań

Proszę się nie uczyć na pamięć odpowiedzi czy wyjaśnienie! Takich pytań można ułożyć znacznie więcej i służą one jedynie sprawdzeniu podstawowego zrozumienia treści zajęć. Jeśli odpowiedź na dowolne z poniższych pytań sprawia Państwu problem, proszę pomyśleć który fragment notatek dotyczy tego pytania, wrócić do ich lektury, a w razie dalszych trudności poprosić prowadzącego o wskazówkę.

1. Czy suma wielomianów niezerowego stopnia może być wielomianem stopnia 0?
2. Wielomiany $f, g \in K[x]$ są stopnia n . Czy wielomian $f + g$ jest stopnia n ? Czy wielomian $f \cdot g$ jest stopnia $2n$?
3. Czy wielomian $x^2 - 1$ ma dwa różne pierwiastki w ciele \mathbb{Z}_2 ?
4. Czy wielomian $x^2 - 1$ ma dwa różne pierwiastki w ciele \mathbb{Z}_p dla $p > 2$?
5. Czy wielomian $x^2 - 2$ ma pierwiastek w ciele \mathbb{Z}_5 ?
6. Czy wielomian $x^3 - 2$ ma pierwiastek w ciele \mathbb{Z}_5 ?
7. Czy wielomian $x^{100} - 2$ ma pierwiastek w ciele \mathbb{Z}_5 ?
8. Czy wielomian $x^2 + x + 1$ ma pierwiastek w ciele \mathbb{Z}_2 ?
9. Rozłóż na czynniki stopnia 1 wielomian $x^2 + x + 1$ o współczynnikach odpowiednio w ciele \mathbb{Z}_3 , \mathbb{Z}_7 , \mathbb{Z}_{13} .
10. Czy istnieje wielomian w stopnia 3 o współczynnikach w ciele \mathbb{Z}_5 taki, że $w(a) = 0$ dla wszystkich $a \in \mathbb{Z}_5$?
11. Wyznacz wszystkie wielomiany w o współczynnikach rzeczywistych, które dla każdego $x \in \mathbb{R}$ spełniają warunek $w(x+3) = w(x)$. Co można powiedzieć o funkcji $f(x) = w(x+3) - w(x)$?
12. Niech w będzie wielomianem o współczynnikach w ciele \mathbb{Z}_2 . Założymy, że $w(0) = w(1) = 0$. Czy w jest wielomianem zerowym lub wielomianem $x^2 + x$?
13. Jedynym pierwiastkiem wielomianu o współczynnikach w \mathbb{C} jest liczba i . Czy wielomian ten może być stopnia 2?
14. Wszystkie pierwiastki wielomianu są liczbami należącymi do $\mathbb{C} \setminus \mathbb{R}$. Czy wielomian ten może mieć współczynniki rzeczywiste?
15. Wiadomo, że $z_1^2 = z_2^2 = z_3^2$, dla pewnych $z_1, z_2, z_3 \in \mathbb{C}$. Czy liczby z_1, z_2, z_3 mogą być parami różne?
16. Rozważmy $w \in \mathbb{C}[x]$, przy czym $w(z) = 0$ dla pewnej liczby zespolonej z . Czy $w(\bar{z}) = 0$?
17. Założymy, że $1+2i$ oraz $-1+2i$ są pierwiastkami wielomianu o współczynnikach rzeczywistych. Czy stopień tego wielomianu może być równy 3?
18. Liczba zespolona $1+2i$ jest pierwiastkiem wielomianu stopnia 8 o współczynnikach rzeczywistych. Czy pozostałe pierwiastki tego wielomianu mogą być rzeczywiste?
19. Wielomian $w \in \mathbb{R}[x]$ spełnia $w(i) = 0$. Czy istnieje taki wielomian $f \in \mathbb{R}[x]$, że

$$w = (x^2 + 1) \cdot f$$

20. Liczby zespolone z, w spełniają $z \cdot w \in \mathbb{R}$. Czy wielomian

$$f = (x - z)(x - w)$$

musi mieć współczynniki rzeczywiste?

21. Wielomian o współczynnikach w pewnym ciele K jest postaci $w = x^2 + ax + b$ oraz

$$w(x_1) = w(x_2) = 0.$$

Czy $x_1 + x_2 = -a$? Czy $x_1 \cdot x_2 = b$?

5.3 Zadania do samodzielnej pracy

Zadania ilustrujące umiejętności niezbędne do zaliczenia przedmiotu oznaczone są symbolem ♠.

Przy tych zadaniach dodany jest opis umiejętności, które sprawdzają.

1. (♠ Wyznaczanie pierwiastków stopnia n , wskazywanie pierwiastków pierwotnych z jedynki)
Znajdź wszystkie pierwiastki, wyznacz ich postaci ogólne (nie tylko trygonometryczne)

- (a) stopnia 4 z liczby -4 ,
- (b) stopnia 4 z liczby $-\sqrt{3} + 3i$,
- (c) stopnia 6 z liczby $-27i$,
- (d) stopnia 3 z liczby $5 + 5i$.

Znajdź liczbę pierwiastków pierwotnych stopnia 12 z jedynki i wypisz ich postaci trygonometryczne.

2. Znajdź sumę oraz iloczyn wszystkich pierwiastków stopnia n z jedynki oraz iloczyn wszystkich pierwiastków pierwotnych stopnia n z 1.

3. Niech $\{1, z_1, \dots, z_{2022}\}$ będzie zbiorem zespolonych pierwiastków stopnia 2023 z jedynki. Wykaż, że

$$(1 - z_1)(1 - z_2)(1 - z_3) \dots (1 - z_{2022}) = 2023.$$

Jaka jest interpretacja geometryczna tej równości?

4. Rozważmy zbiór $\mathcal{A} = \{\varepsilon_0, \dots, \varepsilon_{11}\}$ pierwiastków zespolonych stopnia 12 z 1. Uzasadnij, że

$$(\sqrt{3} + i - \varepsilon_0) \cdot (\sqrt{3} + i - \varepsilon_1) \cdot \dots \cdot (\sqrt{3} + i - \varepsilon_{11}) = 2^{12} - 1.$$

5. (♠ Rozkład wielomianu o współczynnikach rzeczywistych na czynniki stopnia ≤ 2)

Dla każdego z poniższych wielomianów znajdź jego rozkład na czynniki będące wielomianami rzeczywistymi stopnia ≤ 2 .

- (a) $x^4 - 2x^2 + 4$,
- (b) $x^6 + x^2$,
- (c) $x^7 - x$,
- (d) $x^4 + 4$,
- (e) $x^7 + 8x^4 + 4x^3 + 32$,
- (f) $x^6 + 27$,
- (g) $x^6 - x^3 + 1$.

6. Znajdź rozkład wielomianu rzeczywistego $x^4 - 4x^3 + 2x^2 + 4x + 4$ na czynniki rzeczywiste stopnia 2.
Wskazówka: skorzystaj z podstawienia $x = t + 1$.

7. (♠ Znajdowanie pierwiastków równań wielomianowych o współczynnikach w ciele \mathbb{Z}_p)

Znajdź pierwiastki wielomianów

- $x^5 - x \in \mathbb{Z}_5[x]$,
- $x^3 - 1 \in \mathbb{Z}_7[x]$,
- $x^3 - 2 \in \mathbb{Z}_7[x]$.

8. Znajdź wszystkie wspólne pierwiastki wielomianów $10x^{15} + 9x^2 + 1$ oraz $10x^{15} + 8x^2 + 2$ o wspólnikach w ciele \mathbb{Z}_{19} .

9. (♠ Wyznaczanie pierwiastków zespolonych prostych równań wielomianowych)

Rozłóż na czynniki liniowe wielomiany o współczynnikach zespolonych

- (a) $z^2 + 5 - 12i$,
- (b) $(1 + i)z^3 + (3 - 5i)z^2 - 6z$,
- (c) $z^4 + 2z^2 + 2$.

10. (♠ Sprzężenie pierwiastka nierzeczywistego wielomianu z $\mathbb{R}[x]$ jest też pierwiastkiem)

Wyznacz wszystkie zespolone pierwiastki wielomianu $z^4 - 6z^3 + 18z^2 - 30z + 25$, wiedząc że jednym z nich jest $2 - i$.

11. Liczba $\cos \phi + i \cdot \sin \phi$ spełnia równanie $z^n + a_1 z^{n-1} + \dots + a_n = 0$, gdzie $a_1, a_2, \dots, a_n \in \mathbb{R}$. Wykaż, że zachodzi równość $a_1 \sin \phi + a_2 \sin 2\phi + \dots + a_n \sin n\phi = 0$.

5.4 Uzupełnienie. Elementy teorii podzielności wielomianów

W tym podrozdziale omawiamy podstawowe fakty związane z podzielnością wielomianów o współczynnikach w ciele. Rezultaty te przypominają wyniki dotyczące liczb całkowitych. Wspólną teorię tych obiektów daje dział algebry zwany teorią pierścieni.

Uzasadnimy najpierw twierdzenie o dzieleniu z resztą wielomianów o współczynnikach w ciele.

Twierdzenie 5.4.1: O dzieleniu z resztą

Niech f, g będą wielomianami o współczynnikach z ciała K . Założymy ponadto, że g nie jest wielomianem zerowym. Wówczas istnieją wielomiany q i r takie, że:

$$g = q \cdot g + r, \quad \deg(r) < \deg(g).$$

Ponadto wielomiany q, r są wyznaczone jednoznacznie.

Dowód. Niech $f, g \in K[x]$, $g \neq 0$. Zauważmy, że jeśli $\deg(g) > \deg(f)$, to za szukane wielomiany q, r można wziąć $q = 0$ oraz $r = f$. Wówczas oczywiście $\deg(r) = \deg(f) < \deg(g)$. Założymy dalej, że $\deg(f) \geq \deg(g)$. Dowód istnienia wielomianów q, r spełniających równość $g = q \cdot g + r$ jest indukcją ze względu na $\deg(f)$. Z założenia $\deg(f) \geq 0$, a zatem w bazowym kroku indukcji rozważamy sytuację, gdy $\deg(f) = 0$. Skoro $g \neq 0$, to $\deg(g) = 0$ i wystarczy wziąć $q = f/g$ oraz $r = 0$. Wtedy $\deg(r) < \deg(g)$. Przechodzimy wreszcie do kroku indukcyjnego. Niech $n = \deg(f)$ oraz $m = \deg(g) \leq n$. Niech:

$$f = a_0 + a_1x + \dots + a_nx^n, \quad g = b_0 + b_1x + \dots + b_mx^m.$$

Rozważmy wielomian

$$\tilde{f} = b_m \cdot f - a_nx^{n-m} \cdot g.$$

Nietrudno widzieć, że odejmujemy od siebie dwa wielomiany stopnia n o tym samym współczynniku wiodącym. Wielomian \tilde{f} ma zerowy współczynnik przy x^n , a zatem $\deg(\tilde{f}) \leq n-1$. Z założenia indukcyjnego zastosowanego do \tilde{f} wynika, że istnieją wielomiany \tilde{q} oraz \tilde{r} takie, że $\tilde{f} = \tilde{q}\tilde{r}$, $\deg(\tilde{q}) > \deg(\tilde{r})$ oraz:

$$b_m f - a_n x^{n-m} g = g\tilde{q} + \tilde{r}.$$

W szczególności mamy też:

$$f = g \left(\frac{a_n x^{n-m} + \tilde{q}}{b_m} \right) + \frac{\tilde{r}}{b_m}.$$

A zatem dla pary wielomianów f, g definiujemy szukane wielomiany q, r jako

$$q := \frac{a_n x^{n-m} + \tilde{q}}{b_m} \quad \text{oraz} \quad r := \frac{\tilde{r}}{b_m}.$$

Oczywiście spełnione jest założenie $\deg(g) > \deg(r) = \deg(\tilde{r})$. Krok indukcyjny jest zatem zakończony.

Pozostaje pokazać jednoznaczność istnienia wielomianów q, r spełniających $g = q \cdot g + r$ dla danej pary wielomianów f, g . Będzie to (jak zwykle w takich problemach) rozumowanie nie wprost. Założymy, że dla pewnej pary f, g wielomianów istnieją wielomiany q, q', r, r' takie, że $\deg(g) > \deg(r)$, $\deg(g) > \deg(r')$ oraz

$$f = qg + r = q'g + r'.$$

Wynika stąd, że:

$$(q - q')g = r' - r.$$

Przypuśćmy, nie wprost, że $q \neq q'$. Mamy zatem $\deg(q - q')g \geq \deg(g)$. W rezultacie $\deg(r - r') \geq \deg(g)$. Mamy jednak

$$\deg(r' - r) \leq \max\{\deg(r'), \deg(r)\} < \deg(g)$$

(założenie o stopniach r, r' i g). Otrzymaliśmy sprzeczność. A zatem musi zachodzić równość $q = q'$. Wtedy jednak zachodzi także równość $r = r'$. Dowód jednoznaczności przedstawienia $g = q \cdot g + r$ jest zatem zakończony. \square

Definicja 5.4.2

Wielomian $f = a_n x^n + \dots + a_1 x + a_0 \in K[x]$ stopnia n UNORMOWANYM, jeśli $a_n = 1$. Powiemy, że niezerowy wielomian $g \in K[x]$ dzieli f lub jest DZIELNIKIEM f , jeśli istnieje $h \in K[x]$ taki, że $f = g \cdot h$, co oznaczamy $g | f$. Jeśli dla układu wielomianów $p_1, \dots, p_n \in K[x]$ nie istnieje wielomian f stopnia > 0 , że dla każdego i mamy $f | p_i$, to wielomiany p_i nazywamy WZGLEDNIE PIERWSZYMI.

Jeśli $\deg(f) \geq 1$, to wielomian f nazwiemy NIEROZKŁADALNYM, jeśli nie istnieją (niestale) wielomiany $g, h \in K[x]$, $\deg(f), \deg(h) \geq 1$ takie, że $f = g \cdot h$. Wielomian stopnia ≥ 1 , który nie jest nieroziadalny nazywamy ROZKŁADALNYM.

Rozkładalność wielomianu zależy oczywiście od ciała współczynników. Na wykładzie pokazaliśmy, że każdy wielomian $f \in \mathbb{R}[x]$ stopnia co najmniej 3 jest rozkładalny. Tymczasem wielomian $x^3 + 2$ nie jest rozkładalny jako element $\mathbb{Q}[x]$. Po odpowiednim rozszerzeniu współczynników, na przykład w ciele $\mathbb{Q}(\sqrt[3]{2})$, wielomian ten można już jednak rozłożyć na czynniki stopnia ≥ 2 postaci: $x - \sqrt[3]{2}$ oraz $x^2 + \sqrt[3]{2}x + \sqrt[3]{2}x^2$.

Twierdzenie 5.4.3: Lemat Bezout

Dla dowolnych niezerowych względnie pierwszych wielomianów $p_1, \dots, p_n \in K[x]$ o współczynnikach w ciele K istnieją wielomiany $q_1, \dots, q_n \in K[x]$ takie, że

$$q_1 p_1 + \dots + q_n p_n = 1.$$

Dowód. Niech

$$I = \{q_1 p_1 + \dots + q_n p_n \mid q_i \in K[x]\} \subseteq K[x].$$

Niech d będzie wielomianem unormowanym najmniejszego możliwego stopnia należącym do I . Pokażemy, że $\deg d = 0$. Zauważmy, że gdyby któryś z wielomianów p_i nie był podzielny przez d , to korzystając z twierdzenia o dzieleniu z resztą mielibyśmy $p_i = h_i d + r_i$, przy czym $\deg(r_i) < \deg(d)$. Ale skoro $d \in I$, to dla pewnych $q'_1, \dots, q'_n \in K[x]$ mielibyśmy

$$r_i = p_i - h_i d = p_i - (q'_1 p_1 + \dots + q'_n p_n),$$

zatem $r_i \in I$, co jest niemożliwe. Zatem d dzieli wszystkie p_i . Skoro jednak wielomiany te są względnie pierwsze, to $\deg d = 1$. \square

Twierdzenie 5.4.4: O jednoznaczności rozkładu wielomianów

Niech p będzie wielomianem stopnia ≥ 1 w $K[x]$, gdzie K – ciało. Wówczas p można zapisać w postaci:

$$p = a \cdot q_1 \cdot \dots \cdot q_k, \quad (\diamond)$$

gdzie a jest współczynnikiem wiodącym p oraz q_1, \dots, q_k są unormowanymi nieroziadalnymi wielomianami w $K[x]$. Co więcej, rozkład ów jest jednoznaczny z dokładnością do porządku występowania czynników.

Dowód. Dowód ma dwie części. Pierwsza to uzasadnienie istnienia rozkładu (\diamond) , a druga to dowód jego jednoznaczności.

Istnienie rozkładu (\diamond) uzasadniamy przez indukcję ze względu na $\deg p$. Jeśli p jest nieroziadalny, a w szczególności, jeśli p jest stopnia 1, to $p = a \cdot q$, gdzie a jest wiodącym współczynnikiem p i jest jasne, że q jest unormowany i nieroziadalny.

Możemy zatem przejść do kroku indukcyjnego i jednocześnie założyć, że p jest rozkładalny. W takim przypadku $p = p_1 p_2$, dla pewnych $p_1, p_2 \in K[x]$, przy czym $\deg(p) > \deg(p_i) \geq 1$ oraz z założenia indukcyjnego:

$$p_1 = a_1 \cdot q_1 \dots q_l, \quad p_2 = a_2 \cdot q_{l+1} \dots q_k,$$

gdzie q_i są unormowane i nieroziadalne oraz a_i są współczynnikami wiodącymi w p_i . W szczególności $p = a_1 a_2 \cdot q_1 \dots q_l \cdot q_{l+1} \dots q_k$, jest rozkładem typu (\diamond) .

Aby udowodnić jednoznaczność, wykażemy najpierw pewną obserwację: jeśli wielomian $f \in K[x]$ jest nieroziadalny oraz $f | gh$, gdzie $g, h \in K[x]$, to $f | g$, lub $f | h$. Jeśli bowiem f nie dzieli g , to wobec nieroziadalności f wielomiany f, g są względnie pierwsze. Z Lematu Bezout istnieją więc takie $a, b \in K[x]$, że $af + bg = 1$. Stąd $h = afh + bgh$, czyli $f | h$.

Przejdzmy do dowodu jednoznaczności rozkładu (\diamond). Niech:

$$p = a \cdot q_1 \dots q_k = a \cdot q'_1 \dots q'_r$$

przedstawia dwa rozkłady p na czynniki nieroziadalne. Na mocy uwagi wyżej, skoro q_1 dzieli $q'_1 \dots q'_r$, to q_1 dzieli jeden z czynników q'_i . Skoro obydwa te wielomiany są unormowane i nieroziadalne, to $q_1 = q'_i$. Skracamy te dwa czynniki i powtarzamy ten proces, aż wszystkie q_1, \dots, q_k zostaną skrócone. \square

Z powyższego rezultatu wynika oczywiście Twierdzenie 2.3.6 mówiące między innymi, że suma krotności pierwiastków wielomianu o współczynnikach w ciele jest nie większa niż jego stopień, gdyż wielomiany $x - a$ oraz $x - b$ są względnie pierwsze, dla $a \neq b$. Twierdzenie 2.3.6 jest prawdziwe także wtedy, gdy współczynniki wielomianu należą do dziedziny całkowitości.

Jeśli współczynniki wielomianu nie są w ciele, a nawet – nie są w dziedzinie całkowitości, wówczas ani Twierdzenie 2.3.6, ani Twierdzenie 2.8.4 nie musi być prawdziwe. Stosownego przykładu dostarcza wielomian stopnia 2 o współczynnikach w pierścieniu \mathbb{Z}_6 , który ma dwa rozkłady na czynniki postaci $(x - 1)(x - 2) = (x - 4)(x - 5)$.

* * *

W kontekście wiedzy szkolnej warto przypomnieć jeszcze jeden fakt w ogólnym jego sformułowaniu.

Twierdzenie 5.4.5: Wzory Viete'a

Niech K będzie ciałem oraz niech $x_1, x_2, x_3, \dots, x_n$ będą pierwiastkami wielomianu

$$f = a_n x^n + a_{n-1} x^{n-1} + \dots + a_1 x + a_0 \in K[x],$$

gdzie $a_n \neq 0$ (tzn. $\deg f = n > 0$). Wówczas zachodzą równości^a:

$$\begin{cases} \sum_{i=1}^n x_i &= x_1 + x_2 + \dots + x_n &= \frac{-a_{n-1}}{a_n} \\ \sum_{1 \leq i < j \leq n} x_i x_j &= x_1 x_2 + x_1 x_3 + \dots + x_{n-1} x_n &= \frac{a_{n-2}}{a_n} \\ \sum_{1 \leq i < j < k \leq n} x_i x_j x_k &= x_1 x_2 x_3 + x_1 x_2 x_4 + \dots + x_{n-2} x_{n-1} x_n &= \frac{-a_{n-3}}{a_n} \\ && \vdots \\ x_1 x_2 x_3 \dots x_n && = (-1)^n \frac{a_0}{a_n}. \end{cases}$$

^aNie wypisujemy dokładnie równości dotyczących współczynników stojących przy kolejnych potęgach. Po lewych stronach stoją tzw. elementarne wielomiany symetryczne stopnia n od pierwiastków wielomianu.

Dowód. Indukcja ze względu na n . Dla $n = 1$ dowód jest oczywisty. Założymy, że dla każdego wielomianu stopnia n , wzory te są prawdziwe. Rozważmy wielomian stopnia $n+1$, o pierwiastkach: $x_1, x_2, \dots, x_n, x_{n+1}$. Zgodnie z twierdzeniem Bezout istnieje wielomian $g(x)$ (którego wiodący współczynnik to 1), taki, że:

$$f(x) = a_{n+1} \cdot (x - x_1) \cdot g(x).$$

Wielomian g jest stopnia n , i jego pierwiastkami są $x_2, x_3, \dots, x_n, x_{n+1}$. Więcej pierwiastków, zgodnie z udowodnionym wcześniej faktem, mieć nie może. Zatem są to jego wszystkie pierwiastki. Z założenia indukcyjnego mamy zatem:

$$g(x) = x^n - (x_2 + x_3 + \dots + x_{n+1})x^{n-1} + \dots + (-1)^{n+1}(x_2 x_3 \dots x_{n+1}).$$

Wymnażając g w takiej postaci przez $a_{n+1} \cdot (x - x_1)$ dostajemy tezę. \square

5.5 Dodatek. Rozkładanie na czynniki i jego jednoznaczność

Poświęcimy teraz trochę miejsca wielomianom o współczynnikach całkowitych². Przypomnijmy najpierw szkolny rezultat.

Twierdzenie 5.5.1: O wymiernych pierwiastkach wielomianu o współczynnikach w \mathbb{Z}

Załóżmy, że liczby a_0, a_1, \dots, a_n są całkowite, $a_n \neq 0$, $n \neq 1$ oraz

$$a_0 + a_1x + a_2x^2 + \dots + a_nx^n = 0$$

dla pewnej liczby x . Jeśli $x = \frac{p}{q}$ i liczby całkowite p, q są względnie pierwsze, to

$$p | a_0 \quad \text{oraz} \quad q | a_n.$$

Dowód. Pomóżmy równość

$$a_0 + a_1\frac{p}{q} + a_2\left(\frac{p}{q}\right)^2 + \dots + a_n\left(\frac{p}{q}\right)^n = 0$$

przez q^n otrzymując:

$$a_0q^n + a_pq^{n-1} + a_2p^2q^{n-1} + \dots + a_{n-1}p^{n-1}q + a_np^n = 0.$$

Zauważmy, że:

$$q | a_np^n = -q(a_0q^{n-1} + a_1pq^{n-2} + a_2pq^{n-3} + \dots + a_{n-1}p^{n-1}),$$

a ponieważ liczby p, q są względnie pierwsze, to p^n, q też są względnie pierwsze, a z tego wynika, że q jest dzielnikiem a_n . W ten sam sposób widzimy, że $p | a_0p^n$, co oznacza, że p jest dzielnikiem a_0 . \square

Powysze twierdzenie mówi między innymi o całkowitych pierwiastkach wielomianów o współczynnikach całkowitych. Może ono służyć do wyznaczania rozkładów wielomianów (jak w szkole). Warto jednak przyjrzeć się tej własności w nieco ogólniejszym kontekście, rozszerzając nieco pojęcie liczby całkowitej.

Rozważmy liczby zespolone $z_1 = a + bi$, $z_2 = c + di$ oraz $z_3 = e + fi$, gdzie $a, b, c, d, e, f \in \mathbb{Z}$ i założymy, że:

$$(a + bi) = (c + di)(e + fi).$$

Oczywiście mamy stąd, że $|z_1| = |z_2| \cdot |z_3|$, a więc $a^2 + b^2$ jest wielokrotnością $c^2 + d^2$ oraz $e^2 + f^2$. A zatem jeśli rozważymy liczby zespolone, których części: rzeczywista i urojona są liczbami całkowitymi, wówczas zachodzić się zdają pewne związki pomiędzy rozkładem tych liczb na czynniki, a rozkładem na czynniki kwadratów ich modułów – czyli zwykłych liczb całkowitych. Ta prosta obserwacja ma, jak się okazuje, niezwykle daleko idące konsekwencje. Zaczniemy od banalnego zastosowania. Rozwiążmy zadanie.

Zadanie. Niech a, b oraz n będą liczbami naturalnymi. Udowodnić, że istnieją liczby całkowite x, y , dla których zachodzi równość

$$(a^2 + b^2)^n = x^2 + y^2.$$

Jest to zadanie egzaminacyjne (na ocenę celującą) z książki *Algebra liniowa 1 Kolokwia i egzaminy*, autorstwa M. Gewerta i Z. Skoczyłasa. Nie jestem pewien czy Czytelnik od razu wskazałby rozwiązanie bez użycia liczb zespolonych. Tymczasem stosując argumentację podaną wyżej naszą równość przepisujemy do postaci:

$$(a + bi)^n(a - bi)^n = (x + yi)(x - yi).$$

A zatem rozwiązanie, to $x = \operatorname{Re}(a + bi)^n$ oraz $y = \operatorname{Im}(a + bi)^n$. Są to oczywiście liczby naturalne.

Zadanie. Rozwiązać w liczbach zespolonych równanie: $z^4 - 6z^3 + 18z^2 - 30z + 25 = 0$.

W tym przypadku możliwe są różne podejścia, na przykład korzystając z twierdzenia z wykładu można argumentować, że istnieją $a, b, c, d \in \mathbb{R}$, że:

$$z^4 - 6z^3 + 18z^2 - 30z + 25 = (z^2 + az + b)(z^2 + cz + d).$$

²Opieram się o tekst dr. Michała Krycha: „Po co komu wymierność?”, dostępny pod adresem: <https://www.mimuw.edu.pl/~krych/odczyty/18-09-13-warszawa.pdf>.

Trzeba zatem rozwiązać układ równań:

$$a + c = -6, \quad b + d + ac = 18, \quad ad + bc = -30, \quad 25 = bd.$$

W tym przypadku akurat kładąc $b = d = 5$ dostajemy układ $a + c = -6, ac = 8, 5(a + c) = -30$, co daje $a^2 + 6a + 8 = (a + 4)(a + 2) = 0$, czyli $a = -4, c = -2$ (lub odwrotnie). A zatem:

$$z^4 - 6z^3 + 18z^2 - 30z + 25 = (z^2 - 2z + 5) = 0.$$

Widać, że te równania kwadratowe „da” się dalej rozwiązać. Udało się. Czy można było to zrobić inaczej? Może, ale będzie trzeba trochę „gdybać”. Spróbujmy. „Gdyby” istniał pierwiastek postaci $z = a + bi$, gdzie a, b są całkowite, również $a - bi$ byłby pierwiastkiem, a więc mielibyśmy $25 = (a^2 + b^2)z_3z_4$, gdzie z_3, z_4 to pozostałe pierwiastki. „Gdyby” jeszcze z_3, z_4 również miały całkowite części rzeczywiste i ujęte wówczas $a^2 + b^2$ byłby dzielnikiem 25. To teoretycznie nie musi się zdarzyć (nie mamy narzędzi, by stwierdzić czy tak musi być), ale niewiele jest liczb całkowitych a, b , że $a^2 + b^2$ dzieli 25, więc warto „pogdybać”. Może jednym z pierwiastków jest liczba „całkowita” $a + bi$ postaci:

$$\pm 1, \quad \pm i, \quad \pm 5, \quad \pm 5i, \quad \pm 2 \pm i, \quad \pm 1 \pm 2i, \quad \pm 5 \pm 5i.$$

Nietrudno sprawdzić, że wśród tych liczb właśnie liczby $2 \pm i$ oraz $1 \pm 2i$ są rozwiązaniami naszego równania. To rodzi rozmaite domysły: czy przypadkiem nie mamy tu do czynienia z jakąś wersją twierdzenia o pierwiastku całkowitym/wymiernym wielomianu o współczynnikach całkowitych? Tak rzeczywiście jest i wiąże się to z faktem, że podzbiór liczb zespolonych postaci

$$\mathbb{Z}[i] = \{a + bi \mid a, b \in \mathbb{Z}\},$$

czyli tzw. **pierścień liczb całkowitych Gaussa**, ma jednoznaczność rozkładu na czynniki pierwsze. Co to znaczy? Czym są te czynniki? Czytelnik zainteresowany bliższym poznaniem tych liczb i związanej z nimi teorii podzielności zechce zajrzeć do:

M. Krych: Skąd się wzięła liczba i , <https://smp.uph.edu.pl/msn/34/krych.pdf>.

W tekście tym znajdują się informacje nie tylko o liczbach całkowitych Gaussa, ale też o tzw. **liczbach całkowitych Eisensteina** $\mathbb{Z}[\omega_3]$, złożonych z liczb postaci $a + bw_3$, gdzie $a, b \in \mathbb{Z}$ oraz w_3 jest nierzeczywistym pierwiastkiem stopnia 3 z 1. Również w tym zbiorze zachodzi teoria podzielności, a nawet twierdzenie o jednoznaczny rozkładzie... Dlaczego ten jednoznaczny rozkład jest tak istotny?

Prawie 400 lat temu Pierre de Fermat stwierdził, że znalazł „niezwykły dowód” następującego twierdzenia:

Twierdzenie 5.5.2: Wielkie Twierdzenie Fermata

Równanie diofantyczne: $x^n + y^n = z^n$, gdzie x, y, z, n są niezerowymi liczbami całkowitymi, nie ma rozwiązań, dla $n > 2$.

Niestety, Fermat nie był w stanie przedstawić rozwiązania, ponieważ swoje odkrycie zapisał na marginesie kopii starożytnego dzieła *Arytmetyka* Diofantosa. Stwierdził jedynie, że *margines jest zbyt mały, by pomieścić dowód*. Notatka Fermata stała się jedną z wielu nieudowodnionych obserwacji, zostawionych kolejnym pokoleniom. Jak się okazało, wiele przypuszczeń Fermata zostało z czasem rozstrzygniętych. Jedną z osób, która poświęciła im sporo miejsca był sam Euler. Nie był on jednak w stanie pokazać ogólnego dowodu powyższego rezultatu. Z trudem znalazł niełatwą uzasadnienie dla $n = 3$ (używając liczb zespolonych, o czym można przeczytać w tekście dr. Krycha). Problem stał się jednym z naj słynniejszych w historii matematyki, a także źródłem rozwoju licznych jej dziedzin. Twierdzenie Fermata zostało udowodnione dopiero w 1994 roku przez Andrew Wilesa.

Przez stulecia szukano błyskotliwych, krótkich dowodów hipotezy Fermata. Jedna z takich nieudanych prób warta jest jednak przypomnienia, ponieważ dała początek rozwojowi współczesnej teorii liczb. Cofniemy się do roku 1847. Problem Fermata był już wówczas jednym z największych wyzwań matematycznych. Centrum matematycznego świata wciąż jeszcze leżało w Paryżu (niedługo potem trafic miało do Gatyngi, a potem za Ocean Atlantycki). Francuska Akademia Nauk oferowała (od 31 lat) złoty medal i nagrodę 3000 franków za rozwiązanie problemu Fermata. Na posiedzeniu 1 marca, z propozycją dowodu wystąpił znany matematyk Gabriel Lamé. Twierdził, że znalazł cudowne rozwiązanie, bardzo krótkie. Idea dowodu była rzeczywiście niezwykle prosta, a dla jej przedstawienia oprzemy się na następującej obserwacji (związanej bezpośrednio z treścią wykładu).

Uwaga 5.5.3

Niech $n \geq 1$ będzie liczbą całkowitą oraz niech $\zeta = \cos \frac{2pi}{n} + i \cdot \sin \frac{2\pi}{n}$ będzie pierwiastkiem pierwotnym stopnia n z 1. Wówczas

$$z^n - 1 = \prod_{k=0}^{n-1} (z - \zeta^k) = (z - 1)(z - \zeta)(z - \zeta^2) \dots (z - \zeta^{n-1}).$$

Dowód jest oczywisty, bowiem dla $0 \leq k \leq n-1$ liczba ζ^k podniesiona do potęgi n równa jest 1, zaś wszystkie te liczby są parami różne.

Wróćmy do argumentu Lame. Przedstawmy $x^n + y^n = z^n$ jako iloczyn n czynników „całkowitych” na dwa sposoby. Jak? Weźmy $\zeta \in \mathbb{C}$ takie, że $\zeta^n = 1$, $\zeta \neq 1$ oraz n – nieparzyste. Dostaniemy wówczas:

$$x^n + y^n = (x + y)(x + \zeta y)(x + \zeta^2 y) \dots (x + \zeta^{n-1} y) = z \cdot z \cdot \dots \cdot z$$

Formuła ta wynika natychmiast z powyższej uwagi, dla $z = -x/y$. Co tu widzimy? Lamé wysnuwa stąd wniosek, że $x + y$ oraz z mają wspólny dzielnik, co prowadziłoby do sprzeczności. Opiera się przy tym na „względnej pierwszości” czynników uzyskanego rozkładu (można się do niej ograniczyć).

Nie tylko Lamé był niezwykle przejęty zaproponowanym dowodem. Również Cauchy wystąpił i stwierdził, że od dłuższego czasu pracuje nad dowodem, w zasadzie opartym na analogicznych obserwacjach. Obydwie część „zasługi” oddawali Josephowi Liouillowi, który zasugerował im rozważanie liczb ze spolonych w kontekście problemu Fermata. Paradoksalnie, to właśnie Liouville zwrócił uwagę na pewien problem. Zaproponowany wyżej rozkład wyrażenia $x^n + y^n$ na „czynniki względnie pierwsze” dokonuje się w zbiorze liczb $\mathbb{Z}[\zeta]$ postaci:

$$a_1 + a_2 \zeta + a_3 \zeta^2 + \dots + a_{n-1} \zeta^{n-1}, a_i \in \mathbb{Z}.$$

Nie ma gwarancji, że w zbiorze tym zachodzi jednoznaczność rozkładu na czynniki. Gdyby jej nie było, wówczas wyciągnięcie wniosku, że każdy czynnik $x^n + y^n$ jest n -tą potęgą nie jest możliwe... Do tego momentu Czytelnik ma prawo być już poważnie zniecierpliwiony: ani nie powiedzieliśmy czym jest „całkowitość” w \mathbb{C} , ani czym są czynniki pierwsze, nieroziądalne czy względnie pierwsze w $\mathbb{Z}[\zeta]$. Jeśli tak jest, to być może osiągnąłem swój cel. Dokładny opis tego problemu przekracza ów skromny dodatek, ale jest absolutnie w zasięgu. Proszę jedynie o kontynuowanie lektury przy bardziej kompetentnym źródle: artykuły prof. Balcerzyka i dr. Szurka: „*Nieco historii matematyki w wykładzie algebra*”: <http://www.deltami.edu.pl/temat/matematyka/2016/05/30/1981-05-Fermat.pdf>.

Na koniec warto dodać jeszcze jeden komentarz dotyczący wielomianów $x^n - 1$, które pojawiły się po drodze. Można zapytać: jak wyglądają rozkłady tych wielomianów na iloczyny wielomianów niższych (dodatnich) stopni o współczynnikach całkowitych? Nie jest to proste zagadnienie. Gdybyśmy rozkładaли na wielomiany o współczynnikach rzeczywistych, to nie ma żadnego problemu, bo $\overline{\zeta^k} = \zeta^{n-k}$, co oznacza, że $x^n - 1$ ma czynniki liniowe postaci $(x - 1)$ (zawsze), $(x + 1)$ (jeśli n jest liczbą parzystą) oraz (dla $n > 2$) czynniki kwadratowe (o współczynnikach rzeczywistych) postaci

$$x^2 - 2x \cos \frac{2k\pi}{n} + 1 = (x - \zeta^k)(x - \zeta^{n-k}),$$

gdzie $k \notin \{0, \frac{n}{2}\}$. Jeśli jednak ograniczymy się tylko do czynników w $\mathbb{Z}[x]$, to np. dla $n = 15$ mamy:

$$x^{15} - 1 = (x - 1)(x^2 + x + 1)(x^4 + x^3 + x^2 + x + 1)(x^8 - x^7 + x^5 - x^4 + x^3 - x + 1).$$

Czym są cztery czynniki, występujące w tym rozkładzie? Są to tak zwane: pierwszy, trzeci, piąty i piętnasty WIELOMIAN CYKLOTOMICZNY. Z definicji, pierwiastkami n -tego wielomianu cyklotomicznego są pierwiastki pierwotne stopnia n z jedynki. Krótko mówiąc: stopnie tych wielomianów wynoszą odpowiednio: 1, 2, 4, 8, bowiem wśród pierwiastków stopnia 15 z 1 jest jeden pierwiastek pierwotny stopnia 1, dwa pierwiastki pierwotne stopnia 2, cztery pierwiastki pierwotne stopnia 5 oraz 8 pierwiastków pierwotnych stopnia 15. Jak się okazuje, wielomiany cyklotomiczne są nieroziądalne na iloczyn wielomianów niższych (dodatnich) stopni o współczynnikach w \mathbb{Z} .

Temat wielomianów cyklotomicznych jest niezwykle pięknym i ciekawym fragmentem teorii wielomianów będącym na styku algebry i teorii liczb. Czytelnika zainteresowanego tym zagadnieniem odsyłam do tekstu prof. Andrzeja Nowickiego: <https://www-users.mat.umk.pl/~anow/imperium/wlm12.pdf>.

5.6 Dodatek. Kwaterniony

W tym podrozdziale przyjrzymy się pewnej próbie uogólnienia definicji liczb zespolonych, zaprezentowanej na wykładzie 4. Definicja ta określała działania dodawania i mnożenia punktów na płaszczyźnie, czyli na uporządkowanych parach liczb rzeczywistych. Prowadzi ona do naturalnego pytania: czy możliwe jest określenie analogicznej konstrukcji na trójkach lub czwórkach punktów? Jak się okazuje, nie jest możliwe wprowadzenie struktury ciała na zbiorze $\{(x_1, x_2, x_3) \mid x_i \in \mathbb{R}\}$ (uzasadnimy ten fakt w jednym z kolejnych rozdziałów). Istnieje natomiast stosowna, choć niezupełnie skuteczna, konstrukcja działań dla zbiorze czwórek liczb rzeczywistych, odkryta w roku 1843 przez Rowana Hamiltona.

Definicja 5.6.1: Kwaterniony

Niech \mathbb{H} będzie zbiorem czwórek uporządkowanych liczb rzeczywistych postaci

$$\mathbb{H} = \{(x_1, x_2, x_3, x_4) \mid x_1, x_2, x_3, x_4 \in \mathbb{R}\}.$$

Dla dowolnych dwóch elementów $h = (a, b, c, d), h' = (a', b', c', d')$ ze zbioru \mathbb{H} określamy operacje dodawania i mnożenia wzorem:

$$h + h' = (a + a', b + b', c + c', d + d')$$

$$h \cdot h' = (aa' - bb' - cc' - dd', ab' + ba' + cd' - dc', ac' - bd' + ca' + db', ad' + bc' - cb' + da').$$

Przyjmując $1 = (1, 0, 0, 0), i = (0, 1, 0, 0), j = (0, 0, 1, 0)$ oraz $k = (0, 0, 0, 1)$ nietrudno sprawdzić, że dowolny element \mathbb{H} zapisać można w postaci

$$a \cdot 1 + bi + cj + dk,$$

gdzie $a, b, c, d \in \mathbb{R}$, przy czym zwykle pomijamy pisanie czynnika 1 oraz zakładamy, że elementy

$$h = a + bi + cj + dk \quad \text{oraz} \quad h' = a' + b'i + c'j + d'k$$

są równe wtedy i tylko wtedy, gdy $a = a', b = b', c = c', d = d'$. Wprowadzone wyżej działania pozwalają dodawać i mnożyć liczby w powyższej postaci podobnie jak liczby zespolone, zgodnie z regułami:

$$ij = k, \quad jk = i, \quad ki = j, \quad ji = -k, \quad kj = -i, \quad ik = -j, \quad i^2 = j^2 = k^2 = -1.$$

Przykładowo, mamy $i \cdot (-j) = (0, 1, 0, 0) \cdot (0, 0, -1, 0) = (0, 0, 0, -1) = -k$.

Zauważmy, że $(0, 0, 0, 0)$ oraz $(1, 0, 0, 0)$ są elementami neutralnymi dodawania i mnożenia w \mathbb{H} , w rozumieniu aksjomatów ciała przedstawionych w podrozdziale 2.1. Jak się okazuje, co sprowadzić można do żmudnego przeliczenia, wprowadzone działania wraz z wyróżnionymi wyżej elementami spełniają wszystkie aksjomaty ciała, poza aksjomatem przemienności mnożenia. Jak się okazuje, radykalnej zmianie ulega w związku z tym szereg zagadnień, choćby problem rozwiązywania układów równań o współczynnikach kwaternionowych, czy też problem wyznaczania pierwiastków wielomianów o współczynnikach kwaternionowych (definicje tych obiektów są analogiczne, jak dla dowolnego ciała).

Rozważmy przykład związany z wielomianem $x^2 + 1$, traktując jego współczynniki jako kwaterniony. Zauważmy, że funkcja $f(x) = x^2 + 1$ ma co najmniej trzy kwaternionowe zera. Są to oczywiście elementy i, j, k . Okazuje się, że wielomian ten ma więcej pierwiastków, choćby

$$\begin{aligned} f\left(\frac{i}{\sqrt{3}} + \frac{j}{\sqrt{3}} + \frac{k}{\sqrt{3}}\right) &= \frac{i^2}{3} + \frac{j^2}{3} + \frac{k^2}{3} + \frac{ij}{3} + \frac{ik}{3} + \frac{jk}{3} + \frac{ji}{3} + \frac{ki}{3} + \frac{kj}{3} + 1 \\ &= \frac{k + j + i - k - j - i}{3} = 0. \end{aligned}$$

W podobny sposób możemy w istocie otrzymać nieskończenie wiele pierwiastków. Prosty rachunek pokazuje, że wśród kwaternionów postaci $a + bi + cj + dk$ funkcja $f(x) = x^2 + 1$ przyjmuje wartość zero dokładnie na tych argumentach, których współczynnik a równy jest 0 (tzw. kwaterniony czyste), a które leżą na sferze opisanej równaniem $b^2 + c^2 + d^2 = 1$.

Innym przykładem znaczenia przemienności kwaterionów (lub jej braku) jest problem prawdziwości twierdzenia Bézout. Rozważmy iloczyn wielomianów

$$h(x) = (x - i)(x - j).$$

Jeśli wykonamy iloczyn zgodnie z zasadą mnożenia wielomianów opisaną w podrozdziale 2.4, dostaniemy

$$h(x) = x^2 - ix - xj + ij = x^2 - (i + j)x + ij.$$

W ten sposób poznajemy współczynniki wielomianu $h(x) \in \mathbb{H}[x]$. Podstawiając jednak do odpowiadającej mu funkcji wielomianowej kwaterion $s = i$, dostajemy

$$h(i) = i^2 - (i + j)i + ij = 2k \neq 0.$$

Wydaje się to niezgodne z intuicją, skoro naturalnie wyglądające przeniesienie twierdzenia Bézout na przypadek kwaterionowy sugerowałoby, że wielomian $h(x)$ ma pierwiastek $s = i$, skoro ma dzielnik $x - i$. Okazuje się jednak, że w przypadku struktury nieprzemiennej, z rozkładu wielomianu $f = gh$ na czynniki nie wynika, że wartość funkcji wielomianowej $f(s)$ rozkłada się dla każdego s na czynniki $g(s) \cdot h(s)$.

Czytelnik zechce jednak sprawdzić, że dla wielomianu wyżej j jest pierwiastkiem, to znaczy $h(j) = 0$. Okazuje się, że wykorzystując wielokrotnie założenie $xs = sx$, dla każdego $s \in \mathbb{H}$ można udowodnić (identycznie jak w podrozdziale 2.4) następującą „jednostronną” wersję twierdzenia Bézout: dla niezerowego wielomianu $f \in \mathbb{H}[x]$ równoważne są warunki: istnienie pierwiastka s wielomianu f oraz istnienie rozkładu $f = g(x - s)$, dla pewnego wielomianu $g \in \mathbb{H}[x]$.

Do poniższego dodatku wyjątkowo dodaję zestaw zadań. Zawiera on między innymi plan dowodu twierdzenia pokazującego w jaki sposób uogólnić na przypadek kwaterionów twierdzenie mówiące, że wielomian stopnia n ma co najwyżej n różnych pierwiastków.

1. Uzasadnij, że jeśli $a, b \in \mathbb{H}$, to warunek $ab = 0$ implikuje $a = 0$ lub $b = 0$.
2. Niech $h = a + bi + cj + dk \in \mathbb{H}$. Sprzężeniem kwaterionu h nazywamy kwaterion $\bar{h} = a - bi - cj - dk$, zaś modelem kwaterionu h nazywamy liczbę rzeczywistą $|h| = \sqrt{a^2 + b^2 + c^2 + d^2}$. Uzasadnij, że dla dowolnych $h, h' \in \mathbb{H}$ mamy $h \cdot \bar{h} = \bar{h} \cdot h = |h|^2$, $|hh'| = |h| \cdot |h'|$, $|h + h'| \leq |h| + |h'|$.
3. Wykaż, że jeśli każdą z liczb całkowitych n, m można przedstawić w postaci sumy czterech kwadratów liczb całkowitych, to również liczbę nm można przedstawić w postaci sumy czterech kwadratów liczb całkowitych.
4. Uzasadnij, że dla każdego niezerowego kwaterionu h istnieje kwaterion h' , że $hh' = h'h = 1$, który nazywamy elementem odwrotnym do h i oznaczamy h^{-1} .
5. Uzasadnij sformułowaną wyżej „jednostronną” wersję twierdzenia Bézout.
6. Dla $s \in \mathbb{H}$ przez klasę sprzężoności s rozumiemy zbiór elementów postaci $\{asa^{-1} \mid a \in \mathbb{H} \setminus \{0\}\}$. Wykaż, że dla dowolnych dwóch kwaterionów a, b klasy sprzężoności tych elementów są równe lub rozłączne.
7. Kwaterion $h = a + bi + cj + dk$ nazywamy czystym, jeśli $a = 0$. Wykaż, że klasa sprzężoności kwaterionu czystego zawiera wyłącznie kwateriony czyste.
8. Niech $f = g \cdot h$, gdzie $f, g, h \in \mathbb{H}[x]$. Niech $s \in \mathbb{H}$ będzie taki, że $a = h(s) = 0$. Wówczas

$$f(s) = g(asa^{-1})h(s).$$

W szczególności, jeśli s jest pierwiastkiem wielomianu f , ale nie jest pierwiastkiem wielomianu h , to asa^{-1} jest pierwiastkiem wielomianu g .

9. Niech f będzie niezerowym wielomianem stopnia n w $\mathbb{H}[x]$. Udowodnij, że wszystkie pierwiastki f należą do co najwyżej n klas sprzężoności elementów w \mathbb{H} . Co więcej, jeśli

$$f = (t - s_1) \cdot \dots \cdot (t - s_n),$$

dla pewnych $s_1, \dots, s_n \in \mathbb{H}$, to każdy pierwiastek wielomianu f jest w klasie sprzężoności któregoś z elementów s_i .

5.7 Coda. Wokół rozkładu na czynniki wielomianów i ich funkcji

Po dość rozbudowanych dodatkach traktujących o rozkładalności, warto powiedzieć kilka słów o ogólnej koncepcji stojącej za rozkładem, umieszczając ją w odpowiednim kontekście historycznym. W powyższych rozważaniach wspomnieliśmy bowiem tak o istnieniu, jak i o jednoznaczności rozkładu liczb całkowitych (różnych od $-1, 0, 1$) na czynniki pierwsze, o analogicznej własności wielomianów o współczynnikach w ciele oraz o przedziwnych własnościach rozkładów w pierścieniach typu $\mathbb{Z}[\epsilon_n]$, gdzie ϵ_n jest pierwiastkiem stopnia n z 1 (dla $n = 2$ otrzymujemy pierścień liczb całkowitych Gaussa, dla $n = 3$ — liczby Eisensteina, a dla większych n — liczby rozważane w kontekście Wielkiego Twierdzenia Fermata). Rozważania te mają już bardzo współczesny charakter. Poprzedziło je jednak kilka stuleci rozważań nad problemami bardziej chyba dla nas namacalnymi. Wyznaczyły one nurt ważnej koncepcji matematycznej.

Wspomnieliśmy w poprzednim dodatku o przełomie, jaki dokonał Clavius, a po nim Kartezjusz, kojarząc związek rozkładalności na czynniki z istnieniem rozwiązań równań wielomianowych. Zamiast stosować metody geometryczne do rozwiązywania równania typu $x^2 = 9x + 70$, można przenieść wszystkie wyrazy na jedną stronę i przedstawić równanie $x^2 - 9x - 70 = 0$ w postaci równoważnej: $(x - 14)(x + 5) = 0$, uzyskując dwa rozwiązania. Kluczowa koncepcja polega na (niełatwiej często) redukcji złożonego problemu do układu problemów łatwych: równanie stopnia drugiego sprowadzamy do dwóch łatwych równań stopnia pierwszego. Tego typu sposób działania ma fundamentalne znaczenie. Nie bez przesady będzie stwierdzenie, że na kursie algebry liniowej wielomiany potrzebować będziemy właśnie po to (w drugim semestrze), aby skomplikowaną naturę przekształcenia geometrycznego „rozkładać” na składowe mające bardzo czytelną interpretację geometryczną. Zwrócimy więc jedynie uwagę na kilka wątków, w których rozkład na czynniki liniowe odkrywał ważną rolę. Zaczniemy jednak od kwestii zasadniczej, czyli twierdzenia Bezout, należącego w istocie do Kartezjusza.

Aż do XVII wieku teoria równań wielomianowych była, jak wspominaliśmy w rozdziale o wzorach skróconego mnożenia, zagadnieniem rozważanym w języku geometrycznym. *Geometria* Kartezjusza z roku 1637 zawierała ona dwa istotne wzbogacenia dotychczasowej teorii i notacji, wprowadzonej częściowo już w XVI wieku przez Vite'a, a mianowicie czytelną notację wykładniczą: x^3, x^4, x^5 itd. (choć nie x^2 , które pozostało jako xx aż do XVIII wieku) i właśnie owo twierdzenie Bezout. Zobaczmy nieco szerszy kontekst.

Gdy mówimy o rozkładzie $w(x) = (x - c) \cdot v(x)$, to mamy na myśli, że po wymnożeniu $(x - c) \cdot v(x)$ otrzymamy wielomian, który ma takie same współczynniki, jak $w(x)$. Dlaczego to jest ważne? Chcemy bowiem korzystać z następującej implikacji: jeśli $w(x), v(x) \in K[x]$ oraz mamy **rozkład wielomianu na czynniki**: $h(x) = w(x) \cdot v(x)$, to dla każdego $s \in K$ mamy **rozkład wartości funkcji wielomianowej**

$$h(s) = w(s) \cdot v(s).$$

Cóż to za szaleństwo, czy to nie jest oczywiste? Dla wielomianów o współczynnikach w ciele, istotnie jest to prawda. Przyjęcie współczynników innego typu może to zaburzyć, z czym spotkaliśmy się w dodatku o kwaterionach.

Rozważmy przyporządkowanie działające w następujący sposób: dla każdego $a \in K$ rozważamy funkcję $v_a : K[x] \rightarrow K$, która przyporządkowuje wielomianowi postaci $w(x) = r_0 + r_1x + r_2x^2 + \dots + r_nx^n \in K[x]$ wartość odpowiadającej mu funkcji wielomianowej w punkcie a , a więc element

$$v_a(w) = r_0 + r_1a + r_2a^2 + \dots + r_na^n \in K.$$

Funkcja ta nazywa się **EWALUACJĄ WIELOMIANU** w punkcie a . Przyzwyczailiśmy się do pewnych własności ewaluacji, na przykład do następujących. Dla każdego $a \in K$ mamy (tzw. homomorfizm pierścienia):

$$v_a(w + w') = v_a(w) + v_a(w'), \quad v_a(w \cdot w') = v_a(w) \cdot v_a(w').$$

Jak się jednak okazuje, tak być nie musi, jeśli zbiór współczynników wielomianu nie jest przemienny, czego przykład mamy w przypadku kwaterionów! Czytelnik może odczuwać pewną konsternację, dochodząc do tej konkluzji. Wydaje mi się jednak ważne, by pokazać, że stwierdzenie „funkcja wielomianowa iloczynu to iloczyn funkcji wielomianowych” ma głębokie podłożę.

Proszę zauważyc, że ewaluacja może zdecydowanie ułatwić wykonywanie rachunków. Oto przykład.

Zadanie. Rozważmy zbiór $\mathcal{A} = \{\varepsilon_0, \dots, \varepsilon_{11}\}$ pierwiastków zespolonych stopnia 12 z 1. Uzasadnij, że

$$(\sqrt{3} + i - \varepsilon_0) \cdot (\sqrt{3} + i - \varepsilon_1) \cdot \dots \cdot (\sqrt{3} + i - \varepsilon_{11}) = 2^{12} - 1.$$

Proszę zauważać, że pierwiastki stopnia 12 z 1 można wyznaczyć, wyliczając ich postaci ogólne. W ten sposób możliwe jest policzenie powyższego iloczynu przez wymnożenie 12 nawiasów, odpowiednio grupując czynniki zawierające pierwiastki sprzężone. Znacznie łatwiej jest jednak zauważać, że wobec rozkładu

$$x^{12} - 1 = (x - \varepsilon_0)(x - \varepsilon_1) \cdot \dots \cdot (x - \varepsilon_{11})$$

możemy wyznaczyć żądaną iloczyn, poprzez wyznaczenie wartości funkcji wielomianowej odpowiadającej wielomianowi $x^{12} - 1$ w punkcie $s = \sqrt{3} + i$. Innymi słowy, mamy:

$$(\sqrt{3} + i)^{12} - 1 = (\sqrt{3} + i - \varepsilon_0) \cdot (\sqrt{3} + i - \varepsilon_1) \cdot \dots \cdot (\sqrt{3} + i - \varepsilon_{11}) = 2^{12} - 1.$$

Nie tylko przemienność mnożenia w zbiorze współczynników, ale i inna kwestia wchodzi w grę. Chcemy bowiem z tego, że $0 = h(s) = w(s) \cdot v(s)$ wnioskować, że $w(s) = 0$ lub $v(s) = 0$. Bez tej własności rozwiązywanie równań będzie często trudne. Wystarczy odejść niedaleko od definicji ciała, by znaleźć zbiory współczynników nie mające takiej własności. Rozważmy zbiór reszt z dzielenia przez 6, czyli \mathbb{Z}_6 . Jego definicja jest analogiczna, jak dla ciał \mathbb{Z}_p , ale nie jest to jednak ciało, ponieważ nie każdy element \mathbb{Z}_6 ma odwrotność — choćby element 2. Mamy wręcz $2 \cdot 3 = 0$, co dla ciał nie może mieć miejsca. Wystarczy zobaczyć, że zwykły wielomian $x^2 + 5x \in \mathbb{Z}_6[x]$ o współczynnikach w pierścieniu reszt z dzielenia przez 6 ma więcej niż dwa pierwiastki, bowiem

$$x(x+5) = (x+2)(x+3) = x^2 + 5x,$$

co dla ciał nie jest możliwe.

Przejdźmy do ważnych zastosowań historycznych rozmaitych aspektów rozkładu wielomianów na czynniki. Pierwsze wydawać się może zaskakujące, dotyczy bowiem teorii liczb i znany jest pod nazwą małego twierdzenia Fermata. Twierdzenie to pochodzi z roku 1640 i mówi, że jeśli p jest liczbą pierwszą oraz n jest dodatnią liczbą całkowitą względnie pierwszą z p , to liczba $n^{p-1} - 1$ jest podzielna przez p , lub równoważnie — liczba $n^p - n$ jest podzielna przez p .

Twierdzenie to stało się współcześnie jednym z podstawowych narzędzi kryptograficznych, stąd wydaje się ciekawe wspomnienie, że Fermata w istocie interesował się kiedy liczba $2^m - 1$ ma dzielniaki pierwsze. W istocie, największe znajdowane dziś liczby pierwsze są tej postaci (tzw. liczby pierwsze Mersenne'a). Z punktu widzenia teorii rozkładu wielomianów, małe twierdzenie Fermata mówi, że w ciele \mathbb{Z}_p wielomian $x^p - x$ rozkłada się na iloczyn czynników liniowych:

$$x^p - x = \prod_{a \in \mathbb{Z}_p} (x - a),$$

czyli mówiąc inaczej — każdy element tego ciała jest jego pierwiastkiem (a jednak nie jest to wielomian zerowy). Faktoryzacja wielomianów nad ciałami skończonymi ma fundamentalne znaczenie w kriptografii. Do jej zastosowań należy wyznaczanie tzw. dyskretnego logarytmu, niezbędnego przy konstrukcji szyfrowania klucza publicznego. Idee znajdowania tego typu rozkładów pochodząły od Legendre'a i Gaussa, a w XX wieku rozwinięte zostały przez Berlekampa, czy współcześnie przez Cantora i Zassenhausa.

Drugie spojrzenie również pochodzi z pierwszej połowy XVII wieku, gdy Fermat i Kartezjusz niezależnie budowali podstawy geometrii analitycznej. Motywacją było badanie krzywych opisanych przez równania. Krzywym przypisywano stopień: proste to krzywe stopnia 1, krzywe stożkowe (okrąg, parabola, hiperbola) to krzywe stopnia 2, tzw. koniki to krzywe stopnia 3, w tym na przykład krzywa $y^2 = x^3$, niebędąca wykresem funkcji. Podstawowym celem było osiągnięcie następującego rezultatu, zwanego twierdzeniem Bezout: krzywa stopnia m przecina zawsze krzywą stopnia n w nie więcej niż mn punktach. Dla przykładu: prosta przecina okrąg w nie więcej niż dwóch punktach, ale już dwie elipsy mogą się przeciąć w czterech punktach, podobnie okrąg i parabola. Dlaczego twierdzenie to było tak istotne?

W okolicach roku 1620 Kartezjusz zorientował się, że dowolne rozwiązywanie równania wielomianowego stopnia 3 lub 4 może skonstruować poprzez przecinanie krzywych stopnia 2. Co więcej, w swoim fundamentalnym dziele *Geometria* (1637), wskazał on krzywą stopnia 3, zwaną parabolą kartezjańską, której

przecięcie z odpowiednim okręgiem dawało rozwiązywanie dowolnego równania wielomianowego stopnia 5 lub 6. Oczywiście dało mu to niezachwaną wiarę w możliwość uogólnienia tych rezultatów. Jak się okazało, stosowne uogólnienie nie było banalne i odpowiednie konstrukcje rozwiązań wielomianów stopnia n ustalonono dopiero około roku 1750. Co to wszystko ma wspólnego z rozkładem?

Problem wyznaczania liczby przecięć krzywych sprowadzić można do problemu rozwiązywania równań wielomianowych stopnia n . W prostych przypadkach jest to oczywiste. Badając liczbę przecięć elipsy $x^2 + 2y^2 = 1$ oraz paraboli $y = x^2$ należy jedynie wstawić w pierwszym równaniu x^2 w miejsce y i rozwiązać równanie stopnia 4. Otrzymamy zatem nie więcej niż 4 punkty przecięcia.

Czasem proste rozdzielenie zmiennych nie jest zupełnie jasne, zwłaszcza gdy w grę wchodzą tzw. punkty wielokrotne czy punkty w nieskończoności. Niemniej jednak dowód zasadniczego twierdzenia algebry oraz rozwój metod geometrii rzutowej (i związanych z nimi tzw. wielomianów jednorodnych) pozwolił ustalić, że takie rozdzielenie zmiennych jest w istocie zawsze możliwe. Rezultat ten uzyskano ostatecznie dopiero pod koniec XIX wieku, wykorzystując teorię wyznaczników.

Co natomiast można powiedzieć o historii samego dowodu zasadniczego twierdzenia algebry? Rezultat z wykładu mówiący, że wielomian o współczynnikach rzeczywistych mający pierwiastek zespolony $z = a + bi$ ma również pierwiastek sprzężony $\bar{z} = a - bi$ pochodzi od d'Alemberta (1746). Wraz z nim uzyskano oczywiście postawiony przez nas wniosek: zasadnicze twierdzenie algebry jest dla wielomianów rzeczywistych równoważne możliwości rozłożenia każdego takiego wielomianu (dodatniego stopnia) na iloczyn czynników rzeczywistych stopnia 1 lub 2. W ten sposób twierdzenie to formułowało przez większość XVIII stulecia, co pozwoliło unikać wspomnień o (wciąż podejrzany) $\sqrt{-1}$ oraz zezwalało na użytkę metod analizy funkcji rzeczywistych.

Dowody zasadniczego twierdzenia algebry, zarówno te proponowane przez d'Alemberta, jak i pierwotny dowód Gaussa, miały luki, które naprawione zostały dopiero pod koniec XIX wieku i miały charakter analityczny. Jakie było podejście algebraiczne? Główne nurty pochodząły od Eulera, w którego czasach obok ZTA najsłynniejszym problemem było zagadnienie znalezienia wzorów na pierwiastki wielomianów wyższych stopni (ostatecznie rozstrzygnięte negatywnie, dla stopnia ≥ 5 przez Abela, Ruffiniego i Galois), Euler zajął się problemem znanym mu skądinąd z badań nad wielomianem $x^n - 1$. W tym celu wprowadził jednostkę urojoną i zaczął badać wyrażenia postaci $(\cos \theta + i \sin \theta)$ oraz ich potęgi. To właśnie Euler w istocie jako pierwszy sformułował w pełnej wersji znany nam już wzór Moivre'a. W 1749 roku wyznaczył wzory na pierwiastki stopnia n -tego z liczby zespolonej i postulował, że zbiór liczb zespolony jest zamknięty na branie pierwiastków — co jest własnością nieznaną ani liczbom naturalnym, ani całkowitym, wymiernym czy rzeczywistym. Pomijamy w tym miejscu ogromne zastosowania, jakie wniosły prace Eulera do analizy, w tym do przedstawień funkcji w postaci szeregow.

Z punktu widzenia rozkładu na czynniki liniowe, Euler ustalił tożsamość

$$x^n - 1 = (x - \omega_0)(x - \omega_1)(x - \omega_2) \cdots (x - \omega_{n-1}),$$

gdzie $\omega_0, \dots, \omega_{n-1}$ są pierwiastkami stopnia n z 1.

Eulerowi udało się udowodnić, że każdy wielomian rzeczywistych stopnia $n \leq 6$ ma dokładnie n pierwiastków zespolonych. W tym samym roku 1749 podjął próbę ogólnego dowodu, opartą o rozkład wielomianu unormowanego stopnia 2^n na iloczyn wielomianów stopni 2^{n-1} . Ideą była sztuczka znana już z prac Cardano mówiąca, że przez odpowiednie podstawienie można pozbyć się z wielomianu stopnia n wyrazu przy potędze $n - 1$. Planował też udowodnić istnienie rozkładu:

$$x^{2m} + Ax^{2m-2} + Bx^{2m-3} + \dots = (x^m + tx^{m-1} + gx^{m-2} + \dots)(x^m - tx^{m-1} + hx^{m-2} + \dots).$$

Twierdził, że współczynniki A, B będą funkcjami wymiernymi od A, B, \dots, t , ale ogólny przypadek uzasadnił jedynie w formie szkicu, podważonego przez Lagrange'a (zachodziła obawa, że niektóre funkcje wymierne są postaci $0/0$). Dowody przedstawiane przez kolejnych autorów (Laplace, Gauss, Argand) zawierały wciąż luki. Drugie podejście Gaussa z roku 1816 uznane zostało jednak za całkowicie poprawne, i stąd przypisujemy mu pierwszeństwo. Rozumowanie jest w pełni algebraiczne, za wyjątkiem użycia szczególnego przypadku twierdzenia o wartości średniej, które dla wielomianów uzasadnił Bolzano (1817), a dla funkcji ciągły Weierstrass (1874). W 1849 roku, Gauss przedstawił dowód ZTA dla wielomianów zespolonych. Późniejsi autorzy, w tym Frobenius, uznawali zasługi Eulera, którego od pełnego dowodu dzieliła nieumiejętność wykazania, że wielomian rzeczywisty nieparzystego stopnia ma pierwiastek.

Rozdział 6

Przestrzenie liniowe

6.1 Wykład 6

Wykład ten poświęcony będzie pojęciu przestrzeni liniowej nad ciałem. Jest to fundamentalne pojęcie dla całego naszego wykładu i jedno z najważniejszych w całej matematyce. Mówiąc będziemy o strukturze określonej jednocześnie na dwóch typach obiektów: wektorach i skalarach. Struktura ta jest w swojej istocie „geometryczna”, choć odnaleźć ją można w bardzo odległych z pozoru dziedzinach matematyki.

Definicja 6.1.1: Przestrzeń liniowa nad ciałem K

PRZESTRZENIĄ LINIOWĄ NAD CIAŁEM $(K, +, \cdot, 0, 1)$ nazywamy zbiór V , wraz z:

- odwzorowaniem: $\oplus : V \times V \longrightarrow V$, zwany DODAWANIEM WEKTORÓW,
- odwzorowaniem: $\otimes : K \times V \longrightarrow V$, zwany MNOŻENIEM WEKTORA PRZEZ SKALAR,
- wyróżnionym elementem Θ w V zwany WEKTOREM ZEROWYM,

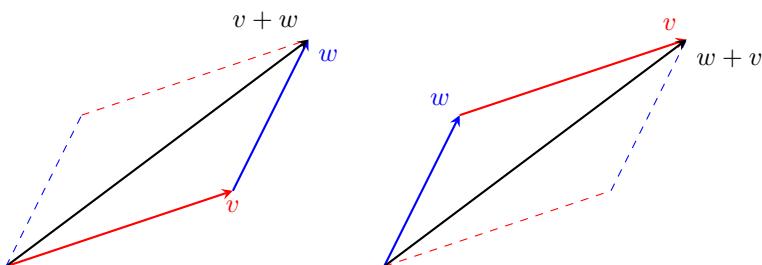
przy czym spełnione są następujące aksjomaty przestrzeni liniowej:

1)	$\alpha \oplus (\beta \oplus \gamma) = (\alpha \oplus \beta) \oplus \gamma$	$\forall_{\alpha, \beta, \gamma \in V}$	łączność dodawania wektorów
2)	$\alpha \oplus \beta = \beta \oplus \alpha$	$\forall_{\alpha, \beta \in V}$	przemienność dodawania wektorów
3)	$\alpha \oplus \Theta = \alpha$	$\forall_{\alpha \in V}$	Θ jest elem. neutralnym \oplus
4)	$\alpha \oplus \gamma = \Theta$	$\forall_{\alpha \in V} \exists_{\gamma \in V}$	istnienie wekt. przeciwnego
5)	$1 \otimes \alpha = \alpha$	$\forall_{\alpha \in V}$	mnożenie wektora przez 1
6)	$(a \cdot b) \otimes \alpha = a \otimes (b \otimes \alpha)$	$\forall_{\alpha, \beta, \gamma \in V}$	zgodność \cdot z \otimes
7)	$(a + b) \otimes \alpha = (a \otimes \alpha) \oplus (b \otimes \alpha)$	$\forall_{\alpha \in V}, \forall_{a, b \in K}$	rozdzielność \otimes względem $+$
8)	$a \otimes (\alpha \oplus \beta) = (a \otimes \alpha) \oplus (a \otimes \beta)$	$\forall_{\alpha, \beta \in V}, \forall_{a \in K}$	rozdzielność \otimes względem \oplus

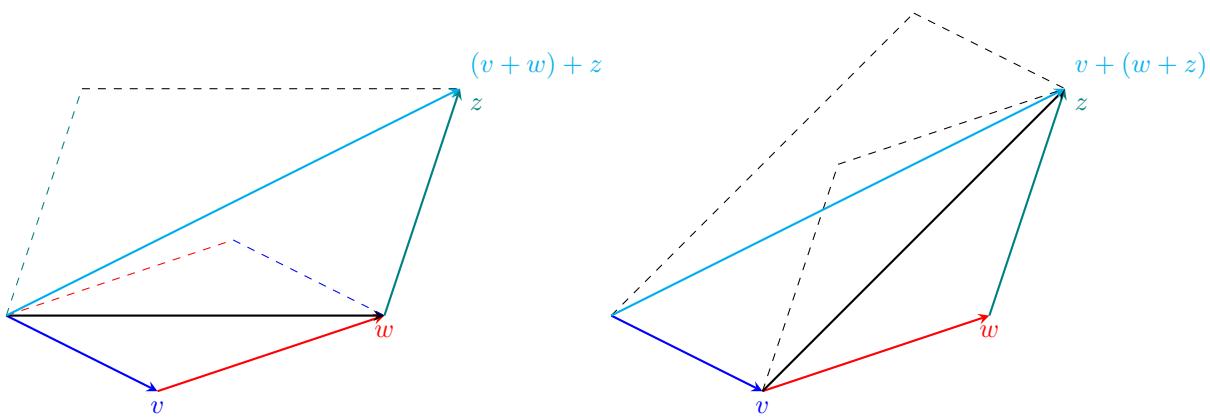
Elementy przestrzeni liniowej V nazywamy WEKTORAMI.

Jak widać, w definicji wystąpiło mnóstwo oznaczeń, zwłaszcza odnośnie działań. W dalszej części wszystkie symbole dodawania $\oplus, +$ będą zamienione na $+$ oraz wszystkie symbole mnożenia \cdot, \otimes będą pomijane.

Zilustrujmy aksjomaty poprzez szkolne intuicje, w myśl których wektor wyznacza kierunek przesunięcia (np. działanie siły), a dodawanie wektorów umożliwia składanie przesunięć (np. wypadkowa układu sił). Dla wektorów v, w , wektor $v + w$ wyznacza przekątną równolegloboku wyznaczonego przez v oraz w .



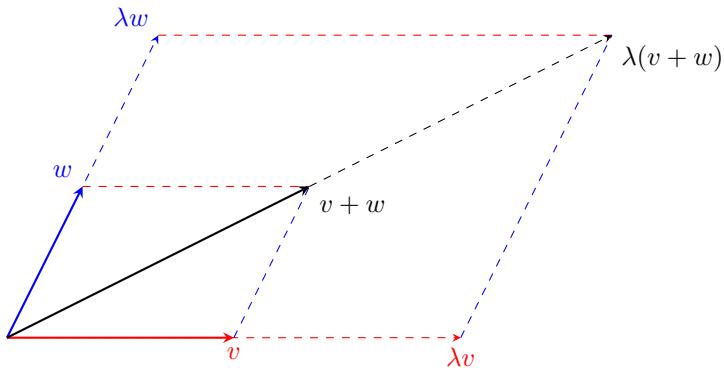
Powyższa ilustracja tłumaczy aksjomat (2) przemienności dodawania wektorów. Oto interpretacja szkolna aksjomatu (1) o łączności dodawania.



Należy pamiętać, że powyższe ilustracje mają charakter komentarza ilustrującego sensowność aksjomatów. W formułowanej przez nas definicji przestrzeni liniowej nie występuje choćby pojęcie punktu, a więc i początku lub końca wektora, które wprowadzimy na gruncie przestrzeni afinicznych w drugim semestrze.

Wektor przeciwny do v , czyli $-v$, reprezentuje kierunek przesunięcia przeciwnego do v tak, by złożenie przesunięć o wektor v oraz $-v$ było identycznością, czyli przesunięciem o wektor 0.

Mnożenie przez skalar interpretujemy jako skalowanie (jednokładność). Szczególne znaczenie ma aksjomat (8) rozdzielności mnożenia przez skalar względem dodawania wektorów. Czy Czytelnik widzi w nim abstrakcyjną formę (bez definiowania czym jest długość czy proporcja wektorów) twierdzenia Talesa?



Powyższe aksjomaty, choć nie będą dotyczyć jedynie (a na starcie: w zasadzie w ogóle) wektorów w takim sensie, w jakim poznaliśmy je w szkole, to będą zachowywać własności, jakie mają wektory. Może to się wydawać niejasne, więc użyjmy następującej analogii: założymy, że zapomnieliśmy czym są „szkolne wektory”, i jakie mają geometryczne własności, ale zapisaliśmy sobie na wszelki wypadek kilka kluczowych informacji, z których „chcemy odzyskać” tę wiedzę. Tak można interpretować aksjomaty (również aksjomaty ciała, poznane na wykładzie drugim oraz na Analizie) — zapominamy czym jest choćby prosta rzeczywista \mathbb{R} (a raczej — zapominamy o modelu, którym się posługiwać mówiąc o niej) i zostawiamy jedynie kluczowe własności. Za ich pomocą budujemy pewną teorię, która nie tylko pozwoli nam odzyskać wiedzę ilustrowaną poprzez model, ale pozwoli wyprowadzić nowe własności natury geometrycznej.

Ale czym są w końcu wektory? To jest kluczowe — wektory stanowić mogą elementy dowolnego zbioru, który wraz z odpowiednimi działaniami spełnia listę aksjomatów wypisanych wyżej. Jak się okaże dalej, mogą to być nie tylko ciągi, ale i macierze, wielomiany, funkcje, podzbiory i wiele innych obiektów. Przy odpowiedniej interpretacji, wszystkie one mają te naturalne własności, które przypisywaliśmy wektorom „szkolnym” tak, że na zbiorach tych uprawiamy w istocie geometrię.

Zanim wyprowadzimy podstawowe własności przestrzeni liniowych musimy zapoznać się z dostateczną liczbą przykładów, przekonujących nas do zasadności wprowadzenia tak abstrakcyjnej definicji. Przykłady te pochodzą, jak się okazuje, z wielu różnych gałęzi matematyki.

Definicja 6.1.2: Przestrzeń współrzędnych K^n

Niech K^n oznacza zbiór wszystkich ciągów n -elementowych o wyrazach z ciała K , tzn.:

$$K^n = \{(x_1, x_2, \dots, x_n) \mid x_i \in K, i = 1, 2, \dots, n\}.$$

Wyraz x_i w ciągu (x_1, x_2, \dots, x_n) nazywamy i -TĄ WSPÓŁRZĘDΝĄ tego wektora.

Wprowadzamy działania w K^n . Dla dowolnych $x_1, \dots, x_n, y_1, \dots, y_n, a \in K$ definiujemy:

- dodawanie wektorów: $(x_1, x_2, \dots, x_n) + (y_1, y_2, \dots, y_n) = (x_1 + y_1, x_2 + y_2, \dots, x_n + y_n)$,
- mnożenie wektora przez skalar: $a \cdot (x_1, x_2, \dots, x_n) = (a \cdot x_1, a \cdot x_2, \dots, a \cdot x_n)$.

Przykłady działań w przestrzeni K^n :

- dla $K = \mathbb{Z}_3$ i $n = 4$ mamy np. $(1, 2, 1, 2) + (0, 2, 2, 1) = (1, 1, 0, 0)$, $2 \cdot (2, 2, 1, 1) = (1, 1, 0, 0)$,
- dla $K = \mathbb{C}$ i $n = 2$ mamy np. $(1, i) + i(i, 0) = (1, i) + (-1, 0) = (0, i)$.

Definicja 6.1.3: Przestrzeń liniowa macierzy rozmiaru $m \times n$ o wyrazach w ciele K

Niech $M_{m \times n}(K)$ oznacza zbiór wszystkich macierzy $m \times n$ o wyrazach z ciała K .

- **Sumą** macierzy $[a_{ij}]$ oraz $[b_{ij}]$ z $M_{m \times n}(K)$ nazywamy macierz $[c_{ij}] \in M_{m \times n}(K)$, której wyrazy spełniają warunek $c_{ij} = a_{ij} + b_{ij}$:

$$\begin{bmatrix} & \vdots \\ \cdots & a_{ij} & \cdots \\ & \vdots \end{bmatrix} + \begin{bmatrix} & \vdots \\ \cdots & b_{ij} & \cdots \\ & \vdots \end{bmatrix} = \begin{bmatrix} & \vdots \\ \cdots & a_{ij} + b_{ij} & \cdots \\ & \vdots \end{bmatrix}.$$

- **Iloczynem** macierzy $[d_{ij}] \in M_{m \times n}(K)$ przez skalar $c \in K$ nazywamy macierz $[c \cdot d_{ij}]$:

$$c \cdot \begin{bmatrix} & \vdots \\ \cdots & d_{ij} & \cdots \\ & \vdots \end{bmatrix} = \begin{bmatrix} & \vdots \\ \cdots & c \cdot d_{ij} & \cdots \\ & \vdots \end{bmatrix}.$$

Wektorem zerowym w przestrzeni liniowej $M_{m \times n}(K)$ jest MACIERZ ZEROWA rozmiarów $m \times n$.

Przykładowo, w przestrzeni liniowej $M_{2 \times 3}(\mathbb{Z}_5)$ (ponownie \oplus i \otimes zastępujemy symbolami $+$ oraz \cdot):

$$\begin{bmatrix} 1 & 3 & 2 \\ 0 & 0 & 2 \end{bmatrix} + 2 \cdot \begin{bmatrix} 4 & 4 & 0 \\ 0 & 4 & 2 \end{bmatrix} = \begin{bmatrix} 1 & 3 & 2 \\ 0 & 0 & 2 \end{bmatrix} + \begin{bmatrix} 3 & 3 & 0 \\ 0 & 3 & 4 \end{bmatrix} = \begin{bmatrix} 4 & 1 & 2 \\ 0 & 3 & 1 \end{bmatrix}.$$

Definicja 6.1.4: Przestrzeń liniowa wielomianów o współczynnikach w ciele K

Niech $K[x]$ będzie zbiorem wszystkich wielomianów zmiennej x o współczynnikach w ciele K , czyli

$$K[x] = \{a_0 + a_1x + \dots + a_nx^n \mid n = \mathbb{N} \cup \{0\}, a_0, a_1, \dots, a_n \in K\}.$$

Dodawanie i mnożenie przez skalar pochodzą od omawianych tydzień wcześniej operacji na wielomianach. Wektorem zerowym w przestrzeni liniowej $K[x]$ jest wielomian zerowy.

Zauważmy, że w definicji przestrzeni liniowej wielomianów rozróżniamy działanie mnożenia skalara przez wielomian oraz działanie mnożenia wielomianów. Drugie z tych działań nie jest częścią definicji przestrzeni liniowej (choć będzie nam w różnych sytuacjach potrzebne). Struktura przestrzeni liniowej V z dodatkowym działaniem mnożenia wektorów (zgodnym z działaniami w V) nazywa się algebrą.

Definicja 6.1.5: Przestrzeń liniowa ciągów nieskończonych o wyrazach w ciele K

Oznaczmy przez K^∞ zbiór wszystkich ciągów o wyrazach z ciała K , to znaczy:

$$K^\infty = \{(x_i) \mid x_i \in K, i = 1, 2, \dots\}.$$

Ciągi $x = (x_i)$ oraz $y = (y_i)$ dodajemy i mnożymy przez skalary według zasady:

$$(x \oplus y)_i = x_i + y_i, \quad (a \otimes x)_i = a \cdot x_i.$$

Wektorem zerowym w przestrzeni liniowej K^∞ jest ciąg, którego wszystkie wyrazy są zerem w K .

Przykładowo, z równości:

$$\frac{1}{n} + (-1)\frac{1}{n+1} = \frac{1}{n(n+1)},$$

zachodzącej dla każdej dodatniej liczby całkowitej n mamy równość w \mathbb{Q}^∞ postaci

$$\left(\frac{1}{1}, \frac{1}{2}, \frac{1}{3}, \dots\right) - \left(\frac{1}{2}, \frac{1}{3}, \frac{1}{4}, \dots\right) = \left(\frac{1}{2}, \frac{1}{6}, \frac{1}{12}, \dots\right).$$

Definicja 6.1.6: Przestrzeń liniowa funkcji ze zbioru X do ciała K

Niech $F(X, K)$ będzie zbiorem wszystkich funkcji z danego niepustego zbioru X do ciała K . Dla $f, g \in F(X, K)$ i dla $a \in K$ funkcje $f + g$ oraz af określone są warunkami:

$$(f + g)(x) = f(x) + g(x), \quad (af)(x) = af(x).$$

Wektor zerowy przestrzeni liniowej $F(X, K)$ to funkcja stała równa 0.

Podajmy jeszcze jeden przykład o dużym znaczeniu w kombinatoryce.

Definicja 6.1.7: Przestrzeń podzbiorów zbioru niepustego X nad ciałem \mathbb{Z}_2

Niech X będzie zbiorem niepustym, zaś $P(X)$ – niech będzie zbiorem podzbiorów zbioru X . Na zbiorze $P(X)$ określmy strukturę przestrzeni liniowej nad ciałem \mathbb{Z}_2 .

Operacja Δ dodawania wektorów określona jest w sposób następujący dla dowolnych $A, B \in P(X)$ jako ich tzw. różnica symetryczna $A \Delta B = A \cup B \setminus (A \cap B)$.

Dla każdego $A \in P(X)$ definiujemy mnożenie wektora A przez skalar (jeden z dwóch w \mathbb{Z}_2):

- $0 \otimes A = \emptyset$ – zbiór pusty
- $1 \otimes A = A$.

Odnosimy kilka istotnych własności wynikających z aksjomatów przestrzeni liniowej.

Obserwacja 6.1. W każdej przestrzeni liniowej V nad ciałem K zachodzi:

- (a) dla każdego $\alpha \in V$ istnieje tylko jeden taki wektor $\delta \in V$, że $\alpha \oplus \delta = 0$. Wektor ten oznaczamy $-\alpha$ i nazywamy WEKTOREM PRZECIWNYM do α .
- (b) $0 \otimes \alpha = 0$, dla każdego $\alpha \in V$ oraz $a \otimes 0 = 0$, dla każdego $a \in K$.
- (c) Jeśli $\alpha \in V$ oraz $a \in K$, to $a \otimes \alpha = 0$, to $a = 0$ lub $\alpha = 0$.
- (d) $-\alpha = (-1) \otimes \alpha$, dla każdego $\alpha \in V$.

Dowód. Zaczniemy od (a). Jeśli $\alpha \oplus \delta_1 = 0$ oraz $\alpha \oplus \delta_2 = 0$, to:

$$\delta_1 = \delta_1 \oplus 0 = \delta_1 \oplus (\alpha \oplus \delta_2) = (\delta_1 + \alpha) + \delta_2 = 0 + \delta_2 = \delta_2.$$

Dowodzimy (b). W ciele K mamy $0 + 0 = 0$. Stąd $0 \otimes \alpha = (0 + 0) \otimes \alpha = 0 \otimes \alpha \oplus 0 \otimes \alpha$. Dodając do obu stron tej równości wektor $-0 \otimes \alpha$ otrzymujemy $0 = a \otimes \alpha$. Dowód $a \otimes 0 = 0$ jest analogiczny.

Uzasadnijmy teraz (c). Jeśli $a \neq 0$, to

$$\alpha = 1 \otimes \alpha = a^{-1}a \otimes \alpha = a^{-1} \otimes 0 = 0.$$

Dowód (d) pozostawiamy jako ćwiczenie. \square

Podobnie jak w przypadku ciał, podstawowym narzędziem do uzyskiwania kolejnych przykładów przestrzeni liniowych jest pojęcie podprzestrzeni liniowej.

Definicja 6.1.8: Podprzestrzeń przestrzeni liniowej

Niepusty podzbiór $W \subset V$ nazywamy PODPRZESTRZENIĄ PRZESTRZENI LINIOWEJ V jeśli dla każdych $\alpha, \beta \in W$ oraz każdego $a \in K$ zachodzi:

- (i) $\alpha + \beta \in W$,
- (ii) $a \cdot \alpha \in W$.

W każdej przestrzeni liniowej V podzbiór $\{0\}$, złożony tylko z wektora zerowego, jest jej podprzestrzenią. Nazywamy ją PODPRZESTRZENIĄ ZEROWĄ. Mówimy, że przestrzeń liniowa V jest PRZESTRZENIĄ ZEROWĄ, jeśli składa się tylko z wektora zerowego.

UWAGA: Podprzestrzeń przestrzeni liniowej jest przestrzenią liniową (z działaniami pochodząymi z V , w tym z odziedziczonym wektorem zerowym). **Wektor zerowy należy do każdej podprzestrzeni!**

Przejdzmy do kluczowego przykładu podprzestrzeni, który już poznaliśmy:

Uwaga 6.1.9

Rozpatrzmy jednorodny układ równań liniowych o współczynnikach w ciele K .

$$U : \begin{cases} a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n = 0 \\ \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \dots + a_{mn}x_n = 0. \end{cases} \quad (*)$$

Zbiór wszystkich rozwiązań układu U jest podprzestrzenią przestrzeni liniowej K^n .

Dowód. Niech $W \subseteq K^n$ będzie zbiorem rozwiązań układu (*). Niech $(s_1, \dots, s_n), (r_1, \dots, r_n) \in W$. Należy pokazać, że do W należą także wektory:

$$(s_1, \dots, s_n) + (r_1, \dots, r_n) = (s_1 + r_1, \dots, s_n + r_n)$$

oraz, że dla każdego $a \in K$ do zbioru W należą również wektory:

$$a \cdot (r_1, r_2, \dots, r_n) = (ar_1, ar_2, \dots, ar_n).$$

Wystarczy sprawdzić, że wektory te spełniają każde równanie układu (*). Rzeczywiście, dla każdego $1 \leq i \leq m$ mamy $a_{i1}s_1 + a_{i2}s_2 + \dots + a_{in}s_n = 0$ oraz $a_{i1}r_1 + a_{i2}r_2 + \dots + a_{in}r_n = 0$, a zatem

$$a_{i1}(s_1 + r_1) + a_{i2}(s_2 + r_2) + \dots + a_{in}(s_n + r_n) = 0.$$

Widzimy więc, że $(s_1, \dots, s_n) + (r_1, \dots, r_n) \in W$. Podobnie, dla każdego $a \in K$:

$$a \cdot (a_{i1}s_1 + a_{i2}s_2 + \dots + a_{in}s_n) = a_{i1}as_1 + a_{i2}as_2 + \dots + a_{in}as_n = 0.$$

A zatem $a \cdot (s_1, s_2, \dots, s_n) \in W$, co oznacza, że W jest podprzestrzenią K^n . \square

Przykład. Rozwiązaniami układu jednorodnego o współczynnikach w \mathbb{R} :

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

jest zbiór wektorów $(0, -s - t, s, t)$, gdzie $s, t \in \mathbb{R}$. Każdy element tego zbioru można zapisać w postaci:

$$s(0, -1, 1, 0) + t(0, -1, 0, 1)$$

gdzie $s, t \in \mathbb{R}$. Tu jest fundamentalna nowość, która jest sercem „geometrycznej strony” naszych rozważań – zbiór rozwiązań możemy teraz zapisywać jako zbiór sum wektorów z odpowiedniej przestrzeni liniowej.

Uwaga 6.1.10

Dla każdej liczby naturalnej m niech $K_{\leq m}[x]$ oznacza zbiór wszystkich wielomianów stopnia co najwyżej m w $K[x]$. Jest to podprzestrzeń $K[x]$.

Interpretując zbiór wielomianów $K[x]$ jako podzbiór K^∞ złożony z ciągów mających jedynie skończenie niezeroowych wyrazów możemy zauważyc, że $K[x]$ można w istocie traktować jako podprzestrzeń K^∞ . Oczywiście przykładów podprzestrzeni K^∞ można wskazać więcej.

Na ćwiczeniach omawiać Państwo będą, w ramach rozmaitych przykładów, szereg podprzestrzeni w przestrzeni \mathbb{R}^∞ oraz $F(\mathbb{R}, \mathbb{R})$, mających związki z analizą. Warto zwrócić uwagę na kilka z nich.

Uwaga 6.1.11

W przestrzeni ciągów \mathbb{R}^∞ wskazać można bardzo wiele podprzestrzeni, np.:

- ciągi mające skończenie wiele niezeroowych wyrazów,
- ciągi ograniczone,
- ciągi zbieżne,
- ciągi $(x_i)_{i=1}^\infty$ spełniające $\sum_{i=1}^\infty x_i^2 < \infty$.
- ciągi $(x_i)_{i=1}^\infty$ spełniające określone rekurencje liniowe, np. $x_{n+2} = x_{n+1} + x_n$.

Uwaga 6.1.12

Przykłady podprzestrzeni w przestrzeni funkcji $F(K, K)$:

- funkcje parzyste, spełniające równanie $f(x) = f(-x)$, dla $x \in K$,
- funkcje nieparzyste, spełniające równanie $f(x) = -f(-x)$, dla $x \in K$,
- nad \mathbb{R} (i nie tylko): funkcje ograniczone, monotoniczne itd.
- funkcje będące rozwiązaniami równania Cauchy'ego^a, tzn. dla każdych $x, y \in K$:

$$f(x + y) = f(x) + f(y),$$

^aTo słynne równanie funkcyjne rozważane dla funkcji rzeczywistych badane było przez wielkich matematyków, jak Cauchy, Darboux, d'Alembert i inni. Przy niewielu dodatkowych założeniach można pokazać, że jego rozwiązaniami są jedynie funkcje postaci $f(x) = ax$, dla $a \in \mathbb{R}$. Do tych „drobnych” dodatkowych założeń należą: ciągłość (Cauchy, 1821), ciągłość w punkcie (Darboux, 1875), monotoniczność lub ograniczoność na dowolnym przedziale (Darboux, 1880). W 1905 roku Georg Hamel pokazał, używając aksjomatu wyboru, że bez przyjęcia tego typu założeń o regularności wskazać można znacznie bardziej skomplikowane i egzotyczne funkcje spełniające równanie Cauchy'ego.

6.2 Wybór przykładowych pytań

Proszę się nie uczyć na pamięć odpowiedzi czy wyjaśnienie! Takich pytań można ułożyć znacznie więcej i służą one jedynie sprawdzeniu podstawowego zrozumienia treści zajęć. Jeśli odpowiedź na dowolne z poniższych pytań sprawia Państwu problem, proszę pomyśleć który fragment notatek dotyczy tego pytania, wrócić do ich lektury, a w razie dalszych trudności poprosić prowadzącego o wskazówkę.

1. Wyznacz taki wektor $\alpha \in \mathbb{R}^4$, że spełniony jest warunek:

$$(3, 1, 2, 1) + 3\alpha = (3, 2, 4, 0).$$

2. Wyznacz taki wektor $\alpha \in \mathbb{Z}_7^4$, że spełniony jest warunek:

$$(3, 1, 2, 1) + 3\alpha = (3, 2, 4, 0).$$

3. Wyznacz taki wektor $\alpha \in \mathbb{Q}[x]$, że spełniony jest warunek:

$$(x^3 + 2x^2 + x + 3) + 3\alpha = 4x^2 + 2x + 3.$$

4. Wyznacz taki wektor $\alpha \in M_{2 \times 2}(\mathbb{R})$, że spełniony jest warunek:

$$\begin{bmatrix} 3 & 1 \\ 2 & 1 \end{bmatrix} + 3\alpha = \begin{bmatrix} 3 & 2 \\ 4 & 0 \end{bmatrix}.$$

5. Rozstrzygnij, czy istnieje taki skalar $a \in \mathbb{C}$, że w przestrzeni liniowej \mathbb{C}^3 zachodzi równość

$$a(2 - 3i, 5 + 4i, -6 + 7i) = (12 - 5i, 7 + 22i, -32 - 9i)?$$

6. Wypisz wszystkie elementy zbioru wektorów

$$\{s \cdot (1, 2, 3, 4) \in \mathbb{Z}_5^4 : s \in \mathbb{Z}_5\}.$$

7. Wypisz wszystkie elementy zbioru wektorów

$$\{s \cdot (1, 2, 0, 1) + t \cdot (0, 1, 2, 1) \in \mathbb{Z}_3^4 : s, t \in \mathbb{Z}_3\}.$$

8. Wskaż takie wektory $\alpha, \beta \in \mathbb{R}^4$, że zbiór wektorów

$$\{(s + t, s - 2t, s, t) \in \mathbb{R}^4 : s, t \in \mathbb{R}\}$$

równy jest zbiorowi

$$\{s \cdot \alpha + t \cdot \beta \in \mathbb{R}^4 : s, t \in \mathbb{R}\}.$$

9. Czy przestrzeń \mathbb{Z}_3^{100} ma skończenie wiele elementów?

10. Niech K będzie ciałem \mathbb{Z}_p . Ile wektorów ma przestrzeń K^n , a ile przestrzeń $M_{m \times n}(K)$?

11. Czy zbiór pusty jest przestrzenią liniową?

12. Czy podprzestrzeń jest przestrzenią liniową?

13. Niech U będzie podprzestrzenią przestrzeni liniowej V . Niech W będzie takim podzbiorem V , że $W \cap U = \emptyset$. Czy W może być podprzestrzenią przestrzeni V ?

14. Niech U, W będą podprzestrzeniami przestrzeni liniowej V . Czy zbiór $U \cup W$ jest podprzestrzenią? Czy $U \cap W$ jest podprzestrzenią?

15. Czy \mathbb{C} jest przestrzenią liniową nad ciałem liczb rzeczywistych?

16. Czy zbiór rozwiązań równania $x_1 + x_2 = 1$ jest podprzestrzenią przestrzeni liniowej \mathbb{R}^2 ?

17. Czy podzbiór $\{(x_1, x_2) \in \mathbb{R}^2 : x_1 \geq 0, x_2 \geq 0\}$ jest podprzestrzenią przestrzeni liniowej \mathbb{R}^2 ?

18. Czy zbiór wielomianów stopnia 2 z dodatkowym wielomianem zerowym jest podprzestrzenią przestrzeni liniowej wielomianów $\mathbb{R}[x]$?

19. Czy zbiór funkcji $f : \mathbb{R} \rightarrow \mathbb{R}$ o wszystkich wartościach dodatnich jest podprzestrzenią przestrzeni liniowej $F(\mathbb{R}, \mathbb{R})$?

20. Czy niezerowa podprzestrzeń przestrzeni \mathbb{R}^3 może zawierać skończenie wiele elementów?

21. Ile podprzestrzeni zawiera przestrzeń liniowa \mathbb{C}^1 ?

6.3 Zadania do samodzielnej pracy

Zadania ilustrujące umiejętności niezbędne do zaliczenia przedmiotu oznaczone są symbolem ♠.

Przy tych zadaniach dodany jest opis umiejętności, które sprawdzają.

1. (♠ Sprawdzanie czy dana trójka spełnia aksjomaty przestrzeni liniowej)
 - (a) Niech V będzie przestrzenią liniową nad \mathbb{C} . Uzasadnij, że V z tym samym dodawaniem a mnożeniem przez skalary określonym jako $a \cdot \alpha = \bar{a}\alpha$ jest też przestrzenią liniową nad \mathbb{C} .
 - (b) Niech $X = \{x \in \mathbb{R} : x > 0\}$. Zdefiniujmy dodawanie \oplus elementów zbioru X oraz mnożenie \odot elementów zbioru X przez liczby rzeczywiste wzorami $x \oplus y = xy$ oraz $\lambda \odot x = x^\lambda$. Wykaż, że trójka (X, \oplus, \odot) jest, przy odpowiednim wyborze wektora zerowego, przestrzenią liniową nad ciałem \mathbb{R} .
 - (c) Uzasadnij, że zbiór postaci $V_1 \times V_2 = \{(\alpha_1, \alpha_2) \mid \alpha_1 \in V_1, \alpha_2 \in V_2\}$ z działaniami określonymi wzorami:

$$(\alpha_1, \alpha_2) + (\beta_1, \beta_2) = (\alpha_1 + \beta_1, \alpha_2 + \beta_2) \quad \text{oraz} \quad a(\alpha_1, \alpha_2) = (a\alpha_1, a\alpha_2)$$

jest przestrzenią liniową nad ciałem K .

2. (♠ Sprawdzanie czy podzbiór przestrzeni liniowej spełnia warunki z definicji podprzestrzeni)
Dla każdego z poniższych podzbiorów \mathbb{R}^2 sprawdź, czy spełnia on warunek (i) oraz czy spełnia on warunek (ii) z definicji podprzestrzeni.

- (a) $\{(x_1, x_2) : x_1, x_1 \in \mathbb{Z}\}$,
- (b) $\{(x_1, x_2) : x_1 = 0 \text{ lub } x_2 = 0\}$,
- (c) $\{(x_1, x_2) : |x_1| - |x_2| = 1\}$,
- (d) $\{(x_1, x_2) : x_1^2 + x_2^2 = 2x_1x_2\}$.

3. Rozważmy przestrzeń liniową $V = \mathbb{R}^n$ i niech W będzie podzbiorem V składającym się z wektorów (x_1, \dots, x_n) , takich że

- (a) $x_n = 0$,
- (b) $x_1 + \dots + x_n = 1$,
- (c) $x_1 + \dots + x_n = 0$,
- (d) $x_1 + \dots + x_n \geq 0$,
- (e) $x_i = x_{n+1-i}$, dla $i = 1, 2, \dots, n$.

W którym z powyższych przypadków W jest podprzestrzenią V ?

4. Dla jakich wartości $s \in \mathbb{R}$ następujący podzbiór $W \subseteq \mathbb{R}^4$ jest podprzestrzenią?

$$W = \{(x_1, x_2, x_3, x_4) \in \mathbb{R}^4 : x_1 + x_2 + 2x_3 + 3x_4 = s^2 - s - 2, x_1 + s^2 x_3^2 + x_4 = x_3^2\}$$

5. Rozważmy podzbiór \mathcal{S} przestrzeni liniowej $F(\mathbb{R}, \mathbb{R})$, złożony z takich funkcji $f(x)$, dla których istnieją $a, b \in \mathbb{R}$, że $f(x) = a \sin(x + b)$. Czy jest to podprzestrzeń przestrzeni liniowej $F(\mathbb{R}, \mathbb{R})$?

6. Niech V będzie przestrzenią liniową nad ciałem \mathbb{R} złożoną z ciągów o wyrazach rzeczywistych, czyli: $V = \{(a_i)_{i \in \mathbb{N}} : a_i \in \mathbb{R}\}$. Niech $W = \{(a_i) : \forall_{i \in \mathbb{N}} a_{i+1} \leq a_i\} \subset V$. Czy W jest podprzestrzenią przestrzeni liniowej V ?

7. Niech W_1, W_2 będą podprzestrzeniami przestrzeni liniowej V . Wykaż, że zbiór $W_1 \cup W_2$ jest podprzestrzenią przestrzeni V wtedy i tylko wtedy, gdy $W_1 \subset W_2$ lub $W_2 \subset W_1$.

8. Rozstrzygnij, czy dla zbioru \mathbb{Z} ze zwykłym działaniem dodawania liczb całkowitych istnieje takie ciało K oraz takie działanie mnożenia liczby całkowitej przez skalar z ciała K , tak by \mathbb{Z} stało się przestrzenią liniową nad K ?

9. Niech $V = \{z \in \mathbb{C} : |z| = 1\}$. Zdefiniujmy działanie dodawania \oplus w zbiorze V wzorem $z_1 \oplus z_2 = z_1 z_2$. Wykaż, że jeżeli K jest (dowolnym) ciałem, to nie istnieje takie mnożenie \odot elementów zbioru V przez skalary z K , aby trójka (V, \oplus, \odot) była przestrzenią liniową nad K .

6.4 Trivia. Kody samokorekcyjne.

Nie sposób opisać wszystkich zastosowań przestrzeni liniowych. W naszym wykładzie zajmować się będziemy w dużej mierze przestrzenią współrzędnych K^n , jej podprzestrzeniami, rozmaitymi opisami tych podprzestrzeni itd. Warto przekonać się od razu, że te podstawowe i bardzo elementarne przestrzenie mają ważne zastosowania. Opowieść przedstawiona poniżej mówi o przykładach tzw. kodów liniowych.

W 1948 roku Claude Shannon, amerykański inżynier i matematyk, wydał artykuł „A Mathematical Theory of Communication”, który uważany jest za początek tzw. teorii informacji i teorii kodowania. Podstawowym celem jest efektywne i wiarygodne przesyłanie komunikatów w niekooperacyjnym (być może wrogim) środowisku. Aby być efektywne – komunikaty nie mogą wymagać nadawania przez zbyt długi czas lub zbyt duży koszt. Aby transmisja była wiarygodna potrzebne jest by otrzymywany sygnał przypominał ten wyemitowany, przynajmniej w ramach pewnej z góry określonej tolerancji. Wysiłki matematyków poszły w dwóch kierunkach. Shannon, ojciec teorii informacji, studiował osiągalne ograniczenia komunikacyjne głównie metodami analitycznymi i probabilistycznymi. Jego kolega – Richard Hamming, pracował nad poprawianiem kodów pierwszych komputerów i stosował głównie metody algebraiczne.

Informacja nadana ze źródła trafia do „przewodu”, „przestrzeni” „kanału”, którym podróżuje do odbiorcy. Nasz model komunikacji oparty jest o założenie, że informacja poddana jest zgodnie z naszą wolą pewnej strukturze u źródła oraz pewnej metodzie odczytu u odbiorcy, ale nie mamy żadnej kontroli nad przestrzenią pomiędzy nadawcą, a odbiorcą. W ten sposób wiadomość ulec może zniekształceniu. Prostym przykładem jest rozmowa w bardzo głośnej kawiarni, pisanie książki, która ma być odczytana lata później. Jest też wiele sposobów radzenia sobie z możliwymi zaburzeniami przekazu. Osobę, której nie dosłyszałem mogę poprosić o powtórzenie, a w przypadku znalezienia zniszczonego manuskryptu mogę próbować poszukiwać innej jego kopii. Tu jednak zaburzone są: efektywność (*Ile razy mam powtarzać?!*) i wiarygodność (*może nie ma innego manuskryptu, a może obydwa są fałszywe?*).



Zakłady Bell Telephone Laboratories w latach 50-tych XX wieku

Shannon i Hamming, a także wielu innych ojców teorii komunikacji, pracowali dla Bell Telephone Laboratories. Byli szczególnie zainteresowani radzeniem sobie z błędami, które powstają gdy wiadomość podróżuje kablem telefonicznym i zostanie zniekształcona przez uderzenie pioruna lub przez nałożenie się na siebie dwóch rozmów. Komunikacja w przestrzeni kosmicznej zaburzana jest przez atmosferę ziemską i aktywność słoneczną. Podczas misji Galileo, gdy padła jedna z anten sondy, naukowcy przeprogramowali komputer pokładowy sondy tak, by w sposób bardziej intensywny przetwarzał kod wysyłany na Ziemię i w ten sposób byli w stanie odzyskać część pierwotnej efektywności przekazu wiadomości. Dyski twarde naszych komputerów wyposażone są w CRC, czyli *Cyclic Redundancy Check*, z uwagi na konieczność wykrywania zaburzeń w przechowywaniu danych wystawionych na działanie promieni gamma czy interferencji magnetycznej. Gdy Phillips wprowadził technologię płyt CD reklamował ją jako niewrażliwą na wiele typów zniszczenia – nawet z porysowanej (nieznacznie) płyty jesteśmy (byliśmy?) w stanie odczytywać informacje. Jest to zasługa teorii kodowania. Można podać wiele więcej przykładów.

Informację można zapisać na wiele sposobów. Używamy w tym celu najczęściej słów zbudowanych z liter określonego alfabetu. W informatyce najczęściej są to bity, a więc ciągi zer i jedynek. **Kodowanie wiadomości polega na dodaniu do niej pewnego dodatkowego zestawu bitów** służącego do jej odczytania w sytuacji, gdy wiemy, że wystąpić może błąd. Można tego dokonywać na wiele sposobów.

Załóżmy, że chcecie Państwo przesłać Komuś wiadomość złożoną z trzech liter ze zbioru $\{0, 1\}$ postaci $v = abc$. Między emiterem a odbiornikiem wiadomość może ulec zniekształceniu i dojdzie do Kogoś nie-właściwe słowo. Czy ów Ktoś zdola wykryć taki błąd i odczytać poprawną wiadomość, jeśli wiemy na przykład, że błąd zwykle nie dotyczy więcej niż jednej litery?

Do opisu rozwiązania zastosujemy algebrę liniową. W tym celu zakłada się, że zakodowana wiadomość, którą przesyłamy, jest podprzestrzenią przestrzeni liniowej. Kodem liniowym długości n nad ciałem F nazywamy podprzestrzeń przestrzeni F^n . Zakodowane słowa to wektory.

Najpierw naiwne rozwiązanie problemu. Dla każdego 0 w planowanej wiadomości, wysyłamy dwa zera. Podobnie dla jedynek. A zatem jeśli oryginalna wiadomość miała na przykład postać $[010]$, to zakodujemy ją jako $[00 : 11 : 00]$. A zatem nasz kod to element przestrzeni $(\mathbb{Z}_2)^6$. Czy możemy traktować go jako podprzestrzeń? Zauważmy, że $s(0, 0, 1, 1, 0, 0)$, $s \in \mathbb{Z}_2$ to po prostu $\{(0, 0, 0, 0, 0), (0, 0, 1, 1, 0, 0)\}$, a więc niezerowe elementy tej podprzestrzeni tworzą zakodowaną wiadomość. Odbiorca jest w stanie stwierdzić czy jest ona poprawna, a więc czy wiadomość przyszła z jednym błędem (nie rozważamy tutaj dla uproszczenia innych sytuacji). Dla przykładu: jeśli wiadomość jest postaci $[00 : 11 : 10]$ odbiorca wie, że jest błąd w trzecim segmencie wiadomości. Może zatem wydedukować, że oryginalna wiadomość miała postaci $[010]$ lub $[011]$. Naiwne podejście ma dwie wady: po pierwsze przesyłamy dwa razy więcej danych, niż potrzeba, a po drugie – odbiorca nie ma dość informacji, by poprawić błąd, który wykrył.

Naiwne rozwiązanie problemu niemożności naprawienia (pojedynczego) błędu w transmisji jest proste: wysyłać trzy razy więcej danych. A więc na przykład oryginalną wiadomość postaci $[010]$ przesłać możemy jako $[000 : 111 : 000]$. Jeśli odbiorca otrzyma, powiedzmy, wiadomość postaci $[000 : 111 : 010]$ to wie już nie tylko, że błąd wystąpił w trzecim segmencie ale też, że w oryginalnej wiadomości ten segment miał postać 000. Widzimy jednak, że nie jest to efektywne przesyłanie danych. Oto inna propozycja, pochodząca od Hamminga.

Jeśli chcemy wysłać wiadomość postaci $[c_1 : c_2 : c_3]$, gdzie $c_1, c_2, c_3 \in \{0, 1\}$, to wysyłamy ciąg złożony z pięciu znaków postaci: $[c_1 : c_2 : c_3 : c_1 + c_2 : c_2 + c_3]$, przy czym operacje dodawania wykonujemy nad ciałem \mathbb{Z}_2 , czyli $c_1 + c_2$ jest równe 0 lub 1 w zależności od składników c_1, c_2 . Okazuje się, że w tym kodowaniu jesteśmy w stanie wykryć nawet dwa błędy, a jeśli jest tylko jeden – to możemy go naprawić. Zobaczmy przykłady.

Wysyłamy $[100]$, a więc po zakodowaniu dostajemy słowo $[10010]$. Założmy, że wystąpi dokładnie jeden błąd przy transmisji i otrzymamy jedną z wiadomości: $[00010]$, $[01010]$, $[10110]$, $[10000]$, $[10011]$. Czy Czytelnik widzi, że w każdym wypadku możemy nie tylko wykryć błąd, ale i go naprawić? W pierwszym przypadku $[00010]$ nie spełnia na czwartej współrzędnej warunku $c_1 + c_2 = 1$, ale spełnia na piątej warunek $c_2 + c_3 = 0$. A zatem skoro jest dokładnie jeden błąd, to c_2, c_3 są przesłane dobrze, a błędny jest przekaz c_1 . Oczywiście umiemy też wykryć wiadomość poprawnie odebraną.

Czy wykrywanie pojedynczego błędu w ogóle może kogoś interesować? Nie tylko może, ale jest powszechnie. Nie ma dwóch numerów kont, które różniłyby się tylko jedną lub dwiema cyframi. Jeśli wysyłając przelew pomylimy się o jedną lub dwie cyfry w numerze konta, to przelew zostanie odrzucony. Kod Hamminga stosuje się dla wiadomości dowolnej długości. Do zakodowania słowa długości n potrzeba $2n - 1$ znaków (oczywiście chodzi o słowo zerojedynkowe).

Być może Czytelnik nie dostrzega jeszcze żadnej wielkiej „matematyki” w tej opowieści, ale zapewniam, że dzieje się tak tylko dlatego, że niemal zmuszam się do unikania wprowadzania jakiejkolwiek terminologii, a dzieje się tu bardzo dużo. Mówiąc o kodach wspomnialiśmy zaraz o odległości Hamminga, problemie pakowania sfer, macierzach generujących, wielomianach kodujących słowa itd. Zainteresowanych odsyłam do bardzo ciekawych notatek J. Halla z teorii kodowania (polecam zwłaszcza wstępny rozdział – kolejne mogą być za trudne na razie – tylko na razie) dostępnych pod adresem:

<https://users.math.msu.edu/users/jhall/classes/CODENOTES/CODING-NOTES.HTML>

Kto by chciał poczytać (w języku polskim) więcej o kodach, szyfrach i ogólnie o teorii informacji, czy też przekonać się wszechstronnym występowaniem kodowania, np. w numerach PESEL, ISBN, IBAN, polecam tekst dr. Grzegorza Szkibielu „Wstęp do teorii informacji i kodowania”, dostępny online.

6.5 Coda. O kształtowaniu się pojęcia wektora

Pojęcie wektora kształtało się w nauce przez stulecia¹ i proces ten miał istotny wpływ na jej współczesny język. Nie chodzi jedynie o matematykę, ale też astronomię, fizykę, chemię, informatykę, ekonomię czy nauki techniczne. Słowo *wektor* pochodzi od łacińskiego *vexus*, znaczącego dosłownie „przewóz”.

Historycznie rzecz biorąc intuicje wektorowe związane były najpierw przede wszystkim z reprezentacją sił działających na obiekt za pomocą skierowanych odcinków oraz z obserwacją, że składanie tych sił spełnia tzw. prawo równoległoboku. Idee te wysłowił bezpośrednio już Arystoteles w czasach antycznych, w dziele *Questiones Mechanicae*. Dzieło to znali autorzy renesansowi, łącznie z Galileuszem, nie zawsze doceniając znaczenie samej reguły, a nawet nie uwzględniając jej wcale w swoich badaniach². Dyskuję w kierunku wysłownienia tej reguły rozpoczętą dopiero siedemnastowieczni autorzy tacy jak Fermat, Hobbes czy Mersenne, głównie w oparciu o próbę zrozumienia praw optyki (odbiicia i załamania) Kartezjusza. Należy jednak pamiętać, że uczeni ci nie określali pojęcia wektora. Formułowali jedynie pewne obserwacje w języku geometrii. Podejście w zasadzie „istotowo wektorowe” stosuje dopiero Newton w *Philosophiae naturalis principia mathematica* (1687), gdzie prawo równoległoboku jest sformułowane wprost, wciąż jednak jednak bez użycia wektorów a jedynie w oparciu o geometrię Euklidesa³.

Samo pojęcie wektora stosowane było najpierw w astronomii, w kontekście, w jakim dziś rozumiemy pojęcie *wektora wodzącego* (mówiąc mało precyzyjnie chodzi o wektor o ustalonym początku i końcu poruszającym się według pewnych zasad, na przykład po okręgu, elipsie itd.) i pojawiło się po raz pierwszy w 1704 roku. Pojęcie to (rayon vecteur) stosują również⁴ Laplace w swoim *Traktacie o mechanice niebieskiej* (1798) oraz Andre Ampère w *Traktacie o matematycznej teorii zjawisk elektrodynamicznych* z roku 1826. Wcześniej pojęcia tego używał też de la Lande w słynnej *Wielkiej encyklopedii francuskiej* (1776).

W geometrii pojęcie wektora pojawiło się wraz z geometryczną interpretacją pojęcia liczby zespolonej, proponowaną już przez Wessela (1797) i Arganda (1806). Podejście to poznaliśmy w ujęciu zaproponowanym przez Hamiltona, jest jednak pewne, że już Gauss posługiwał się nim swobodnie na początku XIX-tego wieku. W starych podręcznikach (np. autorstwa Webera z 1925 roku) znajdziemy zresztą informację, że reprezentacja wektorowa liczb zespolonych pochodzi od Gaussa. Również w geometrii używano pojęcia *rayon vecteur*: robił to zarówno Möbius (1827) w swoim rachunku barycentrycznym, jak i Cauchy, we wstępie do ważnego traktatu *Leçons sur les applications de calcul infinitesimal* (1826).

Do około 1830 roku liczby zespolone były w zasadzie reprezentowane jako wektory, choć nazewnictwo to wprowadził Hamilton i w to w kontekście swojego największego odkrycia — kwaternionów. Jednym z istotnych problemów matematyki początku XIX-tego stulecia było przeniesienie teorii liczb zespolonych w trzy wymiary, tak by na trójkach postaci $a + bi + cj$, gdzie i, j są pierwiastkami z -1 , określić mnożenie, mające sens geometryczny i porządne własności algebraiczne (łączność, przemienność, rozdzielność itd.).

W 1837 roku Hamilton publikuje długą i niezwykle ważną pracę interpretującą liczby zespolone jako uporządkowane pary liczb rzeczywistych, wprowadzając znaną nam zasadę mnożenia owych par. Jednocześnie rozpoczyna poszukiwania „teorii trójek”, wspomnianej już wyżej. Jak się okazało jest to głęboki problem, który doprowadza w 1843 roku do odkrycia kwaternionów, czyli liczb zapisywanych w postaci

$$a + xi + yj + zk,$$

gdzie $a, x, y, z \in \mathbb{R}$ oraz gdzie spełnione są następujące reguły:

$$ij = k, \quad jk = i, \quad ki = j, \quad ji = -k, \quad kj = -i, \quad ik = -j, \quad i^2 = j^2 = k^2 = -1.$$

Co ciekawe, dodawanie (po współrzędnych) i mnożenie kwaternionów są łączne (Hamilton pokazał to w roku 1844, po raz pierwszy używając tego terminu) oraz rozdzielne, dzielenie przez niezerowy element jest zawsze możliwe, dodawanie jest przemienne. Mnożenie kwaternionów jest jednak nieprzemienne.

¹ Skrócone opisy historii pojęć matematycznych znaleźć można w *Earliest Uses of Some Words of Mathematics* w ramach serwisu MacTutor History of Mathematics Archive: <https://mathshistory.st-andrews.ac.uk/Miller/mathword/>.

² David Marshall Miller, *The Parallelogram Rule from Pseudo-Aristotle to Newton*, Archive for History of Exact Sciences 71 (2017), 157-191, <https://www.jstor.org/stable/24913240>.

³ Sformułowanie w języku angielskim: <https://archive.org/details/100878576/page/84/mode/2up?view=theater>

⁴ Gregory H. Moore, *The axiomatization of linear algebra*, Historia Mathematica 22 (1995), 262-303, <https://www.sciencedirect.com/science/article/pii/S0315086085710257>.

Odkrycie kwaternionów miało znaczenie nie tylko naukowe, ale też filozoficzne. Połowa XIX-tego wieku to czasy, gdy Bolyai i Łobaczewski odkrywali pierwsze modele geometrii nieeuklidesowych. Hamilton pokazał, że możliwe jest stworzenie systemu algebraicznego, który łamie jedną z podstawowych reguł (przemienność mnożenia) mając sens także w kontekście naturalnych operacji geometrycznych. Takich systemów powstało więcej (najpierw rozważano tzw. liczby hiperzespolone, a później tzw. algebry).

W 1846 roku Hamilton publikuje pracę, w której wprowadza pojęcie *skalara* i *wektora*, mając na myśli „rzeczywistą” i „urojoną” część swoich kwaternionów. To znaczy: częścią skalarną kwaternionu $a + bi + cj + dk$ jest a , zaś częścią wektorową: $bi + cj + dk$. Jak się okazało, prawa mnożenia kwaternionów o części skalarnej zero miały ważne znaczenie w geometrii. W fizyce, fundamentalne znaczenie miało odkrycie przez Hamiltona tzw. operatora nabla, uczytelnione jeszcze przez Taita do postaci

$$\nabla = i \frac{d}{dx} + j \frac{d}{dy} + k \frac{d}{dz}.$$

W istocie, pierwotne sformułowanie słynnych równań elektrodynamiki klasycznej Maxwella z 1865 roku miało postać kwaternionową⁵. To równania Maxwella i próba ich uczytelnienia doprowadziły do nowoczesnego ujęcia pojęcia wektora i stało się to w pierwszej kolejności za sprawą fizyków.

W 1881 roku J.W. Gibbs opublikował pierwszą połowę *Elementów analizy wektorowej*, gdzie wektor opisywany jest zarówno za pomocą pojedynczego symbolu, jak i za pomocą współrzędnych. Gibbs sformułował, niezależnie od Grassmanna, o którym powiemy niżej, podstawowe działania na wektorach, a także pojęcia iloczynu skalarnego i iloczynu wektorowego. Prace Gibbasa (oraz Taita) rozwinął następnie Heaviside, dodając do niej wyniki Grassmanna. Widzimy zatem, że istotne motywacje do badania wektorów i operacji na nich nie pochodziły w XIX-tym wieku z matematyki, ale właśnie z fizyki. Podobnie było na początku wieku XX-tego, gdy sformułowana została teoria względności Einsteina.

Wracając do matematyki, można powiedzieć, że pod koniec wieku XIX-tego i na początku XX-tego pojęcie wektora traktowano wciąż geometrycznie – jako odcinek skierowany \vec{AB} lub jako formalną różnicę punktów $B - A$. Fizycy traktowali wektor jako obiekt (np. moment lub siłę) mający kierunek i długość. Zakładano też, że wektory mają nie więcej niż trzy współrzędne. Istniał już jednak od ponad 50 lat grunt pod ogólne podejście, pochodzący od mało znanego ówczesnym matematykom uczonego — Grassmanna.

Za twórcę pojęcia przestrzeni liniowej⁶, uważa się właśnie Hermanna Gunthera Grassmanna, urodzonego w Szczecinie w roku 1809. Był jednym z dwanaściorga dzieci. Sam wziął ślub po czterdziestym roku życia, mając ich siedmioro. Trzy lata życia Grassmann spędził w Berlinie, studując teologię i filologię. Nie miał żadnego wykształcenia matematycznego i nigdy nie zajmował pozycji na uniwersytecie (choć starał się o nie wielokrotnie). Życie spędził jako nauczyciel gimnazjalny. Zmarł w roku 1877, nigdy nie otrzymał uznanie jako twórca jakiekolwiek teorii matematycznej. Jego prace matematyczne odkryli później.

Rozważania Grassmanna zawierały idee przekraczające epokę, w której żył. Zanim przejdziemy do algebry liniowej warto wspomnieć, że w dziedzinie arytmetyki, już w 1861 roku Grassmann zdefiniował operacje arytmetyczne w zbiorze liczb naturalnych za pomocą pojęcia indukcji i dowód takie ich własności, jak przemienność, łączność, rozdzielność. Zdołał więc przewidzieć założenia teorii Peano czy Dedekinda opublikowanych niemal 30 lat później. Autorzy ci są wyraźnie zainspirowani wpływem Grassmanna, choć nie nie pomijają trudności w lekturze filozoficznego (miejscami nieprecyzyjnego) języka jego prac.

Najważniejszym dziełem Grassmanna była *Ausdehnungslehre* z 1844 roku, czyli Teoria Rozszerzeń, w zasadzie niezauważona aż do czasu publikacji dzieł zebranych Grassmanna pomiędzy rokiem 1894 oraz 1911, i to mimo tego, że autor przesyłał swoje prace między innymi do Möbiusa, Gaussa, Kummera, Cauchy'ego. Z ostatnim łączył go zresztą wieloletni spór o pierwszeństwo wyników, nierostrzygnięty przed Francuską Akademią Nauk. Grassmann zdefiniował pojęcie kombinacji liniowej, podprzestrzeni, liniowej niezależności, podprzestrzeni rozpinającej, wymiaru (w tym sumy i przecięcia podprzestrzeni) oraz rzutu na podprzestrzenie. Otrzymał również wzory na zamianę współrzędnych przy zmianie bazy w postaci iloczynu operacji elementarnych (fakty te będziemy stopniowo poznawać). Wprowadził również pojęcia, które dały początek iloczynowi zewnętrzemu oraz iloczynowi skalarnemu.

⁵ On the Notation of Maxwell's Field Equations, http://www.zpenergy.com/downloads/Orig_maxwell_equations.pdf.

⁶Tekst na podstawie artykułów D. Fearnley-Sander, *Hermann Grassmann and the Creation of Linear Algebra*, The American Mathematical Monthly , Vol. 86, No. 10 (1979), str. 809-817 oraz W. Wiesław, *Drogi i manowce początków algebry*, Szkoła Matematyki Poglądowej, <https://smp.uph.edu.pl/msn/15/16-26.pdf>.

Prace Grassmanna nie zostały na początku zauważone. Stosunkowo niewielką zmianę wniosły prace Peano, który w 1887 roku zaczął rozważać n -tki (wektory o n współrzędnych) wraz z operacjami dodawania i mnożenia przez skalar. To właśnie Peano, inspirowany pracami Grassmanna, wprowadził pojęcie systemu liniowego (obecnie przestrzeni liniowej) za pomocą aksjomatów (w swoim trzecim podejściu do tego tematu) w 1898 roku. Należy jednak podkreślić, że nawet Peano pisał o wektorach jedynie w kontekście geometrycznym. Robił to mimo tego, że to właśnie on był autorem dowodu istnienia rozwiązania układu n liniowych równań różniczkowych pierwszego rzędu o n zmiennych. Oczywiście Peano zajmował się przestrzeniami liniowymi nad \mathbb{R} , nie znając ogólnej teorii ciał. Jednym z głównych nowych pomysłów Peano było zrozumienie, że wielomiany jednej zmiennej rzeczywistej, a także wielomiany ograniczonego stopnia, tworzą przestrzeń liniową. Właściwe ujęcie prac Grassmana znajdą dopiero wielcy geometrzy różniczkowie początku XX wieku, przede wszystkim Henri Cartan oraz Henri Poincaré.

Alternatywne podejście do aksjomatyzacji pojęcia wektora zaproponował Gaston Darboux. W 1875 roku opublikował pracę analizującą różne dowody prawa składania sił statycznych (np. prawo równolegoboku), rozpoczynając od dowodu Daniela Bernoulliego z 1726 roku. Celem Darboux było uzyskanie uzasadnienia tych prac zawartych wewnętrz geometrii i ustalenia jakie założenia wymagane są do tego, by prawa te zachodziły. Zaproponował cztery takie aksjomaty (których w tym miejscu nie wysławiamy). Prace Darboux podjęli młodzi matematycy Schimmack i Hamel, którzy badali między innymi formalną zależność tych aksjomatów. Jak się okazało, niezależność czwartego aksjomatu Darboux od trzech wcześniejszych wymagała znalezienia nieciągłej funkcji rzeczywistej f , spełniającej dla dowolnych $x, y \in \mathbb{R}$ równanie funkcyjne Cauchy'ego $f(x+y) = f(x) + f(y)$. Hamel znalazł przykład takiej funkcji i jego praca doktoratu opublikowana została przez samego Hilberta w *Mathematische Annalen* w 1905 roku.

Odkrycia Hamały miały fundamentalne znaczenie dla teorii zbiorów, bowiem wymagały nowego wówczas rezultatu Zermelo mówiącego, że każdy zbiór można dobrze uporządkować. Wyniki Hamały, o których wspomnimy w komentarzach do kolejnych wykładów, wymagały skonstruowania bazy przestrzeni liczb rzeczywistych traktowanych jako przestrzenie liniowa nad ciałem liczb wymiernych. Istnienie takiej bazy wymaga aksjomatu wyboru, o czym powiemy w dodatku do kolejnego wykładu. To, co jest na ten moment istotne w podejściu Hamały, to zauważenie, że same liczby rzeczywiste traktować można jak wektory nad ciałem skalarów \mathbb{Q} . Co więcej, problem stwierdzania czy liczba rzeczywista jest skończoną kombinacją liniową (o współczynnikach w \mathbb{Q}) innych liczb rzeczywistych ma głębokie zastosowania.

Pod koniec Pierwszej Wojny Światowej sytuacja przestrzeni liniowych była następująca — ogólne pojęcie przestrzeni liniowej nad \mathbb{R} znane było, ale nie powszechnie, we Włoszech, wśród spadkobierców Peano. Jako jeszcze mniej znane pojęcie, przestrzenie te znano we Francji i Niemczech za sprawą prac Darboux, a potem Schimmacka i Hamały. Pojęcie to doprowadziło do powstania „baz Hamały” badanych intensywnie w kontekście analizy i teorii zbiorów (np. przez Waclawa Sierpińskiego). Kluczowym momentem dla przyjęcia przestrzeni liniowych jako pełnoprawnych obiektów matematycznych były prace Hahna, Banacha i Wienera, związane z tzw. unormowanymi przestrzeniami liniowymi. Na nasz użytku powiedzmy, że chodzi o takie przestrzenie liniowe, gdzie można za pomocą pewnej funkcji (zwanej normą) wprowadzić odległość. Pojęcie normy wektora poznamy w drugim semestrze (w ograniczonym kontekście).

W 1922 roku wiedeński matematyki Hans Hahn sformułował pojęcie unormowanej przestrzeni liniowej i przedstawił 21 przykładów przestrzeni tego typu. Wszystkie one były przestrzeniami funkcji, co miało przełomowe znaczenie. Przykładem były badane już wcześniej przez Schura przestrzenie ciągów nieskończonych (i ich przekształceń). Hahn zajmował się też układami równań liniowych w tych przestrzeniach.

Niezależnie od Hahna, pojęcie przestrzeni unormowanej wprowadził w 1922 roku Stefan Banach⁷. Prace Banacha były o tyle przełomowe, że wprowadzały na dobre metodę aksjomatyczną do analizy. Przestrzenie Banacha określone były najpierw za pomocą 13 aksjomatów, określających w istocie rzeczywistą przestrzeń liniową. Banach cytował, w ramach przykładów, Grassmanna, prace Hamiltona, teorie wektorów Peano itd. Druga grupa aksjomatów dotyczyła normy, a trzecia — pojęcia zupełności. Prace Banacha wywołyły pozytywne reakcje wielkich matematyków, między innymi Norberta Wienera i Maurice'a Frecheta, a pojęcie przestrzeni liniowej trafiło na Międzynarodowy Kongres Matematyków.

W międzyczasie do gry wkroczyło największe nazwisko matematyki początku XX-tego wieku — Dawida Hilberta. W roku 1904 Hilbert opublikował pracę dotyczącą liniowych równań całkowych, gdzie badano między innymi ciągi liczb rzeczywistych, których (nieskończona) suma kwadratów była skończona. Po-

⁷Chodzi o tzw. zupełne unormowane przestrzenie liniowe, zwane przestrzeniami Banacha.

dejście geometryczne do tych badań zaproponowali między innymi Schmidt (1908) oraz Riesz (1913), badający między innymi układy równań liniowych o nieskończonym wielu zmiennych. Badając strukturę rozwiązań tych układów, Riesz wprowadził pojęcie przestrzeni Hilberta. W 1927 roku pojęcie to zostało sformułowane aksjomatycznie przez von Neumanna, w celu zbudowania matematycznych mechaniki kwantowej Heisenberga i Schrödingera. Już wcześniej, pojęcie przestrzeni liniowej aksjomatyzowało dla potrzeb zbudowania matematycznej teorii względności Hermann Weyl (1918). Pojęcie wektora i przestrzeni liniowej było zatem dobrze umotywowane przez teorie fizyczne. Czy istniało jakieś algebraiczne źródło?

Algebraiczne źródło pojęcia przestrzeni liniowej wywodzi się z prac grupy niemieckich matematków, Dirichleta, Kummera, Kroneckera, Dedekinda i Webera, związanych z próbą dowodu Wielkiego Twierdzenia Fermata. Dedekind wprowadził pojęcie ideału, czyli podzbioru A w zbiorze B (np. w \mathbb{C}) zamkniętego na dodawanie, odejmowanie i mnożenie przez element z B . Dedekind sformułował również pojęcie modułu, mając na myśli podzbiór M zbioru \mathbb{C} zamknięty na dodawanie i odejmowanie. Dedekind wprowadził notację $a \equiv b \pmod{M}$ mając na myśli $a - b \in M$. Pojęcie to w latach 70-tych XIX-tego wieku było bardzo ogólne — obejmowało bowiem ideały Dedekinda, a naśladowało przy tym teorię kongruencji Gaussa, uogólniając jednak relację przystawania modulo z pojedynczej liczby całkowitej do całego zbioru.

Dedekind zorientował się, że istnieje związek pomiędzy badanymi przez niego liczbami algebraicznymi Ω , np. postaci $a + b\sqrt{2} + c\sqrt{-3} + d\sqrt{5}$, gdzie $a, b, c, d \in \mathbb{Q}$, a pojęciem „baz” i „liniowej niezależności”. Z obecnej perspektywy można rozumieć, że Dedekind umiał pokazać, że Ω jest przestrzenią skończonego wymiaru nad \mathbb{Q} mimo, że pojęcie to jeszcze nie funkcjonowało. Prace Dedekinda dotyczyły też sytuacji, gdy współczynniki były całkowite, a nie wymierne, co wyprowadza nas z algebry liniowej w kierunku tzw. teorii pierścieni. Dedekind współpracował ściśle z Heinrichem Weberem, z którym rozszerzył w 1882 roku pojęcie modułu do kontekstu funkcyjnego, definiując obiekt nazywany dziś modułem nad pierścieniem wielomianów $\mathbb{C}[z]$ i badając takie moduły, mające skończoną „bazę”. Dekadę później, w 1892 roku, Weber zunifikował rozmaite pojęcia ciała (algebraiczne ciało liczbowe, algebraiczne ciało funkcyjne, ciało skończone) i sformułował abstrakcyjną definicję, znaną do dziś. Prace te podjął w 1910 roku Ernst Steinitz. Pojęcie modułu zwróciło uwagę wielkich algebraików, m.in. Hilberta i Noether, którzy użyli go do zbudowania podstaw teorii pierścieni oraz ich ideałów. Po 1945 roku rozważania te nabraly nowego kontekstu w świetle teorii kategorii.

Podejście algebraiczne zostało dojrzałe ukształtowane w przełomowym podręczniku *Moderne Algebra* van der Waerdena w latach 1930-1931. Po kilku latach książka ta trafiła z Niemiec do Ameryki, a nowoczesne podejście do algebry, uwzględniające przestrzenie liniowe, trafiło do najsłynniejszego przedwojennego podręcznika algebry — *Przeglądu algebry współczesnej* Birkhoffa i Mac Lane'a (1941). W nauczaniu akademickim spopularyzował je ważny podręcznik Mirsky'ego z 1955 roku⁸. W Polsce pojęcie przestrzeni liniowej upowszechnił w nauczaniu akademickim Profesor Andrzej Mostowski z Uniwersytetu Warszawskiego. Kolejne wydania podręcznika, zwłaszcza pisane wspólnie z Marcelim Starkiem, były w zasadzie podstawą wykładu akademickiego przez niemal pół wieku. Po czasie dołączyła do nich wspaniała *Algebra liniowa z geometrią* Profesora Andrzeja Białynickiego-Biruli, która była podstawą do opracowania obecnego programu nauczania tego przedmiotu na naszym Wydziale. Warto tu przywołać fragment recenzji tej ostatniej pozycji z Wiadomości Matematycznych (1976), autorstwa Profesora Narkiewicza

Nowy program studiów matematycznych zlikwidował wykładaną tradycyjnie na I roku geometrię analityczną, łącząc ten przedmiot z algebrą liniową. Recenzowana książka jest pierwszym podręcznikiem powstającego w ten sposób przedmiotu, dopasowanym ściśle do wymogów programowych. W istocie swej jest to podręcznik algebry liniowej w klasycznym ujęciu, z dodaniem elementów teorii przestrzeni afiničnych i przekształceń afiničnych. Czytelnik, przyzwyczajony do tradycyjnej geometrii analitycznej, nie znajdzie jej tu wcale. Jedynie jej ślad przewija się tu i ówdzie w zadaniach. Taki jest los tej archaicznej dyscypliny, zdaniem recenzenta, w pełni zasłużony.

Podsumowując, widzimy jak skomplikowane są dzieje pojęć matematycznych. Powyższy tekst przedstawia i tak jedynie wierzchołek góry lodowej, zarzucając Czytelnika nazwiskami wielkich matematyków, których poznawanie zajmie większość studiów. Warto jednak rozumieć, że uporządkowana i sterylnie wręcz wygładzająca teoria ma korzenie dotykające niemal każdej dziedziny matematyki — i nie tylko matematyki, ale także fizyki czy astronomii. Historia wektorów jest znacznie starsza niż historia teorii, która je opisuje.

⁸https://mathshistory.st-andrews.ac.uk/Extras/Mirsky_books/.

Rozdział 7

Kombinacje liniowe. Podprzestrzeń rozpięta na układzie wektorów

7.1 Wykład 7

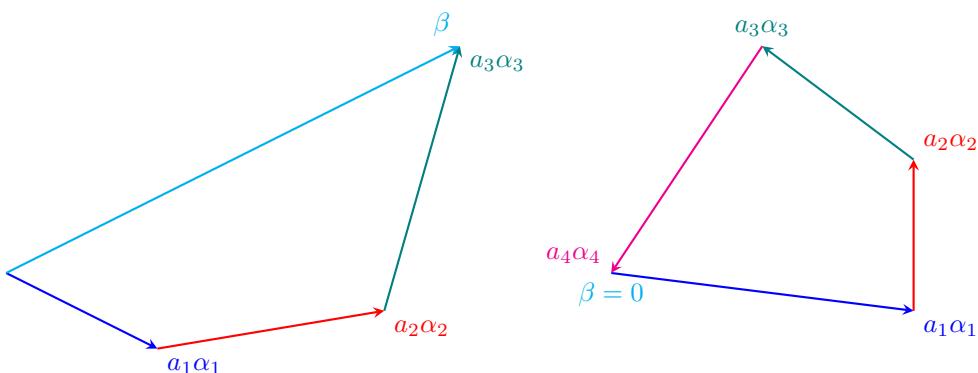
Powiemy teraz o bardzo ważnym typie konstrukcji związkanych z podprzestrzeniami. Chodzi o sytuację, gdy mamy przestrzeń liniową i szukamy takiej jej podprzestrzeni, która zawierałaby z góry określone przez nas wektory – wybranej przy tym możliwie oszczędnie.

Definicja 7.1.1: Kombinacja liniowa

Niech V będzie przestrzenią liniową nad ciałem K . KOMBINACJĄ LINIOWĄ układu wektorów $\alpha_1, \dots, \alpha_k$ o współczynnikach $a_1, \dots, a_k \in K$ nazywamy wektor:

$$\beta = a_1\alpha_1 + \dots + a_k\alpha_k = \sum_{i=1}^k a_i\alpha_i.$$

Na poziomie intuicji geometrycznej wektor β jest kombinacją liniową $\sum_{i=1}^k a_i\alpha_i$ układu wektorów, gdy złożenie przesunięć o wektory $a_1\alpha_1, a_2\alpha_2, \dots, a_n\alpha_n$ jest równe przesunięciu o wektor β . W szczególności, gdy $\beta = 0$ oznacza to, że złożenie tych przesunięć jest identycznością, czyli przesunięciem o wektor zerowy.



Przykłady.

- W przestrzeni $V = \mathbb{R}^4$ kombinacją liniową wektorów

$$(2, 1, -3, 4), (0, 2, 5, 1), (7, 4, 3, 2)$$

ze współczynnikami $2, -1, 1$ jest wektor

$$2 \cdot (2, 1, -3, 4) - 1 \cdot (0, 2, 5, 1) + 1 \cdot (7, 4, 3, 2) = (11, 4, -8, 9).$$

- W przestrzeni funkcji $F(\mathbb{R}, \mathbb{R})$ kombinacją liniową wektorów $\sin(x)$ oraz $\cos(x)$ o współczynnikach $\frac{1}{\sqrt{2}}$ oraz $-\frac{1}{\sqrt{2}}$ jest funkcja

$$\frac{1}{\sqrt{2}} \sin(x) - \frac{1}{\sqrt{2}} \cos(x) = \sin\left(x - \frac{\pi}{4}\right).$$

- Wektor $(0, 3, 1) \in \mathbb{R}^3$ nie jest kombinacją liniową wektorów $(0, 1, 1), (-1, 0, 1)$, bo założenie, że

$$(0, 3, 1) = a(0, 1, 1) + b(-1, 0, 1)$$

prowadzi do układu równań $0 = -b, 3 = a, 1 = a + b$, który nie ma rozwiązań.

- Wektory $(1, 1, -2), (1, 0, -1) \in \mathbb{R}^3$ są rozwiązaniami równania jednorodnego $x_1 + x_2 + x_3 = 0$, skąd wynika, że każda ich kombinacja liniowa $a(1, 1, -2) + b(1, 0, -1) = (a + b, a, -2a - b)$ jest również rozwiązaniem tego równania.

- Jeśli $\beta_1, \beta_2, \dots, \beta_r \in K^n$ są rozwiązaniami układu liniowych równań jednorodnych U , to również

$$a_1\beta_1 + a_2\beta_2 + \dots + a_r\beta_r$$

są rozwiązaniami tego układu, dla dowolnych układów skalarów $a_1, a_2, \dots, a_r \in K$.

Uwaga 7.1.2

Niech $\alpha_1, \dots, \alpha_k$ będą wektorami przestrzeni liniowej V nad K . Jeśli wektory β, γ są kombinacjami liniowymi wektorów $\alpha_1, \dots, \alpha_k$, to wektory $\beta + \gamma$ oraz $a\beta$, dla każdego $a \in K$, również są kombinacjami liniowymi wektorów $\alpha_1, \dots, \alpha_k$.

Dowód. Niech a_1, \dots, a_k oraz b_1, \dots, b_k będą elementami ciała K oraz niech

$$\beta = a_1\alpha_1 + \dots + a_k\alpha_k, \quad \gamma = b_1\alpha_1 + \dots + b_k\alpha_k.$$

Wówczas

$$\beta + \gamma = (a_1 + b_1)\alpha_1 + \dots + (a_k + b_k)\alpha_k$$

oraz dla każdego $a \in K$ mamy

$$a\beta = aa_1\alpha_1 + \dots + aa_k\alpha_k.$$

□

Definicja 7.1.3: Podprzestrzeń rozpięta na układzie wektorów

Niech V będzie przestrzenią liniową nad ciałem K i niech $\alpha_1, \dots, \alpha_k \in V$. Wówczas przez

$$\text{lin}(\alpha_1, \dots, \alpha_k)$$

oznaczamy zbiór wszystkich kombinacji liniowych wektorów $\alpha_1, \dots, \alpha_k$.

Poprzednią obserwację możemy teraz wyrazić w następujący sposób.

Uwaga 7.1.4

Zbiór $\text{lin}(\alpha_1, \dots, \alpha_k)$ jest podprzestrzenią przestrzeni V . Podprzestrzeń ta jest najmniejszą podprzestrzenią V (względem inkluzyji) zawierającą wektory $\alpha_1, \dots, \alpha_k$.

Dowód. Z Uwagi 7.1.2 wynika, że $\text{lin}(\alpha_1, \dots, \alpha_k)$ jest podprzestrzenią w V . Niech W będzie dowolną podprzestrzenią zawierającą wektory $\alpha_1, \dots, \alpha_k$. Z definicji podprzestrzeni W zawiera każdą kombinację liniową wektorów $\alpha_1, \dots, \alpha_k$, czyli każdy wektor z $\text{lin}(\alpha_1, \dots, \alpha_k)$. Stąd $\text{lin}(\alpha_1, \dots, \alpha_k) \subseteq W$. □

Definicja 7.1.5: Układ wektorów rozpinający podprzestrzeń

Niech $\alpha_1, \dots, \alpha_k$ będzie układem wektorów w V . Wówczas podprzestrzeń liniową $\text{lin}(\alpha_1, \dots, \alpha_k)$ nazywamy PRZESTRZENIĄ ROZPIĘTĄ NA UKŁADZIE $\alpha_1, \dots, \alpha_k$. Mówimy, że układ $\alpha_1, \dots, \alpha_k$ ROZPINIA PRZESTRZEŃ V , jeśli $V = \text{lin}(\alpha_1, \dots, \alpha_k)$, to znaczy każdy wektor z V jest kombinacją liniową wektorów $\alpha_1, \dots, \alpha_k$.

Rozważmy układ równań liniowych o współczynnikach w \mathbb{R} postaci:

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

możemy zauważyć, że zbiór rozwiązań tego układu jest podprzestrzenią \mathbb{R}^4 rozpiętą przez wektory $(0, -1, 1, 0), (0, -1, 0, 1)$, czyli jest to zbiór

$$\text{lin}((0, -1, 1, 0), (0, -1, 0, 1)).$$

Zauważmy, że zbiór rozwiązań powyższego układu można zapisać na wiele innych sposobów. Można взять np. rozwiązania $(0, 0, -1, 1)$ oraz $(0, -2, 1, 1)$ i zauważyc, że zbiór rozwiązań powyższego układu jest równy:

$$\text{lin}((0, 0, -1, 1), (0, -2, 1, 1)).$$

Co więcej, nic nie stoi na przeszkodzie, by rozważyć zbiór wszystkich kombinacji liniowych postaci:

$$s(0, 0, -1, 1) + t(0, -1, 1, 0) + r(0, -1, 0, 1), \quad s, t, r \in \mathbb{R}$$

i jest to również zbiór rozwiązań układu powyżej! Innymi słowy mamy równości:

$$\text{lin}((0, -1, 1, 0), (0, -1, 0, 1)) = \text{lin}((0, 0, -1, 1), (0, -2, 1, 1)) = \text{lin}((0, 0, -1, 1), (0, -1, 1, 0), (0, -1, 0, 1)).$$

Przykład. Rozważmy układ równań liniowych o współczynnikach w \mathbb{R}

$$\begin{cases} x_1 + -x_3 = 1 \\ 2x_1 + x_2 = 0 \\ 3x_1 + x_2 + x_3 = 0 \end{cases}$$

Zauważmy, że problem istnienia rozwiązań tego układu równań równoważny jest ze znalezieniem takich $s_1, s_2, s_3 \in \mathbb{R}$, że

$$s_1 \cdot \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} + s_2 \cdot \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix} + s_3 \cdot \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}.$$

Rozstrzygnięcie istnienia rozwiązań układu równań sprowadza się do sprawdzenia, czy pewna macierz rozmiaru 3×1 jest kombinacją liniową pewnych trzech macierzy ze współczynnikami s_1, s_2, s_3 . To pytanie można z kolei przeformułować do pytania dotyczącego kombinacji liniowych w przestrzeni \mathbb{R}^3 . Jest ono równoważne z rozstrzygnięciem, czy $(1, 0, 0) \in \text{lin}((1, 2, 3), (0, 1, 1), (-1, 0, 1))$.

Przy badaniu przestrzeni rozpiętych na układach wektorów w K^n użyteczna jest prosta obserwacja.

Uwaga 7.1.6

Niech $A, A' \in M_{m \times n}(K)$ oraz niech

- $\alpha_1, \dots, \alpha_m$ – wiersze macierzy A traktowane jako wektory w K^n ,
- $\alpha'_1, \dots, \alpha'_m$ – wiersze macierzy A' traktowane jako wektory w K^n .

Jeśli założymy, że A' może być otrzymywana z A za pomocą ciągu operacji elementarnych na wierszach, to wynika stąd, że

$$\text{lin}(\alpha_1, \dots, \alpha_m) = \text{lin}(\alpha'_1, \dots, \alpha'_m).$$

Zanim pokażemy dowód, przedstawmy przykład zaczerpnięty z przestrzeni \mathbb{R}^4 . Weźmy układ wektorów

$$(2, 1, 1, 1), (4, 2, 2, 0), (0, 0, 0, 1)$$

z przestrzeni \mathbb{R}^4 . Wektory te traktować możemy jako wiersze macierzy o czterech kolumnach:

$$\begin{bmatrix} 2 & 1 & 1 & 1 \\ 4 & 2 & 2 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

Po dodaniu dwukrotności trzeciego wiersza do drugiego wiersza, a następnie po przemnożeniu pierwszego wiersza przez 2 otrzymujemy macierz

$$\begin{bmatrix} 4 & 2 & 2 & 2 \\ 4 & 2 & 2 & 2 \\ 0 & 0 & 0 & 1 \end{bmatrix},$$

a zatem dostajemy równość:

$$\text{lin}((2, 1, 1, 1), (4, 2, 2, 0), (0, 0, 0, 1)) = \text{lin}((4, 2, 2, 2), (4, 2, 2, 2), (0, 0, 0, 1)).$$

Zauważmy teraz, że układ rozpinający $\text{lin}((2, 1, 1, 1), (4, 2, 2, 0), (0, 0, 0, 1))$ można pomniejszyć. Po odjęciu pierwszego wiersza od drugiego, a następnie po zamianie drugiego i trzeciego wiersza mamy:

$$\begin{bmatrix} 4 & 2 & 2 & 2 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix},$$

A zatem możemy napisać (i bez macierzy mogliśmy):

$$\text{lin}((2, 1, 1, 1), (4, 2, 2, 0), (0, 0, 0, 1)) = \text{lin}((4, 2, 2, 2), (0, 0, 0, 1)).$$

Czytelnik zechce zauważyć, że powyższa podprzestrzeń nie jest rozpięta przez jeden wektor. Dlaczego?

Dowód. Wystarczy pokazać tezę w przypadku, gdy A' powstaje z A przez wykonanie pojedynczej operacji elementarnej na wierszach. Wykażemy tezę jedynie w najtrudniejszym przypadku. Pokazujemy mianowicie, że dla dowolnych $1 \leq i, j \leq m$ oraz dowolnego $a \in K$ mamy:

$$\text{lin}(\alpha_1, \dots, \color{red}{\alpha_i}, \dots, \color{blue}{\alpha_j}, \dots, \alpha_m) = \text{lin}(\alpha_1, \dots, \alpha_i, \dots, a \cdot \color{red}{\alpha_i} + \color{blue}{\alpha_j}, \dots, \alpha_m).$$

Weźmy $\beta \in \text{lin}(\alpha_1, \dots, \color{red}{\alpha_i}, \dots, \color{blue}{\alpha_j}, \dots, \alpha_m)$. Istnieją $b_1, b_2, \dots, b_m \in K$, że:

$$\begin{aligned} \beta &= b_1\alpha_1 + b_2\alpha_2 + \dots + b_i\alpha_i + \dots + b_j\alpha_j + \dots + b_m\alpha_m = \\ &= b_1\alpha_1 + b_2\alpha_2 + \dots + (b_i - a \cdot b_j)\alpha_i + \dots + b_j(a\alpha_i + \alpha_j) + \dots + b_m\alpha_m. \end{aligned}$$

Zatem $\beta \in \text{lin}(\alpha_1, \dots, \alpha_i, \dots, a \cdot \color{red}{\alpha_i} + \color{blue}{\alpha_j}, \dots, \alpha_m)$.

Weźmy teraz $\gamma \in \text{lin}(\alpha_1, \dots, \alpha_i, \dots, a \cdot \color{red}{\alpha_i} + \color{blue}{\alpha_j}, \dots, \alpha_m)$. Istnieją $c_1, c_2, \dots, c_m \in K$, że:

$$\begin{aligned} \gamma &= c_1\alpha_1 + c_2\alpha_2 + \dots + c_i\alpha_i + \dots + c_j(a \cdot \alpha_i + \alpha_j) + \dots + c_m\alpha_m = \\ &= c_1\alpha_1 + c_2\alpha_2 + \dots + (c_i + a \cdot c_j)\alpha_i + \dots + c_j\alpha_j + \dots + c_m\alpha_m. \end{aligned}$$

Zatem $\gamma \in \text{lin}(\alpha_1, \dots, \color{red}{\alpha_i}, \dots, \color{blue}{\alpha_j}, \dots, \alpha_m)$. □

Pamiętajmy, że rozważać można przykłady podprzestrzeni rozpiętych na układach wektorów w innych przestrzeniach niż K^n , np. $K[x], K^\infty, M_{n \times m}(K)$ czy $F(K, K)$ i wtedy korzystać należy raczej z definicji wprowadzonych wcześniej, nie zaś z metody opisanej wyżej. Dla przykładu, w przestrzeni $\mathbb{R}[x]$ zachodzi równość:

$$\text{lin}(7x^2 + 4x - 3, 9x - 12) = \text{lin}(4x^2 + x, x^2 - 2x + 3).$$

Rzeczywiście, mamy $\text{lin}(7x^2 + 4x - 3, 9x - 12) \subseteq \text{lin}(4x^2 + x, x^2 - 2x + 3)$, ponieważ wektor $7x^2 + 4x - 3 = 2 \cdot (4x^2 + x) - 1 \cdot (x^2 - 2x + 3)$ należy do $\text{lin}(4x^2 + x, x^2 - 2x + 3)$, zaś wektor $9x - 12 = 1 \cdot (4x^2 + x) - 4 \cdot (x^2 - 2x + 3)$ należy do $\text{lin}(4x^2 + x, x^2 - 2x + 3)$. Każda kombinacja liniowa wektorów $7x^2 + 4x - 3$ oraz $9x - 12$ jest kombinacją liniową wektorów $4x^2 + x, x^2 - 2x + 3$, czyli należy do $\text{lin}(4x^2 + x, x^2 - 2x + 3)$.

Podobnie dowodzimy przeciwną inkluzyję: $\text{lin}(7x^2 + 4x - 3, 9x - 12) \supseteq \text{lin}(4x^2 + x, x^2 - 2x + 3)$, ponieważ wektor $4x^2 + x = 1 \cdot (9x - 12) + 4 \cdot (x^2 - 2x + 3)$ należy do $\text{lin}(7x^2 + 4x - 3, 9x - 12)$, a wektor $x^2 - 2x + 3 = 2 \cdot (4x^2 + x) - 1 \cdot (7x^2 + 4x - 3)$ należy do $\text{lin}(7x^2 + 4x - 3, 9x - 12)$. Stąd każda kombinacja liniowa wektorów $4x^2 + x$ oraz $x^2 - 2x + 3$ jest kombinacją liniową wektorów $7x^2 + 4x - 3, 9x - 12$, czyli należy do $\text{lin}(7x^2 + 4x - 3, 9x - 12)$.

Rozważmy jeszcze inny przykład. W przestrzeni \mathbb{R}^∞ rozważmy wszystkie ciągi zadane rekurencją:

$$x_{n+2} = x_{n+1} + x_n$$

Łatwo widzieć, że zbiór rozwiązań tej rekurencji tworzy podprzestrzeń w \mathbb{R}^∞ . Jeśli ciągi $a = (a_1, a_2, \dots)$ oraz $b = (b_1, b_2, \dots)$ są rozwiązaniami równania wyżej, to także ciągi $a + b$ oraz λa są w sposób oczywisty jej rozwiązaniami, dla każdego $\gamma \in \mathbb{R}$. Zauważmy, że dowolne dwa ciągi $(a_1, a_2), (b_1, b_2)$ rozpinające podprzestrzeń \mathbb{R}^2 wyznaczają układ ciągów a, b rozpinający przestrzeń rozwiązań powyższego równania.

Oto inny przykład. Rozważmy podzbiór S zbioru macierzy $M_{2 \times 2}(R)$ postaci zawierający elementy postaci:

$$\begin{bmatrix} a & b \\ b & a \end{bmatrix}.$$

Ponownie, nietrudno przekonać się, że zbiór S jest w istocie podprzestrzenią. Zauważmy też, że

$$\begin{bmatrix} a & b \\ b & a \end{bmatrix} = a \cdot \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} + b \cdot \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}.$$

Zatem

$$X = \text{lin}\left(\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}\right).$$

Pojęcie przestrzeni rozpiętej na układzie wektorów można rozszerzyć na układy niekoniecznie skończone.

Definicja 7.1.7

Niech $X = \{\alpha_t\}_{t \in T}$ będzie dowolnym układem wektorów przestrzeni V . Wówczas przez $\text{lin}(X)$ oznaczamy zbiór wszystkich kombinacji liniowych SKOŃCZONYCH PODUKŁADÓW układu X . To znaczy:

$$\beta \in \text{lin}(X) \iff \beta = \sum_{i=1}^k a_i \alpha_{t_i}, \text{ dla pewnych } a_1, \dots, a_k \in K, \alpha_{t_1}, \dots, \alpha_{t_k} \in X.$$

Jeśli $V = \text{lin}(X)$ to mówimy, że układ X ROZPINIA V i przestrzeń V JEST ROZPIĘTA na X .

Dla układu pustego $X = \emptyset$ przyjmujemy $\text{lin}(X) = \{0\}$.

Przykłady.

- Niech $V = K[x]$. Dla danej liczby naturalnej m niech X będzie układem w V złożonym ze wszystkich wielomianów postaci x^n dla $n \geq m$. Wówczas $\text{lin}(X)$ jest zbiorem wszystkich wielomianów w $K[x]$ podzielnych przez x^m .
- W przestrzeni K^∞ złożonej ze wszystkich ciągów nieskończonych o wyrazach z K rozpatrzmy układ $X = \{e_i\}_{i \in \mathbb{N}}$, gdzie e_i oznacza ciąg mający i -ty wyraz równy 1, a wszystkie pozostałe wyrazy równe 0. Wówczas $\text{lin}(X) = K_c^\infty =$ zbiór wszystkich ciągów prawie stale równych 0.

W sposób analogiczny do przypadku układu skończonego dowodzi się, że $\text{lin}(X)$ jest podprzestrzenią liniową przestrzeni V i że jest to najmniejsza podprzestrzeń w V zawierająca wszystkie wektory układu X .

Pojęcie przestrzeni liniowej to pierwszy krok w kierunku uzyskania nowej geometrycznej perspektywy na rozmaite obiekty matematyczne. Na kolejnym wykładzie zastanowimy się nad fundamentalnym problemem: ile elementów z przestrzeni liniowej rozpiętej przez n wektorów musimy znać, aby przestrzeń ta była wyznaczona jednoznacznie oraz jakie własności mają takie „minimalne układy rozpinające”.

7.2 Wybór przykładowych pytań

Proszę się nie uczyć na pamięć odpowiedzi czy wyjaśnienie! Takich pytań można ułożyć znacznie więcej i służą one jedynie sprawdzeniu podstawowego zrozumienia treści zajęć. Jeśli odpowiedź na dowolne z poniższych pytań sprawia Państwu problem, proszę pomyśleć który fragment notatek dotyczy tego pytania, wrócić do ich lektury, a w razie dalszych trudności poprosić prowadzącego o wskazówkę.

- Chcemy przedstawić wektor $(2, 1)$ jako kombinację liniową wektorów $(3, 2)$ oraz $(5, -1)$ w przestrzeni liniowej \mathbb{R}^2 . Jaki układ równań liniowych jest równoważny temu zadaniu?

- Dany jest układ równań liniowych o macierzy

$$\begin{bmatrix} a & b & c \\ d & e & f \end{bmatrix} \in M_{3 \times 2}(\mathbb{R}).$$

Wiadomo przy tym, że wektor (c, f) nie jest kombinacją liniową wektorów (a, d) oraz (b, e) w \mathbb{R}^2 . Czy powyższy układ równań liniowych ma rozwiązanie?

- Czy wektor $(1, 1, 1)$ należy do podprzestrzeni przestrzeni liniowej \mathbb{R}^3 rozpiętej przez wektory $(2, 1, 0)$ oraz $(0, 5, 5)$?

- Czy wektor $x + 2x^2$ należy do podprzestrzeni przestrzeni liniowej $R[x]$ rozpiętej przez wektory $1 + x$ oraz $1 + x^2$?

- Czy każdy element ciała \mathbb{C} traktowanego jako przestrzeń liniowa nad \mathbb{R} jest kombinacją liniową wektorów $1 + i$ oraz $1 - i$?

- Czy funkcja $\sin(x + \frac{\pi}{2})$ należy do podprzestrzeni przestrzeni liniowej $\mathbb{F}(\mathbb{R}, \mathbb{R})$ rozpiętej przez funkcje $\sin(x), \cos(x)$?

- Czy zachodzi należenie

$$\begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix} \in \text{lin}\left(\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}\right)?$$

- Czy ciąg stały o wyrazach rzeczywistych należy do dowolnej podprzestrzeni rozpiętej przez układ (niekoniecznie skończony) ciągów rosnących w przestrzeni liniowej \mathbb{R}^∞ ?

- Czy wektor zerowy może być kombinacją liniową wektorów niezerowych?

- Czy $\text{lin}(\alpha) = \text{lin}(a\alpha)$, dla dowolnego $a \in K$?

- Wiadomo, że $\text{lin}(\alpha) = \text{lin}(\beta)$. Czy $\alpha = \beta$?

- Niech $U = \text{lin}(\alpha_1, \alpha_2)$ oraz $V = \text{lin}(\beta)$. Czy podprzestrzeń U może być równa V ?

- Wiadomo, że $\text{lin}(\alpha_1, \alpha_2) = \text{lin}(\beta)$. Czy α_1 lub α_2 musi być wektorem zerowym?

- Wektory u, v, w spełniają warunek

$$u + w = 2(v + w).$$

Czy wektor w jest kombinacją liniową wektorów u, v ?

- Wiadomo, że α, β należą do podprzestrzeni liniowej U . Czy $\text{lin}(\alpha, \beta) \subset U$?

- Niech U będzie przestrzenią rozwiązań równania liniowego $x_1 + x_2 = 0$. Czy istnieje wektor $\alpha \in \mathbb{R}^2$ taki, że $U = \text{lin}(\alpha)$?

- Niech U będzie przestrzenią rozwiązań równania liniowego $x_1 + x_2 + x_3 = 0$. Czy istnieje wektor $\alpha \in \mathbb{R}^3$ taki, że $U = \text{lin}(\alpha)$?

- Niech

$$A = \{(x_1, x_2) \in \mathbb{R}^2 : x_1 x_2 = 0\}.$$

Czym jest $\text{lin}(A)$?

- Niech

$$A = \{(x_1, x_2) \in \mathbb{R}^2 : 1 \leq x_1 = x_2 \leq 2\}.$$

Czym jest $\text{lin}(A)$?

- Niech X będzie takim podzbiorem przestrzeni liniowej V , że $\text{lin}(X) = X$. Czy zbiór X jest podprzestrzenią V ? Przypuśćmy, że $\text{lin}(X) \neq V$. Czy jest możliwe, że zbiór $V \setminus \text{lin}(X)$ jest jednoelementowy?

7.3 Zadania do samodzielnej pracy

Zadania ilustrujące umiejętności niezbędne do zaliczenia przedmiotu oznaczone są symbolem ♠. Przy tych zadaniach dodany jest opis umiejętności, które sprawdzają.

1. (♠) Sprawdzanie czy wektor jest kombinacją liniową innych
 - (a) Czy wektor $(2, 1, 2, 1)$ należy do $\text{lin}((1, 2, 0, 2), (1, 0, 3, 1), (0, 1, 0, 2), (1, 1, 2, 0)) \subseteq \mathbb{R}^4$?
 - (b) Czy wektor $(1, 1, 1, 1)$ należy do $\text{lin}((1, 1, 0, 0), (0, 1, 1, 0), (0, 0, 1, 1)) \subset \mathbb{Z}_2^4$?
 - (c) Czy wielomian $x - x^3$ należy do $\text{lin}(x^2, 2x + x^2, x + x^3) \subset R[x]$?
 - (d) Czy macierz $\begin{bmatrix} -2 & -4 \\ -2 & -6 \end{bmatrix}$ należy do $\text{lin}\left(\begin{bmatrix} 1 & 2 \\ -1 & 3 \end{bmatrix}, \begin{bmatrix} -2 & 1 \\ 1 & -1 \end{bmatrix}\right) \subset M_{2 \times 2}(\mathbb{Q})$?
 - (e) Czy funkcja $\cos(3x)$ należy do $\text{lin}(1, \sin(x), \sin^2(x), \sin^3(x)) \subset F(\mathbb{R}, \mathbb{R})$?
2. (♠) Niech

$$\alpha_1 = (3, 2, 1, 1), \quad \alpha_2 = (2, 7, 2, 1), \quad \alpha_3 = (1, 3, 1, 3)$$
 oraz

$$\beta_1 = (2, -2, 0, 3), \quad \beta_2 = (1, 1, 1, 1), \quad \beta_3 = (-1, 3, 1, 10)$$
 będą wektorami przestrzeni \mathbb{R}^4 . Które z wektorów β_i są kombinacjami liniowymi układu $\alpha_1, \alpha_2, \alpha_3$?
3. (♠) Dla jakich wartości parametru $r \in \mathbb{R}$ wektor $(r, 8, 6) \in \mathbb{R}^3$ jest kombinacją liniową wektorów

$$(3, 4, 5), (1, 4, 4), (7, 4, 7)?$$
4. (♠) Rozpatrzmy macierze

$$A_1 = \begin{bmatrix} 1 & 3 & 1 \\ 2 & 5 & 3 \end{bmatrix}, \quad A_2 = \begin{bmatrix} 4 & 5 & 1 \\ 3 & 3 & 2 \end{bmatrix}, \quad B_1 = \begin{bmatrix} 1 & 2 & 1 \\ 2 & 1 & 3 \end{bmatrix}, \quad B_2 = \begin{bmatrix} 7 & 7 & 1 \\ 4 & 1 & 1 \end{bmatrix}.$$
 Niech $W \subset M_{m \times n}(\mathbb{R})$ będzie podprzestrzenią podprzestrzenią rozpiętą na A_1, A_2 . Które z macierzy B_1, B_2 należą do W ?
5. (♠) Podaj przykład wektora $\alpha \in \mathbb{R}^4$, który nie leży w podprzestrzeni

$$\text{lin}((1, 3, 1, 3), (3, 8, 2, 9), (0, 3, 2, 3)).$$
6. (♠) Opis zbioru rozwiązań jednorodnego układu równań liniowych jako podprzestrzeni K^n rozpiętej na układzie wektorów) Zapisz zbiór rozwiązań poniższego układu równań jako podprzestrzeń $\text{lin}(v_1, v_2, v_3)$ przestrzeni \mathbb{R}^5

$$\begin{cases} x_1 + 2x_2 - x_3 + x_4 = 0 \\ 2x_1 - x_2 - 2x_3 + x_4 + x_5 = 0 \\ -x_1 + x_3 + 4x_4 + x_5 = 0 \end{cases}.$$
7. Wykaż, że każdy wektor $(x_1, x_2, x_3, x_4) \in \mathbb{C}^4$ leżący w podprzestrzeni \mathbb{C}^4 rozpiętej przez wektory $\alpha_1 = (i, 1, -i, -1)$, $\alpha_2 = (i, -i, 1, -1)$ oraz $\alpha_3 = (1, 0, 0, -1)$ spełnia równanie $x_1 + x_2 + x_3 + x_4 = 0$, ale nie każdy spełnia równanie $x_4 = -1$.
8. Niech $W = \{f \in \mathbb{R}[x]_2 \mid f(1) = f(2) = 0\}$. Uzasadnij, że $W = \text{lin}(x^2 - 3x + 2)$.
9. Wykazać, że niepusty podzbiór W przestrzeni liniowej V jest jej podprzestrzenią wtedy i tylko wtedy, gdy każda kombinacja liniowa wektorów z W należy do W .
10. W przestrzeni liniowej V dane są wektory u, v, w . Rozstrzygnij, czy zachodzi równość

$$\text{lin}(u, v, w) = \text{lin}(u + v, v + w, w + u)?$$
11. Niech $\alpha_1, \dots, \alpha_n$ będzie układem wektorów w przestrzeni liniowej V . Dla $m = 1, \dots, n$ niech

$$\beta_m = \alpha_1 + \dots + \alpha_m.$$
 Wykaż, że $\text{lin}(\alpha_1, \dots, \alpha_n) = \text{lin}(\beta_1, \dots, \beta_n)$.
12. Niech V będzie przestrzenią liniową, zaś $A \subset V$ pewnym jej podzbiorem. Niech $\alpha, \beta \in V$. Przypuśćmy że $\alpha \notin \text{lin}(A)$, ale $\alpha \in \text{lin}(A \cup \{\beta\})$. Czy wynika stąd, że $\text{lin}(A \cup \{\beta\}) = \text{lin}(A \cup \{\alpha\})$?

7.4 Uzupełnienie. Kombinacje liniowe i układy równań

Jedną z przestrzeni liniowych poznanych na wykładzie jest przestrzeń macierzy o m wierszach i n kolumnach o wyrazach z ciała K . Nietrudno zauważać, że dodawanie macierzy lub mnożenie ich przez skalar są w zasadzie identyczne z operacjami wprowadzonymi w przestrzeni $K^{m \times n}$. Aby to unaoczyćć weźmy na przykład sumę macierzy w $M_{2 \times 3}(\mathbb{Q})$ oraz sumę wektorów w \mathbb{Q}^6 postaci:

$$\begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix} + \begin{bmatrix} 6 & 5 & 4 \\ 3 & 2 & 1 \end{bmatrix} = \begin{bmatrix} 7 & 7 & 7 \\ 7 & 7 & 7 \end{bmatrix}, \quad (1, 2, 3, 4, 5, 6) + (6, 5, 4, 3, 2, 1) = (7, 7, 7, 7, 7, 7).$$

Wkrótce poznamy język, który pozwoli nam powiedzieć, że z punktu widzenia „struktury” przestrzeni liniowych przestrzenie $M_{2 \times 3}(\mathbb{Q})$ oraz \mathbb{Q}^6 w zasadzie niczym się nie różnią – są IZOMORFICZNE. Dlaczego więc rozróżniamy te dwie przestrzenie? Macierze okazały się wygodnym narzędziem do badania układów równań. Jak niedługo zobaczymy, są one również wygodnym narzędziem do badania przekształceń pomiędzy przestrzeniami liniowymi. Jest jeden przypadek, gdy utożsamienie wektorów z macierzami wykonać można bez żadnych dodatkowych umów: gdy rozważamy macierze o jednym wierszu lub jednej kolumnie. Zajmiemy się teraz drugą sytuacją.

Zapiszmy równań liniowych nad \mathbb{R} za pomocą operacji w $M_{3 \times 1}(\mathbb{R})$:

$$\begin{cases} x - z = 0 \\ 2x + y = 0 \\ 3x + y + z = 0 \end{cases} \Rightarrow x \cdot \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} + y \cdot \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix} + z \cdot \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}.$$

Widzimy zatem, że rozwiązywanie układu równań sprowadza się do sprawdzenia, czy pewna macierz rozmiaru 3×1 jest kombinacją liniową pewnych trzech macierzy ze współczynnikami x, y, z . Jest jednak jasne, że w istocie jest to zagadnienie równoważne z przedstawieniem wektora $(0, 0, 0) \in \mathbb{R}^3$ jako kombinacji liniowej wektorów $(1, 2, 3), (0, 1, 1), (-1, 0, 1)$. Często mówimy nawet, że wektory te zapisane zostały w równaniu wyżej w notacji kolumnowej. A zatem w dalszym ciągu często dokonywać będziemy utożsamienia elementów K^n oraz przestrzeni macierzy $M_{1 \times n}(K)$ oraz $M_{n \times 1}(K)$ mówiąc przy tym, że wektor $v \in K^n$ zapisujemy w formie kolumnowej v^T lub wierszowej v .

Rozwiązywanie układów równań przez poszukiwanie kombinacji liniowych nie przyspieszy samego procesu rozwiązywania (dalej stosować będziemy metodę Gaussa), ale pozwoli nam zadać kilka istotnych pytań. Wróćmy do układu wyżej i zapytajmy: czy jeśli zamienimy wektor $(0, 0, 0)$ na dowolny inny, układ pozostanie niesprzeczny? A zatem: czy dowolny wektor $(a, b, c) \in \mathbb{R}^3$ jest kombinacją liniową wektorów $(1, 2, 3), (0, 1, 1), (-1, 0, 1)$? Zupełnie wprost: czy dla każdych a, b, c istnieją x, y, z takie, że

$$x \cdot \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} + y \cdot \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix} + z \cdot \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} a \\ b \\ c \end{bmatrix}?$$

Skąd mamy wiedzieć coś takiego i jak wyznaczyć x, y, z ? Okazuje się, że nie jest to trudne. W języku kombinacji liniowych nasze pytanie brzmi: czy dowolny wektor z \mathbb{R}^3 jest kombinacją liniową wektorów $(1, 2, 3), (0, 1, 1), (-1, 0, 1)$? W skrócie, pytamy o prawdziwość równości:

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

Czy to może być prawda? Nietrudno się przekonać, że tak jest: twierdzenie wykazane na wykładzie mówi, że wpisując powyższe trzy wektory w wiersze możemy wykonywać operacje wierszowe i przekonać się, że ciągiem operacji elementarnych na wierszach macierz

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

sprowadzić można do macierzy:

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

Jest natomiast jasne, że $\text{lin}((1, 0, 0), (0, 1, 0), (0, 0, 1)) = \mathbb{R}^3$, bo dla każdego $(a, b, c) \in \mathbb{R}^3$ mamy $(a, b, c) = a(1, 0, 0) + b(0, 1, 0) + c(0, 0, 1)$.

A zatem odpowiedzieliśmy na pytanie o rozwiązywalność dowolnego układu niejednorodnego o pewnej konkretnej macierzy współczynników. A jak wygląda rozwiązanie dla konkretnych a, b, c ? Zobaczmy nasz układ w jeszcze innej postaci:

$$x \cdot \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} + y \cdot \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix} + z \cdot \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix} = a \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} + b \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} + c \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

Przypomnijmy, że jeśli zacznimy wykonywać jednocześnie te same operacje na wierszach następujących macierzy:

$$\begin{bmatrix} 1 & 0 & -1 \\ 2 & 1 & 0 \\ 3 & 1 & 1 \end{bmatrix}, \quad \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix},$$

wówczas kombinacje liniowe ich kolumn ze współczynnikami x, y, z oraz a, b, c będą nadal równe! Zobaczmy to. Wykonajmy dwie operacje na obydwu macierzach:

$$\begin{bmatrix} 1 & 0 & -1 \\ 2 & 1 & 0 \\ 3 & 1 & 1 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 0 & -1 \\ 0 & 1 & 2 \\ 3 & 1 & 1 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 0 & -1 \\ 0 & 1 & 2 \\ 0 & 1 & 4 \end{bmatrix}, \quad \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ -3 & 0 & 1 \end{bmatrix}$$

Mamy:

$$x \cdot \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} + y \cdot \begin{bmatrix} 0 \\ 2 \\ 1 \end{bmatrix} + z \cdot \begin{bmatrix} -1 \\ 2 \\ 4 \end{bmatrix} = a \begin{bmatrix} 1 \\ -2 \\ -3 \end{bmatrix} + b \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} + c \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

Czy Czytelnik widzi, że kontynuując proces schodkowania macierzy wyjściowego układu równań dojdziemy w końcu do postaci pozwalającej wyznaczyć x, y, z za pomocą a, b, c ? Kontynuujmy eliminację, tym razem zapisując już macierze obok siebie:

$$\begin{array}{c|ccc} 1 & 0 & -1 & 1 & 0 & 0 \\ 0 & 1 & 2 & -2 & 1 & 0 \\ 0 & 1 & 4 & -3 & 0 & 1 \end{array} \rightarrow \begin{array}{c|ccc} 1 & 0 & -1 & 1 & 0 & 0 \\ 0 & 1 & 2 & -2 & 1 & 0 \\ 0 & 0 & 6 & -1 & -1 & 1 \end{array} \\ \rightarrow \begin{array}{c|ccc} 1 & 0 & -1 & 1 & 0 & 0 \\ 0 & 1 & 2 & -2 & 1 & 0 \\ 0 & 0 & 1 & -\frac{1}{6} & -\frac{1}{6} & \frac{1}{6} \end{array} \\ \rightarrow \begin{array}{c|ccc} 1 & 0 & 0 & \frac{5}{6} & -\frac{1}{6} & \frac{1}{6} \\ 0 & 1 & 0 & -\frac{10}{6} & \frac{4}{6} & \frac{2}{6} \\ 0 & 0 & 1 & -\frac{1}{6} & -\frac{1}{6} & \frac{1}{6} \end{array}$$

A zatem mamy:

$$x \cdot \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} + y \cdot \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} + z \cdot \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} = a \begin{bmatrix} \frac{5}{6} \\ -\frac{5}{3} \\ -\frac{1}{6} \end{bmatrix} + b \begin{bmatrix} -\frac{1}{6} \\ \frac{2}{3} \\ -\frac{1}{6} \end{bmatrix} + c \begin{bmatrix} \frac{1}{6} \\ \frac{1}{3} \\ \frac{1}{6} \end{bmatrix}$$

Po uproszczeniu:

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} \frac{5}{6}a - \frac{1}{6}b + \frac{1}{6}c \\ -\frac{5}{3}a + \frac{2}{3}b + \frac{1}{3}c \\ -\frac{1}{6}a - \frac{1}{6}b + \frac{1}{6}c \end{bmatrix}.$$

Wracając do wyjściowego problemu widzimy, że rozwiązaniem układu:

$$\begin{cases} x - z = a \\ 2x + y = b \\ 3x + y + z = c \end{cases}$$

jest trójka:

$$\left(\frac{5}{6}a - \frac{1}{6}b + \frac{1}{6}c, \quad -\frac{5}{3}a + \frac{2}{3}b + \frac{1}{3}c, \quad -\frac{1}{6}a - \frac{1}{6}b + \frac{1}{6}c \right).$$

Jeśli Czytelnik dotrwał do tego momentu, to gratuluję: odwróciliśmy właśnie pierwszą macierz. Nie wiemy na razie co to znaczy, ale sam termin „odwrócenia” powinien rodzić jasne skojarzenia. Rozpisaliśmy ustalony wektor jako kombinację liniową trzech zadanych z góry wektorów. Nie zawsze będzie to jednak możliwe. Proszę zauważać, że gdyby zamiast wektorów $(1, 2, 3), (0, 1, 1), (-1, 0, 1)$ szukać kombinacji liniowych wektorów: $(1, 2, 3), (2, 4, 6), (-1, 0, 1)$, to nie każdy wektor \mathbb{R}^3 byłby ich kombinacją liniową. Inaczej mówiąc $\text{lin}((1, 2, 3), (2, 4, 6), (-1, 0, 1)) \neq \mathbb{R}^3$. Sprawom tym przyjrzymy się już na następnym wykładzie.

7.5 Dodatek. Ciało jako przestrzeń liniowa nad podciałem

Pojęcie wielomianu o współczynnikach w ciele K oraz pojęcie pierwiastka wielomianu, pozwalają na istotne wzbogacenie naszego zasobu przykładów ciał. Rozważmy następującą sytuację. W zbiorze liczb rzeczywistych wybieramy wszystkie liczby postaci:

$$\{a + b\sqrt{2} \mid a, b \in \mathbb{Q}\}.$$

Zbiór ten oznaczamy jako $\mathbb{Q}(\sqrt{2})$. Czytelnik zechce zauważyć, że $a + b\sqrt{2} = c + d\sqrt{2}$ wtedy i tylko wtedy, gdy $a = c$ oraz $b = d$ (wynika to z niewymierności liczby $\sqrt{2}$). Co więcej, wprowadzenie w powyższym zbiorze działań dodawania i mnożenia liczb rzeczywistych prowadzi do zauważenia, że wniosku, że zbiór ten jest ciałem. Rzeczywiście, dla dowolnych $a + b\sqrt{2}$ oraz $c + d\sqrt{2}$ należących do $\mathbb{Q}(\sqrt{2})$ liczby

$$(a + b\sqrt{2}) + (c + d\sqrt{2}) = (a + c) + (b + d)\sqrt{2}, \quad (a + b\sqrt{2})(c + d\sqrt{2}) = (ac + 2bd) + (ad + bc)\sqrt{2}$$

są elementami $\mathbb{Q}(\sqrt{2})$. Odrobina wysiłku, między innymi zauważenie równości:

$$\frac{1}{a + b\sqrt{2}} = \frac{a - b\sqrt{2}}{a^2 - 2b^2}$$

pokazuje, że $\mathbb{Q}(\sqrt{2})$ jest ciałem. Jest to przykład tzw. podciała ciała \mathbb{R} . Przejdźmy do ogólnej definicji.

Definicja 7.5.1: Podciął i rozszerzenie ciał

Mówimy, że piątka $(K', +', \cdot', 0', 1')$ jest PODCIĄŁEM ciała $(K, +, \cdot, 0, 1)$ jeśli K' jest podzbiorem ciała K , $0' = 0$, $1' = 1$ oraz działania $+$ i \cdot powstają przez ograniczenie działań $+$, \cdot określonych na $K \times K$ do zbioru $K' \times K'$.

Parę $K' \subset K$, gdzie K' jest podciąłem ciała K nazywamy ROZSZERZENIEM CIAŁ.

Najbardziej znanym podciąłem ciała liczb rzeczywistych jest ciało liczb wymiernych \mathbb{Q} ze zwykłymi działaniami dodawania, mnożenia oraz z wyróżnionymi elementami 0 i 1. Innym przykładem podciąża jest $\mathbb{Q}(\sqrt{2}) = \{a + b\sqrt{2} \mid a, b \in \mathbb{Q}\}$. Jest to tak zwane rozszerzenie kwadratowe ciała \mathbb{Q} o $\sqrt{2}$. Aby zdefiniować czym jest owo rozszerzenie powiedzmy kilka słów o podciążach ustalonego ciała.

Uwaga 7.5.2

Rozważmy dowolną rodzinę podciąża K_t ciała L , gdzie $t \in T$. Wówczas część wspólna wszystkich ciał K_t jest podciąłem ciała K .

Uwaga 7.5.3

Dla każdego podciąża K ciała L oraz podzbioru S zbioru L istnieje najmniejsze podciąłko $K(S)$ ciała L , które zawiera jednocześnie ciało K oraz zbiór S .

Oto przykłady podciąża ciała liczb rzeczywistych, utworzone w oparciu o powyższe obserwacje:

- ciała $\mathbb{Q}(\sqrt{p})$, gdzie p jest liczbą pierwszą,
- ciała $\mathbb{Q}(\sqrt{p}, \sqrt{q})$, gdzie p, q są liczbami pierwszymi,
- najmniejsze ciało zawierające \mathbb{Q} i pierwiastki z wszystkich liczb pierwszych,
- ciała typu $\mathbb{Q}(\pi)$, $\mathbb{Q}(\sqrt{2}, \pi)$, $\mathbb{Q}(\pi, \pi^2, \pi^3, \dots)$ itd.

Jak ująć powyższe zagadnienia w języku przestrzeni liniowych? Otóż należy zdać sobie sprawę, że jeśli ciało K jest podciąłem ciała L , to ciało L traktować można jako przestrzeń liniową nad ciałem K . Zobaczmy kilka przykładów.

- Ciało \mathbb{C} jest przestrzenią liniową nad ciałem \mathbb{R} . A zatem liczby zespolone traktujemy jako wektory, które mnożymy jedynie przez liczby rzeczywiste (zapominamy o tym, że liczby zespolone można też mnożyć). Zauważmy, że biorąc w tym ujęciu wektory $1, i \in \mathbb{C}$ widzimy, że każda liczba zespolona jest kombinacją liniową wektorów α oraz β , czyli

$$z = a \cdot 1 + b \cdot i, \quad a, b \in \mathbb{R}$$

Innymi słowy: $\mathbb{C} = \text{lin}(1, i) = \mathbb{R}(i)$.

- Ciało $\mathbb{Q}(\sqrt{2})$ jest przestrzenią liniową nad ciałem \mathbb{Q} . Bardzo podobnie jak wyżej widzimy, że $\mathbb{Q}(\sqrt{2}) = \text{lin}(1, \sqrt{2})$, przy czym teraz skalarami są liczby wymierne, a wektorami — liczby postaci

$$a \cdot 1 + b \cdot \sqrt{2}, \quad a, b \in \mathbb{Q}.$$

- Ciało czteroelementowe wprowadzone w uzupełnieniu do wykładu pierwszego zawiera ciało \mathbb{Z}_2 jako podciało. Co więcej, ciało to jest w istocie postaci $\mathbb{Z}_2(\zeta)$, gdzie ζ jest pierwiastkiem wielomianu $x^2 + x + 1 \in \mathbb{Z}_2[x]$.
- Zupełnie innym przykładem jest ciało \mathbb{Q} traktowane jako podciało ciała \mathbb{R} . Można liczby rzeczywiste traktować jako wektory, a liczby wymierne jako skalary. Nie jest jednak możliwe wskazanie takiego skończonego (ani nawet przeliczalnego — to zrozumieją Państwo na wstępnie do matematyki) układu wektorów r_1, r_2, \dots, r_n takiego, by \mathbb{R} było równe $\text{lin}(r_1, r_2, \dots, r_n)$. Jak się okazuje wiąże się to z tym, że istnieją liczby rzeczywiste, które nie są pierwiastkami wielomianów rzeczywistych.

Definicja 7.5.4: Rozszerzenie algebraiczne

Niech $K \subset L$ będą ciałami. Powiemy, że element $a \in L$ jest algebraiczny nad ciałem K , jeśli istnieje wielomian $f \in K[x]$ taki, że $f(a) = 0$. Jeśli element $a \in L$ nie jest ALGEBRAICZNY nad ciałem K , wówczas element ten nazywamy PRZESTĘPNYM nad ciałem K .

Powiemy, że para ta jest ROZSZERZENIEM ALGEBRAICZNYM CIAŁ, jeśli każdy element ciała L jest pierwiastkiem pewnego wielomianu o współczynnikach ciała K . Rozszerzenie $K \subset L$ nazywamy PRZESTĘPNYM, jeśli nie jest ono algebraiczne.

Zobaczmy kilka przykładów.

- Liczby rzeczywiste $\sqrt{2}, i, \sqrt[3]{3}, \sqrt{1+\sqrt{2}}$ są algebraiczne nad \mathbb{Q} , bowiem są pierwiastkami wielomianów wymiernych $x^2 - 2, x^2 + 1, x^3 - 3, x^4 - 2x^2 - 1$.
- Liczby rzeczywiste π, e są przestępne nad \mathbb{Q} (choć dowód nie jest łatwy).
- Liczba π jest algebraiczna nad \mathbb{R} — jest pierwiastkiem wielomianu $x - \pi$.

Rozważania dotyczące rozszerzeń algebraicznych leżą u podstaw teorii równań, a także teorii liczb. W jednym z kolejnych dodatków poznacie Państwo dowód następującego podstawowego rezultatu.

Twierdzenie 7.5.5

Niech $K \subseteq L$ będzie rozszerzeniem ciał. Niech $\alpha \in L$ będzie elementem algebraicznym nad K i niech $f \in K[x]$ będzie wielomianem nierozkładalnym stopnia n takim, że $f(\alpha) = 0$. Wówczas ciało $K(\alpha)$ jest przestrzenią liniową nad ciałem K oraz

$$K(\alpha) = \text{lin}(1, \alpha, \alpha^2, \dots, \alpha^{n-1}).$$

Innymi słowy każdy element ciała $K(\alpha)$ jest postaci:

$$\beta = b_0 + b_1\alpha + \dots + b_{n-1}\alpha^{n-1}, \quad \text{gdzie } b_0, b_1, \dots, b_{n-1} \in L.$$

W szczególności, jeśli L jest ciałem skończonym, to $|K(\alpha)| = |K|^n$, gdzie $|X|$ — moc zbioru X .

Z historycznego punktu widzenia ważną rolę wśród rozszerzeń grają tzw. ROZSzerzenia KWADRATOWE, a więc takie rozszerzenia $K \subseteq K(\alpha)$, że $\alpha \notin K$ jest rozwiązaniem pewnego równania wielomianowego stopnia 2 o współczynnikach w ciele K . Podstawowym przykładem są tu ciała $\mathbb{Q}(\sqrt{p})$, gdzie p jest liczbą pierwszą. Liczba \sqrt{p} jest pierwiastkiem wielomianu $x^2 - p \in \mathbb{Q}[x]$.

Rozszerzenia kwadratowe wiążą się ze starożytnym zagadnieniem tzw. liczb konstruowalnych (nad ciałem \mathbb{Q}), czyli takich długości odcinków, które można skonstruować za pomocą cyrkla i linijki, mając do dyspozycji odcinek długości 1 (czyli też wszystkie odcinki długości $n \in \mathbb{N}$). Zagadnienie to pytało między innymi czy można za pomocą cyrkla i linijki¹:

- dokonać trysekcji dowolnego kąta, a więc np. czy można skonstruować kąt o mierze 20° ,
- skonstruować odcinek o tej własności, że sześcian, którego krawędzią jest ten odcinek ma objętość równą 2,
- skonstruować siedmiokąt foremny?

Jak się okazuje, opisane problemy dotyczą liczb algebraicznych (nad ciałem \mathbb{Q}). Liczba $\cos 20^\circ$ jest, jak się okazuje, pierwiastkiem wielomianu $4x^3 - 3x - \frac{1}{2}$. Liczba $\sqrt[3]{2}$ jest pierwiastkiem wielomianu $x^3 - 2$. Liczba $\cos \frac{2\pi}{7}$ jest, jak się okazuje, pierwiastkiem wielomianu

$$64x^7 - 112x^5 + 56x^3 - 7x - 1.$$

Nie jest to zupełnie elementarny wynik, ale żadna z powyższych liczb nie jest konstruowalna. Co więcej, zachodzą następujące twierdzenie, udowodnione w wieku XIX-tym.

Twierdzenie 7.5.6

Liczby konstruowalne nad \mathbb{Q} tworzą podciało ciała liczb rzeczywistych. Liczba rzeczywista x jest konstruowalna nad \mathbb{Q} wtedy i tylko wtedy, gdy istnieje ciąg rozszerzeń kwadratowych:

$$\mathbb{Q} \subseteq K_1 \subseteq K_2 \subseteq \dots \subseteq K_n$$

takich, że $x \in K_n$.

Mówiąc nieco nieprecyzyjnie, każda liczba niewymienna jest „iterowaną niewymiernością kwadratową”. Na czym to polega? Tak, jak rozważamy np. „niewymierność kwadratową” postaci $1 + \sqrt{2}$ należącą do ciała $\mathbb{Q}(\sqrt{2})$, tak można rozważyć element postaci $a + b\sqrt{3}$, gdzie $a, b \in \mathbb{Q}(\sqrt{2})$, na przykład element postaci:

$$(2 + 3\sqrt{2}) + (3 - \sqrt{2}) \cdot \sqrt{3}.$$

Liczba ta nie jest pierwiastkiem żadnego równania kwadratowego stopnia 2 o współczynnikach wymiernych, ale jest pierwiastkiem wielomianu stopnia 2 o współczynnikach w $\mathbb{Q}(\sqrt{3})$ postaci:

$$x^2 - (4 + 6\sqrt{2})x + (-11 + 6\sqrt{2}),$$

a więc można powiedzieć, że przywołana liczba jest „dwukrotnie iterowaną” niewymiernością kwadratową i jest konstruowalna. Odpowiednim ciągiem rozszerzeń kwadratowych jest tu:

$$\mathbb{Q} \subseteq \mathbb{Q}(\sqrt{2}) \subseteq \mathbb{Q}(\sqrt{2}, \sqrt{3}).$$

Więcej rachunków i przykładów, między innymi wyjaśnienia klasycznych problemów konstruowalności zarysowanych wyżej, znajdą Państwo w podręczniku prof. Guzickiego, natomiast warto wspomnieć, że zarysowana tu tematyka jest fragmentem tzw. teorii Galois związaną z zagadnieniem rozwiązywalności równań wielomianowych stopnia $n > 4$ przez tak zwane pierwiastniki. Tematy te poznają Państwo na Algebrze II (zachęcam do wyboru tego przedmiotu). Do rozważania powyższych tematów potrzebne są zarówno elementy teorii ciał i przestrzeni liniowych, jak również elementy teorii grup i pierścieni, które poznają Państwo na II roku studiów.

¹Więcej o tym zagadnieniu przeczytać można w rozdziale 13. podręcznika prof. Wojciecha Guzickiego *Geometria analityczna. Rozszerzony program matematyki w liceum*, wyd. Omega 2022.

Rozdział 8

Liniowo niezależne układy wektorów

8.1 Wykład 8

Poprzedni wykład poświęcony był pojęciu i podstawowym przykładom przestrzeni rozpiętych na układzie wektorów. Kontynuując te rozważania, wprowadzimy sposób opisu tych przestrzeni w sposób możliwe „oszczędny”. Podstawową motywację daje pytanie: jaka jest geometryczna struktura podprzestrzeni $\text{lin}(\alpha_1, \dots, \alpha_n)$? Jest jasne, że poniższe dwie podprzestrzenie \mathbb{R}^3 , choć rozpięte na układach dwóch wektorów, są „diametralnie różne”:

$$W_1 = \text{lin}((1, 1, 1), (2, 2, 2)), \quad W_2 = \text{lin}((1, 1, 1), (1, 2, 1)).$$

Pierwszą z tych podprzestrzeni można przedstawić w postaci $W_1 = \text{lin}((1, 1, 1))$. Drugiej natomiast nie można przedstawić w postaci $\text{lin}(\alpha)$, gdzie $\alpha \in \mathbb{R}^3$. To jest jasne, bo wektory $(1, 1, 1), (1, 2, 1)$ nie są proporcjonalne. Gdy liczba wektorów rozpinających podprzestrzeń wzrasta, analiza robi się bardziej skomplikowana i sam test proporcjonalności jest niewystarczający. Narzędziem właściwym dla rozstrzygnięcia tego problemu jest fundamentalne dla całej matematyki pojęcie liniowej niezależności układu wektorów.

Definicja 8.1.1: Liniowo zależny i liniowo niezależny układ wektorów (skończony)

Niech V będzie przestrzenią liniową nad ciałem K i niech 0_V będzie wektorem zerowym w V .

- Układ wektorów β_1, \dots, β_m przestrzeni V nad ciałem K nazwiemy LINIOWO ZALEŻNYM, jeśli istnieją elementy a_1, \dots, a_m ciała K , nie wszystkie równe 0, spełniające:

$$a_1\beta_1 + \dots + a_m\beta_m = 0_V.$$

- Układ wektorów $\alpha_1, \dots, \alpha_m$ przestrzeni V nazwiemy LINIOWO NIEZALEŻNYM, jeśli nie jest liniowo zależny. Równoważnie — układ ten jest liniowo niezależny, gdy dla dowolnych skalarów $a_1, \dots, a_m \in K$ zachodzi implikacja

$$a_1\alpha_1 + \dots + a_m\alpha_m = 0_V \implies a_1 = \dots = a_m = 0.$$

Pusty układ wektorów uważamy za liniowo niezależny.

Przykład 1. Układ złożony z jednego niezerowego wektora α jest liniowo niezależny, bowiem z równości $a\alpha = 0_V$ wynika, że $a = 0$ lub $\alpha = 0_V$. Skoro $\alpha \neq 0_V$, to $a = 0$.

Przykład 2. Układ wektorów $\alpha_1, \dots, \alpha_n$ zawierający wektor zerowy — powiedzmy α_n jest liniowo zależny, bo mamy $0 \cdot \alpha_1 + \dots + 0 \cdot \alpha_{n-1} + 1 \cdot 0_V = 0_V$. Podobnie pokazujemy, że układ zawierający choćby dwa identyczne (czy też proporcjonalne) wektory jest liniowo zależny.

Przykład 3. Układ $(1, 0, 0), (2, 0, 0), (3, 0, 0), (4, 0, 0), (5, 0, 0)$ jest liniowo zależny w \mathbb{R}^3 , bo:

$$1(1, 0, 0) + 1(2, 0, 0) + 1(3, 0, 0) + 1(4, 0, 0) + (-2)(5, 0, 0) = (0, 0, 0),$$

oraz

$$\text{lin}((1, 0, 0), (2, 0, 0), (3, 0, 0), (4, 0, 0), (5, 0, 0)) = \text{lin}((1, 0, 0)).$$

Przykład 4. Układ

$$\alpha_1 = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}, \quad \alpha_2 = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}, \quad \alpha_3 = \begin{bmatrix} 0 & 0 \\ 1 & 1 \end{bmatrix}$$

jest liniowo niezależny w przestrzeni liniowej $V = M_{2 \times 2}(\mathbb{R})$, bowiem dla dowolnych $a, b, c \in \mathbb{R}$:

$$a\alpha_1 + b\alpha_2 + c\alpha_3 = 0_V \iff a \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} + b \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix} + c \begin{bmatrix} 0 & 0 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix},$$

co oznacza, że

$$a\alpha_1 + b\alpha_2 + c\alpha_3 = \begin{bmatrix} a+b & a+b \\ b+c & a+c \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}.$$

A zatem mamy:

$$\begin{cases} a+b = 0 \\ b+c = 0 \\ a+c = 0 \end{cases} \quad (\dagger)$$

W rezultacie dostajemy $a = -b = c = -c = 0$. *Uwaga.* Ten sam układ macierzy $\alpha_1, \alpha_2, \alpha_3$ traktowanych jako elementy $M_{2 \times 2}(\mathbb{Z}_2)$ jest liniowo zależny, bowiem układ (\dagger) ma niezerowe rozwiązanie $(1, 1, 1)$ w \mathbb{Z}_2^3 .

Poniższy przykład zostanie szczegółowo omówiony podczas ćwiczeń.

Uwaga 8.1.2

Niech $0 \neq A = [a_{ij}] \in M_{m \times n}(K)$ będzie w postaci schodkowej oraz niech $\alpha_1, \dots, \alpha_r \in K^n$ – niezerowe wiersze macierzy A . Wówczas układ $\alpha_1, \dots, \alpha_r$ jest liniowo niezależny.

Dowód to indukcja po liczbie niezerowych wierszy r macierzy A . Krok bazowy: układ złożony z jednego niezerowego wektora jest liniowo niezależny. Przejdzmy do kroku indukcyjnego. Rozważmy macierz A w postaci schodkowej o r niezerowych wierszach $\alpha_1, \dots, \alpha_r$. Jeśli dla pewnych $\lambda_1, \dots, \lambda_r \in K$ mamy:

$$\lambda_1\alpha_1 + \lambda_2\alpha_2 + \dots + \lambda_r\alpha_r = (0, \dots, 0), \quad (\diamond)$$

to niech pierwszy niezerowy wyraz w wierszu α_1 stoi na k -tym miejscu.

$$\begin{bmatrix} 0 & \dots & 0 & \textcolor{red}{a_{1k}} & \dots & a_{1n} \\ 0 & \dots & 0 & \textcolor{red}{0} & \dots & a_{2n} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & \dots & 0 & \textcolor{red}{0} & \dots & a_{rn} \end{bmatrix}$$

Suma k -tych współrzędnych wektorów $\lambda_1\alpha_1, \dots, \lambda_r\alpha_r$ równa jest k -tej współrzędnej wektora zerowego, czyli $\lambda_1a_{1k} + \lambda_2a_{2k} + \dots + \lambda_ra_{rk} = 0$. Jednak $a_{2k} = \dots = a_{rk} = 0$, bo A jest schodkowa. Co więcej, $a_{1k} \neq 0$. A zatem mamy $\lambda_1a_{1k} = 0$, czyli $\lambda_1 = 0$. A zatem w równości (\diamond) dostajemy: $\lambda_2\alpha_2 + \dots + \lambda_r\alpha_r = (0, \dots, 0)$. Skoro $\alpha_2, \dots, \alpha_r$ są kolejnymi wierszami macierzy schodkowej, to z założenia indukcyjnego wektory te tworzą układ liniowo niezależny, czyli mamy $\lambda_2 = \lambda_3 = \dots = \lambda_r = 0$. Pokazaliśmy, że $\lambda_1\alpha_1 + \dots + \lambda_r\alpha_r = (0, \dots, 0)$ implikuje $\lambda_1 = \dots = \lambda_r = 0$.

Widzimy zatem, że ostatni przykład pozwala rozwiązać następujące zagadnienie w przestrzeni K^n : dana jest podprzestrzeń $W = \text{lin}(\beta_1, \dots, \beta_m)$ w K^n . Znajdź układ liniowo niezależny $\alpha_1, \dots, \alpha_r$ taki, że $W = \text{lin}(\alpha_1, \dots, \alpha_r)$. Rozwiązanie jest takie: traktujemy wektory β_1, \dots, β_m jako wiersze macierzy $A \in M_{m \times n}(K)$ i doprowadzamy A do postaci schodkowej. Zgodnie z powyższą obserwacją niezerowe wiersze $\alpha_1, \dots, \alpha_r$ macierzy A' są liniowo niezależne. Co więcej, na poprzednim wykładzie pokazaliśmy, że wiersze macierzy A' rozpinają tę samą podprzestrzeń K^n , co wiersze macierzy A . Widzimy więc, że problem jest rozwiązyany, bo A' ma r niezerowych wierszy i $m - r$ wierszy zerowych, oraz:

$$\text{lin}(\alpha_1, \dots, \alpha_r) = \text{lin}(\alpha_1, \dots, \alpha_r, \underbrace{0, \dots, 0}_{m-r}) = \text{lin}(\beta_1, \dots, \beta_m).$$

W dalszych rozważaniach wektor 0_V oznaczamy po prostu za pomocą symbolu 0.

Uwaga 8.1.3

Niech V będzie przestrzenią liniową nad ciałem K i niech $\beta_1, \dots, \beta_k \in V$. Następujące warunki są równoważne:

- układ β_1, \dots, β_k jest liniowo zależny,
- jeden z wektorów β_1, \dots, β_k jest kombinacją liniową pozostałych.

Intuicja jest następująca: liniowo zależny układ rozpinający jest „nadmiarowy” – można go „uszczerbić” do podukładu, który rozpinia tę samą podprzestrzeń. Należy też zauważać delikatność założenia: nie twierdzimy, że każdy wektor w układzie liniowo zależnym musi być kombinacją liniową pozostałych. Twierdzimy tylko, że w układzie liniowo zależnym istnieje taki wektor.

Przykład. Układ $(1, 0, 0), (2, 0, 0), (1, 1, 1)$ jest liniowo zależny w \mathbb{R}^3 , bo

$$2(1, 0, 0) + (-1)(2, 0, 0) + 0(1, 1, 1) = (0, 0, 0)$$

ale

- $(1, 1, 1)$ nie jest kombinacją liniową $(1, 0, 0), (2, 0, 0)$,
- $(1, 0, 0) = \frac{1}{2}(2, 0, 0) + 0(1, 1, 1)$.
- $(2, 0, 0) = 2(1, 0, 0) + 0(1, 1, 1)$.

Dowód. Przypuśćmy, że układ wektorów β_1, \dots, β_k jest liniowo zależny. Istnieją zatem $a_1, \dots, a_k \in K$, nie wszystkie równe 0, że $a_1\beta_1 + \dots + a_k\beta_k = 0$. Po ewentualnym przenumerowaniu wektorów możemy zakładać, że $a_1 \neq 0$ (tu nie ma żadnego oszustwa – proszę się nad tym chwilę zastanowić). Wtedy:

$$a_1\beta_1 = -a_2\beta_2 - \dots - a_k\beta_k,$$

czyli

$$\beta_1 = -\frac{a_2}{a_1}\beta_2 - \frac{a_3}{a_1}\beta_3 - \dots - \frac{a_k}{a_1}\beta_k.$$

Zatem β_1 jest kombinacją liniową pozostałych wektorów układu.

Na odwrót: jeśli jeden z wektorów układu jest kombinacją liniową pozostałych, to po ewentualnym przenumerowaniu możemy zakładać, że $\beta_1 = b_2\beta_2 + \dots + b_k\beta_k$. Wtedy $\beta_1 - b_2\beta_2 - \dots - b_k\beta_k = 0$, przy czym współczynnik przy β_1 jest równy 1, a więc jest niezerowy. Stąd układ β_1, \dots, β_k jest liniowo zależny. \square

Powyższe stwierdzenie sugeruje następujący, kluczowy wniosek.

Wniosek 8.1.4

Jeśli wektor β jest kombinacją liniową wektorów β_1, \dots, β_n , to

$$\text{lin}(\beta, \beta_1, \dots, \beta_n) = \text{lin}(\beta_1, \dots, \beta_n).$$

W szczególności, jeśli $V = \text{lin}(\beta_1, \dots, \beta_n)$, to z układu β_1, \dots, β_n wybrać można liniowo niezależny podukład $\alpha_1, \dots, \alpha_k$ taki, że

$$V = \text{lin}(\alpha_1, \dots, \alpha_k) = \text{lin}(\beta_1, \dots, \beta_n).$$

Przymajemy też $\{0\} = \text{lin}(\emptyset)$ (aby objąć $V = \{0\}$).

Dowód. Wystarczy udowodnić pierwszą część tezy. W tym celu dowodzimy dwie inkluze. Z jednej strony, skoro dla pewnych $a_1, \dots, a_n \in K$ mamy $\beta = a_1\beta_1 + \dots + a_n\beta_n$, to także $\beta = 0 \cdot \beta + a_1\beta_1 + \dots + a_n\beta_n$, czyli $\text{lin}(\beta, \beta_1, \dots, \beta_n) \subseteq \text{lin}(\beta_1, \dots, \beta_n)$. Przeciwna inkluza jest natomiast oczywista. \square

Pokazaliśmy, że każda przestrzeń rozpięta na skończonym układzie wektorów jest rozpięta przez pewien układ liniowo niezależny. Zobaczmy teraz, że przestrzeń rozpięta przez układ liniowo niezależny \mathcal{B} nie można rozpięć przez układ liniowo niezależny $\mathcal{C} \supsetneq \mathcal{B}$.

Uwaga 8.1.5

Niech $\alpha_1, \dots, \alpha_k$ będzie układem liniowo niezależnym w przestrzeni V i niech wektor $\beta \in V$. Następujące warunki są równoważne:

- (a) $\beta \in \text{lin}(\alpha_1, \dots, \alpha_k)$,
- (b) układ $\alpha_1, \dots, \alpha_k, \beta$ jest liniowo zależny.

Dowód. Jeśli $\beta \in \text{lin}(\alpha_1, \dots, \alpha_k)$, to układ $\alpha_1, \dots, \alpha_k, \beta$ jest liniowo zależny na mocy Uwagi 8.1.3. Na odwrót: przypuśćmy, że układ $\alpha_1, \dots, \alpha_k, \beta$ jest liniowo zależny. Istnieją wówczas $b, a_1, \dots, a_k \in K$, nie wszystkie równe 0, spełniające $b\beta + a_1\alpha_1 + \dots + a_k\alpha_k = 0$. Rozważamy dwa przypadki.

- Przypadek 1: $b = 0$. Wówczas $a_1\alpha_1 + \dots + a_k\alpha_k = 0$, przy czym pewne a_i jest niezerowe, co przeczy liniowej niezależności układu $\alpha_1, \dots, \alpha_k$.
- Przypadek 2: $b \neq 0$. Mamy $\beta = -\frac{a_1}{b}\alpha_1 - \dots - \frac{a_k}{b}\alpha_k$, czyli $\beta \in \text{lin}(\alpha_1, \dots, \alpha_k)$.

□

Twierdzenie 8.1.6: Steinitz (1910)

Jeśli układ wektorów $\alpha_1, \dots, \alpha_k$ leżących w przestrzeni $V = \text{lin}(\beta_1, \dots, \beta_m)$ jest liniowo niezależny, to:

- (a) $k \leq m$,
- (b) z układu β_1, \dots, β_m można wybrać taki podukład $\beta_{i_1}, \dots, \beta_{i_{m-k}}$, że:

$$\text{lin}(\beta_1, \dots, \beta_m) = \text{lin}(\alpha_1, \dots, \alpha_k, \beta_{i_1}, \dots, \beta_{i_{m-k}}).$$

Twierdzenie powyższe mówi o tym, że liczność rozpinających układów liniowo niezależnych respektuje porządek wyznaczony przez inkluzyję. Jeśli układ k wektorów rozpinia przestrzeń liniową, to nie można w niej „zmieścić” układu złożonego z więcej niż k liniowo niezależnych wektorów (punkt (a)). Co więcej, odpowiedni fragment dowolnego układu rozpinającego daną przestrzeń liniową można zastąpić dowolnym równolicznym układem liniowo niezależnym zawartym w tej przestrzeni (punkt (b)). Stąd też rezultat ten nazywa się w wielu źródłach **twierdzeniem o wymianie**.

Dowód. Pokazujemy najpierw punkt (a) przez indukcję ze względu na m . Teza indukcji brzmi: każdy układ liniowo niezależny zawarty w przestrzeni liniowej rozpiętej przez m wektorów ma nie więcej niż m elementów. Dla $m = 1$ twierdzenie jest oczywiste, bo jeśli układ niezerowych wektorów $\alpha_1, \dots, \alpha_k$ zawarty jest w $\text{lin}(\beta_1)$, to każdy z elementów tego układu jest postaci $a_i \cdot \beta_1$, gdzie a_1, \dots, a_k to niezerowe elementy K . Z drugiej strony jest jasne, że układ $a_1\beta_1, a_2\beta_1, \dots, a_k\beta_1$ może być liniowo niezależny tylko dla $k = 1$.

Przechodzimy do kroku indukcyjnego. Niech $m > 1$. Założmy, że teza jest prawdziwa dla każdej przestrzeni liniowej rozpiętej przez $m - 1$ wektorów. W przestrzeni liniowej $V = \text{lin}(\beta_1, \dots, \beta_m)$ rozpiętej przez m wektorów rozważamy układ liniowo niezależny $\alpha_1, \dots, \alpha_k$. Po ewentualnym przenumerowaniu¹ β_1, \dots, β_m weźmy takie a_{ij} , dla $1 \leq i \leq k, 1 \leq j \leq m$, że $a_{11} \neq 0$ oraz:

$$\begin{aligned} \alpha_1 &= a_{11}\beta_1 + a_{12}\beta_2 + \dots + a_{1m}\beta_m, \\ &\dots \\ \alpha_k &= a_{k1}\beta_1 + a_{k2}\beta_2 + \dots + a_{km}\beta_m. \end{aligned}$$

Teraz poprawiamy podukład $\{\alpha_2, \dots, \alpha_k\}$ układu $\{\alpha_1, \dots, \alpha_k\}$ do takiego układu $\{\gamma_2, \dots, \gamma_k\}$, który jest rozpięty tylko przez β_2, \dots, β_m . Dokładniej, określamy dla $i = 2, 3, \dots, k$ określamy układ wektorów $\gamma_2, \dots, \gamma_k \subseteq \text{lin}(\beta_2, \dots, \beta_m)$:

$$\gamma_i = \alpha_i - \frac{a_{i1}}{a_{11}}\alpha_1 = \underbrace{a_{i1}\beta_1 + a_{i2}\beta_2 + \dots + a_{im}\beta_m}_{\alpha_i} - \frac{a_{i1}}{a_{11}} \underbrace{(a_{11}\beta_1 + a_{12}\beta_2 + \dots + a_{1m}\beta_m)}_{\alpha_1}.$$

¹Czy Czytelnik widzi, dlaczego? W układzie liniowo niezależnym żaden wektor nie może być zerowy, czyli $\alpha_1 \neq 0$.

Nowy układ składa się z $k - 1$ wektorów i każdy jest rzeczywiście kombinacją jedynie wektorów postaci β_2, \dots, β_m (po powyższym rozpisaniu γ_i przy β_1 stoi 0). Przekonajmy się natomiast, że wektory $\gamma_2, \dots, \gamma_k$ są liniowo niezależne. Istotnie, gdybyśmy dla pewnych $c_2, \dots, c_k \in K$, nie wszystkich równych 0, mieli:

$$c_2\gamma_2 + \dots + c_k\gamma_k = 0 \Leftrightarrow c_2 \left(\alpha_2 - \frac{a_{21}}{a_{11}}\alpha_1 \right) + c_3 \left(\alpha_3 - \frac{a_{31}}{a_{11}}\alpha_1 \right) + \dots + c_k \left(\alpha_k - \frac{a_{k1}}{a_{11}}\alpha_1 \right) = 0,$$

czyli równoważnie:

$$-\frac{c_2a_{21} + c_3a_{31} + \dots + c_ka_{k1}}{a_{11}}\alpha_1 + c_2\alpha_2 + \dots + c_k\alpha_k = 0.$$

Skoro jednak układ $c_2 = 0, \dots, c_k = 0$, sprzeczność. Zatem układ $\gamma_2, \dots, \gamma_k$ jest liniowo niezależny.

Podsumujmy: okazało się, że jeśli mamy układ k liniowo niezależnych wektorów w przestrzeni rozpiętej przez m wektorów, to możemy wskazać układ $k - 1$ liniowo niezależnych wektorów zawarty w przestrzeni rozpiętej przez $m - 1$ wektorów. Z założenia indukcyjnego mamy więc $k - 1 \leq m - 1$. A zatem $k \leq m$.

Dowodzimy punkt (b). Rozważmy układ liniowo niezależny $\alpha_1, \dots, \alpha_k$ w przestrzeni $\text{lin}(\beta_1, \dots, \beta_m)$. Niech $\beta_{i_1}, \dots, \beta_{i_s}$ będzie najdłuższym podukładem w β_1, \dots, β_m , że układ

$$\alpha_1, \dots, \alpha_k, \beta_{i_1}, \dots, \beta_{i_s}$$

jest liniowo niezależny (na mocy punktu (a) takie s istnieje). W szczególności, dla każdego $1 \leq j \leq m$, dłuższy układ wektorów

$$\alpha_1, \dots, \alpha_k, \beta_{i_1}, \dots, \beta_{i_s}, \beta_j$$

jest już liniowo zależny. Na mocy Uwagi 8.1.5 mamy zatem:

$$\beta_j \in \text{lin}(\alpha_1, \dots, \alpha_k, \beta_{i_1}, \dots, \beta_{i_s}).$$

W szczególności $\text{lin}(\beta_1, \dots, \beta_m) \subseteq \text{lin}(\alpha_1, \dots, \alpha_k, \beta_{i_1}, \dots, \beta_{i_s})$. Oczywiście wszystkie α_i są kombinacjami liniowymi β_j więc $\text{lin}(\beta_1, \dots, \beta_m) = \text{lin}(\alpha_1, \dots, \alpha_k, \beta_{i_1}, \dots, \beta_{i_s})$. Z udowodnionego już punktu (a) wynika, że

$$k + s \leq m,$$

a więc

$$s \leq m - k.$$

Stąd dołączając do układu $\alpha_1, \dots, \alpha_k, \beta_{i_1}, \dots, \beta_{i_s}$ dowolne $m - k - s$ wektorów $\gamma_1, \dots, \gamma_{m-k-s}$ spośród β_i , dla $i \neq i_1, \dots, i_s$, otrzymujemy układ spełniający

$$\text{lin}(\beta_1, \dots, \beta_m) = \text{lin}(\alpha_1, \dots, \alpha_k, \beta_{i_1}, \dots, \beta_{i_s}, \gamma_1, \dots, \gamma_{m-k-s}).$$

□

Pojęcie liniowej niezależności uogólnia się na dowolne, niekoniecznie skończone, układy wektorów.

Definicja 8.1.7

Układ $X = \{\alpha_t\}_{t \in T}$ wektorów przestrzeni V nazywamy liniowo niezależnym, jeśli każdy jego skończony podukład jest liniowo niezależny (czyli dla każdych $\alpha_{t_1}, \dots, \alpha_{t_k} \in X$ oraz $a_1, \dots, a_k \in K$ z równości $a_1\alpha_{t_1} + \dots + a_k\alpha_{t_k} = 0$ wynika, że $a_1 = 0, \dots, a_k = 0$).

Przykłady, których uzasadnienie zostawiam jako ćwiczenie.

- (a) układ $\{1, x, x^2, x^3, \dots\}$ jest liniowo niezależny w $K[x]$,
- (b) układ ciągów $a_1 = (1, 0, 0, \dots), a_2 = (0, 1, 0, \dots), a_3 = (0, 0, 1, \dots), \dots$ jest liniowo niezależny w K^∞ ,
- (c) układ ciągów $\{(1, t, t^2, t^3, \dots), t \in (0, 1)\}$ jest liniowo niezależny w \mathbb{R}^∞ ,
- (d) układ $\{\sqrt{2}, \sqrt{3}, \sqrt{5}, \sqrt{7}, \sqrt{11}, \dots\} = \{\sqrt{p}, p \in P\}$, gdzie P — zbiór liczb pierwszych, jest liniowo niezależny w przestrzeni \mathbb{R} nad ciałem \mathbb{Q} ,
- (e) układ $\{\sin(x), \sin^2(x), \sin^3(x), \dots\} = \{\sin(x)^n, n \in \mathbb{N}_+\}$ jest liniowo niezależny w przestrzeni $F(\mathbb{R}, \mathbb{R})$.

8.2 Wybór przykładowych pytań

Proszę się nie uczyć na pamięć odpowiedzi czy wyjaśnienie! Takich pytań można ułożyć znacznie więcej i służą one jedynie sprawdzeniu podstawowego zrozumienia treści zajęć. Jeśli odpowiedź na dowolne z poniższych pytań sprawia Państwu problem, proszę pomyśleć który fragment notatek dotyczy tego pytania, wrócić do ich lektury, a w razie dalszych trudności poprosić prowadzącego o wskazówkę.

1. Czy układ wektorów $(1, 0, 0), (2, 0, 0)$ jest liniowo zależny w \mathbb{R}^3 ?
2. Czy układ wektorów $(1, 1, 0), (1, 2, 0), (2, 1, 0)$ jest liniowo zależny w \mathbb{R}^3 ?
3. Czy układ wektorów $(0, 0, 1), (0, 1, 0), (0, 0, 1)$ jest liniowo zależny?
4. Czy poniższy układ macierzy jest liniowo zależny w $M_{2 \times 2}(\mathbb{R})$?

$$\begin{bmatrix} 1 & 1 \\ 2 & 2 \end{bmatrix}, \quad \begin{bmatrix} 2 & 2 \\ 3 & 3 \end{bmatrix}, \quad \begin{bmatrix} 3 & 3 \\ 1 & 1 \end{bmatrix}?$$

5. Niech ξ_1, ξ_2, ξ_3 będą parami różnymi pierwiastkami zespolonymi stopnia 3 z 1. Czy układ ξ_1, ξ_2, ξ_3 jest liniowo zależny w przestrzeni liniowej \mathbb{C} nad ciałem \mathbb{R} ?
6. Czy układ wektorów α, α jest liniowo niezależny?
7. Czy układ wektorów $\alpha, -\alpha$ jest liniowo niezależny?
8. Czy układ wektorów $\alpha, \beta, \alpha + \beta$ jest liniowo niezależny?
9. Suma wektorów α, β, γ równa jest 0. Czy jest to liniowo zależny układ wektorów?
10. Suma wektorów α, β, γ równa jest 0. Czy układ $\alpha, \alpha + \beta, \gamma$ jest liniowo zależny?
11. Układ wektorów α, β, γ jest liniowo zależny. Czy liniowo zależny jest układ $\alpha + \beta, \beta + \gamma, \alpha + \beta$?
12. Układ wektorów α, β, γ jest liniowo zależny. Czy liniowo zależny jest układ $\alpha + 2\beta, \beta + 2\gamma, \alpha + 2\beta$?
13. Układ wektorów α, β, γ jest liniowo zależny. Niech $\delta, \mu, \xi \in \text{lin}(\alpha, \beta, \gamma)$. Czy układ δ, μ, ξ jest liniowo zależny?
14. Układ wektorów α, β, γ jest liniowo niezależny. Niech $\delta, \mu, \xi \in \text{lin}(\alpha, \beta, \gamma)$. Czy układ δ, μ, ξ jest liniowo niezależny?
15. Czy jeśli układ wektorów α, β jest liniowo zależny, to $\beta = a\alpha$, dla pewnego $a \in \mathbb{K}$?
16. Czy układ zawierający wektor zerowy może być liniowo niezależny?
17. Układy α_1, α_2 oraz β_1, β_2 są liniowo niezależne. Czy układ złożony z ich sum $\alpha_1 + \beta_1, \alpha_2 + \beta_2$ jest liniowo niezależny?
18. Czy przestrzeń liniowa \mathbb{R}^3 zawiera układ wektorów $\alpha_1, \alpha_2, \alpha_3$, który jest liniowo zależny, ale każdy jego podkład jest liniowo niezależny?
19. Czy istnieją wektory $\alpha, \beta, \gamma \in \mathbb{R}^4$ takie, że układ $\{\alpha - \beta, \beta - \gamma, \gamma - \alpha\}$ jest liniowo niezależny?
20. Dany jest układ liniowo zależny $\alpha_1, \alpha_2, \alpha_3$ w przestrzeni liniowej V . Czy każdy z wektorów z tego układu da się zapisać jako kombinacja liniowa dwóch pozostałych?
21. Niech V_1, V_2 będą różnymi podprzestrzeniami przestrzeni liniowej V . Niech $\alpha_1 \in V_1 \setminus V_2$ oraz $\alpha_2 \in V_2 \setminus V_1$. Czy układ $\{\alpha_1, \alpha_2\}$ może być liniowo zależny?
22. Wiadomo, że układ złożony z wektorów α, β jest liniowo niezależny, zaś układ złożony z wektorów $\alpha + \gamma$ i $\beta + \gamma$ jest liniowo zależny. Czy $\gamma \in \text{lin}(\alpha, \beta)$?
23. Wiadomo, że wektory $\alpha + \gamma$ i $\beta + \gamma$ są liniowo zależne. Czy $\gamma \in \text{lin}(\alpha, \beta)$?
24. Czy układ złożony ze wszystkich ciągów stałych o wyrazach rzeczywistych jest liniowo zależny?
25. Czy układ złożony ze wszystkich ciągów rosnących o wyrazach rzeczywistych jest liniowo zależny?
26. Wskaż przykład nieskończonego układu liniowo niezależnego w \mathbb{R}^∞ złożonego z takich ciągów, których żaden wyraz nie jest równy 0.

8.3 Zadania do samodzielnej pracy

Zadania ilustrujące umiejętności niezbędne do zaliczenia przedmiotu oznaczone są symbolem ♠.

Przy tych zadaniach dodany jest opis umiejętności, które sprawdzają.

- (♠) Rozstrzyganie czy układ wektorów jest liniowo niezależny)

Dla jakich wartości parametrów $s, t \in \mathbb{R}$ wektory

$$(5, 7, s, 2), \quad (1, 3, 2, 1), \quad (2, 2, 4, t)$$

przestrzeni liniowej \mathbb{R}^4 tworzą układ liniowo niezależny?

- (♠) Czy wektory

$$\alpha = (0, 1, 1), \quad \beta = (1, 0, 1), \quad \gamma = (1, 1, 0)$$

z przestrzeni \mathbb{Z}_2^3 tworzą układ liniowo niezależny?

- (♠) Sprawdź czy układ wielomianów

$$w_1 = 1 + t, \quad w_2 = 1 + t^2, \quad w_3 = t + t^2$$

jest liniowo niezależny w przestrzeni liniowej $\mathbb{R}[t]$. Czy odpowiedź zmienia się, jeśli układ ten rozpatrujemy w przestrzeni liniowej $\mathbb{Z}_2[t]$?

- (♠) Wykaż, że funkcje

$$\sin(x), \quad \cos(x)$$

tworzą liniowo niezależny układ wektorów przestrzeni $F(\mathbb{R}, \mathbb{R})$.

- (♠) Niech

$$A_1 = \begin{bmatrix} a & 2a \\ 2 & 3a \end{bmatrix}, \quad A_2 = \begin{bmatrix} 1 & 2 \\ a & 3 \end{bmatrix}, \quad A_3 = \begin{bmatrix} 1 & 2a \\ a+1 & a+2 \end{bmatrix}, \quad A_4 = \begin{bmatrix} 1 & a+1 \\ 2 & 2a+1 \end{bmatrix}.$$

Dla jakich $a \in \mathbb{R}$ układ $\{A_1, A_2, A_3, A_4\}$ jest liniowo niezależny w przestrzeni liniowej $M_2(\mathbb{R})$?

- (♠) Rozstrzyganie czy układ wektorów jest liniowo niezależny)

Wektory $\alpha_1, \alpha_2, \alpha_3$ tworzą układ liniowo niezależny w przestrzeni liniowej V nad ciałem \mathbb{R} . Rozpatrzmy wektor

$$\beta = -\alpha_1 + 7\alpha_2 - 3\alpha_3.$$

Czy wektory $\alpha_1, \alpha_2, \beta$ tworzą układ liniowo niezależny?

- Rozpatrzmy \mathbb{R} jako przestrzeń liniową nad \mathbb{Q} . Wykaż, że układ $1, \sqrt{2}, \sqrt{3}$ w tej przestrzeni jest liniowo niezależny.

- Niech $\alpha_1, \alpha_2, \dots, \alpha_k$ będzie liniowo niezależnym układem wektorów przestrzeni V nad ciałem K . Czy układ β_1, \dots, β_k jest liniowo niezależny, jeśli:

- (a) $\beta_1 = \alpha_1$ oraz $\beta_i = \alpha_i + \alpha_{i-1}$, dla $i = 2, 3, \dots, k$,
- (b) $\beta_i = \alpha_1 + \dots + \alpha_i$, dla $i = 1, 2, \dots, k$,
- (c) $\beta_i = \alpha_i + \alpha_{i+1}$, dla $i = 1, 2, \dots, k-1$ oraz $\beta_k = \alpha_k + \alpha_1$.

- Niech V_1, V_2 będą różnymi podprzestrzeniami przestrzeni liniowej V . Załóżmy, że $\alpha_1 \in V_1 \setminus V_2$ oraz $\alpha_2 \in V_2 \setminus V_1$. Rozstrzygnij, czy układ $\{\alpha_1, \alpha_2\}$ może być liniowo zależny?

- Dany jest liniowo zależny układ wektorów $\{\alpha_1, \dots, \alpha_n\}$ w przestrzeni liniowej V . Czy każdy z wektorów α_i tego układu można zapisać jako kombinację liniową pozostałych?

- Układ wektorów v_1, \dots, v_n przestrzeni liniowej V jest liniowo niezależny. Wykaż, że jeśli dla pewnego $w \in V$ układ $v_1 + w, \dots, v_n + w$ jest liniowo zależny, to $w \in \text{lin}(v_1, \dots, v_n)$.

- Wykaż, że istnieje nieskończony podzbiór wektorów przestrzeni \mathbb{R}^n , taki że każde n wektorów w tym podzbiorze tworzy układ liniowo niezależny.

- Niech f_1, \dots, f_k będzie takim układem wielomianów w $K[x]$, nie zawierającym wielomianu zerowego, że $\deg f_i \neq \deg f_j$ dla każdego $i \neq j$. Wykaż, że układ f_1, \dots, f_k jest liniowo niezależny.

- Niech K będzie ciałem i niech f_1, \dots, f_n będzie układem funkcji należących do przestrzeni $F(X, K)$. Wykazać, że układ ten jest liniowo niezależny wtedy i tylko wtedy, gdy istnieją takie elementy x_1, \dots, x_n zbioru X , że układ wektorów w przestrzeni K^n złożony z wektorów $\alpha_1, \alpha_2, \dots, \alpha_n$, gdzie $\alpha_i = (f_1(x_i), f_2(x_i), \dots, f_n(x_i))$ jest liniowo niezależny.

8.4 Dodatek. Nieprzeliczalne układy. Algebraiczna niezależność.

Poniższe zadanie ilustruje jak skomplikowana jest struktura ciała \mathbb{R} traktowanego jako przestrzeń liniowa nad ciałem \mathbb{Q} . Wypisanie wprost układu rozpinającego \mathbb{R} nad \mathbb{Q} nie jest możliwe. Można jednak wypisać równoliczny z \mathbb{R} układ liniowo niezależny. Za udostępnienie materiału dziękuje dr. Ł. Kubatowi.

Zadanie. Ustawmy liczby wymierne \mathbb{Q} w ciąg $(q_n)_{n \in \mathbb{N}}$. Dla dowolnego $t \in \mathbb{R}$ niech

$$a(t) = \sum_{n \in N(t)} \frac{1}{n!}, \quad \text{gdzie } N(t) = \{n \in \mathbb{N} : q_n < t\}.$$

- (1) Sprawdź, że szereg definiujący $a(t)$ jest zbieżny (czyli definicja jest poprawna).
- (2) Wykaż, że dla dowolnych $s, t \in \mathbb{R}$ zachodzi $s \neq t \implies a(s) \neq a(t)$.
- (3) Udowodnij, że zbiór $A = \{a(t) : t \in \mathbb{R}\} \subseteq \mathbb{R}$ jest liniowo niezależny nad \mathbb{Q} .

Rozwiązanie. Zauważmy, że

$$a(t) = \sum_{n \in N(t)} \frac{1}{n!} < \sum_{n=0}^{\infty} \frac{1}{n!} = e,$$

co dowodzi (1). Gdy $s, t \in \mathbb{R}$ spełniają $s < t$, to $(s, t) \cap \mathbb{Q} \neq \emptyset$. Istnieje więc takie $n \in \mathbb{N}$, że $s < q_n < t$. Zatem $a(s) < a(s) + \frac{1}{n!} < a(t)$, co dowodzi (2). Aby udowodnić (3) założymy, dla dowodu nie wprost, że

$$q_1 a(t_1) + \cdots + q_k a(t_k) = 0 \quad (*)$$

dla pewnych $t_1, \dots, t_k \in \mathbb{R}$ spełniających $t_1 > \cdots > t_k$, gdzie liczby $q_1, \dots, q_k \in \mathbb{Q}$ nie są wszystkie równe zeru. Wśród równości typu (*) możemy wybrać najkrótszą, czyli taką, w której k jest najmniejsze. Oczywiście musi być $k \geq 2$. Ponadto, dzięki minimalności k , koniecznie $q_1, \dots, q_k \neq 0$. Mnożąc (*) przez stosowną liczbę naturalną możemy założyć, że $q_1, \dots, q_k \in \mathbb{Z}$. Dla $t \in \mathbb{R}$ oraz $m \in \mathbb{N}$ niech

$$L_m(t) = \{n \in N(t) : n \leq m\} \quad \text{oraz} \quad R_m(t) = \{n \in N(t) : n > m\}.$$

Mnożąc (*) przez $m!$ otrzymujemy $L(m) = -R(m)$, gdzie

$$\begin{aligned} L(m) &= q_1 \left(\sum_{n \in L_m(t_1)} \frac{m!}{n!} \right) + \cdots + q_k \left(\sum_{n \in L_m(t_k)} \frac{m!}{n!} \right), \\ R(m) &= q_1 \left(\sum_{n \in R_m(t_1)} \frac{m!}{n!} \right) + \cdots + q_k \left(\sum_{n \in R_m(t_k)} \frac{m!}{n!} \right). \end{aligned}$$

Oczywiście $L(m) \in \mathbb{Z}$. Ponadto

$$\begin{aligned} |R(m)| &\leq |q_1| \left(\sum_{n \in R_m(t_1)} \frac{m!}{n!} \right) + \cdots + |q_k| \left(\sum_{n \in R_m(t_k)} \frac{m!}{n!} \right) \\ &\leq |q_1| \left(\sum_{n>m} \frac{m!}{n!} \right) + \cdots + |q_k| \left(\sum_{n>m} \frac{m!}{n!} \right) \\ &\leq \frac{|q_1| + \cdots + |q_k|}{m+1} \left(\sum_{n>m} \frac{1}{(n-m)!} \right) \\ &\leq \frac{|q_1| + \cdots + |q_k|}{m+1} e. \end{aligned}$$

Wynika stąd, że gdy m jest duże, to $|R(m)| < 1$. Skoro $R(m) = -L(m) \in \mathbb{Z}$, to musi zachodzić równość $L(m) = R(m) = 0$. Ponieważ zbiór $(t_2, t_1) \cap \mathbb{Q}$ jest nieskończony, to znajdziemy takie $m \in \mathbb{N}$ by jednocześnie $|R(m)| < 1$ (wtedy, jak wiemy, $L(m) = R(m) = 0$) oraz $t_2 < q_m < t_1$. W tej sytuacji mamy $m \in L_m(t_1) \setminus (L_m(t_2) \cup \cdots \cup L_m(t_k))$. Zatem równość $L(m) = 0$ implikuje

$$-q_1 = q_1 \left(\sum_{\substack{n \in L_m(t_1) \\ n \neq m}} \frac{m!}{n!} \right) + q_2 \left(\sum_{n \in L_m(t_2)} \frac{m!}{n!} \right) + \cdots + q_k \left(\sum_{n \in L_m(t_k)} \frac{m!}{n!} \right). \quad (**)$$

Obie strony równania (**) są liczbami całkowitymi. Ponadto prawa strona jest podzielna przez m . Zatem także q_1 jest podzielne przez m . W takim razie musi być $q_1 = 0$, gdyż m można wybrać tak, by $m > |q_1|$. Uzyskana sprzeczność ($q_1 = 0$) prowadzi do wniosku, że zbiór A jest liniowo niezależny nad \mathbb{Q} . \square

Uwaga. Dowodzi się, że choć zbiór A jest tej samej mocy co \mathbb{R} , to nie rozpinie on \mathbb{R} nad \mathbb{Q} . Można także wykazać (patrz J. von Neumann, *Ein System algebraisch unabhängiger Zahlen*, Math. Ann. **99** (1928), pp. 134–141), że liczby postaci

$$b(t) = \sum_{n=0}^{\infty} \frac{2^{2^{[nt]}}}{2^{2^{n^2}}} \quad \text{dla } t > 0$$

($[x]$ oznacza część całkowitą liczby $x \in \mathbb{R}$)

są nie tylko liniowo niezależne nad \mathbb{Q} , ale nawet **algebraicznie niezależne** nad \mathbb{Q} , tzn. dla dowolnego $n \geq 1$, dowolnych $0 < t_1 < \dots < t_n$ oraz dowolnego wielomianu zmiennych x_1, \dots, x_n , czyli dla pewnego $0 \neq f \in \mathbb{Q}[x_1, \dots, x_n]$ zachodzi $f(b(t_1), \dots, b(t_n)) \neq 0$.

Przykład ilustrujący algebraiczną zależność. Liczby $\sqrt{\pi}$ oraz $2\pi + 1$ są liniowo niezależne nad \mathbb{Q} , ale są algebraicznie zależne, ponieważ wielomian $2x^2 - y - 1 \in \mathbb{Q}[x, y]$ zeruje się dla $x = \sqrt{\pi}$ oraz $y = 2\pi + 1$.

Prof. J. Mycielski pokazał następujące twierdzenie (*Algebraic independence and measure*, Fund. Math. **61** (1967), pp. 165–169) dla dowolnego doskonałego podzbioru \mathbb{R} (tzn. niepstego, domkniętego oraz bez punktów izolowanych), patrz: <http://matwbn.icm.edu.pl/ksiazki/fm/fm61/fm61117.pdf>.

Twierdzenie 8.4.1

Każdy doskonały podzbiór zbioru liczb rzeczywistych zawiera doskonały podzbiór, który jest algebraicznie niezależny nad \mathbb{Q} .

Pojęcie algebraicznej niezależności elementów \mathbb{R} nad \mathbb{Q} , a także elementów \mathbb{C} nad \mathbb{Q} , związane jest ściśle z pojęciem liczb algebraicznych i przestępnych, z którymi spotkaliście się Państwo (lub spotkacie) na Analizie Matematycznej. Pojęcie to jest bardzo subtelne i tajemnicze. Przestępność liczby π została udowodniona po raz pierwszy właśnie dzięki badaniu algebraicznej niezależności (dowód dla e dokonał elementarnymi metodami analitycznymi Hermite w 1873 roku). Zachodzi mianowicie następujący rezultat.

Twierdzenie 8.4.2: Lindemann-Weierstrass, 1885

Jeśli $\alpha_1, \dots, \alpha_n$ są liczbami algebraicznymi liniowo niezależnymi nad \mathbb{Q} , to liczby $e^{\alpha_1}, \dots, e^{\alpha_n}$ są algebraicznie niezależne nad \mathbb{Q} .

Aby zrozumieć jakie są związki tego wyniku z przestępcością odnotujmy inne, równoważne sformułowanie.

Twierdzenie 8.4.3: Baker, 1966

Jeśli $\alpha_1, \dots, \alpha_n$ są parami różnymi liczbami algebraicznymi, to liczby $e^{\alpha_1}, \dots, e^{\alpha_n}$ są liniowo niezależne nad ciałem liczb algebraicznych $\overline{\mathbb{Q}}$ (algebraiczne domknięcie \mathbb{Q} w \mathbb{C}).

Jak można stosować to twierdzenie? Jeśli α jest niezerową liczbą algebraiczną to zbiór $\{0, \alpha\}$ zawiera różne elementy algebraiczne, więc zbiór $\{e^0, e^\alpha\}$, czyli $\{1, e^\alpha\}$ jest liniowo niezależny nad ciałem liczb algebraicznych, w szczególności e^α nie jest algebraiczna. Gdy udowodnimy przestępcość liczby e możemy z niej łatwo wywnioskować przestępcość liczby π , korzystając ze słynnej tożsamości algebraicznej Eulera $e^{\pi i} + 1 = 0$. Dokładniej, gdyby π była liczbą algebraiczną to πi również, a wtedy przestępna musi być, na mocy poprzedniego argumentu liczba $e^{\pi i} = -1$, co jest niemożliwe. Zatem π jest przestępna. Prosty wariant tego argumentu pokazuje również, że dla niezerowej liczby algebraicznej α liczby $\sin(\alpha), \cos(\alpha), \operatorname{tg}(\alpha)$ i ich hiperboliczne odpowiedniki są liczbami przestępnnymi.

Dowód Twierdzenia Lindemanna jest skomplikowany ale o zagadnieniach tego typu i szeregu innych rezultatów dotyczących liczb przestępnych: <http://www.math.leidenuniv.nl/~evertse/dio15-4.pdf>.

8.5 Trivia. Wektory przynależności do klubów

Wiele ciekawych zastosowań algebry liniowej odnaleźć można w kombinatoryce². Weźmy zbiór n elementowy, na przykład złożony z mieszkańców pewnego miasta. Założymy dalej, że w mieście tym są pewne kluby, na przykład Wisła i Cracovia. Każdemu z tych klubów przypisujemy wektor v_1, v_2 o n współrzędnych w zbiorze $\{0, 1\}$ w następujący sposób: i -ta współrzędna wektora v_i jest równa

- 1, jeśli i -ty mieszkaniec naszego kraju jest kibicem tego klubu
- 0, jeśli i -ty mieszkaniec naszego kraju nie jest kibicem tego klubu,

Gdyby miasto to miało $n = 4$ mieszkańców 1, 2, 3, 4, z których 1 i 4 są kibicami Wisły, a 2, 3 nimi nie są, wówczas mielibyśmy $v_1 = (1, 0, 0, 1)$.

Wprowadzamy operację na parach wektorów przynależności, która ma czytelną interpretację kombinatoryczną. Biorąc dwa wektory $x = (x_1, \dots, x_n)$ oraz $y = (y_1, \dots, y_n)$, reprezentujące przynależność do grona kibiców pewnych dwóch klubów i rozważając operację:

$$\langle x, y \rangle = x_1 y_1 + x_2 y_2 + \dots + x_n y_n$$

widzimy, że wartość $\langle x, y \rangle$ równa jest po prostu liczbie kibiców, wspierających jednocześnie obydwa kluby, a liczba $\langle x, x \rangle$ to liczba kibiców Wisły. Przykładowo, jeśli wektor kibiców Wisły ma postać $(1, 0, 0, 1)$, a wektor kibiców Cracovii ma postać $(0, 0, 0, 1)$, to

$$\langle (1, 0, 0, 1), (0, 0, 0, 1) \rangle = 1 \cdot 0 + 0 \cdot 0 + 0 \cdot 0 + 1 \cdot 1 = 1,$$

a zatem tylko jeden kibic należy jednocześnie do grona sympatyków Wisły i Cracovii.

Być może Czytelnik nie dostrzega jeszcze do czego mogłyby nam się tu przydać kombinacje liniowe czy liniowa niezależność. Sformalizujmy nasze pojęcia, przechodząc do przestrzeni współrzędnych nad dowolnym ciałem. Aby to zrobić, wprowadzimy tak zwany **wektor przynależności** do klubu w mieście o n mieszkańcach, który będzie elementem przestrzeni liniowej K^n nad ciałem K .

Dokładniej, niech m będzie liczbą klubów C_1, C_2, \dots, C_m w mieście o n mieszkańcach $\{1, 2, \dots, n\}$. Dla każdego $i = 1, 2, \dots, m$, niech $n_i \in K^n$ będzie wektorem kodującym przynależność mieszkańców do i -tego klubu, to znaczy $v_i = (v_{i1}, \dots, v_{in})$, gdzie:

$$v_{ij} = \begin{cases} 1, & j \in C_i \\ 0, & j \notin C_i \end{cases}.$$

W przykładach wyżej mamy $m = 2$, C_1 to klub Wisła, C_2 to klub Cracovia, a $n = 4$. Dla wektorów przynależności do klubów C_i oraz C_j , czyli $v_i = (v_{i1}, \dots, v_{in})$ oraz $v_j = (v_{j1}, \dots, v_{jn})$ należących do K^n , definiujemy operację:

$$\langle v_i, v_j \rangle = \langle (v_{i1}, \dots, v_{in}), (v_{j1}, \dots, v_{jn}) \rangle = v_{i1} v_{j1} + v_{i2} v_{j2} + \dots + v_{in} v_{jn} \in K.$$

Zauważmy, że dla $K = \mathbb{R}$ wartość powyższej liczby równa jest liczbie wspólnych członków obydwu klubów. Dla $K = \mathbb{Z}_2$ dostajemy jedynie informacje, czy liczby te są parzyste, czy nieparzyste. Czytelnikowi zostawiamy uzasadnienie jednej ważnej własności wprowadzonej przez nas operacji (wynika ona wprost z definicji). Jeśli wektory v, w, z należą do K^n , wówczas dla dowolnych $c_1, c_2 \in K$ mamy:

$$\langle v, c_1 w + c_2 z \rangle = c_1 \langle v, w \rangle + c_2 \langle v, z \rangle. \quad (\dagger)$$

Obserwacja 8.5.1: Problem Parzystkowa

W mieście o n mieszkańcach i m klubach, w którym każdy klub ma nieparzystą liczbę członków i każde dwa kluby mają parzystą liczbę wspólnych członków, zachodzi $m \leq n$.

²Na podstawie: Y. Zhao: Linear algebra tricks for the Putnam, yufeizhao.com/olympiad/putnam_linear_algebra.pdf.

Dowód. Rozważmy wektory przynależności $v_1, \dots, v_m \in \mathbb{Z}_2^n$. Zauważmy, że skoro wektor v_i ma nieparzyste wiele niezerowych współrzędnych, to $\langle v_i, v_i \rangle = 1$. Co więcej, dla każdych $i \neq j$ mamy $\langle v_i, v_j \rangle = 0$, ponieważ dowolne dwa kluby mają parzystą liczbę wspólnych członków. Twierdzimy, że wynika stąd, że wektory v_1, \dots, v_m są liniowo niezależne w \mathbb{Z}_2^n . Istotnie, jeśli istnieją $c_1, \dots, c_m \in \mathbb{Z}_2$, takie że

$$c_1v_1 + c_2v_2 + \dots + c_mv_m = 0,$$

to mamy również, na mocy (†):

$$0 = \langle v_i, 0 \rangle = \langle v_i, c_1v_1 + c_2v_2 + \dots + c_mv_m \rangle = c_1\langle v_i, v_1 \rangle + c_2\langle v_i, v_2 \rangle + \dots + c_m\langle v_i, v_m \rangle = c_i.$$

Stąd układ wektorów v_1, \dots, v_m jest liniowo niezależny. Jednak układ m wektorów w przestrzeni K^n jest liniowo niezależny jedynie, gdy $m \leq n$, co wynika z twierdzenia Steinitza. \square

Obserwacja 8.5.2: Nierówność Fishera

Niech k będzie dodatnią liczbą całkowitą. W mieście o n mieszkańców utworzono m klubów, przy czym każde dwa kluby mają dokładnie k wspólnych członków. Wykaż, że $m \leq n$.

Idea dowodu jest podobna, choć tym razem wektory v_1, \dots, v_m reprezentujące kluby C_1, \dots, C_m należeć będą do \mathbb{R}^n . W takim przypadku $\langle v_i, v_i \rangle$ opisuje liczbę członków klubu C_i , a liczba $\langle v_i, v_j \rangle$ opisuje liczbę wspólnych członków klubów C_i oraz C_j , czyli k .

Dowód. Rozumując podobnie jak poprzednio stwierdzamy, że jeśli v_1, \dots, v_m są liniowo niezależne, to $m \leq n$. Przypuśćmy, że układ ten nie jest liniowo niezależny. Istnieją zatem $c_1, \dots, c_m \in \mathbb{R}$, że

$$c_1v_1 + \dots + c_mv_m = 0$$

Niech $v_i = (v_{i1}, \dots, v_{in})$. Zatem korzystając z tego, że $\langle v, w \rangle = \langle w, v \rangle$ oraz korzystając wielokrotnie z (†), uzyskujemy

$$\begin{aligned} \langle c_1v_1 + \dots + c_mv_m, c_1v_1 + \dots + c_mv_m \rangle &= \sum_{i=1}^m c_i^2 \langle v_i, v_i \rangle + 2 \sum_{i < j} c_i c_j \langle v_i, v_j \rangle = \\ &= \sum_{i=1}^m c_i^2 \langle v_i, v_i \rangle + 2k \sum_{i < j} c_i c_j = \\ &= \sum_{i=1}^m c_i^2 (\langle v_i, v_i \rangle - k) + \sum_{i=1}^n \sum_{j=1}^n c_i c_j = \\ &= \sum_{i=1}^m c_i^2 (\langle v_i, v_i \rangle - k) + \left(\sum_{i=1}^n c_i \right)^2. \end{aligned}$$

Skoro dowolne dwa kluby mają dokładnie k wspólnych członków, to każdy klub ma ich co najmniej k . Stąd $\langle v_i, v_i \rangle \geq k$. W szczególności wszystkie wyrazy końcowej sumy są nieujemne, a stąd muszą być równe 0. Skoro pewne c_i jest niezerowe, to pewne $\langle v_i, v_i \rangle - k$ jest równe 0, a więc klub C_i ma dokładnie k członków. Stąd jednak wynika, zgodnie z założeniem twierdzenia, że wszystkie kluby zawierają wszystkich k członków klubu C_i , a poza tym są rozłączne (zawierają unikatowych mieszkańców). Stąd wynika natychmiast, że klubów tych jest nie więcej niż mieszkańców, czyli n .

* * *

O operacji $\langle v_i, v_j \rangle$ powiemy więcej w kolejnych rozdziałach oraz w drugim semestrze. Ogólnie mówiąc jest to rodzaj tzw. formy dwuliniowej. Dla dowolnej przestrzeni liniowej K^n oraz wektorów $\alpha = (a_1, \dots, a_n)$, $\beta = (b_1, \dots, b_n)$ można określić operację

$$\langle \alpha, \beta \rangle = a_1b_1 + \dots + a_nb_n$$

analogicznie jak wyżej. Co więcej, jeśli $K = \mathbb{R}$, to operację zdefiniowaną wyżej nazywamy STANDARDOWYM ILOCZYNEM SKALARNYM. Nad tym ciałem można wykazać, że jeśli pewien układ niezerowych wektorów $\alpha_1, \dots, \alpha_n \in K^n$ spełnia warunek $\langle \alpha_i, \alpha_j \rangle = 0$, to jest to układ liniowo niezależny. Czym jest ogólny iloczyn skalarny i jak go można definiować (nie tylko w \mathbb{R}^n , ale dowolnej rzeczywistej przestrzeni liniowej), powiemy innym razem. W kontekście zarysowanym wyżej (nad ciałem \mathbb{Z}_2) teoria ta również może być zarysowana w większej ogólności, co pozwala wyprowadzić więcej obserwacji kombinatorycznych. \square

8.6 Coda. Kombinacje, czyli o przestrzeni barw

Pojęcia przestrzeni liniowej, kombinacji, podprzestrzeni rozpiętej na układzie i wreszcie układów liniowo zależnych i niezależnych — to spora dawka abstrakcji na przestrzeni dwóch tygodni zajęć. Warto poświęcić moment na kilka przyjaznych intuicji, niezupełnie zresztą odległych od rzeczywistości, mających bowiem głębokie podłożę historyczne, związane także z początkami algebry liniowej³.

Mówiąc w skrócie chodzi o mieszanie kolorów. Kolory możemy traktować jak wektory i uzyskiwać rozmaite ich kombinacje. Oto przykład kilku „kombinacji liniowych” koloru zielonego i czerwonego.



Rys. 2. Kombinacje liniowe $a\mathbf{z} + b\mathbf{r}$, dla $a = \frac{k}{10}$ oraz $b = 1 - a$, dla $a = \frac{1}{10}, \frac{9}{10}, \frac{8}{10}, \dots, \frac{2}{10}, \frac{1}{10}, \frac{0}{10}$.

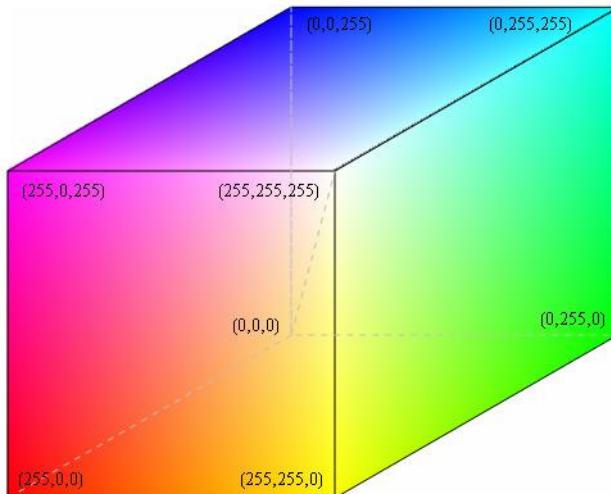
Czy kombinacją zielonego i czerwonego nie jest po prostu kolor żółty? Tak, ale żółty jest w istocie sumą $\mathbf{z} + \mathbf{r}$. Jak to wyjaśnić? Mowa tu mianowicie o modelu RGB „przestrzeni kolorów”, opartego o wyróżnienie trzech barw podstawowych: **czerwonej**, **zielonej** oraz **niebieskiej**, przypisanie im wektorów o współrzędnych

$$\mathbf{r} = (255, 0, 0), \quad \mathbf{z} = (0, 255, 0), \quad \mathbf{n} = (0, 0, 255)$$

lub — jeśli ktoś woli notację w zapisie szesnastkowym (heksagonalnym) — FF0000, 00FF00, 0000FF, i rozważanie wszystkich kombinacji liniowych tych wektorów postaci:

$$a\mathbf{r} + b\mathbf{z} + c\mathbf{n},$$

ale tylko w zakresie $a, b, c \in [0, 1]$. Oto reprezentacja przestrzeni barw RGB.



Rys. 3. Paleta barw RGB, źródło: <https://www.whydmath.org/node/wavlets/imagebasics.html>.

Wierzchołek sześciangu o współrzędnych $(255, 255, 255)$ odpowiada wektorowi który w przestrzeni barw RGB oznacza kolor biały. Wektor $(0, 0, 0)$ oznacza kolor czarny. Kolory mieszane przez nas wyżej „leżą” na przekątnej niewidocznej dla nas podstawy łączącej: czerwony i zielony wierzchołek. Natomiast wektor $\mathbf{z} + \mathbf{r}$ odpowiada wierzchołkowi $(255, 255, 0)$, widocznemu na naszym rysunku jako wierzchołek żółty.

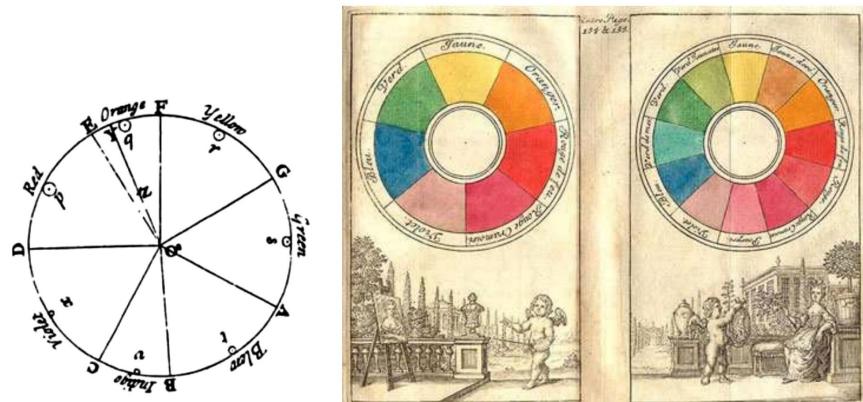
Nasz model posiada wiele wad, odróżniających go od przestrzeni liniowej i nic dziwnego — teoria kolorów i ich postrzegania ma olbrzymią historię i bardzo liczne modele (dziś głównie nielinowe). Mimo wszystko można coś powiedzieć i przy okazji wzmacnić swoje intuicje z algebry liniowej. Po pierwsze: mnożenie wektora przez skalar prowadzi do zmiany jego nasycenia. Podstawowy kolor określony jest przez kolory podstawowe. Zerowe nasycenie oznacza zawsze kolor czarny (a w innych modelach — biały). Oto przykład dla skalarnych wielokrotności wektora \mathbf{r} , postaci $a \cdot \mathbf{r}$, dla $a \in \{\frac{0}{10}, \frac{1}{10}, \frac{2}{10}, \dots, \frac{8}{10}, \frac{9}{10}, \frac{10}{10}\}$.



³Zachęcam też do lektury tekstu: <https://scholar.harvard.edu/files/schwartz/files/lecture17-color.pdf>

W naszym modelu nie istnieje wektor $1,5r$. Liniowa suma dwóch kolorów reprezentowanych jak wyżej nie musi być kolorem. Czy ta teoria ma sens? Przecież mamy doświadczenie mówiące, że każde kolory można zmieszać. Na czym więc polega mieszanie kolorów? Pytania tego typu stawiał już w starożytności Platon i jego uczeń Arystoteles, a nawet i wcześniejszy autorzy: Philolaus ze szkoły pitagorejskiej, Plutarch, Empedokles, a nawet Demokryt, uważany za twórcę teorii atomicznej budowy świata. Chodziło oczywiście o wyróżnienie barw podstawowych, mających również zastosowania sakralne — białego (leukhyn), czarnego (melan), czerwonego (erydron), zielonego (khloron) oraz uzyskiwania innych kolorów jako ich mieszanin. Dla filozofów greckich podstawowe kolory reprezentowały raczej fundamentalne właściwości materii — a ich mieszanie miało odwzorowywać złożone właściwości obiektów mających uzyskany kolor.

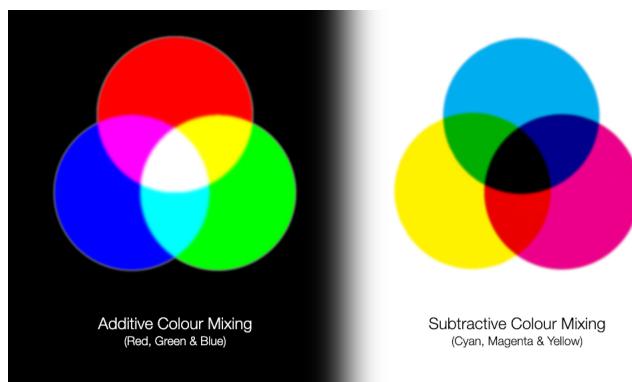
Z naukowego punktu widzenia przełomem były badania Newtona rozpoczęte w roku 1665. W dziele *Optyka* Newton wyjaśnia zjawisko znane ludzkości od zawsze pod symbolem tęczy — światło białe rozszczepia się na siedem rozróżnialnych barw, które reprezentować można na tzw. okręgu barw.



Rys. 4. Tarcza kolorów Newtona (Optyka, 1704) wzorowana na muzycznym kole kwintowym służącym do zmiany tonacji. Zauważcie Państwo, że każdy kolor jest mieszaniną barw. Obok popularyzująca je reprodukcja malarza Claude Bouteta (1708),

Słynny uczony nie tylko dokonał rozszczepienia światła widzialnego za pomocą pryzmatu. To jest nieco bardziej skomplikowane i istotne dla naszych intuicji. Newton rozumiał, że widzialne (białe) światło jest kombinacją fal różnej długości. W swoim przełomowym dziele *Optyka* zaprzeczył powszechniej opinii uczonych, jakoby to zanieczyszczenie pryzmatu powodowało tęczę. Użył w tym celu systemu soczewek, luster i pryzmatów, dzięki którym odizolował światło czerwone i skierował je na pryzmat, który wypuścił światło czerwone. Następnie rozbił światło widzialne na kolory tęczy i skierował je z powrotem na pryzmat, uzyskując z powrotem światło widzialne. Gorące dyskusje wokół teorii Newtona trwały 200 lat.

Co to ma wszystko wspólnego z algebrą liniową? Aby to zrozumieć, trzeba sięgnąć znowu do historii. Demokryt twierdził, iż obserwowane obiekty wysyłają do oka „atomy” wywołujące obraz. Teoria Newtona rozszczepiania światła spotkała się z dużym sprzeciwem m.in. Goethego, który nie przyjmował idei uzyskiwania bieli z barw chromatycznych, myśląc wyłączenie o syntezie subtraktywnej, czyli efektach odbicia światła od powierzchni pokrytych mieszaninami barwników, pochłaniających fale o różnych długościach (nakładanie się tych efektów powoduje wzrost udziału czerni). Różnice między syntezą subtraktywną i addytywną stała się znane dzięki pracom Tobiasa Mayera, autora pracy *De Affinitate Colorum Commentatio* z 1772 roku. Zobaczmy na obrazku czym różni się subtraktywne i addytywne mieszanie.



Rys. 5. Subtraktywne i addytywne mieszanie kolorów w systemach RGB i CMYK.

Aby wylądować wreszcie w algebrze liniowej brakuje nam postaci Grassmanna, dla którego teoria koloru była również osobistą pasją i jedną z motywacji do rozpatrywania przestrzeni liniowych. Skąd to się wzięło? W 1802 roku Thomas Young sformułował tzw. teorię trójchromatyczną, która miała wyjaśniać widzenie kolorów dzięki obecności w siatkówce człowieka i innych naczelnego trzech różnych fotoreceptorów absorbujących światło w różnych zakresach długości. Już wcześniej formułowano empirycznie zasady mieszania barw, głównie w oparciu o syntezę subtraktyną. Odkrycie Newtona było przełomem w tym sensie, że za widzenie barw odpowiadają cechy światła, a nie widzianych obiektów. Young wchodzi do tej opowieści jako fizyk i co ważniejsze — lekarz. Uważał on, że jednakowe pobudzenie trzech rodzajów włókien wywołuje wrażenie bieli (co wspiera rozumienie addytywne — kolokwialnie mówiąc kolory nie składają się przy widzeniu do czarnego — czarny jest wtedy, gdy nie patrzymy). Teorię Younga wsparł w połowie lat 50. tych XX wieku niemiecki fizjolog, fizyk i filozof Hermann von Helmholtz. Uważał on, że reaktywność światłoczułych receptorów zależy od długości fali i jest największa wówczas, gdy długość fali odpowiada barwom podstawowym: czerwonej, zielonej i fioletowej. Helmholtz zdefiniował też trzy stosowane do dziś cechy otrzymywanych kolorów: barwa (hue), nasycenie (saturation) oraz jasność (value), dając źródło dla tzw. przestrzeni barw HSV. Według tego modelu wszystkie barwy wywodzą się ze światła białego, gdzie część widma jest odbita a pozostała pochłonięta przez oświetlane przedmioty.

Wszystkie te idee wsparcie zostały wreszcie przez matematykę. W 1853 roku Hermann Grassmann wydał pracę pt. *Teoria mieszania kolorów*. Zaciekawiony zależnościami mieszania kolorów typu:

$$\text{czerwony} + \text{niebieski} = \text{fioletowy}$$

stwierdził, że gdy dysponujemy takim równaniem, można do każdej jego strony dodać kolor i wciąż uzyskamy prawdziwą równość! Np. do obydwu stron równości wyżej można dodać żółty i dostać:

$$\text{czerwony} + \text{niebieski} + \text{żółty} = \text{fioletowy} + \text{żółty}$$

Prawo to, nam się kojarzące z łącznością, zwane jest w kolorymetrii **trzecim prawem Grassmanna**. Dlaczego to jest takie ważne? Grassmann postulował, że barwa mieszaniny zależy jedynie od barw podstawowych jej składników, a nie od ich składu widmowego oraz, że można wybierać różne „układy kolorów podstawowych”, za pomocą których można uzyskać dowolne kolory. Oto **pierwsze prawo Grassmanna**.

Każda dowolnie wybrana barwa może być określona za pomocą trzech liniowo niezależnych barw.

Inaczej: każde cztery barwy są liniowo zależne.⁴

Czytelnik ewidentnie zobaczy w tych sformułowaniach algebrę liniową. Model Grassmanna został oczywiście poddany wielu ulepszeniom i krytyce, ale na nasz użytek może dobrze opisywać intuicję liniowej niezależności i kombinacji liniowych. Można zastosować intuicyjne analogie uznając, że chcemy uzyskać kolory jako kombinacje liniowe innych (niekoniecznie podstawowych, tak jak w przestrzeni liniowej są różne układy rozpinające). Oto przykład takiej sytuacji w modelu RGB.

Rozważmy zbiór liniowych kombinacji wektorów $\text{lin}(v_1 = (255, 85, 0), v_2 = (85, 85, 0), v_3 = (50, 80, 50))$.



Oto przykład kombinacji liniowej $\frac{1}{3}v_1 + v_2 + \frac{3}{2}v_3$ tych kolorów:

$$\frac{1}{3} \cdot \text{Czerwony} + 1 \cdot \text{Zielony} + \frac{3}{2} \cdot \text{Ciemnoniebieski} = \text{Brąz} + \text{Zielony} + \text{Ciemnoniebieski} = \text{Czerwony}$$

Czy trzy wymienione kolory tworzą układ rozpinający przestrzeń kolorów? Gdybyśmy mogli stosować ujemne nasycenie (a to się robi) — to owszem tak. Są to trzy wektory liniowo niezależne (żaden nie jest kombinacją pozostałych) i gdybyśmy mieli możliwość stosowania kombinacji liniowych z ujemnymi współczynnikami, wówczas wygenerowalibyśmy z nich dowolne kolory. Wszystko to jedynie zbiór intuicji — ważnych jednak z praktycznego punktu widzenia. Dziś teoria kolorów jest znacznie bardziej skomplikowana i stanowi osobną dziedzinę wiedzy (fizyki i chemii). Warto pamiętać, że ma ona duży styk także z algebrą liniową.

⁴Prawa te zostały ustalone dla części środkowej siatkówki oka człowieka. Prawo drugie i trzecie dotyczy również zwierząt. Prawo pierwsze zachowuje słuszność tylko w postaci uogólnionej, ponieważ maksymalna liczba barw liniowo niezależnych może być większa lub mniejsza od trzech. Istnieją ludzie i zwierzęta, dla których rejestrówana maksymalna liczba liniowo niezależnych barw wynosi dwa lub jeden, co powoduje, że niektórzy widzą więcej barw, inni mniej, a jeszcze inni rejestrują tylko szarości. Źródło: <https://slownikzprepressu.weebly.com/grassmanna-prawa.html>.

Rozdział 9

Baza przestrzeni liniowej

9.1 Wykład 9

Na ostatnim wykładzie wyróżniliśmy układy wektorów liniowo niezależnych i pokazaliśmy, że podprzestrzeni rozpiętej przez taki układ wektorów nie można rozpiąć przez żaden jego podkład oraz, że każdy wektor należący do podprzestrzeni rozpiętej na liniowo niezależnym układzie wektorów, tworzy z tymi wektorami układ liniowo zależny. Obserwacje te prowadzą naturalnie do następującej definicji.

Definicja 9.1.1: Baza (skończona) przestrzeni liniowej

Układ $\alpha_1, \dots, \alpha_k$ wektorów przestrzeni V nazywamy BAZĄ PRZESTRZENI V , jeśli spełnia on następujące dwa warunki:

- (a) układ $\alpha_1, \dots, \alpha_k$ jest liniowo niezależny,
- (b) układ $\alpha_1, \dots, \alpha_k$ ROZPINIA V , to znaczy $V = \text{lin}(\alpha_1, \dots, \alpha_k)$.

Wniosek 8.1.4 oznacza, że każda przestrzeń rozpięta na skończonym układzie wektorów posiada bazę. Na kolejnym wykładzie wywnioskujemy z twierdzenia Steinitza, że dwie skończone bazy przestrzeni liniowej są równoliczne, co doprowadzi nas do pojęcia wymiaru przestrzeni liniowej. Za cztery wykłady natomiast przedstawimy twierdzenie mówiące, że każda przestrzeń liniowa ma bazę.

Przykład 1. Układ $1, i$ jest bazą przestrzeni liniowej \mathbb{C} nad ciałem \mathbb{R} . Każda liczba zespolona jest w istocie postaci $a + bi$, gdzie $a, b \in \mathbb{R}$, więc $\text{lin}(1, i) = \mathbb{C}$. Jeśli zaś $a \cdot 1 + b \cdot i = 0$, to $a = 0$ oraz $b = 0$, więc układ $1, i$ jest liniowo niezależny. Analogicznie argumentujemy, że układ $1, \sqrt{2}$ jest bazą $\mathbb{Q}(\sqrt{2})$ nad \mathbb{Q} .

Przykład 2. W przestrzeni K^n rozważmy układ wektorów $\epsilon_1, \dots, \epsilon_n$, zdefiniowany w następujący sposób, dla $i = 1, \dots, n$:

$$\epsilon_i = (a_1, \dots, a_n), \quad \text{gdzie } a_j = \begin{cases} 1, & j = i, \\ 0, & j \neq i. \end{cases}$$

Na przykład dla $n = 3$ mamy $\epsilon_1 = (1, 0, 0), \epsilon_2 = (0, 1, 0), \epsilon_3 = (0, 0, 1)$.

Układ $\epsilon_1, \dots, \epsilon_n$ jest bazą przestrzeni K^n , zwana BAZĄ STANDARDOWĄ przestrzeni K^n . Rzeczywiście:

- układ $\epsilon_1, \dots, \epsilon_n$ jest liniowo niezależny, bowiem jeśli $a_1\epsilon_1 + \dots + a_n\epsilon_n = (0, \dots, 0)$, to zgodnie z działaniami w K^n mamy: $(a_1, a_2, \dots, a_n) = (0, \dots, 0)$. A zatem $a_1 = 0, a_2 = 0, \dots, a_n = 0$.
- Układ $\epsilon_1, \dots, \epsilon_n$ rozpina K^n , gdyż dowolny wektor (x_1, x_2, \dots, x_n) należy do $\text{lin}(\epsilon_1, \dots, \epsilon_n)$, skoro $(x_1, \dots, x_n) = x_1(1, 0, 0, \dots, 0) + x_2(0, 1, 0, \dots, 0) + \dots + x_n(0, 0, 0, \dots, 1) = x_1\epsilon_1 + \dots + x_n\epsilon_n$.

Przykład 3. Niech $V = \{(x_1, x_2, x_3) \in \mathbb{R}^3 \mid x_1 + 2x_2 - x_3 = 0\}$, czyli $(x_1, x_2, x_3) \in V$ wtedy i tylko wtedy, gdy $x_1 = -2x_2 + x_3$. Elementy przestrzeni liniowej V są zatem postaci:

$$(-2x_2 + x_3, x_2, x_3) = (2x_2, x_2, 0) + (x_3, 0, x_3) = x_2(-2, 1, 0) + x_3(1, 0, 1), \quad \text{gdzie } x_2, x_3 \in \mathbb{R}.$$

Stąd $V = \text{lin}((-2, 1, 0), (1, 0, 1))$. Wektory $(-2, 1, 0), (1, 0, 1)$ są oczywiście liniowo niezależne (bo jeśli $a(-2, 1, 0) + b(1, 0, 1) = (0, 0, 0)$, to łatwo widzieć, że $a = b = 0$), a zatem układ ten jest bazą V .

Przykład 4. Niech $W = \text{lin}((1, 2, 1), (0, 1, 1), (1, 3, 2))$ będzie podprzestrzenią \mathbb{R}^3 . Jest to przestrzeń rozpięta przez 3 wektory, ale nie jest to „oszczędny” układ. Wektor $(1, 3, 2)$ jest kombinacją liniową $(1, 2, 1), (0, 1, 1)$. A zatem układ $(1, 2, 1), (0, 1, 1), (1, 3, 2)$ nie jest bazą W . Jest nią natomiast układ $(1, 2, 1), (0, 1, 1)$, układ $(1, 2, 1), (1, 3, 2)$, i wiele innych.

Przykład 5. W przestrzeni liniowej macierzy rozmiaru $m \times n$ o wyrazach w ciele K rozważmy zbiór $m \cdot n$ macierzy E_{ij} , dla $1 \leq i \leq m$ oraz $1 \leq j \leq n$, takich że w i -tym wierszu i j -tej kolumnie macierzy E_{ij} stoi wyraz 1, a na pozostałych miejscach – 0. Innymi słowy jeśli E_{ij} jest macierzą o wyrazach a_{kl} , to:

$$a_{kl} = \begin{cases} 1, & \text{gdy } (i, j) = (k, l), \\ 0, & \text{gdy } (i, j) \neq (k, l) \end{cases}.$$

Wówczas nietrudno pokazać, naśladowując argumentację dla przestrzeni K^n , że układ

$$\{E_{ij}, 1 \leq i \leq m, 1 \leq j \leq n\}$$

jest bazą przestrzeni $M_{m \times n}(K)$. Bazę tę nazywamy JEDYNKAMI MACIERZOWYMI rozmiaru $m \times n$.

Dla przykładu, w przestrzeni $M_{2 \times 3}(K)$ zbiór jedynek macierzowych ma postać:

$$E_{11} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \quad E_{12} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \quad E_{13} = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix},$$

$$E_{21} = \begin{bmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix}, \quad E_{22} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}, \quad E_{23} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

Przykład 6. Rozważmy zbiór F ciągów (x_n) z K^∞ spełniający dla każdego $n \geq 0$ warunki

$$x_{n+2} - x_{n+1} - x_n = 0.$$

Nietrudno widzieć, że F jest podprzestrzenią K^∞ . Zauważmy, że biorąc ciągi $(a_n), (b_n) \in F$ spełniające warunki $a_1 = 1, a_2 = 0$ oraz $b_1 = 0, b_2 = 1$, otrzymujemy układ liniowo niezależny. Co więcej, każdy ciąg w F zadany jest jednoznacznie przez określenie pierwszych dwóch wyrazów, więc $(a_n), (b_n)$ jest bazą F . Warto wskazać inne bazy tej przestrzeni, mające szczególne własności, choćby złożone z ciągów geometrycznych $(\lambda_n), (\gamma_n)$. Do przykładu tego wróćmy w Uzupełnieniu 9.4.

Uwaga 9.1.2

Rozpatrzmy jednorodny układ równań liniowych o n zmiennych i o współczynnikach z ciała K

$$U : \begin{cases} a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n = 0 \\ \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \dots + a_{mn}x_n = 0 \end{cases}$$

Niech $V \subset K^n$ będzie przestrzenią rozwiązań układu U i niech

$$U' : \begin{cases} x_{j_1} = c_{11}x_1 + \dots + c_{1n}x_n \\ \vdots \\ x_{j_k} = c_{k1}x_1 + \dots + c_{kn}x_n \end{cases}$$

zadaje rozwiązanie ogólne układu U , o zmiennych zależnych x_{j_1}, \dots, x_{j_k} , gdzie $j_1 < j_2 < \dots < j_k$ i parametrach $x_{t_1}, \dots, x_{t_{n-k}}$, gdzie $t_1 < t_2 < \dots < t_{n-k}$. Dla każdego $s = 1, \dots, n-k$ niech $\alpha_s = (x_1, \dots, x_n) \in K^n$ będzie rozwiązaniem powyższego układu powstały przez

- podstawienie 1 za parametr x_{t_s} ,
- podstawienie 0 za wszystkie pozostałe parametry x_{t_r} , gdzie $r \neq s$,
- wyliczenie pozostałych zależnych od nich współrzędnych x_{j_i} ze wzorów układu U' .

Wówczas układ $\alpha_1, \dots, \alpha_{n-k}$ jest bazą przestrzeni V .

Dowód. Uzasadnimy, że układ $\alpha_1, \dots, \alpha_{n-k}$ jest liniowo niezależny. Istotnie, jeśli dla $b_1, \dots, b_{n-k} \in K$ mamy $b_1\alpha_1 + \dots + b_{n-k}\alpha_{n-k} = 0$, to dla i -tych współrzędnych $a_{1i}, \dots, a_{n-k,i}$ wektorów $\alpha_1, \dots, \alpha_{n-k}$ mamy

$$b_1a_{1i} + \dots + b_{n-k}a_{n-k,i} = 0.$$

Zatem biorąc i równe kolejno t_1, \dots, t_{n-k} dostajemy $b_1 = \dots = b_{n-k} = 0$.

Układ $\alpha_1, \dots, \alpha_{n-k}$ rozpina zbiór rozwiązań układu U' . Jest bowiem jasne, że wektor

$$\alpha = (a_1, \dots, a_n)$$

spełnia układ U' wtedy i tylko wtedy, gdy każda z jego współrzędnych jest kombinacją liniową wspólnych współrzędnych $a_{t_1}, \dots, a_{t_{n-k}}$. Inaczej mówiąc α jest rozwiązaniem układu U' wtedy i tylko wtedy, gdy

$$\alpha = a_{t_1}\alpha_1 + \dots + a_{t_{n-k}}\alpha_{n-k}.$$

Zatem zbiór rozwiązań układu U' równy jest $\text{lin}(\alpha_1, \dots, \alpha_{n-k})$. \square

Przykład. Rozwiązanie ogólne układu U' zmiennych x_1, x_2, x_3, x_4 o współczynnikach w \mathbb{R} postaci:

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

ma 4 – 2 parametry x_3, x_4 . Każde rozwiązanie U' jest postaci:

$$(-s - t, 0, s, t) = s(-1, 0, 1, 0) + t(-1, 0, 0, 1), \quad s, t \in \mathbb{R},$$

i powstaje przed odpowiedni wybór s, t na współrzędnych $t_1 = 3, t_2 = 4$. W zatem w tym przykładzie:

$$\alpha_1 = (a_{11}, a_{12}, a_{13}, a_{14}) = (-1, 0, 1, 0), \quad \alpha_2 = (a_{21}, a_{22}, a_{23}, a_{24}) = (-1, 0, 0, 1).$$

Twierdzenie 9.1.3

Niech $\alpha_1, \dots, \alpha_k$ będzie układem wektorów przestrzeni V . Wówczas następujące warunki są równoważne:

- (1) układ $\alpha_1, \dots, \alpha_k$ jest bazą przestrzeni V ,
- (2) każdy wektor $\alpha \in V$ można przedstawić w sposób jednoznaczny jako kombinację liniową układu $\alpha_1, \dots, \alpha_k$.

Dowód. Zaczniemy od uzasadnienia implikacji (1) \Rightarrow (2). Niech $\alpha_1, \dots, \alpha_k$ będzie bazą przestrzeni V . Wówczas $V = \text{lin}(\alpha_1, \dots, \alpha_k)$, więc każdy $\alpha \in V$ jest kombinacją układu $\alpha_1, \dots, \alpha_k$. Pozostaje wykazać jednoznaczność. Gdyby:

$$\alpha = a_1\alpha_1 + \dots + a_k\alpha_k = a'_1\alpha_1 + \dots + a'_k\alpha'_k,$$

dla pewnych $a_1, \dots, a_k, a'_1, \dots, a'_k \in K$, to mielibyśmy:

$$(a_1 - a'_1)\alpha_1 + \dots + (a_k - a'_k)\alpha_k = 0.$$

Z liniowej niezależności wektorów $\alpha_1, \dots, \alpha_k$ wynikałoby zatem, że $a_1 - a'_1 = \dots = a_k - a'_k = 0$. A zatem rozkład każdego $\alpha \in V$ jest jednoznaczny.

Dowodzimy odwrotną implikację (2) \Rightarrow (1). Przypuśćmy, że każdy wektor $\alpha \in V$ można jednoznacznie przedstawić jako kombinację układu $\alpha_1, \dots, \alpha_k$. Wykażemy, że $\alpha_1, \dots, \alpha_k$ jest bazą.

Oczywiście skoro każdy wektor z V jest kombinacją liniową wektorów $\alpha_1, \dots, \alpha_k$, to układ ten rozpina V . A zatem warunek (b) z definicji bazy jest spełniony. Pozostaje pokazać, że układ $\alpha_1, \dots, \alpha_k$ jest liniowo niezależny. Przypuśćmy, że $a_1\alpha_1 + \dots + a_k\alpha_k = 0$, dla pewnych $a_1, \dots, a_k \in K$. Wówczas mamy:

$$a_1\alpha_1 + \dots + a_k\alpha_k = 0\alpha_1 + \dots + 0\alpha_k = 0,$$

a skoro także wektor 0 ma jednoznaczny rozkład w V jako kombinacja liniowa $\alpha_1, \dots, \alpha_k$, to uzyskujemy $a_1 = 0, a_2 = 0, \dots, a_k = 0$. Stąd układ $\alpha_1, \dots, \alpha_k$ jest liniowo niezależny. \square

Definicja 9.1.4

Niech V będzie przestrzenią liniową nad ciałem K i niech układ $\alpha_1, \dots, \alpha_k$ będzie bazą przestrzeni V . WSPÓŁRZĘDNYMI WEKTORA $\alpha \in V$ W BAZIE $\alpha_1, \dots, \alpha_k$ nazywamy układ elementów a_1, \dots, a_k ciała K spełniających

$$\alpha = a_1\alpha_1 + \dots + a_k\alpha_k.$$

Przykłady:

- Wektor $(1, 2, 1)$ ma współrzędne $1, 2, 1$ w bazie standardowej przestrzeni \mathbb{R}^3 .
- Wektor $(1, 2, 1)$ ma współrzędne $-1, 1, 1$ w bazie $\{(1, 0, 0), (1, 1, 0), (1, 1, 1)\}$ przestrzeni \mathbb{R}^3 , bo $(1, 2, 1) = -1(1, 0, 0) + 1(1, 1, 0) + 1(1, 1, 1)$.
- Układ $(1, 0), (2, 0)$ nie jest bazą $V = \text{lin}((1, 0))$, bo mamy $(1, 0) = 1 \cdot (1, 0) = 1 \cdot (2, 0) + (-1) \cdot (1, 0)$.

Definicja 9.1.5

Mówimy, że układ $\alpha_1, \dots, \alpha_k$ wektorów przestrzeni V jest

- MAKSYMALNYM UKŁADEM LINIOWO NIEZALEŻNYM, jeśli $\alpha_1, \dots, \alpha_k$ jest liniowo nieszależny i każdy większy układ – to znaczy taki układ wektorów przestrzeni V , który zawiera $\alpha_1, \dots, \alpha_k$ jako podukład właściwy – jest liniowo zależny,
- MINIMALNYM UKŁADEM ROZPINAJĄCYM V , jeśli $\alpha_1, \dots, \alpha_k$ rozpinia V i żaden mniejszy układ – to znaczy żaden podukład właściwy układu $\alpha_1, \dots, \alpha_k$ – nie rozpinia V .

Przykłady.

- Układ $(1, 0, 0), (1, 1, 0)$ nie jest maksymalnym układem liniowo nieszależnym \mathbb{R}^3 . Jest to bowiem podukład właściwy układu liniowo nieszależnego $(1, 0, 0), (1, 1, 0), (1, 1, 1)$.
- Układ $(1, 0, 0), (2, 0, 0)$ nie jest minimalnym układem rozpinającym $V = \text{lin}((1, 0, 0), (2, 0, 0)) \subset \mathbb{R}^3$, bo układ $(1, 0, 0)$ jest jego podukładem właściwym i $V = \text{lin}((1, 0, 0))$.

Twierdzenie 9.1.6

Niech $\alpha_1, \dots, \alpha_k$ będzie układem wektorów przestrzeni liniowej V . Następujące warunki są równoważne.

- $\alpha_1, \dots, \alpha_k$ jest bazą przestrzeni V ,
- $\alpha_1, \dots, \alpha_k$ jest maksymalnym układem liniowo nieszależnym w V ,
- $\alpha_1, \dots, \alpha_k$ jest minimalnym układem rozpinającym V .

Dowód. Zaczniemy od implikacji $(i) \Rightarrow (ii)$. Niech $\alpha_1, \dots, \alpha_k$ będzie bazą przestrzeni V . Gdyby układ $\alpha_1, \dots, \alpha_k$ nie był maksymalny, to istniałby taki wektor $\alpha \in V$, że układ $\alpha_1, \dots, \alpha_k, \alpha$ byłby liniowo nieszależny. Wówczas jednak, zgodnie z Uwagą 8.1.5, wektor α nie mógłby należeć do $\text{lin}(\alpha_1, \dots, \alpha_k)$. To jest jednak niemożliwe, bo $V = \text{lin}(\alpha_1, \dots, \alpha_k)$, zgodnie z definicją bazy. Zatem układ $\alpha_1, \dots, \alpha_k$ jest maksymalnym liniowo nieszależnym układem w V .

Dowód implikacji $(ii) \Rightarrow (i)$. Skoro $\alpha_1, \dots, \alpha_k$ jest maksymalnym układem liniowo nieszależnym w V , to jest on w szczególności liniowo nieszależny. Do pokazania, że $\alpha_1, \dots, \alpha_k$ jest bazą V pozostało wykazać, że $V = \text{lin}(\alpha_1, \dots, \alpha_k)$. Jednak z maksymalności tego układu wynika, że dla każdego wektora $\alpha \in V$ układ $\alpha_1, \dots, \alpha_k, \alpha$ jest liniowo zależny. W szczególności z Uwagi 8.1.5 wynika, że α jest kombinacją liniową wektorów $\alpha_1, \dots, \alpha_k$. A zatem istotnie $V = \text{lin}(\alpha_1, \dots, \alpha_k)$.

Dowód implikacji $(i) \Rightarrow (iii)$. Niech $\alpha_1, \dots, \alpha_k$ będzie (ponownie) bazą V . Gdyby $\alpha_1, \dots, \alpha_k$ nie był minimalnym układem rozpinającym V , to zawierałby podukład właściwy rozpinający V . Weźmy jednak

dowolny wektor spośród $\alpha_1, \dots, \alpha_k$ nie należący do tego podukładu. Jest on kombinacją liniową elementów tego podukładu, bo podukład ten rozpinia (ponoć) przestrzeń V . Daje to sprzeczność z liniową niezależnością układu $\alpha_1, \dots, \alpha_k$.

Dowód implikacji $(iii) \Rightarrow (i)$. Mamy minimalny układ $\alpha_1, \dots, \alpha_k$ rozpinający przestrzeń V . Pokażemy, że jest on bazą tej przestrzeni. Wystarczy pokazać liniową niezależność tego układu. Gdyby układ ten był liniowo zależny, to któryś z $\alpha_1, \dots, \alpha_k$ byłby liniową kombinacją pozostałych, na mocy Uwagi 8.1.3. Na przykład (po ewentualnym przenumerowaniu)

$$\alpha_k = b_1\alpha_1 + \dots + b_{k-1}\alpha_{k-1}.$$

Wówczas jednak

$$V = \text{lin}(\alpha_1, \dots, \alpha_{k-1}),$$

bo dla dowolnego $\alpha \in V$ istniałyby $a_1, \dots, a_k \in K$, że:

$$\begin{aligned} \alpha &= a_1\alpha_1 + \dots + a_k\alpha_k = \\ &= a_1\alpha_1 + \dots + a_k(\underbrace{b_1\alpha_1 + \dots + b_{k-1}\alpha_{k-1}}_{\alpha_k}) = \\ &= (a_1 + a_kb_1)\alpha_1 + \dots + (a_{k-1} + a_kb_{k-1})\alpha_{k-1}. \end{aligned}$$

Jest to jednak sprzeczne z założeniem, że $\alpha_1, \dots, \alpha_k$ jest minimalnym układem rozpinającym V . A zatem układ $\alpha_1, \dots, \alpha_k$ jest liniowo niezależny. \square

Na koniec przyjrzymy się ogólnej definicji bazy przestrzeni liniowej. W rozdziale 14 uzasadnimy jej istnienie, wywodząc ją z tzw. lematu Kuratowskiego-Zorna.

Definicja 9.1.7

Układ $X = \{\alpha_i\}_{i \in T}$ wektorów przestrzeni liniowej V nazywamy BAZĄ, jeśli układ ten jest liniowo niezależny oraz gdy $\text{lin}(X) = V$.

Układ $X = \{x^n \mid n \in \mathbb{N}\}$ jest bazą przestrzeni liniowej $K[x]$. Istotnie, liniowa niezależność wynika stąd, że jeśli $a_1 \cdot x^{i_1} + a_2 \cdot x^{i_2} + \dots + a_n x^{i_n}$ jest wielomianem zerowym, to oczywiście $a_1 = \dots = a_n$. Każdy wielomian jest sumą skończonej kombinacji liniowej elementów powyższego układu, a zatem jest on bazą.

Zauważmy, że ciągu $(1, 1, \dots) \in K^\infty$, którego wszystkie wyrazy są równe 1 nie można przedstawić jako kombinacji liniowej ciągów mających na jednej współrzędnej 1, a na pozostałej — zera. Zasadniczy problem powyższego przykładu polega nie na wskazywaniu dużych układów liniowo niezależnych w K^∞ , ale na tym, że te duże układy wskazane „wprost” są za małe by rozpinać całe K^∞ .

Szczególnym przypadkiem ogólnego twierdzenia o istnieniu bazy przestrzeni liniowej, którego dowód odkładamy, jest rezultat mówiący o istnieniu bazy ciała liczb rzeczywistych \mathbb{R} , jako przestrzeni liniowej nad \mathbb{Q} . Bazy te mają szczególne znaczenie w różnych dziedzinach matematyki i zwane są BAZAMI HAMELA.

Klasycznym i jednym z pierwszych zastosowań baz Hamela był opis wszystkich rozwiązań równania funkcyjnego Cauchy'ego, czyli zagadnienie wyznaczenia wszystkich funkcji $f : \mathbb{R} \rightarrow \mathbb{R}$, które mają własność addytywności, czyli dla każdych $x, y \in \mathbb{R}$ mamy $f(x + y) = f(x) + f(y)$.

W jaki sposób pomagają w tym problemie przestrzenie liniowe? Niech $B = \{b_i \mid i \in I\}$ będzie bazą Hamela. Rozważmy układ liczb rzeczywistych $\{c_i \mid i \in I\}$ indeksowanych tym samym zbiorem I . Twierdzimy, że istnieje dokładnie jedna funkcja addytywna f spełniająca $f(b_i) = c_i$.

Istnienie funkcji f jest jasne: skoro każda liczba rzeczywista zapisuje się jednoznacznie w postaci skończonej sumy $\sum \lambda_i b_i$, to jest jasne, że funkcja f zdefiniowana wzorem $f(\sum \lambda_i b_i) = \sum \lambda_i c_i$ jest addytywna. Z drugiej strony, jeśli funkcja f jest addytywna, to dla każdej liczby naturalnej n mamy $f(nx) = nf(x)$. Mamy też $f(0) = 0$. Skoro $f(-x) = -x$, to równość $f(nx) = nf(x)$ jest spełniona przez wszystkie liczby całkowite. Wreszcie, jeśli $\frac{p}{q}$ jest liczbą wymierną oraz $x \in \mathbb{R}$, to łatwo uzasadnić równość $f(\frac{p}{q}x) = \frac{p}{q}f(x)$. Czyli każda funkcja addytywna ma na bazie Hamela jednoznacznie wyznaczone wartości.

9.2 Wybór przykładowych pytań

Proszę się nie uczyć na pamięć odpowiedzi czy wyjaśnienie! Takich pytań można ułożyć znacznie więcej i służą one jedynie sprawdzeniu podstawowego zrozumienia treści zajęć. Jeśli odpowiedź na dowolne z poniższych pytań sprawia Państwu problem, proszę pomyśleć który fragment notatek dotyczy tego pytania, wrócić do ich lektury, a w razie dalszych trudności poprosić prowadzącego o wskazówkę.

1. Czy układ $(1, 0, 0), (1, 1, 0)$ tworzy bazę przestrzeni liniowej \mathbb{R}^3 ?
2. Czy układ $(1, 1, 2), (2, 2, 4), (3, 3, 6)$ tworzy bazę przestrzeni liniowej \mathbb{R}^3 ?
3. Czy układ $(1, 0, 0), (0, 1, 0), (1, 1, 1)$ tworzy bazę przestrzeni liniowej \mathbb{R}^3 ?
4. Czy układ $(1, 0, 1), (0, 1, 1), (1, 1, 0)$ tworzy bazę przestrzeni liniowej \mathbb{Z}_3 ?
5. Czy istnieje wektor α , który dopełnia układ $(1, 1, 1), (2, 2, 0)$ do bazy przestrzeni liniowej \mathbb{R}^3 ?
6. Czy istnieje wektor α , który dopełnia układ $(1, 1, 1), (2, 2, 2)$ do bazy przestrzeni liniowej \mathbb{R}^3 ?
7. Czy z układu $(1, 1, 1), (2, 2, 3), (3, 3, 4), (4, 4, 5)$ można wybrać bazę przestrzeni liniowej \mathbb{R}^3 ?
8. Jakie są współrzędne wektora $(1, 2, 3)$ w bazie standardowej przestrzeni liniowej \mathbb{R}^3 ?
9. Jakie są współrzędne wektora $(1, 2, 3)$ w bazie $(1, 0, 0), (1, 1, 0), (1, 1, 1)$ przestrzeni liniowej \mathbb{R}^3 ?
10. Czy istnieje baza przestrzeni liniowej \mathbb{R}^3 , w której wektor $(1, 2, 3)$ ma współrzędne $3, 2, -1$?
11. Czy istnieje baza przestrzeni liniowej \mathbb{R}^3 , w której wektor $(1, 2, 3)$ ma współrzędne $0, \pi, 0$?
12. Czy istnieje baza przestrzeni liniowej \mathbb{R}^3 , w której wektor $(1, 2, 3)$ ma współrzędne $1, 1, 1$?
13. Niech $\alpha_1, \alpha_2, \alpha_3$ będzie bazą przestrzeni liniowej \mathbb{R}^3 , w której wektor $(1, 1, 1)$ ma współrzędne $1, 2, 3$.
Jakie współrzędne w tej bazie ma wektor $(2, 2, 2) + \alpha_1$?
14. Zapisz macierz
$$\begin{bmatrix} 1 & 3 & -1 \\ 0 & 1 & -1 \end{bmatrix}$$
jako kombinację liniową jedynek macierzowych.
15. Niech $\alpha_1, \alpha_2, \alpha_3$ będzie bazą przestrzeni liniowej. Jakie współrzędne ma w tej bazie wektor $\alpha_2 - \alpha_1$?
16. Niech układ $\{\alpha_1, \alpha_2\}$ będzie bazą przestrzeni liniowej V . Współrzędne wektora β w tej bazie wynoszą $1, 0$ zaś współrzędne wektora γ wynoszą $2, 0$. Czy wektory β, γ mogą być liniowo niezależne?
17. Niech układ $\{\alpha_1, \alpha_2, \alpha_3\}$ będzie bazą przestrzeni liniowej V . Współrzędne wektora α w tej bazie wynoszą $1, 0, 0$, współrzędne wektora β wynoszą $0, 1, 0$ zaś współrzędne wektora γ wynoszą $1, 1, 0$. Czy wektory α, β, γ mogą być liniowo niezależne?
18. Niech układ $\{\alpha_1, \alpha_2\}$ będzie bazą przestrzeni liniowej V . Suma współrzędnych niezerowego wektora β wynosi 0 . Suma współrzędnych niezerowego wektora γ też wynosi 0 . Czy układ $\{\beta, \gamma\}$ może być bazą V ?
19. Wektory α, β należą do przestrzeni liniowej V . Wiadomo, że $a_1\alpha + b_1\beta = a_2\alpha + b_2\beta$ oraz $a_1 \neq a_2$.
Czy układ $\{\alpha, \beta\}$ może być bazą V ?
20. Niech α_1, α_2 będzie bazą przestrzeni liniowej V . Wektor β nie należy do $\text{lin}(\alpha_2)$. Czy wynika stąd, że układ β, α_2 jest bazą V ?
21. Układ $\{\alpha, \beta\}$ jest bazą przestrzeni liniowej V nad ciałem K . Czy wynika stąd, że układ $\{\alpha, \alpha + \beta\}$ jest bazą V ?
22. Układ $\{\alpha, \beta\}$ jest bazą przestrzeni liniowej V nad ciałem K . Czy wynika stąd, że układ $\{\alpha - \beta, \alpha + \beta\}$ jest bazą V ? Co jeśli $K = \mathbb{Z}_2$?
23. Układ $\{\alpha, \beta\}$ jest bazą przestrzeni liniowej V nad ciałem K . Czy wynika stąd, że układ $\{\alpha, \alpha + \beta, \alpha + \beta + \gamma\}$ jest bazą V ?
24. Czy układ $\{1, 1+x, 1+x+x^2\}$ jest bazą przestrzeni wielomianów nad \mathbb{K} stopnia mniejszego od 3 ?
25. Niech $\mathcal{A} = \{\alpha_1, \alpha_2, \alpha_3\}$ będzie bazą przestrzeni liniowej V . Czy układ $\{\alpha_1 + \alpha_2 + \alpha_3\}$ można uzupełnić do bazy wektorami z \mathcal{A} ? Na ile sposobów można to zrobić?

9.3 Zadania do samodzielnej pracy

Zadania ilustrujące umiejętności niezbędne do zaliczenia przedmiotu oznaczone są symbolem ♠.

Przy tych zadaniach dodany jest opis umiejętności, które sprawdzają.

1. (♠ Wybór bazy jako maksymalnego układu liniowo niezależnego)
Znajdź bazy następujących podprzestrzeni rozpiętych przez układy wektorów w \mathbb{R}^n .
 - a) $\text{lin}((2, 1, 4), (3, 5, -1), (3, -2, 13), (7, 7, 7), (-4, -9, 6))$,
 - b) $\text{lin}((3, 2, 1, 1), (5, 0, 2, 3), (4, 1, 4, 5), (4, 1, -1, -1))$,
 - c) $\text{lin}((2, 7, -1, 2, 6), (3, 1, 4, 2, 2), (4, -5, 9, 2, -2), (5, 15, 2, 6, 14))$.
2. Niech $v_1 = (1, 2, 0), v_2 = (0, 1, 2), v_3 = (2, 0, 1)$. Dla jakich liczb pierwszych p układ $\{v_1, v_2, v_3\}$ jest bazą przestrzeni \mathbb{Z}_p^3 ?
3. Znajdź bazę podprzestrzeni $\mathbb{R}[x]$ postaci $\text{lin}(x^2 - x + 4, x - 1, x^2 + x)$.
4. Znajdź bazę podprzestrzeni $W \subseteq M_{2 \times 2}(\mathbb{R})$ opisanej w następujący sposób:

$$W = \left\{ \begin{bmatrix} a & b \\ c & d \end{bmatrix}, \text{ gdzie } a + b = c, b + c = d, c + d = a \right\}.$$
5. (♠ Znajdowanie bazy przestrzeni rozwiązań jednorodnego układu)
Znajdź bazę przestrzeni rozwiązań następujących układów równań liniowych w \mathbb{R}^n
 - a) $\begin{cases} 9x_1 + 12x_2 + 2x_3 = 0 \\ 5x_1 + 6x_2 + 4x_3 = 0 \\ 2x_1 + 3x_2 - x_3 = 0 \end{cases}$
 - b) $\begin{cases} 5x_1 + 10x_2 + 6x_3 + 3x_4 = 0 \\ 2x_1 + 4x_2 + 4x_3 + 3x_4 = 0 \\ 3x_1 + 6x_2 + 5x_3 + 5x_4 = 0 \end{cases}$
 - c) $\begin{cases} 7x_1 + 3x_2 + 5x_3 + 2x_4 + 8x_5 = 0 \\ 3x_1 + x_2 + x_3 - 4x_4 + 6x_5 = 0 \\ 2x_1 + x_2 + 2x_3 + 3x_4 + x_5 = 0 \end{cases}$
6. (♠ Znajdowanie współrzędnych wektora w danej bazie)
Rozpatrzmy następujące wektory przestrzeni \mathbb{R}^3 :

$$\alpha_1 = (3, 2, 1), \quad \alpha_2 = (7, 3, 1), \quad \alpha_3 = (4, 2, 1), \quad \beta_1 = (0, 2, 1), \quad \beta_2 = (1, 1, 2), \quad \beta_3 = (1, 0, 0).$$
 - a) Wykaż, że $\alpha_1, \alpha_2, \alpha_3$ jest bazą przestrzeni \mathbb{R}^3 . Dla $i = 1, 2, 3$ znajdź współrzędne wektora β_i w tej bazie.
 - b) Podaj przykład takiej bazy przestrzeni, że wektor β_1 ma w tej bazie współrzędne $1, 2, -1$, a wektor β_2 ma współrzędne $0, 0, 1$.
 - c) Czy istnieje taka baza przestrzeni \mathbb{R}^3 w której wektor β_1 ma współrzędne $1, 1, 0$, wektor β_2 ma współrzędne $0, 0, 1$ a wektor β_3 ma współrzędne $1, 1, 1$?
7. Założmy, że wektory $\alpha_1, \alpha_2, \alpha_3$ stanowią bazę pewnej przestrzeni V nad ciałem \mathbb{R} . Rozpatrzmy wektor $\beta = -\alpha_1 + 7\alpha_2 - 3\alpha_3$. Rozstrzygnij, czy wektory $\alpha_1, \alpha_2, \beta$ stanowią bazę przestrzeni V ?
8. Niech $\alpha_1, \dots, \alpha_n$ będzie bazą przestrzeni V i niech wektor $\beta \in V$ ma w tej bazie wszystkie współrzędne równe 1. Czy układ $\alpha_1, \dots, \alpha_{n-1}, \beta$ też jest bazą przestrzeni V ?
9. Niech $\mathcal{A} = \{\alpha_1, \alpha_2, \alpha_3\}$ oraz $\mathcal{B} = \{\beta_1, \beta_2, \beta_3\}$ będą bazami przestrzeni liniowej V . Czy dla każdego wektora $\alpha \in \mathcal{A}$ istnieje wektor $\beta \in \mathcal{B}$, że układ $(\mathcal{A} \setminus \{\alpha\}) \cup \{\beta\}$ jest bazą V ?
10. Niech $\mathcal{A} = \{\alpha_1, \alpha_2, \alpha_3\}$ będzie bazą przestrzeni liniowej V . Czy układ $\{\alpha_1, \alpha_1 + \alpha_2\}$ można uzupełnić do bazy wektorami z \mathcal{A} ? Na ile sposobów można to zrobić?
11. Niech $\mathcal{A} = \{\alpha_1, \alpha_2, \alpha_3\}$ będzie bazą przestrzeni liniowej V . Czy układ $\{\alpha_1 + \alpha_3, \alpha_1 + \alpha_2\}$ można uzupełnić do bazy wektorami z \mathcal{A} ? Na ile sposobów można to zrobić?
12. Niech $\alpha_1, \dots, \alpha_n$ będzie bazą przestrzeni V nad ciałem K . Niech $\beta_j = \sum_{i=1}^j \alpha_i$ dla $j = 1, \dots, n$. Wykaż, że układ β_1, \dots, β_n jest bazą przestrzeni V .

9.4 Uzupełnienie. Rekurencje liniowe. Wzór Bineta.

Czy istnieje odpowiednik pojęcia jednorodnego układu równań liniowych w przypadku przestrzeni nieskończonego wymiaru? Choć często nie myślimy o tym w ten sposób – jest pojęcie bliskie tej intuicji.

Definicja 9.4.1

LINIOWYM JEDNORODNYM RÓWNANIEM REKURENCYJNYM RZĘDU k (lub REKURENCJĄ rzędu k) o współczynnikach nad ciałem K nazywamy równanie postaci:

$$x_{n+k} = c_1 x_{n+k-1} + c_2 x_{n+k-2} + \dots + c_k x_n, \quad (9.1)$$

gdzie $c_1, \dots, c_k \in K$. Rozwiązaniem powyższej rekurencji jest dowolny ciąg $(s_0, s_1, \dots) \in K^\infty$ spełniający równości (9.1) dla każdego $n \geq 0$, nazywany CIĄGIEM REKURENCYJNYM rzędu k .

Nie sposób opisać znaczenia rekurencji dla różnych działów matematyki, zwyczajna dla matematyki dyskretnej, teorii funkcji tworzących, teorii szeregów (np. kryterium wymierności funkcji dającej się rozpisać w szeregu), liniowych równań różniczkowych itd. ale nas rekurencje interesują jako swego rodzaju układy nieskończoności wielu jednorodnych równań o zmiennych ze zbioru $X = \{x_0, \dots\}$, przy czym w każdym z równań występuje skończona (i jednakowa) liczba zmiennych.

Nietrudno uświadomić sobie, że zbiór rozwiązań rekurencji liniowej (9.1) tworzy podprzestrzeń liniową w K^∞ . Jeśli ciągi $a = (a_1, a_2, \dots)$ oraz $b = (b_1, b_2, \dots)$ są rozwiązaniami (9.1), to także ciągi $a+b$ oraz λa są w sposób oczywisty jej rozwiązaniami, dla każdego $\gamma \in K$. Co ciekawego można powiedzieć o przestrzeni rozwiązań rekurencji rzędu k ? Na ten temat można opowiedzieć kilka semestralnych wykładów, ale ograniczymy się do kilku uwag na temat najślynniejszej zapewne rekurencji postaci:

$$x_{n+2} = x_{n+1} + x_n$$

Jednym z rozwiązań tej rekurencji jest słynny ciąg Fibonacciego. Wystarczy określić $s_0 = 0, s_1 = 1$. W istocie: każde rozwiązanie powyższej rekurencji jest wyznaczone jednoznacznie przez pierwsze dwa elementy. Naśladowując język wprowadzony na pierwszym wykładzie: przez operacje elementarne na układzie zadanym przez rekurencję wszystkie równości sprowadzić można do równań postaci $x_m = f_m(s_0, s_1)$, gdzie f_m jest pewną liniową zależnością wiążącą s_0 i s_1 . A więc s_0 i s_1 są parametrami w „rozwiązyaniu ogólnym” tego układu. Wszystkie te intuicje można przerobić na formalne rozumowanie i pokazać, że wymiar podprzestrzeni W opisanej rekurencją $x_{n+2} = x_{n+1} + x_n$ wynosi 2. Przykładowa baza W to

$$(0, 1, 1, 2, 3, \dots), \quad (1, 0, 1, 1, 2, \dots).$$

Jeden z jej elementów to oczywiście ciąg Fibonacciego. Zachodzi następujące twierdzenie, będące (przy najmniej intuicyjnie) uogólnieniem powyższych obserwacji.

Twierdzenie 9.4.2

Zbiór rozwiązań rekurencji liniowej rzędu k ma bazę k -elementową.

Jak wiadomo ciąg Fibonacciego można opisać ogólnym wzorem postaci:

$$F_n = \frac{1}{\sqrt{5}} \left(\left(\frac{1 + \sqrt{5}}{2} \right)^n - \left(\frac{1 - \sqrt{5}}{2} \right)^n \right).$$

To dość niezwykłe, że tak prosta rekurencja opisywalna jest w sposób jawnym tak skomplikowanym wzorem. Czy aby na pewno skomplikowanym? Okazuje się, że jego pochodzenie można łatwo wyjaśnić narzędziami algebry liniowej, choć rozumowanie, które pokaże niżej można jeszcze zdecydowanie uprościć. Pomyśl polega na poszukiwaniu bazy złożonej z ciągów postaci

$$\alpha = (a^0, a^1, a^2, \dots), \quad \beta = (b^0, b^1, b^2, \dots),$$

dla pewnych $a, b \in K$, które należą do W i są liniowo niezależne. Nie dla każdej rekurencji postaci (9.1) da się taką bazę znaleźć, ale w przypadku rozważanej przez nas rekurencji rzędu 2 jest to możliwe (to się

wiąże z zagadnieniami, które dokładniej omawiać będziemy podczas rozważania teorii endomorfizmów).

Wyznaczamy szukaną bazę α, β przestrzeni W jakby „od tyłu”, zapisując ciąg $F = (F_0, F_1, F_2, \dots)$ w tej bazie. A zatem mamy równanie

$$F = c_1\alpha + c_2\beta \in K^\infty.$$

W szczególności mamy układ równań

$$\begin{aligned} c_1 + c_2 &= F_0 \\ c_1a + c_2b &= F_1 \\ c_1a^2 + c_2b^2 &= F_2 \\ c_1a^3 + c_2b^3 &= F_3 \\ &\vdots, \end{aligned}$$

czyli w rezultacie otrzymujemy równości

$$\begin{aligned} c_1 + c_2 &= 0 \\ c_1a + c_2b &= 1 \\ c_1a^2 + c_2b^2 &= 1 \\ c_1a^3 + c_2b^3 &= 2 \\ &\vdots, \end{aligned}$$

A zatem $c_1 = -c_2$, czyli $c_1a - c_1b = 1, c_1a^2 - c_1b^2 = 1$, czyli $c_1(a - b) = c_1(a - b)(a + b) = 1$, czyli $a + b = 1$ itd. Ostatecznie:

$$c_1 = \frac{1}{\sqrt{5}}, \quad c_2 = -\frac{1}{\sqrt{5}}, \quad a = \frac{1 + \sqrt{5}}{2}, \quad b = \frac{1 - \sqrt{5}}{2}.$$

A zatem wzór na F_n pochodzi od rozpisywania n -tej współrzędnej F jako kombinacji liniowej elementów bazowych α, β , czyli $F_n = c_1\alpha^n + c_2\beta^n$. Pozostaje oczywiście sprawdzić, że tak uzyskane ciągi spełniają założenia, to znaczy: należą do W i są liniowo niezależne. Pierwsza obserwacja jest jasna, bo:

$$\left(\frac{1 + \sqrt{5}}{2}\right)^{n+1} = \left(\frac{1 + \sqrt{5}}{2}\right)^{n-1} \left(\frac{1 + \sqrt{5}}{2}\right)^2 = \left(\frac{1 + \sqrt{5}}{2}\right)^{n-1} \left(1 + \frac{1 + \sqrt{5}}{2}\right) = \left(\frac{1 + \sqrt{5}}{2}\right)^{n-1} + \left(\frac{1 + \sqrt{5}}{2}\right)^n,$$

czyli $a^{n+1} = a^n + a^{n-1}$. Podobnie pokazujemy, że $b^{n+1} = b^n + b^{n-1}$. To, że ciągi α, β są liniowo niezależne to łatwe ćwiczenie. A zatem ciągi α, β są rzeczywiście bazą W i zachodzi wzór na ciąg Fibonacciego¹.

Czytelnika nie do końca przekonanego skąd rzeczywiście wzięły się w rozwiązyaniu liczby $(1 \pm \sqrt{5})/2$ polecam nieco ogólniejsze spojrzenie. Rozważmy ciąg (x_n) z K^n spełniający warunki

$$x_{n+2} + ax_{n+1} + bx_n = 0.$$

Interesują nas niezerowe ciągi geometryczne określone wzorem $t_n = q^n$ spełniające to równanie. A zatem iloraz $q \neq 0$ tych ciągów spełnia równanie $q^{n+2} + aq^{n+1} + bq^n = 0$, czyli

$$q^2 + aq + b = 0.$$

Równanie $x^2 + ax + b = 0$ nazywane jest zwykle RÓWNANIEM CHARAKTERYSTYCZNYM równania rekurencyjnego $x_{n+2} + ax_{n+1} + bx_n = 0$. W zależności od tego, czy równanie charakterystyczne ma jedno, czy dwa rozwiązania, postać ogólna ciągów spełniających wyjściowe równanie jest inna. Jeśli równanie to ma dwa różne rozwiązania p oraz q , to wyjściową rekurencję spełnia ciąg

$$x_n = c \cdot p^n + d \cdot q^n,$$

gdzie współczynniki c, d wyznaczamy z układu równań utworzonego przez wstawienie $n = 0$ i $n = 1$ do rozwiązania ogólnego. Czytelnik domyśla się zapewne, że również liniowa rekurencja rzędu k ma równanie charakterystyczne rzędu k i odpowiednio skomplikowane zbiory rozwiązań. Zainteresowanych tym tematem oraz elementarnymi zastosowaniami ciągów rekurencyjnych zachęcam do lektury tekstu prof. Wojciecha Guzickiego: „Równania rekurencyjne”: <https://www.mimuw.edu.pl/~guzicki/materialy/Rekurencja.pdf>.

¹Wzór ten, zwany też wzorem Bineta, znany był już w XVIII wieku Bernoullemu, Eulerowi czy de Moivre'owi.

9.5 Dodatek. Wielomiany ograniczonego stopnia i ich bazy

Z punktu widzenia analizy matematycznej, a także zastosowań matematyki, warto przyjrzeć się pewnym bazom przestrzeni wielomianów², przy czym na razie ograniczymy się do przestrzeni $K_{\leq n}[x]$ złożonej z wielomianów stopnia nie większego niż n . Rozpoczniemy od następującej obserwacji.

Uwaga 9.5.1

Skończony układ wielomianów o parami różnych stopniach jest liniowo niezależny w $K[x]$.

Dowód. Niech $p_1, p_2, \dots, p_m \in K[x]$ oraz niech $\deg(p_i) = d_i$. Po ewentualnym przenumerowaniu możemy założyć, że $d_1 > d_2 > \dots > d_m$. Założymy, że dla pewnych $t_1, t_2, \dots, t_m \in K$ mamy

$$t_1 p_1 + t_2 p_2 + \dots + t_m p_m = 0. \quad (\spadesuit)$$

Skoro $\deg(p_1) = d_1$, to niech $a x^{d_1}$ będzie wyrazem najwyższego stopnia w p_1 , dla pewnego $a \neq 0$. Skoro $d_1 > d_2 > \dots > d_m$, to $t_1 a x^{d_1}$ jest jedynym wyrazem stopnia d_1 w wielomianie (\spadesuit) , będącym zarazem wielomianem zerowym. Zatem $t_1 a x^{d_1} = 0$, skąd $t_1 a = 0$. Skoro $a \neq 0$, to $t_1 = 0$. A zatem została nam kombinacja liniowa wielomianów $t_2 p_2 + \dots + t_m p_m = 0$, dla której możemy powtórzyć powyższy argument, co oznacza, że $t_i = 0$, dla $i = 2, 3, \dots, n$. \square

Nietrudno widzieć, że układ $\{1, x, x^2, \dots, x^n\}$ jest bazą przestrzeni wielomianów stopnia nie większego niż n . Jak już wiemy, układ ten jest liniowo niezależny. Jasne jest jednak, że każdy wielomian stopnia nie większego niż n jest liniową kombinacją powyższych wielomianów. Na kolejnym wykładzie pokażemy, że każda przestrzeń liniowa rozpięta na skończonym układzie wektorów ma równoliczne bazy. Warto uświadomić sobie siłę tego twierdzenia, nawet wtedy gdy chcemy je odnieść do baz $K_{\leq n}[x]$. Pokażmy pewien szczególny przypadek.

Wniosek 9.5.2

Niech $p_0, p_1, \dots, p_n \in K_{\leq n}[x]$ będą wielomianami odpowiednio stopni $0, 1, 2, \dots, n$. Wówczas układ

$$\{p_0, p_1, \dots, p_n\}$$

jest bazą przestrzeni liniowej $K_{\leq n}[x]$.

Dowód. Na mocy wcześniejszej obserwacji wystarczy pokazać, że $\text{lin}(p_0, \dots, p_n) = K_{\leq n}[x]$. Dowód jest indukcją ze względu na n . Dla $n = 0$ teza jest oczywista, bo wielomian stopnia 0 jest niezerowy. Założymy, że układ wielomianów $\{p_0, \dots, p_k\}$ stopni od 0 do k rozpinia $K_{\leq k}[x]$ i weźmy dowolny wielomian p_{k+1} stopnia $k+1$ należący do $K_{\leq k+1}[x]$. Pokażmy, że $\text{lin}(p_0, \dots, p_k, p_{k+1}) = K_{\leq k+1}[x]$. Weźmy dowolny wielomian w taki, że $\deg(w) \leq k+1$. Jeśli $\deg w < k+1$, to w jest z założenia indukcyjnego kombinacją liniową wielomianów p_0, \dots, p_k . Jeśli $\deg(w) = k+1$, to istnieje takie $a \in K$, że

$$w - a \cdot p_{k+1}$$

jest wielomianem stopnia mniejszego niż $k+1$. A zatem $w - a \cdot p_{k+1} \in \text{lin}(p_0, \dots, p_k)$, co oznacza, że $w \in \text{lin}(p_0, \dots, p_k, p_{k+1})$. Zatem $\text{lin}(p_0, \dots, p_k, p_{k+1}) = K_{\leq k+1}[x]$, co kończy dowód. \square

Wniosek 9.5.3

Niech $a \in K$. Układ wielomianów

$$\{1, (x-a), (x-a)^2, \dots, (x-a)^n\}$$

jest bazą przestrzeni $K_{\leq n}[x]$.

²Na podstawie jednego z najlepszych podręczników do algebry jaki znam, autorstwa Keitha Nicholsona, udostępnionego przez Autora do ogólnego użytku. Jest to skarbnica wiedzy o zastosowaniach algebry liniowej i po prostu świetny tekst: <https://lyryx.com/linear-algebra-applications/>.

Wiemy, że jeśli mamy bazę przestrzeni liniowej, to każdy wektor zapisuje się jednoznacznie jako kombinacja liniowa elementów bazowych. Oznacza to, że każdy wielomian $f \in K_{\leq n}[x]$ może być przedstawiony w postaci:

$$f = a_0 + a_1(x - a) + a_2(x - a)^2 + \dots + a_n(x - a)^n.$$

Jak się okazuje, współczynniki a_i mają duże znaczenie w analizie. Są to bowiem współczynniki tzw. wielomianu Taylora funkcji f (dla nas to będzie funkcja wielomianowa, a na analizie to będzie funkcja różniczkowalna odpowiednią liczbę razy). Oczywiście z twierdzenia Bezout wynika, że $a_0 = f(a)$, czyli a_0 jest wartością funkcji wielomianowej odpowiadającej wielomianowi f . Czym są wyższe współczynniki?

Czytelnik znający wzór na pochodną złożenia (lub iloczynu) bez trudu sprawdzi, że pochodna funkcji wielomianowej $f(x)$ odpowiadająca powyższemu wielomianowi równa jest:

$$f^{(1)}(x) = a_1 + 2a_2(x - a) + 3a_3(x - a)^2 + \dots + na_n(x - a)^{n-1}.$$

Oznacza to, że $f^{(1)}(a) = a_1$. Jeśli przez $f^{(n)} = f^{(1)}(f^{(n-1)})$ oznaczymy n -tą pochodną wielomianu f , wówczas biorąc $f = f^{(0)}$ (zerowa pochodna) otrzymujemy następujący wniosek.

Wniosek 9.5.4

Jeśli $f(x)$ jest funkcją wielomianową stopnia n , to

$$f(x) = f(a) + \frac{f^{(1)}(a)}{1!}(x - a) + \frac{f^{(2)}(a)}{2!}(x - a)^2 + \dots + \frac{f^{(n)}(a)}{n!}(x - a)^n.$$

Przykład. Przedstawmy $f(x) = 5x^3 + 10x + 2$ w bazie $\{1, x - 1, (x - 1)^2\}$. Kolejne pochodne funkcji wielomianowej $f(x)$ to:

$$f^{(1)}(x) = 15x^2 + 10, \quad f^{(2)}(x) = 30x, \quad f^{(3)}(x) = 30.$$

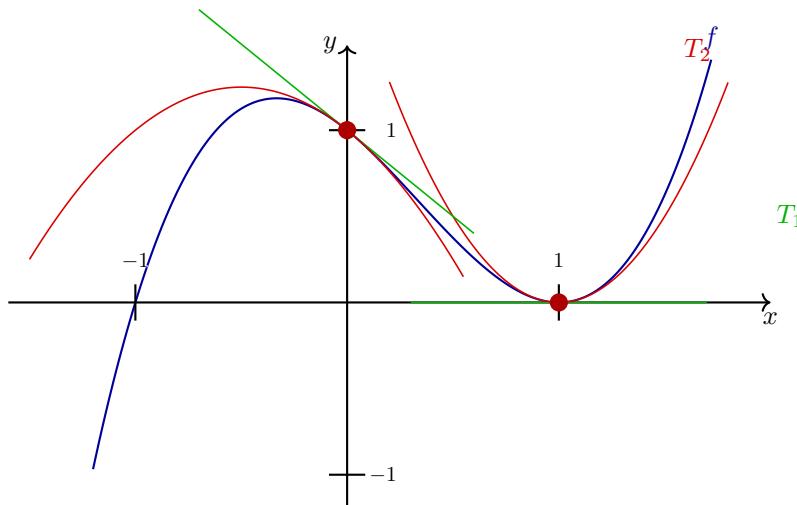
Stąd szukane przedstawienie ma postać:

$$f(x) = f(1) + \frac{f^{(1)}(1)}{1!}(x - 1) + \frac{f^{(2)}(1)}{2!}(x - 1)^2 + \frac{f^{(3)}(1)}{3!}(x - 1)^3 = 17 + 25(x - 1) + 15(x - 1)^2 + 5(x - 1)^3.$$

Wzór wyznaczony wyżej ma ogromne znaczenie w analizie, bowiem umożliwia Państwu badanie własności funkcji różniczkowalnych $n + 1$ razy w sposób ciągły, poprzez zapisanie ich (w otoczeniu punktu x_0) w postaci:

$$f(x_0 + h) = w_n(x_0) + r(x_0, h),$$

gdzie $w_n(x)$ będzie n -tym wielomianem Taylora funkcji f w punkcie x_0 , zaś $r(x_0, h)$ – tak zwaną resztą Peano we wzorze Taylora i w dalszej konsekwencji pozwoli na rozwijanie funkcji w szeregi. Oto przykład takiej sytuacji. Pewną funkcję f , różniczkowalną tyle razy ile trzeba, przybliżamy w punktach 0 oraz 1. Dla każdego z tych punktów wyznaczamy odpowiednie wielomiany Taylora stopnia 1 oraz 2. Ich wartości w tych punktach są oczywiście takie same, jak wartości funkcji f . W niewielkim otoczeniu tych punktów wielomiany te przybliżają z odpowiednio „kontrolowaną” dokładnością przebieg funkcji f .



Rys. 1. Źródło: https://tikz.net/taylor_expansion/.

Drugi ważny przykład dotyczy do pewnego stopnia innej sytuacji. Tym razem nie rozważamy wielomianów różnych stopni, ale szczególny układ $n+1$ wielomianów stopnia n , zwanych wielomianami Lagrange'a, który będzie stanowił bazę $K_{\leq n}[x]$. Założmy, że dany jest układ parami różnych elementów ciała K a_0, a_1, \dots, a_n . Okazuje się, że dla każdego takiego układu argumentów znajdziemy bazę przestrzeni $K_{\leq n}[x]$ złożoną z wielomianów $\delta_0, \delta_1, \dots, \delta_n$ taką, że dla każdego $f \in K_{\leq n}[x]$ mamy:

$$f = f(a_0) \cdot \delta_0 + f(a_1) \cdot \delta_1 + \dots + f(a_n) \cdot \delta_n.$$

Dla przykładu, dla układu trzech punktów a_0, a_1, a_2 będzie to baza:

$$\delta_0 = \frac{(x - a_1)(x - a_2)}{(a_0 - a_1)(a_0 - a_2)}, \quad \delta_1 = \frac{(x - a_0)(x - a_2)}{(a_1 - a_0)(a_1 - a_2)}, \quad \delta_2 = \frac{(x - a_0)(x - a_1)}{(a_2 - a_0)(a_2 - a_1)}.$$

Oto przykład. Dany jest wielomian $f(x) = x^2 - 2x + 1$ i wiemy, że $f(-1) = 4$, $f(0) = 1$ oraz $f(1) = 0$. Wielomiany określone wyżej mają dla $a_0 = -1, a_1 = 0, a_2 = 1$ postać:

$$\delta_0 = \frac{(x - 0)(x - 1)}{(-1 - 0)(-1 - 1)} = \frac{1}{2}(x^2 - x), \quad \delta_1 = \frac{(x + 1)(x - 1)}{(0 + 1)(0 - 1)} = -(x^2 - 1), \quad \delta_2 = \frac{(x + 1)(x - 0)}{(1 + 1)(1 - 0)} = \frac{1}{2}(x^2 + x).$$

Mamy też:

$$f(x) = 4 \cdot \frac{1}{2}(x^2 - x) + 1 \cdot (-1)(x^2 - 1) + 0 \cdot \frac{1}{2}(x^2 + x) = x^2 - 2x + 1.$$

Ktoś zapyta – po co nam takie rozwinięcia? Warto zauważyc, że każdy z trójki wielomianów $\delta_0, \delta_1, \delta_2$ ma tę własność, że $\delta_i(a_j) = 0$, dla $i \neq j$ oraz $\delta_i(a_i) = 1$. Innymi słowy, postulowane przez nas twierdzenie mówi, że jak z góry zadamy wartości $f(a_0), \dots, f(a_n)$, to znajdziemy wielomian stopnia nie większego niż n , który w punktach a_0, \dots, a_n ma dokładnie te zadane z góry wartości. Na przykład chcąc, aby dla $a_0 = -1, a_1 = 0, a_2 = 1$ pewna funkcja kwadratowa przyjmowała wartości 10, 15, 30, wystarczy rozpatrzyć wielomian

$$f = 10\delta_0 + 15\delta_1 + 30\delta_2.$$

W ten sposób skończone układy $n+1$ punktów na płaszczyźnie można **interpolować** wielomianami z $K_{\leq n}[x]$. Przejdzmy do twierdzenia³, które wyjaśnia powyższą sytuację.

Twierdzenie 9.5.5

Założmy, że a_0, a_1, \dots, a_n są parami różne. Rozważmy zbiór wielomianów Lagrange'a $\delta_0, \delta_1, \dots, \delta_n$ postaci:

$$\delta_k = \frac{\prod_{i \neq k} (x - a_i)}{\prod_{i \neq k} (a_k - a_i)}, \quad \text{dla } k = 0, 1, 2, \dots, n.$$

Wówczas układ $\delta_0, \dots, \delta_n$ jest bazą $K_{\leq n}[x]$. Co więcej, dla każdego wielomianu $f \in K_{\leq n}[x]$ mamy:

$$f = f(a_0) \cdot \delta_0 + f(a_1) \cdot \delta_1 + \dots + f(a_n) \cdot \delta_n.$$

Dowód. W wielomianie δ_k licznik jest iloczynem wyrażeń liniowych $x - a_0, x - a_1, \dots, x - a_n$ z pominięciem czynnika $x - a_k$. Podobnie dla mianownika. Widzimy więc, że wielomian δ_k przyjmuje wartość 0, dla wszystkich a_i różnych od a_k . Natomiast $\delta_k(a_k) = 1$.

Układ $\delta_0, \dots, \delta_n$ jest liniowo niezależny. Rzeczywiście, jeśli $r_0 \cdot \delta_0 + r_1 \cdot \delta_1 + \dots + r_n \cdot \delta_n = 0$, dla pewnych $r_1, \dots, r_n \in K$, wówczas skoro porównujemy funkcje wielomianowe, mamy z prawej strony funkcję, która dla każdego x przyjmuje wartość zero. Wartość funkcji po lewej stronie dla a_0 równa jest $r_0 \delta_0(a_0) + r_1 \delta_1(a_0) + \dots + r_n \delta_n(a_0) = r_0$, a zatem $r_0 = 1$. Analogicznie pokazujemy, że $r_2 = \dots = r_n = 0$.

Układ $\delta_0, \dots, \delta_n$ rozpinia $K_{\leq n}[x]$. Istotnie, weźmy dowolny wielomian $f \in K_{\leq n}[x]$ i rozważmy kombinację liniową postaci $w = f(a_0) \cdot \delta_0 + f(a_1) \cdot \delta_1 + \dots + f(a_n) \cdot \delta_n$. Wielomian w jest stopnia nie większego niż n i ma dla a_0, \dots, a_n takie same wartości, jak wielomian f . To znaczy, że wielomian $f - w$ jest stopnia nie większego niż n i ma $n+1$ pierwiastków. Zgodnie z twierdzeniami z rozdziału czwartego (wielomian stopnia n ma nie więcej niż n pierwiastków, licząc krotności) wielomianem zerowym. Zatem $w = f$. \square

³W tekście źródłowym jest ono sformułowane ogólniej i dowód jest identyczny jak poniżej. Źródło zakłada twierdzenie o równoliczności baz. My korzystamy z wiedzy o liczbie pierwiastków wielomianów stopnia n .

9.6 Trivia. Cykle i rozcięcia w grafach

W tym dodatku pokażemy ciekawe zastosowanie algebry liniowej w kombinatoryce, związane z tak zwartymi przestrzeniami cykli i rozcięć w grafach. Ustalmy kilka pojęć wstępnych.

Definicja 9.6.1: Graf nieorientowany (prosty)

Niech X będzie skończonym zbiorem niepustym, E zaś niech będzie podzbiorem zbioru par nieuporządkowanych zbiór X . Parę $G = (X, E)$ nazwiemy GRAFEM NIEZORIENTOWANYM o zbiorze wierzchołków $X = V(G)$ i zbiorze krawędzi $E = E(G)$. Jeśli $\{a, b\} \in E(G)$ to mówimy, że między wierzchołkami a, b grafu G jest KRAWĘDŹ oraz mówimy, że wierzchołki a, b sąsiadują ze sobą. Dodatkowo:

- STOPNIEM WIERZCHOŁKA $x \in V(G)$ w grafie G nazywamy liczbę $\deg(x)$ krawędzi, których jeden z elementów równy jest x ,
- PODGRAFEM grafu $G = (X, E)$ nazywamy graf (X', E') , że $X' \subseteq X$ oraz $E' \subseteq E$, przy czym jeśli $\{a, b\} \in E'$, to $a, b \in X'$,
- ŚCIEŻKĄ nazywamy ciąg wierzchołków x_0, x_1, \dots, x_n , taki że dla każdego $k \in \{0, 1, \dots, n-1\}$ wierzchołki x_k oraz x_{k+1} są sąsiadami,
- DROGĄ nazywamy ciąg krawędzi $e_1 = \{x_0, x_1\}, e_2 = \{x_1, x_2\}, \dots, e_n = \{x_{n-1}, x_n\}$, których zbiór wierzchołków x_0, \dots, x_n tworzy ścieżkę,
- CYKLEM nazywamy drogę zamkniętą, czyli taką, w której $x_0 = x_n$,

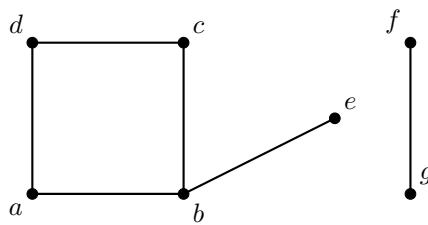
Co więcej, graf nazywamy:

- SPÓJNYM, jeśli każde dwa jego wierzchołki łączy ścieżka,
- ACYKLICZNYM, jeśli nie zawiera on cykli (jako podgrafów),
- DRZEWEM, jeśli jest spójny i acykliczny.

Mówimy też, że podgraf G' grafu G jest jego SPÓJNĄ SKŁADOWĄ, jeśli jest on spójny i nie jest zawarty w sposób właściwy w żadnym podgrafie spójnym grafu G .

Przykład. Rozważmy graf $G = (V(G), E(G))$, gdzie

- zbiór wierzchołków $V(G) = \{a, b, c, d, e, f, g\}$,
- zbiór krawędzi $E(G) = \{\{a, b\}, \{b, c\}, \{c, d\}, \{d, a\}, \{b, e\}, \{f, g\}\}$.



W tym grafie mamy:

- $\deg(e) = \deg(f) = \deg(g) = 1$,
- $\deg(a) = \deg(c) = \deg(d) = 2$,
- $\deg(b) = 3$.

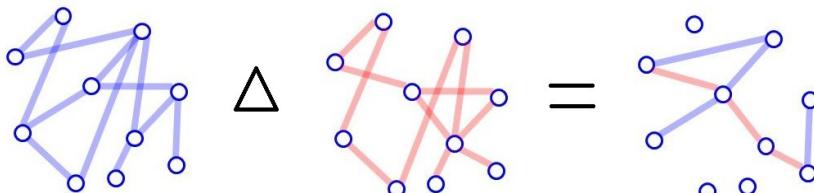
Przykładowa ścieżka w G to: a, b, c , przykładowa droga: $\{a, b\}, \{b, e\}$, przykładowy cykl — $\{a, b\}, \{b, c\}, \{c, d\}, \{d, a\}$. Graf G nie jest spójny. Jego spójnymi składowymi są podgrafy $G' = (V', E')$ oraz $G'' = (V'', E'')$, gdzie

$$V' = \{a, b, c, d, e\}, E' = \{\{a, b\}, \{b, c\}, \{c, d\}, \{d, a\}, \{b, e\}\}, \quad V'' = \{f, g\}, E'' = \{\{f, g\}\}.$$

Wynalazek grafu może nam się kojarzyć głównie z Eulerem (słusznie) i z problemem mostów w Królewcu (ponownie słusznie), ale z punktu widzenia nauki znaczenie tego pojęcia ukazał Kirchoff, zajmujący się badaniem obwodów i sieci elektrycznych. W tym ujęciu interesują nas zwykle grafy z krawędziami zorientowanymi (wskażającymi kierunek przepływu prądu) i z przypisanymi wagami (różne przewody mogą mieć różne właściwości jako przewodniki). Z naszego punktu widzenia ważne jest to, że przy konstruowaniu takiej sieci, napięcie powinno się zbalansować, czyli nie gromadzimy nadmiernego natężenia w żadnym punkcie. Krótko mówiąc: wzdłuż każdego cyklu tego obwodu suma spadków napięć powinna wynosić zero. Wydaje się więc, że chcąc zaplanować taką sieć trzeba sprawdzić każdy cykl w ilustrującym ją grafie G . Jak się jednak okazuje, nie jest to konieczne. Okazuje się, że graf ten spełnia drugie prawo Kirchhoffa wtedy i tylko wtedy, gdy G spełnia to prawo na pewnym podzbiorze cykli, zwany bazą przestrzeni cykli. Czym jest ta przestrzeń liniowa? Zaczniemy od przestrzeni krawędziowej.

Definicja 9.6.2: Przestrzeń krawędziowa grafu

Określamy przestrzeń liniową $P(E)$ nad ciałem \mathbb{Z}_2 , zwaną PRZESTRZENIĄ KRAWĘDZIOWĄ grafu $G = (V(G), E(G))$, jak w przypadku przestrzeni liniowej podzbiorów zbioru niepustego. Zbiorem wektorów jest $P(E)$ — zbiór podzbiorów zbioru krawędzi $E(G)$ grafu G , a sumą $X \oplus Y$ dwóch podzbiorów X i Y należących do $E(G)$ jest różnica symetryczna tych zbiorów $X \Delta Y = (X \cup Y) \setminus (X \cap Y)$. Mnożenie przez skalar jest zdefiniowane jako $1 \cdot X = X$ oraz $0 \cdot X = \emptyset$



Różnica symetryczna dwóch podzbiorów krawędzi grafu

Widzimy więc, że przestrzeń krawędziowa jest pewnym typem przestrzeni podzbiorów. Singletony zawierające pojedyncze wierzchołki grafu G tworzą bazę przestrzeni $P(E)$. Zauważmy, że operacja różnicy symetrycznej zachowuje się dobrze na innych wprowadzonych przez nas strukturach. Kluczowa obserwacja jest taka, że różnica symetryczna dwóch cykli jest cyklem. W ten sposób dochodzimy do definicji.

Definicja 9.6.3: Przestrzeń cykli grafu

Niech $G = (V(G), E(G))$ będzie grafem. Podprzestrzeń $C(E)$ przestrzeni $P(E)$ rozpiętą przez wszystkie cykle nazywamy przestrzenią cykli grafu G . Liczbę elementów bazy $C(E)$ nazywamy LICZBĄ CYKLOMATYCZNĄ grafu G .

O znaczeniu przestrzeni cykli świadczy podstawowa własność szczególnych i najstarszych ich typów — cykli Eulerowskich. Jak się okazuje, różnica symetryczna dwóch takich cykli również jest cyklem Eulerowskim, czyli przechodzącym przez każdą krawędź dokładnie raz. Jak wiadomo, graf F ma cykl Eulerowski, jeśli każdy jego wierzchołek ma parzysty stopień. Różnica symetryczna dwóch grafów, których wierzchołki mają parzyste stopnie oczywiście ma również tę własność.

Skoro dysponujemy przestrzenią cykli można się zastanawiać nad jej bazą. Jest to zagadnienie wymagające odrobinę wiedzy z teorii grafów. Do podstawowych pojęć należy drzewo (las) rozpinające oraz fundamentalne cykle.

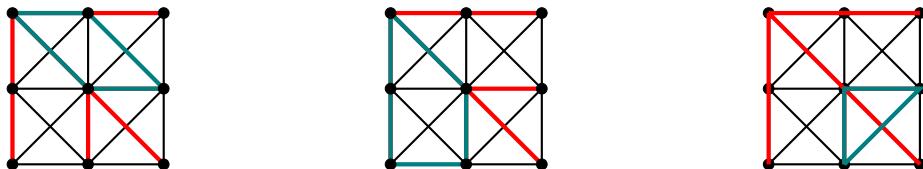
Definicja 9.6.4

Podgraf, który zawiera wszystkie wierzchołki grafu spójnego G i jest drzewem, nazywamy DRZEWEM ROZPINAJĄCYM grafu G . Jeśli G nie jest spójny, wówczas dowolną sumę rozłączną grafów rozpinających spójne składowe tego grafu, nazywamy LASEM ROZPINAJĄCYM G .

Oto przykłady dwóch drzew rozpinających grafu o 9 wierzchołkach (na czerwono). Możemy też przyjąć, że suma rozłączna tych dwóch drzew stanowi las rozpinający graf o 18 wierzchołkach i dwóch składowych.



Niech L będzie lasem rozpinającym grafu G . Wówczas dodanie dowolnej krawędzi e z G nie należącej do L utworzy dokładnie jeden cykl, zwany CYKLEM FUNDAMENTALNYM C_e związanym z lasem rozpinającym L . Cykl C_e jest wyznaczony jednoznacznie, ponieważ biorąc końce x, y krawędzi e wiemy, że w lesie L jest dokładnie jedna droga między x , a y . Oto przykłady cykli fundamentalnych dla lasu L na prawym grafie.



Nietrudno widzieć, że jeśli L jest lasem rozpinającym grafu G , to każdy cykl w G ma wspólną krawędź z dopełnieniem L . Gdyby bowiem cykl nie miał wspólnej krawędzi z dopełnieniem L , to byłby zawarty w L , co przeczy acykliczności L . Okazuje się, że zachodzi następujące twierdzenie.

Twierdzenie 9.6.5

Zbiór cykli fundamentalnych dowolnego lasu rozpinającego L grafu G stanowi bazę podprzestrzeni cykli $C(E)$. W szczególności $\dim C(E) = m + n - c$, gdzie m jest liczbą krawędzi w grafie G , n jest liczbą wierzchołków, a c jest liczbą składowych spójnych.

Dowód. Oznaczmy przez C_e cykl fundamentalny powstały przez dopisanie krawędzi e do lasu L . Rozważmy układ

$$\{C_e, e \in E(G) \setminus E(L)\}$$

Po pierwsze zauważmy, że układ ten jest liniowo niezależny, ponieważ graf $C_{e_1} \Delta C_{e_2} \Delta \dots \Delta C_{e_k}$ zawiera krawędzie e_1, \dots, e_k . Innymi słowy — C_e jest jedynym elementem tego układu zawierającym e .

Z drugiej strony układ ten rozpinia przestrzeń cykli. Istotnie, jeśli $H \in C(E)$, to bierzemy wszystkie krawędzie e_1, \dots, r_k z $V(H)$, które nie są w $V(L)$ i rozważamy $H \Delta C_{e_1} \Delta C_{e_2} \Delta \dots \Delta C_{e_k}$. Rezultat dalej jest w przestrzeni cykli, i jest to podgraf L , ponieważ każda krawędź $e \notin E(L)$ została usunięta przez różnicę symetryczną z C_e . Ale element przestrzeni cykli nie może być podgrafem drzewa, o ile nie jest pusty. Stąd $H = C_{e_1} \Delta C_{e_2} \Delta \dots \Delta C_{e_k}$. \square

Czytelnik może się zastanawiać po co nam taka osobliwa przestrzeń? Istnieje sporo powodów, które należą do bardziej zaawansowanej matematyki — jak choćby zastosowania w teorii grup homologii kompleksów symplektycznych w topologii algebraicznej, czy już wspomniane prawo Kirchoffa. Jest również niezwykłe zastosowanie w samej teorii grafów, które wspominamy bez dowodu.

Przypomnijmy, że GRAFEM PLANARNYM nazywamy graf, który można narysować na płaszczyźnie w taki sposób, by żadne dwie krawędzie się nie przecinały. Klasyczne kryterium planarności pochodzi od polskiego matematyka Kazimierza Kuratowskiego i wskazuje podgrafy, których graf planarny zawierać nie może. Poniższe kryterium dotyczy natomiast samej przestrzeni cykli.

Twierdzenie 9.6.6: Mac Lane 1937, O'Neil 1973

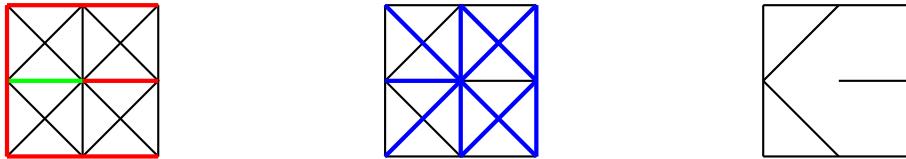
Skończony nieorientowany graf G jest planarny wtedy i tylko wtedy, gdy każda krawędź grafu G jest składnikiem dokładnie dwóch wektorów bazy $C(E)$, czyli dwóch fundamentalnych cykli.

Powiedzieliśmy więc coś o podprzestrzeni cykli, ale skoro tematem wykładu była suma prosta, to pokażemy naturalne dopełnienie prostej przestrzeni cykli, czyli tzw. przestrzeń rozcięć grafu.

Definicja 9.6.7: Przestrzeń rozcięć grafu

Niech $G = (V, E)$. Podprzestrzeń $R(E)$ przestrzeni $P(E)$ rozpięta przez wszystkie rozcięcia, czyli podzbiory $P(E)$, których usunięcie rozspaja graf, nazywamy PRZESTRZENIĄ ROZCIĘĆ grafu G .

Jeśli z lasu L usuniemy dowolną krawędź, to w (odpowiadającej jej spójnej składowej) powstają dwa rozłączne zbiory wierzchołków V_1, V_2 . Zbiór wszystkich krawędzi G takich, że koniec jest w V_1 , a drugi w V_2 tworzy rozcięcie, które nazywamy ROZCIĘCIEM FUNDAMENTALNYM związanym z lasem L . Oto przykłady (na niebiesko) rozcięć fundamentalnych, dla jednego z drzew wskazanych wyżej, z którego usunięto krawędź (zieloną). Spójne grafy po usunięciu rozcięcia fundamentalnego są po prawej.



Dalsza część tego tekstu będzie dla Państwa jasna dopiero po wprowadzeniu sumy prostej za kilka wykładów. Umieścimy ją jednak już teraz. To, co jest ciekawe w tym dodatkowym obiekcie to fakt, że zachodzi równość $P(E) = C(E) \oplus R(E)$. Chodzi o to, że każdy element $P(E)$ można zapisać jednoznacznie za pomocą sumy elementów z bazy $C(E)$ oraz sumy elementów z bazy $R(E)$. Dowodzi się ten fakt na różne sposoby, ale najbardziej popularny korzysta z następującego rezultatu.

Uwaga 9.6.8

Każdy cykl i rozcięcie w grafie G mają parzystą liczbę wspólnych krawędzi.

Dowód. Rozważmy rozcięcie S w spójnym grafie G (to oczywiście wystarczy do dowodu). Założymy, że usunięcie S ze zbioru krawędzi G rozbija zbiór wierzchołków na dwie rozłączne podzbiory V_1 i V_2 . Niech C będzie cyklem w G . Jeśli wszystkie wierzchołki C leżą w jednym ze zbiorów V_1 lub V_2 , to wszystkie krawędzie C są różne od krawędzi w S , co oznacza, że w tym przypadku cykl S i rozcięcie S mają 0 wspólnych krawędzi — czyli liczbę parzystą.

Jeśli pewne wierzchołki C są w V_1 , a pewne w V_2 , to przechodząc cykl przechodzimy między zbiorami V_1 i V_2 . Skoro cykl jest drogą zamkniętą, liczba krawędzi przejścia między V_1 i V_2 musi być parzysta. Każdemu przejściu z jednego zbioru do drugiego odpowiadać musi przejście z powrotem. \square

Rozumowanie pokazujące, że przestrzeń krawędzi grafu nieorientowanego jest sumą prostą przestrzeni cykli i krawędzi można przeprowadzić na wiele sposobów, ale wskażemy uniwersalną drogę — ważną w całej algebrze liniowej. Już w jednym z poprzednich wykładów mówiliśmy o naiwnym ujęciu prostopadłości, pochodząącym od swego rodzaju uogólnienia iloczynu skalarnego. To uogólnienie może pójść bardzo daleko, o czym świadczy poniższa definicja.

Definicja 9.6.9: Forma dwuliniowa dla grafów i podzbiorów

Niech $P(E)$ będzie przestrzenią krawędzi grafu G i weźmy wektory $v_1 = a_1e_1 + \dots + a_m e_m$ oraz $v_2 = b_1e'_1 + \dots + b_m e'_m$, gdzie a_i, b_j należą do \mathbb{Z}_2 , oraz $e_i, e'_j \in E$. Określamy:

$$\langle v_1, v_2 \rangle = a_1b_1 + a_2b_2 + \dots + a_mb_m.$$

Warunek $\langle v_1, v_2 \rangle = 0$ spełniony jest zawsze, gdy zbiory v_1, v_2 mają parzystą liczbę wspólnych krawędzi. Z twierdzenia powyżej nietrudno wywnioskować, że w istocie przestrzeń cykli jest „prostopadła” do przestrzeni rozcięć grafu. Dokładniej, dla dowolnej podprzestrzeni $W \subset P(E)$ określić można

$$W^\perp = \{v \in P(E) : \langle v, w \rangle = 0, \forall_{w \in W}\}.$$

Czytelnikowi pozostawiam pokazanie, że w rozważanym przypadku mamy $C(E)^\perp = R(E)$ oraz, że wynika stąd równość $P(E) = C(E) \oplus R(E)$. Z pewnością w przyszłym semestrze będą Państwo mieli więcej narzędzi do uzasadniania takich wyników.

Rozdział 10

Wymiar przestrzeni liniowej. Rząd macierzy

10.1 Wykład 10

Na ostatnim wykładzie wprowadzone zostało bazy przestrzeni liniowej. Na tym wykładzie wprowadzimy pojęcie wymiaru przestrzeni liniowej, stanowiącego liczbę elementów dowolnej bazy przestrzeni liniowej. Jak się okazuje, na mocy twierdzenia Steinitza pojęcie to jest dobrze zdefiniowane.

Wniosek 10.1.1: O liczbie wektorów rozpinających podprzestrzeń

- (a) Jeśli W jest podprzestrzenią przestrzeni $V = \text{lin}(\beta_1, \dots, \beta_m)$, to w W istnieje taki układ liniowo niezależny $\alpha_1, \dots, \alpha_k$, $k \leq m$, że $W = \text{lin}(\alpha_1, \dots, \alpha_k)$.
- (b) Jeśli $\text{lin}(\alpha_1, \dots, \alpha_k) = \text{lin}(\alpha'_1, \dots, \alpha'_l)$ i układy $\alpha_1, \dots, \alpha_k$ oraz $\alpha'_1, \dots, \alpha'_l$ są liniowo niezależne, to $k = l$.

Dowód. Weźmy najdłuższy liniowo niezależny układ w W (to ma sens, bo wszystkie mają długość $\leq m$). Niech to będzie układ $\alpha_1, \dots, \alpha_k$. Pokażemy, że:

$$W = \text{lin}(\alpha_1, \dots, \alpha_k).$$

Dowodzimy, że mają miejsce dwie inkluze. Jedna z nich: $\text{lin}(\alpha_1, \dots, \alpha_k) \subseteq W$, jest oczywista, bo skoro wektory $\alpha_1, \dots, \alpha_k$ należą do W , to każda ich kombinacja liniowa też (bo W to podprzestrzeń). Dowodzimy teraz, że: $\text{lin}(\alpha_1, \dots, \alpha_k) \supseteq W$.

Weźmy dowolny wektor $\alpha \in W$. Układ $\alpha_1, \dots, \alpha_k, \alpha$ jest dłuższy niż układ $\alpha_1, \dots, \alpha_k$, więc jest liniowo zależny. Korzystając z implikacji $(b) \Rightarrow (a)$ w dowodzie Uwagi 8.1.5, otrzymując $\alpha \in \text{lin}(\alpha_1, \dots, \alpha_k)$. Wobec dowolności α otrzymujemy drugą inklikcję.

Dowód (b). Skoro $\text{lin}(\alpha_1, \dots, \alpha_k) = \text{lin}(\alpha'_1, \dots, \alpha'_l)$ i układy $\alpha_1, \dots, \alpha_k$ oraz $\alpha'_1, \dots, \alpha'_l$ są liniowo niezależne, to wobec $\alpha_1, \dots, \alpha_k \in \text{lin}(\alpha'_1, \dots, \alpha'_l)$ mamy $k \leq l$. Z drugiej strony mamy przecież także symetryczne należenie: $\alpha'_1, \dots, \alpha'_l \in \text{lin}(\alpha_1, \dots, \alpha_k)$, czyli z twierdzenia Steinitza: $l \leq k$. A zatem $k = l$. \square

Twierdzenie 10.1.2

Jeśli przestrzeń liniowa V posiada bazę złożoną z n wektorów, to każda baza przestrzeni V jest złożona z n wektorów.

Dowód. Jeśli $\alpha_1, \dots, \alpha_k$ oraz $\alpha'_1, \dots, \alpha'_l$ są bazami przestrzeni V , to $\text{lin}(\alpha_1, \dots, \alpha_k) = \text{lin}(\alpha'_1, \dots, \alpha'_l) = V$. Układy te są liniowo niezależne, a zatem na mocy Wniosku 10.1.1 mamy $k = l$. \square

Definicja 10.1.3: Wymiar przestrzeni liniowej

Mówimy, że przestrzeń liniowa V jest n WYMIAROWA, jeśli V posiada bazę złożoną z n wektorów. Piszymy wówczas

$$\dim V = n$$

i liczbę n nazywamy WYMIAREM PRZESTRZENI V . Przyjmujemy też $\dim\{0\} = 0$.

Mówimy, że przestrzeń liniowa V jest SKOŃCZENIE WYMIAROWA, jeśli V jest n wymiarowa dla pewnego $n \in \mathbb{N} \cup \{0\}$. Jeśli V nie jest skończenie wymiarowa, to V nazywamy NIESKOŃCZENIE WYMIAROWĄ i piszemy $\dim V = \infty$.

Podajmy kilka ważnych przykładów.

- Oczywiście $\dim K^n = n$, o czym świadczy choćby baza standardowa.
- Jeśli $V = M_{m \times n}(K)$, to baza przestrzeni V złożona jest (na przykład) z macierzy E_{ij} , które poza wyrazem w i -tym wierszu i j -tej kolumnie, równym 1, mają same wyrazy zerowe. Nietrudno zatem widzieć, że $\dim M_{m \times n}(K) = m \cdot n$.
- Niech $V = \{(x_1, x_2, x_3) \in \mathbb{R}^3 \mid x_1 + 2x_2 - x_3 = 0\}$. Wówczas $\dim(V) = 2$, bo V ma bazę postaci $\{(-2, 1, 0), (1, 0, 1)\}$.
- Dla każdego n układ n wektorów x, x^2, \dots, x^n przestrzeni $K[x]$ jest liniowo niezależny. A zatem przestrzeń ta nie może być skończenie wymiarowa. Gdyby jej wymiar wynosił k , to na mocy twierdzenia Steinitza każdy układ liniowo niezależny w $K[x]$ musiałby liczyć nie więcej niż k wektorów. A zatem $\dim K[x] = \infty$. Podobnie nietrudno pokazać, że $\dim K^\infty = \infty$.

Wyznaczając wymiar podprzestrzeni rozpiętych na układach wektorów w K^n korzystać będziemy często z Uwag 7.1.6 oraz 8.1.2. Zobaczmy przykład. Dana jest podprzestrzeń

$$V = \text{lin}((1, 2, 0, 1, 0), (0, 1, 1, 1, 1), (2, 2, 3, 0, 3), (1, 3, 1, 2, 1)) \subseteq K^5.$$

Wykonując operacje elementarne na wierszach macierzy rozmiaru 4×5 , której wiersze stanowią układy rozpinające powyższą podprzestrzeń mamy: Przekształcamy macierz układu wektorów roznajających V przy pomocy operacji elementarnych na wierszach:

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

Na mocy Uwagi 7.1.6 mamy

$$V = \text{lin}((1, 0, -2, -1, -2), (0, 1, 1, 1, 1), (0, 0, 5, 0, 5)).$$

Jeśli K jest np. ciałem \mathbb{Q} , to układ ten jest na mocy Uwagi 8.1.2 liniowo niezależny, a zatem jest bazą V i $\dim V = 3$. Jeśli zaś założymy, że $K = \mathbb{Z}_5$, to $V = \text{lin}((1, 0, 3, -1, 3), (0, 1, 1, 1, 1))$ i $\dim V = 2$.

Wniosek 10.1.4

Podprzestrzeń przestrzeni rozpiętej na skończonym układzie wektorów jest skończenie wymiarowa. Jeśli W jest podprzestrzenią V i $\dim V = n$, to $\dim W \leq n$.

Dowód. Niech $W \subseteq V = \text{lin}(\beta_1, \dots, \beta_m)$. Wówczas $W = \text{lin}(\alpha_1, \dots, \alpha_k)$ dla pewnego układu liniowo niezależnego $\alpha_1, \dots, \alpha_k$ na mocy wniosku. Układ $\alpha_1, \dots, \alpha_k$ jest bazą przestrzeni W , więc W jest skończenie wymiarowa. Jeśli $\dim V = n$, to dla każdej bazy $\gamma_1, \dots, \gamma_n$ przestrzeni V układ $\alpha_1, \dots, \alpha_k$ jest zawarty w $\text{lin}(\gamma_1, \dots, \gamma_n)$, a więc $k \leq n$, z twierdzenia Steinitza. \square

Poniższe wnioski wynika w sposób oczywisty z przedstawionych wyżej rozumowań

Wniosek 10.1.5

Niech V będzie przestrzenią skończenie wymiarową. Wówczas:

- (a) Każdy liniowo niezależny układ wektorów V można, dołączając pewną liczbę wektorów, uzupełnić do bazy przestrzeni V ,
- (b) Z każdego układu β_1, \dots, β_m rozpinającego V można wybrać bazę podprzestrzeni V ,
- (c) Jeśli $\dim(V) = k$ i $\alpha_1, \dots, \alpha_k$ jest liniowo niezależnym układem wektorów przestrzeni V , to $\alpha_1, \dots, \alpha_k$ jest bazą przestrzeni K .
- (d) Niech W będzie podprzestrzenią przestrzeni V . Wówczas $\dim W \leq \dim V$. Przy tym jeśli zachodzi $\dim W = \dim V$, to $W = V$.

Dowód. Ad (a). Niech $\alpha_1, \dots, \alpha_k$ będzie układem liniowo niezależnym wektorów przestrzeni V . Wówczas najdłuższy układ liniowo niezależny postaci $\alpha_1, \dots, \alpha_k, \gamma_1, \dots, \gamma_s$ jest bazą przestrzeni V .

Ad (b). Najdłuższy spośród liniowo niezależnych podukładów układu β_1, \dots, β_m jest bazą przestrzeni V .

Ad (c). Układ $\alpha_1, \dots, \alpha_k$ jest maksymalnym układem liniowo niezależnym w V , więc jest bazą.

Ad (d). Na mocy wcześniejszego wniosku $\dim W \leq \dim V$. Jeśli $\dim W = \dim V$, to każda baza przestrzeni W jest też bazą przestrzeni V , więc $W = V$. \square

Kluczowym na tym wykładzie będzie pojęcie rzędu macierzy. Oparte jest ono o prawdziwość następującego zaskakującego (w pewnym sensie) rezultatu.

Twierdzenie 10.1.6

Niech $A \in M_{m \times n}(K)$ oraz niech:

- $w(A) = \dim \text{lin}(\alpha_1, \dots, \alpha_m)$, gdzie $\alpha_1, \dots, \alpha_m \in K^n$ są wierszami macierzy A ,
- $k(A) = \dim \text{lin}(\beta_1, \dots, \beta_n)$, gdzie $\beta_1, \dots, \beta_n \in K^m$ są kolumnami macierzy A .

Wówczas $w(A) = k(A)$. Innymi słowy: dla każdej macierzy A maksymalna liczba liniowo niezależnych wierszy macierzy A jest równa maksymalnej liczbie liniowo niezależnych kolumn macierzy A .

Przykład:

$$A = \begin{bmatrix} 1 & 0 & 1 & 1 & 2 & 3 & 1 & 2 \\ 4 & 1 & 3 & 1 & 0 & 0 & 1 & 1 \end{bmatrix} \in M_{2 \times 8}(\mathbb{R}).$$

- $w(A) = \dim(\text{lin}((1, 0, 1, 1, 2, 3, 1, 2), (4, 1, 3, 1, 0, 0, 1, 1))) = 2$,
- $k(A) = \dim(\text{lin}((1, 4), (0, 1), (1, 3), (1, 1), (2, 0), (3, 0), (1, 1), (2, 1))) = 2$.

Dowód. Przypomnijmy ponownie, że jeśli A' jest macierzą schodkową otrzymaną z A elementarnymi operacjami na wierszach oraz jeśli $\alpha'_1, \dots, \alpha'_r$ to wszystkie niezerowe wiersze macierzy A' , wówczas:

- (i) $\text{lin}(\alpha_1, \dots, \alpha_m) = \text{lin}(\alpha'_1, \dots, \alpha'_r)$,
- (ii) $\alpha'_1, \dots, \alpha'_r$ jest bazą przestrzeni rozpiętej przez wiersze macierzy A .

Widzimy zatem, że $w(A) = r$. Założmy, że w macierzy schodkowej A' pierwsze niezerowe wyrazy w wierszach $\alpha'_1, \dots, \alpha'_r$ znajdują się odpowiednio w kolumnach o indeksach $s_1 < s_2 < \dots < s_r$. Pokażemy, że $\beta_{s_1}, \dots, \beta_{s_r}$ stanowią bazę $\text{lin}(\beta_1, \dots, \beta_n)$. A zatem trzeba dowieść, że wektory te są liniowo niezależne oraz, że rozpinają podprzestrzeń kolumnową macierzy A .

Niech $a_1, \dots, a_r \in K$ oraz $a_1\beta_{s_1} + \dots + a_r\beta_{s_r} = 0$. Niech układ $\beta'_{s_1}, \dots, \beta'_{s_r}$ powstaje z $\beta_{s_1}, \dots, \beta_{s_r}$ przez wykonanie na A elementarnej operacji σ na wierszach $\begin{bmatrix} \beta_1 & \cdots & \beta_n \end{bmatrix} \xrightarrow{\sigma} \begin{bmatrix} \beta'_1 & \cdots & \beta'_n \end{bmatrix}$ to

$$a_1\beta'_{s_1} + \dots + a_r\beta'_{s_r} = 0.$$

Czy to widać? Przy operacji elementarnej następuje albo zamiana współrzędnych wszystkich powyższych wektorów, albo przemnożenie współrzędnych każdego z powyższych wektorów przez stałą, albo dodanie do współrzędnych o numerze j współrzędnych o numerze i przemnożonych przez stałą. Ilustracja (dodanie do j -tego wiersza i -tego przemnożonego przez a):

$$a_1 \begin{bmatrix} b_{1s_1} \\ b_{2s_1} \\ \vdots \\ \color{red}{b_{is_1}} \\ \vdots \\ b_{js_1} + a \cdot \color{red}{b_{is_1}} \\ \vdots \\ b_{ms_1} \end{bmatrix} + \dots + a_r \begin{bmatrix} b_{1s_r} \\ b_{2s_r} \\ \vdots \\ \color{red}{b_{is_r}} \\ \vdots \\ b_{js_r} + a \cdot \color{red}{b_{is_r}} \\ \vdots \\ b_{ms_r} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ \color{red}{0} \\ \vdots \\ 0 + a \cdot \color{red}{0} \\ \vdots \\ 0 \end{bmatrix}$$

Wniosek: układ $\beta_{s_1}, \dots, \beta_{s_r}$ jest liniowo niezależny wtedy i tylko wtedy, gdy $\beta'_{s_1}, \dots, \beta'_{s_r}$ jest liniowo niezależny. Wykonajmy operacje elementarne na wierszach A aż dostaniemy postać zredukowaną A'' .

I teraz **kluczowy argument** całego rozumowania: kolumna s_i -ta β''_{s_i} macierzy zredukowanej A'' to i -ty wektor bazy standardowej ϵ_i przestrzeni K^m , czyli $a_1\epsilon_1 + \dots + a_r\epsilon_r = 0$. Stąd $a_1 = a_2 = \dots = a_r = 0$. A zatem $\beta_{s_1}, \dots, \beta_{s_r}$ to układ liniowo niezależny. Czy jest to układ rozpinający $\text{lin}(\beta_1, \dots, \beta_n)$?

Niech β będzie dowolną kolumną macierzy A . Szukamy a_1, \dots, a_r takich, że $a_1\beta_{s_1} + \dots + a_r\beta_{s_r} = \beta$. Dostajemy układ równań liniowych o macierzy rozszerzonej:

$$U = [\beta_{s_1} \quad \dots \quad \beta_{s_r} \mid \beta] \quad (*)$$

Sprowadzenie macierzy U do postaci zredukowanej U'' odbywa się przy pomocy tych samych operacji, które sprowadzają A do A'' , a więc pierwsze r kolumn U'' to pierwsze r wektorów bazy standardowej.

$$[\beta_{s_1} \quad \dots \quad \beta_{s_r} \mid \beta] \xrightarrow{\dots} [\epsilon_1 \quad \dots \quad \epsilon_r \mid \beta''] = U''$$

Analogicznie jak w rozumowaniu wyżej mamy:

$$a_1\epsilon_1 + \dots + a_r\epsilon_r = \beta''.$$

Ale β'' jest kolumną macierzy A'' (bo była kolumną A), więc ma tylko pierwsze r niezerowych współrzędnych, co oznacza, że układ $(*)$ ma rozwiązanie. Zatem układ $\beta_{s_1}, \dots, \beta_{s_r}$ rozpinia $\text{lin}(\beta_1, \dots, \beta_n)$. A zatem jest to baza tej przestrzeni i ostatecznie $\dim \text{lin}(\beta_1, \dots, \beta_n) = k(A) = r$. \square

Definicja 10.1.7: Rząd macierzy

RZĘDEM MACIERZY $A \in M_{m \times n}(K)$ nazywamy liczbę $\dim \text{lin}(\alpha_1, \dots, \alpha_m) = \dim \text{lin}(\beta_1, \dots, \beta_n)$, gdzie $\alpha_1, \dots, \alpha_m \in K^n$ są wierszami macierzy A , zaś $\beta_1, \dots, \beta_n \in K^m$ są kolumnami macierzy A . Rząd macierzy oznaczamy przez $r(A)$.

Wniosek 10.1.8

Rząd macierzy A równy jest liczbie niezerowych wierszy po doprowadzeniu A do postaci schodkowej.

Przykłady:

$$r \begin{bmatrix} 1 & 1 & 1 \\ 2 & 2 & 2 \\ 3 & 3 & 3 \end{bmatrix} = 1, \quad r \begin{bmatrix} 1 & 1 \\ 1 & 2 \\ 3 & 1 \end{bmatrix} = 2, \quad r \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} = 0.$$

Wykonywanie operacji elementarnych na wierszach nie zmienia rzędu macierzy. Podobnie jest oczywiście z **elementarnymi operacjami na kolumnach macierzy**: dodaniem do kolumny innej kolumny pomnożonej przez stałą, zamianą kolumn, przemnożeniem kolumny przez niezerowy skalar. Stosowanie obydwu typów operacji, zarówno wierszowych jak i kolumnowych, może uprościć wyznaczanie rzędu.

Przykład. Dla $n > 1$ policzyć rząd macierzy $A = [a_{ij}] \in M_{n \times n}(\mathbb{R})$ postaci:

$$A = \begin{bmatrix} -n+1 & 1 & \dots & 1 & 1 \\ 1 & -n+1 & \dots & 1 & 1 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 1 & 1 & \dots & -n+1 & 1 \\ 1 & 1 & \dots & 1 & -n+1 \end{bmatrix}.$$

Wykonujemy następujące operacje.

- Do ostatniego wiersza dodajemy (kolejno) wszystkie pozostałe wiersze.
- Odejmujemy ostatnią kolumnę (kolejno) od każdej z pozostałych kolumn

W ten sposób otrzymujemy:

$$\begin{bmatrix} -n+1 & 1 & \dots & 1 & 1 \\ 1 & -n+1 & \dots & 1 & 1 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 1 & 1 & \dots & -n+1 & 1 \\ \textcolor{red}{0} & \textcolor{red}{0} & \dots & \textcolor{red}{0} & \textcolor{red}{0} \end{bmatrix} = \begin{bmatrix} -n & 0 & \dots & 0 & \textcolor{blue}{1} \\ 0 & -n & \dots & 0 & \textcolor{blue}{1} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \dots & -n & \textcolor{blue}{1} \\ 0 & 0 & \dots & 0 & \textcolor{blue}{0} \end{bmatrix}.$$

Otrzymaliśmy macierz w postaci schodkowej, która ma dokładnie $n - 1$ niezerowych wierszy. A zatem rząd wyjściowej macierzy A równy jest $n - 1$.

Rząd ma wiele własności związanych z działaniami na macierzach. Oto jeden z prostszych przykładów.

Uwaga 10.1.9

Niech $A, B \in M_{m \times n}(K)$. Wówczas

$$r(A + B) \leq r(A) + r(B).$$

Dowód. Niech $\alpha_1, \dots, \alpha_n \in K^m$ będą kolumnami macierzy A oraz niech $\beta_1, \dots, \beta_n \in K^m$ będą kolumnami macierzy B . Kolumny macierzy $A + B$ mają zatem postać $\alpha_1 + \beta_1, \dots + \alpha_n + \beta_n$. Skoro rząd macierzy równy jest wymiarowi przestrzeni kolumnowej, to należy wykazać, że

$$\dim \text{lin}(\alpha_1 + \beta_1, \dots + \alpha_n + \beta_n) \leq \dim(\text{lin}(\alpha_1, \dots, \alpha_n)) + \dim(\text{lin}(\beta_1, \dots, \beta_n)).$$

Niech $\gamma_1, \dots, \gamma_r$ będzie bazą $\text{lin}(\alpha_1, \dots, \alpha_n)$ oraz niech $\delta_1, \dots, \delta_s$ będzie bazą $\text{lin}(\beta_1, \dots, \beta_n)$. Oczywiście $r = r(A), s = r(B)$ oraz każdy wektor $\text{lin}(\alpha_1 + \beta_1, \dots + \alpha_n + \beta_n)$ jest kombinacją liniową wektorów γ_i, δ_j . Zatem z twierdzenia Steinitza dowolna baza tej przestrzeni ma nie więcej niż $r + s$ elementów. \square

Innym ważnym przykładem operacji na macierzach jest ich transponowanie.

Definicja 10.1.10: Macierz transponowana

Niech $A = [a_{ij}] \in M_{m \times n}(K)$. Macierz $B = [b_{ij}] \in M_{n \times m}(K)$ spełniająca $b_{ij} = a_{ji}$ dla każdego i, j nazywamy MACIERZĄ TRANSPONOWANĄ do A i oznaczamy A^T .

Przykład. Jeśli

$$A = \begin{bmatrix} 1 & 0 & 5 \\ 4 & 2 & 3 \end{bmatrix}, \quad \text{to} \quad A^T = \begin{bmatrix} 1 & 4 \\ 0 & 2 \\ 5 & 3 \end{bmatrix}.$$

Wniosek 10.1.11

Dla każdej macierzy A zachodzi $r(A) = r(A^T)$.

10.2 Wybór przykładowych pytań

Proszę się nie uczyć na pamięć odpowiedzi czy wyjaśnienie! Takich pytań można ułożyć znacznie więcej i służą one jedynie sprawdzeniu podstawowego zrozumienia treści zajęć. Jeśli odpowiedź na dowolne z poniższych pytań sprawia Państwu problem, proszę pomyśleć który fragment notatek dotyczy tego pytania, wrócić do ich lektury, a w razie dalszych trudności poprosić prowadzącego o wskazówkę.

1. Jaki jest wymiar podprzestrzeni $\{(x_1, x_2, x_3) \in K^3 : x_3 = 0\}$?
2. Jaki jest wymiar podprzestrzeni $\{(x_1, x_2, x_3) \in K^3 : x_1 = x_2 = x_3\}$?
3. Jaki jest wymiar przestrzeni liniowej wielomianów stopnia ≤ 5 ?
4. Jak jest wymiar podprzestrzeni przestrzeni liniowej $M_{2 \times 2}(K)$ złożonej z macierzy (a_{ij}) spełniających warunek $a_{11} = 0$?
5. Jak jest wymiar podprzestrzeni przestrzeni liniowej $M_{2 \times 2}(K)$ złożonej z macierzy (a_{ij}) spełniających warunek $a_{11} = a_{44}$?
6. Jaki jest wymiar podprzestrzeni przestrzeni liniowej $F(\mathbb{R}, \mathbb{R})$ złożonej z funkcji stałych?
7. Niech $V = \text{lin}(\alpha_1, \alpha_2, \alpha_3, \alpha_4)$. Czy $\dim V$ może być równy 3?
8. Niech $V = \text{lin}(\alpha_1, \alpha_2, \alpha_3)$ oraz $\alpha_1, \alpha_2, \alpha_3$ są liniowo niezależne. Czy $\dim V$ może być równy 2?
9. W przestrzeni liniowej V wymiaru 2 dany jest układ wektorów $\alpha_1, \alpha_2, \alpha_3$. Czy jeden z tych wektorów jest kombinacją liniową pozostałych?
10. Niech W będzie jednowymiarową podprzestrzenią liniową przestrzeni \mathbb{R}^2 . Czy istnieje baza przestrzeni \mathbb{R}^2 której żaden wektor nie należy do W ?
11. W przestrzeni liniowej V dany jest układ liniowo niezależny $\alpha_1, \alpha_2, \alpha_3$. Wiadomo też, że $\dim V = 4$. Niech $\alpha_4 \notin \text{lin}(\alpha_1, \alpha_2, \alpha_3)$. Czy wynika stąd, że $\alpha_1, \alpha_2, \alpha_3, \alpha_4$ jest bazą V ?
12. Czy w przestrzeni macierzy $M_{2 \times 3}(K)$ istnieje 7 macierzy tworzących układ liniowo niezależny?
13. Suma wektorów $\alpha_1 + \dots + \alpha_n$ jest równa 0. Czy wynika stąd, że wymiar przestrzeni $\text{lin}(\alpha_1, \dots, \alpha_n)$ równy jest $n - 1$?
14. Niech V będzie podprzestrzenią przestrzeni liniowej K^n stanowiącą zbiór rozwiązań układu równań $x_1 = x_2 = x_3 = 0$. Czy wymiar przestrzeni V jest równy $n - 3$?
15. Wiadomo, że $\dim \text{lin}(\alpha_1, \dots, \alpha_k) = k - 1$. Czy $v_1 \in \text{lin}(\alpha_2, \alpha_3, \dots, \alpha_k)$?
16. Wiadomo, że $\dim \text{lin}(\alpha_1, \dots, \alpha_k) = k - 1$. Czy możliwe jest, że $\alpha_1, \alpha_2 \in \text{lin}(\alpha_3, \alpha_4, \dots, \alpha_k)$?
17. Ile może być równy rząd niezerowej macierzy rozmiaru 3×4 ?
18. Ile może być równy rząd macierzy, której wszystkie wyrazy są identyczne?
19. Czy dowolne macierze tego samego układu równań liniowych mają ten sam rząd?
20. Czy jeśli dwie macierze rozmiaru $m \times n$ mają równe rzędy, to są macierzami tego samego układu równań liniowych?
21. Macierz A powstała z macierzy B przez permutację jej kolumn. Czy wynika stąd, że $r(A) = r(B)$?
22. W macierzy A skreślamy wiersz i otrzymujemy macierz A' . Czy możliwe jest $r(A) = r(A') + 2$?
23. Macierz A ma trzy kolumny i dwa wiersze oraz suma kolumn macierzy A jest wektorem zerowym. Jakie są możliwe wartości $r(A)$?
24. Niech $A \in M_{3 \times 3}(K)$ i wiadomo, że ostatnia kolumna jest zerowa. Czy wiersze tej macierzy mogą być liniowo niezależne?
25. Niech $A \in M_{3 \times 4}(K)$ i wiadomo, że ostatnia kolumna jest zerowa. Czy wiersze tej macierzy mogą być liniowo niezależne?

10.3 Zadania do samodzielnej pracy

Zadania ilustrujące umiejętności niezbędne do zaliczenia przedmiotu oznaczone są symbolem ♠.

Przy tych zadaniach dodany jest opis umiejętności, które sprawdzają.

- (♣ Znajdowanie wymiaru przestrzeni rozpiętej na układzie wektorów)

Znajdź $\dim W \subset \mathbb{R}^5$ w zależności od $s \in \mathbb{R}$, gdzie

$$W = \text{lin}((10, 3, 9+s, 1, 2-s), (4, 1, 6, 1, 1), (2, 1, -1, -1, -2)).$$

- Znajdź wymiar poniższej przestrzeni liniowej V nad ciałem \mathbb{C}

$$V = \left\{ \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \in M_2(\mathbb{C}) : a_{11} + a_{12} + a_{21} + a_{22} = 0 \right\}.$$

- (♣ Obliczanie rzędów macierzy) Wyznacz rząd macierzy rzeczywistych:

$$\begin{bmatrix} 1 & -1 & 0 & 3 \\ 2 & -3 & 2 & 1 \\ 1 & 2 & 1 & 3 \\ 0 & 4 & 0 & -2 \end{bmatrix}, \quad \begin{bmatrix} 1 & 2 & 3 & 4 & 5 \\ 0 & -2 & -1 & 3 & 1 \\ 2 & 0 & -5 & 3 & -4 \\ -3 & -1 & -2 & 1 & 3 \end{bmatrix}, \quad \begin{bmatrix} 1+i & -1 & 3-2i \\ 0 & 2+3i & -1 \\ 0 & 5-i & 9 \\ -1 & 8 & 7i \end{bmatrix}.$$

Wyznacz rzędy macierzy rzeczywistych, w zależności od parametrów $a, b, c \in \mathbb{R}$ oraz $s, t \in \mathbb{R}$:

$$\begin{bmatrix} a & -b & 1 \\ b & a & 1 \\ 1 & 1 & c \end{bmatrix}, \quad \begin{bmatrix} 10 & -1 & -1 & 3 \\ 2s & -3 & 2 & 1 \\ 4 & 2 & t+3 & 3 \\ 0 & -2 & 3 & 1 \end{bmatrix}, \quad \begin{bmatrix} 3 & 4 & 2 & 1 & 2 \\ 2 & 3 & 1 & 4 & 6 \\ 1 & 2 & t^2-2t & 7 & 10 \\ 4 & 5 & 3 & -t & -2 \end{bmatrix}$$

- W pewnej macierzy $P \in M_{6 \times 6}(\mathbb{R})$ zasłonięto wszystkie niezerowe wyrazy zastępując je symbolem *.

W rezultacie otrzymano:

$$P = \begin{bmatrix} * & * & * & 0 & 0 & 0 \\ * & * & * & 0 & 0 & 0 \\ * & * & * & 0 & 0 & 0 \\ 0 & * & 0 & * & * & 0 \\ 0 & * & * & 0 & * & * \\ 0 & 0 & * & * & 0 & * \end{bmatrix}$$

Jaki jest najmniejszy możliwy rząd macierzy P ?

- Niech macierz B powstaje z macierzy A poprzez wykreślenie s wierszy i t kolumn. Wykaż, że

$$r(A) \leq s+t+r(B).$$

- Wyznacz bazę i wymiar podprzestrzeni przestrzeni $F(\mathbb{R}, \mathbb{R})$ złożonej z funkcji wielomianowych f spełniających warunek $f(1) = f(2) = 0$.

- Dana jest przestrzeń liniowa V nad ciałem \mathbb{R} oraz jej podprzestrzenie W_1, W_2 , przy czym

$$0 < \dim W_1, \dim W_2 < \dim V < \infty.$$

Wykaż, że istnieje element $\alpha \in V$, taki że $\alpha \notin W_1$ oraz $\alpha \notin W_2$. Wykaż dalej, że istnieje baza przestrzeni V , taka że żaden z jej elementów nie jest zawarty ani w W_1 , ani w W_2 . Czy teza jest prawdziwa dla przestrzeni liniowej nad ciałem \mathbb{Z}_2 ?

- Wykaż, że $\dim V = \infty$ wtedy i tylko wtedy, gdy istnieje ciąg wektorów v_1, v_2, \dots przestrzeni V taki, że dla każdego n układ $\{v_1, \dots, v_n\}$ jest liniowo niezależny.

- Niech $\dim V = \infty$ oraz niech $\mathcal{A} = \{\alpha_i\}_{i \in \mathbb{N}}$ będzie bazą przestrzeni V . Dla każdej liczby $n \in \mathbb{N}$ niech $V_n = \text{lin}(\alpha_1, \dots, \alpha_n)$. Wykaż, że dla każdej skończonej wymiarowej podprzestrzeni $W \subset V$ istnieje $n \in \mathbb{N}$, dla którego W jest podprzestrzenią przestrzeni V_n .

- W przestrzeni liniowej macierzy 2×2 o współczynnikach rzeczywistych rozpatrzmy zbiór macierzy rzędu 2, czyli $\mathcal{A} = \{A \in M_{2 \times 2}(\mathbb{R}) : r(A) = 2\}$. Czy $\text{lin}(\mathcal{A}) = M_{2 \times 2}(\mathbb{R})$? Odpowiedź uzasadnij.

- Wykaż, że jeśli $A = [a_{ij}] \in M_{3 \times 3}(\mathbb{R})$ jest niezerową macierzą spełniającą warunek $a_{ij} = -a_{ji}$, dla każdych i, j , to $r(A) = 2$.

- Niech $n \geq 1$ oraz $a_1, \dots, a_n \in \mathbb{R}$. Rozważmy macierz A rozmiaru $n! \times n$, której wierszami są wszystkie możliwe permutacje ciągu a_1, \dots, a_n . Wyznacz możliwe wartości liczby $r(A)$.

10.4 Uzupełnienie. Zadanie o macierzach półmagicznych

Opowiemy tu o pewnym trudnym zadaniu o wymiarze, które pojawiło się kiedyś na kolokwium.

Zadanie. Niech $I_n \in M_{n \times n}(\mathbb{Q})$ będzie macierzą, której jedyne niezerowe wyrazy znajdują się na przekątnej i są równe 1. Zbiór \mathcal{P}_n zawiera wszystkie macierze powstałe z I_n przez wykonanie dowolnie wielu operacji elementarnych zamiany wierszy (w tym I_n). Niech $V = \text{lin}(\mathcal{P}_n)$. Wykazać, że:

- (a) jeśli $A \in V$, to suma wyrazów w każdym wierszu i w każdej kolumnie macierzy A jest taka sama,
- (b) jeśli W_0 jest podprzestrzenią $M_{n \times n}(\mathbb{Q})$ złożoną z macierzy o zerowej sumie wyrazów w każdym wierszu i w każdej kolumnie oraz jeśli $F_{ij} \in W_0$ jest taką macierzą, która na pozycjach $(i, j), (n, n)$ ma 1, na pozycjach $(i, n), (n, j)$ ma -1 oraz na pozostałych pozycjach ma 0, to $W_0 = \text{lin}(F_{ij}, 1 \leq i, j \leq n-1)$,
- (c) jeśli macierz $R_{(i,j,n)}$ powstaje z I_n przez dwie kolejno wykonane operacje elementarne: zamianę wiersza i -tego z n -tym, a następnie zamianę wiersza j -tego z n -tym, dla $1 \leq i < j < n$, oraz jeśli macierz $S_{(k,n)}$ powstaje z I_n przez zamianę k -tego i n -tego wiersza, dla $1 \leq k < n$, to układ macierzy:

$$\{R_{(i,j,n)}, 1 \leq i < j \leq n-1\} \cup \{S_{(k,n)}, 1 \leq k \leq n-1\} \cup \{I_n\}$$

jest bazą V . W szczególności $\dim(V) = (n-1)^2 + 1$.

Rozwiążanie. Dowód (a). Jest jasne, że każda macierz ze zbioru \mathcal{P}_n posiada dokładnie jeden niezerowy element w każdym wierszu i w każdej kolumnie, który jest równy 1 (można to udowodnić na przykład przez indukcję względem liczby operacji elementarnych zamiany wierszy wykonanych na I_n). A zatem każda macierz z tego zbioru ma jednakową sumę wyrazów w każdym wierszu i w każdej kolumnie. Wynosi ona 1. Zauważmy teraz, że jeśli pewne dwie macierze $A, B \in M_{n \times n}(\mathbb{Q})$ mają tę własność, że suma wyrazów w każdym wierszu i w każdej kolumnie każdej z tych macierzy jest jednakowa i wynosi odpowiednio s_A oraz s_B , to macierz $A + B$ oraz macierz aA , gdzie $a \in \mathbb{Q}$, też mają tę własność, że suma wyrazów w każdym wierszu i w każdej kolumnie jest jednakowa, i wynosi ona odpowiednio $s_A + s_B$ oraz as_A . A zatem dowolna kombinacja liniowa macierzy o tej własności też ma tę własność. W szczególności $V = \text{lin}(\mathcal{P}_n)$ złożona jest z macierzy o tej własności (nie wiemy jednak czy każda macierz o tej własności należy do $\text{lin}(\mathcal{P}_n)$ – to pokażemy dalej).

Dowód (b). Macierz F_{ij} ma postać (przykład dla $i, j > 1$):

$$F_{ij} = \begin{bmatrix} 0 & \dots & \color{red}{0} & \dots & 0 \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ \color{red}{0} & \dots & 1 & \dots & -1 \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ 0 & \dots & -1 & \dots & 1 \end{bmatrix},$$

gdzie na czerwono podkreślone zostały i -ty wiersz i j -ta kolumna. Niech $Z = (z_{ij}) \in M_n(\mathbb{R})$ będzie macierzą, w której sumy wyrazów w każdym wierszu i każdej kolumnie wynoszą 0. Wówczas dla każdego $1 \leq i \leq n$ kombinacja liniowa

$$z_{i1}F_{i1} + z_{i2}F_{i2} + \dots + z_{i,n-1}F_{i,n-1}$$

równa jest (zgodnie z założeniem o sumie wyrazów w i -tym wierszu $z_{i1} + z_{i2} + \dots + z_{i,n-1} + z_{i,n} = 0$):

$$\begin{bmatrix} 0 & 0 & \dots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ z_{i1} & z_{i2} & \dots & z_{i,n-1} & -z_{i1} - z_{i2} - \dots - z_{i,n-1} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ -z_{i1} & -z_{i2} & \dots & -z_{i,n-1} & z_{i1} + z_{i2} + \dots + z_{i,n-1} \end{bmatrix} = \begin{bmatrix} 0 & 0 & \dots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ z_{i1} & z_{i2} & \dots & z_{i,n-1} & z_{i,n} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ -z_{i1} & -z_{i2} & \dots & -z_{i,n-1} & -z_{i,n} \end{bmatrix}$$

Oznacza to, że macierz postaci

$$\sum_{i=1}^{n-1} z_{i1}F_{i,1} + z_{i2}F_{i,2} + \dots + z_{i,n-1}F_{i,n-1}$$

można zapisać jako:

$$\begin{bmatrix} z_{11} & \dots & z_{1,n} \\ 0 & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & 0 \\ -z_{11} & \dots & -z_{1,n} \end{bmatrix} + \begin{bmatrix} 0 & \dots & 0 \\ z_{21} & \dots & z_{2,n} \\ \vdots & \ddots & \vdots \\ 0 & \dots & 0 \\ -z_{21} & \dots & -z_{2,n} \end{bmatrix} + \dots + \begin{bmatrix} 0 & \dots & 0 \\ 0 & \dots & 0 \\ \vdots & \ddots & \vdots \\ z_{n-1,1} & \dots & z_{n-1,n} \\ -z_{n-1,1} & \dots & -z_{n-1,n} \end{bmatrix} = Z.$$

W ostatniej równości korzystamy z założenia, że suma wyrazów każdej z kolumn jest zerowa, a więc mamy na przykład równość

$$-z_{11} - z_{21} - \dots - z_{n-1,1} = z_{n1}.$$

W rezultacie macierz Z jest kombinacją liniową macierzy F_{ij} postaci:

$$Z = \sum_{i=1}^{n-1} \sum_{j=1}^{n-1} z_{ij} F_{ij}.$$

Pokazaliśmy zatem, że każdy element podprzestrzeni W_0 jest kombinacją liniową macierzy F_{ij} .

Dowód (c). Znowu warto się przyjrzeć jakie permutacje opisują macierze $R_{(i,j,n)}$ oraz $S_{(k,n)}$. Pierwsza z nich umieszcza w j -tej kolumnie **i -ty wektor standardowy**, w n -tej kolumnie umieszcza **j -ty wektor standardowy** oraz **n -ty wektor standardowy** umieszcza w i -tej kolumnie (pozostałe są na swoich pozycjach). Podobnie macierz $S_{(k,n)}$ powstaje przez zamianę k -tej oraz **n -tej** kolumny. Oto ilustracja.

$$R_{(i,j,n)} = \begin{bmatrix} 0 & \dots & 0 & \dots & 0 & \dots & 0 \\ \vdots & \ddots & \vdots & \ddots & \vdots & \ddots & \vdots \\ 0 & \dots & 0 & \dots & 1 & \dots & 0 \\ \vdots & \ddots & \vdots & \ddots & \vdots & \ddots & \vdots \\ 0 & \dots & 0 & \dots & 0 & \dots & 1 \\ \vdots & \ddots & \vdots & \ddots & \vdots & \ddots & \vdots \\ 0 & \dots & 1 & \dots & 0 & \dots & 0 \end{bmatrix}, \quad S_{(k,n)} = \begin{bmatrix} 0 & \dots & 0 & \dots & 0 \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ 0 & \dots & 0 & \dots & 1 \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ 0 & \dots & 1 & \dots & 0 \end{bmatrix},$$

Zauważmy, że jeśli $i \neq j$ oraz $i, j < n$, a także jeśli $k < n$, wówczas mamy:

$$F_{ij} = I_n - S_{(i,n)} - S_{(j,n)} + R_{(i,j,n)}, \quad F_{k,k} = I_n - S_{(k,n)}. \quad (\star)$$

Rzeczywiście F_{ij} równe jest (można też to policzyć rozkładając wszystko na jedynki macierzowe):

Zauważmy też, że macierze $R_{(i,j,n)}$ oraz $S_{(k,n)}$ należą do \mathcal{P}_n . A zatem na mocy punktów (a) i (b) mamy:

$$W_0 \stackrel{(b)}{=} \text{lin}(F_{ij}) \stackrel{(\star)}{\subseteq} \text{lin}(\{R_{(i,j,n)}\} \cup \{S_{(k,n)}\} \cup \{I_n\}) \subseteq \text{lin}(\mathcal{P}_n) = V \stackrel{(a)}{\subseteq} W,$$

gdzie W jest podprzestrzenią wszystkich macierzy, które mają taką samą sumę wyrazów w każdym wierszu i w każdej kolumnie (oczywiście $W_0 \subseteq W$).

Zauważmy jednak, że jeśli S jest macierzą, której suma elementów w każdym wierszu i każdej kolumnie równa jest t , to $S - tI_n \in W_0$. Każda macierz z W jest w rezultacie, na mocy punktu (b), kombinacją liniową macierzy F_{ij} oraz I_n . A zatem $W = \text{lin}(F_{ij} \cup \{I_n\})$. Stąd:

$$W = \text{lin}(F_{ij} \cup \{I_n\}) \subseteq \text{lin}(\{R_{(i,j,n)}\} \cup \{S_{(k,n)}\} \cup \{I_n\}) \subseteq V \subseteq W \Rightarrow V = W.$$

Układ $\{F_{ij}, 1 \leq i, j < n\}$ jest liniowo niezależny. Wynika to natychmiast z dowodu (b), gdzie opisaliśmy każdą liniową kombinację wektorów F_{ij} . Jeśli macierz Z z punktu (b) jest zerowa, to wszystkie z_{ij} są zerowe. Skoro więc $I_n \notin W_0$, to widzimy, że układ $\{F_{ij}, 1 \leq i, j < n\}$ tworzy bazę W_0 , a układ

$$\{F_{ij} \cup I_n, 1 \leq i, j < n\}$$

tworzy bazę W . Dostajemy zatem $\dim(V) = \dim(W) = (n-1)^2 + 1$. Jednak zbiór

$$\{R_{(i,j,n)}, 1 \leq i < j \leq n-1\} \cup \{S_{(k,n)}, 1 \leq k \leq n-1\} \cup \{I_n\}$$

rozpinia $W = V$ oraz ma $(n-1)^2 + 1$ elementów. Jest to zatem minimalny układ rozpinający $W = V$, a więc także baza przestrzeni V . Dowód jest zakończony.

10.5 Dodatek. Stopień rozszerzenia ciała

Przypomnijmy, że jeśli dane jest ciało L i jego podciało K , to nietrudno sprawdzić, że L ma strukturę przestrzeni liniowej nad ciałem K . Zajmiemy się teraz własnościami wymiaru tej przestrzeni.

Definicja 10.5.1: Podciało i rozszerzenie ciała

Parę $K \subset L$, gdzie K jest podciąłem ciała L nazywamy ROZSZERZENIEM CIAŁ. Wymiar L jako przestrzeń liniową nad K nazywamy STOPNIEM ROZSZERZENIA K i oznaczamy $[L : K]$. Gdy liczba ta jest skończona, to ciało L nazywamy ROZSZERZENIEM SKOŃCZONYM ciała K .

Przykłady

- Mamy $[\mathbb{C} : \mathbb{R}] = 2$. Wynika to z faktu, że każdy element ciała \mathbb{C} może być jednoznacznie przedstawiony w postaci $a + bi$, gdzie $a, b \in \mathbb{R}$.
- Rozszerzenie $\mathbb{C} \supset \mathbb{Q}$ również zadaje na \mathbb{C} strukturę przestrzeni liniowej nad ciałem K . Nie jest to jednak rozszerzenie skończone.
- Rozważmy ciało czteroelementowe $K = \{0, 1, a, b\}$, skonstruowane w Uzupełnieniu 3.4. Ciało to oczywiście zawiera jako podciąłem ciało \mathbb{Z}_2 . Co więcej, zachodzą równości $a^2 + a + 1 = 0$ oraz $b^2 + b + 1 = 0$. Oznacza to, że wielomian $x^2 + x + 1 \in K[x]$ ma pierwiastki a oraz b , i tylko te. Mamy też $\mathbb{Z}_2(a) = \mathbb{Z}_2(b) = K$. Wreszcie, nietrudno widzieć, że traktując K jako przestrzeń liniową nad ciałem \mathbb{Z}_2 , mamy $\dim K = 2$.
- Każde ciało skończone o p^n elementach jest n -wymiarową przestrzenią liniową nad ciałem \mathbb{Z}_p , co wynika bezpośrednio z Wniosku 3.4.3.

Twierdzenie 10.5.2

Jeśli dane są rozszerzenia ciał $K \subseteq L$ oraz $L \subseteq M$, przy czym $[L : K] = n$ oraz $[M : L] = m$, wówczas $[M : K] = mn$.

Dowód. Rozważmy układ x_1, \dots, x_n elementów L będący bazą L nad K oraz układ y_1, \dots, y_m elementów M będący bazą M nad L . Wykażemy, że układ

$$X = \{x_i y_j \mid 1 \leq i \leq n, 1 \leq j \leq m\}$$

jest bazą M nad K .

Niech $c_{ij} \in K$ będą takie, że $\sum_{i=1}^n \sum_{j=1}^m c_{ij} x_i y_j = 0$. Możemy tę równość zapisać w postaci:

$$\sum_{j=1}^m (c_{1j} x_1 + \dots + c_{nj} x_n) y_j = 0.$$

Współczynniki stojące wyżej przy y_j są elementami ciała L , natomiast y_j są liniowo niezależne nad L , skąd dla każdego j mamy $c_{1j} x_1 + \dots + c_{nj} x_n = 0$. Skoro zaś x_i są liniowo niezależne nad K , to $c_{ij} = 0$. Układ X jest więc liniowo niezależny.

Weźmy dowolny element $y \in M$. Wykażemy, że $y \in \text{lin}(X)$. Skoro y_1, \dots, y_m jest bazą M nad L , to istnieją $c_1, \dots, c_m \in L$, że $y = c_1 y_1 + \dots + c_m y_m$. Każdy z elementów c_i można natomiast przedstawić jako kombinację liniową elementów x_i , co kończy dowód. \square

Wniosek 10.5.3

Jeśli dany jest ciąg ciał $K_0 \subset K_1 \subset K_2 \subset \dots \subset K_m$ oraz $[K_i : K_{i-1}] = n_i$, dla $i = 1, 2, \dots, m$, to $[K_m : K_0] = n_1 n_2 \dots n_m$.

Przejdziemy teraz do pewnej ważnej klasy rozszerzeń, charakteryzującej, jak się okaże, ciała będące skończeniem wymiarowymi przestrzeniami nad swoimi podciążami. Teoria tych rozszerzeń ma fundamentalne znaczenie dla problemu rozwiązywalności równań wielomianowych.

Definicja 10.5.4: Rozszerzenie algebraiczne

Niech $K \subset L$ będą ciałami. Powiemy, że element $a \in L$ jest ALGEBRAICZNY nad K , jeśli istnieje niezerowy wielomian $f \in K[x]$ taki, że $f(a) = 0$. Rozszerzenie $K \subseteq L$ nazywamy ALGEBRAICZNYM, jeśli każdy element L jest algebraiczny nad K .

W dalszym ciągu przyjmiemy następującą notację. Jeśli $K \subseteq L$ jest rozszerzeniem ciała i $a \in L$, to przez $K(a)$ oznaczamy najmniejsze podciało L , które zawiera K oraz element a i nazywamy rozszerzeniem ciała K o element a . W dalszym ciągu przyjrzymy się jak wygląda baza i wymiar $[K(a) : K]$, gdy a jest elementem algebraicznym. Wykorzystamy w tym celu rezultaty uzasadnione w Uzupełnieniu 5.4 dotyczącym teorii podzielności wielomianów.

Przykłady

- Liczby rzeczywiste $\sqrt{2}$, i , $\sqrt[3]{3}$, $\sqrt{1+\sqrt{2}}$ są algebraiczne nad \mathbb{Q} , bowiem są pierwiastkami wielomianów wymiernych $x^2 - 2$, $x^2 + 1$, $x^3 - 3$, $x^4 - 2x^2 - 1$.
- Rozszerzenia $K \subseteq K(a)$, że $a \notin K$ jest rozwiązaniem pewnego równania wielomianowego stopnia 2 o współczynnikach w ciele K nazywamy *kwadratowymi*.
- Rozszerzenie ciała \mathbb{Q} o pierwiastek stopnia n z 1 nazywamy *cyklotomicznym*.

Naszym celem będzie pokazanie, że każdy element algebraiczny nad ciałem K jest pierwiastkiem pewnego wielomianu nierozkładalnego z $K[x]$. W tym celu użyjemy pojęcia największego wspólnego dzielnika elementów z $K[x]$.

Na mocy Twierdzenia 5.4.2 o dzieleniu z resztą, dla wielomianów $f, g \in K[x]$ określić można *największy wspólny dzielnik* wielomianów f, g , a więc taki wielomian $d = NWD(f, g)$ największego możliwego stopnia, który dzieli zarówno f , jak i g . Wielomian ten jest wyznaczony z dokładnością do stałej, więc możemy przyjąć, że jest to wielomian unormowany. Z Twierdzenia 5.4.2 wynika, że jeśli mamy wielomian e , który dzieli zarówno f , jak i g , to e dzieli również $NWD(f, g)$.

Definicję największego wspólnego dzielnika niezerowych wielomianów o współczynnikach w ciele można oczywiście rozszerzyć na dowolny skończony układ wielomianów. Możemy też sformułować uogólnienie lematu Bezout, czyli Twierdzenia 5.4.4: dla dowolnego układu niezerowych wielomianów p_1, \dots, p_n istnieją takie wielomiany q_1, \dots, q_n , że

$$q_1 p_1 + \dots + q_n p_n = NWD(p_1, \dots, p_n).$$

Twierdzenie 10.5.5

Jeśli $K \subseteq L$ i $a \in L$ jest algebraiczny nad K , to istnieje wielomian nierozkładalny $f \in K[x]$, że $\deg f > 0$ i $f(a) = 0$. Ponadto, dla każdego wielomianu $g \in K[x]$, że $g(a) = 0$, wielomian f jest dzielnikiem g .

Dowód. Skoro a jest elementem algebraicznym, to istnieje wielomian $w \in K[x]$, że $w(a) = 0$. Rozłożymy w na iloczyn $f_1 \dots f_k$ czynników nierozkładalnych, zgodnie z Twierdzeniem 2.6.3. Wówczas zachodzi również równość

$$f_1(a) \cdot \dots \cdot f_k(a) = 0,$$

czyli jedna z liczb $f_i(a)$ jest zerem. Weźmy dowolne i o tej własności i określmy $f = f_i$. Twierdzimy, że jest to szukany wielomian.

Niech $g \in K[x]$ spełnia $g(a) = 0$ i niech $d = NWD(f, g)$. Skoro f jest nierozkładalny, to $d = f$ lub $d = 1$. Z uogólnionej wersji Lematu Bezout mamy wielomiany $r, s \in K[x]$, że

$$d = rf + sg.$$

Stąd podstawiając a dostajemy $d(a) = 0$. Stąd $d \neq 1$, czyli $d = f$. \square

Widzimy zatem, że gdy element $a \in L$ jest algebraiczny nad ciałem K , to wielomian nierozkładalny f , taki że $f(a) = 0$, jest wyznaczony z dokładnością do stałej.

Definicja 10.5.6: Stopień elementu algebraicznego

Gdy $K \subseteq L$ oraz $a \in L$ jest elementem algebraicznym nad K , to stopień wielomianu nieroziadalnego f , że $f(a) = 0$ nazywamy STOPNIEM ELEMENTU a nad K . Wielomian f nazywamy WIELOMIANEM MINIMALNYM elementu a .

Przejdziemy teraz do uzasadnienia podstawowego rezultatu teorii rozszerzeń algebraicznych, który pomoże nam też przestawić alternatywny dowód Wniosku 3.4.3 opisującego moc ciał skończonych, pozwalający dodatkowo na konstrukcję ciała skończonego dowolnej mocy.

Twierdzenie 10.5.7

Niech $K \subseteq L$ będzie rozszerzeniem ciał. Niech $\alpha \in L$ będzie elementem algebraicznym stopnia n nad K . Wówczas:

- Każdy element ciała $K(\alpha)$ jest postaci $c_{n-1}\alpha^{n-1} + \dots + c_1\alpha + c_0$, dla pewnych $c_0, \dots, c_{n-1} \in K$.
- Bazą ciała $K(\alpha)$ względem K jest układ $1, \alpha, \dots, \alpha^{n-1}$.
- $[K(\alpha) : K] = n$.

Dowód. Niech $X = \{g(\alpha) \in L \mid g \in K[x], \deg g < n\}$. Jest jasne, że X jest zbiorem zawierającym $K(\alpha)$. Wykażmy, że jest to podział L . Oczywiście suma elementów X jest w tym zbiorze.

Rozważmy natomiast niezerowe elementy $g_1(\alpha), g_2(\alpha)$, gdzie $\deg g_i < n$. Niech f będzie wielomianem minimalnym a . Z założenia ma on stopień n . Na mocy twierdzenia o dzieleniu z resztą istnieje wielomian r stopnia mniejszego od n , który jest resztą z dzielenia wielomianu g_1g_2 przez f . Mamy więc $g_1(\alpha)g_2(\alpha) = r(\alpha)$. Skoro mamy $\deg r < \deg f$, to $r(\alpha) \neq 0$. Zatem $g_1(\alpha)g_2(\alpha)$ należy do X .

Pozostało wykazać, że każdy niezerowy element $g(\alpha) \in X$ jest odwracalny, gdzie $g \in K[x]$ jest wielomianem stopnia $< n$. Mamy $g(\alpha) \neq 0$ oraz $f(\alpha) = 0$, więc f nie jest dzielnikiem g . Skoro jednak f jest nieroziadalny, to $NWD(f, g) = 1$. Z Lematu Bezout mamy więc wielomiany $r, s \in K[x]$, że $1 = rf + sg$. Podstawiając α do uzyskanej równości, dostajemy $s(\alpha)g(\alpha) = 1$.

Wykazaliśmy zatem, że $X = K(\alpha)$. Przeprowadzona konstrukcja pokazuje, że każdy element ciała $K(\alpha)$ jest kombinacją liniową elementów $1, \alpha, \dots, \alpha^{n-1}$ (bierzemy $g = x^i$, dla $i < n$). Pozostaje więc uzasadnić liniową niezależność tego układu. Jeśli istnieją $c_0, \dots, c_{n-1} \in K$, że $c_0 + c_1\alpha + \dots + c_{n-1}\alpha^{n-1} = 0$, to biorąc wielomian $g = c_{n-1}\alpha^{n-1} + \dots + c_1\alpha + c_0$, mamy $g(\alpha) = 0$. Zatem zgodnie z Twierdzeniem 10.5.5 mamy $f \mid g$. Skoro jednak $\deg f = n$, to g jest wielomianem zerowym. □

Wniosek 10.5.8

Niech $K \subseteq L$. Wówczas element $a \in L$ jest algebraiczny nad K wtedy i tylko wtedy, gdy stopień a nad K jest skończony. W szczególności każde rozszerzenie skończone jest algebraiczne.

Dla konstrukcji rozszerzeń algebraicznych przydatny będzie jeszcze jeden wniosek.

Wniosek 10.5.9

Niech $g \in K[t]$ i $\deg g > 0$. Istnieje takie rozszerzenie L ciała K , że wielomian $g \in L[x]$ ma w ciele L pierwiastek.

Dowód. Jeśli f jest dowolnym dzielnikiem nieroziadalnym g , to szukanym ciałem jest zbiór $X_f \subset K[t]$ wielomianów stopnia mniejszego od $\deg f$. Wprowadzamy w nim działania dodawania i mnożenia modulo f , definiowane analogicznie jak działania w ciele \mathbb{Z}_p . Podobnie jak w dowodzie Twierdzenia 10.5.7 pokazujemy, że X_f jest ciałem.

Biorąc wielomian $t \in X_f$ stwierdzamy, że t jest pierwiastkiem wielomianu f . Istotnie, jeśli przyjmiemy $f = a_n x^n + \dots + a_1 x + a_0$, to żeby policzyć $f(t)$, trzeba znać wartość $a_n t^n$, a ten element jest, zgodnie z definicją X_f , resztą z dzielenia wielomianu $a_n t^n$ przez f , a to jest $-a_{n-1} t^{n-1} - \dots - a_1 t - a_0$, czyli

$$f(t) = (-a_{n-1} t^{n-1} - \dots - a_1 t - a_0) + a_{n-1} t^{n-1} + \dots + a_1 t + a_0 = 0.$$

□

Przykłady

- Rozważmy wielomian $f = t^2 + t + 1 \in \mathbb{Z}_2[t]$. Wielomian ten nie ma pierwiastków w ciele \mathbb{Z}_2 . Wiemy natomiast, że ciała czteroelementowego $K = \{0, 1, a, b\}$ mamy $a^2 + a + 1 = b^2 + b + 1 = 0$. Formalnie konstrukcję ciała K można przeprowadzić jednak podobnie jak w twierdzeniu wyżej.

Wielomian f jest nierozkładalny nad \mathbb{Z}_2 . Rozważmy wszystkie wielomiany nad \mathbb{Z}_2 stopnia mniejszego od 2, czyli $X_f = \{0, 1, t, t + 1\}$. Na zbiorze tym definiujemy działania modulo f , czyli na przykład $t(t + 1) = -1$, gdyż $t(t + 1) = t^2 + t + 1 - 1$. Jak się okazuje, przyjmując $a = t$, $b = t + 1$ uzyskujemy w zbiorze X_f strukturę ciała identyczną jak ciała K (o tych samych tabelkach działań).

- Rozważmy dowolny wielomian nierozkładalny stopnia n nad ciałem K . Uzyskane powyżej rezultaty oznaczają, że istnieje rozszerzenie L ciała K , które jest stopnia n , i które można skonstruować jako ciało reszt modulo f .

W szczególności, biorąc ciało \mathbb{Z}_p i dowolny wielomian nierozkładalny f stopnia n nad \mathbb{Z}_p uzyskujemy konstrukcję ciała reszt modulo f nad K , które ma wymiar n nad \mathbb{Z}_p , czyli ma p^n elementów.

Podstawowy problem konstrukcji opisanej wyżej brzmi: czy dla każdego ciała K istnieje wielomian nierozkładalny dowolnego stopnia? Dopiero wtedy wiedzielibyśmy na przykład, że istnieje ciało skończonego dowolnej mocy p^k . To nie jest jednak elementarny problem. Warto poczynić następujący komentarz.

Można udowodnić, choć nie jest to łatwe twierdzenie, że w każdym ciele skończonym K istnieje taki element α , że każdy niezerowy element ciała K jest pewną jego potęgą. W ten sposób ciało K staje się rozszerzeniem algebraicznym ciała \mathbb{Z}_p o pojedynczy element i wielomian minimalny tego elementu jest właśnie szukanym wielomianem nierozkładalnym stopnia n nad \mathbb{Z}_p . Do tego jednak potrzeba mieć ciało, które chcemy dopiero skonstruować.

Na koniec wspomnijmy jeszcze o ważnym historycznym kontekście omawianego tematu. Rozszerzenia kwadratowe wiążą się ze starożytnym zagadnieniem wskazywania tzw. liczb konstruowalnych (nad ciałem \mathbb{Q}), czyli takich długości odcinków, które można skonstruować za pomocą cyrkla i linijki, mając do dyspozycji odcinek długości 1. Chodzi między innymi o konstrukcję kąta o mierze 20° , konstrukcję odcinka o długości $\sqrt[3]{2}$ oraz o konstrukcję siedmiokąta foremnego.

Opisane problemy dotyczą w istocie liczb algebraicznych nad ciałem \mathbb{Q} . Liczba $\cos 20^\circ$ jest pierwiastkiem wielomianu $4x^3 - 3x - \frac{1}{2}$, liczba $\sqrt[3]{2}$ jest pierwiastkiem wielomianu $x^3 - 2$, zaś liczba $\cos \frac{2\pi}{7}$ jest, jak się okazuje, pierwiastkiem wielomianu $64x^7 - 112x^5 + 56x^3 - 7x - 1$. Żadna z powyższych liczb nie jest konstruowalna. Można udowodnić bowiem następujące twierdzenie.

Twierdzenie 10.5.10

Liczby konstruowalne nad ciałem \mathbb{Q} tworzą podciało ciała liczb rzeczywistych. Liczba rzeczywista x jest konstruowalna nad ciałem \mathbb{Q} wtedy i tylko wtedy, gdy istnieje ciąg takich rozszerzeń kwadratowych $\mathbb{Q} \subseteq K_1 \subseteq K_2 \subseteq \dots \subseteq K_n$, że $x \in K_n$.

Wspomniane tu zagadnienia wywodzą się z dwóch ważnych problemów klasycznych: badania rozwiązywalności równań wielomianowych stopnia $n > 4$ przez tak zwane pierwiastniki, które rozstrzygnęła tzw. teoria Galois, oraz prób dowodu Wielkiego Twierdzenia Fermata, które dały początek teorii rozszerzeń całkowitych i w konsekwencji teorii pierścieni. Czytelnika zainteresowanego tymi zagadnieniami, odsyłamy do dedykowanych podręczników teorii ciał i pierścieni, choćby do podręcznika Browkina „Wybrane zagadnienia algebry”.

Więcej rachunków i przykładów na elementarnym poziomie, między innymi wyjaśnienia klasycznych problemów konstruowalności zarysowanych wyżej, znajdą Państwo również w podręczniku prof. Guzickiego „Geometria analityczna. Rozszerzony program matematyki w liceum”.

10.6 Coda. O kształtowaniu się pojęcia wymiaru

W trakcie studiów poznają Państwo szereg podstawowych koncepcji matematycznych, które rozważane będą w coraz ogólniejszym kontekście na kolejnych przedmiotach. Należą do nich chociażby koncepcje liczby, przestrzeni, ciągłości, nieskończoności, niezależności oraz właśnie wymiaru (a także wiele innych). Na kursie algebry liniowej budujemy dość prostą, ale niezwykle użyteczną teorię przestrzeni liniowych, zbudowaną wokół pojęcia wektora i pewnych podstawowych aksjomatów motywowanych geometrycznie. W ten sposób możliwe jest mówienie o pewnych zjawiskach mających podłożem geometryczne i możliwe jest określenie choćby pojęcia wymiaru. Nie jest to pojęcie banalne, a jego historia jest bardzo ciekawa.

Czytelnik pytać może — czy wystarczy mówić o wymiarze przestrzeni liniowych? A jeśli chcemy rozważać ogólniejsze przestrzenie, choćby przestrzenie metryczne lub topologiczne, rozmaitości, czy fraktale — jak dla tych obiektów definiować wymiar? Historycznie rzecz biorąc właściwej dynamiki problem ten nabiera w wieku XIX, wraz z odkryciem przez Cantora, że prosta i płaszczyzna są w istocie równoliczne oraz wraz ze znalezieniem przez Peano ciągłego odwzorowania odcinka na kwadrat. Co było wcześniej?

W świecie Pitagorejczyków wymiary miały znaczenie filozoficzne i do pewnego stopnia religijne. Sam Arystoteles pisze w swoim dziele *O niebie*:

Wielkość, jeśli jest podzielna w jedną stronę, jest linią, jeśli w dwie strony – powierzchnią, a jeśli w trzy – ciałem. Poza nimi nie ma innej wielkości, ponieważ istnieją tylko trzy wymiary, a to, co jest podzielne w trzech kierunkach, jest podzielne we wszystkich [...] Bo, jak mówią pitagorejczycy, wszechświat i wszystko, co się w nim znajduje, jest określone przez liczbę trzy, ponieważ początek, środek i koniec dają liczbę wszechświata, a liczba, którą podają, jest triadą. I tak, wziawszy te trzy z natury jako (że tak powiem) jej prawa, dalej używamy liczby trzy w kulcie Bogów.

Także u Euklidesa występują obiekty wymiaru 1, 2 czy 3, a obiektom trójwymiarowym poświęca się sporo miejsca. Jednym ze szczytowych wszakże osiągnięć *Elementów* i starożytnej matematyki jest klasyfikacja wielościanów foremnych. Poza trzema wymiarami, ani Arystoteles, ani Euklides czy Ptolemeusz nie widzieli innych możliwych opcji. Ten stan rzeczy zachował się w zasadzie aż do wieku XVIII. I tak dla przykładu Kepler argumentował będzie, że liczba wymiarów równa 3 istnieje na chwałę Trójcy Świętej, Wallis w swoim wykładzie algebry z roku 1686 napisze, że wyżej wymiarowa przestrzeń jest to „potwór z natury, mniej możliwy niż chimera czy centaur”, a trójwymiarowość przestrzeni dowodzić będzie próbował (bez powodzenia) w swoim doktoracie sam Immanuel Kant (1747).

Skąd wzięła się potrzeba wyjścia poza trzy wymiary? Z jednej strony, kluczowym elementem było przeniesienie rozważań geometrycznych do układu współrzędnych, czyli osiągnięcie Kartezjusza. W ten sposób teoria krzywych czy powierzchni, stała się teorią równań z wieloma niewiadomymi. Na różne sposoby badano intuicję „stopni swobody”, czyli liczby niezależnych parametrów potrzebnych do opisu ruchu punktu znajdującego się na obiekcie, którego wymiar rozważamy. Także z naszego punktu widzenia — rozwiązywanie układu równań jednorodnych zależnych od k parametrów opisuje podprzestrzeń wymiaru k , i z każdego rozwiązania można „dotrzeć do innego” przy użyciu k liniowo niezależnych rozwiązań.

Na nasze potrzeby wystarczy stwierdzić, że już w połowie XVIII wieku zorientowano się, że do rozwiązywania niektórych zagadnień (choćby pochodzących z mechaniki) nie wystarczy rozważanie nie więcej niż trzech parametrów. Idea ta wyrażona została wprost przez Jeana d'Alemberta w Wielkiej Encyklopedii Francuskiej — pomniku Oświecenia redagowanym wspólnie z Diderotem. We wpisie „Wymiar” wpisana była idea reprezentowania praw mechaniki w czterech wymiarach, z których czwartym był czas. Podobną ideę wyraził w 1797 roku jeden z najważniejszych po odejściu Eulera matematyków w Europie — Joseph-Louis Lagrange w swoich *Mécanique Analytique* (1788) oraz *Théorie des Fonctions Analytiques* (1797).

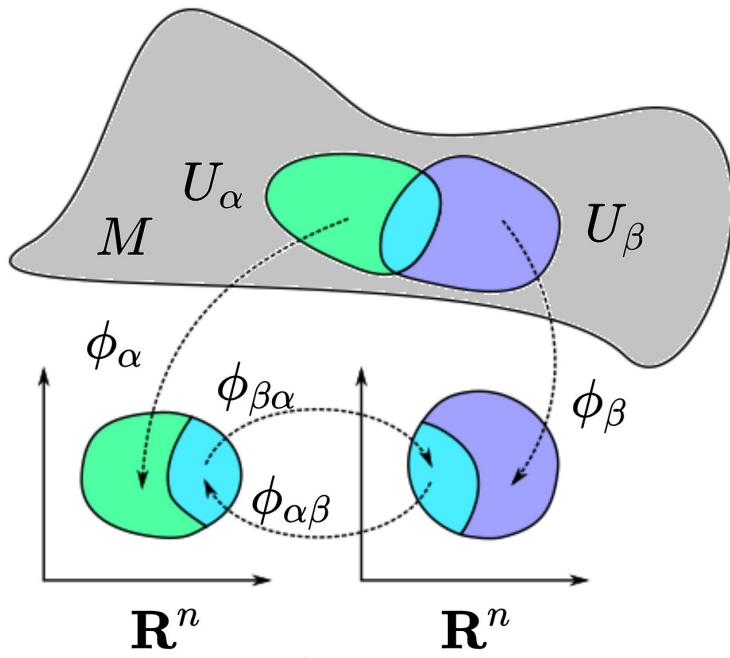
Kolejna ciekawa intuicja czwartego wymiaru pochodziła od Ferdynanda Mobiusa, jednego z twórców geometrii rzutowej, który stwierdził, że symetryczne do siebie figury trójwymiarowe można by na siebie nałożyć, gdyby istniał czwarty wymiar (opowiemy o tym w drugim semestrze — idea jest taka, że tak jak do nałożenia na siebie obiektów symetrycznych na płaszczyźnie typu \mathbb{S}^2 potrzeba wyjścia w trzeci wymiar, tak i podobnie jest np. z prawą i lewą dłonią lub lewym czy prawym butem — jeśli wyjdziemy w czwarty wymiar). Współrzędne jednorodne (barycentryczne), które poznamy w kolejnym semestrze, również wymagały w wersji przestrzennej używania czterech współrzędnych.

Kolejni matematycy używający punktów o więcej niż trzech współrzędnych, wywodzili się nie tylko z geometrii (a jest tu wiele ciekawych wątków, choćby Julius Plücker i współrzędne jednorodne w przestrzeni rzutowej, Ludwig Schläfli i wielościany foremne w czwartym wymiarze, Hamilton i kwaterniony...), ale przede wszystkim z analizy. W 1847 roku Cauchy ogłosił, że „nazywać będziemy zbiór n zmiennych punktem analitycznym, a równanie lub układ równań — miejscem analitycznym”. Zasadnicze jednak i ostateczne przejście do wyżej wymiarowej geometrii dokonało się za sprawą słynnego wykładu habilitacyjnego Riemanna z 1854 roku „O hipotezach leżących u podstaw geometrii”. Wykład ten był jednym z najważniejszych wydarzeń w historii matematyki, wciąż mającym na nią wielki wpływ.

Co było tak istotnego w wykładzie Riemanna? Gdy w 1915 roku Einstein zmienił dzięki ogólnej teorii względności nasze rozumienie wszechświata, sformułował pojęcie czterowymiarowej czasoprzestrzeni, która zgina się i zakrzywia w reakcji na koncentrację masy lub energii. Jest więc obiektem zakrzywionym — jak bardzo? To gigantyczny problem współczesnej matematyki i fizyki (choćby teoria superstrun, przewiduje, że do unifikacji teorii względności i teorii kwantowej należy rozważyć 11-wymiarową czasoprzestrzeń). Pytanie brzmi jednak — skąd możemy wiedzieć, że znajdujemy się na zakrzywionej czasoprzestrzeni, będąc w jej środku? Skąd wiemy, że jest zakrzywiona? To pytanie pięknie spopularyzowano. Abbott we *Flatlandii* (1884) pyta nas jak rozumieją trzeci wymiar „płaszczaki” żyjące na płaszczyźnie?

Ktoś mógłby powiedzieć: bez trudu umiemy odróżnić czy żyjemy na obiekcie zakrzywionym, czy nie, bo przecież mamy równania opisujące różne obiekty. Równanie $x_1^2 + x_2^2 + x_3^2 = 1$ opisuje sferę w przestrzeni trójwymiarowej i nikt nie wątpi, że nie jest to równanie liniowe (dokładniej powiemy o tym w drugim semestrze). Co więcej, сфera, gdy przyjrzymy się jej z bliska, wygląda jak zwykła powierzchnia dwuwymiarowa, a stąd rozsądne jest przypisywanie jej właśnie wymiaru 2. Jest to obiekt dwuwymiarowy w przestrzeni trójwymiarowej. Tak o niej myślimy. Sfera jest przykładem tzw. rozmaitości dwuwymiarowej, tak jak okrąg — jednowymiarowej na płaszczyźnie. Pojęcie rozmaitości pochodzi od Riemanna.

Nie mamy tu narzędzi, by dokładnie opowiedzieć o rozmaitościach, ale idea jest następująca: aby o pewnym obiekcie zawartym w (powiedzmy) \mathbb{R}^N powiedzieć, że jest n -wymiarową rozmaistością, potrzebujemy mieć sposób rysowania map obszarów (otwartych) tej rozmaistości. Sfera jest dwuwymiarowa, bo zaczacząc wokół dowolnego znajdującego się na niej punktu okrąg, możemy uzyskać wycinek sfery przekształcić („zmapować”) jednoznacznie w dysku leżący na płaszczyźnie. Oto intuicyjny obrazek owych map.



Rys 1. Źródło: Wikipedia.

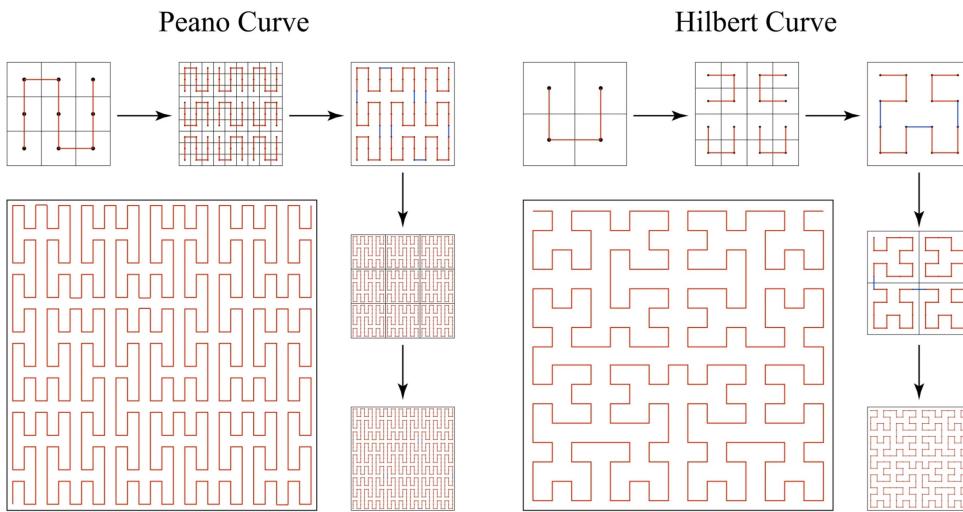
Formalnie są to różniczkowalne i wzajemnie jednoznaczne przekształcenia tych obszarów na odpowiednio otwarte podzbiory przestrzeni \mathbb{R}^n . Muszą one być zgodne, co oznacza, że dla dwóch obszarów U_α oraz U_β i dwóch map ϕ_α , ϕ_β , mapy $\phi_\alpha^{-1} \circ \phi_\beta$ oraz $\phi_\beta^{-1} \circ \phi_\alpha$ muszą również mieć własności map — być różniczkowalne i wzajemnie jednoznaczne. Co to wszystko znaczy, dowiedzą się Państwo na Analizie i Topologii.

Założenie, które przyjmujemy w tym podejściu jest takie, że jesteśmy w stanie widzieć nasz obiekt, np. sferę, jakby od zewnętrz, czyli — zanurzamy go w coś większego i opisujemy np. równaniami czy funkcjami. Jednak, jeśli zaczniemy rozmawiać o kształcie całego wszechświata, lub po prostu obiektu, poza który nie możemy wyjść, to jak mierzyć jego krzywiznę? Na to pytanie próbowały odpowiedzieć Gauss i Riemann.

Riemann był uczniem Gaussa, u którego w 1851 roku przygotował rozprawę o teorii zmiennych zespolonych, mających później stać się podstawą tzw. powierzchni riemannowskich. Gauss opisywał to osiągnięcie jako wzniosłe i niezwykłe płodne. To na prośbę Gaussa Riemann przygotowywał swój wykład inauguracyjny. Celem było właśnie sformułowanie użytecznej definicji miary krzywizny przestrzeni. Teoria ta była rozwijana przez Gaussa dla przestrzeni dwuwymiarowej. Wykazał on, że jedna zmienna potrzebna jest do opisu krzywizny w otoczeniu punktu w przestrzeni dwuwymiarowej (tzw. krzywizna Gaussa). Riemann rozwinał to pojęcie na przestrzenie wyższych wymiarów. Wykazał, że do opisu krzywizny w przestrzeni trójwymiarowej potrzeba sześciu zmiennych (tzw. metryka riemannowska), a w czterowymiarowej przestrzeni — dwudziestu zmiennych. Ogólny obiekt opisany przez Riemanna — tensor krzywizny, jest właśnie podstawą i głównym narzędziem ogólnej teorii Einsteina. O kwestiach tych będziecie Państwo się uczyć na geometrii różniczkowej. Bez algebry liniowej i jej zaawansowanych narzędzi teoria ta nie ma racji bytu.

Z naszej perspektywy ważne jest to, że pierwsza część wykładu Riemanna zawierała zdefiniowaną w sposób jawny i klarowny przestrzeń n -wymiarową. Idee Riemanna zdecydowanie wyprzedzały swoje czasy i zapewne jedynie Gauss był w stanie docenić ich głębio. Niemniej jednak napływające z różnych źródeł matematyki koncepcje wielowymiarowej przestrzeni sprawiły, że już pod koniec XIX wieku pojawiła się mnogość książek, wspomnień i dzieł omawiających i popularyzujących wyższe wymiary. W 1895 roku wielki Henri Poincaré pisał, że „geometria n wymiarów bada rzeczywistość; nikt w to nie wątpi”.

Z matematycznego punktu widzenia dopiero co wyemancypowane pojęcie czecha ogromny kryzys, i to nie w czterech czy więcej wymiarach, ale już w wymiarze 1 czy 2. Wspomniane już wyżej odkrycia Cantora (1878) i Peano (1890) sprawiły, że stosowane dotychczas intuicyjne definicje wymiaru przestały być sensowne.



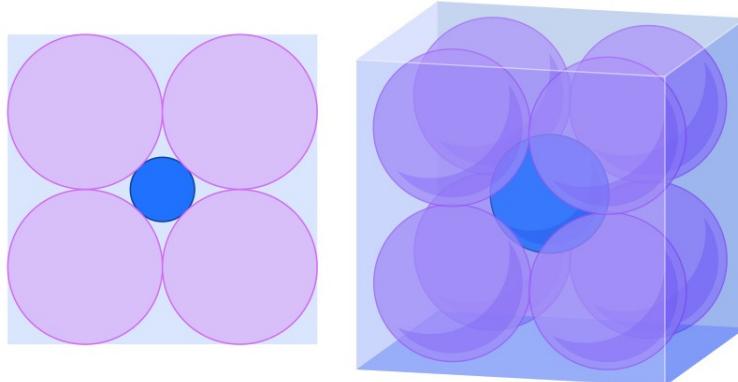
Rys 2. Źródło: <https://galileo-unbound.blog/2023/03/08/a-short-history-of-hyperspace/>.

Głębsze zrozumienie przykładów wyżej otrzymają Państwo na analizie i topologii. Problem jest jasny — jak zdefiniować krzywą wymiaru 1, by definicja nie obejmowała kwadratu? Pod koniec XIX wieku pojawiło się pytanie czy istnieje parametryczna reprezentacja kwadratu na odcinek, która byłaby jednocześnie ciągła i wzajemnie jednoznaczna (czyli homeomorfizm — pojęcie, które poznają Państwo na II roku). Zapytano ogólniej — kiedy iloczyn kartezjański n -kopii odcinka jednostkowego $[0, 1]$, oznaczany przez I^n , jest homeomorficzny z kostką I^m , dla $m \neq n$. Oczekiwano, że nie jest to możliwe i pomiędzy rokiem 1890 i 1910 pojawiło się wiele fałszywych dowodów tego faktu. Poprawny przedstawił dopiero Brouwer w 1911.

Zauważmy, że zbiory typu I^n nie są przestrzeniami liniowymi, więc rezultat ten dał przekonanie, że powinna istnieć funkcja, przypisująca tzw. przestrzeniom topologicznym liczbę zwaną wymiarem. Kroki w tym kierunku poczynił najpierw Poincaré w roku 1912, wywodząc z intuicji „oddzielania” obiektów wyżej wymiarowych obiektami niżej wymiarowymi ideę indukcyjnej definicji wymiaru. Idea opierała się na obserwacji, że obiekty trójwymiarowe oddzielać można dwuwymiarowymi (np. dwa rozłączne wielo-

ściany za pomocą sfery), obiekty dwumiarowe krzywymi itd. Stosowną definicję podał Brouwer w 1913 r.

W przestrzeniach wysokich wymiarów odkryć można wiele nieintuicyjnych zjawisk, które rozważać będą Państwo na wyższych latach studiów (a co dopiero w wymiarze nieskończonym). Oto stosunkowo prosty przykład. Umieścmy 2^n sfer o promieniu 1 wewnątrz n -wymiarowej kostki, której bok ma długość 4. W środku kostki umieścmy kolejną sferę, styczną zewnętrznie do czterech już umieszczonej.

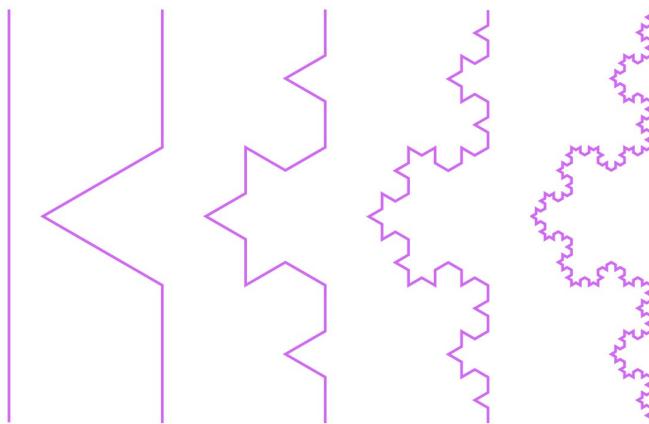


Rys 3. Źródło: <https://www.quantamagazine.org/a-mathematicians-guided-tour-through-high-dimensions-20210913/>.

Gdy wymiar n rośnie, rośnie również promień niebieskiej sfery — wynosi on $\sqrt{n} - 1$. Nietrudno więc stwierdzić, że dla $n \geq 10$ promień tej kuli będzie większy niż długość krawędzi n -wymiarowej kostki, która go zawiera. Innymi słowy sfera ta wystawać będzie poza kostkę! Wydaje się to dziwne.

W międzyczasie wspomniany już wcześniej Felix Haussdorff, sformułował w 1918 roku teorię wymiaru samopodobieństwa, przypisującą niektórym zbiorom wymiar ułamkowy, a nie tylko całkowity. Mandelbrot spopularyzował te zbiory w latach osiemdziesiątych jako fraktale. Pomyśl był prosty, a podamy jedynie jego intuicję — obiekt wymiaru samopodobieństwa d to taki, który poddany jednokładności o skali k , zmienia miarę o czynnik k^d . Co to znaczy? Na razie niewiele wiemy o pojęciu miary, ale oto intuicja. Dla zwykłych obiektów typu odcinek czy kwadrat — wymiar samopodobieństwa działa tak jak zwykły wymiar. Odcinek powiększony trzykrotnie zwiększa długość 3^1 razy, a kwadrat po jednokładności w skali 3 zmienia pole na 3^2 razy większe. Sześcian po jednokładności o skali 3 zmienia objętość na 3^3 razy większą.

Rozważmy jednak np. tak zwaną krzywą Kocha, która konstruujemy poprzez kolejne iteracje. Zaczynamy od odcinka, z którego usuwamy środkową trzecią część (długości $1/3$) i zastępujemy ją dwoma odcinkami o długości równej usuniętemu fragmentowi. Następnie z każdego z czterech powstałych w ten sposób odcinków usuwamy środkową trzecią część i zastępujemy dwiema. Powyższą procedurę kontynuujemy "w nieskończoność" (co da się formalnie opisać — ale to na razie zostawmy, podobne konstrukcje np. zbioru Cantora czy dywanu Sierpińskiego poznacie Państwo na topologii). Jaki ta krzywa ma wymiar?

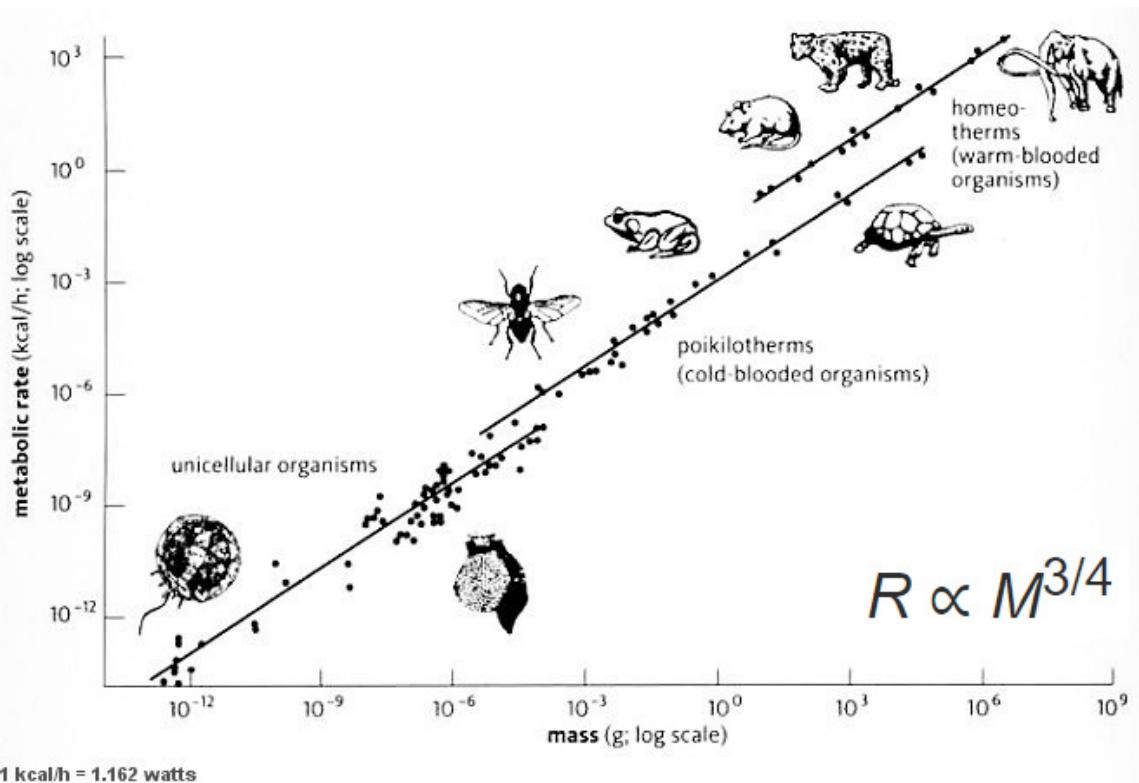


Rys 4. Pierwsze kilka przybliżeń krzywej Kocha. Źródło jak wyżej.

Uzyskana krzywa ma tę własność, że jeśli powiększymy ją trzykrotnie, otrzymamy cztery kopie wyjściowego obiektu. To znaczy, że wymiar Haussdorffa d tej krzywej spełnia równość $3^d = 4$, a więc $d = \log_3 4$.

Czytelnik może uważać podobną konstrukcję jedynie za ciekawostkę. Okazuje się jednak, że ma ona głębokie znaczenie dla współczesnego rozumienia wielu ważnych zależności w przyrodzie, finansach, statystyce i wielu innych dziedzinach.

Świetnym przykładem jest tak zwane prawo Kleibera z roku 1930, które mówi — w największym skrócie, że dla większości wyższych kręgowców metabolizm jest proporcjonalny do masy ciała podniesionej do potęgi 3/4. Co ciekawe, Kleiber uzyskał tę obserwację na podstawie szerokich empirycznych studiów rozmaitych gatunków, a mimo to uzyskany przezeń wykładnik 3/4 nie był intuicyjny. Dlaczego?



Rys 5. Prawo Kleibera.

Intuicyjnie rzecz biorąc, metabolizm, czyli zużycie energii, pochodzi głównie z potrzeby ogrzewania się, a na ogrzewanie główny wpływ ma powierzchnia ciała osobnika. Im większa powierzchnia — tym więcej tracimy ciepła. Im większa masa — tym więcej produkujemy ciepła. Przy takim wyjaśnieniu wydawałoby się, że wykładnik powinien wynosić 2/3, ponieważ (po wzięciu logarytmów) tyle wynosi stosunek między zmianą powierzchni, a zmianą objętości obiektu trójwymiarowego. Jedną z prób wyjaśnienia, dlaczego te wykładniki się różnią jest koncepcja fraktalnej budowy naszych organów krążenia — obejmujących płuc, układ żył i tętnic itd. Nosi ona znamiona struktury samopodobnej, która zdaje się wyjaśniać skąd pojawia się 3/4. Wykładnik ten pojawia się zresztą przy badaniu wielu innych pozornie niepowiązanych ze sobą zjawisk, nie tylko w biologii. Zainteresowanych tym tematem odsyłam do klasycznego artykułu Jamesa Browna, Briana Enquista i Geoffrey Westa z 1999 roku: The Fourth Dimension of Life: Fractal Geometry and Allometric Scaling of Organisms: <https://www.santafe.edu/research/results/working-papers/the-fourth-dimension-of-life-fractal-geometry-and->.

* * *

Pojęcie wymiaru przekroczyło dawno geometrię. Ujęcie prezentowane przez nas w duchu algebrai liniowej dopuszcza o myśleniu o liczbach zespolonych, jako dwuwymiarowej przestrzeni liniowej nad ciałem liczb rzeczywistych, a o liczbach rzeczywistych — jako nieskończoność wymiarowej przestrzeni nad ciałem liczb wymiernych. To ujęcie uogólnione zostało na liczne struktury, takie jak pierścienie, grupy, a także na obiekty kombinatoryczne. Kluczem do sformułowania poprawnej algebraicznej definicji wymiaru jest zastanowienie się jakie będzie on miał własności ze względu na wzajemne położenie obiektów, które mierzy. Podobnie jak w algebrze liniowej, w żadnym ujęciu nie wyobrażamy sobie, by zbiór niskowymiarowy zawierał jako podzbiór zbiór wyżej wymiarowy. Chcemy mieć jakiś sensowny sposób przekształcania na siebie zbiorów o tym samym wymiarze. Chcemy wiedzieć jak wymiar może zachowywać się przy braniu części wspólnej. Niektóre z tych zagadnień, odniesionych do przestrzeni liniowych, rozważamy na kolejnym wykładzie. Inne — te dotyczące przekształceń, poznamy za kilka tygodni mówiąc o izomorfizmach.

Rozdział 11

Twierdzenie Kroneckera-Capellego

11.1 Wykład 11

Poniższe twierdzenie jest jednym z głównych rezultatów tego wykładu.

Twierdzenie 11.1.1: Kroneckera-Capellego

Niech U będzie układem równań liniowych o współczynnikach w ciele K postaci:

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n = b_1 \\ \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \dots + a_{mn}x_n = b_m. \end{cases}$$

o macierzy współczynników A oraz rozszerzonej macierzy współczynników A_u . Wówczas:

- Układ U ma rozwiązanie wtedy i tylko wtedy, gdy $r(A) = r(A_u)$,
- Przestrzeń rozwiązań układu jednorodnego odpowiadającego układowi U ma wymiar $n - r(A)$
- Jeśli α jest rozwiązaniem układu U , a W jest przestrzenią rozwiązań układu jednorodnego odpowiadającego układowi U , to zbiór rozwiązań układu U jest postaci

$$\alpha + W = \{\alpha + \beta \mid \beta \in W\}.$$

Przypomnijmy pojęcia użyte w twierdzeniu. Macierze A oraz A_u układu wyżej to odpowiednio:

$$A = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix}, \quad A_u = \left[\begin{array}{cccc|c} a_{11} & a_{12} & \dots & a_{1n} & b_1 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} & b_m \end{array} \right].$$

Układ jednorodny odpowiadający układowi U to układ jednorodny o macierzy A . Macierz rozszerzoną A_u układu m równań liniowych będziemy też w skrócie zapisywać w postaci $[A|b]$, gdzie $b \in K^m$.

Punkt (c) wykazaliśmy już w uzupełnieniu do pierwszego wykładu (Uwaga 1.4.2), ale uzupełnimy dowód o pojęcie podprzestrzeni. Punkty (a) i (b) opisują problem rozwiązywalności i „rozmiaru” zbioru rozwiązań układu równań liniowych w języku wymiaru. Również te fakty są dla nas w zasadzie intuicyjnie jasne. Wiemy bowiem, że układ równań może okazać się sprzeczny jedynie, gdy w wyniku sprowadzania macierzy A_u do postaci zredukowanej pojawi się wiersz postaci $[00\dots 0|1]$. Nietrudno będzie nam formalnie pokazać, na podstawie posiadanej już wiedzy, że sytuacja ta może wystąpić jedynie, gdy $r(A) < r(A_u)$.

Również punkt (b) jest faktycznie znany. Wykazaliśmy już w Uwadze 9.1.2, że baza przestrzeni rozwiązań jednorodnego układu równań ma tyle elementów ile jest zmiennych niezależnych tego układu. Ta zaś liczba równa jest liczbie wszystkich zmiennych pomniejszonej o liczbę zmiennych zależnych. Wszystkich zmiennych jest n , zaś zmiennych zależnych jest tyle, co schodków macierzy A po sprowadzeniu jej, za pomocą elementarnych operacji wierszowych, do postaci zredukowanej, czyli $r(A)$.

Przykład. Porównajmy układy równań o macierzy współczynników $A \in M_{2 \times 3}(\mathbb{R})$ oraz macierzach rozszerzonych A_u oraz B_u :

$$A_u = \left[\begin{array}{ccc|c} 1 & 1 & 1 & 1 \\ 2 & 2 & 2 & 2 \end{array} \right], \quad B_u = \left[\begin{array}{ccc|c} 1 & 1 & 1 & 1 \\ 2 & 2 & 2 & 4 \end{array} \right], \quad A = \left[\begin{array}{ccc} 1 & 1 & 1 \\ 2 & 2 & 2 \end{array} \right].$$

Macierze współczynników obydwu tych układów mają rzad 1. Macierz A_u ma również rzad 1, natomiast macierz B_u ma rzad 2. Pierwszy układ jest w sposób oczywisty niesprzeczny. Również zgodnie z punktem (a) powyższego twierdzenia układ ten ma rozwiązanie, bo rzad macierzy współczynników tego układu jest równy rzędowi macierzy rozszerzonej. W drugim przypadku rzad macierzy współczynników jest mniejszy — i układ niejednorodny nie ma rozwiązań. Jak znajdujemy rozwiązanie układu o macierzy A_u ?

Rozwiązujeśmy odpowiadający mu układ jednorodny o macierzy (współczynników) A . Na mocy punktu (b) twierdzenia wyżej wiemy, że wymiar W przestrzeni rozwiązań układu równych jednorodnych równy jest 3 (liczba niewiadomych) -1 (rzad macierzy współczynników), czyli 2. Na mocy punktu (a), możemy stwierdzić, że po wyznaczeniu $W = \text{lin}((1, -1, 0), (1, 0, -1))$ wystarczy wziąć dowolne rozwiązanie układu o macierzy A_u , na przykład $\alpha = (1, 0, 0)$ i zbiór rozwiązań całego układu niejednorodnego o macierzy rozszerzonej A_u ma postać:

$$(1, 0, 0) + \text{lin}((1, -1, 0), (1, 0, -1)).$$

* * *

Dowodzimy twierdzenie Kroneckera-Capellego. Punkt (a) wynika ze znanego nam już faktu.

Uwaga 11.1.2

Niech $\alpha_1, \dots, \alpha_n$ będzie układem wektorów w przestrzeni K^m , gdzie

$$\alpha_1 = (a_{11}, \dots, a_{m1}), \dots, \alpha_n = (\alpha_{1n}, \dots, \alpha_{mn}).$$

Wówczas następujące warunki są równoważne:

- wektor $\beta = (b_1, \dots, b_m)$ jest kombinacją liniową wektorów $\alpha_1, \dots, \alpha_n$,
- istnieją $s_1, s_2, \dots, s_n \in K$ takie, że

$$s_1 \cdot \begin{bmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{bmatrix} + s_2 \cdot \begin{bmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{m2} \end{bmatrix} + \dots + s_n \cdot \begin{bmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{mn} \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix}.$$

- układ równań liniowych zmiennych x_1, \dots, x_n nad ciałem K o macierzy rozszerzonej

$$A_u = [A|\beta] = \left[\begin{array}{cccc|c} a_{11} & a_{12} & \dots & a_{1n} & b_1 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} & b_m \end{array} \right]$$

ma rozwiązanie (jest niesprzeczny).

Dowód powyższej uwagi wynika z definicji operacji dodawania wektorów i mnożenia przez skalar w przestrzeni liniowej macierzy. Wniosek: kolumny dowolnej macierzy $A \in M_{m \times n}(K)$ rozpinają podprzestrzeń złożoną ze wszystkich takich wektorów $\beta \in K^m$, że układ o macierzy rozszerzonej $[A|\beta]$ jest niesprzeczny.

Weźmy $\alpha_1, \dots, \alpha_n, \beta \in K^n$, które są kolumnami macierzy A_u . Wówczas mamy ciąg równoważnych stwierdzeń:

$$\begin{aligned} x_1, \dots, x_n \text{ jest rozwiązaniem układu } U &\iff x_1\alpha_1 + \dots + x_n\alpha_n = \beta \\ &\iff \beta \in \text{lin}(\alpha_1, \dots, \alpha_n) \\ &\iff \text{lin}(\alpha_1, \dots, \alpha_n) = \text{lin}(\alpha_1, \dots, \alpha_n, \beta), \\ &\iff \dim \text{lin}(\alpha_1, \dots, \alpha_n) = \dim \text{lin}(\alpha_1, \dots, \alpha_n, \beta) \\ &\iff r(A) = r(A_u). \end{aligned}$$

Punkt (b) wynika, jak już wspomnieliśmy, z Uwagi 9.1.2. Dowodzimy (c), przytaczając jeszcze raz rozumowania z Uzupełnienia 1.4. Przeprowadzimy rozumowanie w dwóch krokach.

Uwaga 11.1.3

Jeśli (s_1, s_2, \dots, s_n) oraz $(s'_1, s'_2, \dots, s'_n)$ są rozwiązaniami układu równań

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + \dots + a_{2n}x_n = b_2 \\ \dots \\ a_{m1}x_1 + a_{m2}x_2 + \dots + a_{mn}x_n = b_m. \end{cases}, \quad (11.1)$$

to ciąg

$$(s_1 - s'_1, s_2 - s'_2, \dots, s_n - s'_n)$$

jest rozwiązaniem układu równań:

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n = \mathbf{0} \\ a_{21}x_1 + a_{22}x_2 + \dots + a_{2n}x_n = \mathbf{0} \\ \dots \\ a_{m1}x_1 + a_{m2}x_2 + \dots + a_{mn}x_n = \mathbf{0}. \end{cases} \quad (11.2)$$

Dowód. Ciągi (s_1, s_2, \dots, s_n) oraz $(s'_1, s'_2, \dots, s'_n)$ są rozwiązaniami każdego z równań

$$a_{i1}x_1 + a_{i2}x_2 + \dots + a_{in}x_n = b_i,$$

gdzie $i = 1, 2, \dots, m$. Pisząc wprost mamy:

$$\begin{aligned} a_{i1}s_1 + a_{i2}s_2 + \dots + a_{in}s_n &= b_i \\ a_{i1}s'_1 + a_{i2}s'_2 + \dots + a_{in}s'_n &= b_i, \end{aligned}$$

czyli po odjęciu stronami widzimy, że $(s_1 - s'_1, s_2 - s'_2, \dots, s_n - s'_n)$ spełnia każde z m równań układu (11.2):

$$a_{i1}(s_1 - s'_1) + a_{i2}(s_2 - s'_2) + \dots + a_{in}(s_n - s'_n) = 0,$$

co oznacza, że jest to rozwiązanie całego układu (11.2). \square

Uwaga 11.1.4

Załóżmy, że $(\mathbf{s}_1, \dots, \mathbf{s}_n)$ jest rozwiązaniem układu równań (11.1). Wówczas każde rozwiązanie układu (11.1) jest postaci:

$$(\mathbf{s}_1, \dots, \mathbf{s}_n) + (\mathbf{u}_1, \dots, \mathbf{u}_n),$$

gdzie $(\mathbf{u}_1, \dots, \mathbf{u}_n)$ jest rozwiązaniem układu (11.2).

Dowód. Weźmy rozwiązanie $(\mathbf{s}_1, \dots, \mathbf{s}_n)$ układu (11.1) i rozwiązanie $(\mathbf{u}_1, \dots, \mathbf{u}_n)$ układu (11.2). Wówczas ciąg $(\mathbf{s}_1 + \mathbf{u}_1, \dots, \mathbf{s}_n + \mathbf{u}_n) = (\mathbf{s}_1, \dots, \mathbf{s}_n) + (\mathbf{u}_1, \dots, \mathbf{u}_n)$ jest rozwiązaniem układu (11.1), bo spełnia dowolne z jego równań:

$$a_{i1}(\mathbf{s}_1 + \mathbf{u}_1) + \dots + a_{in}(\mathbf{s}_n + \mathbf{u}_n) = (a_{i1}\mathbf{s}_1 + \dots + a_{in}\mathbf{s}_n) + (a_{i1}\mathbf{u}_1 + \dots + a_{in}\mathbf{u}_n) = b_i + 0 = b_i.$$

Pozostaje wykazać, że dowolne rozwiązanie (s'_1, \dots, s'_n) układu (11.1) można przedstawić w postaci (\diamond) , dla pewnego rozwiązania $(\mathbf{u}_1, \dots, \mathbf{u}_n)$ układu (11.2). Zgodnie z poprzednią Uwagą wystarczy wziąć:

$$\mathbf{u}_1 = s'_1 - \mathbf{s}_1, \dots, \mathbf{u}_n = s'_n - \mathbf{s}_n$$

i dostajemy $(s'_1, \dots, s'_n) = (s_1 + (s'_1 - s_1), \dots, s_n + (s'_n - s_n)) = (s_1 + \dots + s_n) + (s'_1 - s_1, \dots, s'_n - s_n)$. \square

Dla zilustrowania powyższych faktów rozważmy układ równań liniowych o współczynnikach w \mathbb{R} postaci:

$$\begin{cases} x + 2y + 3z = 1 \\ 4x + 5y + 6z = 1 \quad (\dagger) \\ 7x + 8y + 9z = 1 \end{cases}$$

Zgodnie z procedurą opisaną wyżej do opisania rozwiązań tego układu potrzebujemy dowolne jego rozwiązanie, na przykład $(s_1, s_2, s_3) = (-1, 1, 0)$ oraz wszystkie rozwiązania układu postaci:

$$\begin{cases} x + 2y + 3z = 0 \\ 4x + 5y + 6z = 0 \quad (\dagger) \\ 7x + 8y + 9z = 0 \end{cases}$$

Można sprawdzić, że rozwiązania tego układu są postaci $(z, -2z, z)$, gdzie $z \in \mathbb{R}$. W związku z tym każde rozwiązanie układu (\dagger) ma postać $(-1 + z, 1 - 2z, z)$, gdzie $z \in \mathbb{R}$. Stąd zbiór rozwiązań ma postać:

$$(-1, 1, 0) + \text{lin}((1, -2, 1)).$$

Definicja 11.1.5: Podprzestrzeń opisana układem równań

Jeśli $V \subseteq K^n$ jest przestrzenią rozwiązań jednorodnego układu równań liniowych U , to mówimy, że przestrzeń V jest OPISANA UKŁADEM U .

Możemy teraz sformułować ważny wniosek stanowiący twierdzenie klasyfikacyjne.

Wniosek 11.1.6

Każda podprzestrzeń V przestrzeni K^n jest opisana pewnym jednorodnym układem równań liniowych U . Jeśli $\dim V = k$, to można tak dobrać ten układ U , by składał się z $n - k$ równań. Dla $\dim V = k$ oraz $i < n - k$ nie istnieje złożony z i równań układ równań liniowych opisujący V .

Dowód. Jeśli $(s_{11}, \dots, s_{1n}), \dots, (s_{k1}, \dots, s_{kn})$ są rozwiązaniami układu

$$\begin{cases} a_{11}x_1 + \dots + a_{1n}x_n = 0 \\ \dots \\ a_{m1}x_1 + \dots + a_{mn}x_n = 0 \end{cases},$$

to $(a_{11}, \dots, a_{1n}), \dots, (a_{m1}, \dots, a_{mn})$ są rozwiązaniami układu:

$$\begin{cases} s_{11}x_1 + \dots + s_{1n}x_n = 0 \\ \dots \\ s_{k1}x_1 + \dots + s_{kn}x_n = 0 \end{cases}.$$

Jeśli $(s_{11}, \dots, s_{1n}), \dots, (s_{k1}, \dots, s_{kn})$ jest bazą przestrzeni V rozwiązań układu o macierzy schodkowej

$$\begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix},$$

to na mocy tw. Kroneckera-Capellego $k = n - m$, co więcej wektory $(a_{11}, \dots, a_{1n}), \dots, (a_{m1}, \dots, a_{mn})$ są liniowo niezależne i są rozwiązaniami układu o macierzy rzędu k postaci

$$\begin{bmatrix} s_{11} & s_{12} & \dots & s_{1n} \\ \vdots & \vdots & \ddots & \vdots \\ s_{k1} & s_{k2} & \dots & s_{kn} \end{bmatrix}.$$

Ponownie na mocy tw. Kroneckera-Capellego, powyższy układ ma przestrzeń rozwiązań wymiaru

$$n - k = n - (n - m) = m.$$

Czyli jest to $\text{lin}((a_{11}, \dots, a_{1n}), \dots, (a_{m1}, \dots, a_{mn}))$ – przestrzeń wymiaru $n - k$.

Wykazaliśmy zatem, że jeśli $\alpha_1, \dots, \alpha_k$ jest bazą przestrzeni $V \subseteq K^n$ oraz $A \in M_{k \times n}(K)$ jest macierzą o wierszach $\alpha_1, \dots, \alpha_k$, to przestrzeń V można opisać układem dowolnych $n - k$ równań, których współczynniki tworzą bazę przestrzeni rozwiązań układu danego macierzą A . Równań tych nie może być oczywiście mniej, bowiem przestrzeń rozwiązań układu o mniej niż $n - k$ równaniach ma wymiar większy niż $n - (n - k)$, na mocy twierdzenia Kroneckera-Capellego. \square

Przykład. Rozważmy $V = \text{lin}((1, 2, 0, 1, 0), (0, 0, 1, 1, 1)) \subseteq \mathbb{R}^5$. Rozwiązania układu równań o macierzy

$$\begin{bmatrix} 1 & 2 & 0 & 1 & 0 \\ 0 & 0 & 1 & 1 & 1 \end{bmatrix}$$

stanowi przestrzeń współczynników wszystkich równań liniowych, których rozwiązania zawierają V . Wybierając różne bazy tej przestrzeni dostajemy różne (ale równoważne) układy równań opisujące V , na przykład dla bazy

$$(-2, 1, 0, 0, 0), (-1, 0, -1, 1, 0), (0, 0, -1, 0, 1)$$

mamy następujący układ opisujący V :

$$\begin{cases} -2x_1 + x_2 = 0 \\ -x_1 - x_3 + x_4 = 0 \\ -x_3 + x_5 = 0 \end{cases}.$$

Zadanie. Założmy, że U jest układem pięciu równań z trzema niewiadomymi o współczynnikach z \mathbb{Q} . Założmy ponadto, że rzęd macierzy współczynników układu U jest równy 2 oraz trójkę $(1, 1, 1)$ i $(1, 2, 1)$ są rozwiązaniami tego układu równań. Rozstrzygnij, czy trójka $(1, 1, 2)$ jest rozwiązaniem układu równań U . Odpowiedź uzasadnij.

Wprawdzie umiemy rozwiązać to zadanie przy użyciu intuicji z początku zajęć, warto korzystać z twierdzenia Kroneckera-Capellego. Macierz układu U ma wymiar 5×3 . Skoro rzęd macierzy układu U równy jest 2, to przestrzeń rozwiązań układu jednorodnego odpowiadającego temu układowi ma wymiar $3 - 2 = 1$. W szczególności rozwiązań układu U ma postać

$$\alpha + \text{lin}(\beta),$$

gdzie α jest dowolnym rozwiązaniem układu U , a β jest niezerowym rozwiązaniem układu jednorodnego.

Zauważmy, że wektor $(1, 2, 1) - (1, 1, 1) = (0, 1, 0)$ jest, na mocy Uwagi 11.1.3, rozwiązaniem układu jednorodnego odpowiadającego U . Zatem zbiór rozwiązań układu U ma postać

$$(1, 1, 1) + \text{lin}((0, 1, 0)).$$

Jest więc jasne, że trójka $(1, 1, 2)$ nie jest rozwiązaniem układu U .

Wniosek 11.1.7

Układ U ma jednoznaczne rozwiązanie wtedy i tylko wtedy, gdy $r(A_u) = r(A) = n$.

Dowód. Układ U ma jednoznaczne rozwiązanie wtedy i tylko wtedy, gdy $r(A_u) = r(A)$ oraz $\dim W = 0$, co jest równoważne $r(A_u) = r(A) = n$. \square

Dokonaliśmy zatem klasyfikacji podprzestrzeni w K^n i umiemy je wyrażać zarówno jako przestrzenie rozpięte przez dany układ wektorów, jak i jako przestrzeń opisane przez określony układ równań.

* * *

Na koniec przyjrzymy się nieco innej metodzie wyznaczania układu równań opisujących podprzestrzeń $W = \text{lin}(\alpha_1, \dots, \alpha_k)$ w przestrzeni liniowej K^n .

Załóżmy, że $\alpha_1, \dots, \alpha_k$ jest bazą W . Dla każdego $1 \leq i \leq k$ niech $\alpha_i = (a_{i1}, \dots, a_{in})$, gdzie $a_{ij} \in K$. Element $\beta = (x_1, \dots, x_n)$ należy do W wtedy i tylko wtedy, gdy istnieją takie a_1, \dots, a_k , że

$$\beta = a_1\alpha_1 + \dots + a_k\alpha_k,$$

czyli równoważnie:

$$\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_k \end{bmatrix} = a_1 \cdot \begin{bmatrix} a_{11} \\ a_{12} \\ \vdots \\ a_{1n} \end{bmatrix} + a_2 \cdot \begin{bmatrix} a_{21} \\ a_{22} \\ \vdots \\ a_{2n} \end{bmatrix} + \dots + a_k \cdot \begin{bmatrix} a_{k1} \\ a_{k2} \\ \vdots \\ a_{kn} \end{bmatrix}.$$

Innymi słowy wektor β należy do W wtedy i tylko wtedy, gdy poniższy układ równań liniowych o macierzy rozszerzonej ma rozwiązanie

$$\left[\begin{array}{cccc|c} a_{11} & a_{21} & \dots & a_{k1} & x_1 \\ a_{12} & a_{22} & \dots & a_{k2} & x_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{1n} & a_{2n} & \dots & a_{kn} & x_n \end{array} \right].$$

W kolumny powyższej macierzy wpisane są współrzędne $\alpha_1, \dots, \alpha_k, \beta$. Skoro pierwsze k kolumn powyższej macierzy jest bazą, to powyższy układ ma rozwiązanie wtedy i tylko wtedy, gdy rzad powyższej macierzy jest równy k . Zauważmy, że gdy sprowadzimy tę macierz do postaci schodkowej, uzyskamy macierz postaci:

$$\left[\begin{array}{cccc|c} a'_{11} & a'_{21} & \dots & a'_{k1} & r_1 \\ a'_{12} & a'_{22} & \dots & a'_{k2} & r_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a'_{kn} & a'_{kn} & \dots & a'_{kk} & r_k \\ 0 & 0 & \dots & 0 & r_{k+1} \\ 0 & 0 & \dots & 0 & r_{k+2} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \dots & 0 & r_n \end{array} \right],$$

gdzie r_i są wyrażeniami postaci $c_{1i}x_1 + c_{2i}x_2 + \dots + c_{ni}x_n$. A zatem rzad powyższej macierzy jest równy k wtedy i tylko wtedy, gdy

$$r_{k+1} = \dots = r_n = 0.$$

W ten sposób uzyskujemy układ $n - k$ równań opisujący podprzestrzeń W .

Przykład. Rozważmy ponownie podprzestrzeń $V = \text{lin}((1, 2, 0, 1, 0), (0, 0, 1, 1, 1)) \subseteq \mathbb{R}^5$. Wektor (x_1, \dots, x_5) należy do V wtedy i tylko wtedy, gdy rozwiązanie ma układ równań o macierzy

$$\left[\begin{array}{cc|c} 1 & 0 & x_1 \\ 2 & 0 & x_2 \\ 0 & 1 & x_3 \\ 1 & 1 & x_4 \\ 0 & 1 & x_5 \end{array} \right]$$

Sprowadzając powyższą macierz do postaci schodkowej, mamy:

$$\left[\begin{array}{cc|c} 1 & 0 & x_1 \\ 2 & 0 & x_2 \\ 0 & 1 & x_3 \\ 1 & 1 & x_4 \\ 0 & 1 & x_5 \end{array} \right] \longrightarrow \left[\begin{array}{cc|c} 1 & 0 & x_1 \\ 0 & 0 & x_2 - 2x_1 \\ 0 & 1 & x_3 \\ 1 & 1 & x_4 \\ 0 & 0 & x_5 - x_3 \end{array} \right] \longrightarrow \left[\begin{array}{cc|c} 1 & 0 & x_1 \\ 0 & 0 & -2x_1 + x_2 \\ 0 & 1 & x_3 \\ 0 & 0 & -x_1 - x_3 + x_4 \\ 0 & 0 & -x_3 + x_5 \end{array} \right] \longrightarrow \left[\begin{array}{cc|c} 1 & 0 & x_1 \\ 0 & 1 & x_3 \\ 0 & 0 & -2x_1 + x_2 \\ 0 & 0 & -x_1 - x_3 + x_4 \\ 0 & 0 & -x_3 + x_5 \end{array} \right]$$

W rezultacie wektor $(x_1, x_2, x_3, x_4, x_5)$ należy do V wtedy i tylko wtedy, gdy jego współrzędne spełniają (uzyskany też inną metodą wyżej) układ równań

$$\begin{cases} -2x_1 + x_2 = 0 \\ -x_1 - x_3 + x_4 = 0 \\ -x_3 + x_5 = 0 \end{cases}.$$

11.2 Wybór przykładowych pytań

Proszę się nie uczyć na pamięć odpowiedzi czy wyjaśnienie! Takich pytań można ułożyć znacznie więcej i służą one jedynie sprawdzeniu podstawowego zrozumienia treści zajęć. Jeśli odpowiedź na dowolne z poniższych pytań sprawia Państwu problem, proszę pomyśleć który fragment notatek dotyczy tego pytania, wrócić do ich lektury, a w razie dalszych trudności poprosić prowadzącego o wskazówkę.

1. Macierz jednorodnego układu równań liniowych o 3 zmiennych jest rzędu 1. Jaki jest wymiar przestrzeni rozwiązań tego układu?
2. Wymiar przestrzeni rozwiązań jednorodnego układu trzech równań liniowych o n zmiennych jest równy 20. Jakie są możliwe wartości liczby n ?
3. Rząd macierzy rozszerzonej $[A|b]$ układu równań liniowych jest równy 3. Ile może się równać $r(A)$?
4. Podaj przykład takiego wektora v oraz takiej podprzestrzeni W w przestrzeni liniowej \mathbb{R}^3 , że zbiór rozwiązań równania liniowego $x_1 - x_2 + x_3 = 1$ jest postaci $v + W$.
5. Podaj przykład takiego wektora v oraz takiej podprzestrzeni W w przestrzeni liniowej \mathbb{R}^3 , że zbiór rozwiązań układu równań liniowych $x_1 = 1, x_1 - x_2 + x_3 = 1$ jest postaci $v + W$.
6. Podaj przykład takiego wektora v oraz takiej podprzestrzeni W w przestrzeni liniowej \mathbb{R}^3 , że zbiór rozwiązań układu równań liniowych $x_1 = 1, x_2 = 2, x_1 - x_2 + x_3 = 1$ jest postaci $v + W$.
7. Wiadomo, że podprzestrzeń $\text{lin}((1, 1, 0, 1), (1, 2, 0, 1), (1, 1, 0, 2)) \subset \mathbb{R}^4$ można opisać jednym równaniem. Jakie to równanie?
8. Iloma równaniami liniowymi opisana jest podprzestrzeń $\text{lin}((2180, 3180, 4420, 5440))$?
9. Iloma równaniami liniowymi opisana jest podprzestrzeń $\text{lin}((1, 1, 1, 1), (2, 2, 2, 2), (3, 3, 3, 3)) \subset \mathbb{R}^4$?
10. Niech $\alpha_1, \dots, \alpha_n$ będzie bazą przestrzeni K^n i niech $k \leq n$. Iloma równaniami opisana jest podprzestrzeń $\text{lin}(\{\alpha_i \mid i \geq k\})$?
11. Założmy, że układ równań liniowych o współczynnikach rzeczywistych ma dwa różne rozwiązania. Czy zbiór rozwiązań tego układu może być skończony?
12. Czy układ równań liniowych o współczynnikach w \mathbb{Z}_5 może mieć nieskończenie wiele rozwiązań?
13. Układ równań liniowych o współczynnikach w ciele K postaci

$$\begin{cases} a_{11}x_1 + a_{12}x_2 = 0 \\ a_{21}x_1 + a_{22}x_2 = 0 \end{cases}$$

ma niezeroowe rozwiązanie. Jaki rząd ma macierz współczynników tego układu?

14. O liczbach rzeczywistych $a_1, a_2, a_3, a_4, b_1, b_2$ wiadomo, że $(b_1, b_2) \notin \text{lin}((a_1, a_2), (a_3, a_4))$. Czy układ

 - $$\begin{cases} a_1x_1 + a_3x_2 = b_1 \\ a_2x_1 + a_4x_2 = b_2 \end{cases}$$
 może mieć rozwiązanie?

15. Macierz rozszerzona $[A|b]$ układu równań liniowych U jest w postaci schodkowej i ostatni niezeroowy wiersz ma dwa różne od zera wyrazy. Czy układ U może być sprzeczny?
16. Macierz rozszerzona $[A|b]$ układu równań liniowych U ma tyle samo wierszy, co kolumn. Czy układ U może mieć dokładnie jedno rozwiązanie?
17. Do jednorodnego układu równań liniowych dopisano jedno równanie. Czy wynika stąd, że wymiar przestrzeni rozwiązań nowo otrzymanego układu wzrósł o 1?
18. Rozważmy podprzestrzeń W przestrzeni macierzy rozmiaru 2×3 postaci

$$\begin{bmatrix} a & b & c \\ d & e & f \end{bmatrix}$$

złożonej z takich macierzy, w których druga i trzecia kolumna są identyczne. Jaki układ równań jednorodnych spełniają wyrazy a, b, c, d, e, f dowolnej macierzy należącej do W ? Jaki jest rząd macierzy tego układu? Oblicz na tej podstawie $\dim W$.

11.3 Zadania do samodzielnej pracy

Zadania ilustrujące umiejętności niezbędne do zaliczenia przedmiotu oznaczone są symbolem ♠.
Przy tych zadaniach dodany jest opis umiejętności, które sprawdzają.

1. (♠) Znajdowanie wymiaru przestrzeni rozwiązań układu jednorodnego)

Dla każdego $t \in \mathbb{R}$ niech V_t będzie przestrzenią rozwiązań układu równań o współczynnikach w \mathbb{R}

$$\begin{cases} x_1 - x_2 + 2x_4 = 0 \\ x_1 + 2x_2 + 3x_3 + (2-t)x_4 = 0 \\ tx_1 + tx_2 + tx_3 + tx_4 = 0 \end{cases}.$$

Dla każdego $t \in \mathbb{R}$ znajdź wymiar przestrzeni V_t .

2. (♠) Niech V będzie podprzestrzenią przestrzeni \mathbb{R}^5 opisaną układem równań

$$\begin{cases} x_1 + 2x_2 - x_3 + x_4 + x_5 = 0 \\ 2x_1 + 4x_2 - 2x_3 + 2x_4 + ax_5 = 0 \end{cases}.$$

Wyznacz wymiar V w zależności od parametru a .

3. (♠) Rozstrzyganie rozwiązywalności układu równań

Rozpatrzmy układ równań liniowych

$$\begin{cases} 3x_1 + x_2 + 3x_3 = 2 \\ 4x_1 + 4x_2 + 7x_3 = t \\ 5x_1 + sx_2 + 11x_3 = 0 \\ 2x_1 - 2x_2 - x_3 = 3 \end{cases}$$

Dla jakich $s, t \in \mathbb{R}$ układ ten

- ma dokładnie jedno rozwiązanie?
- ma nieskończenie wiele rozwiązań?
- nie ma rozwiązań?

4. (♠) Opisywanie podprzestrzeni rozpiętej na układzie wektorów układem równań)

Dla każdej z poniższych podprzestrzeni $W \subset \mathbb{R}^n$ znaleźć opisujący ją układ równań liniowych.

- $W = \text{lin}((3, 1, 2, -1), (4, 2, 1, 5), (5, 5, 4, 3)) \subset \mathbb{R}^4$,
- $W = \text{lin}((4, 1, 2, -3), (2, 3, 1, -9), (2, -1, 1, 3), (6, 4, 3, -12)) \subset \mathbb{R}^4$,
- $W = \text{lin}((5, 1, 9, 0, 2), (5, 2, -2, 5, -1), (4, 1, 5, 1, 1)) \subset \mathbb{R}^5$.

5. (♠) Rozpatrzmy niejednorodny układ równań liniowych U o współczynnikach w ciele \mathbb{Q} :

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

- Przedstaw zbiór rozwiązań układu U w postaci $v + W$, dla $v \in \mathbb{Q}^4$ oraz podprzestrzeni $W \subseteq \mathbb{Q}^4$.
- Znajdź bazę podprzestrzeni $\text{lin}(v) + W \subseteq \mathbb{Q}^4$.
- Znajdź jednorodny układ równań liniowych opisujący podprzestrzeń

$$\text{lin}((0, 0, 1, -1), (1, -1, 0, 0), (1, t, 1, 1)) \subseteq \mathbb{Q}^4,$$

gdzie $t \in \mathbb{Q}$ jest parametrem.

6. (♠) Niech $W = \text{lin}((7, 9, 6, 8), (11, u, 12, u+1), (2, 1, 3, 2), (3, -4, 9, 2)) \subset \mathbb{R}^4$. Dla jakich $u \in \mathbb{R}$ można podprzestrzeń W opisać jednym równaniem liniowym?

11.4 Uzupełnienie. Kilka uwag o prostopadłości

Zbiór rozwiązań liniowego równania jednorodnego $a_1x_1 + a_2x_2 = 0$ w \mathbb{R}^2 to prosta przechodząca przez punkt $(0, 0)$, o ile tylko $(a_1, a_2) \neq (0, 0)$. To jest, proszę zauważać, twierdzenie Kroneckera-Capellego. Tylko gdy $(a_1, a_2) \neq (0, 0)$, macierz tego układu ma rzad 1, a zatem i zbiór rozwiązań ma wymiar 1, co interpretujemy geometrycznie używając pojęcia „prostej”. Z punktu widzenia geometrii elementarnej ważne może być stwierdzenie, że zbiór wektorów (x_1, x_2) spełniających powyższe równanie odpowiada zbiorowi wektorów (x_1, x_2) , które są **prostopadłe**¹ do wektora (a_1, a_2) . Kiedyś w szkole uczyono (ale teraz już nie), że prostopadłość wektorów równoważna jest temu, że ich iloczyn skalarny równy jest 0. Natomiast wyrażenie $a_1x_1 + a_2x_2$ opisuje właśnie ów elementarny iloczyn skalarny wektorów (a_1, a_2) oraz (x_1, x_2) . Podobnie zdefiniowany elementarny iloczyn skalarny w przestrzeni \mathbb{R}^3 pozwala stwierdzić, że jeśli tylko $(a_1, a_2, a_3) \neq (0, 0, 0)$, to zbiór rozwiązania równania $a_1x_1 + a_2x_2 + a_3x_3 = 0$ jest płaszczyzną przechodzącą przez punkt $(0, 0, 0)$ i prostopadłą do wektora (a_1, a_2, a_3) . Używam określenia „elementarny”, bowiem w drugim semestrze pojęcie iloczynu skalarnego zdefiniujemy w sposób aksjomatyczny, dla dowolnej przestrzeni liniowej nad \mathbb{R} , a po pewnych umowach i osłabieniu nazwy „iloczyn skalarny” na „funkcjonal dwuliniowy”, badać będziemy choćby prostopadłość wektorów także w przestrzeniach nad innymi ciałami. A zatem okaże się, że prostopadłe mogą być grafy, funkcje zespolone, macierze, wielomiany (to akurat będzie dość ważne już na pierwszym roku Analizy) itd.

Możemy na razie nie martwić się abstrakcją, a zastanowić czym miałyby być prostopadłość układu wektorów w znanej przestrzeni K^n . Powiemy mianowicie, że dwa wektory (a_1, \dots, a_n) oraz (b_1, \dots, b_n) są prostopadłe, ozn. $(a_1, \dots, a_n) \perp (b_1, \dots, b_n)$, jeśli $a_1b_1 + a_2b_2 + \dots + a_nb_n = 0$. Zbiór wektorów w K^n prostopadłych do ustalonego zbioru wektorów X oznaczać będziemy przez X^\perp . Innymi słowy:

$$X^\perp = \{v \in K^n : v \perp x, \forall x \in X\}.$$

Proszę zauważać, że dla każdego podzbioru $X \subseteq K^n$ zbiór X^\perp jest podprzestrzenią w K^n . Istotnie, jeśli (v_1, v_2, \dots, v_n) oraz (w_1, \dots, w_n) są prostopadłe do dowolnego wektora $(x_1, \dots, x_n) \in X$, to są do niego prostopadłe również wektory $(v_1 + w_1, \dots, v_n + w_n)$ oraz (av_1, \dots, av_n) . po prostu dlatego, że

$$(v_1 + w_1)x_1 + \dots + (v_n + w_n)x_n = v_1x_1 + \dots + v_nx_n + w_1x_1 + \dots + w_nx_n = 0.$$

Zauważmy dalej, że jeśli $X = \text{lin}(\alpha_1, \dots, \alpha_n)$, to podprzestrzeń X^\perp opisuje zbiory współczynników wszystkich równań liniowych, których rozwiązaniami są wektory z X . Z drugiej strony: jeśli mamy jednorodny układ równań liniowych o macierzy, której wierszami są wektory $\alpha_1, \dots, \alpha_n$, to zbiór rozwiązań tego układu równy jest... $\text{lin}(\alpha_1, \dots, \alpha_n)^\perp$. Czy Czytelnik widzi dualność, którą tu otrzymujemy? Czy Czytelnik widzi co robi tu Twierdzenie Kroneckera-Capellego? Mówi ono po prostu, że dla podprzestrzeni V przestrzeni liniowej K^n mamy (to się w ogólności zepsuje dla pewnych K i pewnych „niestandardowych prostopadłości”, ale dla tej „elementarnej” – tzw. standardowej wersji to zawsze jest prawda):

$$(V^\perp)^\perp = V.$$

Zachęcam Czytelnika, by zastanowił się nad innymi konsekwencjami naszkicowanych tu definicji. Oczywiście (choć będą „wyjątki”) $\dim V^\perp = n - \dim V$. Oczywiście, jeśli $V \subseteq W$ są podprzestrzeniami K^n , to $W^\perp \subseteq V^\perp$. Jakże istotne jest to odwrócenie kolejności pomiędzy podzbiorami i przestrzeniami prostopadłymi! Jest to bodaj najprostszy przykład odpowiedniości Galois, o której napiszę dalej. To jeszcze nie koniec. Nie wnikając w geometrię warto zauważać, że prostopadłość jest swego rodzaju „lepszą liniową niezależnością”. W przypadku skończonego układu wektorów liniowa niezależność nie wynika z tego, że dowolne dwa elementy układu są liniowo niezależne. Przyjmuje się natomiast następującą definicję.

Definicja 11.4.1

Układ wektorów $X \subseteq K^n$ nazwiemy **PROSTOPADŁYM** (albo **ORTOGONALNYM**), jeśli $\alpha \perp \beta$, dla każdych $\alpha, \beta \in X$. Układ prostopadły będący bazą przestrzeni V nazywamy **BAZĄ PROSTOPADŁĄ** (albo **ORTOGONALNĄ**) przestrzeni V (względem naszego „standardowego” iloczynu skalarnego).

Przykładem bazy prostopadłej jest oczywiście baza standardowa, oczywiście niejedynym. Czytelnikowi zostawiam następujące proste ćwiczenie: dowolny układ prostopadły złożony z niezerowych wektorów jest liniowo niezależny! Konsekwencje tego faktu są bardzo ciekawe, ale na razie nie będziemy eksplorować wątków geometrycznych. Zachęcam Czytelnika do poszukiwania prostopadłości w naszych rozważaniach.

¹Osoby zainteresowane elementarnymi dowodami tych własności odnoszącymi się do twierdzeń szkolnych zachęcam do zatrudnienia do wykładu dra Michała Krycha: *Elementy geometrii analitycznej*, dostępnego na stronie: <https://www.mimuw.edu.pl/~krych/chemia/2016-2017>.

11.5 Dodatek. Odpowiedniość Galois i Nullstellensatz Hilberta

Aby jeszcze lepiej i głębiej zrozumieć dlaczego twierdzenie Kroneckera-Capellego jest istotne, zdefiniujmy dwie operacje \mathcal{R} oraz \mathcal{W} na podzbiorach w K^n .

- Dla podzbioru $S \subseteq K^n$ przez $\mathcal{W}(S) \subseteq K^n$ rozumiemy zbiór złożony z n współczynników każdego takiego jednorodnego równania n zmiennych, którego **rozwiązańem jest każdy element** z S . Innymi słowy, wektor $(\textcolor{red}{a}_1, \dots, \textcolor{red}{a}_n) \in K^n$ należy do $\mathcal{W}(S)$ jeśli dla każdego $(s_1, \dots, s_n) \in S$: zachodzi

$$a_1s_1 + a_2s_2 + \dots + a_ns_n = 0.$$

- Dla podzbioru $T \subseteq K^n$ przez $\mathcal{R}(T) \subseteq K^n$ rozumiemy **zbiór rozwiązań wszystkich jednorodnych równań** liniowych n zmiennych, których n -tki współczynników należą do T . Innymi słowy, wektor $(\textcolor{blue}{s}_1, \dots, \textcolor{blue}{s}_n) \in K^n$ należy do $\mathcal{R}(T)$ jeśli dla każdego $(a_1, \dots, a_n) \in T$ zachodzi równość:

$$a_1s_1 + a_2s_2 + \dots + a_ns_n = 0.$$

Na przykład $(1, 1, -1) \in \mathcal{W}((2, 1, 3))$, ponieważ $1 \cdot 2 + 1 \cdot 1 + (-1) \cdot 3 = 0$, czyli $(2, 1, 3)$ jest rozwiązaniem równania $1 \cdot x_1 + 1 \cdot x_2 + (-1) \cdot x_3 = 0$. Są oczywiście inne elementy $\mathcal{W}(2, 1, 3)$, na przykład $(-2, -2, 2)$. Weźmy jednak odwrotną sytuację: biorę wektor $(1, 1, -1)$ i interesuje mnie jakiś element $\mathcal{R}((1, 1, -1))$. Oczywiście – jednym z nich jest $(2, 1, 3)$, ale nie jedynym. Co to wszystko znaczy? Po co te komplikacje?

Nietrudno widzieć, że mamy dwie zależności (jest ich więcej):

$$\mathcal{W}(\mathcal{R}(S)) \supseteq S, \quad \mathcal{R}(\mathcal{W}(T)) \supseteq T.$$

Pierwsza z nich mówi, że każdy wektor jest elementem (czasami niejedynym, stąd inkluza) zbioru rozwiązań równania, którego jest rozwiązaniem, a druga mówi, że jeśli równanie ma określone rozwiązanie, to rozwiązanie to jest jego rozwiązaniem (niekoniecznie jedynym, więc znowu jest inkluza). Brzmi to niemal banalnie, ale interesujące jest to, że zależności te nie dotyczą jedynie równań liniowych! Zauważmy, że jeśli $S_1 \subseteq S_2$, to $\mathcal{W}(S_1) \supseteq \mathcal{W}(S_2)$, podobnie dla operacji \mathcal{R} . Wszystko, co powiedzieliśmy na dzisiejszym wykładzie można w zasadzie streszczyć prostym i eleganckim stwierdzeniem, że jeśli S jest podprzestrzenią liniową przestrzeni K^n – niezależnie czy rozumianą jako przestrzeń współczynników czy przestrzeń rozwiązań, to mamy $\mathcal{R}(S) = S^\perp$ oraz $\mathcal{W}(S) = S^\perp$, czyli:

$$\mathcal{W}(\mathcal{R}(S)) = S, \quad \mathcal{R}(\mathcal{W}(S)) = S. \quad (*)$$

Możemy też wrócić do wyjściowego przykładu i zapisać wyrażone w nim postulaty w nowym języku. Chcemy znaleźć układ $n-1$ równań, którego zbiorem rozwiązań jest **dokładnie** $\text{lin}(\alpha) \neq 0$. Rzeczywiście:

- $\mathcal{W}(\text{lin}(\alpha))$ jest przestrzenią $n-1$ wymiarową,
- dla $n-1$ liniowo niezależnych elementów r_1, \dots, r_{n-1} z $\mathcal{W}(\text{lin}(\alpha))$ mamy $\mathcal{R}(r_1, \dots, r_{n-1}) = \text{lin}(\alpha)$.

Innymi słowy szukane przez nas $n-1$ równań będzie miało współczynniki będące bazą $\mathcal{W}(\text{lin}(\alpha))$.

Operacje tego typu, co \mathcal{W} i \mathcal{R} „rosziane są” po całej matematyce. Rozważmy jeden ważny przykład:

- podzbiorom X ciała K przyporządkowujemy zbiór $\mathcal{W}(X)$ wszystkich wielomianów o współczynnikach z $K[x]$, których pierwiastkami są wszystkie elementy zbioru X ,
- każdemu zbiorowi wielomianów W można przypisać zbiór $\mathcal{R}(W)$ jego wspólnych pierwiastków w K .

Przykład. Niech $K = \mathbb{C}$. Wówczas:

- $\mathcal{W}(\{-i, i\})$ to zbiór wszystkich wielomianów podzielnych przez $(x^2 + 1)$
- $\mathcal{W}(\{0, -i, i\})$ — zbiór wszystkich wielomianów podzielnych przez $x(x^2 + 1)$.

Zauważmy też, że zbiór $\{-i, i\}$ jest zbiorem wspólnych rozwiązań istotnie różnych zbiorów wielomianów, na przykład zbioru wielomianów podzielnych przez $(x^2 + 1)^5$.

Dokładny opis zbioru wielomianów $\mathcal{W}(\mathcal{R}(W))$ nie jest banalny! Jest on treścią słynnego Nullstellensatz – twierdzenia Hilberta o zerach z 1893 roku, które jest uogólnieniem Zasadniczego Twierdzenia Algebry i punktem wyjścia geometrii algebraicznej. Powiedzmy kilka słów o tym twierdzeniu, bez wchodzenia w techniczne detale. Ograniczymy się jedynie do pokazania w jaki sposób twierdzenie to uogólnia Twierdzenie Kroneckera-Capellego. Chodzi mianowicie o rozwiązywanie układów równań, ale wielomianowych. Wychodzimy od następującej sytuacji. Mamy wielomiany f_1, f_2, \dots, f_m i chcemy coś powiedzieć o zbiorze rozwiązań układu $f_1 = 0, f_2 = 0, \dots, f_m = 0$. I nie chodzi nam tylko o wielomiany w $K[x]$ tak, by zbiór wspólnych rozwiązań leżał w K . Chodzi nam o tzw. wielomiany n zmiennych i (znowu) o podzbiory K^n .

Definicja 11.5.1

WIELOMIANEM ZMIENNYCH x_1, \dots, x_n O WSPÓŁCZYNNIKACH Z CIAŁA K nazywamy wyrażenie postaci:

$$\sum_{i_1, \dots, i_n} a_{i_1 \dots i_n} x_1^{i_1} x_2^{i_2} \dots x_n^{i_n},$$

gdzie i_1, \dots, i_n są liczbami całkowitymi nieujemnymi (suma ta brana jest po wszystkich możliwych układach liczb całkowitych nieujemnych), elementy $a_{i_1 i_2 \dots i_n} \in K$, przy czym zakładamy, że suma ta jest skończona, czyli współczynniki $a_{i_1 i_2 \dots i_n}$ są różne od 0 tylko dla skończonej liczby indeksów i_1, \dots, i_n . Zbiór wszystkich wielomianów zmiennych x_1, \dots, x_n o współczynnikach w ciele K oznaczamy $K[x_1, \dots, x_n]$.

Rozważmy kilka przykładów dla przyswojenia sobie wprowadzonej notacji.

- W wielomianie $w \in \mathbb{R}[x_1, x_2]$ postaci $w = x_1^2 + 4x_1x_2 + 3x_2 + 5$ mamy $a_{20} = 1$, $a_{11} = 4$, $a_{01} = 3$, $a_{00} = 5$ oraz $a_{i_1 i_2} = 0$ dla pozostałych par i_1, i_2 .
- W wielomianie $g \in \mathbb{Q}[x_1, x_2, x_3]$ postaci $g = 7x_1x_2^2x_3^7 - 3x_1x_3^4 + 14x_2^5x_3$ mamy $a_{127} = 7$, $a_{104} = -3$, $a_{051} = 14$ oraz $a_{i_1 i_2 i_3} = 0$, dla pozostałych i_1, i_2, i_3 .

Uwaga. W świetle powyższej definicji wyrażenie postaci x_2x_1 nie jest wielomianem zmiennych x_1, x_2 o współczynnikach w żadnym ciele K , czyli nie należy do $K[x_1, x_2]$. Jest to natomiast wielomian zmiennych x_2, x_1 , czyli element $K[x_2, x_1]$. Wyrażenie $x_1x_2 + x_2x_1$ nie ma sensu ani jako element $K[x_1, x_2]$, ani jako element $K[x_2, x_1]$. W przyszłości poznacie Państwo obiekty zwane NIEPRZEMIENNymi WIELOMIANAMI zmiennych x_1, \dots, x_n , które oznacza się jako $K\langle x_1, \dots, x_n \rangle$ i nazywa czasem algebrą wolną o generatorach x_1, \dots, x_n nad ciałem K .

Definicja 11.5.2

STOPNIEM WIELOMIANU $\sum_{i_1, \dots, i_n} a_{i_1, \dots, i_n} x_1^{i_1} x_2^{i_2} \dots x_n^{i_n} \in K[x_1, \dots, x_n]$ nazywamy największą z liczb $i_1 + i_2 + \dots + i_n$, dla których $a_{i_1 \dots i_n} \neq 0$. Stopień wielomianu f oznaczamy $\deg f$. Jeśli f jest WIELOMIANEM ZEROWYM – to znaczy $a_{i_1 \dots i_n} = 0$, dla wszystkich i_1, \dots, i_n , to piszemy $\deg f = -\infty$.

Definicja 11.5.3

SUMĄ WIELOMIANÓW $\sum_{i_1, \dots, i_n} a_{i_1 \dots i_n} x_1^{i_1} x_2^{i_2} \dots x_n^{i_n}$ oraz $\sum_{i_1, \dots, i_n} b_{i_1 \dots i_n} x_1^{i_1} x_2^{i_2} \dots x_n^{i_n}$ nazywamy wielomian $\sum_{i_1, \dots, i_n} c_{i_1 \dots i_n} x_1^{i_1} x_2^{i_2} \dots x_n^{i_n}$ taki, że $c_{i_1 \dots i_n} = a_{i_1 \dots i_n} + b_{i_1 \dots i_n}$, dla każdych i_1, \dots, i_n .

Dla wielomianów w i g z wcześniejszych przykładów mamy $\deg w = 2$, $\deg g = 10$. Sumą wielomianów $2x_1^2x_2 + 6x_1x_2 - 5x_1$ oraz $7x_1^5 - 2x_1x_2 + 5x_1$ jest wielomian $7x_1^5 + 2x_1^2x_2 + 4x_1x_2$.

Szczególnymi typami wielomianów są wielomiany liniowe, to znaczy wielomiany stopnia 1, np.

$$x_1, x_1 + \dots + x_n, \quad 2x_1 + x_3 - x_4.$$

Rozwiązywanie jednorodnych układów równań liniowych jest z tej perspektywy szczególnym przypadkiem rozwiązywania wielomianowych układów równań. Ich rozwiązaniami są podprzestrzenie liniowe. Rozwiązaniami układów równań wielomianowych są tzw. zbiory algebraiczne. Powiemy o nich więcej na zakończenie drugiego semestru. Istotne jest to, że znamy wiele zbiorów algebraicznych, np. zbiór zer wielomianu dwóch zmiennych $x_1^2 - x_2^2$ to dwie przecinające się proste, a zbiór rozwiązań równania $x_1^2 - x_2$ to parabola (są też sfery, walce, hiperboloidy itd.). Badanie układów równań wielomianowych to punkt wyjścia wielkiego działu matematyki – geometrii algebraicznej. O czym jest zatem² Twierdzenie Hilberta? Zaczniemy od „stosunkowo prostej” sytuacji.

Weźmy element $a = (a_1, \dots, a_n) \in K^n$ i zastanówmy się jak może wyglądać zbiór wielomianów n zmiennych, które zerują się na a , czyli $\mathcal{W}\{a\}$. W przypadku wielomianów jednej zmiennej jest to po prostu zbiór

$$(x - a)f, \quad \text{gdzie } f \in K[x].$$

²Na motywach tekstu prof. Andrzeja Nowickiego: Afiniczne zbiory algebraiczne (do znalezienia na stronie Profesora) oraz tekstu Hilbert's Nullstellensatz w *The Princeton Companion to Mathematics*.

W przypadku wielomianów wielu zmiennych wnioskować należy, że w $\mathcal{W}(\{a\})$ jest każdy wielomian postaci:

$$(x_1 - a_1)f_1 + (x_2 - a_2)f_2 + \dots + (x_n - a_n)f_n,$$

gdzie $f_1, \dots, f_n \in K[x_1, \dots, x_n]$ są dowolne. Można, jak się okazuje pokazać, że powyższy zbiór jest w istocie całym $\mathcal{W}(\{a\})$. Nie jest to bardzo trudne. Trudniej rozwiązywać problem odwrotny.

Twierdzenie Hilberta o zerach podejmuje następujący problem: założmy, że startujemy od pewnego zbioru wielomianów n zmiennych nad k , wyznaczamy wszystkie punkty w K^n , na których mogą się one zerować, a potem dla tych punktów wyznaczamy wszystkie wielomiany, które się na nich zerują. Co dostajemy? Innymi słowy, jeśli $X \subseteq K[x_1, \dots, x_n]$, to czym jest $\mathcal{W}(\mathcal{R}(X))$?

Nawet w przypadku wielomianów jednej zmiennej możemy otrzymać nietrywialne odpowiedzi, jak widzieliśmy wyżej. Warto założyć chociaż, że ciało K jest algebraicznie domknięte, żeby nie martwić się zbiorami pustymi. Założmy, że X jest skończonym zbiorem złożonym z wielomianów f_1, \dots, f_m . Zbiór $\mathcal{R}(X)$ to zbiór jego wspólnych pierwiastków. Czym jest teraz $\mathcal{W}(\mathcal{R}(X))$? Jakie jeszcze wielomiany zerują się na tym samym zbiorze, co wielomiany f_1, \dots, f_m ? Na pewno są to wielomiany postaci:

$$f_1g_1 + \dots + f_mg_m,$$

gdzie g_1, \dots, g_m są dowolnymi wielomianami z $K[x_1, \dots, x_n]$. Co jeszcze? Czasem coś jeszcze, bo np.

$$(x - 1) \in \mathcal{W}(\mathcal{R}(\{(x - 1)^2\})).$$

Okazuje się, że jest to jedyny rodzaj „niespodzianki”, o czym mówi słynny wynik Hilberta.

Twierdzenie 11.5.4: Hilberta o zerach

Niech K będzie ciałem algebraicznie domkniętym oraz niech f_1, \dots, f_m należą do $K[x_1, \dots, x_n]$. Wówczas jeśli $h \in \mathcal{W}(\mathcal{R}(f_1, \dots, f_m))$ (tzn. funkcja wielomianowa odpowiadająca h zeruje się na podzbiorze K^n , będącym częścią wspólną zbiorów zer funkcji wielomianowych odpowiadających f_1, \dots, f_m), to istnieje $r \in \mathbb{Z}_+$ oraz wielomiany g_1, \dots, g_m , że:

$$h^r = f_1g_1 + \dots + f_mg_m.$$

W ten sposób wyróżnione zostają zbiory wielomianów X , dla których

$$\mathcal{W}(\mathcal{R}(X)) = X.$$

Zbiory te są bowiem w odpowiedniości ze zbiorami rozwiązań wielomianowych układów równań nad ciałem algebraicznie domkniętym, w podobny sposób jak w zależności tej są podprzestrzenie liniowe ze zbiorami rozwiązań jednorodnych liniowych układów równań. Dla zainteresowanych zostawiam jedynie hasło: ideał radykalny. Porównajmy tę sytuację z twierdzeniem Kroneckera-Capellego. Czy Czytelnik wie, porównując z (*), że jest ono w jakimś sensie szczególnym przypadkiem twierdzenia Hilberta?

Olimpijczykom znany może być następujący fakt żyjący pod nazwą „kombinatoryczne Nullstellensatz”.

Twierdzenie 11.5.5

Niech p będzie niezerowym wielomianem zmiennych x_1, \dots, x_n stopnia $\sum_{i=1}^n m_i$, w którym współczynnik przy $x_1^{m_1} \cdots x_n^{m_n}$ jest różny od zera. Wówczas dla dowolnych zbiorów S_1, \dots, S_n zawartych w \mathbb{R} spełniających warunki $|S_i| > m_i$, dla $1 \leq i \leq n$, istnieją takie $c_i \in S_i$, że $p(c_1, \dots, c_n) \neq 0$.

Zainteresowanych tym twierdzeniem i jego ładnymi elementarnymi aspektami odsyłam na przykład do artykułu Jacka Dymela „O zastosowaniach Combinatorial Nullstellensatz”, dostępnego na stronach Delty: <http://www.deltami.edu.pl/temat/matematyka/algebra/2017/06/16/2017-07-delta-dymel.pdf>.

11.6 Trivia. Zadanie o wymiarze podprzestrzeni macierzy

Zadanie. Niech K będzie ciałem oraz niech $n > 1$. Podzbiór \mathcal{S} w przestrzeni $M_{n \times n}(K)$ złożony jest ze wszystkich macierzy $S = (s_{ij})$ o wyrazach ze zbioru $\{0, 1\}$, spełniających warunki

$$s_{ii} = 0, \text{ dla } i = 1, 2, \dots, n \quad \text{oraz} \quad s_{ij} + s_{ji} = 1, \text{ dla dowolnych } 1 \leq i < j \leq n.$$

Niech $K = \mathbb{Z}_2$. Rozstrzygnij czy \mathcal{S} jest podprzestrzenią $M_{n \times n}(K)$ oraz wyznacz $\dim \text{lin}(\mathcal{S})$.

Rozwiązanie. Zauważmy, że macierz zerowa nie należy do \mathcal{S} , więc nie jest to podprzestrzeń $M_{n \times n}(K)$.

Pokażemy, że dla dowolnego ciała K wymiar przestrzeni $\text{lin}(\mathcal{S}) \subset M_{n \times n}(K)$ to

$$\frac{n(n-1)}{2} + 1.$$

Zaczniemy od intuicji, a potem przedstawimy formalny dowód. Każda macierz z $\text{lin}(\mathcal{S})$ jest, niezależnie od ciała, wyznaczona jednoznacznie przed określenie czy poszczególne wyrazy pod przekątną równe są 0 czy 1 oraz przez określenie dowolnego wyrazu nad przekątną. Istotnie, wyrazy dowolnej macierzy $A = (a_{ij}) \in \mathcal{S}$ spełniają, dla $i \neq j$, warunki:

$$a_{12} + a_{21} = a_{ij} + a_{ji} = 1,$$

a zatem równości $c_{ji} = c_{ij} - c_{12} - c_{21}$, dla $j > i$, zachodzą dla każdej macierzy $C = (c_{ij})$ z $\text{lin}(\mathcal{S})$. Intuicyjnie zatem układ opisujący macierze z $\text{lin}(\mathcal{S})$ zależy od $\frac{n(n-1)}{2} + 1$ parametrów i to jest właśnie wymiar $\text{lin}(\mathcal{S})$.

Przejdźmy do dowodu (przykładowego, bo wychodząc z powyższych intuicji można wskazać bazę $\text{lin}(\mathcal{S})$). Weźmy $C \in \text{lin}(\mathcal{S})$. Mamy:

$$C = a_1 S_1 + a_2 S_2 + \dots + a_k S_k,$$

gdzie $a_1, \dots, a_k \in K$ oraz $S_1, \dots, S_k \in \mathcal{S}$. A zatem wyrazy macierzy C spełniają warunki

$$c_{ij} + c_{ji} = a_1 + a_2 + \dots + a_k.$$

Oznacza to, że dla $c = a_1 + a_2 + \dots + a_k$ wyrazy c_{ij} macierzy C są rozwiązaniami układu U_c postaci:

$$U_c : \begin{cases} c_{11} &= 0 \\ \vdots & \\ c_{nn} &= 0 \\ c_{12} + c_{21} &= c \\ \vdots & \\ c_{n-1,n} + c_{n,n-1} &= c. \end{cases}$$

Innymi słowy, układ U_c jest złożony z:

- n równań postaci $c_{11} = 0, \dots, c_{nn} = 0$,
- $\frac{n(n-1)}{2}$ równań postaci $c_{ij} + c_{ji} = c$.

Rząd macierzy jednorodnego układu U_0 wynosi $\textcolor{red}{n} + \frac{\textcolor{blue}{n(n-1)}}{2}$, bo każda z $\textcolor{blue}{n^2}$ niewiadomych występuje tylko w jednym równaniu. Stąd wymiar przestrzeni rozwiązań W_0 układu U_0 to, zgodnie z tw. Kroneckera-Capellego

$$\dim W_0 = \textcolor{blue}{n^2} - \textcolor{red}{n} - \frac{\textcolor{red}{n(n-1)}}{2} = \frac{n(n-1)}{2}.$$

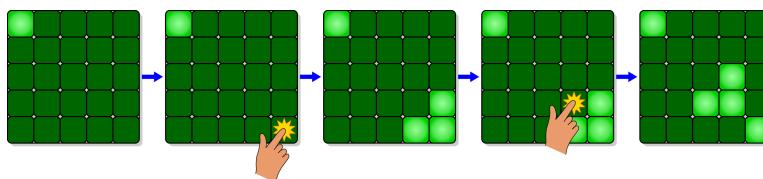
Wiemy z wykładu, że zbiór wszystkich macierzy spełniających niejednorodny układ U_1 ma tę własność, że różnica dowolnych dwóch macierzy z tego zbioru spełnia układ jednorodny U_0 , czyli jest w W_0 . Co więcej, jeśli dla $c \neq 0$ macierz M spełnia U_c , to macierz $c^{-1} \cdot M$ spełnia U_1 . Oznacza to, że biorąc bazę W_0 oraz dowolną macierz $C \in \text{lin}(\mathcal{S}) \setminus W_0$ dostajemy bazę $\text{lin}(\mathcal{S})$. W szczególności:

$$\dim \text{lin}(\mathcal{S}) = \dim W_0 + 1 = \frac{n(n-1)}{2} + 1.$$

■

11.7 Trivia. Lights Out

W 1995 roku Tiger Electronics wydało grę *Lights Out*. Gra składa się z tablicy rozmiaru 5 na 5 złożonej z 25 przycisków, z których każdy jest w jednym z dwóch stanów: włączony (wtedy przycisk jest podświetlony) lub wyłączony. Po rozpoczęciu gry włącza się losowa konfiguracja przycisków. Naciśnięcie dowolnego przycisku spowoduje przełączenie tego przycisku, a także jego sąsiadów (ale nie po przekątnej).



Zadaniem jest wyłączenie wszystkich światel, najlepiej za pomocą jak najmniejszej liczby ruchów.

W 1998 roku Marlow Anderson oraz Todd Fell użyli metod algebry liniowej³ do pokazania, że nie wszystkie konfiguracje prowadzą do rozwiązania oraz, że dla każdej rozwiązywalnej konfiguracji istnieją dokładnie cztery strategie (nie mające zbędnych ruchów). Idea jest prosta: każdą z tablic reprezentować można jednoznacznie jako element przestrzeni $M_5(\mathbb{Z}_2)$. W ten sposób wciśnięcie klawisza w i -tym wierszu i j -tej kolumnie odpowiada dodawaniu pewnej macierzy $t_{ij} \in M_5(\mathbb{Z}_2)$, na przykład (dla ilustracji wyżej):

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

Dla dowolnych macierzy $x, y \in M_5(\mathbb{Z}_2)$ mamy $x + y = y + x$ oraz $x + x = 0$, więc widzimy, że kolejność wciskania przycisków nie ma znaczenia dla strategii oraz żadnego przycisku nie trzeba wciskać więcej niż raz. Założmy, że wyjściową konfigurację światel opisuje macierz b . Problem wyłączenia światel w b równoważny jest zatem pytaniu: czy b należy do $\text{lin}(t_{ij}, 1 \leq i, j \leq 5)$? A zatem jest to w istocie problem rozwiązania układu równań — w tym przypadku o 25 niewiadomych (o współczynnikach w \mathbb{Z}_2). Nietrudno sprawdzić, że macierz $A \in M_{25 \times 25}(\mathbb{Z}_2)$ tego układu ma 25 bloków rozmiaru 5×5 postaci:

$$A = \begin{bmatrix} C & I_5 & 0 & 0 & 0 \\ I_5 & C & I_5 & 0 & 0 \\ 0 & I_5 & C & I_5 & 0 \\ 0 & 0 & I_5 & C & I_5 \\ 0 & 0 & 0 & I_5 & C \end{bmatrix}, \quad \text{gdzie} \quad C = \begin{bmatrix} 1 & 1 & 0 & 0 & 0 \\ 1 & 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 1 & 0 \\ 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 1 & 1 \end{bmatrix} \quad \text{oraz} \quad I_5 = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}.$$

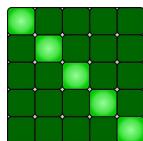
Widzimy zatem, że sprawdzając, czy można „zgasić światła” w macierzy b rozwiązujemy układ równań $Ax = b$, gdzie b jest wektorem o 25 współrzędnych złożonym z kolejnych wyrazów macierzy b . Dla przykładu, gdy b jest macierzą po prawej na ilustracji wyżej, to wektor b ma postać:

$$b = (1, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 1, 0, 0, 0, 1, 1, 0, 0, 0, 0, 0, 1).$$

Wektor $x \in \mathbb{Z}_2^{25}$ wskazuje które przyciski mamy wcisnąć, aby ze zgaszonej macierzy dostać b .

Jak można policzyć (najlepiej przy pomocy komputera), rząd macierzy $A \in M_{25 \times 25}(\mathbb{Z}_2)$ równy jest 23. Oznacza to, że liczba rozwiązywalnych konfiguracji gry równa jest 2^{23} . Prawdopodobieństwo, że losowa konfiguracja jest rozwiązywalna równe jest $2^{23}/2^{25} = 1/4$.

Odnajdujmy dwie ciekawe sytuacje: pierwsza, gdy rozwiązanie układu $Ax = 0$ zwraca strategie, które nie zmieniają oświetlenia na tablicy oraz druga, gdy rozwiązanie układu $Ax = x$ zwraca te strategie, w których wciśnięcie odpowiednich przycisków sprawia, że jedynie wciśnięte przyciski są zapalone. Przykładem tej sytuacji jest wciśnięcie przycisków na przekątnej, startując od pustej tablicy. Jest jeszcze 31 innych.



³Wyjaśnienie tego szkicu i ładny opis problemu można znaleźć w rozdziale 24. pięknej książki „Permutation puzzles”, dostępnej pod adresem <http://www.sfu.ca/~jtmulhol/math302/notes/permutation-puzzles-book.pdf> (strona autora).

11.8 Coda. Bardzo wstępnie o twierdzeniach klasyfikacyjnych

Twierdzenie Kroneckera-Capellego przeprowadza klasyfikację podprzestrzeni przestrzeni liniowej K^n w języku układów równań. Każda taka podprzestrzeń wymiaru k może być opisana za pomocą układu równań liniowych złożonego z $n - k$ elementów, którego macierz ma rząd $n - k$. Tu w istocie zawarta jest bardzo głęboka treść, której jeszcze nie wysłowiliśmy. Twierdzenie to mówi w istocie, że jeśli rozważymy podprzestrzeń wymiaru k i opiszemy ją poprzez układ $n - k$ równań zapisany w postaci ogólnej:

$$\begin{cases} x_{j_1} &= c_{11}x_{t_1} + c_{12}x_{t_2} + \dots + c_{1,k}x_{t_k} \\ &\vdots \\ x_{j_{n-k}} &= c_{n-k,1}x_{t_1} + c_{n-k,2}x_{t_2} + \dots + c_{n-k,k}x_{t_k} \end{cases}$$

to dokonując w tych równaniach **liniowej zamiany zmiennych** postaci:

$$\begin{cases} y_1 &= x_{j_1} - (c_{11}x_{t_1} + c_{12}x_{t_2} + \dots + c_{1,k}x_{t_k}) \\ &\vdots \\ y_{n-k} &= x_{j_{n-k}} - (c_{n-k,1}x_{t_1} + c_{n-k,2}x_{t_2} + \dots + c_{n-k,k}x_{t_k}) \\ y_{n-k+1} &= 0 \\ &\vdots \\ y_n &= 0 \end{cases},$$

to w owym układzie współrzędnych nasza podprzestrzeń opisana jest równaniami $y_1 = 0, \dots, y_{n-k} = 0$.

Oto przykład: weźmy dwuwymiarową podprzestrzeń W w \mathbb{R}^3 opisaną równaniem $x_1 - x_2 + x_3 = 0$. Określenie to mówi, że podprzestrzeń W złożona jest z wektorów

$$v = x_1 \cdot (1, 0, 0) + x_2 \cdot (0, 1, 0) + x_3 \cdot (0, 0, 1),$$

których współrzędne w bazie standardowej przestrzeni \mathbb{R}^3 spełniają równanie wyżej. A co by się stało, gdybyśmy rozważali zbiory wektorów, których współrzędne spełniają równania liniowe, ale nie są to współrzędne w bazie standardowej? Jakie? Możemy mianowicie rozważyć liniową zamianę zmiennych postaci $y_1 = x_1 - x_2 + x_3$ oraz $y_2 = 0, y_3 = 0$. Wówczas nasza podprzestrzeń opisana jest równanie $y_1 = 0$. Liniowa zamiana zmiennych oznacza, że teraz rozważamy współrzędne z innej bazy. Jakiej? Weźmy bazę zawierającą jako pierwszy wektor $(1, -1, 1)$, a dwa pozostałe — to elementy bazy W , na przykład: $(1, -1, 1), (1, 1, 0), (1, 0, -1)$. Innymi słowy, rozważając wektory postaci $y_1(1, -1, 1) + y_2(1, 1, 0) + y_3(1, 0, -1)$ stwierdzamy łatwo, że wektory te należą do W wtedy i tylko wtedy, gdy $y_1 = 0$.

Twierdzenia klasyfikacyjne obecne są w całej nauce. Być może klasycznym dziełem w historii nauki pokazującym jej dążenie do klasyfikowania obiektów w zależności od tego ile wspólnych cech posiadają, była *Systema Naturae* Karola Linneusza z 1735 roku. W tym rozumieniu „charakteryzacja” ukazuje po prostu „wspólny charakter” — w tym przypadku chodzi o klasyfikację biologiczną organizmów. Twierdzenia klasyfikacyjne matematyki, idą nieco dalej. Przechodzą one bowiem do języka niezmienników i funkcji.

Dla dwóch krzywych opisanych równaniem kwadratowym, na przykład elipsy i hiperboli, nie wystarczy powiedzieć, że posiadają one pewne charakterystyczne własności. Wymagamy, aby istniała transformacja/przekształcenie określonego typu, które przeprowadzi jedną krzywą w drugą. Innymi słowy, interesuje nas, by obiekty były identyczne modulo pewna operacja. Być może Czytelnika zaciekawi fakt, że w zależności od przyjętego typu przekształcenia, okrąg i hiperbola mogą być równoważne, lub nie. Powiemy o tym w drugim semestrze: nie istnieje izometria lub izomorfizm afaniczny przeprowadzający elipsę w hiperbolę, ale istnieje przekształcenie rzutowe, które tego dokonuje.

Na najbliższych wykładach poznamy pojęcie przekształcenia liniowego i izomorfizmu przestrzeni liniowych. Będzie to pierwsza kluczowa klasa przekształceń, która będzie utożsamiała przestrzenie liniowe, które uważamy jako takie same. Udoskonimy też, że każda skończona wymiarowa przestrzeń liniowa wymiaru n nad ciałem K jest izomorficzna z przestrzenią K^n . Ten właśnie izomorfizm będzie formalnie zapisywać w sobie liniową zamianę zmiennych za pomocą pewnej macierzy odwracalnej. Wkrótce dowiemy się co to wszystko znaczy. Przedstawmy kilka przykładów z historii matematyki.

Wielkim osiągnięciem była klasyfikacja wielościanów foremnych w przestrzeni trójwymiarowej — jest ich tylko 5. Ten wynik dawał starożytnym silny impuls filozoficzny do pewnej konstrukcji światopoglądowej — mówiącej, że wszystko na świecie ma pierwotną przyczynę — Arché. Nie było zgody, gdzie jest ona umieszczona. Jedni widzieli ją w wodzie, inni w bezkresie, inni w powietrzu inni w ogniu. Później próbowało scalać te koncepcje i u podstaw rzeczywistości widziano kilka podstawowych *elementów*, działających przeciwnie, których wzajemne oddziaływanie miało być źródłem zmiany. Inna koncepcja pochodziła od Demokryta, który wszystkie rzeczy materialne postrzegał jako stworzone z małych, niepodzielnych częstek (atomów). Różne proporcje ich połączeń miały prowadzić do różnorodności rzeczy.

Wielościany foremne nazywany czasem bryłami platońskimi właśnie dlatego, że słynny ateński filozof uważały, że są one budulcem owych „podstawowych elementów” czy „atomów”. Ogień miał być zbudowany z czworościanów foremnych, ziemia z sześciąnów, powietrze z ośmiościanów, woda z dwudziestościanów. Dwudziestościan reprezentować miał żywioł niebieski. Arystoteles nazywał go eterem. Kierunek ten nie pochodził jedynie od uczonych greckich. Niekiedy formułowany był w kulturze starożytnej Chin, Japonii, czy Indii. W czasach nowożytnych koncepcje żywiołów obecne były w badaniach alchemicznych, którego przedmiotem zainteresowania były metale i ich „transmutacje”, których celem miało być otrzymanie złota lub odkrycie kamienia filozoficznego. Rozwój filozofii przyrody, badań empirycznych i aparatury w XVII wieku doprowadził do skupienia się na bardziej ogólnych celach poznanawczych i przyczyniło się stopniowo do sformułowania idei pierwiastka chemicznego. Na koncepcji pięciu żywiołów opierał się jeszcze Kartezjusz — twórca geometrii analitycznej.

Z czego składa się matematyczne twierdzenie klasyfikacyjne? Ma ono dwa elementy:

- Listę normalnych (kanonicznych) form, reprezentantów danej relacji równoważności — w przypadku rozważanej przez nas teorii są to równania postaci, na przykład $x_1 = 0, x_2 = 0, \dots, x_{n-k} = 0$.
- Twierdzenie klasyfikacyjne stwierdzający, że każdy obiekt jest równoważny do jednej z form normalnych (kanonicznych). W naszym przypadku chodzi o stwierdzenie, że dla każdej przestrzeni wymiaru k istnieje liniowa zamiana zmiennych (izomorfizm liniowy) taka, że w odpowiednim układzie współrzędnych podprzestrzeń opisać można dokładnie jednym z równań kanonicznych. Co więcej twierdzenie to orzeka, że każdych dwóch różnych form kanonicznych nie można przekształcić na siebie — w naszym przypadku — za pomocą liniowej zamiany zmiennych.

Zarówno na GALu, jak i na kolejnych przedmiotach poznają Państwo szereg pięknych twierdzeń klasyfikacyjnych. Są one bardzo często celem całego kursu, i często to są właśnie „twierdzenia z nazwiskiem”. W drugim semestrze zmierzać będziemy do następujących twierdzeń klasyfikacyjnych:

- twierdzenie Jordana o klasyfikacji endomorfizmów przestrzeni liniowej skończonego wymiaru nad ciałem algebraicznie domkniętym,
- twierdzenie Cartana o rozkładzie dowolnej izometrii liniowej n -wymiarowej przestrzeni liniowej na złożenie nie więcej niż n symetrii prostopadłych, zawierające w sobie klasyfikację izometrii płaszczyzny i przestrzeni trójwymiarowej, dokonaną już wieki wcześniejsi,
- twierdzenie spektralne o ortogonalnej diagonalizacji, pochodzące w pierwszych wersjach od Kartezjusza i Fermata, w kolejnych od twierdzenie Eulera (twierdzenie o osiach głównych), aż do ogólnych rozważań Lagrange'a i Jacobiego, skumulowanych w pracach Cauchy'ego i Sylvester, które wspólną formę przybrały dzięki Frobeniusowi⁴
- twierdzenia Sylvester'a o inercji, o klasyfikującego rzeczywiste formy dwuliniowe, rozróżniające między innymi czterowymiarową przestrzeń euklidesową i czasoprzestrzeń,
- twierdzenie klasyfikujące kwadryki — powierzchnie stopnia 2 w afinicznej przestrzeni euklidesowej.

Na innych wykładach poznawać będziecie Państwo twierdzenia klasyfikujące grupy przemienne, powierzchnie wyższych stopni (zarówno od strony algebraicznej jak i różniczkowej), rozmaistości na analizie i topologii i wiele innych. Twierdzenia te ukazują jedynie wierzchołek współczesnej konstrukcji matematyki, w której odkrywamy coraz mocniej, jak twierdzenia klasyfikujące obiekty w różnych „kategoriach” mają się do siebie. Będziemy do tego tematu wracać. Morał na dziś brzmi: kluczowym elementem działalności matematycznej jest klasyfikowanie. Wielkie problemy matematyki to między innymi problemy klasyfikacyjne, np. rozwiązana przez Perelmana (w roku 2002) słynna Hipoteza Poincarégo z roku 1904.

⁴L.A. Steen, *Highlight in the history of spectral theory*, The American Mathematical Monthly, Vol. 80, No. 4 (Apr. 1973), pp. 359-381, <https://www.jstor.org/stable/2319079>.

Rozdział 12

Operacje na podprzestrzeniach

12.1 Wykład 12

Dziś zobaczymy w jaki sposób poznane metody pozwalają nam poruszać się w świecie podprzestrzeni przestrzeni liniowej, w szczególności w jaki sposób pozwalają na przypisanie parze czy też całej rodzinie podprzestrzeni dwóch naturalnych obiektów – sumy i części wspólnej. Szczególną rolę grają też rozkłady przestrzeni na sumy proste podprzestrzeni. Zaczniemy od powrotu do przestrzeni K^n i zobaczymy w jaki sposób mając dane jej podprzestrzenie konstruować możemy nowe, związane z nimi podprzestrzenie.

- Jeśli V_1, V_2 są podprzestrzeniami w K^n opisanymi układami równań U_1 oraz U_2 , to podprzestrzeń

$$V_1 \cap V_2$$

jest opisana układem równań złożonym ze wszystkich równań z U_1 oraz wszystkich równań z U_2 . Na przykład, jeśli $W_1, W_2 \subseteq \mathbb{R}^3$ opisane są odpowiednio układami $x_1 + x_2 + x_3 = 0, x_1 - x_3 = 0$ oraz $x_1 - x_2 = 0$, to największa podprzestrzeń zawierająca elementy zarówno z W_1 , jak i W_2 , czyli właśnie $W_1 \cap W_2$, opisana jest układem równań:

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

- Jeśli $V_1 = \text{lin}(\alpha_1, \dots, \alpha_n)$ oraz $V_2 = \text{lin}(\beta_1, \dots, \beta_m)$ są podprzestrzeniami przestrzeni V , to podprzestrzeń:

$$W = \text{lin}(\alpha_1, \dots, \alpha_n, \beta_1, \dots, \beta_m)$$

złożona jest ze wszystkich wektorów postaci $\alpha + \beta$, gdzie $\alpha \in V_1, \beta \in V_2$ jest najmniejszą podprzestrzenią w V , która zawiera jednocześnie V_1 oraz V_2 . Innymi słowy, jest to przestrzeń $\text{lin}(V_1 \cup V_2)$.

Na przykład, jeśli $V_1 = \text{lin}((1, 0, 1)), V_2 = \text{lin}((1, 0, 2)) \subseteq \mathbb{R}^3$, to najmniejsza podprzestrzeń \mathbb{R}^3 zawierająca te dwie podprzestrzenie to $\text{lin}((1, 0, 1), (1, 0, 2))$.

Definicja 12.1.1: Suma podprzestrzeni

Niech $X, Y \subseteq V$. Przez $X + Y$ oznaczać będziemy zbiór

$$\{x + y \mid x \in X, y \in Y\}.$$

Jeśli X, Y są podprzestrzeniami w przestrzeni V , to $X + Y$ też jest podprzestrzenią przestrzeni V zwaną SUMĄ PODPRZESTRZENI X i Y .

Również pojęcie części wspólnej podprzestrzeni przestrzeni K^n przenosi się na dowolne przestrzenie liniowe. Następującą obserwację pozostawiamy jako kolejne proste ćwiczenie.

Uwaga 12.1.2

Część wspólna $X \cap Y$ podprzestrzeni liniowej V jest podprzestrzenią liniową. Nazywamy ją ILOCZYNEM, PRZECIĘCIEM (lub CZEŚCIĄ WSPÓLNAĄ) podprzestrzeni.

Zauważmy zatem, że z wraz dowolnymi dwiema podprzestrzeniami V_1, V_2 przestrzeni V rozważać można dwie podprzestrzenie: $V_1 \cap V_2$ – będącą ich największą wspólną podprzestrzenią oraz $V_1 + V_2$ – najmniejszą przestrzeń liniową, której V_1, V_2 są podprzestrzeniami. Kluczowa dla zrozumienia związku w opisany układzie podprzestrzeni jest następująca formuła.

Twierdzenie 12.1.3: Formuła Grassmanna, 1844

Niech V_1, V_2 będą skończenie wymiarowymi podprzestrzeniami przestrzeni V . Wówczas podprzestrzenie $V_1 \cap V_2$ oraz $V_1 + V_2$ też są skończenie wymiarowe i zachodzi:

$$\dim(V_1 + V_2) = \dim V_1 + \dim V_2 - \dim(V_1 \cap V_2).$$

Dowód. Niech $\gamma_1, \dots, \gamma_m$ będzie bazą przestrzeni $V_1 \cap V_2$. Uzupełniamy ją, na mocy twierdzenia Steinitza, do bazy $\gamma_1, \dots, \gamma_m, \alpha_1, \dots, \alpha_k$ przestrzeni V_1 oraz do bazy $\gamma_1, \dots, \gamma_m, \beta_1, \dots, \beta_l$ przestrzeni V_2 . Wykażemy, że układ $\gamma_1, \dots, \gamma_m, \alpha_1, \dots, \alpha_k, \beta_1, \dots, \beta_l$ jest bazą przestrzeni $V_1 + V_2$. Oczywiście układ ten rozpinia tą przestrzeń. Pozostaje wykazać jego liniową niezależność.

Przypuśćmy, że $c_1\gamma_1 + \dots + c_m\gamma_m + a_1\alpha_1 + \dots + a_k\alpha_k + b_1\beta_1 + \dots + b_l\beta_l = 0$. Stąd wynika, że

$$c_1\gamma_1 + \dots + c_m\gamma_m + a_1\alpha_1 + \dots + a_k\alpha_k = -(b_1\beta_1 + \dots + b_l\beta_l) \in V_1. \quad (12.1)$$

Zauważmy jednak, że jeśli $b_1\beta_1 + b_2\beta_2 + \dots + b_l\beta_l \in V_1$, to kombinacja ta należy w istocie do iloczynu $V_1 \cap V_2$, a więc jest równa pewnej kombinacji postaci $c'_1\gamma_1 + \dots + c'_m\gamma_m$. Wtedy jednak

$$b_1\beta_1 + \dots + b_l\beta_l - c'_1\gamma_1 - \dots - c'_m\gamma_m = 0,$$

co z liniowej niezależności układu $\beta_1, \dots, \beta_l, \gamma_1, \dots, \gamma_m$ implikuje, że $b_1 = b_2 = \dots = b_l = c'_1 = \dots = c'_m = 0$.

Powyższy argument oznacza, że we wzorze (12.1) całą kombinację liniową $-(b_1\beta_1 + \dots + b_l\beta_l)$ możemy zastąpić przez 0. Mamy zatem:

$$c_1\gamma_1 + \dots + c_m\gamma_m + a_1\alpha_1 + \dots + a_k\alpha_k = 0.$$

To jest jednak kombinacja wektorów bazowych z V_1 , co oznacza, że $c_1 = \dots = c_m = a_1 = \dots = a_k = 0$. Istotnie więc

$$\dim(V_1 + V_2) = m + k + l = (m + k) + (m + l) - m = \dim V_1 + \dim V_2 - \dim(V_1 \cap V_2).$$

□

Oto prosty przykład wniosku z powyższej formuły: jeśli V_1, V_2 są podprzestrzeniami wymiaru 4 w przestrzeni 6-wymiarowej V , to $V_1 \cap V_2 \geq 2$.

Istotnie, skoro $V_1 + V_2$ jest podprzestrzenią V , to $\dim(V_1 + V_2) \leq \dim V = 6$. Stąd

$$\dim(V_1 \cap V_2) = \dim V_1 + \dim V_2 - \dim(V_1 + V_2) \geq 4 + 4 - 6 = 2.$$

Definicja 12.1.4: Suma prosta podprzestrzeni

Jeśli dla pewnych podprzestrzeni X, Y przestrzeni liniowej V każdy wektor $\alpha \in V$ da się przedstawić jednoznacznie w postaci sumy wektorów $x \in X$ oraz $y \in Y$ to mówimy, że V jest SUMĄ PROSTĄ podprzestrzeni X, Y , co oznaczamy przez $V = X \oplus Y$.

Przykład 1. Zgodnie z opisem poczynionym wyżej mamy: $\mathbb{R}^2 = \text{lin}(1, 1) \oplus \text{lin}(1, -1)$.

Przykład 2. Każdy ciąg zbieżny o wyrazach rzeczywistych można w jednoznaczny sposób przedstawić jako sumę ciągu stałego i ciągu zbieżnego do 0. A zatem podprzestrzeń \mathcal{C} przestrzeni \mathbb{R}^∞ złożona z ciągów zbieżnych jest sumą prostą podprzestrzeni złożonej z ciągów stałych (jednowymiarowa) i podprzestrzeni \mathcal{C}_0 złożonej z ciągów zbieżnych do zera (nieskończoność wymiarowa – dlaczego?).

Przykład 3. Każdą funkcję $f : \mathbb{R} \rightarrow \mathbb{R}$ można przedstawić w sposób jednoznaczny jako sumę funkcji parzystej i nieparzystej, bo mamy rozkład:

$$f(x) = \frac{f(x) + f(-x)}{2} + \frac{f(x) - f(-x)}{2}$$

Dlaczego rozkłady przedstawione z Przykładach 2 i 3 są jednoznaczne i uzyskane sumy podprzestrzeni są sumami prostymi?

Uwaga 12.1.5

Niech V_1, V_2 będą podprzestrzeniami przestrzeni V . Następujące warunki są równoważne:

- $V = V_1 \oplus V_2$,
- $V = V_1 + V_2$ i $V_1 \cap V_2 = \{0\}$.

Dowód. Założymy najpierw, że $V = V_1 \oplus V_2$. Wówczas dla każdego $\alpha \in V$ mamy $\alpha = \alpha_1 + \alpha_2$, gdzie $\alpha_1 \in V_1, \alpha_2 \in V_2$, więc $V = V_1 + V_2$. Gdyby $V_1 \cap V_2 \neq \{0\}$, to istniałby niezerowy wektor $\alpha \in V_1 \cap V_2$. Co więcej, można by było przedstawić ten wektor na dwa sposoby jako sumę wektora z V_1 i V_2 , mianowicie: $\alpha = \alpha + 0 = 0 + \alpha$. Jest to sprzeczne z jednoznacznością w definicji sumy prostej.

Na odwrót: przypuśćmy, że $V = V_1 + V_2$ oraz $V_1 \cap V_2 = \{0\}$. Musimy wykazać, że każdy wektor $\alpha \in V$ daje się jednoznacznie przedstawić jako suma wektora z V_1 i wektora z V_2 . Z równości $V = V_1 + V_2$ wynika, że istnieją $\alpha_1 \in V_1$ oraz $\alpha_2 \in V_2$ spełniające $\alpha = \alpha_1 + \alpha_2$. Gdyby to przedstawienie nie było jednoznaczne, to zachodziłoby równanie $\alpha = \alpha'_1 + \alpha'_2$, dla pewnych $\alpha'_1 \in V_1, \alpha'_2 \in V_2$, przy czym $\alpha_1 \neq \alpha'_1$ (równoważnie: $\alpha_2 \neq \alpha'_2$). Wtedy jednak

$$0 \neq \alpha_1 - \alpha'_1 = \alpha'_2 - \alpha_2 \in V_1 \cap V_2,$$

co jest sprzeczne z $V_1 \cap V_2 = \{0\}$. □

Poniższy wniosek odnosi się bezpośrednio do formuły Grassmanna.

Wniosek 12.1.6

Niech V_1, V_2 będą podprzestrzeniami skończenia wymiarowej przestrzeni V . Założymy również, że $V_1 \cap V_2 = \{0\}$. Wówczas następujące warunki są równoważne:

- $V = V_1 \oplus V_2$,
- $\dim V = \dim V_1 + \dim V_2$,
- jeśli \mathcal{A} jest bazą V_1 oraz \mathcal{B} jest bazą V_2 , to $\mathcal{A} \cup \mathcal{B}$ jest bazą V .

Na koniec omówimy definicje uogólniające pojęcie sumy i iloczynu podprzestrzeni, a także pojęcie sumy prostej na dowolną rodzinę podprzestrzeni. Wyjściowa sytuacja jest następująca: dana jest rodzina podprzestrzeni $\{V_t\}_{t \in T}$ przestrzeni V , gdzie T może być zbiorem nieskończonym (np. zbiorem liczb naturalnych lub rzeczywistych). Interesuje nas znalezienie najmniejszej podprzestrzeni V , której podprzestrzeniami są wszystkie elementy rozważanej rodziny oraz znalezienie największej podprzestrzeni, będącej jednocześnie podprzestrzenią wszystkich podprzestrzeni należących do rozważanej rodziny.

Definicja 12.1.7

Niech $\{V_t\}_{t \in T}$ będzie rodziną podprzestrzeni przestrzeni V . Wówczas określamy zbiór

$$\sum_{t \in T} V_t = \{\alpha_{t_1} + \dots + \alpha_{t_r} \mid \alpha_{t_i} \in V_{t_i}, t_1, \dots, t_r \in T, r \in \mathbb{N}\},$$

zwany SUMĄ RODZINY PODPRZESTRZENI $\{V_t\}_{t \in T}$. W przypadku $T = \{1, \dots, n\}$ piszemy:

$$\sum_{i=1}^n V_i = V_1 + \dots + V_n.$$

Przykłady:

- $K[x] = \sum_{n \in \mathbb{N}} K_{\leq n}[x]$.
- $\mathbb{R}^3 = \text{lin}((1, 0, 0), (0, 1, 0)) + \text{lin}((1, 0, 0), (1, 1, 0)) + \text{lin}((1, 1, 0), (0, 1, 0))$.
- Przestrzeń $\text{lin}((1, 0, 1), (1, 0, 2)) \subseteq \mathbb{R}^3$ jest sumą rodzin podprzestrzeni indeksowanej (na przykład) liczbami niewymiernymi postaci:

$$V_t = \text{lin}(1, 0, t), \quad t \in T = \mathbb{R} \setminus \mathbb{Q}.$$

Uwaga 12.1.8

Niech $\{V_t\}_{t \in T}$ będzie rodziną podprzestrzeni przestrzeni V . Wówczas

$$\sum_{t \in T} V_t = \text{lin} \left(\bigcup_{t \in T} V_t \right).$$

W szczególności, suma rodzin podprzestrzeni V jest podprzestrzenią V . Jest to najmniejsza podprzestrzeń w V zawierająca wszystkie $\{V_t\}_{t \in T}$.

Wówczas podprzestrzeń

$$\bigcap_{t \in T} V_t$$

nazywamy ILOCZYNEM, PRZECIĘCIEM lub CZEŚCIĄ WSPÓLΝĄ rodziny $\{V_t\}_{t \in T}$.

Rozważmy następujący przykład: niech $V_{[0,1]} \subseteq F(\mathbb{R}, \mathbb{R})$ będzie podprzestrzenią złożoną ze wszystkich funkcji, które przyjmują wartość zero na zbiorze $[0, 1]$. Rozważmy też, dla $x \in [0, 1]$, podprzestrzenie $F(\mathbb{R}, \mathbb{R})$ postaci $V_x = \{f \in F(\mathbb{R}, \mathbb{R}) : f(x) = 0\}$. Wówczas:

$$V_{[0,1]} = \bigcap_{x \in [0,1]} V_x$$

Czytelnik znający zasadę włączeń i wyłączeń, pozwalającą wyznaczyć moc sumy skończenie wielu zbiorów skończonych, patrząc na formułę Grassmanna może dojść do przekonania, że zachodzić musi jej uogólnienie na przypadek wymiaru sumy trzech lub więcej składników. Jest to niestety nieprawda. W szczególności, jeśli V_1, V_2, V_3 są podprzestrzeniami V , to $\dim(V_1 + V_2 + V_3)$ nie jest równy:

$$\dim V_1 + \dim V_2 + \dim V_3 - \dim(V_1 \cap V_2) - \dim(V_1 \cap V_3) - \dim(V_2 \cap V_3) + \dim(V_1 \cap V_2 \cap V_3),$$

Czytelnik zechce sprawdzić to dla $V_t = \text{lin}(1, t)$, gdzie $t = 0, 1, 2$.

Prawidłowe uogólnienie formuły Grassmanna nie jest proste, o ile suma podprzestrzeni nie spełnia jakiegoś dodatkowego warunku. Najistotniejszym przykładem takiego warunku jest oczywiście bycie sumą prostą. Co to oznacza?

Definicja 12.1.9

Mówimy, że przestrzeń V jest SUMĄ PROSTĄ RODZINY PODPRZESTRZENI $\{V_t\}_{t \in T}$ jeśli każdy wektor $\alpha \in V$ daje się przedstawić jednoznacznie jako suma

$$\alpha_{t_1} + \dots + \alpha_{t_r},$$

gdzie $\alpha_{t_i} \in V_{t_i}$, dla pewnych parami różnych $t_i \in T$. Wówczas piszemy:

$$V = \bigoplus_{t \in T} V_t,$$

a w przypadku, gdy $T = \{1, \dots, n\}$ po prostu

$$V = V_1 \oplus \dots \oplus V_n.$$

Przykłady (zachęcam do samodzielnego uzasadnienia):

- $\mathbb{R}^4 = \text{lin}(1, 0, 0, 0) \oplus \text{lin}(0, 1, 0, 0) \oplus \text{lin}((1, 1, 1, 0), (0, 0, 0, 1))$.
- Jeśli $\mathcal{A} = \{\alpha_1, \dots, \alpha_n\}$ jest bazą przestrzeni V , to:

$$V = \text{lin}(\alpha_1) \oplus \text{lin}(\alpha_2) \oplus \dots \oplus \text{lin}(\alpha_n).$$

- Jeśli \mathcal{A} jest bazą V oraz $\mathcal{A} = \mathcal{A}_1 \cup \dots \cup \mathcal{A}_n$ jest rozbiciem \mathcal{A} na parami rozłączne podzbiory, to:

$$V = \bigoplus_{i=1}^n \text{lin}(\mathcal{A}_i).$$

Jak się okazuje, jeśli przestrzeń skończenie wymiarowa V spełnia $V = V_1 \oplus \dots \oplus V_n$, to

$$\dim V = \sum_{i=1}^n \dim V_i.$$

Dowód wymaga uzasadnienia następującej obserwacji.

Uwaga 12.1.10

Niech $\{V_t\}_{t \in T}$ będzie rodziną podprzestrzeni przestrzeni V . Wówczas następujące warunki są równoważne:

- $V = \bigoplus_{t \in T} V_t$
- dla każdych $t_0, t_1, \dots, t_k \in T$ zachodzi: $V_{t_0} \cap \sum_{i=1}^k V_{t_i} = \{0\}$.

W szczególności aby suma algebraiczna była sumą prostą $V = V_1 \oplus V_2 \oplus V_3$ nie wystarczy, aby mieć

$$V = V_1 + V_2 + V_3, \quad \text{oraz} \quad V_1 \cap V_2 \cap V_3 = \{0\}.$$

Dla przykładu weźmy podprzestrzenie \mathbb{R}^2 postaci:

$$V_1 = \text{lin}(1, 0), \quad V_2 = \text{lin}(1, 1), \quad V_3 = \text{lin}(0, 1).$$

Właściwe uogólnienie Obserwacji 12.1.5 wymaga zastąpienia warunku $V_1 \cap V_2 \cap V_3 = \{0\}$ układem warunków:

$$V_1 \cap (V_2 + V_3) = \{0\}, \quad V_2 \cap (V_1 + V_3) = \{0\}, \quad V_3 \cap (V_1 + V_2) = \{0\}.$$

Rozkłady na sumy proste mają wielkie znaczenie dla lepszego zrozumienia wykładu w drugim semestrze, choć nie są one, z uwagi na brak miejsca i godzin wykładowych, szerzej omówione w skrypcie, na którym się opieramy. Nietrudno się o tym przekonać próbując uogólnić przykład podany na zakończenie zasadniczej części wykładu. Gdy zapoznamy się (a zajmie nam to pozostałą część semestru) z językiem niezbędnym do badania przekształceń liniowych (zarówno macierzowym, jak i szkicowo – diagramowym), wówczas przyjdzie czas na badanie niezmienników przekształceń liniowych. Wiele z nich łatwiej będzie zrozumieć wiążąc z przekształceniami liniowymi rozkłady na sumy proste, związane z tzw. przestrzeniami niezmienniczymi. Póki co zajmiemy się jednak inną ważną strukturą związaną z podprzestrzeniami.

12.2 Wybór przykładowych pytań

Proszę się nie uczyć na pamięć odpowiedzi czy wyjaśnienie! Takich pytań można ułożyć znacznie więcej i służą one jedynie sprawdzeniu podstawowego zrozumienia treści zajęć. Jeśli odpowiedź na dowolne z poniższych pytań sprawia Państwu problem, proszę pomyśleć który fragment notatek dotyczy tego pytania, wrócić do ich lektury, a w razie dalszych trudności poprosić prowadzącego o wskazówkę.

1. Czy wektor $(1, 2, 1)$ należy do podprzestrzeni $\text{lin}((1, 1, 0)) + \text{lin}((0, 0, 1))$?
2. Czy wektor $(1, 2, 2)$ należy do podprzestrzeni $\text{lin}((1, 1, 1), (0, 1, 1)) \cap \text{lin}((1, 0, 1), (0, 1, 0))$?
3. Czy istnieje podprzestrzeń W przestrzeni liniowej \mathbb{R}^3 , że $\text{lin}((1, 1, 0)) + W = \text{lin}((1, 1, 1))$?
4. Czy istnieje podprzestrzeń W przestrzeni liniowej \mathbb{R}^3 , że $\text{lin}((1, 1, 1)) + W = \text{lin}((1, 0, 1), (0, 1, 1))$?
5. Wyznacz $\text{lin}((1, 2, 1)) \cap \text{lin}((1, 0, 0), (0, 1, 1))$.
6. Czy $\mathbb{R}^2 = \text{lin}((1, 0)) \oplus \text{lin}((1, 1))$?
7. Czy $\mathbb{K}[x] = \text{lin}(1) \oplus \text{lin}(x, x^2, x^3, \dots)$?
8. Czy $\mathbb{K}[x] = \{\text{wielomiany stałe}\} \oplus \{\text{wielomiany spełniające } f(0) = 0\}$?
9. Niech V będzie przestrzenią liniową nad \mathbb{Z}_2 zaś U jej podprzestrzenią. Czy $U + U$ jest przestrzenią zerową?
10. Czy przestrzeń liniowa $V_1 + V_2$ zawiera podprzestrzeń $V_1 \cap V_2$? Czy zawiera zbiór $V_1 \cup V_2$?
11. Niech $V = \text{lin}(v_1, v_2)$, $W = \text{lin}(w_1, w_2)$. Wiadomo, że v_1, v_2, w_1, w_2 jest bazą przestrzeni liniowej $V + W$. Czy $V + W = V \oplus W$?
12. Niech U, W będą podprzestrzeniami przestrzeni liniowej V . Wiadomo, że dla pewnych $u_1, u_2 \in U$ oraz $w_1, w_2 \in W$ zachodzi $u_1 + w_1 = u_2 + w_2$. Czy $U \cap W$ może być przestrzenią zerową?
13. Niech U, W, Z będą podprzestrzeniami przestrzeni liniowej V oraz niech $V = W \oplus Z$. Uzasadnij, że
$$U = (U \cap W) \oplus (U \cap Z).$$
14. Założmy, że $\dim(U \cap W) = \dim(W)$. Uzasadnij, że $W \subseteq U$.
15. Założmy, że $\dim(U + W) = \dim(W)$. Uzasadnij, że $U \subseteq W$.
16. Założmy, że U jest podprzestrzenią wymiaru 3 w 5-wymiarowej przestrzeni liniowej V . Niech W będzie podprzestrzenią V . Ile wynosić może wymiar $U \cap W$? Ile może wynosić wymiar $U + W$?
17. Założmy, że U, W są podprzestrzeniami zawartymi w 3-wymiarowej przestrzeni liniowej V oraz $6 \leq \dim U + \dim W$. Uzasadnij, że $V = U + W$.
18. V_1, V_2 są podprzestrzeniami wymiaru $n - 1$ przestrzeni n wymiarowej V . Czy przecięcie $V_1 \cap V_2$ może być wymiaru $n - 1$? Czy może być wymiaru $n - 3$?
19. V_1, V_2 są podprzestrzeniami przestrzeni n wymiarowej V , $\dim V_1 = n - 1$ oraz $V_2 \not\subseteq V_1$. Uzasadnij, że $\dim(V_1 \cap V_2) = \dim(V_2) - 1$.
20. Wiadomo, że $V_1 \subset \mathbb{K}^n$ jest opisana jednym niezerowym równaniem, $V_2 \subset \mathbb{K}^n$ jest opisana jednym niezerowym równaniem oraz $V_1 \cap V_2$ jest opisana jednym równaniem. Czy $V_1 = V_2$?
21. Niech U, W_1, W_2 będą podprzestrzeniami przestrzeni liniowej V . Wiadomo, że $U + W_1 = U + W_2$. Czy $W_1 = W_2$? Czy $\dim W_1 = \dim W_2$?
22. Niech U, W_1, W_2 będą podprzestrzeniami przestrzeni liniowej V . Wiadomo, że $U \oplus W_1 = U \oplus W_2$. Czy $W_1 = W_2$? Czy $\dim W_1 = \dim W_2$?
23. Podprzestrzenie V_1, V_2 przestrzeni liniowej V spełniają $\dim(V_1 + V_2) = 2$ oraz $\dim V_1 \cap V_2 = 1$. Czy $V_1 \cup V_2$ jest podprzestrzenią liniową V ?
24. Podprzestrzenie V_1, V_2, V_3 przestrzeni liniowej V spełniają $U_1 \subseteq U_2$, $U_2 \cap U_3 = \{0\}$, oraz $V = U_1 \oplus U_3$. Uzasadnij, że $U_1 = U_2$.

12.3 Zadania do samodzielnej pracy

Zadania ilustrujące umiejętności niezbędne do zaliczenia przedmiotu oznaczone są symbolem ♠.

Przy tych zadaniach dodany jest opis umiejętności, które sprawdzają.

1. (♠ Znajdowanie baz i wymiarów sum i przecięć podprzestrzeni). Niech V_1 i v_2 będą następującymi podprzestrzeniami przestrzeni \mathbb{R}^n . Znajdź bazy i wymiary przestrzeni $V_1 + V_2$ oraz $V_1 \cap V_2$.

- (a) $V_1 = \text{lin}((2, 1, 3, 4), (3, 9, 3, 9), (-1, 7, -3, 1))$, $V_2 = \text{lin}((1, -3, 3, 0), (2, 5, 3, 5), (1, 8, 0, 5))$,
(b) $V_1 = \text{lin}((3, 2, 1, 0), (4, 3, 0, 2), (1, 2, 2, -3))$, zaś V_2 jest opisana układem równań:

$$\begin{cases} x_1 + 2x_2 - x_3 + x_4 = 0 \\ 3x_1 + 5x_2 + x_3 - 5x_4 = 0 \end{cases}.$$

- (c) V_1, V_2 są opisane układami równań liniowych, odpowiednio U_1, U_2 :

$$U_1 : \begin{cases} 2x_1 + x_2 - x_3 + 4x_4 = 0 \\ 3x_1 - x_2 + 2x_3 + x_4 = 0 \end{cases}, \quad U_2 : \begin{cases} -x_1 + 2x_2 - 5x_3 + 3x_4 = 0 \\ 2x_1 - 4x_2 + 10x_3 - 6x_4 = 0 \end{cases}.$$

2. (♠ Rozstrzyganie kiedy suma podprzestrzeni jest prosta).

- (a) Niech $V_1 \subset \mathbb{R}^3$ będzie podprzestrzeniami opisanymi równaniami $x_1 + 2x_2 - x_3 = 0$ i niech $V_2 = \text{lin}((2, -t + 2, 4), (2s, 6, -8)) \subset \mathbb{R}^3$. Dla jakich wartości parametrów $s, t \in \mathbb{R}$ zachodzi $\mathbb{R}^3 = V_1 \oplus V_2$?
(b) Niech V_1, V_2 będą podprzestrzeniami przestrzeni \mathbb{R}^4 , przy czym $V_1 = \text{lin}((1, 1, 1, 2), (2, 0, 1, 3), (0, 2, 1, 1))$ oraz V_2 jest opisana układem równań $x_1 + x_2 - x_3 = 0, x_2 + tx_4 = 0$. Znajdź wszystkie takie $t \in \mathbb{R}$, że $\mathbb{R}^4 = V_1 \oplus V_2$.
(c) Niech $A = \text{lin}(-2, 1, 0, -3), (2, -1, 1, 3)$. Znajdź takie podprzestrzenie A i B przestrzeni \mathbb{R}^4 , by \mathbb{R}^4 było sumą prostą A i V , a nie było sumą prostą podprzestrzeni B i V ani A i B .
(d) Niech $A = \text{lin}((1, 2, 3, 4), (4, 3, 2, 1), (2, 3, 4, 5)) \subseteq \mathbb{R}^4$. Znajdź takie podprzestrzenie $B, C \subseteq \mathbb{R}^4$, że $\mathbb{R}^4 = A \oplus B = B \oplus C = C \oplus A$ lub wykaż, że takie podprzestrzenie nie istnieją.

3. Dla podprzestrzeni V_1 przestrzeni V znaleźć taką podprzestrzeń V_2 , aby $V_1 \oplus V_2 = V$, jeśli

- (a) $V = M_{n \times n}(K)$, $V_1 = \{[a_{ij}] \in M_{n \times n}(K) : a_{ij} = 0 \text{ dla } i > j\}$.
(b) V — ciągi zbieżne o wyrazach w \mathbb{R} , V_1 — ciągi stałe.
(c) $V = F(\mathbb{R}, \mathbb{R})$, $V_1 = \{f \in V : f(0) = f(1) = 0\}$.

4. (♠ Wyciąganie prostych wniosków z formuły Grassmanna).

- (a) Czy istnieje przestrzeń liniowa V wymiaru 7 zawierająca podprzestrzenie W_1, W_2 takie, że $\dim W_1 = 4$, $\dim W_2 = 5$, $\dim(W_1 \cap W_2) = 1$?
(b) Niech W_1 i W_2 będą podprzestrzeniami liniowymi przestrzeni liniowej V oraz $\dim V = 5$, $\dim W_1 = \dim W_2 = 4$. Czy wymiar przestrzeni $W_1 \cap W_2$ może być równy 2?
(c) Czy istnieją podprzestrzenie V_1 i V_2 przestrzeni \mathbb{R}^7 , że $\dim(V_1 \cap V_2) = 2$ i $\dim V_1 = \dim V_2 = 5$?
(d) Niech $V_1, V_2 \subset \mathbb{R}^6$ będą podprzestrzeniami wymiaru 5. Czy możliwe jest, aby $\dim(V_1 \cap V_2) = 1$?
(e) W przestrzeni \mathbb{R}^{11} dane są podprzestrzenie V, W , przy czym $\dim V = 6$ oraz $\dim W = 8$. Czy przestrzeń $V \cap W$ może mieć wymiar 5?
(f) Dane są podprzestrzenie liniowe $V \subseteq W \subseteq \mathbb{R}^6$, przy czym $\dim V = 5$ oraz $W \neq \mathbb{R}^6$. Czy wynika z tego, że $V = W$?

5. Założmy, że $U, W \neq \{0\}$ są podprzestrzeniami przestrzeni liniowej V . Przypuśćmy, że istnieje taka funkcja $f : V \rightarrow \mathbb{R}$, że $f(u) < f(w)$, dla dowolnych niezerowych wektorów $u \in U$ oraz $w \in W$. Uzasadnij, że $\dim U + \dim W \leq \dim V$.

6. Założmy, że U, W skończenie wymiarowymi podprzestrzeniami przestrzeni liniowej V . Pokaż, że:

- gdy $\dim U \leq \dim W$ oraz $\dim(U + W) = \dim(U \cap W) + 1$, to $U \subseteq W$.
- gdy $\dim U < \dim W$ oraz $\dim(U + W) = \dim(U \cap W) + 2$, to $U \subseteq W$.

7. Niech V_1, V_2 będą n wymiarowymi podprzestrzeniami skończenie wymiarowej przestrzeni liniowej V . Założmy, że układ $\{\alpha_1, \alpha_2, \dots, \alpha_n\}$ jest bazą V_1 oraz układ $\{\beta_1, \beta_2, \dots, \beta_n\}$ jest bazą V_2 . Wykaż, że układ $\{\alpha_1, \alpha_2, \dots, \alpha_n, \alpha_1 + \beta_1, \alpha_2 + \beta_2, \dots, \alpha_n + \beta_n\}$ jest bazą $V_1 + V_2$ wtedy i tylko wtedy, gdy $V_1 \cap V_2 = \{0\}$. Wykaż, że istnieje podprzestrzeń W przestrzeni V taka, że $V_1 \oplus W = V_2 \oplus W = V$.

12.4 Dodatek. Jednoznaczność daje wyniki o nieistnieniu

Na wykładzie po raz kolejny pojawiło się pojęcie, które w swojej naturze zawiera koncepcję jednoznaczności przedstawienia pewnego elementu za pomocą innych. Przestrzeń liniowa V jest sumą prostą podprzestrzeni V_1 oraz V_2 , jeśli każdy wektor v z przestrzeni V można zapisać jednoznacznie właśnie za pomocą sumy wektorów $v_1 \in V_1$ oraz $v_2 \in V_2$. O sile tej jednoznaczności przekonamy się wielokrotnie, nie tylko w kontekście sumy prostej dwóch, ale i większej liczby podprzestrzeni. Jaka koncepcja matematyczna stoi za tym pojęciem? Była ona już wspominana — jednoznaczność oznacza, że pewne konfiguracje obiektów matematycznych nie są możliwe do uzyskania. Tego typu argument ma centralne znaczenie w matematyce.

Z jednoznaczności rozkładu liczby całkowitej na czynniki pierwsze wynika, że nie istnieje para liczb całkowitych m, n , spełniających równość $2m^2 = n^2$. Taka para nie istnieje, ponieważ liczba całkowita ma jeden rozkład na czynniki pierwsze. Równość $2m^2 = n^2$ oznacza, że liczba 2 wchodzi do rozkładu liczby całkowitej $2m^2$ nieparzyste wiele razy, a do rozkładu liczby n^2 — parzyste wiele razy. Jednoznaczność tego zabrania. Po obydwu stronach stać muszą liczby, które dzielą się przez tą samą potęgę liczby 2.

Tego typu dowodów można formułować więcej. Wspominaliśmy o jednoznaczności rozkładu na czynniki wielomianów o współczynnikach w ciele. I ponownie można z tego wyciągać rozmaite wnioski — choćby taki (stosunkowo przyziemny), że funkcja rzeczywista \sqrt{x} nie jest wymierna — nie można jej przedstawić jako ilorazu rzeczywistych funkcji wielomianowych. Czy Czytelnik umie to ściśle uzasadnić? Dlaczego nie zakładamy, że ciało jest dowolne, skoro twierdzenie o jednoznaczności jest prawdziwe nad dowolnym ciałem? Utożsamienie wielomianów z funkcjami wielomianowymi nie jest prawdziwe dla ciał skończonych. Funkcja $\sqrt{x} : \mathbb{Z}_2 \rightarrow \mathbb{Z}_2$ jest identycznością, czyli funkcją wielomianową $\sqrt{x} = x$.

W tym dodatku opowiemy o bardzo słynnym zagadnieniu algebraicznym, którego elementarne rozwiążanie dostarcza pojęcie sumy prostej. Chodzi o rozstrzygnięcie, czy w przestrzeni \mathbb{R}^3 można wprowadzić strukturę mnożenia wektorów w taki sposób, aby uzyskać strukturę ciała nad \mathbb{R} , podobnie jak liczby zespolone \mathbb{C} traktować można jako przestrzeń \mathbb{R}^2 z odpowiednio zdefiniowanym dodawaniem i mnożeniem. Pytanie to miało duże znaczenie w wieku XIX-tym, zwłaszcza dla Rowana Hamiltona, który to właśnie dostrzegł strukturę liczb zespolonych jako par liczb rzeczywistych z określonym działaniem.

O co chodzi? Dobre opisuje to tekst prof. Zbigniewa Marciniaka w Delcie pt. *Dlaczego w przestrzeni trójwymiarowej nie ma przywoitego mnożenia?* (https://www.deltami.edu.pl/media/articles/1996/04/delta-1996-04-dlaczego-w-przestrzeni-trojwymiarowej-nie-ma-przywoitego_yKYRwmA.pdf). Profesor w różnych miejscach w Delcie podaje dwa różne dowody. Ja chciałbym jednak przywołać taki, który wydaje mi się jeszcze bardziej naturalny i korzysta z pojęcia sumy prostej.

Przypomnijmy o co chodzi. O każdej liczbie zespolonej $a + bi \in \mathbb{C}$ możemy myśleć jako o parze liczb rzeczywistych (a, b) z działaniami dodawania i mnożenia:

$$(a, b) + (c, d) = (a + c, b + d), \quad (a, b) \cdot (c, d) = (ac - bd, ad + bc).$$

Dodawanie nietrudno uogólnić na przestrzeń trójwymiarową nad \mathbb{R} , czyli:

$$(a_1, a_2, a_3) + (b_1, b_2, b_3) = (a_1 + b_1, a_2 + b_2, a_3 + b_3).$$

Pytanie brzmi: czy \mathbb{R}^3 można wyposażyć w działanie mnożenia, które z tej przestrzeni robiłoby ciało? Może nawet bylibyśmy gotowi zrezygnować z przemienności mnożenia, jak w przypadku kwaternionów, o których można myśleć jak o elementach \mathbb{R}^4 z odpowiednim dodawaniem i mnożeniem. Czytelnik zechce sprawdzić, że zwykłe mnożenie po współrzędnych nie jest dobre, ponieważ w żadnym ciele iloczyn elementów niezerowych nie może być zerowy, a tu mielibyśmy $(1, 0, 0) \cdot (0, 1, 0) = (0, 0, 0)$.

Zakładać będziemy, że D jest ciałem, które jest jednocześnie n -wymiarową przestrzenią liniową nad \mathbb{R} . Dla $\lambda \in \mathbb{R}$ oraz $1 \in D$ pisać będziemy po prostu $\lambda \cdot 1 = \lambda \in D$. Przytoczone rozumowanie pochodzi z książki Mateja Bresara *Introduction to Noncommutative Algebra*.

Uwaga 12.4.1

Dla każdego $s \in D$ istnieje $\lambda \in \mathbb{R}$, że $s^2 + \lambda s \in \mathbb{R}$.

Zanim zobaczymy dowód zauważmy, że stwierdzenie to oznacza, że każdy element D spełnia równanie kwadratowe o współczynnikach w \mathbb{R} .

Dowód. Zauważmy, że jeśli $\dim D = n$, to układ $n + 1$ elementów $1, s, \dots, s^n$ jest liniowo zależny. Oznacza to, że istnieje wielomian $f(x) \in \mathbb{R}[x]$ stopnia co najwyżej n , że $f(s) = 0$. Założymy, że wiodący współczynnik f równy jest 1. Wiemy z wykładu, że $f(x)$ rozkłada się na iloczyn czynników liniowych i kwadratowych w $\mathbb{R}[x]$:

$$f(x) = (x - \lambda_1) \dots (x - \lambda_r)(x^2 + a_1x + b_1) \dots (x^2 + a_sx + b_s),$$

gdzie $\lambda_i, a_i, b_i \in \mathbb{R}$. Skoro $f(s) = 0$, to

$$(x - \alpha_1) \dots (x - \alpha_r)(x^2 + a_1s + b_1) \dots (x^2 + a_ss + b_s) = 0.$$

Skoro D jest ciałem (nawet nieprzemiennym), to jeden z czynników musi być równy 0, a zatem s jest pierwiastkiem wielomianu kwadratowego o współczynnikach w \mathbb{R} . \square

Czytelnika może boleć, że nagle do gry wkracza wynika wymagający zasadniczego twierdzenia algebry. To prawda, ponieważ chcemy coś wykazać dla przestrzeni dowolnego wymiaru n . Gdyby ograniczyć się do $n = 3$, wówczas trzeba tylko wiedzieć, że każdy wielomian rzeczywisty stopnia 3 ma pierwiastek, a tu ZTA nie potrzeba. Przechodzimy do kluczowego argumentu.

Uwaga 12.4.2

Rozważmy podzbiór ciała D postaci:

$$V = \{v \in D : v^2 \in \mathbb{R}, v^2 \leq 0\}.$$

Wówczas V jest podprzestrzenią D (nad \mathbb{R}) oraz

$$D = \mathbb{R} \oplus V.$$

Ów rozkład będzie miał fundamentalne znaczenie dla dalszej klasyfikacji. Naszym celem będzie później pokazanie, że wymiar przestrzeni V wynosić może jedynie 0, 1 lub 3.

Dowód. Zacznijmy od następującej obserwacji. Jeśli $s \in D \setminus V$ spełnia jednocześnie $s^2 \in \mathbb{R}$, to $s^2 > 0$, a stąd $s^2 = \lambda^2$, dla pewnego $\lambda \in \mathbb{R}$. Stąd $(s - \lambda)(s + \lambda) = 0$, skąd $s = \pm\lambda \in \mathbb{R}$.

Jest jasne, że $\mathbb{R} \cap V = \{0\}$ oraz, że V jest zamknięty na mnożenie przez liczby rzeczywiste tzn. jeśli $s \in D$ oraz $\lambda \in \mathbb{R}$, wówczas $\lambda s \in V$. Sprawdzimy teraz, że jeśli $u, v \in V$, to $u + v \in V$. Istotnie, możemy założyć, że u, v są liniowo niezależne (inaczej to oczywiste). Twierdzimy, że wówczas układ $1, u, v$ również jest liniowo niezależny. Rzeczywiście, jeśli dla pewnych $a, b, c \in \mathbb{R}$ mamy $au = bv + c$, to podniesieniu obydwu stron do kwadratu dostajemy $b^2v^2 \in \mathbb{R}$, skąd $b = 0$ lub $c = 0$, skąd $a = b = c = 0$. Jak wykazać teraz, że $u + v \in V$? Skoro $u + v \in D$, to istnieją $a, b \in \mathbb{R}$, że:

$$(u + v)^2 + a(u + v) \in \mathbb{R}, \quad (u - v)^2 + b(u - v) \in \mathbb{R}.$$

Z drugiej strony:

$$(u + v)^2 + (u - v)^2 = 2u^2 + 2v^2 \in \mathbb{R}.$$

Porównując te równości, uzyskujemy $a(u + v) + b(u - v) \in \mathbb{R}$. Skoro jednak $u, v, 1$ są liniowo niezależne, to $a + b = a - b = 0$, skąd $a = b = 0$, więc $u + v \in V$. Stąd V jest podprzestrzenią D .

Pozostało pokazać, że $D = \mathbb{R} + V$. Weźmy $s \in D \setminus \mathbb{R}$. Zgodnie z poprzednią obserwacją mamy $s^2 + \lambda s \in \mathbb{R}$, dla pewnego $\lambda \in \mathbb{R}$. Ponownie używając argumentu z pierwszego akapitu mamy $s + \frac{\lambda}{2} \in V$. Stąd

$$s = -\frac{\lambda}{2} + (s + \frac{\lambda}{2}) \in \mathbb{R} + V.$$

\square

W tym momencie możemy do końca nie widzieć jakie zalety ma wydzielenie jednowymiarowego składnika prostego w naszym ciele D . Wnioskować chcemy, że jeśli D ma wymiar powyżej 2, to ma wymiar co najmniej 4 (a dokładniej – zawiera kwaterniony), a więc nie jest wymiaru 3. Oto dokładne sformułowanie.

Uwaga 12.4.3

Jeśli $\dim D > 2$, to istnieje liniowo niezależny układ wektorów $i, j, k \in V$, że

$$i^2 = j^2 = k^2 = -1, \quad ij = -ij = k, \quad ki = -ik = j, \quad jk = -kj = i.$$

Jak się okazuje, tu wkracza niebanalny pomysł. Rozważamy nowe działanie, dla dowolnych elementów $u, v \in V$ definiujemy

$$u \circ v = uv + vu.$$

Tu jest podkreślony potencjalnie nieprzemieniły charakter D . Zauważmy, że :

$$u \circ v = (u + v)^2 - u^2 - v^2 \in \mathbb{R}, \text{ a jeśli } v \neq 0, \text{ to } v \circ v = 2v^2 \neq 0.$$

Dowód. Uzasadnienie poprzedniego faktu wraz z formułą Grassmanna daje nam, że $\dim V = \dim D - 1 = n - 1 > 1$. Możemy zatem wybrać dwa liniowo niezależne wektory $u, v \in V$. Niech

$$u := w - \frac{w \circ v}{v \circ v} v.$$

Łatwo sprawdzić, że $u \neq 0$ oraz $u \circ v = 0$. Niech:

$$i := \frac{1}{\sqrt{-u^2}} u, \quad j := \frac{1}{\sqrt{-v^2}} v, \quad k = ij.$$

Bezpośredni rachunek pozwala sprawdzić, że zachodzą równości postulowane w tezie. Co więcej, dla dowolnych $a, b, c, d \in \mathbb{R}$ mamy:

$$(a + bi + cj + dk)(a - bi - cj - dk) = a^2 + b^2 + c^2 + d^2,$$

co oznacza, że $1, i, j, k$ są liniowo niezależne. \square

Zostawiłem Czytelnikowi sprawdzenie dokładnych rachunków, ale może wypada powiedzieć, że w drugim semestrze wektory typu $w - \frac{w \circ v}{v \circ v} v$ będą rzutami prostopadłymi wektora w na przestrzeń prostopadłą do $\text{lin}(v)$. W tym sensie zamiast \circ stać będzie iloczyn skalarny. Krótko mówiąc wykaźaliśmy, że jeśli dany jest układ liniowo niezależny, który zawiera wektory z V , to ma on co najmniej 3 elementy. To już nam pokazuje, że w $D = \mathbb{R}^3$ nie ma struktury ciała z kompatybilnym z dodawaniem po współrzędnych mnożeniu. Teraz wykażemy, że jeśli D ma mieć takie *dobre mnożenie* (niekoniecznie przemienne), to $\dim D = 0, 1, 3$. W istocie, D „jest” (z dokładnością do izomorfizmu algebr) jedną z algebr: $\mathbb{R}, \mathbb{C}, \mathbb{H}$.

Twierdzenie 12.4.4: Frobenius (1878)

Przestrzeń liniowa D skończonego wymiaru nad \mathbb{R} z mnożeniem, które spełnia wraz z dodawaniem wektorów aksjomaty ciała, ewentualnie poza przemiennością (tzw. *algebra z dzieleniem* nad \mathbb{R}), jest wymiaru 1, 2 lub 4.

Dowód. Niech $\dim D = n > 4$. Niech i, j, k będą elementami uzyskanymi w poprzednim twierdzeniu. Skoro $\dim V > n - 1 > 3$, to istnieje $v \notin \text{lin}(i, j, k)$. Rozważmy element:

$$e := v + \frac{i \circ v}{2} i + \frac{j \circ v}{2} j + \frac{k \circ v}{2} k.$$

Zauważmy, że jest to element niezerowy (bo inaczej v jest kombinacją liniową i, j, k) oraz mamy:

$$i \circ e = j \circ e = k \circ e = 0.$$

Z pierwszych dwóch równości mamy zatem $i \circ e = ie + ei = 0 = j \circ e = je + ej$, czyli:

$$ie = -ei, \quad je = -ej \Rightarrow -ie\cancel{j} = ei\cancel{j}, \quad \text{oraz} \quad je = -ej \Rightarrow \cancel{ije} = -\cancel{iej}.$$

Stąd $iej = -iej = ije$. Natomiast z równości $k \circ e = 0$ wynika, że $ke = -ek$, czyli wstawiając $ij = k$ mamy $ije = -eij$. To jednak oznacza, że $iej = 0$, co nie jest możliwe. Sprzeczność z założeniem, że $n > 4$. \square

Dowód był dość subtelny, ale wynik jest wysoce niebanalny. Widzimy jak istotne znaczenie miało wydzielenie składnika prostego V . Tego typu myślenie jest charakterystyczne dla zaawansowanej algebry – wydzielić „duży” fragment, który lepiej rozumiemy i w nim prowadzić właściwe rozumowanie.

12.5 Trivia. Krata podprzestrzeni przestrzeni liniowej

Omówione na wykładzie operacje sumy i części wspólnej podprzestrzeni przestrzeni V dają nam lepsze zrozumienie zależności pomiędzy podprzestrzeniami. Jest to jednak z konieczności spojrzenie „lokalne”. Brakuje nam bowiem języka do zrozumienia struktury kombinatorycznej zbioru wszystkich podprzestrzeni przestrzeni liniowej V . Na to wyzwanie odpowiada tzw. TEORIA KRAT, o której powiemy kilka słów.

Zacznijmy od następującego problemu. Niech V będzie skończenie wymiarową przestrzenią liniową nad ciałem K oraz niech U_1, \dots, U_n będzie układem podprzestrzeni przestrzeni V . Niech $\mathcal{L}(U_1, \dots, U_n)$ oznacza zbiór wszystkich podprzestrzeni, które można uzyskać startując z układu U_1, \dots, U_n i używając dowolnie wiele razy operacji sumy i przecięcia podprzestrzeni. Jak wygląda zbiór \mathcal{L} i jaką ma moc?

Oczywiście dla $n = 1$ oraz $n = 2$ moc \mathcal{L} nie przekracza odpowiednio 1 oraz 4. Rzeczywiście, suma lub iloczyn dowolnej podprzestrzeni z samą sobą jest jej równy. Jeśli zaś U_1, U_2 są podprzestrzeniami V , to:

$$\mathcal{L}(U_1, U_2) = \{U_1 \cap U_2, U_1, U_2, U_1 + U_2\}.$$

Dlaczego? Startując od U_1, U_2 i wykonując operację $+$ oraz \cap dostajemy:

$$U_1 + U_1 = U_1, \quad U_2 + U_2 = U_2, \quad U_1 + U_2, \quad U_1 \cap U_1 = U_1, \quad U_2 \cap U_2 = U_2, \quad U_1 \cap U_2.$$

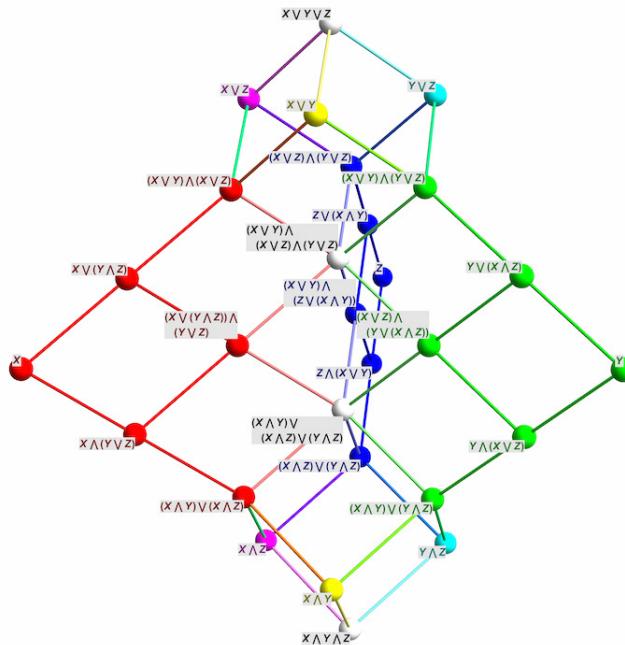
Niech $P = U_1 \cap U_2$ oraz $S = U_1 + U_2$. Ponawiamy stosowanie operacji $+$ oraz \cap dostając:

$$U_i + P = U_i, \quad U_i \cap P = P, \quad U_i + S = S, \quad U_i \cap S = U_i, \quad P + S = S, \quad P \cap S = P.$$

Oczywiście nietrudno podać przykład sytuacji, gdy czwórka U_1, U_2, P, S zawiera parami różne podprzestrzenie. Jak wygląda sytuacja, gdy startujemy od trzech podprzestrzeni? Problem ten należy do klasyki.

Twierdzenie 12.5.1: Dedekind, 1900

Dla przestrzeni liniowej V oraz trójki jej podprzestrzeni X, Y, Z zbiór $\mathcal{L}(X, Y, Z)$ może mieć nie więcej niż 28 elementów, które zaprezentować można na następującym diagramie



Rys. 1. Krata 28-elementowa generowana w sposób wolny (jako krata modularna) przez podprzestrzenie X, Y, Z .

Źródło wraz z ładną wizualizacją 3D pod adresem:

<https://blogs.ams.org/visualinsight/2016/01/01/free-modular-lattice-on-3-generators/>

Przykład sytuacji, gdy $|\mathcal{L}(X, Y, Z)| = 28$ ma miejsce na przykład w niezupełnie banalnej sytuacji, gdy

$$V = \mathbb{R}^8, \quad X = \text{lin}(\epsilon_2, \epsilon_4, \epsilon_5, \epsilon_8), \quad Y = \text{lin}(\epsilon_2, \epsilon_3, \epsilon_6, \epsilon_7), \quad Z = \text{lin}(\epsilon_1, \epsilon_4, \epsilon_6, \epsilon_7 + \epsilon_8).$$

Dedekind pokazał również, że dla $n \geq 4$ istnieje przestrzeń liniowa V oraz podprzestrzenie U_1, \dots, U_n takie, że $\mathcal{L}(U_1, \dots, U_n)$ jest zbiorem nieskończonym. Jeden z poważnych i trudnych problemów teorii krat mający odniesienie do współczesnej matematyki¹, polega na zrozumieniu „geometrycznej” struktury „generowanej” przez cztery podprzestrzenie² znajdujące się „w położeniu ogólnym”.

Spróbujmy objąć rysunek przedstawiony na poprzedniej stronie oraz przekonać Czytelnika, że problem podprzestrzeni ma pewne ogólne metody i wyniki niezwykle istotne w matematyce. Przede wszystkim widzimy na rysunku pewien graf, czy też diagram, którego wierzchołkami są podprzestrzenie. Jak rozumieć krawędzie? Otóż dwie podprzestrzenie $W_1, W_2 \in \mathcal{L}(X, Y, Z)$ połączone są krawędzią, jeśli $W_1 \subseteq W_2$ oraz nie istnieje element $W \in \mathcal{L}(X, Y, Z)$ taki, że $W_1 \subsetneq W \subsetneq W_2$. Innymi słowy, jeśli określmy porządek w zbiorze $\mathcal{L}(X, Y, Z)$ przez relację inkluzyji, to dwie podprzestrzenie połączone są krawędzią, to jedna położona jest w tym porządku „bezpośrednio nad” drugą. Takie diagramy (tzw. diagramy Hassego) można rysować dla każdego zbioru P z częściowym porządkiem \leqslant (parę (P, \leqslant) nazywamy też POSETEM).

Wreszcie, wyjaśnijmy znaczenie symboli \vee, \wedge , które oznaczają tzw. **operacje kratowe**.

Definicja 12.5.2: Krata

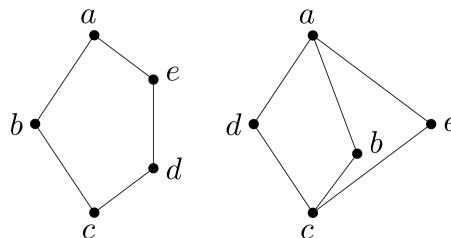
Niech A będzie niepustym zbiorem oraz \vee, \wedge będą działaniami dwuargumentowymi w A . Mówimy, że trójkąt (A, \vee, \wedge) jest KRATĄ, jeżeli dla dowolnych $x, y, z \in A$ następujące warunki:

1. $x \vee x = x, \quad x \wedge x = x,$
2. $(x \vee y) \vee z = x \vee (y \vee z), \quad (x \wedge y) \wedge z = x \wedge (y \wedge z),$
3. $x \vee y = y \vee x, \quad x \wedge y = y \wedge x,$
4. $(x \vee y) \wedge y = y, \quad (x \wedge y) \vee y = y.$

W każdej kracie spełniona jest równoważność $x \vee y = y \iff x \wedge y = x$. Relacja \leqslant , zdefiniowana na A za pomocą równoważności $x \leqslant y \iff x \vee y = y$, jest częściowym porządkiem, w którym każda para x, y ma ograniczenie górne $x \vee y$ oraz ograniczenie dolne $x \wedge y$. Z kratą związany jest więc porządek.

Przykłady:

- Jeśli (A, \vee, \wedge) jest kratą i $B \subseteq A$, to jeśli B jest zamknięty na operacje \vee, \wedge (tzn. dla dowolnych $x, y \in B$ mamy $x \vee y, x \wedge y \in B$), to B nazywamy podkratą kraty (A, \vee, \wedge) .
- Niech X będzie zbiorem oraz $P(X)$ – zbiorem jego podzbiorów. Wówczas zbiór $P(X)$ z działaniami \vee – sumy zbiorów oraz \wedge – części wspólnej zbiorów jest kratą.
- Niech V będzie przestrzenią liniową nad ciałem K , zaś $S(V)$ – zbiorem wszystkich podprzestrzeni przestrzeni V . Wówczas $S(V)$ z operacjami \vee sumy podprzestrzeni oraz \wedge – części wspólnej jest kratą. Jeśli U_1, \dots, U_n należą do $S(V)$, wówczas przez $\mathcal{L}(U_1, \dots, U_n)$ oznaczamy część wspólną wszystkich podkrat $S(V)$, zawierających U_1, \dots, U_n . Jest to, co łatwo pokazać, krata.
- Dwie niezwykle istotnymi przykładami krat są tak zwany PIECIOKĄT i DIAMENT, czyli kraty na zbiorze pięcioelementowym $A = \{a, b, c, d, e\}$ reprezentowane za pomocą następujących diagramów:



Rys. 2. Kraty pięcioelementowe. Źródło: Wikipedia.

¹Patrz artykuł Gian-Carlo Rota, *Ten Mathematics Problems I will never solve*, 1997: <https://www.degruyter.com/document/doi/10.1515/dmvm-1998-0215/html>

²Jest też tzw. problem czterech podprzestrzeni, badany m.in. przez Gelfanda, związany z tzw. teorią reprezentacji. Aby zrozumieć jak wiąże się z zagadniением Dedekinda, polecam: https://golem.ph.utexas.edu/category/2015/09/the_free_modular_lattice_on_3.html oraz <https://www.sciencedirect.com/science/article/pii/S0024379504002575>.

Pierwsza z nich, zwana pięciokątem lub kratą N_5 to krata, w której spełnione są relacje:

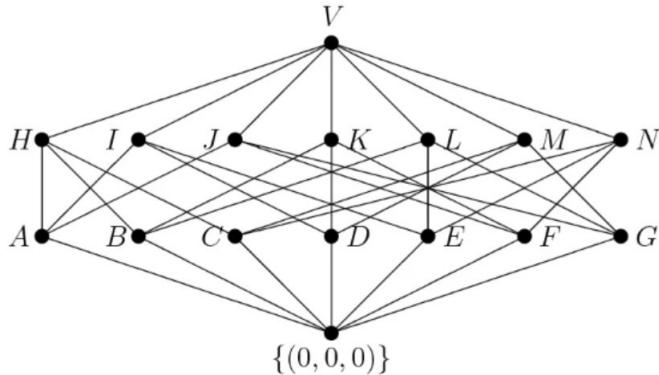
$$c \leq x \leq a, \text{ dla dowolnego } x, \quad d \wedge b = e \wedge b = c, \quad d \vee b = e \vee b = a$$

Diament lub krata M_3 to krata, w której spełnione są relacje:

$$c \leq a \leq x, \text{ dla dowolnego } x, \quad x \wedge y = c \text{ oraz } x \vee y = a, \text{ dla dowolnych } x \neq y \text{ w zbiorze } \{b.d.e\}.$$

Warto zauważyć, że $M_3 = S(\mathbb{Z}_2^2)$ jest kratą podprzestrzeni dwuwymiarowej przestrzeni liniowej nad ciałem skończonym \mathbb{Z}_2 . Zachęcam Czytelnika do narysowania kraty podprzestrzeni w \mathbb{Z}_3^2 .

- Krata podprzestrzeni przestrzeni \mathbb{Z}_2^3 jest bardziej skomplikowana i jej diagram ma postać:



Rys. 3. Krata podprzestrzeni przestrzeni trójwymiarowej $V = \mathbb{Z}_2^3$. Źródło:
<https://link.springer.com/article/10.1007/s00500-019-03866-y>

- Zbiór liczb całkowitych \mathbb{Z} z operacjami $\vee = NWW$ oraz $\wedge = NWD$ jest krata, której odpowiada porządek częściowy wyznaczony przez podzielność.

Gdy rozważamy nową abstrakcyjną strukturę zawsze zastanawiamy się nad tym czy można ją realizować za pomocą szczególnych typów struktur (poprzez tzw. reprezentacje). Na przykład: czy pięciokąt może być krata podprzestrzeni przestrzeni liniowej? Czy może być krata podzbiorów pewnego zbioru? Odpowiedzi na te pytania prowadzą do niezwykle istotnych własności algebraicznych w teorii grup, modułów czy algebr. Powiedzmy o dwóch najbardziej znanych, nawiązujących do arytmetyki.

Definicja 12.5.3: Krata rozdzielna i krata modularna

Powiemy, że A jest KRATĄ ROZDZIELNĄ (DYSTRYBUTYWNA), jeśli dla dowolnych $x, y, z \in A$ mamy:

$$(x \vee y) \wedge z = (x \wedge z) \vee (y \wedge z), \quad (x \wedge y) \vee z = (x \vee z) \wedge (y \vee z). \quad (\spadesuit)$$

Mówimy, że krata A jest MODULARNA, jeśli warunki rozdzielności (\spadesuit) zachodzą dla takich trójkę x, y, z , dla których zachodzi warunek $x \leq z$ (gdzie \leq jest porządkiem wyznaczonym przez kratę A).

Dla każdego zbioru X , krata $(P(X), \cup, \cap)$ jest krata rozdzielnią. Podkrata kraty rozdzielnej też jest zawsze rozdzielna. Co więcej, ważne twierdzenie Birkhoffa-Stone'a z 1934 roku mówi, że krata jest rozdzielna wtedy i tylko wtedy, gdy jest izomorficzna³ z pewną podkratą kraty $(P(X), \cup, \cap)$, dla pewnego zbioru X . Zauważmy, że krata podprzestrzeni przestrzeni liniowej wymiaru większego niż 1 nie jest nigdy rozdzielna. Aby to zrozumieć wystarczy popatrzeć na $S(V)$ dla $V = \mathbb{R}^2$ i rozważyć $x = \text{lin}((1, 0))$, $y = \text{lin}((0, 1))$, $z = \text{lin}((1, 1))$. Jak się natomiast okazuje, krata podprzestrzeni jest zawsze modularna.

Ciekawe, że pięciokąt i diament stanowią niezwykle istotne obiekty dla stwierdzania czy krata jest rozdzielna lub modularna. Żadna z tych dwóch krat nie jest, jak się okazuje, rozdzielna. Okazuje się, że krata jest rozdzielna wtedy i tylko wtedy, gdy żadna z jej podkrat nie zawiera ani diamentu, ani pięciokąta. Krata jest modularna wtedy i tylko wtedy, gdy nie ma podkraty zawierającej pięciokąt. Czytelnika zainteresowanego dowodami tych rezultatów oraz innymi ciekawostkami odsyłam choćby do artykułu dr Małgorzaty Jastrzębskiej *O pewnych kratach testowych* w czasopiśmie Delta (12/2013): <https://www.deltami.edu.pl/temat/matematyka/algebra/2013/12/31/kraty.pdf>.

³Nie mówimy tu czym jest izomorfizm krat – powiedzmy, że diagram dowolnej kraty rozdzielnej jest diagramem pewnej podkraty w kracie $(P(X), \cup, \cap)$, dla pewnego zbioru X . Szczegółowe znalezienie można w dowolnym wykładzie algebra uniwersalnej lub teorii krat, na przykład w https://math.uwb.edu.pl/~mariusz/share/classes/tk/teoria_krat-w.pdf.

Rozdział 13

Wstęp do przestrzeni ilorazowych

13.1 Wykład 13*

W rozmaitych rozważaniach dotyczących własności przestrzeni liniowej V przydatne jest rozważanie struktury ilorazowej przestrzeni liniowej, w której elementami są zbiory wektorów, których różnica należy do ustalonej podprzestrzeni. Zanim podamy definicję tej struktury, przyjrzymy się naturalnej motywacji dla jej wprowadzenia.

Definicja 13.1.1: Układ liniowo niezależny modulo podprzestrzeń

Niech W będzie podprzestrzenią przestrzeni liniowej V nad ciałem K . Powiemy, że układ wektorów $v_1, \dots, v_n \in V$ jest LINIOWO NIEZALEŻNY MODULO W , jeśli dla $a_1, \dots, a_n \in K$ mamy

$$a_1v_1 + \dots + a_nv_n \in W \Rightarrow a_1 = \dots = a_n = 0.$$

Powiemy, że układ liniowo niezależny modulo W jest BAZĄ V MODULO W , jeśli jest maksymalnym liniowo niezależnym układem modulo W .

Liniowa niezależność układu modulo podprzestrzeń zerowa jest tą samą liniową niezależnością, którą wprowadziliśmy w Definicji 8.1.1. Co więcej, jeśli $W' \subseteq W$ jest podprzestrzenią, to układ wektorów jest liniowo niezależny modulo W tylko wtedy, gdy jest liniowo niezależny modulo W' . W szczególności każdy układ liniowo niezależny modulo podprzestrzeń W jest układem liniowo niezależnym w przestrzeni liniowej V .

Przykład 1. Niech $V = \mathbb{R}^4$ oraz

$$W = \{(x_1, x_2, x_3, x_4) \in V \mid x_1 + 2x_4 = 0, x_1 + 2x_2 - 2x_3 + 4x_4 = 0\}.$$

Wówczas przykładem bazy przestrzeni V modulo W jest para wektorów

$$v_1 = (1, 0, 0, 0), \quad v_2 = (0, 1, 0, 0).$$

Istotnie, jeśli kombinacja liniowa $av_1 + bv_2 = (a, b, 0, 0)$ należy do W , to musi spełniać obydwa równania $x_1 + 2x_4 = 0$ oraz $x_1 + 2x_2 - 2x_3 + 4x_4 = 0$, czyli $a = 0$ oraz $a + 2b = 0$, co daje $a = b = 0$. A zatem $\{v_1, v_2\}$ to układ liniowo niezależny modulo W . Gdyby można go było rozszerzyć przy pomocy wektora v_3 , to warunek $a_1v_1 + a_2v_2 + a_3v_3 \in W$ oraz $v_1, v_2 \notin W$ implikowałby $v_3 \in W$, co jest sprzeczne z liniową niezależnością układu $\{v_1, v_2, v_3\}$ modulo W .

Zauważmy, że dla każdego niejednorodnego układu równań liniowych, któremu odpowiada układ jednorodny o zbiorze rozwiązań W , istnieją takie $a, b \in K$, że

$$V_{a,b} = (av_1 + bv_2) + W = \{av_1 + bv_2 + w \mid w \in W\}$$

jest zbiorem rozwiązań tego układu niejednorodnego. Zbiory $V_{1,0}$ oraz $V_{0,1}$ zawierają wektory bazy modulo W uzyskanej wyżej (i tylko one). Każdej kombinacji liniowej $av_1 + bv_2$ wektorów bazy v_1, v_2 przestrzeni V modulo W odpowiadają więc parami rozłączne rozłączne zbiory $V_{a,b}$, których sumą mnogościową jest V , i na których naturalne wydaje się wykonywanie „działan modulo W ” po współrzędnych w bazie modulo M . Konstrukcję zbiorów $V_{a,b}$ uogólnia poniższa uwaga.

Uwaga 13.1.2

Niech $W \subsetneq V$ i niech v_1, \dots, v_n będzie układem liniowo niezależnym modulo W . Wówczas następujące warunki są równoważne:

- (1) układ v_1, \dots, v_n jest bazą V modulo W ,
- (2) $\text{lin}(v_1, \dots, v_n) \oplus W = V$.

Dowód. Założymy (1) i niech w_1, \dots, w_m będzie bazą W (argument dla $\dim V = \infty$ jest podobny). Pokażemy, że układ $v_1, \dots, v_n, w_1, \dots, w_m$ jest bazą V . Weźmy $a_1, \dots, a_n, b_1, \dots, b_m \in K$ takie, że:

$$a_1v_1 + \dots + a_nv_n + b_1w_1 + \dots + b_mw_m = 0.$$

Wówczas $a_1v_1 + \dots + a_nv_n = -(b_1w_1 + \dots + b_mw_m) \in W$, a skoro v_1, \dots, v_n to układ liniowo niezależny modulo W , to $a_1 = \dots = a_n = 0$. Zatem w powyższej kombinacji mamy tylko $b_1w_1 + \dots + b_mw_m = 0$, co oznacza, że $b_1 = \dots = b_m = 0$, skoro w_1, \dots, w_m jest bazą W . A zatem układ $v_1, \dots, v_n, w_1, \dots, w_m$ jest liniowo niezależny. Pokażemy, że rozpinia on V . Weźmy $v \in V$. Twierdzimy, że istnieją c_1, \dots, c_n takie, że

$$v - c_1v_1 - \dots - c_nv_n \in W.$$

Układ v, v_1, \dots, v_n nie może być liniowo niezależny modulo W , więc dla pewnych $d, d_1, \dots, d_n \in K$, nie wszystkich równych zero, zachodzi $dv + d_1v_1 + \dots + d_nv_n \in W$. Mamy zatem dwie możliwości:

- albo $d \neq 0$ i wtedy $v - d_1v_1 - \dots - d_nv_n \in W$, zgodnie z tezą,
- albo $d = 0$ i dla pewnego i mamy $d_i \neq 0$, co oznacza, że układ v_1, \dots, v_n jest liniowo zależny modulo W , sprzeczność.

Pokazaliśmy, że $v_1, \dots, v_n, w_1, \dots, w_m$ jest bazą V . A zatem

$$W + \text{lin}(v_1, \dots, v_n) = V.$$

Oczywiście $\text{lin}(v_1, \dots, v_n) \cap \text{lin}(w_1, \dots, w_m) = 0$, bo jeśli jakiś wektor w należy do części wspólnej, to

$$w = e_1v_1 + \dots + e_nv_n = f_1w_1 + \dots + f_mw_m,$$

czyli $e_1v_1 + \dots + e_nv_n - f_1w_1 - \dots - f_mw_m = 0$, co wobec faktu, że $v_1, \dots, v_n, w_1, \dots, w_m$ jest bazą V oznacza, że $e_1 = \dots = e_n = f_1 = \dots = f_m = 0$.

Pokazaliśmy zatem, że $U \oplus \text{lin}(v_1, \dots, v_n) = V$. Dowód drugiej implikacji zostawiamy jako ćwiczenie. \square

Definicja 13.1.3: Warstwa podprzestrzeni

Niech W będzie podprzestrzenią przestrzeni V nad ciałem K i niech $\alpha \in V$. Zbiór

$$\alpha + W = \{\alpha + \gamma \mid \gamma \in W\}$$

nazywamy WARSTWĄ PODPRZESTRZENI W w przestrzeni V .

Widzimy zatem, że jeśli v_1, \dots, v_n jest bazą przestrzeni V modulo W , to V można rozbić na sumę rozłączną warstw $v + W$, gdzie $v \in \text{lin}(v_1, \dots, v_n)$. Fakt, że warstwy są rozłączne można również uzasadnić osobno, korzystając z poniższej ważnej obserwacji.

Uwaga 13.1.4

Niech W będzie podprzestrzenią przestrzeni V . Dla każdych $\alpha, \beta \in W$ mamy

$$\alpha + W = \beta + W \iff \alpha - \beta \in W$$

Dowód. Założymy, że $v + W = v' + W$. Skoro $v \in v + W$, to $v \in v' + W$. Stąd istnieje $w \in W$ takie, że $v = v' + w$. Stąd $v - v' \in W$. W drugą stronę, założymy, że $v - v' \in W$. Bez straty ogólności wystarczy pokazać, że $v \in v' + W$. Niech $w = v - v' \in W$. Wówczas $v = v' + w$, a stąd $v \in v' + W$, co pokazuje $v + W \subseteq v' + W$. Drugie zawieranie pokazujemy w sposób analogiczny. \square

Uwaga 13.1.5

Niech W będzie podprzestrzenią przestrzeni V .

- (i) Dla warstw $v + W$ oraz $v' + W$ określamy sumę warstw $\textcolor{red}{+}$ oraz iloczyn \cdot skalarza z ciała K przez warstwę:

$$(v + W) \textcolor{red}{+} (v' + W) = (v + v') + W, \quad a \cdot (v + W) = av + W.$$

Działania te są dobrze określone, tzn. jeśli $v + W = v' + W$, to dla każdego $v'' \in V$ oraz dla każdego $a \in K$ mamy: $(v + W) \textcolor{red}{+} (v'' + W) = (v' + W) \textcolor{red}{+} (v'' + W)$ oraz $a \cdot (v + W) = a \cdot (v' + W)$.

- (ii) zbiór $V/W = \{\alpha + W \mid \alpha \in V\}$ z działaniami dodawania i mnożenia przez skalar określonymi wyżej oraz z warstwą $0 + W$ tworzy przestrzeń liniową nad ciałem K .

Dowód. Pokażmy, że dodawanie warstw jest dobrze określonym działaniem. Warunek jest symetryczny, więc wystarczy pokazać jedno zawieranie. Niech

$$u \in (v + W) + (v'' + W) = (v + v'') + W.$$

Istnieje $w \in W$, że $u = (v + v'') + w$. Skoro $v + W = v' + W$, to mamy element $w' = v - v' \in V$. A zatem $v = v' + w'$. Stąd:

$$u = (v + v'') + w = ((v' + w') + v'') + w = (v' + v'') + (w' + w).$$

Rzeczywiście więc $u \in (v' + v'') + W$, skoro $w' + w \in W$. W rezultacie uzyskujemy

$$(v + v'') + W \subseteq (v' + v'') + W.$$

Dodawanie warstw jest dobrze określone. Analogicznie dowodzimy, że mnożenie warstwy przez skalar jest dobrze określone.

Sprawdzenie, że zbiór V/W spełnia aksjomaty przestrzeni liniowej sprowadza się dla większości aksjomatów do skorzystania z tego, że sama przestrzeń V jest liniowa. Sprawdzimy jedynie istnienie elementu zerowego i przeciwnego.

Twierdzimy, że $W = 0 + W$ jest zerem w V/W . Istotnie, niech $v + W \in V/W$. Wówczas:

$$(v + W) \textcolor{red}{+} W = (v + 0) + W = v + W, \quad W \textcolor{red}{+} (v + W) = (0 + v) + W = v + W,$$

co załatwia sprawę. Oczywiście biorąc $v + W \in V/W$ widzimy, że warstwą przeciwną jest $-v + W$. \square

Definicja 13.1.6: Przestrzeń ilorazowa

Przestrzeń V/W określoną w poprzedniej uwadze nazywamy PRZESTRZENIĄ ILORAZOWĄ przestrzeni V przez podprzestrzeń W .

Powyższa definicja przestrzeni ilorazowej nie wymaga wykorzystania pojęcia bazy V modulo W . Uzasadnimy teraz, że każda taka baza wyznacza jednoznacznie pewną bazę V/W (i odwrotnie), będąc z nią zarazem równoliczna.

Uwaga 13.1.7

Niech W będzie podprzestrzenią przestrzeni liniowej V . Jeśli $\mathcal{A} = \{v_1, \dots, v_n\}$ jest bazą W oraz $\mathcal{B} = \{y_1, \dots, y_m\}$ ma tę własność, że $\{y_1 + W, \dots, y_m + W\}$ to baza V/W , wówczas $\mathcal{A} \cap \mathcal{B} = \emptyset$ oraz $\mathcal{A} \cup \mathcal{B}$ jest bazą V . W szczególności, jeśli przestrzeń V jest skończenie wymiarowa, to przestrzeń V/W też jest skończenie wymiarowa i zachodzi równość

$$\dim V = \dim W + \dim(V/W).$$

Dowód. Weźmy element v z bazy $\mathcal{A} \subset W$. Wówczas mamy $v + W = W$, czyli jest to warstwa zerowa. Nie może ona należeć do żadnej bazy V/W , czyli $\mathcal{A} \cap \mathcal{B} = \emptyset$.

Pokażmy, że zbiór $\mathcal{A} \cup \mathcal{B}$ rozpina V . Niech $v \in V$. Zatem $v + W \in V/W$ i istnieją skalary t_1, \dots, t_m takie, że

$$v + W = t_1(y_1 + W) + \dots + t_m(y_m + W) = (t_1y_1 + \dots + t_my_m) + W.$$

Mamy więc $v - t_1y_1 + \dots + t_my_m \in W$, czyli istnieją takie elementy $s_1, \dots, s_n \in K$, że:

$$v - t_1y_1 + \dots + t_my_m = s_1v_1 + \dots + s_nv_n.$$

Widzimy zatem, że $v \in \text{lin}(\mathcal{A} \cup \mathcal{B})$.

Pokażmy wreszcie, że $\mathcal{A} \cup \mathcal{B}$ to zbiór liniowo niezależny. Założymy, że dla pewnych s_1, \dots, s_n oraz t_1, \dots, t_m z ciała K mamy

$$s_1v_1 + \dots + s_nv_n + t_1y_1 + \dots + t_my_m = 0.$$

Zatem

$$s_1v_1 + \dots + s_nv_n = -(t_1y_1 + \dots + t_my_m) \in W,$$

a skoro v_1, \dots, v_n to baza W , dostajemy $s_1 = \dots = s_n = 0$. A zatem mamy

$$t_1y_1 + \dots + t_my_m = 0.$$

To oznacza, że $t_1(y_1 + W) + \dots + t_m(y_m + W) = 0 + W$. Ale $y_1 + W, \dots, y_m + W$ to baza V/W , czyli $t_1 = \dots = t_m = 0$, co kończy dowód. \square

Wniosek 13.1.8

Niech W będzie podprzestrzenią przestrzeni liniowej V . Następujące warunki są równoważne:

- (i) Układ v_1, \dots, v_k jest bazą V modulo W ,
- (ii) układ $v_1 + W, \dots, v_k + W$ jest bazą V/W ,
- (iii) $V/W = \{v + W \mid v \in \text{lin}(v_1, \dots, v_k)\}$.

Zachęcam Czytelnika do pokazania w analogiczny sposób wariantu tezy postawionej w (iv): jeśli \mathcal{C} jest bazą V taką, że $\mathcal{A} \subseteq \mathcal{C}$ jest bazą W , to układ $\{v + W \mid v \in \mathcal{C} \setminus \mathcal{A}\}$ jest bazą V/W .

Definicja 13.1.9: Kowymiar

Niech $U \subseteq V$. Liczbę

$$\text{codim } U := \dim V/U$$

nazywamy KOWYMIAREM przestrzeni U .

Na koniec rozważymy przykład ilorazu przestrzeni nieskończonego wymiaru. Niech $X = \mathbb{R}[x]$ oraz

$$U = \{f \in X : f(0) = f(1) = 0\}, \quad V = \{f \in X : f(0) = f(1)\}.$$

Zarówno U , jak i V są nieskończonymi wymiarami. Istotnie, mają one postać:

$$U = \{x(x-1)f(x) : f \in K[x]\}, \quad V = \{a + x(x-1)f(x) : a \in \mathbb{R} \text{ oraz } f \in K[x]\}.$$

Układem liniowo niezależnym modulo U jest choćby $\{1, x\}$. Czy można go rozszerzyć? Jeśli tak, to wielomian rozszerzający ten układ musi być stopnia co najmniej drugiego (układ liniowo niezależny modulo U musi być liniowo niezależny). Jednak dla każdego wielomianu f stopnia większego niż 1 istnieje kombinacja liniowa $a_1 + a_2x + a_3f$ taka, że powstały wielomian jest podzielny przez $x(x-1)$. Oznacza to, że mamy

$$\dim X/U = 2, \quad \dim X/V = 1, \quad \dim V/U = 1.$$

13.2 Wybór przykładowych pytań

Proszę się nie uczyć na pamięć odpowiedzi czy wyjaśnienie! Takich pytań można ułożyć znacznie więcej i służą one jedynie sprawdzeniu podstawowego zrozumienia treści zajęć. Jeśli odpowiedź na dowolne z poniższych pytań sprawia Państwu problem, proszę pomyśleć który fragment notatek dotyczy tego pytania, wrócić do ich lektury, a w razie dalszych trudności poprosić prowadzącego o wskazówkę.

1. Czy wektor $(1, 3, 2)$ należy do warstwy $(1, 1, 1) + \text{lin}((1, 3, 2))$ w przestrzeni liniowej \mathbb{R}^3 ?
2. Czy wektor $(1, 2, 1)$ należy do warstwy $(1, 1, 1) + \text{lin}((1, 2, 1))$ w przestrzeni liniowej \mathbb{R}^3 ?
3. Czy wektor $(1, 3, 2)$ należy do warstwy $(1, 1, 1) + \text{lin}((1, 2, 1))$ w przestrzeni liniowej \mathbb{R}^3 ?
4. Czy wektor $(1, 3, 3)$ należy do warstwy $(1, 1, 1) + \text{lin}((1, 2, 1))$ w przestrzeni liniowej \mathbb{R}^3 ?
5. Czy warstwy wektorów $(1, 1, 1), (1, 2, 3)$ względem podprzestrzeni $\text{lin}((0, 1, 2)) \subseteq \mathbb{R}^3$ są równe?
6. Czy warstwy wektorów $(1, 1, 1), (1, 2, 3)$ względem podprzestrzeni $\text{lin}((1, 2, 3)) \subseteq \mathbb{R}^3$ są równe?
7. Ile elementów ma przestrzeń ilorazowa \mathbb{Z}_3^4/W , gdzie $W = \text{lin}(1, 1, 1, 1)$?
8. Czy w przestrzeni ilorazowej $\mathbb{R}^2/\text{lin}((1, 1))$ prawdziwa jest równość

$$((1, 1) + \text{lin}((1, 1))) + ((4, 1) + \text{lin}((1, 1))) = (6, 3) + \text{lin}((1, 1))?$$

9. Czy w przestrzeni ilorazowej $\mathbb{R}^2/\text{lin}((1, 1))$ prawdziwa jest równość

$$((1, 2) + \text{lin}((1, 1))) + ((4, 1) + \text{lin}((1, 1))) = (3, 0) + \text{lin}((1, 1))?$$

10. Niech W będzie podprzestrzenią przestrzeni $\mathbb{R}[x]$ złożoną z wielomianów podzielnych przez $x^2 + 1$. Czy w przestrzeni ilorazowej $\mathbb{R}[x]/W$ prawdziwa jest równość

$$(x^4 + W) + (x^2 + W) = 0 + W?$$

11. Niech W będzie podprzestrzenią przestrzeni $\mathbb{R}[x]$ złożoną z wielomianów podzielnych przez $x^2 + 1$. Czy w przestrzeni ilorazowej $\mathbb{R}[x]/W$ prawdziwa jest równość

$$(x + W) + (x^2 + W) = x - 1 + W?$$

12. Niech $W = \text{lin}((1, 1, 1, 1))$. Czy poniższy układ jest liniowo niezależny w \mathbb{R}^4/W ?

$$(1, 1, 0, 0) + W, \quad (0, 0, 1, 1) + W$$

13. Niech $W = \text{lin}(x^2)$. Czy poniższy układ jest liniowo niezależny w $\mathbb{R}[x]/W$?

$$x^2 + 1 + W, \quad x^4 + 1 + W, \quad x^6 + 1 + W, \quad x^8 + 1 + W.$$

14. Ile jest równy $\dim V/V$?

15. Załóżmy, że $\dim V/U = 0$. Czy $U = V$?

16. Policz wymiar oraz wskaż dowolną bazę przestrzeni liniowej $\mathbb{R}^4/\text{lin}((1, 1, 0, 0), (0, 0, 1, 1))$.

17. Policz wymiar oraz wskaż dowolną bazę przestrzeni liniowej $\mathbb{R}^3/\{(x_1, x_2, x_3) \in \mathbb{R}^3 : x_1 + x_2 + x_3 = 0\}$.

18. Policz wymiar oraz wskaż dowolną bazę przestrzeni liniowej

$$M_{2 \times 2}(\mathbb{R}) / \text{lin} \left(\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \right)$$

19. Niech $U \subseteq W$ będą podprzestrzeniami przestrzeni V . Czy V/W jest podprzestrzenią V/U ?

20. Niech U, W będą podprzestrzeniami przestrzeni V , przy czym $\dim U + \dim W = \dim V$. Czy

$$\dim V/(U \cap W) = \dim(U + W)?$$

13.3 Zadania do samodzielnej pracy

Zadania ilustrujące umiejętności niezbędne do zaliczenia przedmiotu oznaczone są symbolem ♠. Przy tych zadaniach dodany jest opis umiejętności, które sprawdzają.

- Niech $V = \mathbb{R}^4$ oraz

$$U = \{(x_1, x_2, x_3, x_4) \in V : x_1 + 2x_4 = 0, x_1 + 2x_2 - 2x_3 + 4x_4 = 0\}.$$

Wyznacz bazę i wymiar przestrzeni V/U .

- Niech $V = \mathbb{R}^5$ oraz

$$W = \text{lin}((1, 2, 3, 2, 1), (5, 1, 7, 10, 1), (-1, 2, 1, -2, 1)).$$

Wyznacz bazę i wymiar przestrzeni V/U .

- W przestrzeni liniowej $V = \mathbb{R}[x]$ rozpatrzmy podprzestrzeń

$$W = \{w \in V \mid \forall_{s \in \mathbb{R}} w(s) = w(-s)\}.$$

Wyznacz bazę przestrzeni W oraz bazę przestrzeni V/W . Wywnioskuj stąd, że

$$\dim V = \dim W = \dim V/W.$$

- W przestrzeni liniowej $V = \mathbb{Z}_2[x]$ rozważmy wielomian $w = x^2 + x + 1$. Uzasadnij, że podzbiór W przestrzeni V złożony z wielomianów podzielnych przez w jest podprzestrzenią. Wyznacz V/W . Wykaż, że przestrzeń ta ma strukturę ciała czteroelementowego.
- W przestrzeni liniowej $K[x]$ rozważmy podprzestrzeń W_f złożoną ze wszystkich wielomianów podzielnych przez ustalony niezerowy wielomian f . Uzasadnij, że $\dim K[x]/W_f = \deg f$.
- W przestrzeni liniowej $V = K^\infty$ rozważmy podprzestrzeń W złożoną z takich ciągów (x_n) , że $x_n = 0$, dla $n > N$, dla pewnego ustalonego N . Uzasadnij, że $\dim V/W = \infty$.
- W przestrzeni liniowej $V = M_{n \times n}(K)$ rozważamy podprzestrzeń S złożoną z macierzy A spełniających warunek $A = A^T$ (tzw. macierze symetryczne). Wyznacz $\dim V/S$.
- Niech V będzie przestrzenią liniową skończonego wymiaru, a U, W – jej podprzestrzeniami. Uzasadnij, że

$$\dim(U + W)/W = \dim U/(U \cap W).$$

- Niech U, W będą podprzestrzeniami przestrzeni liniowej V , takimi że

$$\dim V/W < \infty \quad \text{oraz} \quad \dim V/U < \infty.$$

Uzasadnij, że $\dim V/(U \cap W) < \infty$.

- Udowodnij implikację $(2) \Rightarrow (1)$ z Uwagi 13.1.5. Wykaż, że jeśli $U \oplus W = V$ oraz układ v_1, \dots, v_k jest bazą U , to układ $v_1 + W, \dots, v_k + W$ jest bazą V/W .
- Niech U będzie taką podprzestrzenią przestrzeni liniowej V , że V/U jest przestrzenią skończonego wymiaru.
 - Wykaż, że jeśli W jest skończenie wymiarową podprzestrzenią V spełniającą $V = U + W$, to $\dim W \geq \dim V/U$.
 - Wykaż, że istnieje skończenie wymiarowa podprzestrzeń W przestrzeni V , że $\dim W = \dim V/U$ oraz $V = U \oplus W$.
- Niech U będzie podprzestrzenią przestrzeni liniowej V .
 - Wykaż, że każda podprzestrzeń przestrzeni V/U jest postaci W/U , gdzie W jest podprzestrzenią V spełniającą warunek $U \subseteq W \subseteq V$.
 - Wykaż, że przyporządkowanie $W \mapsto W/U$ zadaje wzajemnie jednoznaczna (i zachowującą inkluzję) odpowiedniość pomiędzy zbiorem wszystkich podprzestrzeni przestrzeni V zawierających U oraz zbiorem wszystkich podprzestrzeni V/U .

Rozdział 14

Każda przestrzeń liniowa ma bazę

14.1 Wykład 14*

Celem tego wykładu jest przybliżenie Czytelnikowi zagadnienia istnienia bazy w dowolnej przestrzeni liniowej, niekoniecznie skończonego wymiaru. Uzasadnić zatem chcemy następujące twierdzenie.

Twierdzenie 14.1.1: Haussdorff, 1932

Każda przestrzeń liniowa V nad ciałem K ma bazę.

Podstawowa idea polega na konstrukcji bazy jako maksymalnego liniowo niezależnego zbioru, nawiązując w ten sposób do równoważności stwierdzonej przez nas dla baz przestrzeni skończonego wymiaru. Aby zrozumieć dobrze o co chodzi musimy powiedzieć kilka słów o porządkach i elementach maksymalnych. Dlaczego to podejście zadziała i jakie są wyzwania?

Klasycznym przykładem, który obrazuje złożoność rozważanego przez nas problemu jest ciało liczb rzeczywistych traktowane jako przestrzeń liniowa nad ciałem liczb wymiernych. Jest to oczywiście przestrzeń nieskończona wymiaru. Przykładowym nieskończonym układem liniowo niezależnym jest zbiór pierwiastków ze wszystkich liczb pierwszych $\{\sqrt{2}, \sqrt{3}, \sqrt{5}, \dots\}$. Nietrudno się o tym przekonać pokazując, że \sqrt{p} nie jest kombinacją liniową elementów \sqrt{q} , dla $q < p$, tzn. nie istnieją liczby wymierne a_q takie, że:

$$\sum_{q < p, q \in \mathbb{P}} a_q \cdot \sqrt{q} = \sqrt{p}.$$

Nietrudno się jednak przekonać, że wypisany zbiór liniowo niezależny nie jest bazą \mathbb{R} nad \mathbb{Q} . Żadna liczba postaci $\sqrt[p]{p}$ nie jest kombinacją liniową (o współczynnikach w \mathbb{Q}) pierwiastków z liczb pierwszych. A nawet gdyby rozważać zbiór pierwiastków dowolnego stopnia ze wszystkich liczb pierwszych – to również nie jest maksymalny zbiór liniowo niezależny – nie należy do niego choćby liczba π czy e . Widzimy więc, że wypisanie maksymalnego zbioru liniowo niezależnego „wprost” jest w zasadzie niemożliwe. W dodatku zobaczymy przykład nieprzeliczalnego podzbioru \mathbb{R} , który jest liniowo niezależny nad \mathbb{Q} i który też nie jest bazą. Co więcej, wydaje się, że możemy wystartować z rozłącznych nieskończonych zbiorów liniowo niezależnych i „nadbudowywać” na nich różne większe zbiory liniowo niezależne. Jak owe „nadbudowane” zbiory miałyby się do potencjalnej bazy? Potrzebne są pewne narzędzia, by doprecyzować te problemy.

Udowodnimy następujący rezultat, korzystając z rezultatów ze wstępów do matematyki.

Twierdzenie 14.1.2

Niech V będzie niezerową przestrzenią liniową i niech \mathcal{S} będzie zbiorem złożonym z liniowo niezależnych podzbiorów w V . Wówczas \mathcal{S} zawiera PODZBIÓR MAKSYMALNY ze względu na inklucję – to znaczy taki podzbiór $M \in \mathcal{S}$, że nie istnieje zbiór $N \in \mathcal{S}$, który zawiera M jako podzbiór właściwy. Zbiór M jest bazą przestrzeni V .

Dlaczego istnieje element maksymalny w \mathcal{S} ? Nie unikniemy w tym miejscu użycia rezultatu, znanego jako lemat Kuratowskiego-Zorna, stanowiącego zwieńczenie podstawowego kursu ze wstępów do matematyki.

Twierdzenie 14.1.3: Lemat Kuratowskiego-Zorna

Niech \mathcal{S} będzie niepustym zbiorem częściowo uporządkowanym. Jeśli każdy liniowo uporządkowany podzbiór (łańcuch) w \mathcal{S} ma ograniczenie górne w \mathcal{S} , wówczas \mathcal{S} zawiera element maksymalny.

Aby zrozumieć lemat Kuratowskiego-Zorna, potrzebujemy czterech pojęć: częściowego porządku w zbiorze, zbioru liniowo uporządkowanego, ograniczenia górnego oraz elementu maksymalnego.

Definicja 14.1.4: Relacja częściowego porządku

Mówimy, że RELACJA \leqslant na zbiorze \mathcal{S} (czyli podzbiór zbioru $\mathcal{S} \times \mathcal{S}$) jest CZEŚCIOWYM PORZĄDKIEM, jeśli spełnione są następujące warunki:

- zwrotność — dla każdego $s \in \mathcal{S}$ mamy $s \leqslant s$,
- antysymetryczność — dla każdych $s, s' \in \mathcal{S}$ jeśli $s \leqslant s'$ oraz $s' \leqslant s$, to $s = s'$,
- przechodniość — dla każdych $s, s', s'' \in \mathcal{S}$ jeśli $s \leqslant s'$ oraz $s' \leqslant s''$, to $s \leqslant s''$.

Mówimy, że \leqslant jest LINIOWYM PORZĄDKIEM, jeśli dla każdych $s, s' \in \mathcal{S}$ mamy $s \leqslant s'$ lub $s' \leqslant s$.

Jeśli \leqslant jest częściowym porządkiem w zbiorze \mathcal{S} , a po ograniczeniu do niepustego podzbioru $\mathcal{T} \subseteq \mathcal{S}$ jest ona porządkiem liniowym, wówczas podzbiór \mathcal{T} nazywamy ŁAŃCUCHEM w \mathcal{S} .

Zobaczmy trzy przykłady, z których trzeci jest kluczowy dla naszych rozważań.

- Porządek liniowy \leqslant w \mathbb{R} – czyli zwykła relacja nierówności pomiędzy liczbami rzeczywistymi jest porządkiem liniowym.
- W zbiorze dodatnich liczb całkowitych \mathbb{Z}_+ wprowadzamy relację częściowego porządku $a \leqslant b$ postaci $a | b$, czyli RELACJĘ PODZIELNOŚCI. Oczywiście jest to częściowy porządek, ale nie jest to porządek liniowy, bowiem (choćby) 2 nie dzieli 3, ani 3 nie dzieli 2. Dla każdej liczby pierwszej p wskazać możemy łańcuch $\{p^n \mid n \in \mathbb{N}_+\} = \{p, p^2, p^3, \dots\}$.
- Niech \mathcal{S} będzie podzbiorem zbioru wszystkich podzbiorów $P(X)$ niepustego zbioru X . Wprowadzamy RELACJĘ INKLUSJI w \mathcal{S} postaci $A \leqslant B$ wtedy i tylko wtedy, gdy $A \subseteq B$. Gdy X jest zbiorem co najmniej dwuelementowym oraz $\mathcal{S} = P(X)$, to relacja ta nie jest relacją liniowego porządku, bowiem podzbiory jednoelementowe (różne) są nieporównywalne.

Definicja 14.1.5: Ograniczenie górne

Niech \leqslant będzie częściowym porządkiem w zbiorze \mathcal{S} oraz niech \mathcal{T} będzie podzbiorem \mathcal{S} . Powiemy, że element $s \in \mathcal{S}$ jest OGRANICZENIEM GÓRNYM zbioru \mathcal{T} , jeśli dla każdego $t \in \mathcal{T}$ mamy $t \leqslant s$.

Oczywiście ograniczenie górne nie musi należeć do zbioru \mathcal{T} . W zbiorze \mathbb{R} z relacją liniowego porządku \leqslant element 1 jest ograniczeniem górnym zbioru $(0, 1)$, ale do niego nie należy.

Definicja 14.1.6: Element maksymalny i element największy

Niech \leqslant będzie relacją częściowego porządku w zbiorze \mathcal{S} . Powiemy, że element $x \in \mathcal{S}$ jest

- MAKSYMALNY, jeśli dla każdego $y \in \mathcal{S}$ takiego, że $y \geqslant x$ mamy $y = x$,
- NAJWIĘKSZY, jeśli dla każdego $y \in \mathcal{S}$ mamy $x \geqslant y$.

Kluczowe jest zauważenie, że element maksymalny w zbiorze \mathcal{S} nie musi być większy od każdego innego elementu tego zbioru (czyli największy), ale jedynie od elementów, z którymi można go porównać – to właśnie znaczy, że jest maksymalny.

Z naszej perspektywy kluczowa jest sytuacja, gdy rozważamy podzbiór \mathcal{S} zbioru podzbiorów przestrzeni liniowej V złożony z podzbiorów liniowo niezależnych. Dla przykładu, dla $V = \mathbb{R}^3$ zbiór:

$$B_1 = \{(1, 0, 0), (0, 1, 0), (0, 0, 1)\}$$

jest elementem maksymalnym w \mathcal{S} , to znaczy – nie istnieje układ liniowo niezależny B_2 , który zawiera B_1 oraz pewien element, który nie należy do B_1 . Wskazać można wiele innych elementów maksymalnych \mathcal{S} (czyli innych baz \mathbb{R}^3). A czym jest liniowo uporządkowany zbiór baz? Zróbmy krok do tyłu.

Aby zrozumieć jak działa Lemat Kuratowskiego-Zorna warto przyjrzeć się relacji częściowego porządku inkluzyji w zbiorze podzbiorów $P(X)$ zbioru niepustego X . Czym jest liniowo uporządkowany ciąg elementów $P(X)$? Jest to na przykład (czy to jedyna możliwość?) ciąg wstępujących podzbiorów A_1, A_2, A_3, \dots spełniający warunek:

$$A_1 \subseteq A_2 \subseteq A_3 \subseteq \dots$$

Czy umiemy wskazać w $P(X)$ ograniczenie górne podzbioru $\{A_1, A_2, A_3, \dots\}$? Tak, jest to nieskończona suma tych zbiorów

$$A = A_1 \cup A_2 \cup A_3 \cup \dots = \bigcup_{i=1}^{\infty} A_i.$$

Zauważmy, że zbiór A nie jest żadnym ze zbiorów A_i , a jednak jest zawsze podzbiorem $P(X)$. To jest prawdą niezależnie od tego, czy łańcuch jest „wstępujący” (czy może być inaczej?) i ponumerowany liczbami naturalnymi (a jeśli jest nieprzeliczalny?). Sumę rodziny zbiorów można zdefiniować niezależnie od tego jaki zbiór ją indeksuje. W tym przypadku lemat Kuratowskiego-Zorna można wysławić prościej.

Wniosek 14.1.7

Niech \mathcal{S} będzie niepustą rodziną podzbiorów zbioru X uporządkowaną przez inkluzyję. Jeśli dla każdego łańcucha $\{A_t\}_{t \in T}$ elementów \mathcal{S} wiadomo, że

$$\bigcup_{t \in T} A_t \in \mathcal{S},$$

to w \mathcal{S} istnieje element maksymalny.

Przejdzmy do pokazania, że każda przestrzeń liniowa ma bazę. Sprowadza się to do uzasadnienia poniższego rezultatu.

Twierdzenie 14.1.8

Niech V będzie niezerową przestrzenią liniową nad ciałem K . W zbiorze \mathcal{S} wszystkich podzbiorów liniowo niezależnych V istnieje element maksymalny.

Dowód. Zakładamy, że $V \neq 0$, więc $\mathcal{S} \neq 0$. Na mocy wniosku wystarczy pokazać, że dla dowolnego łańcucha $\{B_t\}_{t \in T}$ liniowo niezależnych podzbiorów B_t przestrzeni V również zbiór $B = \bigcup_t B_t$ jest liniowo niezależny.

Biorąc dowolny skończony układ elementów v_1, \dots, v_n zbioru B wiemy, że istnieją elementy B_{t_1}, \dots, B_{t_k} łańcucha B_t , że każdy v_i należy do jednego ze zbiorów B_{t_i} . Skoro zbiory te tworzą skończony łańcuch, to istnieje $N \in \{1, \dots, k\}$ takie, że $v_1, \dots, v_k \in B_{t_N}$. A zatem v_1, \dots, v_n jest podukładem układu liniowo niezależnego B_{t_N} — czyli jest to układ liniowo niezależny. A zatem B jest układem liniowo niezależnym. Zgodnie z Lematem Kuratowskiego-Zorna w \mathcal{S} istnieje element maksymalny. \square

W dowodzie skorzystaliśmy z tego, że jeśli $\{s_1, \dots, s_n\}$ jest skończonym łańcuchem, to istnieje s_i takie, że $s_j \leq s_i$, dla $j \neq i$. Dowód to prosta indukcja. Możemy przejść do uzasadnienia Twierdzenia 14.1.2.

Dowód twierdzenia o istnieniu bazy. Niech M będziemy maksymalnym elementem \mathcal{S} . Rozpatrując podprzestrzeń $W = \text{lin}(M)$ pokażemy, że $W = V$, co pokaże, że M jest bazą. Gdyby M nie rozpiniał V , wówczas $W \neq V$ i można wskazać wektor $v \in V$ taki, że $v \notin W$. W szczególności M jest podzbiorem

właściwym zbioru $M \cup \{v\}$. Pokażemy jednak, że $M \cup \{v\}$ jest układem liniowo niezależnym.

Aby pokazać, że $M \cup \{v\}$ jest liniowo niezależny, założymy przeciwnie, że dla pewnego skończonego podzbioru $\{v_1, \dots, v_k\}$ zbioru $M \cup \{v\}$ mamy

$$c_1v_1 + c_2v_2 + \dots + c_kv_k = 0,$$

przy czym $c_i \in K$ nie są wszystkie zerowe. Skoro elementy M tworzą układ liniowo niezależny, to także układ $\{v_1, \dots, v_k\}$ jest liniowo niezależny i koniecznie jednym z elementów v_i musi być v . Możemy więc przenumerować elementy układu v_1, \dots, v_k tak, by założyć $v_k = v$. Co więcej, musimy mieć także $c_k \neq 0$. Mamy też $k \geq 2$, bo inaczej $c_1v = 0$, co jest niemożliwe, bo $v \neq 0$ oraz $c_1 \neq 0$. A zatem mamy:

$$v = -\frac{c_1}{c_k}v_1 - \frac{c_2}{c_k}v_2 - \dots - \frac{c_{k-1}}{c_k}v_{k-1}.$$

A zatem $v \in \text{lin}(v_1, \dots, v_{k-1}) \subseteq W$. Ale zakładaliśmy, że $v \notin W$, więc $M \cup \{v\}$ okazuje się zbiorem liniowo niezależnym, co przeczy maksymalności M w zbiorze \mathcal{S} . A zatem $W = \text{lin}(M) = V$ i M jest bazą V .

Wniosek 14.1.9

Każdy podzbiór liniowo niezależny niezerowej przestrzeni liniowej V można rozszerzyć do bazy V . W szczególności, każda podprzestrzeń W przestrzeni V jest składnikiem prostym V .

Dowód. Rozważamy liniowo niezależny podzbiór L przestrzeni V . Wskażemy bazę V zawierającą L poprzez uzasadnienie istnienia maksymalnego podzbioru liniowo niezależnego zawierającego L .

Niech \mathcal{S} będzie zbiorem wszystkich liniowo niezależnych podzbiorów V zawierających L . Podobnie jak w dowodzie Twierdzenia 14.1.8 uzasadniamy, że zbiór ten ma element maksymalny ze względu na inklucję. Z lematu Kuratowskiego-Zorna istnieje zatem element maksymalny zbioru \mathcal{S} . Dalsza część dowodu jest identyczna, jak w uzasadnieniu Twierdzenia 14.1.2. \square

Kolejny wniosek dotyczy zbiorów rozpinających.

Wniosek 14.1.10

Każdy zbiór rozpinający niezerową przestrzeń liniową V zawiera bazę V .

Dowód. Niech L będzie zbiorem rozpinającym V . Rozważmy zbiór liniowo niezależnych podzbiorów zbioru L . Jest on niepusty, gdyż zbiór L jest niepusty, a każdy pojedynczy wektor z L tworzy układ liniowo niezależny.

Podobnie jak w poprzednich rozumowaniach wykazujemy, że zbiór ten spełnia warunki lematu Kuratowskiego-Zorna, skąd istnieje maksymalny podzbiór liniowo niezależny B zawarty w L . Zbiór B jest oczywiście szukaną bazą V .

Istotnie, skoro L rozpinia V , wystarczy wykazać, że każdy element L jest w $\text{lin}(B)$. Gdyby jednak pewien element $v \in L$ nie był w $\text{lin}(B)$, wtedy $B \cup \{v\}$ byłby liniowo niezależnym podzbiorem L , który ściśle zawiera B , co przeczyłoby maksymalności tego zbioru. \square

Warto dodać drobną uwagę do ostatniego dowodu. Mogłoby się wydawać, że do znalezienia bazy wewnętrz układowu rozpinającego wystarczy znaleźć minimalny zbiór rozpinający, zamiast konstruować w nim maksymalny podzbiór liniowo niezależny. Nie jest to jednak skuteczna droga!

Wydaje się, że dowód istnienia minimalnego rozpinającego mógłby przebiegać w następujący sposób: bierzemy wszystkie podzbiory rozpinające zbioru rozpinającego L i porządkujemy przez tzw. odwrotną inklucję tzn. dla dwóch podzbiorów rozpinających A, B zbioru L mamy $A \leq B$ wtedy i tylko wtedy, gdy $A \supseteq B$. To jest oczywiście porządek częściowy, a przecięcie dowolnej rodziny podzbiorów wydaje się być ich naturalnym ograniczeniem górnym, przy takim porządku. Czy więc nie można użyć lematu Kuratowskiego-Zorna by uznać, że taki minimalny ze względu na inklikcję podzbiór jest bazą V ? Niestety ograniczenie górnego w postaci przecięcia może nie istnieć! Zachęcam Czytelnika do wskazania przykładu takiej sytuacji.

14.2 Dodatek. Kilka zadań ilustrujących użycie Lematu K-Z

Zadanie 14.1. Niech \mathbb{R}_+ będzie zbiorem liczb rzeczywistych dodatnich. Zbiór ten można rozbić na sumę dwóch rozłącznych, niepustych podzbiorów, z których każdy jest zamknięty na dodawanie. (*) Pokazać, że podobna własność ma zbiór wszystkich liczb niewymiernych.

Rozwiązańe. Wprowadzamy porządek częściowy na zbiorze par (A, B) , gdzie A, B są rozłącznymi podzbiorami \mathbb{R}_+ , każdy domknięty ze względu na dodawanie, tzn. jeśli $a_1, a_2 \in A$, to $a_1 + a_2 \in A$, podobnie dla zbioru B . Nazwijmy ten zbiór par rozłącznych podzbiorów \mathbb{R}_+ przez \mathcal{S} . Jest to zbiór niepusty. Możemy wybrać na przykład $A = \mathbb{N} \setminus \{0\}$ oraz $B = \{n \cdot \pi \mid n \in \mathbb{N} \setminus \{0\}\}$. W tym przypadku $A \cup B \neq \mathbb{R}_+$.

Powiemy, że $(A, B) \leqslant (A', B')$, dla pewnych par $(A, B), (A', B') \in \mathcal{S}$, jeśli $A \subseteq A'$ oraz $B \subseteq B'$. Jest to oczywiście porządek częściowy w \mathcal{S} . Chcemy użyć lematu Kuratowskiego-Zorna i wybrać w \mathcal{S} element maksymalny ze względu na relację \leqslant . Czego nam brakuje? Trzeba pokazać, że każdy łańcuch elementów w \mathcal{S} ma ograniczenie górne. Niech A_t, B_t będą, dla $t \in T$ i pewnego zbioru T , takimi podzbiorami \mathbb{R}_+ , że dla każdego t mamy $(A_t, B_t) \in \mathcal{S}$ oraz rodzinę $\{A_t\}_{t \in T}, \{B_t\}_{t \in T}$ są łańcuchami. Zauważmy, że zbiory

$$A = \bigcup_{t \in T} A_t, \quad B = \bigcup_{t \in T} B_t$$

są rozłączne. Jesli bowiem $x \in A \cap B$, to z definicji sumy zbiorów dla pewnych $s, t \in T$ mamy $x \in A_s \cap B_t$. Mamy jednak $A_t \subseteq A_s$ lub $A_s \subseteq A_t$ (bo $\{A_t\}$ to łańcuch). A zatem $x \in A_s \cap B_s$ lub $x \in A_t \cap B_t$, co jest niemożliwe, bo $(A_s, B_s), (A_t, B_t) \in \mathcal{S}$, czyli zbiory A_s oraz B_s są rozłączne. Podobnie $A_t \cap B_t = \emptyset$.

Czy zbiory A, B są zamknięte na dodawanie? Oczywiście tak. Dowodzimy to podobnie jak wyżej. Czy widzimy, że dla każdego $t \in T$ zachodzi warunek $(A_t, B_t) \leqslant (A, B)$? Skoro tak, to element (A, B) jest ograniczeniem górnym łańcucha $\{(A_t, B_t)\}_{t \in T}$. Spełnione są zatem założenia lematu Kuratowskiego-Zorna i w zbiorze \mathcal{S} istnieje element maksymalny (X, Y) .

Chcemy teraz pokazać, że $X \cup Y = \mathbb{R}^+$. Gdyby istniał element $r \in \mathbb{R}^+ \setminus (X \cup Y)$, to bierzemy $X \cup \{r\}$ i domykamy ze względu na dodawanie dostając X' . Innymi słowy – X' jest maksymalnym podzbiorem \mathbb{R}_+ zawierającym X oraz r i zamkniętym na dodawanie. Czy taki zbiór istnieje? Owszem – na mocy Lematu Kuratowskiego-Zorna (proszę to sprawdzić). Bierzemy też Y' jako domknięcie addytywne $Y \cup \{r\}$. Twierdzimy, że jeden ze zbiorów $X' \cap Y$ lub $X \cap Y'$ jest pusty. Gdyby było inaczej, to mielibyśmy elementy $x_0, x_1 \in X, y_0, y_1 \in Y, m, n \geq 1$ takie, że

$$X' \ni x_0 + nr = y_0 \in Y \quad \text{oraz} \quad Y' \ni y_1 + mr = x_1 \in X.$$

Musimy mieć $m, n \geq 1$, bo $X \cap Y = \emptyset$. Wtedy jednak

$$mx_0 - nx_1 = m(y_0 - nr) - n(x_1 - mr) = my_0 + ny_1 \in X \cap Y.$$

Ale $X \cap Y = \emptyset$, sprzeczność. Zatem jeden ze zbiorów $X' \cap Y$ lub $X \cap Y'$ jest pusty. Ale to oznacza, że para (X', Y) lub (X, Y') jest ściśle większa niż (X, Y) , sprzeczność z maksymalnością (X, Y) . Dowód jest zakończony. Co ciekawe wiedząc, że istnieje baza \mathbb{R} nad \mathbb{Q} można wskazać jednolinijkowe uzasadnienie.

W części gwiazdkowej: niech $B = \{b_i, i \in I\}$ będzie bazą Hamela \mathbb{R} nad \mathbb{Q} przy czym $b_j = 1$. Niech \prec będzie porządkiem na I takim, że j jest maksymalny. Jeśli $x \in \mathbb{R} \setminus \mathbb{Q}$ to piszemy $x = \lambda_0 b_{i_0} + \dots + \lambda_n b_{i_n}$, gdzie $i_0 \prec \dots \prec i_n$ i żaden z wymiernych współczynników λ_i nie jest zerem. Niech $x \in A$ wtedy i tylko wtedy, gdy $x \notin \mathbb{Q}$ oraz $\lambda_0 > 0$. Niech $B = (\mathbb{R} \setminus \mathbb{Q}) \setminus A$. Zbiory A, B są zamknięte na dodawanie, bo jesli λ_0 oraz λ'_0 są minimalnymi (w porządku \prec) współczynnikami x oraz y oraz są obydwa dodatnie/ujemne, to minimalny współczynnik $x + y$ to jedna z liczb λ_0, λ'_0 albo $\lambda_0 + \lambda'_0$.

Zadanie 14.2. Niech X będzie zbiorem nieskończonym. Pokazać, że istnieje bijekcja $f : X \rightarrow X$ o tej własności, że dla każdego $x \in X$ oraz każdego $n > 0$ mamy $f^n(x) \neq x$.

Szkic. Po pierwsze można zdefiniować zbiór częściowo uporządkowany P złożony z par (U, f) , gdzie $f : U \rightarrow U$ jest bijekcją oraz $U \subset X$ przy czym dla każdego $n > 0$ i dla każdego $x \in U$ mamy $f^n(x) \neq x$ oraz $(U, f) \leqslant (V, g)$ wtedy i tylko wtedy, gdy $U \subseteq V$ oraz u jest obcięciem v do U . Zbiór P jest niepusty, bo w X istnieje zawsze podzbiór przeliczalny $\{u_1, \dots\}$ i można wziąć $f(u_i) = u_{i+1}$ na tym podzbiiorze. Sprawdzamy łatwo, że założenia Lematu K-Z są spełnione.

Niech (U, f) będzie maksymalny. Jeśli $U \neq X$, to $X \setminus U$ jest skończony lub nieskończony. Jeśli jest skończony to biorę bijekcję $g : X \rightarrow U$ (istnieje) oraz rozważamy ciąg $X \xrightarrow{g} U \xrightarrow{f} U \xrightarrow{g^{-1}} X$. Wtedy $h = g^{-1}fg$ jest bijekcją taką, że $h^n(x) \neq x$. Jeśli $X \setminus U$ jest nieskończony to można znaleźć bijekcję g na zbiorze $V \subseteq X \setminus U$ równolicznym z \mathbb{Z} i potem określić h jako rozszerzenie f przez g . To przeczy maksymalności (U, f) w P .

Zadanie 14.3. Pokazać, że funkcję $f(x) = x$ (dla $x \in \mathbb{R}$) można przedstawić jak sumę dwóch funkcji okresowych. (*) Pokazać, że funkcję $g(x) = x^2$ można przedstawić jako sumę trzech funkcji okresowych.

Rozwiążanie. Niech a, b będą dwiema liczbami rzeczywistymi liniowo niezależnymi nad \mathbb{Q} . Wiadomo, że zbiór $\{a, b\}$ można dopełnić do bazy Hamela B . Każda liczba $x \in \mathbb{R}$ może być przedstawiona jako $\lambda a + \text{reszta} := f(x) + g(x)$. Jako, że pierwszy wyraz $x + b$ to $\lambda a = f(x)$ oraz pierwszy wyraz $x + a$ to jest $(\lambda + 1)a$, przy pozostałych wyrazach niezmienionych to widzimy, że $f(x)$ jest okresowa z okresem b , zaś $g(x)$ jest okresowa z okresem a .

Niech a, b, c będą trzema liczbami rzeczywistymi liniowo niezależnymi nad \mathbb{Q} . Podobnie jak wcześniej dopełniamy $\{a, b, c\}$ do bazy Hamela i rozważamy $x = \lambda_1 a + \lambda_2 b + \text{reszta} = f(x) + g(x) + h(x)$. Łatwo widzieć (jak wyżej), że $f(x)$ ma okres b oraz c , $g(x)$ ma okres a oraz c oraz $h(x)$ ma okres a oraz v . A więc x^2 można zapisać jako sumę dziewięciu składników postaci $f(x)g(x)$ itd. Każdy z tych dziewięciu składników ma okres a, b lub c (na przykład $f(x)h(x)$ ma okres b). Grupując te dziewięć wyrazów w trójki z okresami a, b, c dostajemy żądany rozkład.

Zadanie 14.4. Danych jest 17 liczb rzeczywistych o następującej własności: jeśli usuniemy z tego zbioru dowolny element, to pozostałe szesnaście można pogrupować w dwa ośmioelementowe podzbiory o równych sumach. Pokazać, że wszystkie te 17 liczb musi być równe.

Szkic. Jest jasne, że jeśli pewne $\{a_1, \dots, a_{17}\}$ spełniają warunki zadania, to spełniają je także liczby $\{a_1 - b, \dots, a_{17} - b\}$ oraz $\{ca_1, \dots, ca_{17}\}$, gdzie b, c są liczbami rzeczywistymi. Pokażemy najpierw, że rezultat ten jest prawdziwy, gdy nasze 17 liczb jest całkowite. Weźmy liczby całkowite $\{a_1, \dots, a_{17}\}$ spełniające warunki zadania. Dodając do nich pewien element możemy założyć, że jeden z nich to 0. Usunięcie dowolnego elementu sprawia, że pozostałe 16 mają sumę parzystą, więc wszystkie te liczby są tej samej parzystości. W tym przypadku (zawieranie 0) są to liczby parzyste. Dzieląc je wszystkie przez 2 dostajemy rodzinę 17 liczb całkowitych o własności z zadania i jedno z nich to zero. A więc znowu muszą być wszystkie parzyste i możemy znowu dzielić przez 2. I tak w nieskończoność...

A zatem ciąg liczb całkowitych $\{a_1, \dots, a_{17}\}$ spełniających warunki zadania, o ile zawiera 0, to jest ciągiem zer. A więc ogólnie ciąg liczb całkowitych spełniających warunki zadania składa się z różnych elementów. Jeśli $\{a_1, \dots, a_{17}\}$ są liczbami wymiernymi i spełniają warunki zadania, to można je wszystkie przemościć przez „wspólny mianownik” i dostać liczby całkowite spełniające warunki zadania. Te zaś muszą być równe, jak wyżej.

Załóżmy wreszcie, że mamy układ liczb rzeczywistych $\{a_1, \dots, a_{17}\}$ spełniających warunki zadania. Jeśli $\{b_i \mid i \in I\}$ jest bazą Hamela, to nasze liczby mogą być zapisane w postaci $a_j = \sum_i \lambda_{i,j} b_i$. A zatem dla każdego $i \in I$ układ $\{\gamma_{i,j} \mid 1 \leq j \leq 17\}$ jest układem 17 liczb wymiernych o własności z zadania (dlaczego?) W szczególności $\lambda_{i,j}$ są wszystkie równe pewnemu λ_i .

Zadanie 14.5. Pokazać, że istnieje podzbiór $A \subset \mathbb{R}$ różny od \emptyset, \mathbb{R} taki, że dla każdego $x \in R$ tylko skończenie wiele ze zbiorów $A, A + x, A + 2x, A + 3x, \dots$ jest różnych (przez zbiór $A + B$ rozumiemy podzbiór \mathbb{R} złożony z sum $a + b$, gdzie $a \in A, b \in B$)

Rozwiążanie. Niech B będzie bazą Hamela oraz niech A będzie zbiorem takich liczb $y \in \mathbb{R}$, których zapis $y = \gamma_1 b_1 + \dots + \gamma_m b_m$ w bazie Hamela ma współczynniki γ_i całkowite. Oczywiście $A \neq \mathbb{R}$. Weźmy dowolny $x \in \mathbb{R}$ i niech $x = \gamma_1 b'_1 + \dots + \gamma_n b'_n$ będzie rozkładem x w bazie Hamela B z pewnymi niezerowymi współczynnikami wymiernymi γ_i . Jeśli N jest wspólnym mianownikiem γ_i , to $Nx \in A$, a skoro A jest zamknięty na dodawanie, to $A + Nx = A$. A zatem $A + (k + N)x = A + kx$, dla wszystkich $k \in \mathbb{Z}$. W szczególności tylko zbiory $A, A + x, A + 2x, \dots, A + (N-1)x$ mogą być różne w ciągu $A, A + x, \dots$

14.3 Trivia. Podział prostokąta na kwadraty

Problem – Zadanie. Prostokąt R o bokach długości 1 oraz x , gdzie x jest liczbą niewymierną, nie może być złożony ze skończenie wielu kwadratów.

Załóżmy przeciwnie, że takie rozcięcie prostokąta o rozmiarach $1 \times x$ jest możliwe. Dzielimy go na kwadraty Q_1, \dots, Q_n , gdzie s_i jest długością boku każdego z kwadratów Q_i , dla $1 \leq i \leq n$. UWAGA: to wcale nie muszą (nie mogą wszystkie) być liczby wymierne! Potraktujemy te liczby jako... wektory!

Rozważać będziemy ciało \mathbb{R} jako przestrzeń liniową nad ciałem \mathbb{Q} . Mówiliśmy już kilkakrotnie, że to jest dość niezwykła, nieskończona wymiarowa przestrzeń, kryjąca wiele niespodzianek. Niech $V \subseteq \mathbb{R}$ będzie podprzestrzenią rozpiętą przez liczby s_1, \dots, s_n . Czyli: V to zbiór kombinacji liniowych (o współczynnikach w \mathbb{Q}) tego układu liczb. Skoro (jak twierdzimy) możliwy jest podział prostokąta R na sumę kwadratów o bokach s_i , to mamy $1, x \in \text{lin}(s_1, s_2, \dots, s_n)$, bo $1, x$ są po prostu sumami pewnych s_i .

Teraz pojawia się sprytne (ale jakże często stosowana w matematyce) sztuczka. Określamy funkcję $f : V \rightarrow \mathbb{R}$ spełniającą warunki $f(1) = 1$, $f(x) = -1$ oraz taką, że dla każdych $x, y \in V$ mamy $f(x+y) = f(x) + f(y)$ oraz $f(qx) = qf(x)$, dla $q \in \mathbb{Q}$. Czy taka funkcja istnieje? Przecież to musiaby być jedno z tych dziwnych rozwiązań równania Cauchy'ego postawionego na poprzednim wykładzie! Mamy tu funkcję nie z \mathbb{R} do \mathbb{R} , tylko z V do \mathbb{R} . Dziwne, prawda? Funkcja taka jednak istnieje.

Owszem, skoro $1, x$ są liniowo niezależne nad \mathbb{Q} (a to łatwo sprawdzić), to układ ten możemy na mocy tw. Steinitza dopełnić do bazy $1, x, b_3, \dots, b_k$ przestrzeni V (niekoniecznie $k = n$, bo może niektóre s_i to kombinacje liniowe pozostałych?). Kładziemy dalej $f(1) = 1$, $f(x) = -1$ oraz $f(b_i) = 0$, dla $i = 3, 4, \dots$ Następnie mając funkcję f określona na samej tylko bazie V bierzemy dowolny wektor $v \in V$ i rozpisujemy go (jednoznacznie!) w bazie $1, x, b_3, \dots, b_k$ w postaci $v = a_1 + a_2x + a_3b_3 + \dots + a_kb_k$, gdzie $a_i \in \mathbb{Q}$. Definiujemy teraz $f(v) := a_1f(1) + a_2f(x) + a_3f(b_3) + \dots + a_kf(b_k) = a_1 - a_2$. Zachęcam każdego do sprawdzenia, że teraz nasza funkcja spełnia warunki $f(x+y) = f(x) + f(y)$ oraz $qf(x) = f(qx)$. Tego typu funkcje wprowadzimy niedługo na wykładzie w większej ogólności. Na razie ograniczamy się do powyższych wyjaśnień. Zauważmy też, że absolutnie nie pojawił się wzór na f (w zwykłym sensie).

Rozważmy teraz, dla każdego prostokąta A o bokach a, b , gdzie $a, b \in V$, liczbę $v(A) = f(a)f(b)$. Jeśli prostokąt R rozmiarów $1 \times x$ byłby złożony z kwadratów Q_1, \dots, Q_n , to ze wzoru na sumę pól mamy:

$$v(R) = v(Q_1) + v(Q_2) + \dots + v(Q_n).$$

Jak to jest jednak możliwe, skoro $v(R) = f(1)f(x) = -1$, zaś $v(Q_i) = f(s_i)^2 \geq 0$, dla wszystkich i ? To jest sprzeczność. A zatem R nie może być rozcięty na kwadraty Q_1, \dots, Q_n . Problem rozwiązany.

Rozumowanie to może budzić wiele pytań. Wydaje się, że jest ono przesadnie skomplikowane i wymaga jakiejś strasznej maszynerii. Dlaczego to było konieczne? Oczywiście problemem jest fakt, że postulowany podział kwadratu jest stosunkowo dowolny, liczba składników jest duża, nie muszą to być kwadraty o bokach wymiernej długości – mimo wszystko mamy prawo być zaskoczeni. Użyliśmy poważnej technologii z wykładu, a nawet przemyciliśmy po cichu pojęcie przekształcenia liniowego. To powinno zastanowić.

Rzecz jasna istnieje drugie dno całego tego problemu. Tak naprawdę korzystaliśmy tu po cichu z własności addytywności pola na płaszczyźnie. Nie definiowaliśmy zbyt ściśle co oznacza rozbicie na kwadraty itd. To oczywiście można doprecyzować. Ale jest pewien ogólniejszy problem. Nietrudno pokazać, że dowolny wielokąt na płaszczyźnie można pociąć na części, z których ułoży się prostokąt (a nawet kwadrat) – mówimy, że dowolny wielokąt jest **równoważny przez pocięcie** z prostokątem. Stąd sposób obliczenia pola dowolnego wielokąta jest wyznaczony jednoznacznie. Pytanie: **czy można dowolny wielościan pociąć na skończoną liczbę mniejszych wielościanów, z których ułożyć się prostopadłościan?**

To zaskakujące pytanie było jednym z tzw. 23 problemów Hilberta ogłoszonych w 1900 jako najpoważniejsze problemy matematyczne na nadchodzący XX wiek. Przynajmniej cztery są otwarte do dzisiaj. Problem, który rozważamy rozwiązano jako jeden z pierwszych. Jeszcze w tym samym roku Max Dehn udowodnił, że istnieją pary wielościanów... które nie są równoważne przez pocięcie! Dowód ten jest zrozumiałym dla wytrwałych i opisany bardzo przejrzystie w artykule prof. Marka Kordosa „Pole i objętość” na łamach czasopisma Delta (jest dostępny online – wystarczy wyszukać na stronie <http://www.deltami.edu.pl/>).

Rozdział 15

Przekształcenia liniowe

15.1 Wykład 15

Ostatni wykłady pokazały jak rzeczywistość algebraiczną można oglądać w rzeczywistości geometrycznej, i odwrotnie. Algebraicznemu układowi równań liniowych przypisaliśmy geometrycznie opisywalny zbiór rozwiązań, ale i odwrotnie – pokazaliśmy, że dla każdego zbioru możemy wskazać odpowiedni układ równań jednorodnych. Tego typu ODPOWIEDNIOŚCI są kluczowe w naszym spojrzeniu. Drugą, po układach równań – znacznie ogólniejszą klasą obiektów algebraicznych prowadzącą do rozważania konfiguracji geometrycznych, będą przekształcenia liniowe, którymi zaczniemy się dziś zajmować.

Definicja 15.1.1: Przekształcenie liniowe

Niech V, W będą przestrzeniami liniowymi nad ciałem K . Funkcję $\phi : V \rightarrow W$ nazwiemy PRZEKSZTAŁCENIEM LINIOWYM, jeśli dla dowolnych $\alpha, \beta \in V$ oraz dla każdego $a \in K$ zachodzi:

- (i) $\phi(\alpha + \beta) = \phi(\alpha) + \phi(\beta)$,
- (ii) $\phi(a \cdot \alpha) = a \cdot \phi(\alpha)$. (W szczególności: $\phi(0_V) = 0_W$, czyli zero przechodzi w zero.)

Zwróćmy uwagę na to, że w powyższych warunkach z lewej strony mamy do czynienia z dodawaniem i mnożeniem w przestrzeni V , a po prawej – z dodawaniem i mnożeniem w przestrzeni W .

Poniżej znajduje się lista przykładów przekształceń liniowych (zachęcam do sprawdzenia). Zaczniemy od przykładów algebraicznych. W dalszym ciągu powiemy więcej o klasach mających czytelne interpretacje geometryczne.

- (a) Przekształcenie $\phi : V \rightarrow W$ nazwiemy ZEROVYM, jeśli dla każdego $\alpha \in V$ mamy

$$\phi(\alpha) = 0.$$

- (b) Przekształcenie $\phi : V \rightarrow W$ nazwiemy IDENTYCZNOŚCIĄ, jeśli dla każdego $\alpha \in V$ mamy

$$\phi(\alpha) = \alpha.$$

- (c) Przekształcenie $\phi : K^3 \rightarrow K^3$ dane wzorem

$$\phi((x_1, x_2, x_3)) = (x_1, x_2, x_2).$$

- (d) Przekształcenie $\phi : K^\infty \rightarrow K^\infty$, przypisujące wektorowi o i -tej współrzędnej x_i wektor o i -tej współrzędnej x_{i+1}

$$\phi((x_1, x_2, \dots)) = (x_2, x_3, \dots).$$

- (e) Odwzorowanie $d : K[x] \rightarrow K[x]$, zwane POCHODNĄ, zadane wzorem

$$\phi(a_0 + a_1x + \dots + a_nx^n) = a_1 + 2a_2x + 3a_3x^2 + \dots + na_nx^{n-1}.$$

(f) Odwzorowanie $tr : M_{n \times n}(K) \rightarrow K$ przypisujące macierzy $A = [a_{ij}]$ sumę elementów na przekątnej:

$$\phi(A) = a_{11} + \dots + a_{nn},$$

zwane ŚLADEM.

(g) $\phi : F(\mathbb{R}, \mathbb{R}) \rightarrow F(\mathbb{R}, \mathbb{R})$ przyporządkowujące funkcji f funkcję parzystą $\phi(f)$ daną wzorem

$$(\phi(f))(x) = \frac{f(x) + f(-x)}{2},$$

(h) Niech $(\mathbb{R}_+, \boxplus, \boxtimes, 1)$ będzie przestrzenią liniową nad \mathbb{R} , gdzie $x \boxplus y = xy$ oraz $a \boxtimes x = x^a$, dla $a \in \mathbb{R}$. Przekształcenie $l : \mathbb{R}_+ \rightarrow \mathbb{R}$ dane wzorem

$$l(x) = \ln(x).$$

Przekształcenia liniowe to jedyne funkcje pomiędzy przestrzeniami liniowymi, które ZACHOWUJĄ KOMBINACJE LINIOWE.

Uwaga 15.1.2

Następujące warunki są równoważne:

- (i) $\phi : V \rightarrow W$ jest przekształceniem liniowym,
- (ii) dla każdych $a_1, \dots, a_k \in K$ oraz $\alpha_1, \dots, \alpha_k \in V$ zachodzi

$$\phi(a_1\alpha_1 + a_2\alpha_2 + \dots + a_k\alpha_k) = a_1\phi(\alpha_1) + a_2\phi(\alpha_2) + \dots + a_k\phi(\alpha_k).$$

Dowód. Indukcja ze względu na k . Dla $k = 2$ teza wynika z definicji przekształcenia liniowego. Najpierw korzystamy z tego, że ϕ zachowuje dodawanie, a potem mnożenie przez skalar.

$$\phi(a_1\alpha_1 + a_2\alpha_2) = \phi(a_1\alpha_1) + \phi(a_2\alpha_2) = a_1\phi(\alpha_1) + a_2\phi(\alpha_2).$$

Niech $k > 2$. Z definicji przekształcenia liniowego mamy:

$$\phi(a_1\alpha_1 + a_2\alpha_2 + \dots + a_k\alpha_k) = \phi(a_1\alpha_1 + a_2\alpha_2 + \dots + a_{k-1}\alpha_{k-1}) + a_k\phi(\alpha_k).$$

Korzystając z założenia indukcyjnego dostajemy implikację (i) \Rightarrow (ii). Odwrotna implikacja jest oczywista. \square

Przykład. Jeśli o przekształceniu liniowym $\phi : \mathbb{R}^3 \rightarrow \mathbb{R}^2$ wiemy, że

$$\phi((1, 1, 0)) = (2, 1), \quad \phi((0, 1, 1)) = (3, 1), \quad \phi((2, 1, 1)) = (1, 0),$$

to możemy wyznaczyć $\phi((1, 0, 0)), \phi((0, 1, 0)), \phi((1, 0, 0))$. Istotnie, skoro

$$\begin{aligned} (1, 0, 0) &= 0 \cdot (1, 1, 0) + \left(-\frac{1}{2}\right) \cdot (0, 1, 1) + \frac{1}{2} \cdot (2, 1, 1) \\ (0, 1, 0) &= 1 \cdot (1, 1, 0) + \frac{1}{2} \cdot (0, 1, 1) + \left(-\frac{1}{2}\right) \cdot (2, 1, 1) \\ (0, 0, 1) &= (-1) \cdot (1, 1, 0) + \frac{1}{2} \cdot (0, 1, 1) + \frac{1}{2} \cdot (2, 1, 1) \end{aligned}$$

to

$$\begin{aligned} \phi((1, 0, 0)) &= 0 \cdot (2, 1) + \left(-\frac{1}{2}\right) \cdot (3, 1) + \frac{1}{2} \cdot (1, 0) = \left(-1, -\frac{1}{2}\right) \\ \phi((0, 1, 0)) &= 1 \cdot (2, 1) + \frac{1}{2} \cdot (3, 1) + \left(-\frac{1}{2}\right) \cdot (1, 0) = \left(3, \frac{3}{2}\right) \\ \phi((0, 0, 1)) &= (-1) \cdot (2, 1) + \frac{1}{2} \cdot (3, 1) + \frac{1}{2} \cdot (1, 0) = \left(0, -\frac{1}{2}\right) \end{aligned}$$

Zauważmy, że powyższe wyliczenia pozwalają nam określić wzór przekształcenia ϕ . Mówi o tym w sposób ogólny poniższa uwaga.

Wniosek 15.1.3

Niech $\epsilon_1, \dots, \epsilon_n$ będzie bazą standardową przestrzeni liniowej K^n . Dowolne przekształcenie liniowe $\phi : K^n \rightarrow K^m$ jest zadane wzorem:

$$\begin{aligned}\phi((x_1, \dots, x_n)) &= \phi(x_1(1, 0, \dots, 0) + x_2(0, 1, \dots, 0) + \dots + x_n(0, 0, \dots, 1)) = \\ &= \phi(x_1\epsilon_1 + \dots + x_n\epsilon_n) = \\ &= x_1\phi(\epsilon_1) + \dots + x_n\phi(\epsilon_n) = \\ &= x_1(a_{11}, \dots, a_{m1}) + \dots + x_n(a_{1n}, \dots, a_{mn}) = \\ &= (a_{11}x_1 + \dots + a_{1n}x_n, a_{21}x_1 + \dots + a_{2n}x_n, \dots, a_{m1}x_1 + \dots + a_{mn}x_n),\end{aligned}$$

dla dowolnych elementów $a_{ij} \in K$, spełniających warunek $\phi(\epsilon_i) = (a_{1i}, \dots, a_{mi})$.

Uzasadnienie, że wszystkie przekształcenia rozważane w powyższym wniosku są liniowe przedstawimy w ogólnym kontekście w dowodzie ważnego Twierdzenia 15.1.7 (choć jest ono bezpośrednie).

Przykład. Wracając do powyższego przykładu widzimy, że przekształcenie liniowe $\phi : \mathbb{R}^3 \rightarrow \mathbb{R}^2$ spełniające warunki

$$\phi((1, 0, 0)) = \left(-1, -\frac{1}{2}\right), \quad \phi((0, 1, 0)) = \left(3, \frac{3}{2}\right), \quad \phi((0, 0, 1)) = \left(0, -\frac{1}{2}\right)$$

ma wzór

$$\phi((x_1, x_2, x_3)) = \left(-x_1 + 3x_2, -\frac{1}{2}x_1 + \frac{3}{2}x_2 - \frac{1}{2}x_3\right).$$

Przejdziemy do podania ważnych klas przekształceń liniowych, mających odniesienia geometryczne.

Definicja 15.1.4: Homotetia/jednokładność

Przekształcenie liniowe $\phi : V \rightarrow V$ przestrzeni liniowej w siebie dane wzorem

$$\phi(\alpha) = a\alpha$$

nazywamy HOMOTETIĄ (albo JEDNOKŁADNOŚCIĄ) o skali a .

Oczywiście homotetia jest przekształceniem liniowym. Jeśli $\alpha, \beta \in V$, to

$$\phi(\alpha + \beta) = a(\alpha + \beta) = a\alpha + a\beta = \phi(\alpha) + \phi(\beta)$$

Podobnie sprawdzamy, że dla każdego $c \in K$ mamy $\phi(c\alpha) = c\phi(\alpha)$.

Definicja 15.1.5: Obrót

OBROTEM O KĄT $\theta \in \mathbb{R}$ nazywamy przekształcenie $\phi : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ zadane wzorem

$$\phi((x_1, x_2)) = (\cos \theta \cdot x_1 - \sin \theta \cdot x_2, \sin \theta \cdot x_1 + \cos \theta \cdot x_2).$$

Zauważmy, że $\phi((1, 0)) = (\cos \theta, \sin \theta)$, $\phi((0, 1)) = (-\sin \theta, \cos \theta)$.

Obrót jest przekształceniem liniowym na mocy Wniosku 15.1.3.

Przykład. Dla liczby zespolonej $u = \cos \theta + i \sin \theta$ o module 1 definiujemy przekształcenie $\phi : \mathbb{C} \rightarrow \mathbb{C}$ wzorem:

$$\phi(z) = uz$$

Jest to przekształcenie liniowe \mathbb{C} jako dwuwymiarowej przestrzeni liniowej nad \mathbb{R} , a dokładniej jest to obrót o kąt θ . Rzeczywiście, biorąc $z = x_1 + ix_2$ mamy

$$\phi(z) = (\cos \theta + i \sin \theta)(x_1 + ix_2) = (\cos \theta \cdot x_1 - \sin \theta \cdot x_2) + i \cdot (\sin \theta \cdot x_1 + \cos \theta \cdot x_2).$$

Czytelnik bez trudu wywnioskuje stąd, że jeśli w powyższej definicji pominiemy warunek $|u| = 1$, to uzyskane przekształcenie ϕ jest złożeniem jednokładności i obrotu.

Przejdziemy teraz do niezwykle ważnych dla całego kursu klas przekształceń — rzutu i symetrii. Będą one zdefiniowane w sposób niezupełnie być może zgodny z intuicją geometryczną Czytelnika. Będą to bowiem rzuty i przekształcenia „w kierunku” wyznaczonym przez pewną podprzestrzeń.

Definicja 15.1.6: Rzut i symetria wzdłuż podprzestrzeni

Jeśli $V = V_1 \oplus V_2$, to dla każdego $\alpha \in V$ istnieją jednoznacznie wyznaczone $\alpha_1 \in V_1$ i $\alpha_2 \in V_2$, że

$$\alpha = \alpha_1 + \alpha_2.$$

Definiujemy:

- RZUT $\phi : V \rightarrow V$ przestrzeni V na V_1 wzdłuż V_2 dany wzorem $\phi(\alpha) = \alpha_1$.
- SYMETRIE $\psi : V \rightarrow V$ przestrzeni V względem V_1 wzdłuż V_2 daną wzorem $\psi(\alpha) = \alpha_1 - \alpha_2$.

Przykład 1. Mamy $\mathbb{R}^2 = \text{lin}((1,1)) \oplus \text{lin}((1,0))$. W szczególności każdy wektor (x_1, x_2) zapisuje się jednoznacznie jako suma elementów tej podprzestrzeni

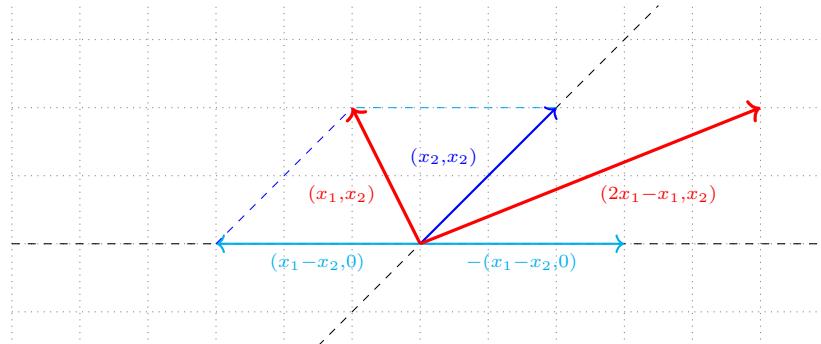
$$(x_1, x_2) = (x_2, x_2) + (x_1 - x_2, 0).$$

W szczególności rzut ϕ przestrzeni \mathbb{R}^2 na $\text{lin}((1,1))$ wzdłuż $\text{lin}((1,0))$ określony jest zgodnie z definicją wyżej wzorem

$$\phi((x_1, x_2)) = (x_2, x_2)$$

natomiast symetria ψ względem $\text{lin}((1,1))$ wzdłuż $\text{lin}((1,0))$ określona jest wzorem

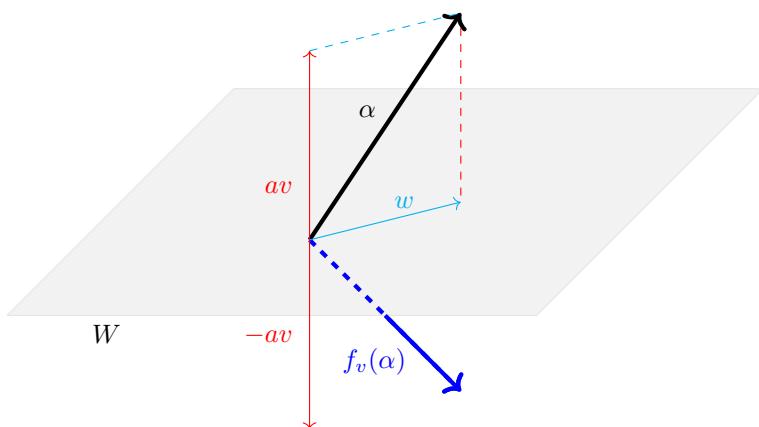
$$\psi((x_1, x_2)) = (x_2, x_2) - (x_1 - x_2, 0) = (2x_1 - x_1, x_2).$$



Rys. 1. Obraz wektora $(-1, 2) = (2, 2) + (-3, 0)$ przy rzucie ϕ przestrzeni \mathbb{R}^2 na $\text{lin}((1, 1))$ wzdłuż $\text{lin}((1, 0))$ to $(2, 2)$, natomiast obraz $(-1, 2)$ przy symetrii względem $\text{lin}((1, 1))$ wzdłuż $\text{lin}((1, 0))$ to $(2, 2) - (-3, 0) = (5, 2)$. Oznajmy też, że obraz tego samego wektora $(-1, 2)$ przy rzucie na $\text{lin}((1, 0))$ wzdłuż $\text{lin}((1, 1))$ to $(-3, 0)$.

Przykład 2. Niech $V = \text{lin}(v) \oplus W$, gdzie $v \neq 0$ oraz W jest podprzestrzenią wymiaru $n - 1$, która nie zawiera wektora v . Wówczas można określić ODBICIE wektora α względem podprzestrzeni W , będące po prostu symetrią V względem W wzdłuż $\text{lin}(v)$.

Oto ilustracja w przestrzeni trójwymiarowej. Jest to więc przekształcenie $f_v : \mathbb{R}^3 \rightarrow \mathbb{R}^3$, które bierze wektor α , rozkłada go na sumę $\alpha = av + w$, gdzie $w \in W$, i przeprowadza go na $f_v(\alpha) = -av + w$.



Na zakończenie wstępnych rozważań uzasadnimy twierdzenie mówiące, że zadanie wartości przekształcania liniowego na bazie, określa je jednoznacznie (także, dla przestrzeni nieskończonego wymiaru).

Twierdzenie 15.1.7: O jednoznaczności na bazie

Niech V, W będą przestrzeniami liniowymi nad ciałem K . Niech $\alpha_1, \dots, \alpha_n$ będzie bazą przestrzeni V , zaś β_1, \dots, β_n niech będzie dowolnym układem wektorów przestrzeni W . Wówczas istnieje **dokładnie jedno** takie przekształcenie liniowe $\phi : V \rightarrow W$, że

$$\phi(\alpha_1) = \beta_1, \quad \phi(\alpha_2) = \beta_2, \quad \dots, \quad \phi(\alpha_n) = \beta_n. \quad (*)$$

Dowód. Pokażemy najpierw, że istnieje przekształcenie ϕ spełniające (*). Dla każdego $\alpha \in V$ istnieją $a_1, \dots, a_n \in K$, że $\alpha = a_1\alpha_1 + \dots + a_n\alpha_n$ (czyli: współrzędne α w bazie $\alpha_1, \dots, \alpha_n$). Oznacza to, że poniższe przekształcenie $\phi : V \rightarrow W$ jest dobrze określone (czyli — jest funkcją):

$$\phi(\alpha) = a_1\beta_1 + \dots + a_n\beta_n.$$

Podstawiając za α kolejne α_i , dla $1 \leq i \leq n$, dostajemy oczywiście $\phi(\alpha_i) = \beta_i$, bo jedyną niezerową współrzędną wektora α_i w bazie $\alpha_1, \dots, \alpha_n$ jest i -ta współrzędna równa 1. A zatem ϕ spełnia (*). Dlaczego jest to przekształcenie liniowe? Jeśli $\alpha' = a'_1\alpha_1 + \dots + a'_n\alpha_n$, dla pewnych współrzędnych a'_1, \dots, a'_n , to

$$\phi(\alpha + \alpha') = (a_1 + a'_1)\beta_1 + \dots + (a_n + a'_n)\beta_n = a_1\beta_1 + \dots + a_n\beta_n + a'_1\beta_1 + \dots + a'_n\beta_n = \phi(\alpha) + \phi(\alpha').$$

Podobnie pokazujemy, że dla każdego $a \in K$ mamy $\phi(a\alpha) = a\phi(\alpha)$. A zatem ϕ jest liniowe.

Założymy, że istnieje przekształcenie liniowe $\psi : V \rightarrow W$ takie, że $\psi(\alpha_i) = \beta_i$, dla $1 \leq i \leq n$. Wówczas dla każdego $\alpha \in V$ mamy, na mocy Uwagi 15.1.3:

$$\begin{aligned} \psi(\alpha) &= \psi(a_1\alpha_1 + \dots + a_n\alpha_n) = a_1\psi(\alpha_1) + \dots + a_n\psi(\alpha_n) = \\ &= a_1\beta_1 + \dots + a_n\beta_n = a_1\phi(\alpha_1) + \dots + a_n\phi(\alpha_n) = \phi(a_1\alpha_1 + \dots + a_n\alpha_n) = \phi(\alpha). \end{aligned}$$

A zatem przekształcenie liniowe ϕ spełniające (*) jest wyznaczone jednoznacznie. \square

Przykład. Niech $\phi : \mathbb{R}^2 \rightarrow \mathbb{R}^4$ spełnia następujące warunki:

$$\phi((1, 0)) = (1, 1, 2, 1), \quad \phi((0, 1)) = (0, 3, 1, -2).$$

Wówczas z liniowości ϕ mamy:

$$\begin{aligned} \phi((x, y)) &= \phi(x(1, 0) + y(0, 1)) = x \cdot \phi((1, 0)) + y \cdot \phi((0, 1)) = x(1, 1, 2, 1) + y(0, 3, 1, -2) = \\ &= (x, x + 3y, 2x + y, x - 2y). \end{aligned}$$

Warto jest zrozumieć konsekwencje zaprzeczenia założeń powyższego twierdzenia.

- W przypadku, gdy układ $\alpha_1, \dots, \alpha_n$ jest liniowo niezależny, ale nie rozpina V , dla ustalonego układu β_1, \dots, β_n w przestrzeni W może istnieć wiele różnych przekształceń liniowych spełniających (*).

Weźmy $\alpha_1 = (1, 0, 0), \alpha_2 = (0, 1, 0), \beta_1 = (1, 0, 0), \beta_2 = (0, 1, 0)$. Układ $((1, 0, 0), (0, 1, 0))$ jest liniowo niezależny w \mathbb{R}^3 , ale nie rozpina tej przestrzeni. Istnieje więc więcej niż jedno przekształcenie liniowe $\phi : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ takie, że $\phi(\alpha_1) = \beta_1$ oraz $\phi(\alpha_2) = \beta_2$. Jednym z nich jest identyczność, a drugim symetria względem płaszczyzny $\text{lin}((1, 0, 0), (0, 1, 0))$ wzdłuż $\text{lin}(0, 0, 1)$,

- Jeśli $\alpha_1, \dots, \alpha_n$ nie jest liniowo niezależny, to dla pewnego układu β_1, \dots, β_n może nie istnieć przekształcenie liniowe $\phi : V \rightarrow W$ spełniające (*).

Weźmy $\alpha_1 = (1, 0), \alpha_2 = (0, 1), \alpha_3 = (1, 1)$ oraz $\beta_1 = \beta_2 = \beta_3 = (1, 0, 0)$. Teraz układ $\alpha_1, \alpha_2, \alpha_3$ rozpina \mathbb{R}^3 , ale jest liniowo zależny. Gdyby istniało takie przekształcenie liniowe $\phi : \mathbb{R}^2 \rightarrow \mathbb{R}^3$, że $\phi(\alpha_1) = \phi(\alpha_2) = \phi(\alpha_3)$, to z liniowości ϕ mielibyśmy

$$(0, 0, 0) = \phi((0, 0)) = \phi(\alpha_1 + \alpha_2 - \alpha_3) = \phi(\alpha_1) + \phi(\alpha_2) - \phi(\alpha_3) = (1, 0, 0).$$

Zauważmy, że nie postawiliśmy żadnych warunków odnośnie układu wektorów $\{\beta_i\}$ — poza równolicznością z bazą $\alpha_1, \dots, \alpha_n$. Wektory te mogą być nawet wszystkie równe wektorowi zerowemu przestrzeni W .

15.2 Wybór przykładowych pytań

Proszę się nie uczyć na pamięć odpowiedzi czy wyjaśnienie! Takich pytań można ułożyć znacznie więcej i służą one jedynie sprawdzeniu podstawowego zrozumienia treści zajęć. Jeśli odpowiedź na dowolne z poniższych pytań sprawia Państwu problem, proszę pomyśleć który fragment notatek dotyczy tego pytania, wrócić do ich lektury, a w razie dalszych trudności poprosić prowadzącego o wskazówkę.

1. Czy przekształcenie które każdemu wektorowi przyporządkowuje wektor przeciwny jest przekształceniem liniowym?
2. Dana jest przestrzeń liniowa V i wektor $\alpha \in V$. Czy przekształcenie zadane wzorem $f(v) = v + \alpha$ dla każdego $v \in V$ jest przekształceniem liniowym?
3. Niech V będzie przestrzenią liniową nad ciałem \mathbb{K} . Dla jakich $a \in \mathbb{K}$ przekształcenie przypisujące wszystkim wektorom z V element a jest liniowe?
4. Przekształcenie liniowe $\phi : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ spełnia $\phi((1, 0, 2)) = (1, 2)$ oraz $\phi((1, 0, 3)) = (1, 3)$.
 - Czy $\phi((1, 0, 4)) = (1, 4)$?
 - Czy $\phi((2, 0, 4)) = (2, 4)$?
 - Czy możliwe jest wyznaczenie $\phi((0, 0, 2))$?
 - Czy możliwe jest wyznaczenie $\phi((0, 1, 0))$?
 - Czy możliwe jest wskazanie takiego $\alpha \in \mathbb{R}^3$, że $\phi(\alpha) = (1, 0)$?
5. Niech $\phi_a : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ będzie homotetią o skali a .
 - Wyznacz $\phi_{-\frac{1}{2}}((1, 1, 1))$.
 - Wyznacz wszystkie wektory $\alpha \in \mathbb{R}^3$, że $\phi_{-\frac{1}{2}}(\alpha) = (1, 1, 1)$.
 - Wyznacz $\phi_2(\phi_{-2}((1, 2, 3)))$.
 - Wyznacz $\phi_2(\phi_{\frac{1}{2}}((1, 2, 3)))$.
 - Czy istnieje takie a , że $f_a((1, 2, 3)) = (3, 2, 1)$?
6. Niech $O_\theta : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ będzie obrotem o kąt θ .
 - Wyznacz $O_{-\frac{\pi}{2}}((1, 1))$.
 - Wyznacz $O_{\frac{\pi}{3}}((\sqrt{3}, 1))$.
 - Czy istnieje $\theta \in (0, 2\pi)$ taki, że $O_\theta((1, 1)) = (-\sqrt{2}, 0)$?
 - Czy istnieje $\theta \in (0, 2\pi)$ taki, że $O_\theta((3, 2)) = (2, -3)$?
 - Uzasadnij, że dla każdego α $O_{\theta+\zeta}(\alpha) = O_\theta(O_\zeta(\alpha))$.
7. Niech $W = \text{lin}((1, 1))$ oraz $U = \text{lin}((1, -1))$.
 - Znajdź takie $\alpha \in W$ oraz $\beta \in U$, że $(2, 1) = \alpha + \beta$.
 - Wyznacz obraz wektora $(2, 1)$ przy rzucie \mathbb{R}^2 na W wzdłuż U .
 - Wyznacz obraz wektora $(2, 1)$ przy rzucie \mathbb{R}^2 na U wzdłuż W .
 - Wyznacz obraz wektora $(2, 1)$ przy symetrii \mathbb{R}^2 względem W wzdłuż U .
 - Wyznacz obraz wektora $(2, 1)$ przy symetrii \mathbb{R}^2 względem U wzdłuż W .
8. Czy istnieje przekształcenie liniowe $\phi : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ takie, że $\phi((1, 0)) = (1, 0)$ oraz $\phi((2, 0)) = (1, 1)$?
9. Czy istnieje przekształcenie liniowe $\phi : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ takie, że $\phi((1, 0)) = (1, 0)$ oraz $\phi((1, 1)) = (2, 0)$?
10. Niech $(\mathbb{R}_+, \boxplus, \boxtimes, 1)$ będzie przestrzenią liniową nad \mathbb{R} , gdzie $x \boxplus y = xy$ oraz $a \boxtimes x = x^a$, dla $a \in \mathbb{R}$. Uzasadnij, że przekształcenie $l : \mathbb{R}_+ \rightarrow \mathbb{R}$ dane wzorem

$$l(x) = \ln(x)$$
 jest liniowe.
 11. Ile jest przekształceń liniowych $\phi : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ takich, że $\phi((1, 0)) = (1, 0)$ oraz $\phi((1, 1)) = (1, 0)$?
 12. Ile jest przekształceń liniowych $\phi : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ takich, że $\phi((1, 0)) = (1, 0)$?

15.3 Zadania do samodzielnej pracy

Zadania ilustrujące umiejętności niezbędne do zaliczenia przedmiotu oznaczone są symbolem ♠. Przy tych zadaniach dodany jest opis umiejętności, które sprawdzają.

1. Które z poniższych odwzorowań $\phi : \mathbb{R}^3 \rightarrow \mathbb{R}^2$ są przekształceniemi liniowymi:
 - a) $\phi((x_1, x_2, x_3)) = (x_1 + 3x_2 - 1, 4x_1 + 2x_2 + 6)$,
 - b) $\phi((x_1, x_2, x_3)) = (x_1 + 3x_2 - x_3, 4x_1 + 2x_2 + 6x_3)$,
 - c) $\phi((x_1, x_2, x_3)) = (x_1 + 3x_2 - x_3, 4|x_1| + 2|x_2| + 6|x_3|)$,
 - d) $\phi((x_1, x_2, x_3)) = ((x_1 + 2)^2 - x_1^2 - x_3 - 4, 4x_1 + 2x_2 + 6x_3)$.
2. Dla jakich wartości parametru $t \in \mathbb{R}$ przekształcenie $\phi : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ dane wzorem

$$\phi((x_1, x_2)) = (x_1 + x_2 + (t^2 - 9)x_1x_2, 5x_1 + 3(x_2 - 1) + t)$$
 jest przekształceniem liniowym?
3. (♠) Znajdowanie wzoru na przekształcenie liniowe zadane na bazie)
 Znajdź wzory na przekształcenia liniowe $\phi : \mathbb{R}^n \rightarrow \mathbb{R}^m$ zadane następującymi warunkami:
 - (a) $\phi : \mathbb{R}^2 \rightarrow \mathbb{R}^3$, gdzie

$$\phi((3, 1)) = (4, 5, -1), \quad \phi((7, 2)) = (-3, 0, 5),$$
 - (b) $\phi : \mathbb{R}^3 \rightarrow \mathbb{R}^2$, gdzie

$$\phi((1, 2, 1)) = (7, 2), \quad \phi((3, 2, 4)) = (20, 17), \quad \phi((5, 1, 2)) = (17, 12),$$
 - (c) $\phi : \mathbb{R}^3 \rightarrow \mathbb{R}^3$, gdzie

$$\phi((1, 0, 1)) = (5, 1, 3), \quad \phi((0, 1, 1)) = (2, 3, 4), \quad \phi((1, 0, 0)) = (6, 7, 7).$$
4. (♠) Znajdowanie wzoru na rzut/symetrię)
 Niech $\alpha_1 = (1, 2, 2), \alpha_2 = (1, 2, 1), \alpha_3 = (1, 1, 2)$. Niech $V = \text{lin}(\alpha_1, \alpha_2)$ oraz $W = \text{lin}(\alpha_3)$. Znajdź wzór na $\phi : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ będący symetrią \mathbb{R}^3 względem podprzestrzeni V wzdłuż podprzestrzeni W .
5. Niech $\phi : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ będzie rzutem na $\text{lin}((1, 1))$ wzdłuż $\text{lin}((1, -1))$. Oblicz (bez znajdowania wzoru przekształcenia ϕ) ile wynosi $\phi(1, 0)$.
6. Znajdź wzór rzutu $\phi : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ na $\text{lin}((1, 1))$ wzdłuż $\text{lin}((1, -1))$.
7. Niech $\psi : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ będzie symetrią względem $\text{lin}((-1, 0, 0))$ wzdłuż $\text{lin}((0, -1, 0), (0, 1, 1))$. Oblicz (bez znajdowania wzoru przekształcenia ϕ) $\psi((3, 2, 1))$.
8. Wyznacz wzór symetrii $\psi : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ $\text{lin}((-1, 0, 0))$ wzdłuż $\text{lin}((0, -1, 0), (0, 1, 1))$.
9. Rozstrzygnij, czy przekształcenie $f : \mathbb{C} \rightarrow \mathbb{C}$ dane wzorem $f(z) = \bar{z}$ jest liniowe?
10. Uzasadnij, że jeśli $\phi : V \rightarrow V$ jest przekształceniem liniowym oraz $\dim V = 1$, to ϕ jest homotetią.
11. Uzasadnij, że jeśli $\{\alpha_1, \dots, \alpha_k, \beta_1, \dots, \beta_m\}$ jest bazą przestrzeni V , to rzut ϕ na $\text{lin}(\alpha_1, \dots, \alpha_k)$ wzdłuż $\text{lin}(\beta_1, \dots, \beta_m)$ spełnia $\phi(\alpha_i) = \alpha_i$ oraz $\phi(\beta_j) = 0$.
12. Niech $\psi : V \rightarrow V$ będzie symetrią względem $\text{lin}(\alpha_1, \dots, \alpha_k)$ wzdłuż $\text{lin}(\beta_1, \dots, \beta_m)$. Wyznacz wartości $\psi(\alpha_i)$ oraz $\psi(\beta_j)$.
13. Niech $\phi : V \rightarrow W$ będzie przekształceniem liniowym i niech $\alpha_1, \dots, \alpha_m$ będzie takim układem wektorów z V , że układ $\phi(\alpha_1), \dots, \phi(\alpha_m)$ jest liniowo niezależny. Wykaż, że układ $\alpha_1, \dots, \alpha_m$ jest liniowo niezależny.
14. Niech V, W będą przestrzeniami liniowymi nad ciałem K . Wykresem funkcji $\phi : V \rightarrow W$ nazywamy zbiór

$$G_\phi = \{(\alpha, \phi(\alpha)) \mid \alpha \in V\}.$$
 Wykaż, że ϕ jest przekształceniem liniowym wtedy i tylko wtedy, gdy G_ϕ jest podprzestrzenią przestrzeni $V \times W$.
15. Wskaż przykład funkcji $f : \mathbb{R}^2 \rightarrow \mathbb{R}$, która spełnia warunek $f(a\alpha) = af(\alpha)$, dla każdego $a \in \mathbb{R}$ oraz $\alpha \in \mathbb{R}$, ale f nie jest przekształceniem liniowym.

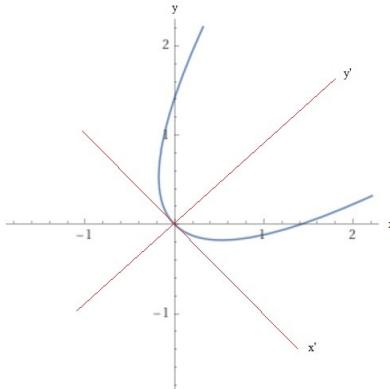
15.4 Uzupełnienie. Liniowa zamiana współrzędnych

W rozdziale wprowadzającym przekształcenia liniowe wskazaliśmy kilka razy na aspekt geometryczny. W tym uzupełnieniu uczynimy to ponownie, odnosząc się do elementarnej geometrii płaszczyzny i jej podzbiorów. Temat ten w odpowiedniej ogólności zaprezentowany zostanie pod koniec GALu II.

Rozważmy następujący przykład. Na płaszczyźnie, której punkty mają współrzędne x, y rozważamy podzbiór (możemy zamiast współrzędnych punktów myśleć o współrzędnych wektorów w bazie standardowej) złożony z takich (x, y) , które spełniają równanie

$$x^2 - 2xy + y^2 - \sqrt{2}(x + y) = 0.$$

Być może ktoś z Państwa wpisze to równanie do programu graficznego i przekona się, że szukany zbiór punktów wygląda jak parabola.



Czy mamy jakiś algebraiczne wyjaśnienie uzasadniające, że w istocie uzyskaliśmy parabolę? Oto jeden z możliwych sposobów patrzenia. Rozważmy przekształcenie liniowe $\phi : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ dane wzorem:

$$\phi((x, y)) = \left(\frac{\sqrt{2}}{2}x - \frac{\sqrt{2}}{2}y, \frac{\sqrt{2}}{2}x + \frac{\sqrt{2}}{2}y \right).$$

Jest to w istocie obrót o $\frac{\pi}{4}$ (przeciwne do kierunku wskazówek zegara). Przyglądając się rysunkowi powyżej spodziewamy się, że przekształcenie to wyprostuje naszą parabolę, a proste zaznaczone na czerwono „przeniesie” na osie. Istotnie: możemy myśleć o naszym przekształceniu jako o „zmianie współrzędnych”. A zatem przyjmujemy:

$$x' = \frac{\sqrt{2}}{2}x - \frac{\sqrt{2}}{2}y, \quad y' = \frac{\sqrt{2}}{2}x + \frac{\sqrt{2}}{2}y$$

i próbujemy wyrazić zbiór wyżej za pomocą nowych zmiennych. Nietrudno widzieć, że mamy:

$$x = \frac{\sqrt{2}}{2}x' + \frac{\sqrt{2}}{2}y', \quad y = -\frac{\sqrt{2}}{2}x' + \frac{\sqrt{2}}{2}y'.$$

Wstawiając uzyskane wyrażenia do równości $x^2 - 2xy + y^2 - \sqrt{2}(x + y) = 0$ otrzymujemy

$$\frac{1}{2}(x' + y')^2 - (x' + y')(-x' + y') + \frac{1}{2}(-x' + y')^2 - y' = 0.$$

Po redukcji, otrzymujemy już bez żadnych wątpliwości równanie paraboli postaci:

$$y' = 2x'^2.$$

Co oznacza nasze rozwiązanie? Mówiąc, że jeśli zamiast patrzeć na rozważany zbiór wektorów (x, y) w zapisanych we współrzędnych w bazie $(1, 0), (0, 1)$ spełniających pewne równanie warto patrzeć na ten sam zbiór wektorów zapisanych we współrzędnych (x', y') w bazie

$$\left(\frac{\sqrt{2}}{2}, -\frac{\sqrt{2}}{2} \right), \quad \left(\frac{\sqrt{2}}{2}, \frac{\sqrt{2}}{2} \right)$$

Wtedy współrzędne tych samych wektorów spełniają już równanie paraboli. A skąd jest baza powyżej? Geometrycznie to jest jasne (patrz czerwone osie), a algebraicznie mamy:

$$(x, y) = \left(\frac{\sqrt{2}}{2}x' + \frac{\sqrt{2}}{2}y', -\frac{\sqrt{2}}{2}x' + \frac{\sqrt{2}}{2}y' \right) = x' \left(\frac{\sqrt{2}}{2}, -\frac{\sqrt{2}}{2} \right) + y' \left(\frac{\sqrt{2}}{2}, \frac{\sqrt{2}}{2} \right).$$

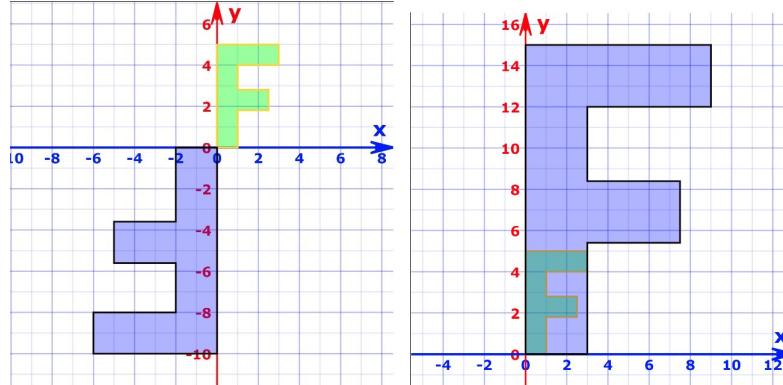
15.5 Trivia. Geometria przekształceń liniowych

Aby przekonać się, że przekształcenia liniowe mają sporo wspólnego z podstawowymi przekształceniiami geometrycznymi płaszczyzny znanyimi ze szkoły, rozważmy funkcje $f : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ postaci:

$$f(x, y) = (ax + by, cx + dy),$$

gdzie $a, b, c, d \in \mathbb{R}$. Jak wiemy tylko funkcje zadane tymi wzorami mogą opisywać przekształcenia liniowe z K^2 do K^2 . Jeśli spojrzymy na owe funkcje¹ z punktu widzenia geometrii analitycznej, a więc jako funkcje z (płaszczyzny kartezjańskiej) \mathbb{R}^2 do \mathbb{R}^2 , wówczas wśród f znajdują się między innymi:

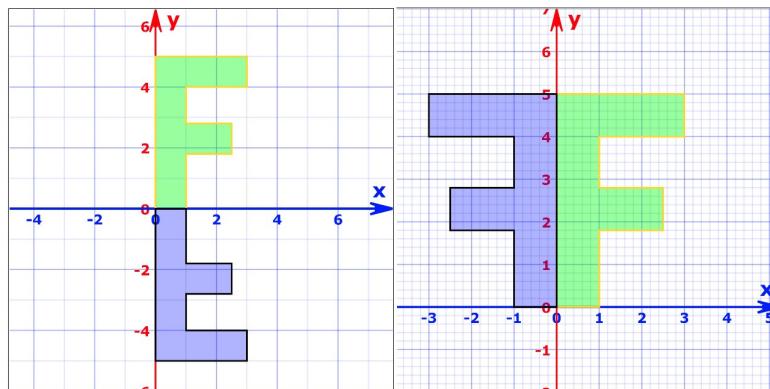
- jednokładność o skali λ i środku $(0, 0)$ postaci $f(x, y) = (\lambda x, \lambda y)$



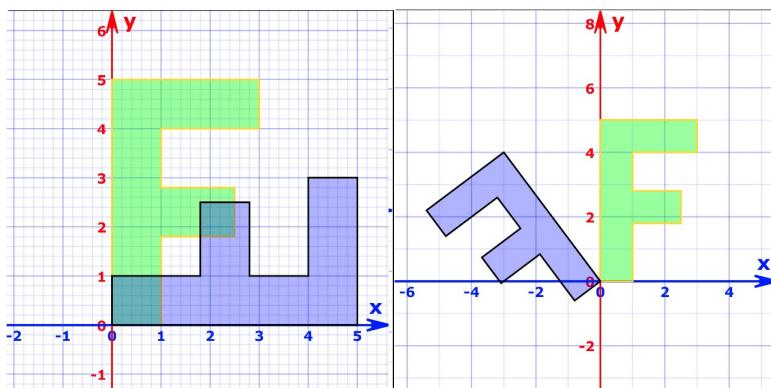
Rys. 1. Jednokładność o skalach odpowiednio $\lambda = -2$ (z lewej) oraz $\lambda = 3$.

- symetria prostopadła względem prostej $y = \tan(\theta/2)$ dana wzorem

$$f(x, y) = (\cos \theta \cdot x + \sin \theta \cdot y, \sin \theta \cdot x - \cos \theta \cdot y).$$



Rys. 2. Symetria prostopadła względem prostej $x = 0$ (z lewej) oraz $y = 0$ (z prawej).

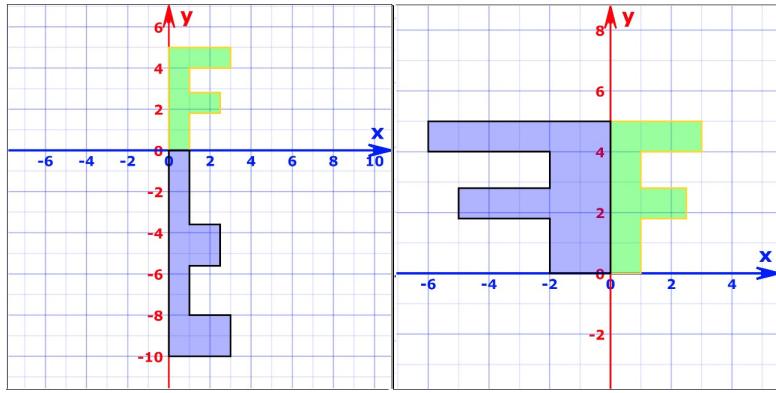


Rys. 3. Symetria prostopadła względem prostej $y = x$ (z lewej) dana wzorem $(x, y) \mapsto (y, x)$ oraz względem prostej $y = -3x$ (z prawej) dana wzorem $(x, y) \mapsto (-\frac{4}{5}x - \frac{3}{5}y, -\frac{3}{5}x + \frac{4}{5}y)$.

¹Rysunki uzyskane za pomocą portalu <https://www.mathsisfun.com/algebra/matrix-transform.html>, gdzie można samodzielnie poeksperymentować — choćby po to, by przekonać się, że nie dostajemy tylko izometrii czy podobieństw.

- powinowactwa prostokątne o osiach OX lub OY oraz skali λ , czyli przekształcenia dane wzorami:

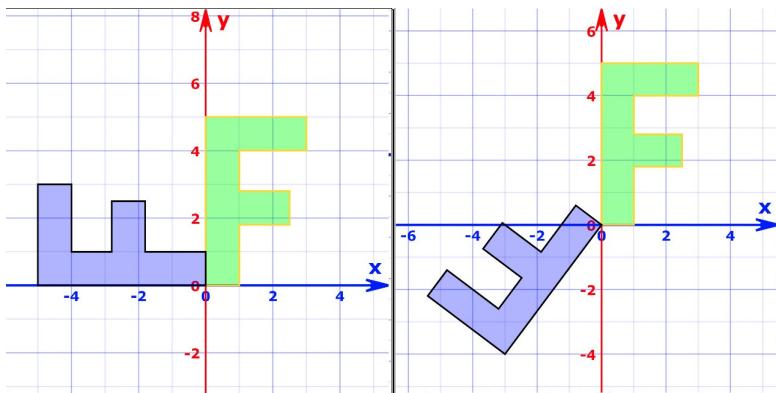
$$f(x, y) = (x, \lambda y), \quad g(x, y) = (\lambda x, y).$$



Rys. 4. Powinowactwa względem osi OX oraz OY o skali $\lambda = -2$.

- obroty o kąt θ względem punktu $(0, 0)$ dane wzorem:

$$f(x, y) = (\cos \theta x - \sin \theta y, \sin \theta x + \cos \theta y).$$



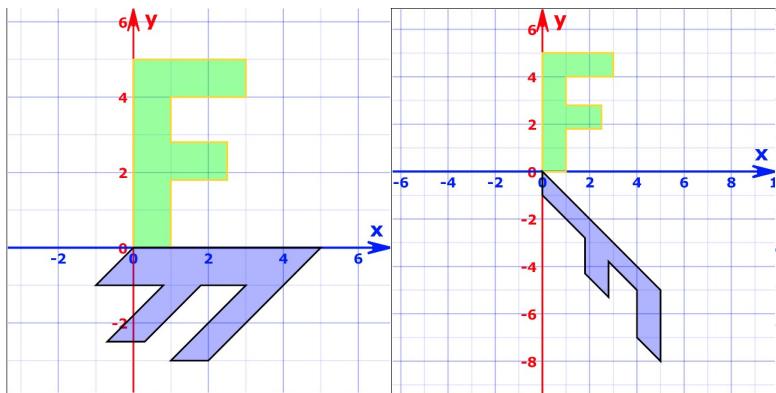
Rys. 5. Obroty o kąty $\pi/2$ oraz ?

- złożenia wyżej wymienionych w pewnej kolejności, np. dla przekształceń danych wzorami:

$$f(x, y) = (x + y, y), \quad g(x, y) = (y, -x),$$

mamy:

$$f(g(x, y)) = (-x + y, -x), \quad g(f(x, y)) = (y, -x - y)$$



Rys. 6. Funkcja $(x, y) \mapsto (-x + y, -x)$ (z lewej) oraz $(x, y) \mapsto (y, -x - y)$ (z prawej).

Zachęcam do zabawy apletem <https://www.mathsisfun.com/algebra/matrix-transform.html>. Należy oczywiście pamiętać, że obrazem przestrzeni wymiaru 2 przy przekształceniu liniowym nie musi być przestrzeń dwuwymiarowa, co można zobaczyć np. dla odwzorowania $(x, y) \mapsto (x + y, 2x + 2y)$. W drugim semestrze w odwzorowaniach opisanych wyżej rozpoznamy przekształcenia afinczne płaszczyzny \mathbb{R}^2 , które przeprowadzają punkt $(0, 0)$ na $(0, 0)$. Można je utożsamić z przekształceniami liniowymi. Póki co — zostańmy przy intuicjach pamiętając, że budujemy ogólną teorię „geometryczną”, nie tylko dla K^n .

Rozdział 16

Jądro i obraz. Monomorfizm, epimorfizm, izomorfizm

16.1 Wykład 16

Na poprzednim wykładzie wprowadziliśmy definicję przekształceń liniowych oraz wskazaliśmy ważne przykłady tych obiektów, także z geometrycznego punktu widzenia. Przekonamy się teraz, że można za pomocą tych pojęć wyrazić w nowym języku znane już fakty dotyczące przestrzeni liniowych. Wskażemy również ważne klasy przekształceń liniowych, pozwalających na utożsamianie przestrzeni o tej samej strukturze. Rozpoczniemy od prostego wniosku dotyczącego zachowania się podprzestrzeni przy przekształceniach liniowych.

Uwaga 16.1.1

Jeśli $\phi : V \rightarrow W$ jest przekształceniem liniowym. Wówczas:

- jeśli A jest podprzestrzenią V , to $\phi(A)$ jest podprzestrzenią W ,
- jeśli B jest podprzestrzenią W , to $\phi^{-1}(B)$ jest podprzestrzenią V .

Dowód. Pokażemy, że $\phi(A)$ jest podprzestrzenią W . Jeżeli $\beta_1, \beta_2 \in \phi(A)$, to istnieją $\alpha_1, \alpha_2 \in A$ takie, że $\beta_1 = \phi(\alpha_1)$ oraz $\beta_2 = \phi(\alpha_2)$. Skoro $\alpha_1 + \alpha_2 \in A$, to $\phi(\alpha_1 + \alpha_2) \in \phi(A)$. A zatem z definicji przekształcenia liniowego mamy $\phi(\alpha_1) + \phi(\alpha_2) \in \phi(A)$. Podobnie pokazujemy, że jeśli $\lambda \in K$ oraz $\beta = \phi(\alpha)$, dla pewnego $\alpha \in A$, to oczywiście $\lambda\alpha \in A$, czyli $\phi(\lambda\alpha) = \lambda\phi(\alpha) \in \phi(A)$. A zatem $\phi(A)$ jest podprzestrzenią W . Analogicznie pokazujemy, że $\phi^{-1}(B)$ jest podprzestrzenią V . \square

Szczególnie istotna jest sytuacja, gdy mowa jest o obrazie całej przestrzeni V zawartym w W oraz o przeciwbieżnie przestrzeni zerowej zawartym w V . Zgodnie z powyższym wnioskiem są to podprzestrzenie.

Definicja 16.1.2: Jądro i obraz przekształcenia liniowego

Niech $\phi : V \rightarrow W$ będzie przekształceniem liniowym.

- **JĄDREM** przekształcenia ϕ nazywamy zbiór $\ker(\phi) = \{\alpha \in V \mid \phi(\alpha) = 0\} \subseteq V$.
- **OBRAZEM** przekształcenia ϕ nazywamy zbiór $\text{im}(\phi) = \{\phi(\alpha) \mid \alpha \in V\} \subseteq W$.

Oczywiście jądro i obraz są podprzestrzeniami, odpowiednio $\ker(\phi) = \phi^{-1}(\{0\})$ oraz $\text{im}(\phi) = \phi(V)$. Zobaczmy kilka przykładów.

- Jeśli przekształcenie $\phi : \mathbb{R}^3 \rightarrow \mathbb{R}$ dane jest wzorem $\phi((x_1, x_2, x_3)) = x_1 + x_2$, to

$$\text{im}(\phi) = \mathbb{R} \quad \text{oraz} \quad \ker(\phi) = \{(x_1, x_2, x_3) \in \mathbb{R}^3 : x_1 + x_2 = 0\}.$$

- Jeśli $\psi : \mathbb{R} \rightarrow \mathbb{R}^3$ dane jest wzorem $\psi(x) = (x, x, x)$, to $\ker(\psi) = \{0\}$ oraz $\text{im}(\psi) = \text{lin}((1, 1, 1))$.

- Niech $\phi : V \rightarrow V$ będzie homotetią, przy czym $\phi(v) = av$, dla każdego $v \in V$ oraz pewnego ustalonego $a \in K$. Wówczas:

$$\ker(\phi) = \begin{cases} \{0\}, & a \neq 0 \\ V, & a = 0 \end{cases}, \quad \text{im}(\phi) = \begin{cases} V, & a \neq 0 \\ \{0\}, & a = 0 \end{cases}.$$

- Niech $\phi : V \rightarrow V$ będzie rzutem na V_1 wzduż V_2 . Wówczas $\ker(\phi) = V_2$ oraz $\text{im}(\phi) = V_1$.
- Niech $\phi : \mathbb{R}[x] \rightarrow \mathbb{R}[x]$ jest pochodną, to: $\ker(\phi) = \{w \in \mathbb{R}[x] : \deg(w) \leq 0\}$ oraz $\text{im}(\phi) = \mathbb{R}[x]$.

Kluczowy przykład

Niech $a_{ij} \in K$, gdzie $1 \leq i \leq m, 1 \leq j \leq n$. Niech $\phi : K^n \rightarrow K^m$ będzie przekształceniem liniowym postaci:

$$\phi((x_1, \dots, x_n)) = (a_{11}x_1 + \dots + a_{1n}x_n, \dots, a_{m1}x_1 + \dots + a_{mn}x_n). \quad (\dagger)$$

Wówczas $\ker(\phi) \subseteq K^n$ jest zbiorem rozwiązań jednorodnego układu równań postaci:

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n = 0 \\ \dots \\ a_{m1}x_1 + a_{m2}x_2 + \dots + a_{mn}x_n = 0. \end{cases} \quad (\ddagger)$$

A przestrzeń $\text{im}(\phi) \subseteq K^m$? Jest to w istocie przestrzeń kolumnowa macierzy $A = [a_{ij}]$. Mamy:

$$\text{lin}(\phi(\epsilon_1), \dots, \phi(\epsilon_n)) = \text{lin}((a_{11}, \dots, a_{1n}), \dots, (a_{m1}, \dots, a_{mn})).$$

Rzeczywiście, $\text{im}(\phi)$ jest rozpięta przez obrazy wektorów bazy standardowej przestrzeni K^n . Innymi słowy $\text{im}(\phi)$ rozpięty jest przez kolumny macierzy układu (\ddagger) zapisanego wyżej.

Z twierdzenia Kroneckera-Capellego wynika ważny wniosek dotyczący dowolnego przekształcenia liniowego $f : K^n \rightarrow K^m$.

$$n = \dim(K^n) = n - r(A) + r(A) = \dim \ker(f) + \dim \text{im}(f).$$

Zauważmy, że wymiar obrazu dowolnego przekształcenia liniowego można wyznaczyć korzystając z następującej uwagi.

Obserwacja 16.1.3: Obraz jest rozpięty przez obrazy układu rozpinającego

Niech $V = \text{lin}(\alpha_1, \dots, \alpha_n)$ oraz niech $f : V \rightarrow W$ będzie przekształceniem liniowym. Wówczas

$$\text{im}(\phi) = \text{lin}(\phi(\alpha_1), \dots, \phi(\alpha_n)).$$

Odnoszącmy, że we wzorze powyżej $\dim \text{im}(f)$ równe jest rzędowi macierzy rozważanego układu. Stąd bierze nazwę poniższa definicja.

Definicja 16.1.4: Rząd przekształcenia liniowego

Wymiar przestrzeni $\text{im}(\phi)$ nazywamy RZĘDEM PRZEKSZTAŁCENIA, ozn. $r(\phi)$.

Przejdziemy teraz do twierdzenia będącego uzasadnieniem poczynionej wyżej obserwacji o wymiarach.

Twierdzenie 16.1.5: Wymiary obrazu i jądra przekształcenia liniowego

Niech $\phi : V \rightarrow W$ będzie przekształceniem liniowym. Wówczas:

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

Dowód. Pokażemy najpierw następujący istotny fakt, wiążący tezę twierdzenia z rozkładem przestrzeni liniowej V na sumę prostą.

Uwaga 16.1.6

Niech $\phi : V \rightarrow W$ będzie przekształceniem liniowym. Niech U będzie taką podprzestrzenią przestrzeni V , że

$$V = \ker(\phi) \oplus U.$$

Niech $\alpha_1, \dots, \alpha_k$ będzie bazą U . Wówczas układ $\phi(\alpha_1), \dots, \phi(\alpha_k)$ jest bazą przestrzeni $\text{im}(\phi)$.

Pokażemy, że dla każdego dopełnienia prostego przestrzeni $\ker(\phi)$ i każdej jego bazy $\alpha_1, \dots, \alpha_k$ zbiór wektorów $\phi(\alpha_1), \dots, \phi(\alpha_k)$ rozpina $\text{im}(\phi)$. Następnie pokażemy, że układ ten jest liniowo niezależny.

Niech $\beta \in \text{im}(\phi)$. Chcemy pokazać, że $\beta \in \text{lin}(\phi(\alpha_1), \dots, \phi(\alpha_k))$. Wiadomo, że $\beta = \phi(\alpha) = \phi(\alpha' + \alpha'')$, gdzie $\alpha' \in \ker(\phi)$ oraz $\alpha'' \in U$. A zatem $\alpha'' = a_1\alpha_1 + \dots + a_k\alpha_k$, dla pewnych $a_1, \dots, a_k \in K$. Zatem:

$$\begin{aligned}\beta &= \phi(\alpha) = \phi(\alpha' + \alpha'') \\ &= \phi(\alpha') + \phi(a_1\alpha_1 + \dots + a_k\alpha_k) \\ &= 0 + a_1\phi(\alpha_1) + \dots + a_k\phi(\alpha_k) \\ &\in \text{lin}(\phi(\alpha_1), \dots, \phi(\alpha_k)).\end{aligned}$$

Dowodzimy liniowej niezależności tego układu. Przypuśćmy, że $a_1\phi(\alpha_1) + \dots + a_k\phi(\alpha_k) = 0$. Wówczas $\phi(a_1\alpha_1 + \dots + a_k\alpha_k) = 0$, a zatem $a_1\alpha_1 + \dots + a_k\alpha_k \in \ker(\phi)$. Ale przecież $\alpha_1, \dots, \alpha_k$ jest bazą U . A zatem $a_1\alpha_1 + \dots + a_k\alpha_k \in \ker(\phi) \cap U = \{0\}$. A szczególnie $a_1\alpha_1 + \dots + a_k\alpha_k = 0$, czyli $a_1 = \dots = a_k = 0$, bo $\alpha_1, \dots, \alpha_k$ jest bazą U . Układ $\phi(\alpha_1), \dots, \phi(\alpha_k)$ jest zatem liniowo niezależny.

Dowód Twierdzenia 16.1.5 jest teraz natychmiastowy. Na mocy twierdzenia o wymiarze sumy prostej (wniosek z formuły Grassmanna) mamy:

$$\dim(V) = \dim \ker(\phi) + \dim(U) = \dim \ker(\phi) + k = \dim \ker(\phi) + \dim \text{im}(\phi).$$

□

Rezultat nasz ma sens również w przypadku, gdy V jest przestrzenią nieskończonym wymiaru. Dowód wymaga pewnej modyfikacji, ale w rezultacie okazuje się, że jeśli $\phi : V \rightarrow W$ jest liniowe i $\dim(V) = \infty$, to wymiary przestrzeni $\ker(\phi)$ oraz $\text{im}(\phi)$ nie mogą być jednocześnie skończenie wymiarowe.

Patrząc na formułę wiążącą wymiar jądra i obrazu przekształcenia liniowego warto pochylić się dłużej nad przypadkami, gdy $\dim \ker(\phi) = \{0\}$ oraz, gdy $\dim \text{im}(\phi) = \dim W$.

Definicja 16.1.7: Monomorfizm, epimorfizm, izomorfizm

Przekształcenie liniowe $\phi : V \rightarrow W$ nazywamy:

- MONOMORFIZMEM, gdy ϕ jest różnowartościowe, tzn. $\phi(\alpha) = \phi(\beta) \Rightarrow \alpha = \beta$, dla $\alpha, \beta \in V$.
- EPIMORFIZMEM, gdy jest „na”, tzn. gdy dla każdego $\gamma \in W$ istnieje $\alpha \in V$ takie, że $\phi(\alpha) = \gamma$.
- IZOMORFIZMEM, gdy ϕ jest różnowartościowe i „na” (to znaczy, gdy ϕ jest bijekcją).

Przykłady

- Przekształcenie liniowe $\phi : \mathbb{R} \rightarrow \mathbb{R}^3$ dane wzorem $\phi(x) = (x, x, x)$, jest monomorfizmem, ale nie jest epimorfizmem.
- Przekształcenie liniowe $\phi : \mathbb{R}^3 \rightarrow \mathbb{R}$ dane wzorem $\phi(x, y, z) = x$ jest epimorfizmem, ale nie jest monomorfizmem.
- Przekształcenie liniowe $\phi : \mathbb{R}^4 \rightarrow M_{2 \times 2}(\mathbb{R})$ dane poniższym wzorem jest izomorfizmem:
$$\phi((x_1, x_2, x_3, x_4)) = \begin{bmatrix} x_1 & x_2 \\ x_3 & x_4 \end{bmatrix}.$$

Uwaga 16.1.8

Niech $\phi : V \rightarrow W$ będzie przekształceniem liniowym. Wówczas:

- ϕ jest monomorfizmem wtedy i tylko wtedy, gdy $\ker(\phi) = \{0\}$,
- ϕ jest epimorfizmem wtedy i tylko wtedy, gdy $\text{im}(\phi) = W$.

Dowód. Tylko pierwsza równoważność wymaga dowodu. Jeśli ϕ jest monomorfizmem oraz dla pewnego $\alpha \in V$ mamy $\phi(\alpha) = 0$, to skoro $\phi(0) = 0$, z równowartościowości ϕ wynika, że $\alpha = 0$. A zatem $\ker(\phi) = 0$. Na odwrót: jeśli $\ker(\phi) = \{0\}$ oraz dla pewnych $\alpha, \beta \in V$ mamy $\phi(\alpha) = \phi(\beta)$, to z liniowości $\phi(\alpha - \beta) = 0$. Skoro $\ker(\phi) = \{0\}$, to $\alpha - \beta = 0$, czyli $\alpha = \beta$. W szczególności ϕ to monomorfizm. \square

Dla przekształcenia liniowego $\phi : K^n \rightarrow K^m$ danego wzorem (\dagger) wnioskujemy, że jest ono monomorfizmem, gdy odpowiadający mu jednorodny układ równań liniowych $(\dagger\dagger)$ ma jedynie rozwiązanie będące wektorem zerowym w K^n , a jest ono epimorfizmem, gdy rzad macierzy $m \times n$ tego układu równy jest m .

Wniosek 16.1.9

Niech $\phi : V \rightarrow W$ będzie przekształceniem liniowym, przy czym $\dim(V), \dim(W) < \infty$. Wówczas:

- jeśli ϕ jest monomorfizmem, to $\dim V \leq \dim W$,
- jeśli ϕ jest epimorfizmem, to $\dim W \leq \dim V$,
- jeśli ϕ jest izomorfizmem, to $\dim W = \dim V$.

Co ważne, tezę punktu trzeciego można dla przestrzeni skończenie wymiarowych łatwo odwrócić.

Wniosek 16.1.10

Niech $\phi : V \rightarrow W$ będzie przekształceniem liniowym i niech $\dim V = \dim W < \infty$. Wówczas następujące warunki są równoważne:

- (a) ϕ jest monomorfizmem,
- (b) ϕ jest epimorfizmem,
- (c) ϕ jest izomorfizmem.

Definicja 16.1.11: Izomorfizm przestrzeni liniowych

Mówimy, że przestrzenie V i W nad ciałem K są IZOMORFICZNE, jeśli istnieje izomorfizm $\phi : V \rightarrow W$. Oznaczenie: $V \simeq W$.

To właśnie izomorfizm przestrzeni liniowych jest pojęciem, które mówi o tym, że jakieś dwie przestrzenie są „jednakowe” z punktu widzenia algebry liniowej, czyli mają tą samą strukturę. Co to znaczy jednakowe? Przytoczymy teraz kilka rezultatów, które o tym mówią.

Twierdzenie 16.1.12

Niech $\phi : V \rightarrow W$ będzie przekształceniem liniowym. Następujące warunki są równoważne:

- (a) ϕ jest izomorfizmem,
- (b) ϕ przeprowadza każdą bazę przestrzeni V na bazę przestrzeni W ,
- (c) ϕ przeprowadza pewną bazę przestrzeni V na bazę przestrzeni W .

Dowód. Pokażemy tezę w sytuacji, gdy $\dim(V) < \infty$. Ogólny przypadek uzasadnia się analogicznie.

Wiemy już, że dla dowolnego przekształcenia liniowego $\phi : V \rightarrow W$ i każdej podprzestrzeni U w W takiej, że $V = \ker(\phi) \oplus U$ przekształcenie ϕ przeprowadza bazę przestrzeni U na bazę przestrzeni $\text{im}(\phi)$. Jeśli ϕ jest izomorfizmem, to $\text{im}(\phi) = W$, a także $\ker(\phi) = \{0\}$, więc $V = U$. Zatem ϕ przeprowadza bazę przestrzeni V na bazę przestrzeni W . Pokazaliśmy $(i) \Rightarrow (ii)$. Implikacja $(ii) \Rightarrow (iii)$ jest oczywista.

Przypuśćmy, że pewne przekształcenie liniowe ϕ przeprowadza bazę $\alpha_1, \dots, \alpha_n$ przestrzeni V na bazę β_1, \dots, β_n przestrzeni W . Pokażemy, że ϕ jest izomorfizmem czyli, że $\ker(\phi) = \{0\}$ oraz $\text{im}(\phi) = W$.

Jeśli $\alpha = a_1\alpha_1 + \dots + a_n\alpha_n \in \ker(\phi)$, to $0 = \phi(\alpha) = a_1\beta_1 + \dots + a_n\beta_n$, co wobec liniowej niezależności układu $(\beta_1, \dots, \beta_n)$ oznacza, że $a_1 = \dots = a_n = 0$. A zatem $\alpha = 0$. A zatem wobec dowolności wyboru α mamy $\ker(\phi) = \{0\}$.

Weźmy $\beta \in W$ i niech $\beta = a_1\beta_1 + \dots + a_n\beta_n$. Wówczas $\beta = \phi(\alpha)$, dla pewnego $\alpha = a_1\alpha_1 + \dots + a_n\alpha_n$. A zatem z dowolności wyboru β mamy $W = \text{im}(\phi)$. \square

Rezultat ten pozwala nam udowodnić kluczowy wniosek.

Wniosek 16.1.13: Izomorfizm z K^n

Niech V, W będą przestrzeniami skończonego wymiaru nad ciałem K . Wówczas następujące warunki są równoważne:

- (i) $V \simeq W$,
- (ii) $\dim V = \dim W$,

W konsekwencji, mamy izomorfizm $V \simeq K^{\dim V}$.

Implikacja $(i) \Rightarrow (ii)$ została pokazana już wcześniej w oparciu o formułę

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

gdzie $\phi : V \rightarrow W$ jest izomorfizmem. Mamy bowiem $\ker(\phi) = 0$ oraz $\text{im}(\phi) = W$. Uzyskujemy zatem $\dim V = \dim W$, czyli (ii).

Implikacja $(ii) \Rightarrow (i)$ wymaga twierdzenia o jednoznaczny definiowaniu na bazie z poprzedniego rozdziału. Niech $\alpha_1, \dots, \alpha_n$ będzie bazą V oraz niech β_1, \dots, β_n będzie bazą W . Definiujemy $\phi(V)$ warunkiem $\phi(\alpha_i) = \beta_i$, dla $1 \leq i \leq n$. Wiadomo, że takie przekształcenie istnieje dla każdego układu wektorów W równoliczniego z bazą $\alpha_1, \dots, \alpha_n$. Takie przekształcenie ϕ , które wybraliśmy, musi być jednak izomorfizmem, bo przeprowadza bazę V na bazę W .

A zatem pokazaliśmy, że każda przestrzeń n -wymiarowa nad ciałem K jest izomorficzna z przestrzenią liniową K^n . Izomorfizmy te można uzyskać na wiele sposobów, o czym powiemy następnym razem.

Kończymy wstępnią część naszych rozważań przykładem odsyłającym nas w daleką na razie przyszłość. Rozważmy przekształcenie liniowe $\phi : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ dane wzorem:

$$\phi((x_1, x_2, x_3)) = (2x_1, 2x_2, 4x_3).$$

Z geometrycznego punktu widzenia przekształcenie to nie jest ani obrotem, ani symetrią czy rzutem, ale biorąc rozkład:

$$\mathbb{R}^3 = \text{lin}((1, 0, 0), (0, 1, 0)) \oplus \text{lin}(0, 0, 1) \quad (*)$$

widzimy, że ϕ ograniczone do każdego ze składników jest na nim homotetią, bo:

- dla każdego $v \in \text{lin}((1, 0, 0), (0, 1, 0))$ mamy $\phi(v) = 2v$,
- dla każdego $w \in \text{lin}(0, 0, 1)$ mamy $\phi(w) = 4w$.

Idea jest zatem następująca: znając ϕ na każdym ze składników prostych, „wiemy co robi” ϕ na całej przestrzeni liniowej! To nie przypadek, ale zwiastun wielkiej i ważnej teorii, którą zajmiemy się w kolejnym semestrze. Kluczową kwestią jest wskazywanie rozkładów takich, jak (*), dla innych przekształceń liniowych. Na razie jednak zajmiemy się zbudowaniem podstaw, zwłaszcza opisem przekształceń liniowych pomiędzy przestrzeniami skończonego wymiaru w języku macierzowym.

16.2 Wybór przykładowych pytań

Proszę się nie uczyć na pamięć odpowiedzi czy wyjaśnienie! Takich pytań można ułożyć znacznie więcej i służą one jedynie sprawdzeniu podstawowego zrozumienia treści zajęć. Jeśli odpowiedź na dowolne z poniższych pytań sprawia Państwu problem, proszę pomyśleć który fragment notatek dotyczy tego pytania, wrócić do ich lektury, a w razie dalszych trudności poprosić prowadzącego o wskazówkę.

- Przekształcenie liniowe $\phi : \mathbb{R}^4 \rightarrow \mathbb{R}^2$ dane jest wzorem

$$\phi((x_1, x_2, x_3, x_4)) = (x_1 - x_2, x_1 + x_2 + x_3).$$

- Czy wektor $(1, 1, 0, 0)$ należy do $\ker \phi$?
- Czy wektor $(0, 0, 0, 1)$ należy do $\ker \phi$?
- Czy wektor $(0, 0)$ należy do $\text{im } \phi$?
- Czy wektor $(1, 1)$ należy do $\text{im } \phi$?

- Podaj przykład przekształcenia liniowego $\phi : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ takiego, że

$$\ker \phi = \text{lin}((1, 0)) \quad \text{oraz} \quad \text{im } \phi = \text{lin}((1, 0)).$$

- Niech V będzie przestrzenią liniową nad ciałem \mathbb{K} , zaś $\phi : V \rightarrow \mathbb{K}$ niezerowym przekształceniem liniowym. Czy ϕ może nie być "na"?

- Niech $\phi : \mathbb{R}^3 \rightarrow \mathbb{R}$ będzie przekształceniem liniowym. Jaki jest możliwy wymiar przestrzeni $\ker \phi$?

- Niech $\phi : \mathbb{R}^3 \rightarrow \mathbb{R}^5$ będzie przekształceniem liniowym. Jaki jest możliwy wymiar przestrzeni $\text{im } \phi$?

- Założmy, że $\phi : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ spełnia $\text{im}(\phi) = \mathbb{R}^3$. Czy istnieją parami różne wektory $\alpha, \beta, \gamma, \delta \in \mathbb{R}^3$, że $\phi(\alpha + \beta + \gamma + \delta) = 0$?

- Czy istnieje przekształcenie liniowe $\phi : \mathbb{R}^5 \rightarrow \mathbb{R}^5$ takie, że $\text{im } \phi = \ker \phi$?

- Czy istnieje przekształcenie liniowe $\phi : K^{10} \rightarrow K^{11}$, że $\dim \ker \phi = 2$, $\dim \text{im } \phi = 9$?

- Czy istnieje epimorfizm $\phi : M_{3 \times 2}(\mathbb{R}) \rightarrow \mathbb{R}_{\leq 6}[x]$?

- Dane są przekształcenia liniowe $\phi_1, \phi_2 : \mathbb{R}^5 \rightarrow \mathbb{R}^3$. Czy musi istnieć niezerowy wektor $\alpha \in \mathbb{R}^5$ taki, że $\phi_1(\alpha) = \phi_2(\alpha)$?

- Która z poniższych przestrzeni liniowych jest izomorficzna z przestrzenią $M_{4 \times 6}(\mathbb{R})$ macierzy rzeczywistych rozmiaru 4×6 nad ciałem \mathbb{R} ?

- $\mathbb{R}_{\leq 23}[x]$ — przestrzeń wielomianów rzeczywistych stopnia nie większego niż 23
- $M_{12 \times 2}(\mathbb{R})$ — przestrzeń macierzy rzeczywistych rozmiaru 12×2 ,
- przestrzeń rozwiązań równania liniowego o 25 zmiennych postaci $x_1 + x_2 + \dots + x_{25} = 0$ o współczynnikach rzeczywistych?

- Która z poniższych przestrzeni jest izomorficzna z przestrzenią liniową \mathbb{Z}_2^{16} ?

- $M_{4 \times 4}(\mathbb{Z}_2)$ — przestrzeń macierzy rozmiaru 4×4 o wyrazach w ciele \mathbb{Z}_2 ,
- $M_{1 \times 16}(\mathbb{Z}_2)$ — przestrzeń macierzy rozmiaru 1×16 o wyrazach w ciele \mathbb{Z}_2
- przestrzeń podzbiorów zbioru czteroelementowego nad ciałem \mathbb{Z}_2

- Która z poniższych przestrzeni liniowych jest izomorficzna z przestrzenią rozwiązań równania liniowego o 25 zmiennych postaci $x_1 + x_2 + \dots + x_{25} = 0$ nad ciałem \mathbb{C} ?

- $\mathbb{C}_{\leq 25}[x]$ — przestrzeń wielomianów stopnia ≤ 25 o współczynnikach zespolonych
- $M_{5 \times 5}(\mathbb{C})$ — przestrzeń macierzy rozmiaru 5×5 o wyrazach zespolonych
- \mathbb{C}^{25} ?

- Niech V będzie przestrzenią liniową nad ciałem \mathbb{Z}_2 . Czy V może mieć 6 elementów?

- Założymy, że $\dim V = \infty$. Niech $\phi : V \rightarrow V$ będzie przekształceniem liniowym, spełniającym warunek $\ker(\phi) = 1$. Uzasadnij, że przestrzenie liniowe V oraz $\text{im } (\phi)$ są izomorficzne.

16.3 Zadania do samodzielnej pracy

Zadania ilustrujące umiejętności niezbędne do zaliczenia przedmiotu oznaczone są symbolem ♠.

Przy tych zadaniach dodany jest opis umiejętności, które sprawdzają.

- Czy istnieje (jeśli tak, podaj przykład) przekształcenie liniowe $\phi : \mathbb{R}^4 \rightarrow \mathbb{R}^3$ spełniające warunki:

- (a) $\ker(\phi) = \{(x_1, x_2, x_3, x_4) : x_1 - x_2 + 6x_3 + 2x_4 = 0\}$, $\text{im}(\phi) = \text{lin}((2, 3, 1))$.
- (b) $\ker(\phi) = \text{lin}((1, 0, 3, 3))$, $\text{im}(\phi) = \{(x_1, x_2, x_3) : 4x_1 + 5x_2 - x_3 = 0\}$.
- (c) $\ker(\phi) = \text{lin}((1, 1, 1, 1), (1, 1, 1, 0))$, $\text{im}(\phi) = \text{lin}((1, 1, 1), (1, 1, 0))$.

- (♠) Znajdowanie bazy oraz wymiaru obrazu i jądra przekształcenia liniowego)

Dla każdego z przekształceń znajdź bazę i wymiar jego obrazu oraz bazę i wymiar jego jądra.

- (a) $\phi : \mathbb{R}^3 \rightarrow \mathbb{R}^2$ dane wzorem

$$\phi((x_1, x_2, x_3)) = (2x_1 + x_2 - 3x_3, x_1 + 4x_2 + 2x_3),$$

- (b) $\phi : \mathbb{R}^3 \rightarrow \mathbb{R}^4$ dane wzorem

$$\phi((x_1, x_2, x_3)) = (4x_1 + 3x_2 + 5x_3, x_1 + 2x_2 + x_3, 2x_1 - x_2 + 3x_3, 6x_1 + 7x_2 + 7x_3).$$

- (♠) Stwierdzanie, kiedy przekształcenie liniowe jest monomorfizmem, epimorfizmem, izomorfizmem).

Dla jakich wartości parametru $r \in \mathbb{R}$ przekształcenie

- (a) $\phi : \mathbb{R}^3 \rightarrow \mathbb{R}^5$ dane wzorem

$$\phi((x_1, x_2, x_3)) = (x_1 + x_2 + 2x_3, 2x_1 + x_2 + x_3, x_1 + 3x_2 + rx_3, 5x_1 + 3x_2 + 4x_3, x_1 + 2x_2 + 5x_3)$$

jest monomorfizmem?

- (b) $\phi : \mathbb{R}^4 \rightarrow \mathbb{R}^3$ dane wzorem

$$\phi((x_1, x_2, x_3, x_4)) = (4x_1 + x_2 + rx_3 + x_4, 3x_1 + 2x_2 + x_3 + x_4, 2x_1 + 3x_2 + 3x_3 + x_4)$$

jest epimorfizmem?

- (c) $\phi : \mathbb{R}^4 \rightarrow \mathbb{R}^4$ dane wzorem

$$\phi((x_1, x_2, x_3, x_4)) = (5x_1 - x_2 + rx_3 + 5x_4, 2x_1 - 3x_2 - 6x_3 + rx_4, 3x_1 + 2x_2 + x_3 + 4x_4, x_1 + 5x_2 + 7x_3 + 3x_4)$$

jest izomorfizmem?

- Niech $T : M_{2 \times 2}(\mathbb{R}) \rightarrow M_{2 \times 2}(\mathbb{R})$ będzie przekształceniem określonym wzorem

$$T \left(\begin{bmatrix} a & b \\ c & d \end{bmatrix} \right) = \begin{bmatrix} a+c & b+d \\ a+c & b+d \end{bmatrix}.$$

Wykaż, że T jest przekształceniem liniowym. Znajdź wymiary obrazu i jądra tego przekształcenia.

- Niech $\mathbb{R}_{\leq n}[x]$ będzie podprzestrzenią przestrzeni wielomianów złożoną z wielomianów stopnia $\leq n$.

Niech $f : \mathbb{R}_{\leq n}[x] \rightarrow \mathbb{R}_{\leq n}[x]$ będzie funkcją zadaną wzorem $f(w(x)) = w(x+1) - w(x)$. Wykaż, że f jest przekształceniem liniowym, znajdź bazy i wymiary jądra oraz obrazu f .

- Niech V będzie przestrzenią liniową. Wykaż, że V jest nieskończenie wymiarowa wtedy i tylko wtedy, gdy V zawiera podprzestrzeń W taką, że $W \simeq V$ i $W \neq V$.

- Niech V będzie skończenie wymiarową przestrzenią liniową i niech $\phi_1, \phi_2 : V \rightarrow V$ będą przekształceniami liniowymi.

- (a) Udowodnij, że $\ker \phi_1 \cap \ker \phi_2 \subseteq \ker \phi_1 + \ker \phi_2$. Kiedy zachodzi równość?

- (b) Niech $\text{im } \phi_1 + \text{im } \phi_2 = V = \ker \phi_1 + \ker \phi_2$. Wykaż, że $\text{im } \phi_1 \cap \text{im } \phi_2 = \{0\} = \ker \phi_1 \cap \ker \phi_2$.

- Niech V_0, \dots, V_{n+1} będą przestrzeniami liniowymi nad ciałem K , przy czym $V_0 = V_{n+1} = \{0\}$. Niech $f_i : V_i \rightarrow V_{i+1}$, dla $i = 0, \dots, n$ będzie ciągiem przekształceń liniowych, takim że $\text{im } f_i = \ker f_{i+1}$. Wykaż, że

$$\sum_{i=1}^n (-1)^i \dim V_i = 0$$

16.4 Uzupełnienie. Kojądro. Twierdzenie o homomorfizmie.

W tym dodatku, i kolejnych dotyczących przekształceń liniowych, wprowadzać będziemy kolejne elementy teorii przekształceń liniowych związanych z przestrzeniami ilorazowymi.

Definicja 16.4.1: Naturalne rzutowanie

Odwzorowanie $\pi : V \rightarrow V/U$ przypisujące każdemu elementowi $v \in V$ warstwę $v + U$ nazywane jest NATURALNYM RZUTOWANIEM na V/U . Obraz $v \in V$ jest często oznaczany w skrócie jako \bar{v} .

Przykład. Biorąc $\pi : \mathbb{R}^2 \rightarrow \mathbb{R}^2 / \text{lin}(0, 1)$ mamy: $\overline{(1, 0)} = \overline{(1, 1)}$.

Uwaga 16.4.2

Niech U, W będą takimi podprzestrzeniami przestrzeni liniowej V , że $V = U \oplus W$. Wówczas przekształcenie $\pi : W \rightarrow V/U$ przypisujące każdemu $w \in W$ warstwę $w + U$ jest izomorfizmem.

Dowód. Zauważmy, że każdy wektor $v \in V$ może być zapisany w postaci $v = u + w$, gdzie $u \in U$ oraz $w \in W$. Zatem $v + U = u + w + U = w + U = \pi(w)$. W rezultacie każdy element przestrzeni V/U należy do $\text{im}(\pi)$ i przekształcenie to jest epimorfizmem. To, że π jest monomorfizmem uzasadniamy w następujący sposób. Jeśli $\pi(w) = \pi(w')$, dla pewnych $w, w' \in W$, to $w + U = w' + U$. Stąd $w - w' \in U$, a stąd $w - w' \in U \cap W = \{0\}$, gdyż $V = U \oplus W$. Zatem $w = w'$. \square

Definicja 16.4.3: Kojądro

Niech $\phi : V \rightarrow W$ będzie przekształceniem liniowym. Przestrzeń ilorazowa $W / \text{im } \phi$ nazywamy KOJĄDREM ϕ i oznaczamy przez $\text{coker } \phi$.

Oczywiście przekształcenie liniowe jest epimorfizmem wtedy i tylko wtedy, gdy jego kojądro jest zerowe.

Twierdzenie 16.4.4: O homomorfizmie

Jeśli $\phi : V \rightarrow W$ jest przekształceniem liniowym, to odwzorowanie $\bar{\phi} : V / \ker \phi \rightarrow \text{im } \phi$ przypisujące warstwie $v + \ker \phi$ element $\phi(v)$ jest dobrze określone i jest izomorfizmem przestrzeni liniowych.

Dowód. Oczywiście $\bar{\phi}$ jest dobrze określone, ponieważ jeśli $v + \ker \phi = v' + \ker \phi$, to $v - v' \in \ker \phi$, a stąd $\phi(v) = \phi(v')$. To, że $\bar{\phi}$ jest monomorfizmem jest jasne, gdyż jeśli $v + \ker \phi$ należy do $\ker \bar{\phi}$, to $\phi(v) = 0$, czyli $v + \ker \phi$ jest warstwą zerową. Oczywiście $\bar{\phi}$ jest na. \square

Powyższe twierdzenie daje nam natychmiast alternatywny dowód jednego z punktów twierdzenia Kroneckera-Capellego, tylko w nieco ogólniejszym kontekście.

Wniosek 16.4.5

Niech $\phi : V \rightarrow W$ będzie przekształceniem liniowym i niech $w_0 \in W$. Jeśli v_0 jest pewnym rozwiązaniem równania $\phi(v) = w_0$, gdzie $v \in V$, to każde rozwiązanie równania $\phi(v) = w_0$ należy do zbioru $v_0 + \ker \phi$.

Dowód. Element $v \in V$ spełnia $\phi(v) = w_0$ wtedy i tylko wtedy, gdy $\bar{\phi}(v + \ker \phi) = \bar{\phi}(v_0 + \ker \phi)$. Skoro $\bar{\phi}$ jest 1-1, to powyższa równość zachodzi wtedy i tylko wtedy, gdy $v + \ker \phi = v_0 + \ker \phi$. Oczywiście wszystkie wektory v spełniające powyższy warunek stanowią warstwę $v_0 + \ker \phi$. \square

Zauważmy, że rozważając podprzestrzeń $C^1(\mathbb{R}, \mathbb{R})$ przestrzeni $F(\mathbb{R}, \mathbb{R})$ złożoną z funkcji mających pochodną na całej prostej możemy rozważyć przekształcenie liniowe $d : C^1(\mathbb{R}, \mathbb{R}) \rightarrow F(\mathbb{R}, \mathbb{R})$ przypisujące funkcji jej pochodną. Wiadomo, że jądro d składa się jedynie z funkcji stałych. W ten sposób uzyskany rezultat mówi, że jeśli F jest dowolną funkcją taką, że $F' = f_0$, to każda inna funkcja, której pochodną jest f_0 ma postać $F + c$, gdzie c jest stałą.

16.5 Dodatek. Różniczkowanie i pierwiastki wielokrotne

W tym dodatku przyjrzymy się nieco bliżej algebraicznym własnościom przekształcenia liniowego przestrzeni wielomianów $K[x]$ zwanego pochodną lub różniczkowaniem. Jest to, jak wiemy, przekształcenie liniowe $d : K[x] \rightarrow K[x]$, zadane wzorem

$$\phi(a_0 + a_1x + \dots + a_nx^n) = a_1 + 2a_2x + 3a_3x^2 + \dots + na_nx^{n-1}.$$

Jak się okaże, przekształcenie to ma związek z problemem istnienia pierwiastka wielokrotnego wielomianu f . Przypomnijmy, że element $a \in K$ jest pierwiastkiem m -krotnym $f \in K[x]$ wtedy i tylko wtedy, gdy $f = (x - a)^m g$, gdzie $g \in K[x]$ oraz $g(a) \neq 0$.

Obserwacja 16.5.1: Jądro różniczkowania

Niech $d : K[x] \rightarrow K[x]$ będzie różniczkowaniem. Wówczas

$$\ker d = \begin{cases} K, & \text{jeśli charakterystyka } K \text{ równa jest } 0 \\ K[x^p], & \text{jeśli charakterystyka } K \text{ równa jest } p > 0 \end{cases},$$

przy czym przez $K[x^p]$ rozumiemy zbiór sum postaci $\sum_{i=0}^n a_i x^{ip}$, gdzie $a_i \in K$.

Dowód. Warunek $w = a_0 + a_1x + \dots + a_nx^n \in \ker d$ jest równoważny temu, że dla każdego i spełniającego $a_i \neq 0$ mamy $ix_{i-1} = 0$. W przypadku, gdy charakterystyka ciała K jest 0, mamy $i = 0$. Wtedy w jest wielomianem stałym. Gdy charakterystyka K jest równa $p > 0$, $w \in \ker d$ wtedy i tylko wtedy, gdy p jest dzielnikiem i , dla każdego i spełniającego $a_i \neq 0$. To jest równoważne stwierdzeniu, że $w \in K[x^p]$. \square

Obserwacja 16.5.2: różniczka iloczynu

Dla dowolnych $f, g \in K[x]$ mamy $d(fg) = d(f) \cdot g + f \cdot d(g)$.

Dowód. Jeśli $f = x^n$ oraz $g = x^m$, to oczywiście bezpośrednie sprawdzenie daje:

$$\begin{aligned} d(x^n y^m) &= d(x^{m+n}) = (m+n)x^{m+n-1} \\ d(x^n)x^m + x^n d(x^m) &= nx^{n-1}x^m + mx^{m-1}x^n = (m+n)x^{m+n-1} \end{aligned}$$

W ogólnym przypadku iloczyn fg jest sumą jednomianów postaci $a_i b_j x^i y_j$. Teza wynika więc z liniowości przekształcenia d . \square

Przez indukcję uzyskujemy następujący wniosek, przypominający wniosek o pochodnej funkcji złożonej w szczególnym przypadku.

Wniosek 16.5.3: różniczka potęgi

Dla każdego $f \in K[x]$ oraz $n \in \mathbb{N}$ mamy $d(f^n) = n f^{n-1} d(f)$.

Jesteśmy gotowi do uzasadnienia głównego rezultatu.

Twierdzenie 16.5.4: Warunek istnienia pierwiastka wielokrotnego

Niech $f \in K[x]$ będzie wielomianem stopnia > 0 . Następujące warunki są równoważne:

- Element $a \in K$ jest pierwiastkiem wielokrotnym f .
- $f(a) = df(a) = 0$.

Dowód. Niech $f = (x - a)^m g$, gdzie m jest krotnością a w f oraz niech $g \in K[x]$ będzie taki, że $g(a) \neq 0$. Jeśli $m = 0$, to $f(a) = g(a) \neq 0$. W przeciwnym przypadku: $df = m(x - a)^{m-1}g + (x - a)^m d(g)$. Jeśli $m = 1$, to $df(a) = g(a) \neq 0$. Jeśli natomiast $m \geq 2$, to $f(a) = df(a) = 0$. Zatem $a \in K$ jest pierwiastkiem wielokrotnym f wtedy i tylko wtedy, gdy $m \geq 2$, co jest równoważne $f(a) = df(a) = 0$. \square

16.6 Trivia. Układy współrzędnych i układy równań

Pokazaliśmy właśnie izomorfizm przestrzeni skończenie wymiarowej V nad ciałem K z przestrzenią współrzędnych $K^{\dim V}$. Jak wspomnieliśmy, takich izomorfizmów jest wiele i polegają one na wyborze bazy \mathcal{A} przestrzeni V i zadaniu funkcji $\phi_{\mathcal{A}}$ postaci:

$$V \ni \alpha \xrightarrow{\phi_{\mathcal{A}}} (a_1, \dots, a_n) \in K^n,$$

gdzie $n = \dim V$ oraz a_1, \dots, a_n są współrzędnymi α w bazie \mathcal{A} . Zobaczmy dwa przykłady.

- Przestrzeń wielomianów $K[x]_{\leq 3}$ stopnia nie większego niż 3 nad ciałem K jest wymiaru 4. Wybierzmy dwie bazy tej przestrzeni postaci:

$$\mathcal{A} = (1, x, x^2, x^3), \quad \mathcal{B} = (1, x + 1, (x + 1)^2, (x + 1)^3).$$

Mamy zatem dwa izomorfizmy $\phi_{\mathcal{A}}, \phi_{\mathcal{B}} : K[x]_{\leq 3} \rightarrow K^4$ zadane warunkami:

$$\begin{aligned} \phi_{\mathcal{A}}(1) &= (1, 0, 0, 0), & \phi_{\mathcal{A}}(x) &= (0, 1, 0, 0), & \phi_{\mathcal{A}}(x^2) &= (0, 0, 1, 0), & \phi_{\mathcal{A}}(x^3) &= (0, 0, 0, 1), \\ \phi_{\mathcal{B}}(1) &= (1, 0, 0, 0), & \phi_{\mathcal{B}}(x + 1) &= (0, 1, 0, 0), & \phi_{\mathcal{B}}((x + 1)^2) &= (0, 0, 1, 0), & \phi_{\mathcal{B}}((x + 1)^3) &= (0, 0, 0, 1). \end{aligned}$$

Biorąc np. wielomian $w(x) = x^3 + 1 = (x + 1)^3 - 3(x - 1)^2 + 3(x + 1) \in K[x]$ mamy:

$$\phi_{\mathcal{A}}(w(x)) = (1, 0, 0, 1), \quad \phi_{\mathcal{B}}(w(x)) = (1, 3, -3, 1).$$

Zauważmy, że przy ustalonym utożsamieniu, np. przy $\phi_{\mathcal{A}}$ możemy opisywać podprzestrzenie $K[x]_{\leq 3}$ jako zbiory rozwiązań układów równań o czterech zmiennych, utożsamianych ze współrzędnymi wektorów $\phi_{\mathcal{A}}(w)$, dla $w \in K[x]_{\leq 3}$. W ten sposób możemy więc uznać, że dowolna podprzestrzeń przestrzeni skończonego wymiaru może być opisana układem równań liniowych, o ile wybierzymy wcześniej układ współrzędnych, które to współrzędne traktować będziemy dalej jako zmienne.

Oto przykład. Rozważmy równanie o czterech zmiennych postaci $x_1 - x_4 = 0$. Przy izomorfizmie $\phi_{\mathcal{A}}$ można uznać, że równanie to opisuje te wielomiany w $K[x]_{\leq 3}$, których współrzędne $\phi_{\mathcal{A}} = (x_1, x_2, x_3, x_4)$ spełniają równanie $x_1 - x_4 = 0$. A zatem są to wielomiany postaci $a + bx + cx^2 + ax^3$, dla dowolnych $a, b, c \in K$.

Jeśli jednak rozważymy izomorfizm $\phi_{\mathcal{B}}$, wówczas należy dowolny wielomian $w \in K[x]_{\leq 3}$ przedstawić w bazie \mathcal{B} i szukać wielomianów, które przy 1 oraz $(x + 1)^3$ mają te same współczynniki. Będzie to oczywiście inny zbiór wielomianów niż te, uzyskane przy izomorfizmie $\phi_{\mathcal{A}}$. To zagadnienie będzie dla nas ważne w drugim semestrze i rozważać je będziemy w znacznie większej ogólności.

- Przestrzeń $W \subseteq K^4$ rozwiązań układu $x_1 + x_2 + x_3 + x_4 = 0$ jest trójwymiarowa i biorąc jej bazy:

$$\mathcal{A} = ((1, -1, 0, 0), (1, 0, -1, 0), (1, 0, 0, -1)), \quad \mathcal{B} = ((-1, 1, 0, 0), (0, -1, 1, 0), (0, 0, -1, 1))$$

możemy zadań przykładowe dwa izomorfizmy $\phi_{\mathcal{A}}, \phi_{\mathcal{B}} : W \rightarrow K^3$ warunkami:

$$\begin{aligned} \phi_{\mathcal{A}}((1, -1, 0, 0)) &= (1, 0, 0), & \phi_{\mathcal{A}}((1, 0, -1, 0)) &= (0, 1, 0), & \phi_{\mathcal{A}}((1, 0, 0, -1)) &= (0, 0, 1), \\ \phi_{\mathcal{B}}((-1, 1, 0, 0)) &= (1, 0, 0), & \phi_{\mathcal{B}}((0, -1, 1, 0)) &= (0, 1, 0), & \phi_{\mathcal{B}}((0, 0, -1, 1)) &= (0, 0, 1). \end{aligned}$$

Definicja 16.6.1: Układ współrzędnych

Izomorfizmy n -wymiarowej przestrzeni V nad K na przestrzeń K^n nazywamy UKŁADAMI WSPÓŁRZĘDNYCH w V . UKŁADEM WSPÓŁRZĘDNYCH ZWIĄZANYCH Z BAZĄ (v_1, \dots, v_n) w V nazywamy izomorfizm $\sigma : V \rightarrow K^n$ przeprowadzający v_j na j -ty wektor bazy standardowej ϵ_j .

W szczególności, biorąc dowolną bazę \mathcal{A} w K^n można zadać na tej przestrzeni układ współrzędnych odpowiadający tej bazie. Warto jednak patrzeć na to ogólnie: tzw. współrzędne wektora v przestrzeni skończenie wymiarowej V w bazie \mathcal{A} , które już jakiś czas rozważamy, to nic innego jak obraz tego wektora w odpowiednim układzie współrzędnych wyznaczonym przez tę bazę, czyli $\phi_{\mathcal{A}}(v)$. Widzimy zatem, że mamy różne sposoby zadawania współrzędnych na przestrzeniach skończonego wymiaru. Jest to istotne, jak się przekonamy w drugim semestrze, choćby dlatego, że przekształcenia liniowe mogą wyznaczać pewne układy współrzędnych, w których geometria przekształcenia liniowego jest szczególnie dobrze widoczna.

Rozdział 17

Działania na przekształceniach liniowych. Diagramy przekształceń

17.1 Wykład 17

Aby zrozumieć lepiej rolę przekształceń liniowych w teorii przestrzeni liniowych i związanymi z nią zagadnieniami algebraicznymi i geometrycznymi, odwołamy się do pojęć znanych z ogólnej teorii funkcji, związanych z operacjami na przekształceniach liniowych. Zasygnalizujemy też ważny sposób patrzenia na funkcje poprzez tzw. diagramy przemienne.

Powiemy na początek o strukturze przestrzeni liniowej złożonej ze wszystkich przekształceń liniowych pomiędzy ustalonimi przestrzeniami liniowymi V, W (nad tym samym ciałem K). Przypomnijmy w tym celu definicję działania dodawania przekształceń i działania mnożenia przekształcenia przez skalar.

Definicja 17.1.1: Działania na przekształceniach liniowych

Niech V, W będą przestrzeniami liniowymi nad ciałem K i niech $\phi, \psi : V \rightarrow W$ będą przekształceniemi liniowymi.

- SUMĄ ϕ i ψ nazywamy odwzorowanie $\phi + \psi : V \rightarrow W$ zadane wzorem:

$$(\phi + \psi)(v) = \phi(v) + \psi(v), \text{ dla każdego } v \in V,$$

- ILOCZYNEM ϕ przez skalar $a \in K$ nazywamy odwzorowanie $a \cdot \phi : V \rightarrow W$ postaci:

$$(a \cdot \phi)(v) = a \cdot \phi(v), \text{ dla każdego } v \in V.$$

Przykład. Jeśli $V = \mathbb{R}^3$ oraz $W = \mathbb{R}^2$, to dla przekształceń liniowych $\phi, \psi : V \rightarrow W$ zadanych wzorami:

$$\phi((x_1, x_2, x_3)) = (x_1 + x_3, x_1 - x_2 + x_3), \quad \psi((x_1, x_2, x_3)) = (0, 2x_1)$$

mamy:

$$(\phi + \psi)((x_1, x_2, x_3)) = (x_1 + x_3, 3x_1 - x_2 + x_3), \quad (2 \cdot \psi)((x_1, x_2, x_3)) = (0, 4x_1).$$

Oczywiście jeśli $\phi, \psi : V \rightarrow W$ są przekształceniemi liniowymi przestrzeni liniowych nad ciałem K oraz jeśli $a \in K$, to funkcje $\phi + \psi$ oraz $a \cdot \phi$ są przekształceniemi liniowymi z V do W .

Definicja 17.1.2: Przestrzeń przekształceń liniowych ustalonych przestrzeni liniowych

Niech V, W będą przestrzeniami liniowymi nad K . Przestrzeń liniową złożoną ze wszystkich przekształceń liniowych $\phi : V \rightarrow W$ będziemy oznaczać symbolem^a $L(V, W)$. Zerem tej przestrzeni liniowej jest przekształcenie zerowe.

^aCzęsto stosuje się także ogólniejszą notację: $\text{Hom}(V, W)$.

W przypadku przestrzeni V, W skończonego wymiaru opis przestrzeni $L(V, W)$ jest nietrudny. Zobaczmy najpierw przykład. Jeśli rozważymy przestrzeń $L(K^3, K^2)$, to jej wymiar wynosi 6, ponieważ bez trudu jesteśmy w stanie wskazać bazę tej przestrzeni. Jest ona złożona z sześciu przekształceń liniowych postaci:

$$\phi_{11}((x_1, x_2, x_3)) = (x_1, 0), \quad \phi_{12}((x_1, x_2, x_3)) = (x_2, 0), \quad \phi_{13}((x_1, x_2, x_3)) = (x_3, 0),$$

$$\phi_{21}((x_1, x_2, x_3)) = (0, x_1), \quad \phi_{22}((x_1, x_2, x_3)) = (0, x_2), \quad \phi_{23}((x_1, x_2, x_3)) = (0, x_3).$$

Funkcje te tworzą układ liniowo niezależny, gdyż gdyby dla pewnych $a_{11}, a_{12}, a_{13}, a_{21}, a_{22}, a_{23}$ funkcja

$$a_{11}\phi_{11} + a_{12}\phi_{12} + a_{13}\phi_{13} + a_{21}\phi_{21} + a_{22}\phi_{22} + a_{23}\phi_{23}$$

przyjmowała wartość $(0, 0)$, to przyjmowałaby ją w szczególności dla każdego wektora bazy st. Stąd

$$(a_{11}\phi_{11} + a_{12}\phi_{12} + a_{13}\phi_{13} + a_{21}\phi_{21} + a_{22}\phi_{22} + a_{23}\phi_{23})((1, 0, 0)) = a_{11}(1, 0) + a_{21}(0, 1) = (0, 0)$$

$$(a_{11}\phi_{11} + a_{12}\phi_{12} + a_{13}\phi_{13} + a_{21}\phi_{21} + a_{22}\phi_{22} + a_{23}\phi_{23})((0, 1, 0)) = a_{12}(1, 0) + a_{22}(0, 1) = (0, 0)$$

$$(a_{11}\phi_{11} + a_{12}\phi_{12} + a_{13}\phi_{13} + a_{21}\phi_{21} + a_{22}\phi_{22} + a_{23}\phi_{23})((0, 0, 1)) = a_{13}(1, 0) + a_{23}(0, 1) = (0, 0)$$

czyli $a_{11} = a_{21} = 0$, $a_{12} = a_{22} = 0$ oraz $a_{13} = a_{23} = 0$.

Wiemy, że każde przekształcenie liniowe $\phi : K^3 \rightarrow K^2$ jest postaci:

$$\phi((x_1, x_2, x_3)) = (a_{11}x_1 + a_{12}x_2 + a_{13}x_3, a_{21}x_1 + a_{22}x_2 + a_{23}x_3).$$

Zatem:

$$\phi = a_{11}\phi_1 + a_{12}\phi_2 + a_{13}\phi_3 + a_{21}\phi_4 + a_{22}\phi_5 + a_{23}\phi_6.$$

Rozważane funkcje rozpinają więc $L(K^3, K^2)$, która to jest przestrzenią wymiaru 6. A zatem są one bazą. Poniżej znajduje się twierdzenie opisujące w ogólną sytuację, z identycznym co do istoty dowodem.

Twierdzenie 17.1.3

Niech V, W będą skończenie wymiarowymi przestrzeniami liniowymi nad ciałem K , przy czym $\dim V = n$ oraz $\dim W = m$. Ma miejsce izomorfizm przestrzeni liniowych $L(V, W) \cong M_{m \times n}(K)$. W szczególności $\dim L(V, W) = m \cdot n$.

Dowód. Niech $\mathcal{A} = (\alpha_1, \dots, \alpha_n)$ oraz $\mathcal{B} = (\beta_1, \dots, \beta_m)$ będą odpowiednio bazami przestrzeni liniowych V, W . Określamy przekształcenia liniowe $\phi_{ij} : V \rightarrow W$ na bazie \mathcal{A} wzorem:

$$\phi_{ij}(\alpha_k) = \begin{cases} \beta_j, & \text{gdy } k = i \\ 0, & \text{gdy } k \neq i. \end{cases}$$

Innymi słowy określamy $m \cdot n$ funkcji z V do W . Przekształcenie ϕ_{ij} przypisuje wektor β_j wektorowi α_i , a pozostałym wektorów bazy \mathcal{A} przypisuje 0. Oczywiście ϕ_{ij} są parami różnymi przekształceniami liniowymi, zgodnie z Twierdzeniem 15.1.7.

Zauważmy, że układ funkcji ϕ_{ij} jest liniowo niezależny. Jeśli bowiem dla pewnych $a_{ij} \in K$ mamy

$$\sum_{i,j} a_{ij}\phi_{ij} = 0,$$

to wartość funkcji $\sum_{i,j} a_{ij}\phi_{ij}$ na wektorze α_i równa jest

$$a_{i1}\beta_1 + a_{i2}\beta_2 + \dots + a_{im}\beta_m$$

Skoro powyższy wektor jest zerowy (jak każda wartość rozważanej funkcji), a układ \mathcal{B} jest bazą W , to mamy $a_{ij} = 0$, dla każdego j . Wobec dowolności i widzimy, że wszystkie współczynniki a_{ij} są równe 0.

Dowolny element $\phi \in L(V, W)$ jest kombinacją liniową ϕ_{ij} . Istotnie, dla dowolnego α_k z bazy \mathcal{A} mamy $\phi(\alpha_k) \in W$, a także $\phi_{kj}(\alpha_k) = \beta_j$ oraz $\phi_{ij}(\alpha_k) = 0$, dla $k \neq i$. Istnieją więc takie $a_{k1}, \dots, a_{km} \in K$, że

$$\phi(\alpha_k) = a_{k1}\beta_1 + \dots + a_{km}\beta_m = a_{k1}\phi_{k1}(\alpha_k) + \dots + a_{km}\phi_{km}(\alpha_k) = \left(\sum_{i,j} a_{ij}\phi_{ij} \right)(\alpha_k).$$

Stąd ϕ oraz przekształcenie $\sum_{i,j} a_{ij}\phi_{ij}$ są identyczne na bazie \mathcal{A} przestrzeni liniowej V , czyli są równe. Zatem układ ϕ_{ij} rozpina $L(V, W)$. Jest to więc mn -elementowa baza $L(V, W)$. Stąd $\dim L(V, W) = mn$. \square

Definicja 17.1.4: Złożenie przekształceń liniowych

Niech $\phi : V \rightarrow W$ oraz $\psi : W \rightarrow Z$ będą przekształceniami liniowymi przestrzeni nad ciałem K . **ZŁOŻENIEM PRZEKSZTAŁCEŃ** ϕ i ψ nazywamy odwzorowanie $\psi \circ \phi : V \rightarrow Z$ zadane wzorem:

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

Oczywiście $\phi : V \rightarrow W$ oraz $\psi : W \rightarrow Z$ są przekształceniami liniowymi, to także $\psi \circ \phi$ jest przekształceniem liniowym, bo dla dowolnych $\alpha, \beta \in V$ mamy:

$$(\psi \circ \phi)(\alpha + \beta) = \psi(\phi(\alpha + \beta)) = \psi(\phi(\alpha) + \phi(\beta)) = \psi(\phi(\alpha)) + \psi(\phi(\beta)) = (\psi \circ \phi)(\alpha) + (\psi \circ \phi)(\beta).$$

Rozważmy kilka przykładów.

- Dla $\phi, \psi : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ zadanych wzorami:

$$\phi((x, y)) = (x, 0), \quad \psi((x, y)) = (x, x)$$

mamy

$$(\psi \circ \phi)((x, y)) = \psi(\phi((x, y))) = \psi((x, 0)) = (x, x), \quad (\phi \circ \psi)((x, y)) = \phi(\psi((x, y))) = \phi((x, x)) = (x, 0).$$

Widzimy zatem, że kolejność składania przekształceń ma znaczenie.

- Niech $\phi_a : V \rightarrow V$ będzie homotetią o skali a przestrzeni liniowej V nad ciałem K . Wówczas dla dowolnych $a_1, a_2 \in K$ mamy

$$\phi_a \circ \phi_b = \phi_b \circ \phi_a = \phi_{ab}.$$

- Niech $O_\theta : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ będzie obrotem o kąt θ . Wówczas dla dowolnych $\theta_1, \theta_2 \in \mathbb{R}$ mamy

$$O_{\theta_1} \circ O_{\theta_2} = O_{\theta_2} \circ O_{\theta_1} = O_{\theta_1 + \theta_2}.$$

- Niech $V = U \oplus W$. Niech ϕ będzie rzutem na U wzdluz W . Wówczas $\phi \circ \phi = \phi$. Można wykazać, że każde przekształcenie liniowe $\phi : V \rightarrow V$ spełniające równość $\phi \circ \phi = \phi$ jest rzutem na pewną podprzestrzeń V (patrz zadania do samodzielnej pracy).

- Niech $V = U \oplus W$. Niech ψ będzie symetrią względem U wzdluz W . Wówczas $\psi \circ \psi = \text{id}_V$.

- Niech $\phi : V \rightarrow V$. Przez n -krotną ITERACJĘ lub n -tą POTĘGĘ przekształcenia ϕ rozumiemy n -krotne złożenie ϕ ze sobą, czyli endomorfizm

$$\phi^n = \underbrace{\phi \circ \phi \circ \dots \circ \phi}_n.$$

Przyjmujemy przy tym, że $\phi^0 = \text{id}_V$. Niech $\phi : K^n \rightarrow K^n$ dany będzie wzorem

$$\phi((x_1, x_2, \dots, x_n)) = (x_2, x_3, \dots, x_n, 0).$$

Wówczas $\phi^n = 0$ oraz $\phi^{n-1} \neq 0$.

* * *

Gdy rozważamy złożenia przekształceń liniowych, obok zapisu funkcyjnego, stosować będziemy notację **diagramową**. Jest to spojrzenie charakterystyczne dla nowoczesnej matematyki, czerpiącej silnie z tak zwanej teorii kategorii (powiemy o niej w dalszych wykładach). Naszym celem jest zwrócenie uwagi na tak zwane MYŚLENIE DIAGRAMOWE, pozwalające rozważać całe układy przekształceń i ich złożień.

Czym więc są owe diagramy¹? Zaczniemy od przykładu. Fakt istnienia złożenia przekształceń $\phi : V \rightarrow W$ oraz $\psi : W \rightarrow Z$ postaci $\psi \circ \phi : V \rightarrow Z$ opisujemy na diagramie w następujący sposób:

$$\begin{array}{ccc} V & \xrightarrow{\phi} & W \\ & \searrow \psi \circ \phi & \downarrow \psi \\ & & Z \end{array}$$

¹Do końca oczywiście nie opowiemy tego precyzyjnie bez języka teorii kategorii, a jedynie podamy pewien uproszczony model (aby zdefiniować graf potrzebny jest zbiór, co w ogólnym kontekście teorii kategorii nie jest potencjalnie mile widziane, ale jeszcze się nie spotkałem z tym, by komuś to realnie przeszkadzało). W istocie diagram to teoriokategoryjny odpowiednik rodziny indeksowanej zbiorów z teorii mnogości i definiuje się go jako pewien funktor, ale to nam jest zupełnie niepotrzebne.

Definicja 17.1.5: Diagram przekształceń liniowych, diagram przemienny

DIAGRAMEM PRZEKSZTAŁCEŃ LINIOWYCH nazywać będziemy graf skierowany, którego wierzchołki etykietowane są przestrzeniami liniowymi, a krawędzie – przekształceniami liniowymi pomiędzy nimi. Przekształcenie liniowe postaci $\phi : V \rightarrow W$ oznaczamy zatem $V \xrightarrow{\phi} W$. Jeśli w diagramie występuje podgraf typu $V \xrightarrow{\phi} W \xrightarrow{\psi} Z$ to znaczy, że istnieje złożenie $\psi \circ \phi$.

Powiemy, że diagram przekształceń jest PRZEMIENNY, jeśli dla dowolnych dwóch wierzchołków V, W tego diagramu, złożenia wzdłuż dowolnych dwóch ścieżek tego diagramu (traktowanego jako graf) o początkach w V i końcach w W są sobie równe (jako przekształcenia).

Dla przykładu, poniższy diagram jest przemienny, o ile $\phi_2 \circ \phi_1 = \psi_2 \circ \psi_1$.

$$\begin{array}{ccc} V & \xrightarrow{\phi_1} & W \\ \downarrow \psi_1 & & \downarrow \phi_2 \\ X & \xrightarrow{\psi_2} & Z \end{array}$$

Kluczowe własności algebraiczne – bycie izomorfizmem, monomorfizmem czy epimorfizmem odczytać można w języku złożen przekształceń liniowych. Podstawowej intuicji dostarczają nam tu izomorfizmy. Teoria funkcji podpowiada bowiem, że dla każdej bijekcji istnieje odwrotna bijekcja. Okazuje się, że dla bijekcji-izomorfizmów, bijekcja odwrotna jest również izomorfizmem.

Twierdzenie 17.1.6

Niech $\phi : V \rightarrow W$ będzie przekształceniem liniowym. Następujące warunki są równoważne:

- (1) ϕ jest izomorfizmem,
- (2) istnieje takie przekształcenie liniowe $\psi : W \rightarrow V$, że: $\psi \circ \phi = \text{id}_V$ oraz $\phi \circ \psi = \text{id}_W$.

Dowód. Niech ϕ będzie izomorfizmem. Określamy $\psi : W \rightarrow V$ warunkiem $\psi(\beta) = \alpha$, gdzie $\phi(\alpha) = \beta$. Jeśli $\psi(\beta_1) = \alpha_1$ oraz $\psi(\beta_2) = \alpha_2$, to $\phi(\alpha_1) = \beta_1$, $\phi(\alpha_2) = \beta_2$. Z liniowości ϕ mamy

$$\phi(\alpha_1 + \alpha_2) = \beta_1 + \beta_2.$$

A zatem

$$\psi(\beta_1 + \beta_2) = \alpha_1 + \alpha_2 = \psi(\beta_1) + \psi(\beta_2).$$

Analogicznie sprawdzamy $\psi(c\beta) = c\psi(\beta)$, dla każdych $\beta \in W$, $c \in K$. Zatem ψ jest liniowe i $\psi \circ \phi = \text{id}_V$ oraz $\phi \circ \psi = \text{id}_W$. Stąd (1) \Rightarrow (2).

Przechodzimy do implikacji (2) \Rightarrow (1). Weźmy $\alpha, \beta \in V$ i niech $\phi(\alpha) = \phi(\beta)$. Wówczas

$$\alpha = \text{id}_V(\alpha) = (\psi \circ \phi)(\alpha) = \psi(\phi(\alpha)) = \psi(\phi(\beta)) = (\psi \circ \phi)(\beta) = \text{id}_V(\beta) = \beta.$$

Zatem ϕ jest różnowartościowe. Mamy $\phi \circ \psi = \text{id}_W$, a więc dla każdego $\gamma \in W$ mamy

$$\gamma = \text{id}_W(\gamma) = (\phi \circ \psi)(\gamma) = \phi(\psi(\gamma)).$$

A więc $\gamma = \phi(\psi(\gamma))$, czyli ϕ jest „na”. □

Powyższy dowód pokazuje, że może być tylko jedno ψ spełniające warunek (*). Co więcej, ψ to izomorfizm.

Istnienie **dokładnie jednego** przekształcenia liniowego, które sprawia, że diagram złożen przekształceń jest przemienny oznaczamy zaznaczając to przekształcenie przerywaną linią.

$$\begin{array}{ccc} V & \xrightarrow{\phi} & W \\ id_V \searrow & \downarrow \psi & \searrow id_W \\ V & \xrightarrow{\phi} & W \end{array}$$

Definicja 17.1.7: Przekształcenie odwrotne

Jeśli dla przekształcenia liniowego $\phi : V \rightarrow W$ istnieje przekształcenie liniowe $\psi : W \rightarrow V$ spełniające (2), to ψ nazywamy PRZEKSZTAŁCENIEM ODWROTNYM do ϕ i oznaczamy przez ϕ^{-1} .

Jak widzimy ϕ^{-1} istnieje wtedy i tylko wtedy, gdy ϕ jest izomorfizmem.

Wniosek 17.1.8

Niech V, W będą przestrzeniami liniowymi. Wówczas

- przekształcenie liniowe $\phi : V \rightarrow W$ jest monomorfizmem wtedy i tylko wtedy, gdy istnieje takie przekształcenie liniowe $\psi : W \rightarrow V$, że $\psi \circ \phi = \text{id}_V$.
- przekształcenie liniowe $\phi : W \rightarrow V$ jest epimorfizmem wtedy i tylko wtedy, gdy istnieje takie przekształcenie liniowe $\psi : W \rightarrow V$, że $\phi \circ \psi = \text{id}_W$.

Wypowiedzieliśmy pewne uwagi o przekształceniach liniowych, których złożenia są identycznościami. Drugą ważną sytuacją jest ta, gdy złożenie przekształceń liniowych jest zerowe.

Uwaga 17.1.9

Niech $\phi : V \rightarrow W$ oraz $\psi : W \rightarrow Z$ będą przekształceniami liniowymi. Wówczas jeśli $\psi \circ \phi = 0$, to

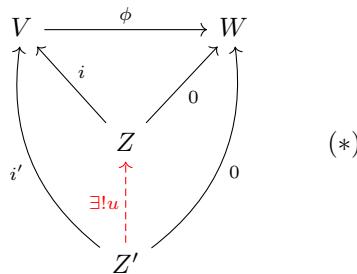
$$\text{im } \phi \subseteq \ker \psi.$$

Na końcu powiemy nieco o tzw. własnościach uniwersalnych. Pojęcie to wywodzi się z tzw. teorii kategorii, która teorie matematyczne buduje nie fundując ich na pojęciu zbioru. Jest to podejście ważne zwłaszcza w algebrze i topologii. Omówimy przykład takiej własności dla jądra przekształcenia liniowego.

Definicja 17.1.10: Własność uniwersalna jądra

Niech $\phi : V \rightarrow W$ będzie przekształceniem liniowym przestrzeni liniowych nad ciałem K . Jeśli przestrzeń liniowa Z nad ciałem K oraz przekształcenie liniowe $i : Z \rightarrow V$ spełniają warunki:

- $\phi \circ i = 0$,
- dla dowolnej przestrzeni liniowej Z' oraz dowolnego przekształcenia liniowego $i' : Z' \rightarrow V$ spełniającego $\phi \circ i' = 0$ istnieje dokładnie jedno przekształcenie liniowe $u : Z' \rightarrow Z$ takie, że $i \circ u = i'$



to mówimy, że para Z oraz $i : Z \rightarrow V$ SPEŁNIA WŁASNOŚĆ UNIwersalną JĄDRA ϕ .

Definicja wygląda bardzo pokrętnie (i pewnie ktoś mógłby powiedzieć, że w ogóle tak się nie powinno jej wprowadzać), ale jej sens jest następujący: interesuje nas czy istnieją takie pary Z oraz $i : Z \rightarrow V$, które spełniają własność wyżej i czy można się wypowiedzieć o jednoznaczności. Po czym miałoby to nas interesować? Otóż posiadanie takiej własności oznacza, że jakiś obiekt (na przykład jądro) można scharakteryzować jedynie na podstawie własności złożen przekształceń liniowych, a bez działań na elementach (tzn. stwierdzania, że badamy zbiór zer jakiejś funkcji). Jest jeszcze drugi ważny aspekt, który tylko sygnalizuję. Gdyby odwrócić wszystkie strzałki w powyższej definicji, również uzyskalibyśmy własność uniwersalną pewnego obiektu Z , zwanego **kojądrem** ϕ . Z definicji jest to przestrzeń ilorazowa $W / \text{im } \phi$.

Twierdzenie 17.1.11: Jądro spełnia swoją własność uniwersalną

Niech $\phi : V \rightarrow W$ będzie przekształceniem liniowym. Niech $i : \ker \phi \rightarrow V$ będzie dane wzorem $i(v) = v$. Wówczas para $Z = \ker \phi, i$ spełnia własność uniwersalną jądra ϕ .

Dowód. Rozważmy dowolną przestrzeń liniową Z' oraz przekształcenie liniowe $i' : Z' \rightarrow V$, że $\phi \circ i' = 0$. Ponieważ $\phi \circ i' = 0$, to obraz przekształcenia i' zawiera się w jądrze ϕ — zgodnie z Uwagą 17.1.9. Zatem można zdefiniować przekształcenie $u : Z' \rightarrow \ker \phi$, takie że dla każdego $v' \in Z'$ mamy:

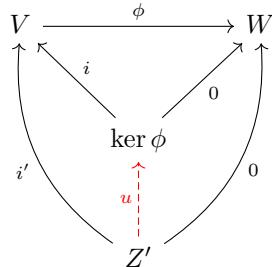
$$u(v') = i'(v').$$

Widać, że u jest dobrze określone, ponieważ $i'(v') \in \ker \phi$ dla każdego $v' \in Z'$.

Oczywiście u jest liniowe, gdyż dla dowolnych $v'_1, v'_2 \in Z'$ oraz $a, b \in K$:

$$u(av'_1 + bv'_2) = i'(av'_1 + bv'_2) = ai'(v'_1) + bi'(v'_2) = au(v'_1) + bu(v'_2).$$

Należy uzasadnić, że poniższy diagram jest przemienny dla tak określonego u .



Dla każdego $v' \in Z'$, mamy:

$$(i \circ u)(v') = i(u(v')) = i(i'(v')) = i'(v').$$

Zatem $i \circ u = i'$, co oznacza, że dolny trójkąt diagramu jest przemienny. Co więcej, mamy $\phi \circ i' = 0$.

Pozostała kwestia jednoznaczności u . Założymy, że istnieje inne przekształcenie $u' : Z' \rightarrow \ker \phi$, takie że $i \circ u' = i'$. Dla dowolnego $v' \in Z'$ mamy:

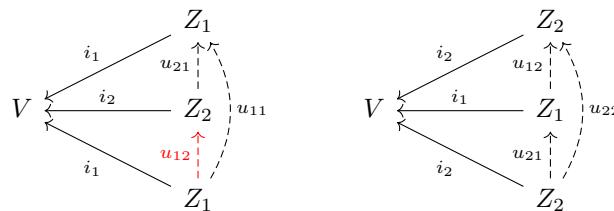
$$i(u(v')) = i'(v') = i(u'(v')).$$

Ponieważ i jest monomorfizmem, wynika stąd, że $u(v') = u'(v')$ dla każdego $v' \in Z'$. Zatem $u = u'$. \square

Twierdzenie 17.1.12

Niech $\phi : V \rightarrow W$ będzie przekształceniem liniowym. Jeśli przestrzenie liniowe Z_1, Z_2 wraz z przekształceniemi liniowymi $i_1 : Z_1 \rightarrow V$ oraz $i_2 : Z_2 \rightarrow V$ spełniają własność uniwersalną jądra ϕ , to istnieje izomorfizm $f : Z_1 \rightarrow Z_2$.

Dowód. Rozważmy następujące diagramy, w których dla $i, j \in 1, 2$ cztery przekształcenia $u_{ij} : Z_i \rightarrow Z_j$ spełniają warunki własności uniwersalnych dla przestrzeni Z_j .



Oto konstrukcje u_{21} oraz u_{22} . Skoro Z_1 oraz i_1 spełnia własność uniwersalną, to wstawiając w miejsce Z' oraz i' w diagramie (*) w Definicji 17.1.10 odpowiednio Z_2 oraz i_2 wiemy, że istnieje dokładnie jedno przekształcenie $u_{21} : Z_2 \rightarrow Z_1$, że $i_2 = i_1 \circ u_{21}$. Podobnie wstawiając w miejsce Z' oraz i' tę samą przestrzeń Z_2 oraz funkcję i_2 mamy dokładnie jedno przekształcenie liniowe $u_{22} : Z_2 \rightarrow Z_2$, że $i_2 = i_2 \circ u_{22}$.

Skoro jednak $u_{11} : Z_1 \rightarrow Z_1$ ma jako jedyne spełniać warunek $i_1 = i_1 \circ u_{11}$, to $u_{11} = u_{21} \circ u_{12} = \text{id}_{Z_1}$. Analogicznie $u_{22} = u_{12} \circ u_{21} = \text{id}_{Z_2}$. Na mocy Twierdzenia 17.1.6 szukanym izomorfizmem f jest u_{12} .

17.2 Wybór przykładowych pytań

Proszę się nie uczyć na pamięć odpowiedzi czy wyjaśnienie! Takich pytań można ułożyć znacznie więcej i służą one jedynie sprawdzeniu podstawowego zrozumienia treści zajęć. Jeśli odpowiedź na dowolne z poniższych pytań sprawia Państwu problem, proszę pomyśleć który fragment notatek dotyczy tego pytania, wrócić do ich lektury, a w razie dalszych trudności poprosić prowadzącego o wskazówkę.

1. Dane są przekształcenia liniowe $\phi((x_1, x_2)) = (-x_1, x_2)$ oraz $\psi((x_1, x_2)) = (x_1, -x_2)$. Znajdź $\phi + \psi$.
2. Dane jest przekształcenie liniowe $\phi((x_1, x_2)) = (-2x_1, x_2)$. Znajdź $-\phi$ oraz ϕ^{-1} .
3. Niech $\phi : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ będzie obrotem o kąt θ . Znajdź ϕ^{-1} .
4. Niech $\phi : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ będzie jednokładnością o skali $a \neq 0$. Znajdź ϕ^{-1} .
5. Niech $\phi : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ będzie rzutem na $\text{lin}((1, 1))$ wzdłuż $\text{lin}(1, -1)$. Czy przekształcenie $-\phi$ jest rzutem? Czy przekształcenie $\text{id}_{\mathbb{R}^2} - \phi$ jest rzutem?
6. Niech $V = U \oplus W$ oraz niech ϕ, ψ będą odpowiednio rzutem na U względem W i symetrią względem U wzdłuż W . Uzasadnij, że przekształcenia $\phi, \psi, \text{id}_{\mathbb{R}^2}$ tworzą układ liniowo zależny.
7. Uzasadnij, że podzbiór przestrzeni $L(V, V)$ złożony ze wszystkich jednokładności jest podprzestrzenią. Jakiego jest ona wymiaru?
8. Podaj przykład takiego niezerowego przekształcenia liniowego $\phi : \mathbb{R}^2 \rightarrow \mathbb{R}^2$, że $\phi^2 = 0$.
9. Dane jest przekształcenie liniowe $\phi : K^\infty \rightarrow K^\infty$

$$\phi((x_1, x_2, \dots)) = (x_2, x_3, \dots).$$

Wskaż przekształcenie $\psi : K^\infty \rightarrow K^\infty$, że $\phi \circ \psi = \text{id}$. Czy istnieje przekształcenie $\psi : K^\infty \rightarrow K^\infty$, że $\psi \circ \phi = \text{id}$?

10. Niech $\phi : V \rightarrow W$ oraz $\psi : W \rightarrow Z$. Uzasadnij, że

$$\ker(\psi \circ \phi) \supseteq \ker \phi, \quad \text{im}(\psi \circ \phi) \subseteq \text{im } \psi.$$

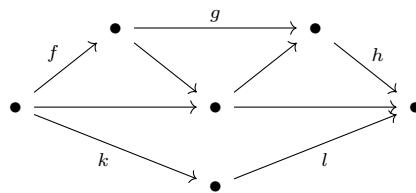
11. Poniższe trzy diagramy przekształceń liniowych są przemienne

$$\begin{array}{ccc} \begin{array}{ccc} A & \xrightarrow{a} & B \\ \downarrow b & & \downarrow d \\ D & \xrightarrow{f} & E \end{array} & \begin{array}{ccc} B & \xrightarrow{c} & C \\ \downarrow d & & \downarrow e \\ E & \xrightarrow{g} & F \end{array} & \begin{array}{ccc} A & \xrightarrow{a} & B & \xrightarrow{c} & C \\ \downarrow b & & \downarrow d & & \downarrow e \\ D & \xrightarrow{f} & E & \xrightarrow{g} & F \end{array} \end{array}$$

Dla diagramu po lewej oznacza to, że $d \circ a = f \circ b$.

- Wypisz wszystkie równości wynikające z przemienności środkowego i prawego diagramu.
- Uzasadnij, że przemienność diagramów lewego i środkowego implikuje przemienność prawego.

12. Rozważmy diagram przekształceń



Uzasadnij, że jeśli cztery wewnętrzne trójkąty tego diagramu są przemienne, to zewnętrzna komórka również jest przemienne (czyli $l \circ k = h \circ g \circ f$).

13. Dany jest ciąg przekształceń liniowych (zerowe, f , zerowe), o którym wiadomo, że obraz przekształcenia opisanego poprzednią strzałką jest równy jądro przekształcenia opisanego kolejną strzałką.

$$0 \xrightarrow{0} V \xrightarrow{f} W \xrightarrow{0} 0$$

Uzasadnij, że f jest izomorfizmem.

17.3 Zadania do samodzielnej pracy

Zadania ilustrujące umiejętności niezbędne do zaliczenia przedmiotu oznaczone są symbolem ♠.

Przy tych zadaniach dodany jest opis umiejętności, które sprawdzają.

1. (♠) Wykonywanie działań na przekształceniach

Niech $\phi : \mathbb{R}^3 \rightarrow \mathbb{R}^2$, $\psi : \mathbb{R}^3 \rightarrow \mathbb{R}^2$ będą przekształceniami liniowymi spełniającymi warunki

$$\phi((1, 1, 1)) = (3, 7), \quad \phi((1, 1, 0)) = (2, 5), \quad \phi((1, 0, 0)) = (1, 6)$$

$$\psi((2, 2, 1)) = (3, 3), \quad \psi((2, 1, 0)) = (5, 0), \quad \psi((2, 1, 1)) = (4, 2).$$

Znajdź wzór na przekształcenie $\phi + \psi$.

2. (♠) Wykonywanie działań na przekształceniach

Niech $\phi : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ będzie jednokładnością o skali -2 oraz niech $\psi : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ niech będzie symetrią prostopadłą względem $\text{lin}((1, 1))$ wzdłuż $\text{lin}((1, -1))$

- Znajdź wartość przekształcenia $\phi \circ \psi$ na wektorze $(2, 1)$.
- Znajdź wzór przekształcenia $2\phi - \psi$.
- Rozstrzygnij, czy przekształcenia ϕ, ψ są liniowo niezależne w $L(\mathbb{R}^2, \mathbb{R}^2)$?
- Znajdź wzory przekształceń $\psi \circ \phi, \phi \circ \psi, \phi^3, \psi^3$.
- Znajdź przekształcenia odwrotne do $\psi \circ \phi$ oraz $\phi \circ \psi$.
- Niech $\phi, \psi : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ będą obrotami o kąty odpowiednio θ_1, θ_2 . Uzasadnij, że $\phi \circ \psi = \psi \circ \phi$ jest obrotem o kąt $\theta_1 + \theta_2$.

4. (♠) Rozważmy przekształcenie liniowe $\phi : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ dane wzorem

$$\phi_\theta((x_1, x_2)) = (\cos 2\theta \cdot x_1 + \sin 2\theta \cdot x_2, \sin 2\theta \cdot x_1 - \cos 2\theta \cdot x_2).$$

- Wyznacz $\phi_\theta((\cos \theta, \sin \theta))$ oraz $\phi_\theta((- \sin \theta, \cos \theta))$. Jaka jest interpretacja geometryczna przekształcenia ϕ_θ ?
- Wykaż, że dowolny obrót o kąt θ jest złożeniem dwóch przekształceń liniowych postaci $\phi_{\theta_1}, \phi_{\theta_2}$ dla pewnych θ_1, θ_2 .
5. (♠) Rozstrzyganie czy złożenie jest monomorfizmem/epimorfizmem/izomorfizmem

- Przekształcenia liniowe $\phi : \mathbb{R}^4 \rightarrow \mathbb{R}^5$ i $\psi : \mathbb{R}^5 \rightarrow \mathbb{R}^4$ spełniają $r(\phi) = 4 = r(\psi)$. Czy wynika stąd, że przekształcenie $\psi \circ \phi : \mathbb{R}^4 \rightarrow \mathbb{R}^4$ jest izomorfizmem?
- Niech $\phi : \mathbb{R}^3 \rightarrow \mathbb{R}^2$ oraz $\psi : \mathbb{R}^2 \rightarrow \mathbb{R}^4$ będą przekształceniami liniowymi. Czy przekształcenie liniowe $\psi \circ \phi : \mathbb{R}^3 \rightarrow \mathbb{R}^4$ będące ich złożeniem może być monomorfizmem?
- Czy istnieją przekształcenia liniowe $f : \mathbb{R}^4 \rightarrow \mathbb{R}^3$, $g : \mathbb{R}^3 \rightarrow \mathbb{R}^4$ takie, że zachodzi równość $f \circ g = \text{id}_{\mathbb{R}^3}$?

6. Niech $\phi : V \rightarrow V$ będzie takim przekształceniem liniowym, że $\phi \circ \phi = \phi$. Wykaż, że istnieją takie podprzestrzenie V_1, V_2 w V , że ϕ jest rzutem na V_1 wzdłuż V_2 .

7. Niech $\phi : V \rightarrow V$ będzie takim przekształceniem liniowym, że $\phi \circ \phi = \text{id}$, gdzie V jest przestrzenią liniową nad ciałem K , w którym $1 + 1 \neq 0$ (np. $K = \mathbb{R}$ lub \mathbb{C} lub \mathbb{Q}). Wykaż, że istnieją takie podprzestrzenie V_1, V_2 w V , że ϕ jest symetrią względem V_1 wzdłuż V_2 .

8. Wykaż, że jeśli przekształcenie $\phi \in L(V, V)$ jest przemienne z dowolnym rzutem $\psi \in L(V, V)$, to znaczy: $\phi \circ \psi = \psi \circ \phi$, wówczas przekształcenie ϕ jest homotetią

9. Niech V, W będą przestrzeniami liniowymi, zaś niech $\phi \in L(V, W)$ oraz $\psi \in L(W, W)$ spełniają $\psi \circ \phi = 0$ oraz $\text{im}(\text{id}_W - \psi) \subseteq \text{im}(\phi)$. Wykaż, że $W = \text{im } \phi \oplus \text{im } g$.

10. Niech V będzie skończenie wymiarową przestrzenią liniową. Założmy, że $\phi, \psi \in L(V, V)$ spełniają $\phi \circ \psi = \psi \circ \phi$, oraz że $\phi - \psi$ jest monomorfizmem. Wykaż, że $\ker(\phi \circ \psi) = \ker \phi \oplus \ker \psi$.

11. Dane są przestrzenie liniowe V, W nad ciałem K , przy czym $\dim V = n$, $\dim W = 2$. Dane jest również $\phi \in L(V, V)$. Definiujemy funkcję $\Phi : L(V, W) \rightarrow L(V, W)$ wzorem $\Phi(\psi) = \psi \circ \phi$. Wykaż, że Φ jest przekształceniem liniowym. Czy jeśli ϕ jest izomorfizmem, to również Φ jest izomorfizmem?

12. Wykaż, że jeśli $\phi_1 : V \rightarrow W, \phi_2 : V \rightarrow W$ oraz $\psi : W \rightarrow Z$ są przekształceniami liniowymi, to zachodzi równość

$$\psi \circ (\phi_1 + \phi_2) = \psi \circ \phi_1 + \psi \circ \phi_2.$$

17.4 Uzupełnienie. Faktoryzacje i przekształcenia ilorazowe

Jakiś czas temu w notatkach do wykładu pojawiła się definicja przestrzeni ilorazowej V/W , gdzie $W \subseteq V$ jest podprzestrzenią V . Wspomnieliśmy wówczas o formule $\dim V/W = \dim V - \dim W$. Jak ją wyprowadzić, do czego używa się takich przestrzeni, skąd taka nazwa i dlaczego warto o nich powiedzieć coś właśnie teraz, gdy zajmujemy się przekształceniami liniowymi? Kluczem do sprawy jest pojęcie faktoryzowania się przekształcenia liniowego, na swój sposób odwrotne do pojęcia złożenia. Oto definicja.

Definicja 17.4.1: Faktoryzowanie się przekształcenia przez inne

Niech $\phi : V \rightarrow W$ będzie przekształceniem liniowym. Powiemy, że ϕ FAKTORYZUJE SIĘ PRZEZ PRZEKSZTAŁCENIE liniowe $\psi : V' \rightarrow W$, jeśli istnieje przekształcenie $\pi : V \rightarrow V'$, że $\phi = \psi \circ \pi$, czyli gdy następujący diagram jest przemienny.

$$\begin{array}{ccc} V & \xrightarrow{\phi} & W \\ & \searrow \pi & \nearrow \psi \\ & V' & \end{array}$$

Oczywiście faktoryzacja zawsze jest możliwa, jeśli weźmiemy $V' = V$ oraz $\psi = \phi$. Sprawa jest jednak nieco ciekawsza. Patrzmy najpierw na przykłady:

- Rozważmy przekształcenie $\phi : K^4 \rightarrow K$ dane wzorem $\phi((x_1, x_2, x_3, x_4)) = x_4$. Jest ono oczywiście liniowe. Przekształcenie to faktoryzuje się przez $\psi' : K^2 \rightarrow K$ dane wzorem $\psi(y_1, y_2) = y_2$. Istotnie, jeśli $\pi : K^4 \rightarrow K^2$ dane jest wzorem: $\pi(z_1, z_2, z_3, z_4) = (z_2, z_4)$, to mamy: $\psi(\pi((x_1, x_2, x_3, x_4))) = \psi((x_2, x_4)) = x_4 = \phi((x_1, x_2, x_3, x_4))$. Mamy więc:

$$\begin{array}{ccc} K^4 & \xrightarrow{\phi} & K \\ & \searrow \pi & \nearrow \psi \\ & K^2 & \end{array}$$

- Rozważmy podprzestrzeń C przestrzeni \mathbb{R}^∞ złożoną z wszystkich ciągów zbieżnych i rozważmy przekształcenie $\phi : C \rightarrow \mathbb{R}$ dane wzorem $\phi((x_1, \dots)) = \lim_{n \rightarrow \infty} x_n$. Jest to oczywiście przekształcenie liniowe. Czy znajdziemy dla niego jakąś faktoryzację? Może ktoś uzna to za trywialne – ale owszem, jesteśmy w stanie to zrobić. Rozważmy podprzestrzeń D wszystkich ciągów stałych. Bierzemy teraz przekształcenie $\psi : D \rightarrow \mathbb{R}$ dane wzorem $\psi((x, x, x, \dots)) = x$. Czy widzimy, że ϕ faktoryzuje się przez ψ ? Jak wygląda przekształcenie π ? I skąd wiedzieliśmy, żeby szukać właśnie ciągów stałych?

Kluczem jest pojęcie przestrzeni ilorazowej. Zauważmy, że jeśli $W \subseteq V$, to mamy naturalne przekształcenie $\pi : V \rightarrow V/W$ zadane wzorem: $\pi(\alpha) = \alpha + W$ (przyporządkujemy wektorowi jego warstwę). Jest to dobrze określone przekształcenie liniowe. Zachodzi następujące twierdzenie.

Twierdzenie 17.4.2: Własność uniwersalna przestrzeni ilorazowej

Niech V będzie przestrzenią liniową, zaś U – jej podprzestrzenią. Wówczas dla każdego przekształcenia liniowego $\phi : V \rightarrow W$ takiego, że $\ker(\phi)$ zawiera U , istnieje dokładnie jedno przekształcenie liniowe $\psi : V/U \rightarrow W$ takie, że $\phi = \pi \circ \psi$, czyli następujący diagram jest przemienny.

$$\begin{array}{ccc} V & \xrightarrow{\phi} & W \\ & \searrow \pi & \nearrow \psi \\ & V/U & \end{array}$$

W szczególności dowolne przekształcenie liniowe $\phi : V \rightarrow W$ faktoryzuje się przez odpowiednie przekształcenie $\psi : V/\ker(\phi) \rightarrow W$ dane wzorem: $\psi(\alpha + \ker(\phi)) = \phi(\alpha)$.

Dowód. Określamy przekształcenie liniowe ψ na dowolnej warstwie $\alpha + U$ wektora $\alpha \in V$:

$$\psi(\alpha + U) = \phi(\alpha).$$

Przekształcenie to jest dobrze określone na V/U , bo dla $\alpha, \alpha' \in U$ takich, że $\alpha + U = \alpha' + U$ mamy

$$\phi(\alpha) = \phi(\alpha') \iff \phi(\alpha - \alpha') = 0 \iff \alpha - \alpha' \in \ker(\phi).$$

Skoro jednak $\alpha + U = \alpha' + U$, to $\alpha - \alpha' \in U$. Skoro zaś $U \subseteq \ker(\phi)$, to $\phi(\alpha - \alpha') = 0$, czyli $\phi(\alpha) = \phi(\alpha')$.

Liniowość ψ jest prostą konsekwencją liniowości ϕ oraz działań w V/U :

$$\psi((\alpha + U) + (\alpha' + U)) = \psi((\alpha + \alpha' + U)) = \phi(\alpha + \alpha') = \phi(\alpha) + \phi(\alpha') = \psi(\alpha + U) + \psi(\alpha' + U).$$

$$\psi(\lambda(\alpha + U)) = \psi(\lambda\alpha + U) = \phi(\lambda\alpha) = \lambda\phi(\alpha) = \lambda\psi(\alpha + U).$$

Oczywiście co najwyżej jedna funkcja spełniać może warunek $\phi = \pi \circ \psi$, co kończy dowód. \square

Z powyższego rezultatu płyną następujące wnioski, zwane twierdzeniami o izomorfizmie.

Twierdzenie 17.4.3: Pierwsze twierdzenie o izomorfizmie

Niech V, W będą przestrzeniami liniowymi nad ciałem K oraz niech $\phi : V \rightarrow W$ będzie przekształceniem liniowym. Wówczas przekształcenie $\psi : V/\ker(\phi) \rightarrow \text{im}(\phi)$ dane wzorem

$$\psi(\alpha + \ker(\phi)) = \phi(v).$$

jest izomorfizmem przestrzeni liniowych $V/\ker(\phi)$ oraz $\text{im}(\phi)$.

Twierdzenie 17.4.4: Drugie twierdzenie o izomorfizmie

Niech V będzie przestrzenią liniową, zaś $U, W \subseteq V$ będą jej podprzestrzeniami. Wówczas mamy izomorfizm przestrzeni liniowych:

$$U/(U \cap W) \simeq (U + W)/W.$$

Twierdzenie 17.4.5: Trzecie twierdzenie o izomorfizmie

Niech V będzie przestrzenią liniową, W — będzie podprzestrzenią V , zaś U — podprzestrzeń W . Wówczas W/U jest podprzestrzenią V/U i mamy izomorfizm przestrzeni liniowych:

$$(V/U)/(W/U) \simeq V/W.$$

Twierdzenia te mają spore znaczenie na przykład w teorii Jordana lub teorii iloczynów tensorowych. Zostawiam Czytelnikowi ich dowody, między innymi sprawdzenie, że przekształcenie liniowe ψ określone w pierwszym twierdzeniu tak, jak we własności uniwersalnej, jest rzeczywiście izomorfizmem. To łatwe ćwiczenie. Warto natomiast wspomnieć jeszcze o wyniku, zwanym twierdzeniem o odpowiedniości.

Twierdzenie 17.4.6: Twierdzenie o odpowiedniości

Niech V będzie przestrzenią liniową oraz W jej podprzestrzenią. Przez $S(X)$ oznaczmy zbiór wszystkich podprzestrzeni przestrzeni liniowej X . Wówczas ma miejsce bijekcja

$$S(V/W) \longleftrightarrow \{U \in S(V) : W \subseteq U \subseteq V\}$$

polegająca na przypisaniu podprzestrzeni U spełniającej warunek $W \subseteq U \subseteq V$ podprzestrzeni U/W przestrzeni V/W . Bijekcja ta zachowuje sumy i przecięcia podprzestrzeni (jest izomorfizmem krat).

17.5 Dodatek. Ciągi dokładne

Opowiemy w tym dodatku o specjalnym typie ciągów, mającym ogromne zastosowania w algebrze, topologii i geometrii. Powiemy też więcej o rozumowaniach, które są typowe dla diagramów przekształceń.

Definicja 17.5.1: Ciąg dokładny

Ciąg przekształceń liniowych

$$V_1 \xrightarrow{\phi_1} V_2 \xrightarrow{\phi_2} V_3 \xrightarrow{\phi_3} \cdots \xrightarrow{\phi_{n-2}} V_{n-1} \xrightarrow{\phi_{n-1}} V_n \quad (\star)$$

nazywamy DOKŁADNYM, jeśli dla każdego $1 \leq i \leq n-2$ mamy $\text{im } \phi_i = \ker \phi_{i+1}$.

Przykłady.

- Ciąg przekształceń zerowych jest zawsze dokładny. Będziemy też stosowali następującą konwencję: pisząc $0 \longrightarrow V$ lub $V \longrightarrow 0$ mamy na myśli przekształcenia zerowe z przestrzeni zerowej lub do przestrzeni zerowej (nad ustalonym ciałem).
- Rozważmy ciąg

$$0 \rightarrow \mathbb{R} \xrightarrow{f} \mathbb{R}^2 \xrightarrow{g} \mathbb{R}^2 \xrightarrow{h} \mathbb{R} \rightarrow 0$$

gdzie

$$f(t) = (t, 0), \quad g(x, y) = (0, y), \quad h(x, y) = y.$$

Ciąg też jest dokładny, gdyż

- $\ker(f) = \{(0, 0)\}$, co jest obrazem pierwszego przekształcenia zerowego,
- $\text{im}(f) = \{(t, 0) \mid t \in \mathbb{R}\} = \ker(g)$,
- $\text{im}(g) = \{(0, y) \mid y \in \mathbb{R}\} = \ker(h)$,
- $\text{im}(h) = \mathbb{R}$, co jest jądrem ostatniego przekształcenia zerowego.

- Niech U, W będą podprzestrzeniami przestrzeni liniowej V oraz $V = U \oplus W$, czyli dla każdego $v \in V$ mamy $v = u + w$, gdzie $u \in U, w \in W$. Wówczas ciąg

$$0 \rightarrow U \xrightarrow{\phi} V \xrightarrow{\psi} W \rightarrow 0$$

$\phi(u) = u$ oraz $\psi(v) = w$ jest ciągiem dokładnym.

- Niech $\phi : V \rightarrow W$ będzie przekształceniem liniowym. Ciąg

$$0 \rightarrow \ker \phi \xrightarrow{i} V \xrightarrow{\pi} V/\ker \phi \rightarrow 0$$

gdzie i jest inkluzją, a π — naturalnym rzutowaniem jest również ciągiem dokładnym. Jak wiemy, przestrzeń $V/\ker \phi$ jest izomorficzna z $\text{im } \phi$.

Szczególnie ważne w zastosowaniach jest rozważanie diagramów, których wiersze lub kolumny są dokładne. Oto przykład problemu tego typu.

Zadanie. Załóżmy, że w diagramie przemiennym odwzorowań liniowych

$$\begin{array}{ccccccc} U_1 & \xrightarrow{r_1} & U_2 & \xrightarrow{r_2} & U_3 & \xrightarrow{r_3} & U_4 & \xrightarrow{r_4} & U_5 \\ \downarrow f_1 & & \downarrow f_2 & & \downarrow f_3 & & \downarrow f_4 & & \downarrow f_5 \\ V_1 & \xrightarrow{s_1} & V_2 & \xrightarrow{s_2} & V_3 & \xrightarrow{s_3} & V_4 & \xrightarrow{s_4} & V_5 \end{array}$$

wiersze są ciągami dokładnymi. Wykaż, że gdy f_1, f_2, f_4, f_5 są izomorfizmami, to f_3 jest izomorfizmem.

Mówiąc, że górnny (i dolny) wiersz omawianego diagramu jest dokładny mamy na myśli to, że wszystkie poniższe ciągi $U_1 \xrightarrow{r_1} U_2 \xrightarrow{r_2} U_3$, $U_2 \xrightarrow{r_2} U_3 \xrightarrow{r_3} U_4$, $U_3 \xrightarrow{r_3} U_4 \xrightarrow{r_4} U_5$ są dokładne (analogicznie w dolnym wierszu).

Przedstawmy rozwiązańe, uzasadniając że f_3 jest monomorfizmem i epimorfizmem. Przypomnijmy diagram i przemienności, które mają miejsce

$$\begin{array}{ccccccc} U_1 & \xrightarrow{r_1} & U_2 & \xrightarrow{r_2} & U_3 & \xrightarrow{r_3} & U_4 & \xrightarrow{r_4} & U_5 \\ \downarrow f_1 & & \downarrow f_2 & & \downarrow f_3 & & \downarrow f_4 & & \downarrow f_5 \\ V_1 & \xrightarrow{s_1} & V_2 & \xrightarrow{s_2} & V_3 & \xrightarrow{s_3} & V_4 & \xrightarrow{s_4} & V_5 \end{array}$$

- Odwzorowanie f_3 jest monomorfizmem. Niech $x_3 \in U_3$ i $f_3(x_3) = 0$. Z przemienności diagramu mamy:

$$f_4(r_3(x_3)) = s_3(f_3(x_3)).$$

Ponieważ $f_3(x_3) = 0$, to $s_3(0) = 0$, a więc:

$$f_4(r_3(x_3)) = 0.$$

Ponieważ f_4 jest izomorfizmem, mamy $r_3(x_3) = 0$. Z dokładności górnego wiersza $\ker(r_3) = \text{im}(r_2)$, istnieje więc $x_2 \in U_2$, takie że $x_3 = r_2(x_2)$. Korzystając dalej z przemienności diagramu:

$$f_3(r_2(x_2)) = s_2(f_2(x_2)).$$

widzimy, że ponieważ $f_3(x_3) = 0$, to $s_2(f_2(x_2)) = 0$. Z dokładności dolnego wiersza $\ker(s_2) = \text{im}(s_1)$, więc $f_2(x_2) \in \text{im}(s_1)$. Ponieważ f_2 jest izomorfizmem, mamy $x_2 \in \text{im}(r_1)$, a więc $x_2 = r_1(x_1)$ dla pewnego $x_1 \in U_1$. Wówczas:

$$x_3 = r_2(r_1(x_1)) = 0.$$

Zatem f_3 jest monomorfizmem

- Odwzorowanie f_3 jest epimorfizmem. Niech $y_3 \in V_3$. Z dokładności dolnego wiersza $\ker(s_3) = \text{im}(s_2)$, więc istnieje $y_2 \in V_2$, takie że $s_2(y_2) = y_3$. Ponieważ f_2 jest izomorfizmem, istnieje $x_2 \in U_2$, takie że $f_2(x_2) = y_2$. Wówczas:

$$f_3(r_2(x_2)) = s_2(f_2(x_2)) = s_2(y_2) = y_3.$$

Zatem f_3 jest epimorfizmem.

Jak widać rozumowania te mogą być całkiem delikatne. Z czasem jednak uzasadnienia stają się bardziej czytelne, gdy korzystamy z gotowych znanych już faktów o ciągach dokładnych. Powyższy zadanie jest w istocie przypadkiem szczególnym tzw. lematu o piątce — podstawowego narzędzia omawianej teorii. Jest on o tyle ogólniejszy, że o f_1 zakładamy tylko, że jest epimorfizmem, a o f_5 — że jest monomorfizmem. Wyników tego typu jest zresztą więcej.

W zbiorze zadań dra Kubata znajdują Państwo szereg kolejnych zadań związanych z ciągami dokładnymi. Przytoczę dwa z nich, które mają szczególnie istotne zastosowania. Pierwszy dotyczy zagadnienia tzw. rozszczepialności.

Twierdzenie 17.5.2

Założmy, że

$$0 \longrightarrow U \xrightarrow{f} V \xrightarrow{g} W \longrightarrow 0$$

jest ciągiem dokładnym przestrzeni liniowych. Następujące warunki są równoważne:

- istnieje takie $r \in L(V, U)$, że $r \circ f = \text{id}_U$.
- istnieje takie $s \in L(W, V)$, że $g \circ s = \text{id}_W$.
- istnieją takie $r \in L(V, U)$ oraz $s \in L(W, V)$, że $r \circ f = \text{id}_U$, $g \circ s = \text{id}_W$, $r \circ s = 0$ oraz $f \circ r + s \circ g = \text{id}_V$.

Gdy ciąg dokładny spełnia powyższe warunki, to mówimy, że ROZSZCZEPIA SIĘ. Pojęcie to ma duże znaczenie np. w teorii grup czy modułów, gdzie nie każdy ciąg się rozszczepia (tak jest, jak mówi twierdzenie, dla ciągów dokładnych przestrzeni liniowych).

Przytaczam jeszcze dwa zadania, których dowody mogą być pouczające dla osób rozpoczynających pracę z diagramami.

Twierdzenie 17.5.3: Lemat dziewiątkowy

Założymy, że w diagramie przemiennym odwzorowań liniowych

$$\begin{array}{ccccccc}
 & 0 & & 0 & & 0 & \\
 & \downarrow & & \downarrow & & \downarrow & \\
 0 & \longrightarrow & U_1 & \longrightarrow & U_2 & \longrightarrow & U_3 \longrightarrow 0 \\
 & \downarrow & & \downarrow & & \downarrow & \\
 0 & \longrightarrow & V_1 & \xrightarrow{f} & V_2 & \xrightarrow{g} & V_3 \longrightarrow 0 \\
 & \downarrow & & \downarrow & & \downarrow & \\
 0 & \longrightarrow & W_1 & \longrightarrow & W_2 & \longrightarrow & W_3 \longrightarrow 0 \\
 & \downarrow & & \downarrow & & \downarrow & \\
 & 0 & & 0 & & 0 &
 \end{array}$$

kolumny są krótkimi ciągami dokładnymi. Wówczas

- a) gdy pierwszy i drugi wiersz są krótkimi ciągami dokładnymi, to trzeci wiersz także.
- b) gdy drugi i trzeci wiersz są krótkimi ciągami dokładnymi, to pierwszy wiersz także.
- c) gdy pierwszy i trzeci wiersz są krótkimi ciągami dokładnymi oraz $g \circ f = 0$, to drugi wiersz jest również krótkim ciągiem dokładnym.

Drugie słynne zadanie wiąże się z pojęciem kojądra przekształcenia liniowego $\phi : V \rightarrow W$, które to określamy jako iloraz $W/\text{im } \phi$ (wspomniamy o nim przy własności uniwersalnej jądra, jako o obiekcie, którego własność uniwersalna zadana jest przez graf z odwróconymi strzałkami).

Twierdzenie 17.5.4: Lemat o węźlu

Przypuśćmy, że w diagramie przemiennym

$$\begin{array}{ccccccc}
 U_1 & \xrightarrow{g_1} & U_2 & \xrightarrow{g_2} & U_3 & \longrightarrow & 0 \\
 \downarrow f_1 & & \downarrow f_2 & & \downarrow f_3 & & \\
 0 & \longrightarrow & V_1 & \xrightarrow{h_1} & V_2 & \xrightarrow{h_2} & V_3
 \end{array}$$

wiersze są ciągami dokładnymi przestrzeni liniowych nad ciałem K . Wówczas istnieje ciąg dokładny

$$\begin{array}{ccccccc}
 0 & \longrightarrow & \ker g_1 & \longrightarrow & \ker f_1 & \xrightarrow{\bar{g}_1} & \ker f_2 \xrightarrow{\bar{g}_2} \ker f_3 \longrightarrow \\
 & & & & \downarrow & & \\
 & & & & \text{coker } f_1 & \xrightarrow{\bar{h}_1} & \text{coker } f_2 \xrightarrow{\bar{h}_2} \text{coker } f_3 \longrightarrow \text{coker } h_2 \longrightarrow 0,
 \end{array}$$

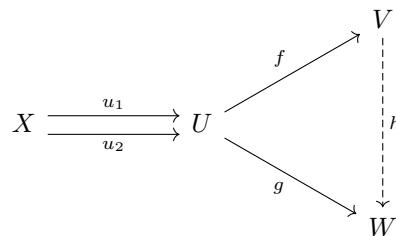
W zadaniu tym przekształcenia $\bar{g}_1, \bar{g}_2, \bar{h}_1, \bar{h}_2$ nazywanymi indukowanymi przez, odpowiednio, g_1, g_2, h_1, h_2 (warto pomyśleć co to w istocie oznacza). Warto też pomyśleć czym są nienazwane odwzorowania z powyższego ciągu dokładnego? Tyle uwag, a te i znacznie więcej znajdą Państwo w zbiorze dra Kubata.

17.6 Trivia. Być epimorfizmem, być monomorfizmem

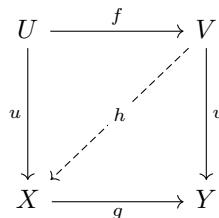
Wspomnieliśmy na wykładzie o własnościach uniwersalnych. Będą one wracać zwłaszcza w kursie na poziomie gwiazdkowym, choćby w definicji iloczynu tensorowego. Zobaczmy, że jeden obiekt można definiować za pomocą różnych własności. Zostawiam Czytelnikowi rozwiązanie następującego problemu.

Zadanie. Niech $f \in L(U, V)$. Wykaż, że następujące warunki są równoważne:

- odwzorowanie f jest epimorfizmem,
- istnieją takie $u_1, u_2 \in L(X, U)$, że $f \circ u_1 = f \circ u_2$ oraz dla dowolnego $g \in L(U, W)$ spełniającego $g \circ u_1 = g \circ u_2$ istnieje dokładnie jeden taki $h \in L(V, W)$, że $g = h \circ f$ (patrz diagram poniżej).



- dla dowolnych $u \in L(U, X)$, $v \in L(V, Y)$ i monomorfizmu $g \in L(X, Y)$, jeśli $v \circ f = g \circ u$, to istnieje taki $h \in L(V, X)$, że $u = h \circ f$ oraz $v = g \circ h$ (patrz diagram poniżej),



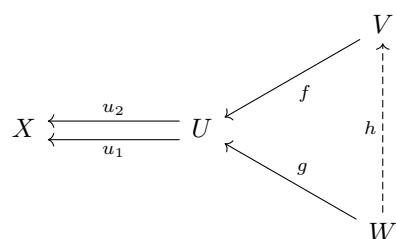
- dla dowolnych $p \in L(U, X)$ oraz $u \in L(X, V)$ z faktu, że u jest monomorfizmem oraz $f = u \circ p$ wynika, że u jest izomorfizmem.

Epimorfizm $f \in L(U, V)$ spełniający (2), (3) lub (4) nazywamy, odpowiednio, *regularnym*, *silnym* lub *ekstremalnym*. Zadanie to pokazuje, że w kategorii przestrzeni liniowych (nad ustaloną ciałem) pojęcia te są równoważne. Nie jest tak jednak w innych kategoriach (np. w kategorii grup).

Jeśli Czytelnikowi to zadanie nie wydaje się porywające, warto zastanowić się co się stanie, gdy w zadaniu tym zamienimy zwrot wszystkich strzałek (i kierunek wszystkich złożień), a w sformułowaniu zamienimy wszystkie słowa „monomorfizm” na „epimorfizm” i odwrotnie. Czy teza pozostanie prawdziwa?

A zatem, dla przykładu, proszę rozstrzygnąć czy następujące warunki są równoważne.

- Odwzorowanie f jest monomorfizmem.
- Istnieją takie $u_1, u_2 \in L(U, X)$, że $u_1 \circ f = u_2 \circ f$ oraz dla dowolnego $g \in L(W, U)$ spełniającego $u_1 \circ g = u_2 \circ g$ istnieje dokładnie jeden taki $h \in L(W, V)$, że $g = f \circ h$ (patrz diagram poniżej).



Rozważanie takich dualnych konstrukcji jest jednym z podstawowych sposobów operowania teorii kategorii — gdzie do różnych obiektów istnieją obiekty dualne (czasem istnieją, czasem nie). Ta wewnętrzna symetria, mająca fundamentalne skutki dla całej matematyki, nie jest wcale czymś banalnym, bowiem czasem o obiekcie umiemy dużo powiedzieć (jak o jądrze), a z obiektem dualnym (np. kojadrzem) mamy znacznie mniej do czynienia i nie rozumiemy go tak dobrze. Ma to skutki nie tylko w matematyce, ale i fizyce.

Rozdział 18

Mnożenie macierzy Macierz odwrotna

18.1 Wykład 18

Na poprzednim wykładzie wykazaliśmy, że przestrzeń przekształceń liniowych z przestrzenią n -wymiarowej do przestrzeni m -wymiarowej jest izomorficzna z przestrzenią macierzy rozmiaru $m \times n$. To oznacza, że każdemu przekształceniu liniowemu przestrzeni skończenie wymiarowych przypisać można pewną macierz. Jak się przekonamy na kolejnym wykładzie, można to zrobić na wiele sposobów. Zaczniemy jednak od najprostszego, wymagającego jedynie baz standardowych przestrzeni współrzędnych. Wcześniej jednak wprowadzimy kluczowe dla całego kursu pojęcie iloczynu macierzy.

Definicja 18.1.1: Mnożenie macierzy

Niech $A \in M_{1 \times n}(K)$ oraz $B \in M_{n \times 1}(K)$ będą postaci:

$$A = [a_1 \ a_2 \ a_3 \ \dots \ a_n], \quad B = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \\ \vdots \\ b_n \end{bmatrix}.$$

Wówczas iloczynem $A \cdot B$ macierzy A, B nazywamy macierz rozmiaru 1×1 , której jedyny wyraz ma postać:

$$\sum_{i=1}^n a_i b_i = a_1 b_1 + a_2 b_2 + \dots + a_n b_n.$$

ILOCZYNEM MACIERZY

$$A = [a_{ij}] \in M_{m \times k}(K) \quad \text{przez} \quad B = [b_{ij}] \in M_{k \times n}(K)$$

nazywamy taką macierz $C = [c_{ij}] \in M_{m \times n}(K)$ rozmiarów $m \times n$, że dla każdych i, j mamy:

$$c_{ij} = \sum_{l=1}^k a_{il} b_{lj} = a_{i1} b_{1j} + a_{i2} b_{2j} + \dots + a_{ik} b_{kj}$$

Innymi słowy: wyraz w i -tym wierszu i j -tej kolumnie macierzy C to jedyny wyraz macierzy będącej iloczynem i -tego wiersza macierzy A (rozmiaru $1 \times k$) i j -tego wiersza macierzy B (rozmiaru $k \times 1$).

Przykład. Jeśli $A = \begin{bmatrix} 1 & 0 & 2 \\ 1 & 3 & 1 \end{bmatrix}$, $B = \begin{bmatrix} 1 & 3 & 1 & 0 \\ 2 & 1 & 0 & 0 \\ 1 & 0 & 0 & 1 \end{bmatrix}$ to nie istnieje iloczyn macierzy postaci BA , zaś

$$AB = \begin{bmatrix} 1 \cdot 1 + 0 \cdot 2 + 2 \cdot 1 & 1 \cdot 3 + 0 \cdot 1 + 2 \cdot 0 & 1 \cdot 1 + 0 \cdot 0 + 2 \cdot 0 & 1 \cdot 0 + 0 \cdot 0 + 2 \cdot 1 \\ 1 \cdot 1 + 3 \cdot 2 + 1 \cdot 1 & 1 \cdot 3 + 3 \cdot 1 + 1 \cdot 0 & 1 \cdot 1 + 3 \cdot 0 + 1 \cdot 0 & 1 \cdot 0 + 3 \cdot 0 + 1 \cdot 1 \end{bmatrix} = \begin{bmatrix} 3 & 3 & 1 & 2 \\ 8 & 6 & 1 & 1 \end{bmatrix}.$$

Składanie funkcji nie jest przemienne. Nie jest przemienne również mnożenie macierzy. Na przykład, jeśli

$$A = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \quad B = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} \quad \text{, to} \quad AB = \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}, \quad BA = \begin{bmatrix} 0 & 1 \\ 0 & 1 \end{bmatrix}.$$

Ważna motywacja 1. Jak niemal nigdy na tym wykładzie, wykorzystamy motywację praktyczną. Oto „zadanie z treścią”.

Ania prowadzi kawiarnię, oferując różne napoje. Do każdego napoju używa kombinacji składników: liść herbaciany, plaster cytryny, łyżeczka ziarenek kawy oraz małe opakowanie mleka. Oto przepisy na napoje:

- szklanka mleka: wymaga jednego małego opakowania mleka,
- herbatka cytrynowa: wymaga dwóch listków herbaty i jednego plastra cytryny,
- kawa: wymaga dwóch łyżeczek ziarenek.

Aby napoje sprzedawały się lepiej, Ania oferuje je w zestawach:

- zestaw 1 — dwie szklanki mleka i jedna herbatka,
- zestaw 2 — jedna szklanka mleka, dwie kawy i jedna herbatka.

Pytanie brzmi: ile składników potrzeba do przygotowania każdego zestawu? Ktoś powie, to przecież banalny rachunek. Oto zaś stosowny obrazek (zapozyczam go z ciekawych wykładów Qirui Li — <http://qirui.li/>)

	🥛	☕	🍵
🌿	0	0	2
🍋	0	0	1
☕	0	2	0
🐄	1	0	0

•

	🍱	🍜
🥛	2	1
☕	0	2
🍵	1	1
🐄	0	4

=

	🍱	🍜
🌿	2	1
🍋	1	1
☕	0	4
🐄	2	1

Czytelnik bez trudu sprawdzi, że rachunek wyżej rozwiązuje zadanie. Wyjaśnienie dodatkowej wartości podejścia do problemu przez mnożenie odpowiednio sformułowanych macierzy dają poniższe dwie obserwacje, które uzasadnimy dalej.

- Gdy pytamy ile składników potrzebujemy, aby przygotować zestaw 1, patrzymy na pierwszą kolumnę uzyskanego iloczynu i widzimy, że jest ona kombinacją liniową kolumn pierwszej macierzy ze współczynnikami będącymi wyrazami pierwszej kolumny drugiej macierzy (odpowiadającą zestawowi 1).

	🥛	☕	🍵
🌿	0	0	2
🍋	0	0	1
☕	0	2	0
🐄	1	0	0

•

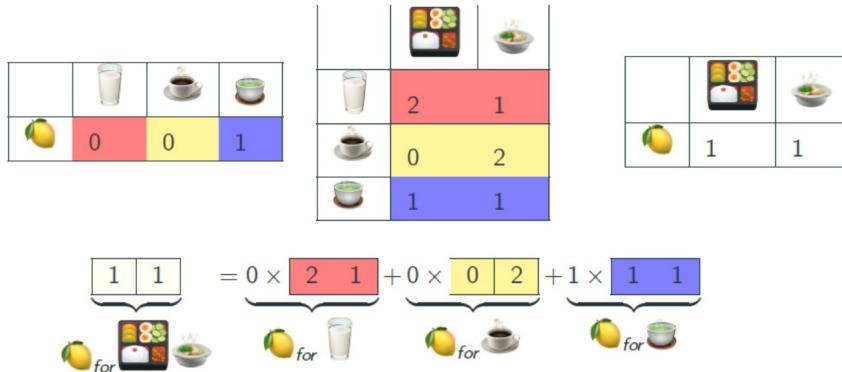
	🍱	🍜
🥛	2	1
☕	0	2
🍵	1	1
🐄	0	4

=

	🍱	🍜
🌿	2	1
🍋	1	1
☕	0	4
🐄	2	1

🍱 :	<table border="1" style="display: inline-table; vertical-align: middle;"> <tr><td>2</td></tr> <tr><td>1</td></tr> <tr><td>0</td></tr> <tr><td>2</td></tr> </table>	2	1	0	2	= 2 × <table border="1" style="display: inline-table; vertical-align: middle;"> <tr><td>0</td></tr> <tr><td>0</td></tr> <tr><td>1</td></tr> </table>	0	0	1	+ 0 × <table border="1" style="display: inline-table; vertical-align: middle;"> <tr><td>0</td></tr> <tr><td>2</td></tr> <tr><td>0</td></tr> </table>	0	2	0	+ 1 × <table border="1" style="display: inline-table; vertical-align: middle;"> <tr><td>1</td></tr> <tr><td>0</td></tr> <tr><td>0</td></tr> </table>	1	0	0
2																	
1																	
0																	
2																	
0																	
0																	
1																	
0																	
2																	
0																	
1																	
0																	
0																	
		2 ×	0 ×	1 ×													

- Gdy pytamy, ile cytryn potrzeba do przygotowania odpowiedniego dania, patrzymy na wiersz drugi iloczynu i przekonujemy się, że jest on kombinacją liniową wierszy drugiego czynnika ze współczynnikami z wiersza pierwszego czynnika odpowiadającego cytrynie.



Obserwacja 18.1.2: Wiersze i kolumny iloczynu

Niech $A \in M_{m \times n}(K)$ oraz $B \in M_{n \times k}(K)$. Wówczas:

- i -ta kolumna macierzy AB traktowana jako macierz rozmiaru $m \times 1$ powstaje przez iloczyn macierzy A oraz i -tej kolumny macierzy B (traktowanej jako macierz $n \times 1$), czyli jest kombinacją liniową kolejnych kolumn macierzy A ze współczynnikami będącymi kolejnymi wyrazami i -tej kolumny macierzy B ,
- i -ty wiersz macierzy AB traktowany jako macierz rozmiaru $1 \times k$ powstaje przez iloczyn i -tego wiersza macierzy A (traktowanego jako macierz $1 \times n$) oraz macierzy B , czyli jest kombinacją liniową kolejnych wierszy macierzy B ze współczynnikami będącymi kolejnymi wyrazami i -tego wiersza macierzy A .

Dowód. Przeprowadzimy uzasadnienie pierwszego punktu, drugi zostawiając już Czytelnikowi. Niech $A = [a_{ij}]$ oraz $B = [b_{kl}]$. Wówczas i -ta kolumna macierzy AB ma z definicji postać (jako macierz $m \times 1$)

$$\begin{bmatrix} a_{11}b_{1i} + a_{12}b_{2i} + \dots + a_{1n}b_{ni} \\ \vdots \\ a_{m1}b_{1i} + a_{m2}b_{2i} + \dots + a_{mn}b_{ni} \end{bmatrix} = b_{1i} \cdot \begin{bmatrix} a_{11} \\ \vdots \\ a_{m1} \end{bmatrix} + b_{2i} \cdot \begin{bmatrix} a_{12} \\ \vdots \\ a_{m2} \end{bmatrix} + \dots + b_{ni} \cdot \begin{bmatrix} a_{1n} \\ \vdots \\ a_{mn} \end{bmatrix}.$$

Z drugiej strony i -ta kolumna macierzy AB powstaje jako iloczyn A oraz i -tej kolumny B (traktowanej jako macierz), co widać po lewej stronie równości wyżej, gdzie zapisany jest wynik tegoż mnożenia. \square

Wniosek 18.1.3

Niech $A \in M_{m \times n}(K)$ oraz $B \in M_{n \times k}(K)$. Wówczas $r(AB) \leq \min\{r(A), r(B)\}$.

Dowód. Rząd macierzy X równy jest z definicji $\dim w(X) = \dim k(X)$, gdzie $w(X)$ jest przestrzenią wierszową, a $k(X)$ — przestrzenią kolumnową macierzy X . Z Obserwacji 18.1.2 mamy:

$$k(AB) \subseteq k(A) \Rightarrow r(AB) \leq r(A) \quad \text{oraz} \quad w(AB) \subseteq w(B) \Rightarrow r(AB) \leq r(B).$$

\square

Ważna motywacja 2. Niech $\phi : K^n \rightarrow K^m$ będzie przekształceniem liniowym postaci:

$$\phi((x_1, \dots, x_n)) = (a_{11}x_1 + \dots + a_{1n}x_n, a_{21}x_1 + \dots + a_{2n}x_n, \dots, a_{m1}x_1 + \dots + a_{mn}x_n).$$

Równość wyżej jest spełniona wtedy i tylko wtedy, gdy dla każdego (x_1, \dots, x_n) zachodzi równość postaci

$$\begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix} \cdot \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} a_{11}x_1 + \dots + a_{1n}x_n \\ \vdots \\ a_{m1}x_1 + \dots + a_{mn}x_n \end{bmatrix}.$$

Warto zauważyć, że w powyższym równoważnym sformułowaniu wzoru ϕ kolumny rozważanej macierzy $[a_{ij}]$ są po prostu obrazami wektorów bazy standardowej przestrzeni K^n przy przekształceniu ϕ , a więc:

$$\phi(\epsilon_1) = \begin{bmatrix} \color{red}{a_{11}} & a_{12} & \dots & a_{1n} \\ \vdots & \vdots & \ddots & \vdots \\ \color{red}{a_{m1}} & a_{m2} & \dots & a_{mn} \end{bmatrix} \cdot \begin{bmatrix} 1 \\ \vdots \\ 0 \end{bmatrix} = \begin{bmatrix} \color{red}{a_{11}} \\ \vdots \\ \color{red}{a_{m1}} \end{bmatrix}, \dots, \phi(\epsilon_n) = \begin{bmatrix} a_{11} & a_{12} & \dots & \color{blue}{a_{1n}} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & \color{blue}{a_{mn}} \end{bmatrix} \cdot \begin{bmatrix} 0 \\ \vdots \\ 1 \end{bmatrix} = \begin{bmatrix} \color{blue}{a_{1n}} \\ \vdots \\ \color{blue}{a_{mn}} \end{bmatrix}$$

Zauważmy również, że jeśli przekształcenie liniowe $\psi : K^m \rightarrow K^r$ dane jest wzorem

$$\psi((y_1, \dots, y_m)) = (b_{11}y_1 + \dots + b_{1m}y_m, b_{21}y_1 + \dots + b_{2m}y_m, \dots, b_{r1}y_1 + \dots + b_{rm}y_m),$$

$\psi(\phi((x_1, \dots, x_n))) = (z_1, \dots, z_r)$ wtedy i tylko wtedy, gdy dla każdego (x_1, \dots, x_n) mamy:

$$\begin{bmatrix} b_{11} & b_{12} & \dots & b_{1m} \\ \vdots & \vdots & \ddots & \vdots \\ \vdots & \vdots & \ddots & \vdots \\ b_{r1} & b_{r2} & \dots & b_{rm} \end{bmatrix} \cdot \left(\begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix} \cdot \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} \right) = \begin{bmatrix} z_1 \\ \vdots \\ z_r \end{bmatrix}.$$

Niech $A = [a_{ij}]$ oraz $B = [b_{ij}]$. Każde przekształcenie liniowe zadane jest jednoznacznie na bazie, a wstawiając za (x_1, \dots, x_n) wektory bazy standardowej widzimy, że wartość $\psi \circ \phi$ na i -tym wektorze bazy standardowej dostajemy w istocie mnożąc macierz B przez i -tą kolumnę macierzy A . Na mocy Obserwacji 18.1.2 jest to zarazem i -ta kolumna macierzy BA . Stąd $\psi \circ \phi((x_1, \dots, x_n))$ wyznaczamy licząc

$$B \cdot \left(A \cdot \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} \right) = \left(\begin{bmatrix} b_{11} & b_{12} & \dots & b_{1m} \\ \vdots & \vdots & \ddots & \vdots \\ \vdots & \vdots & \ddots & \vdots \\ b_{r1} & b_{r2} & \dots & b_{rm} \end{bmatrix} \cdot \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix} \right) \cdot \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = (BA) \cdot \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}.$$

Wniosek 18.1.4: Łączność mnożenia macierzy

Jeśli dane są macierze $B \in M_{r \times m}(K)$, $A \in M_{m \times n}(K)$, $X \in M_{n \times s}(K)$, tak że iloczyny BA i AX mają sens, to mamy:

$$B(AX) = (BA)X.$$

Dowód. Niech i -ta kolumna X ma wyrazy x_{1i}, \dots, x_{ni} . Zgodnie z rozważaniami przedstawionymi wyżej mamy równość

$$B \cdot \left(A \cdot \begin{bmatrix} x_{1i} \\ \vdots \\ x_{ni} \end{bmatrix} \right) = (BA) \cdot \begin{bmatrix} x_{1i} \\ \vdots \\ x_{ni} \end{bmatrix} \quad (*)$$

Teza wynika zatem z Obserwacji 18.1.2 mówiącej, że dla dowolnych macierzy P, Q , dla których iloczyn PQ ma sens, i -ta kolumna macierzy PQ powstaje przez pomnożenie macierzy P przez i -tą kolumnę macierzy Q (traktowaną jako macierz). W ten sposób:

- po lewej stronie równości $(*)$ mnożymy B przez i -tą kolumnę AX dostając i -tą kolumnę macierzy $B(AX)$,
- po prawej stronie równości $(*)$ mnożymy BA przez i -tą kolumnę X , dostając i -tą kolumnę $(BA)X$.

Stąd i -te kolumny macierzy $B(AX)$ oraz $(BA)X$ są równe, dla dowolnego i , co kończy dowód. \square

Aby ukazać w jaki sposób łączność mnożenia wykorzystana być może do udowodnienia czegoś ciekawego, rozważmy następujący przykład. Jak wiemy, ciąg Fibonacciego określa się przez rekurencję $F_0 = 0$, $F_1 = 1$, $F_{n+2} = F_{n+1} + F_n$. Zauważmy, że jeśli

$$A = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}, \text{ to } A^n = \begin{bmatrix} F_{n+1} & F_n \\ F_n & F_{n-1} \end{bmatrix}$$

Kolejne ważne zastosowanie łączności mnożenia macierzy pokażemy w motywacji 3.

Ważna motywacja 3. Mnożenie macierzy pozwala sprowadzić rozwiązywanie układu równań do pewnego równania macierzowego.

Rozważmy układ równań o współczynnikach w \mathbb{R} postaci:

$$\begin{cases} x_1 + x_2 + x_3 = 6 \\ 2x_2 + 5x_3 = -4 \\ 2x_1 + 5x_2 - x_3 = 27 \end{cases}.$$

Zauważmy, że zgodnie z definicją mnożenia macierzy, układ ten można zapisać w postaci

$$\begin{bmatrix} 1 & 1 & 1 \\ 0 & 2 & 5 \\ 2 & 5 & -1 \end{bmatrix} \cdot \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 6 \\ -4 \\ 27 \end{bmatrix}.$$

Definicja 18.1.5: Postać macierzowa układu równań liniowych

Rozważmy układ U złożony m równań liniowych o n zmiennych i współczynnikach w ciele K postaci

$$\begin{cases} a_{11}x_1 + \dots + a_{1n}x_n = b_1 \\ \vdots \\ a_{m1}x_1 + \dots + a_{mn}x_n = b_m \end{cases}$$

Niech $A \in M_{m \times n}(K)$ będzie macierzą współczynników tego układu. Wówczas POSTACIA MACIERZOWĄ układu U nazywamy równanie

$$A \cdot \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} b_1 \\ \vdots \\ b_m \end{bmatrix}.$$

W skrócie równanie to zapisujemy często w postaci $Ax = b$, gdzie $x \in M_{n \times 1}(K)$ ma kolejne wyrazy równe kolejnym współrzędnym wektora (x_1, \dots, x_n) , zaś $B \in M_{m \times 1}(K)$ ma kolejne wyrazy równe kolejnym współrzędnym wektora (b_1, \dots, b_m) .

Załóżmy dalej, że A jest macierzą rozmiaru $n \times n$ oraz, że układ $Ax = b$ ma dokładnie jedno rozwiązanie. Zauważmy, że byłoby bardzo interesujące, gdyby istniała taka macierz B , że

$$B(Ax) = x = Bb$$

W takim bowiem przypadku istnienie macierzy B dawałoby nam rozwiązywanie układu równań. Przykładowo, dla układu równań $Ax = b$ postaci

$$\begin{bmatrix} 1 & 2 \\ 3 & 0 \end{bmatrix} \cdot \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

biorąc macierz $B = \begin{bmatrix} 0 & \frac{1}{3} \\ \frac{1}{2} & -\frac{1}{6} \end{bmatrix}$, dostajemy:

$$BAx = \begin{bmatrix} 0 & \frac{1}{3} \\ \frac{1}{2} & -\frac{1}{6} \end{bmatrix} \cdot \begin{bmatrix} 1 & 2 \\ 3 & 0 \end{bmatrix} \cdot \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

Stąd rozwiązanie układu równań wyliczamy wyznaczając iloczyn Bb postaci

$$\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 & \frac{1}{3} \\ \frac{1}{2} & -\frac{1}{6} \end{bmatrix} \cdot \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} \frac{1}{3} \\ \frac{1}{3} \end{bmatrix}.$$

W przykładzie pojawiła się macierz mająca duże znaczenie w całej teorii mnożenia macierzy i przekształceń liniowych.

Definicja 18.1.6: Macierz identycznościowa rozmiaru n

Niech $I_n \in M_{n \times n}(K)$ będzie macierzą, której wszystkie wyrazy ii (czyli w i -tym wierszu i i -tej kolumnie) są równe 1, dla $1 \leq i \leq n$, a pozostałe są równe 0, postaci

$$I_n = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{bmatrix}.$$

Macierz tę nazywać będziemy MACIERZĄ IDENTYCZNOŚCIOWĄ (rozmiaru n), ozn. I lub I_n , gdy chcemy podkreślić jej rozmiar.

Nietrudno widzieć, że jeśli $A \in M_{n \times m}(K)$ oraz $B \in M_{r \times n}(K)$, to

$$AI_n = A \quad I_n B = B.$$

Definicja 18.1.7: Macierz odwrotna

Powiemy, że macierz $B \in M_{n \times n}(K)$ jest ODWROTNĄ do macierzy $A \in M_{n \times n}(K)$, jeśli

$$AB = BA = I_n.$$

Macierz odwrotną do macierzy A oznaczamy, o ile istnieje, jako A^{-1} . Macierz, która ma odwrotną nazywamy MACIERZĄ ODWRACALNĄ.

Wyznaczanie macierzy odwrotnej do macierzy $A \in M_{n \times n}(K)$ można przeprowadzić, startując od macierzy o $2n$ kolumnach, której pierwsze n kolumn to kolejne kolumny macierzy A , a kolejne n kolumn to kolejne kolumny macierzy I_n . Oznaczamy ją przez $[A | I_n]$. Jeśli za pomocą elementarnych operacji wierszowych sprowadzimy taką macierz do postaci, w której pierwsze n kolumn to kolejne kolumny macierzy I_n , to n kolejnych kolumn powstałej macierzy to kolejne kolumny macierzy A^{-1} . Schematycznie algorytm przedstawia się następująco:

$$[A | I_n] \longrightarrow [I_n | A^{-1}].$$

Uzasadnienie: rozważmy równanie

$$AX = I_n,$$

gdzie $A \in M_{n \times n}(K)$ jest dana, natomiast $X \in M_{n \times n}(K)$ to szukana macierz odwrotna. Wówczas na mocy Obserwacji 18.1.2, i -ta kolumna macierzy X jest rozwiązaniem układu równań o macierzy rozszerzonej

$$[A | \epsilon_i],$$

gdzie ϵ_i to i -ty wektor bazy standardowej K^n . Innymi słowy — algorytm opisany wyżej jest w istocie algorytmem jednoczesnego rozwiązania n układów równań takich, jak wyżej.

Przykład. Wyznaczmy macierz odwrotną do macierzy

$$A = \begin{bmatrix} 1 & 2 \\ 3 & 0 \end{bmatrix} \in M_{2 \times 2}(\mathbb{R}).$$

Mamy

$$\left[\begin{array}{cc|cc} 1 & 2 & 1 & 0 \\ 3 & 0 & 0 & 1 \end{array} \right] \longrightarrow \left[\begin{array}{cc|cc} 1 & 2 & 1 & 0 \\ 0 & -6 & -3 & 1 \end{array} \right] \longrightarrow \left[\begin{array}{cc|cc} 1 & 2 & 1 & 0 \\ 0 & 1 & \frac{1}{2} & -\frac{1}{6} \end{array} \right] \longrightarrow \left[\begin{array}{cc|cc} 1 & 0 & 0 & \frac{1}{3} \\ 0 & 1 & \frac{1}{2} & -\frac{1}{6} \end{array} \right].$$

Rzeczywiście więc:

$$\begin{bmatrix} 1 & 2 \\ 3 & 0 \end{bmatrix} \cdot \begin{bmatrix} 0 & \frac{1}{3} \\ \frac{1}{2} & -\frac{1}{6} \end{bmatrix} = \begin{bmatrix} 0 & \frac{1}{3} \\ \frac{1}{2} & -\frac{1}{6} \end{bmatrix} \cdot \begin{bmatrix} 1 & 2 \\ 3 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = I_2.$$

* * *

Na kolejnym wykładzie powiążemy macierze w głębszy sposób z przekształceniemi liniowymi. Uzasadnimy również, że macierze odwracalne rozmiaru $n \times n$ są w istocie macierzami izomorfizmów przestrzeni n -wymiarowych, czyli macierzami rzędu n (w zasadzie już to wiemy — być może Czytelnik to wyjaśni).

18.2 Wybór przykładowych pytań

Proszę się nie uczyć na pamięć odpowiedzi czy wyjaśnienie! Takich pytań można ułożyć znacznie więcej i służą one jedynie sprawdzeniu podstawowego zrozumienia treści zajęć. Jeśli odpowiedź na dowolne z poniższych pytań sprawia Państwu problem, proszę pomyśleć który fragment notatek dotyczy tego pytania, wrócić do ich lektury, a w razie dalszych trudności poprosić prowadzącego o wskazówkę.

1. Macierz A rozmiaru 2×3 pomnożono przez macierz B , uzyskując macierz rozmiaru 2×3 . Jakich rozmiarów jest macierz B ?
2. Macierz C pomnożono przez macierz D rozmiaru 2×3 , uzyskując macierz rozmiaru 2×3 . Jaki jest rozmiar macierzy C ?
3. Dane są macierze $A \in M_{2 \times 4}(K)$, $B \in M_{4 \times 3}(K)$, $C \in M_{3 \times 2}(K)$. Jaki jest rozmiar macierzy $CABC$?
4. Wykonaj iloczyn:

$$\begin{bmatrix} 1 & n \\ 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} 1 & m \\ 0 & 1 \end{bmatrix}.$$

5. Macierze A, B spełniają warunek $\begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix} \cdot A = B \cdot \begin{bmatrix} 4 & 0 \\ 0 & 4 \end{bmatrix}$. Czy $2B = A$?

6. Niech

$$A^2 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}.$$

Jaki jest rozmiar macierzy A ? Podaj przykład macierzy $A \neq \pm I_2$, która spełnia powyższe równanie.

7. Podaj przykład dwóch różnych niezerowych macierzy $A, B \in M_{2 \times 2}(K)$ spełniających $AB = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$.
8. Wykonaj iloczyn

$$\begin{bmatrix} 4 & 1 \\ 5 & 1 \\ 6 & 1 \end{bmatrix} \cdot \begin{bmatrix} 1 \\ 3 \end{bmatrix}$$

Wynik jest macierzą 3×1 , której jedyna kolumna jest kombinacją liniową kolumn macierzy po prawej. Jakie są współczynniki tej kombinacji?

9. Macierz $A \in M_{3 \times 3}(K)$ ma identyczne kolumny: pierwszą i trzecią. Czy również AB ma identyczne kolumny: pierwszą i trzecią?
10. Macierz $A \in M_{3 \times 3}(K)$ ma identyczne wiersze: pierwszy i trzeci. Czy również AB ma identyczne wiersze: pierwszy i trzeci?
11. Dla dowolnego wektora $(x_1, x_2, x_3) \in \mathbb{R}^3$ istnieje taki wektor (y_1, y_2) , że

$$\begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & -1 \end{bmatrix} \cdot \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}.$$

Czy odwzorowanie przypisujące wektorowi (x_1, x_2, x_3) wektor (y_1, y_2) jest liniowe? Jaki jest wzór tego przekształcenia?

12. Rozważmy układ równań liniowych o 5 zmiennych i współczynnikach rzeczywistych postaci

$$\begin{cases} 3x_1 + 3x_2 + 2x_3 + 6x_4 + 4x_5 = 4 \\ x_1 + 2x_2 + x_3 + 4x_4 + 3x_5 = 1 \\ 2x_2 + 3x_3 + 4x_4 + x_5 = 4 \\ x_2 + 2x_4 + 2x_5 = -1 \end{cases}.$$

Zapisz ten układ w postaci macierzowej. Sprawdź przy pomocy mnożenia macierzy, czy wektor $(1, 1, 1, 0, -1)$ jest rozwiązaniem tego układu.

13. Znajdź macierze odwrotne do macierzy $M_{2 \times 2}(K)$, gdzie

$$\begin{bmatrix} 2 & 0 \\ 0 & 3 \end{bmatrix}, \quad \begin{bmatrix} 2 & 1 \\ 0 & 3 \end{bmatrix},$$

gdzie $K = \mathbb{R}$ lub $K = \mathbb{Z}_5$.

18.3 Zadania do samodzielnej pracy

Zadania ilustrujące umiejętności niezbędne do zaliczenia przedmiotu oznaczone są symbolem ♠. Przy tych zadaniach dodany jest opis umiejętności, które sprawdzają.

- (♣ Mnożenie macierzy) Wykonaj iloczyny macierzy o współczynnikach rzeczywistych:

$$\begin{bmatrix} \cos \alpha & -\sin \alpha \\ \sin \alpha & \cos \alpha \end{bmatrix} \cdot \begin{bmatrix} \cos \beta & -\sin \beta \\ \sin \beta & \cos \beta \end{bmatrix} \quad \text{oraz} \quad \begin{bmatrix} \cos \alpha & -\sin \alpha \\ \sin \alpha & \cos \alpha \end{bmatrix}^n, \quad \begin{bmatrix} 0 & 1 \\ -1 & 1 \end{bmatrix}^n.$$

- (♣) Wyznacz macierz $X \in M_{2 \times 2}(\mathbb{R})$ spełniającą:

$$\text{a)} \begin{bmatrix} 1 & 3 \\ 1 & 2 \end{bmatrix} \cdot X = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}, \quad \text{b)} X \cdot \begin{bmatrix} 2 & -1 \\ 4 & -2 \end{bmatrix} = \begin{bmatrix} 1 & 3 \\ 6 & 2 \end{bmatrix}.$$

- (♣ Mnożenie macierzy) Niech $A = \begin{bmatrix} 4 & 2 \\ 4 & 1 \end{bmatrix} \in M_{2 \times 2}(\mathbb{Z}_5)$. Wyznacz A^5 .

- (♣) Czy istnieje macierz $A \in M_{2 \times 2}(\mathbb{R})$, że $A^2 = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$?

- (♣ Znajdowanie macierzy odwrotnej)

Dla każdej z poniższych macierzy znajdź macierz odwrotną:

$$A_1 = \begin{bmatrix} 2 & 3 \\ 7 & 9 \end{bmatrix}, \quad A_2 = \begin{bmatrix} 3 & 9 & 8 \\ 2 & 7 & 8 \\ 1 & 3 & 2 \end{bmatrix}, \quad A_3 = \begin{bmatrix} 2 & 5 & 0 & 0 \\ 1 & 3 & 0 & 0 \\ 7 & 9 & 3 & 4 \\ 4 & 6 & 2 & 3 \end{bmatrix}$$

- Dla każdej z poniższych macierzy rozmiaru $n \times n$ znajdź macierz odwrotną:

$$\begin{bmatrix} 1 & 1 & \dots & 1 & 1 \\ 0 & 1 & \dots & 1 & 1 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \dots & 1 & 1 \\ 0 & 0 & \dots & 0 & 1 \end{bmatrix}, \quad \begin{bmatrix} 1 & 1 & 1 & \dots & 1 \\ 1 & 2 & 2 & \dots & 2 \\ 1 & 2 & 3 & \dots & 3 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & 2 & 3 & \dots & n \end{bmatrix}, \quad \begin{bmatrix} 0 & 1 & 1 & \dots & 1 \\ 1 & 0 & 1 & \dots & 1 \\ 1 & 1 & 0 & \dots & 1 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & 1 & 1 & \dots & 0 \end{bmatrix}, \quad \begin{bmatrix} 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & 1 \\ 1 & 0 & 0 & \dots & 0 \end{bmatrix}.$$

- Które z poniższych macierzy 2×2 o wyrazach rzeczywistych mają macierz odwrotną? Wyznacz tę macierz, jeśli istnieje.

$$A = \begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix}, \quad B = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}, \quad C = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}, \quad D = \begin{bmatrix} \cos \alpha & -\sin \alpha \\ \sin \alpha & \cos \alpha \end{bmatrix}.$$

- Niech $A \in M_{3 \times 2}(K)$ oraz $B \in M_{2 \times 3}(K)$ oraz

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

- Co można powiedzieć o trzecim wierszu macierzy A oraz trzeciej kolumnie macierzy B ?

- Wyznacz BA .

- Niech $A \in M_{m \times n}(K)$ oraz $B \in M_{m \times k}(K)$. Wykaż, że równanie $AX = B$ ma rozwiązanie, dla pewnej macierzy $X \in M_{n \times k}(K)$ wtedy i tylko wtedy, gdy $r(A) = r([A|B])$, gdzie $[A|B]$ jest macierzą o $n+k$ kolumnach, gdzie pierwsze n z nich to kolumny macierzy A , a ostatnie k — to kolumny macierzy B .

- Wykaż, że jeśli macierz $A \in M_{n \times n}(K)$ spełnia $AC = CA$, dla każdej macierzy $C \in M_{n \times n}(K)$, to $A = aI$, dla pewnego $a \in K$. Wskazówka: jakie macierze są przemienne z jedynkami macierzowymi?

- Niech $A \in M_{4 \times 2}(\mathbb{R})$ oraz $B \in M_{2 \times 4}(\mathbb{R})$ będą takimi macierzami, że:

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

Znajdź macierz BA (wskazówka: podziel macierz AB na cztery bloki).

Rozdział 19

Macierz przekształcenia liniowego. Zmiana bazy

19.1 Wykład 19

Kontynuujemy dziś w większej niż ostatnio ogólności wątek związków między macierzami, a przekształceniemi liniowymi. Wiemy, że przekształcenie liniowe $\phi : V \rightarrow W$ przestrzeni skończonego wymiaru określić można jednoznacznie na dowolnej bazie \mathcal{A} przestrzeni V . Możemy więc przyjrzeć się obrazom elementów bazy \mathcal{A} i odczytywać ich współrzędne w (dowolnej) bazie przestrzeni W .

Definicja 19.1.1: Macierz przekształcenia liniowego

Niech V, W będą przestrzeniami liniowymi nad ciałem K i niech $\phi : V \rightarrow W$ będzie przekształceniem liniowym. Niech

- $\mathcal{A} = (\alpha_1, \dots, \alpha_n)$ będzie (uporządkowaną) bazą przestrzeni V ,
- $\mathcal{B} = (\beta_1, \dots, \beta_m)$ będzie (uporządkowaną) bazą przestrzeni W .

MACIERZĄ PRZEKSZTAŁCENIA ϕ w bazach \mathcal{A}, \mathcal{B} nazywamy taką macierz $A = [a_{ij}] \in M_{m \times n}(K)$, że dla każdego $1 \leq j \leq n$:

$$\phi(\alpha_j) = a_{1j}\beta_1 + a_{2j}\beta_2 + \dots + a_{mj}\beta_m = \sum_{i=1}^m a_{ij}\beta_i,$$

Taką macierz A oznaczamy $M(\phi)_{\mathcal{A}}^{\mathcal{B}}$. Innymi słowy:

w j -tej kolumnie macierzy $M(\phi)_{\mathcal{A}}^{\mathcal{B}}$ stoją współrzędne wektora $\phi(\alpha_j)$ w bazie \mathcal{B} .

Przykład. Przekształcenie liniowe $\phi : \mathbb{R}^3 \rightarrow \mathbb{R}^2$ dane jest wzorem

$$\phi((x_1, x_2, x_3)) = (2x_1 + x_2 - x_3, x_1 - x_2 + x_3).$$

Rozważamy bazy:

- $\mathcal{A} = (\alpha_1, \alpha_2, \alpha_3)$, gdzie $\alpha_1 = (1, 0, 1), \alpha_2 = (0, 1, 2), \alpha_3 = (2, 1, 0)$ – baza przestrzeni \mathbb{R}^3 ,
- $\mathcal{B} = (\beta_1, \beta_2)$, gdzie $\beta_1 = (0, 1), \beta_2 = (2, 0)$ – baza przestrzeni \mathbb{R}^2 .

Wówczas:

$$\phi(\alpha_1) = (1, 2) = 2 \cdot \beta_1 + \frac{1}{2} \cdot \beta_2, \quad \phi(\alpha_2) = (-1, 1) = 1 \cdot \beta_1 - \frac{1}{2} \cdot \beta_2, \quad \phi(\alpha_3) = (5, 1) = 1 \cdot \beta_1 + \frac{5}{2} \cdot \beta_2.$$

Możemy zatem przypisać przekształceniu ϕ macierz, która w kolumnach będzie miała współrzędne obrazów kolejnych wektorów z bazy \mathcal{A} , ale współrzędne te będą w bazie \mathcal{B} . Macierz tą oznaczać będziemy jako: $M(\phi)_{\mathcal{A}}^{\mathcal{B}}$, tzn.

$$M(\phi)_{\mathcal{A}}^{\mathcal{B}} = \begin{bmatrix} 2 & 1 & 1 \\ \frac{1}{2} & -\frac{1}{2} & \frac{5}{2} \end{bmatrix}.$$

Oczywiście mamy też

$$M(\phi)_{\text{st}}^{\text{st}} = \begin{bmatrix} \color{red}{2} & \color{blue}{1} & \color{cyan}{-1} \\ \color{red}{1} & \color{blue}{-1} & \color{cyan}{1} \end{bmatrix},$$

bowiem

$$\phi((1, 0, 0)) = \color{red}{2} \cdot (1, 0) + \color{red}{1} \cdot (0, 1), \quad \phi((0, 1, 0)) = \color{blue}{1} \cdot (1, 0) - \color{blue}{1} \cdot (0, 1), \quad \phi((0, 0, 1)) = \color{cyan}{-1} \cdot (1, 0) + \color{cyan}{1} \cdot (0, 1).$$

W kolumnach macierzy $M(\phi)_{\text{st}}^{\text{st}}$ stoją wektory zawierające współrzędne obrazów wektorów z bazy standardowej w \mathbb{R}^3 , a te współrzędne są w bazie standardowej \mathbb{R}^2 .

Ogólnie, jeśli $f : K^n \rightarrow K^m$ jest przekształceniem liniowym zadanym wzorem

$$f((x_1, \dots, x_n)) = (a_{11}x_1 + \dots + a_{1n}x_n, a_{21}x_1 + \dots + a_{2n}x_n, \dots, a_{m1}x_1 + \dots + a_{mn}x_n),$$

wówczas

$$M(\phi)_{\text{st}}^{\text{st}} = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix}.$$

* * *

Po co nam takie macierze przekształceń liniowych? Dobrą motywacją może być następujący przykład. Założymy, że mamy bazę

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

przestrzeni \mathbb{R}^3 i rozważmy przekształcenie $\phi : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ o następującej macierzy $M(\phi)_{\mathcal{A}}^{\mathcal{A}}$.

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

Zgodnie z definicją, w pierwszej kolumnie są współrzędne wektora $\phi((1, 0, 1))$ w bazie \mathcal{A} , w drugiej kolumnie są współrzędne wektora $\phi((2, 0, -1))$ w bazie \mathcal{A} , zaś w trzeciej kolumnie są współrzędne wektora $(5, 1, 3)$ w bazie \mathcal{A} , czyli:

$$\begin{aligned} \phi((1, 0, 1)) &= 1 \cdot (1, 0, 1) + 0 \cdot (2, 0, -1) + 0 \cdot (5, 1, 3) \\ \phi((2, 0, -1)) &= 0 \cdot (1, 0, 1) - 1 \cdot (2, 0, -1) + 0 \cdot (5, 1, 3) \\ \phi((5, 1, 3)) &= 0 \cdot (1, 0, 1) + 0 \cdot (2, 0, -1) - 1 \cdot (5, 1, 3) \end{aligned}$$

Czy Czytelnik widzi, że z tej konkretnej postaci macierzy możemy odczytać informację, że ϕ jest w istocie symetrią względem $\text{lin}((1, 0, 1))$ względem $\text{lin}((2, 0, -1), (5, 1, 3))$? Tak właśnie jest! Z drugiej strony zupełnie nie „widać” tego rozważając tylko wzór tego przekształcenia liniowego. Nietrudno go wyznaczyć.

Skoro $\phi((1, 0, 1)) + \phi((2, 0, -1)) = \phi((3, 0, 0)) = (1, 0, 1) - (2, 0, -1) = (-1, 0, 2)$, to

$$\phi((1, 0, 0)) = \left(-\frac{1}{3}, 0, \frac{2}{3} \right).$$

Stąd

$$\phi((0, 0, 1)) = \phi((1, 0, 1)) - \phi((1, 0, 0)) = (1, 0, 1) - \left(-\frac{1}{3}, 0, \frac{2}{3} \right) = \left(\frac{4}{3}, 0, \frac{1}{3} \right).$$

Natomiast

$$\phi((0, 1, 0)) = \phi((5, 1, 3)) - \phi((5, 0, 0)) - \phi((0, 0, 3)) = (-5, -1, -3) - \left(-\frac{5}{3}, 0, \frac{10}{3} \right) - \left(\frac{12}{3}, 0, \frac{3}{3} \right) = \left(-\frac{22}{3}, -1, -\frac{22}{3} \right).$$

W szczególności wzór przekształcenia ϕ to:

$$\phi((x_1, x_2, x_3)) = \left(-\frac{1}{3}x_1 - \frac{22}{3}x_2 + \frac{4}{3}x_3, -x_2, \frac{2}{3}x_1 - \frac{22}{3}x_2 + \frac{1}{3}x_3 \right).$$

Czytelnik zechce sprawdzić, że po wstawieniu do uzyskanego wzoru kolejno wektorów $(1, 0, 1)$, $(2, 0, -1)$ oraz $(5, 1, 3)$ otrzymamy odpowiednio wektory $(1, 0, 1)$, $(-2, 0, 1)$, $(-5, -1, -3)$. Czy znając jedynie wzór przekształcenia ϕ lub macierz $M(\phi)_{\text{st}}^{\text{st}}$ umielibyśmy stwierdzić, że jest to „w istocie” pewna symetria?

Dalsze przykłady.

- Niech $\text{id} : K^n \rightarrow K^n$ będzie identycznością, zaś $\mathcal{A} = (\alpha_1, \dots, \alpha_n)$ niech będzie bazą K^n . Wówczas:

$$M(\text{id})_{st}^{st} = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{bmatrix}.$$

Macierz tę nazywać będziemy MACIERZĄ IDENTYCZNOŚCIOWĄ (rozmiaru n), ozn. I (lub I_n , gdy chcemy podkreślić jej rozmiar). Zauważmy jednak, że w innych niż standardowe bazach, macierze przekształcenia id nie muszą być wcale identycznościowe.

- Niech $\phi_a : K^n \rightarrow K^n$ będzie jednokładnością o skali a . W bazach standardowych macierz ϕ_a to:

$$M(\phi_a)_{st}^{st} = a \cdot I_n.$$

- Niech $V = W \oplus U$ i niech

- $\phi : V \rightarrow V$ będzie **rzutem** na W wzduż U .
- ψ będzie **symetrią** względem W wzduż U

Niech $\mathcal{A} = (\alpha_1, \dots, \alpha_n)$ będzie taką bazą przestrzeni V , że

- $(\alpha_1, \dots, \alpha_k)$ (dla pewnego $1 < k < n$) jest bazą przestrzeni W ,
- $(\alpha_{k+1}, \dots, \alpha_n)$ jest bazą przestrzeni U .

Wówczas macierz $M(\phi)_{\mathcal{A}}^{\mathcal{A}}$ ma w pierwszych **k kolumnach** pierwsze k wektorów bazy standardowej K^n , zaś dalej **kolumny zerowe**, natomiast macierz $M(\psi)_{\mathcal{A}}^{\mathcal{A}}$ ma w pierwszych **k kolumnach** pierwsze k wektorów, zaś dalej **$n - k$ wektorów przeciwnych** do wektorów z bazy standardowej K^n :

$$M(\phi)_{\mathcal{A}}^{\mathcal{A}} = \begin{bmatrix} 1 & 0 & 0 & \cdots & 0 & 0 & \cdots & 0 \\ 0 & 1 & 0 & \cdots & 0 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 1 & 0 & \cdots & 0 \\ 0 & 0 & 0 & \cdots & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 0 & 0 & \cdots & 0 \end{bmatrix}, \quad M(\psi)_{\mathcal{A}}^{\mathcal{A}} = \begin{bmatrix} 1 & 0 & 0 & \cdots & 0 & 0 & \cdots & 0 \\ 0 & 1 & 0 & \cdots & 0 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 1 & 0 & \cdots & 0 \\ 0 & 0 & 0 & \cdots & 0 & -1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 0 & 0 & \cdots & -1 \end{bmatrix},$$

* * *

Wykonywanie przekształcenia liniowego pomiędzy przestrzeniami skończenie wymiarowymi można realizować za pomocą mnożenia macierzy tego przekształcenia przez odpowiedni wektor współrzędnych. Widzieliśmy to na poprzednim wykładzie w przypadku przekształceń liniowych z pomiędzy przestrzeniami współrzędnych. Dokładniej — stwierdziliśmy ostatnio, że dla każdej macierzy $A \in M_{m \times n}(K)$ przekształcenie $\phi : K^n \rightarrow K^m$ przypisujące wektorowi (x_1, \dots, x_n) wektor $\phi((x_1, \dots, x_n)) = (y_1, \dots, y_m)$, przy czym

$$A \cdot \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} y_1 \\ \vdots \\ y_m \end{bmatrix}$$

jest liniowe. W powyższym obliczeniu wykorzystaliśmy współrzędne wektorów w bazie standardowej. Innymi słowy, tę samą równość możemy zapisać w postaci:

$$M(\phi)_{st}^{st} \cdot \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} y_1 \\ \vdots \\ y_m \end{bmatrix}$$

Po wprowadzeniu pojęcia macierzy przekształcenia liniowego wniosek ten sformułować można ogólniej i niekoniecznie tylko w kontekście przestrzeni K^n , ale dowolnej skończenie wymiarowej przestrzeni liniowej nad ciałem K .

Uwaga 19.1.2

Niech $\phi : V \rightarrow W$ będzie przekształceniem liniowym. Niech $\mathcal{A} = (\alpha_1, \dots, \alpha_n)$ będzie bazą przestrzeni liniowej V oraz $(\beta_1, \dots, \beta_m)$ niech będzie bazą przestrzeni W . Jeśli

- a_1, \dots, a_n są współrzędnymi wektora α w bazie \mathcal{A} ,
- b_1, \dots, b_m są współrzędnymi wektora $\phi(\alpha)$ w bazie \mathcal{B} , to

$$M(\phi)_{\mathcal{A}}^{\mathcal{B}} \cdot \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix}.$$

Dowód. Niech $M(\phi)_{\mathcal{A}}^{\mathcal{B}} = [a_{ij}]$, dla $1 \leq i \leq m$ oraz $1 \leq j \leq n$. Wówczas z definicji tej macierzy mamy:

$$\phi(\alpha_j) = a_{1j}\beta_1 + a_{2j}\beta_2 + \dots + a_{mj}\beta_m.$$

Stąd dla dowolnego $\alpha = a_1\alpha_1 + \dots + a_n\alpha_n$ mamy:

$$\begin{aligned} \phi(\alpha) &= \phi(a_1\alpha_1 + \dots + a_n\alpha_n) = a_1\phi(\alpha_1) + \dots + a_n\phi(\alpha_n) \\ &= a_1(a_{11}\beta_1 + a_{21}\beta_2 + \dots + a_{m1}\beta_m) + \dots + a_n(a_{1n}\beta_1 + a_{2n}\beta_2 + \dots + a_{mn}\beta_m) \\ &= (a_{11}a_1 + \dots + a_{1n}a_n)\beta_1 + \dots + (a_{m1}a_1 + \dots + a_{mn}a_n)\beta_m. \end{aligned}$$

Zapisaliśmy więc $\phi(\alpha)$ w bazie \mathcal{B} . Z drugiej strony wiemy, że współrzędne te równe są b_1, \dots, b_m . A zatem oczywiście b_i powstaje przez przemnożenie i -tego wiersza macierzy $M(\phi)_{\mathcal{A}}^{\mathcal{B}}$ przez macierz mającą kolumnę a_1, \dots, a_n , czyli $b_i = a_{i1}a_1 + \dots + a_{in}a_n$. \square

Przykład. Niech V będzie przestrzenią liniową wielomianów stopnia ≤ 3 nad \mathbb{R} , W — przestrzeń liniową $M_{2 \times 2}(\mathbb{R})$, a przekształcenie liniowe $\phi : V \rightarrow W$ niech dane będzie wzorem

$$\phi(a + bx + cx^2 + dx^3) = \begin{bmatrix} a & 2b \\ 3c & 4d \end{bmatrix}.$$

Rozważmy bazy $\mathcal{A} = (1, x, x^2, x^3)$ oraz $\mathcal{B} = \left(\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \right)$. Wówczas

- współrzędne wielomianu $w = a + bx + cx^2 + dx^3$ w bazie \mathcal{A} tworzą wektor $(a, b, c, d) \in \mathbb{R}^4$,

- oczywiście $M(\phi)_{\mathcal{A}}^{\mathcal{B}} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 3 & 0 \\ 0 & 0 & 0 & 4 \end{bmatrix}$ (policz współrzędne $\phi(1), \phi(x), \phi(x^2), \phi(x^3)$ w bazie \mathcal{B}),

- skoro

$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 3 & 0 \\ 0 & 0 & 0 & 4 \end{bmatrix} \cdot \begin{bmatrix} a \\ b \\ c \\ d \end{bmatrix} = \begin{bmatrix} a \\ 2b \\ 3c \\ 4d \end{bmatrix},$$

to macierz $\phi(w)$ ma w bazie \mathcal{B} współrzędne $a, 2b, 3c, 4d$, a zatem $\phi(w) = \begin{bmatrix} a & 2b \\ 3c & 4d \end{bmatrix}$.

* * *

Co mówi powyższy rezultat na poziomie diagramowym? Niech $\Psi_{\mathcal{X}} : X \rightarrow K^r$ będzie izomorfizmem przypisującym wektorowi $\alpha \in X$ wektor (a_1, \dots, a_r) jego współrzędnych w bazie \mathcal{X} . Wykonywanie przekształcenia $\phi : V \rightarrow W$ opisane w Uwadze 19.1.2 wygląda w sposób następujący:

$$\begin{array}{ccc} V & \xrightarrow{\phi} & W \\ \downarrow \Psi_{\mathcal{A}} & & \uparrow \Psi_{\mathcal{B}}^{-1} \\ K^n & \xrightarrow{M(\phi)_{\mathcal{A}}^{\mathcal{B}}} & K^m \end{array}$$

Warto odnotować wniosek dotyczący tzw. MACIERZY ZAMIANY WSPÓŁRZĘDNYCH.

Wniosek 19.1.3

Jeśli \mathcal{A}, \mathcal{B} są bazami przestrzeni liniowej V , zaś $C = M(\text{id})_{\mathcal{A}}^{\mathcal{B}}$, gdzie $\text{id} = \text{id}_V$ jest identycznością na V , to dla każdego $\alpha \in V$: jeśli a_1, \dots, a_n są współrzędnymi α w bazie \mathcal{A} , zaś b_1, \dots, b_n są jego współrzędnymi w bazie \mathcal{B} , to:

$$C \cdot \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix}.$$

Przechodzimy teraz do rezultatu uogólniającego obserwację poczynioną na ostatnim wykładzie w bardzo szczególnym przypadku — składaniu przekształceń liniowych odpowiada mnożenie macierzy.

Twierdzenie 19.1.4: Składanie przekształceń, a mnożenie ich macierzy

Jeśli V, W, Z są przestrzeniami liniowymi nad K z bazami odpowiednio $\mathcal{A}, \mathcal{B}, \mathcal{C}$, oraz $\phi : V \rightarrow W$, $\psi : W \rightarrow Z$ są przekształceniemi liniowymi, to:

$$M(\psi \circ \phi)_{\mathcal{A}}^{\mathcal{C}} = M(\psi)_{\mathcal{B}}^{\mathcal{C}} \cdot M(\phi)_{\mathcal{A}}^{\mathcal{B}}$$

Dowód. Rozważmy następujące bazy odpowiednio przestrzeni V, W, Z :

$$\mathcal{A} = (\alpha_1, \dots, \alpha_n), \quad \mathcal{B} = (\beta_1, \dots, \beta_m), \quad \mathcal{C} = (\gamma_1, \dots, \gamma_k).$$

Niech też dane będą macierze:

$$M(\phi)_{\mathcal{A}}^{\mathcal{B}} = [a_{ij}], \quad M(\psi)_{\mathcal{B}}^{\mathcal{C}} = [b_{ij}], \quad M(\psi \circ \phi)_{\mathcal{A}}^{\mathcal{C}} = [c_{ij}],$$

dla odpowiednich zakresów i, j w każdej z macierzy. Z definicji macierzy przekształceń liniowych $\phi, \psi, \psi \circ \phi$:

$$\begin{aligned} \phi(\alpha_j) &= a_{1j} \cdot \beta_1 + a_{2j} \cdot \beta_2 + \dots + a_{mj} \cdot \beta_m, \\ \psi(\beta_l) &= b_{1l} \cdot \gamma_1 + b_{2l} \cdot \gamma_2 + \dots + b_{kl} \cdot \gamma_k, \\ (\psi \circ \phi)(\alpha_j) &= c_{1j} \cdot \gamma_1 + c_{2j} \cdot \gamma_2 + \dots + c_{kj} \cdot \gamma_k. \end{aligned}$$

Z definicji złożenia oraz liniowości ψ mamy jednak:

$$\begin{aligned} (\psi \circ \phi)(\alpha_j) &= \psi(\phi(\alpha_j)) = \psi(a_{1j} \cdot \beta_1 + a_{2j} \cdot \beta_2 + \dots + a_{mj} \cdot \beta_m) = \\ &= a_{1j} \cdot \psi(\beta_1) + a_{2j} \cdot \psi(\beta_2) + \dots + a_{mj} \cdot \psi(\beta_m). \end{aligned}$$

Rozkładamy każdy z wektorów $\psi(\beta_l)$ w bazie \mathcal{C} :

$$\begin{aligned} (\psi \circ \phi)(\alpha_j) &= \psi(\phi(\alpha_j)) = a_{1j} \cdot (b_{11} \cdot \gamma_1 + b_{21} \cdot \gamma_2 + \dots + b_{k1} \cdot \gamma_k) \\ &\quad + a_{2j} \cdot (b_{12} \cdot \gamma_1 + b_{22} \cdot \gamma_2 + \dots + b_{k2} \cdot \gamma_k) \\ &\quad + \dots \\ &\quad + a_{mj} \cdot (b_{1m} \cdot \gamma_1 + b_{2m} \cdot \gamma_2 + \dots + b_{km} \cdot \gamma_k). \end{aligned}$$

Grupujemy teraz wszystkie wyrazy stojące przy wektorach z bazy \mathcal{C} :

$$\begin{aligned} (\psi \circ \phi)(\alpha_j) &= \psi(\phi(\alpha_j)) = (a_{1j}b_{11} + a_{2j}b_{12} + \dots + a_{mj}b_{1m})\gamma_1 \\ &\quad + (a_{1j}b_{21} + a_{2j}b_{22} + \dots + a_{mj}b_{2m})\gamma_2 \\ &\quad + \dots \\ &\quad + (a_{1j}b_{k1} + a_{2j}b_{k2} + \dots + a_{mj}b_{km})\gamma_k. \end{aligned}$$

A zatem wyraz c_{ij} , stojący przy wektorze γ_i w powyższym przedstawieniu, równy jest

$$a_{1j}b_{i1} + a_{2j}b_{i2} + \dots + a_{mj}b_{im},$$

czyli jest on *iloczynem* i -tego wiersza $M(\psi)_{\mathcal{B}}^{\mathcal{C}}$ i j -tej kolumny $M(\phi)_{\mathcal{A}}^{\mathcal{B}}$. Uzyskaliśmy tezę. \square

Odnajdujmy ważny wniosek, fundamentalny dla naszych rozważań.

Wniosek 19.1.5: Zmiana baz w macierzy przekształcenia

Jeśli $\phi : V \rightarrow W$ jest przekształceniem liniowym, zaś $\mathcal{A}, \mathcal{A}'$ są bazami przestrzeni V oraz jeśli $\mathcal{B}, \mathcal{B}'$ są bazami przestrzeni W , to:

$$M(\phi)_{\mathcal{A}'}^{\mathcal{B}'} = M(\text{id}_W)_{\mathcal{B}}^{\mathcal{B}'} \cdot M(\phi)_{\mathcal{A}}^{\mathcal{B}} \cdot M(\text{id}_V)_{\mathcal{A}'}^{\mathcal{A}}. \quad (*)$$

Dowód. Mamy równość

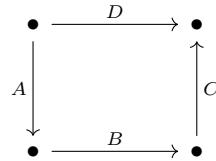
$$\phi = \text{id}_W \circ \phi \circ \text{id}_V,$$

a zatem korzystając z Twierdzenia 19.1.4 mamy

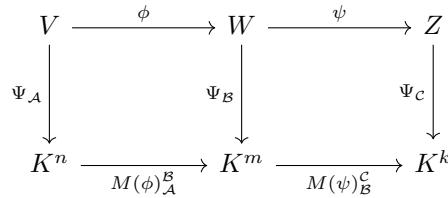
$$M(\phi)_{\mathcal{A}'}^{\mathcal{B}'} = M(\text{id}_W \circ \phi \circ \text{id}_V)_{\mathcal{A}'}^{\mathcal{B}'} = M(\text{id}_W)_{\mathcal{B}}^{\mathcal{B}'} \cdot M(\phi)_{\mathcal{A}}^{\mathcal{B}} \cdot M(\text{id}_V)_{\mathcal{A}'}^{\mathcal{A}}.$$

□

Diagramowe interpretacje powyższych obserwacji mogą nam ułatwić zrozumienie wzoru na zmianę baz. Możemy patrzeć na diagramy przemienne, jako na diagramy przekształceń liniowych, ale możemy za ich pomocą ilustrować też równości iloczynów macierzy (na przykład macierzy tych przekształceń). A zatem na przykład równość macierzy postaci $D = CBA$ zapisać możemy na diagramie w postaci:



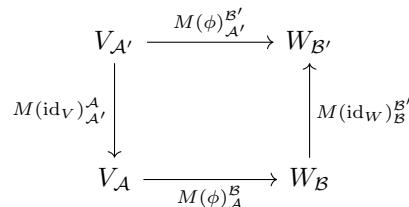
Samo Twierdzenie 19.1.4 zapisać można w postaci diagramu przemennego (przypomnijmy, że izomorfizm $\Psi_{\mathcal{X}}$ przypisuje wektorowi z przestrzeni liniowej X wektor jego współrzędnych w bazie \mathcal{X}):



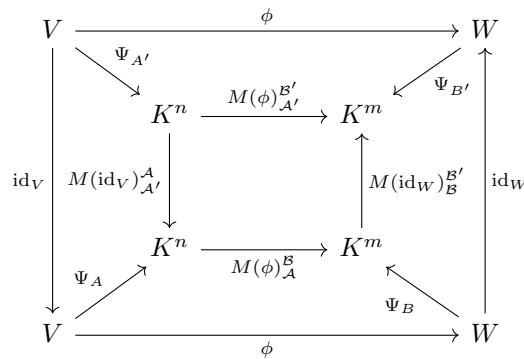
Powyższy diagram będziemy zapisywać krócej, utożsamiając górnego i dolnego wiersz i pisząc:

$$V_{\mathcal{A}} \xrightarrow{M(\phi)_{\mathcal{A}}^{\mathcal{B}}} W_{\mathcal{B}} \xrightarrow{M(\psi)_{\mathcal{B}}^{\mathcal{C}}} Z_{\mathcal{C}}$$

Stosując powyższą skróconą notację łatwiej nam będzie ilustrować wzór (*) na zmianę bazy:



Gdyby nie stosować uproszczenia notacji, diagram miałby nieco bardziej skomplikowaną postać:



Przykład. Niech

$$U = \text{lin}((10, 14, -4)) \subseteq \mathbb{R}^3 \quad \text{oraz} \quad W = \text{lin}((0, 1, 0), (0, 1, 1))$$

tak, że $\mathbb{R}^3 = U \oplus V$. Wyznaczmy wzór na symetrię \mathbb{R}^3 względem U i wzdłuż W . Nazwijmy tę symetrię ϕ . Innymi słowy, szukamy $a_{11}, a_{12}, a_{13}, a_{21}, a_{22}, a_{23}, a_{31}, a_{32}, a_{33} \in \mathbb{R}$ takich, że:

$$\phi((x_1, x_2, x_3)) = (\color{red}{a_{11}}x_1 + \color{blue}{a_{12}}x_2 + \color{teal}{a_{13}}x_3, \color{red}{a_{21}}x_2 + \color{blue}{a_{12}}x_2 + \color{teal}{a_{23}}x_3, \color{red}{a_{31}}x_1 + \color{blue}{a_{32}}x_2 + \color{teal}{a_{33}}x_3).$$

Powyższy warunek jest równoważny temu, że:

$$\phi((1, 0, 0)) = (\color{red}{a_{11}}, \color{red}{a_{21}}, \color{red}{a_{31}}), \quad \phi((0, 1, 0)) = (\color{blue}{a_{12}}, \color{blue}{a_{22}}, \color{blue}{a_{32}}), \quad \phi((0, 0, 1)) = (\color{teal}{a_{13}}, \color{teal}{a_{23}}, \color{teal}{a_{33}}).$$

Innymi słowy, szukamy macierzy:

$$M(\phi)^{st}_{st} = \begin{bmatrix} \color{red}{a_{11}} & \color{blue}{a_{12}} & \color{teal}{a_{13}} \\ \color{red}{a_{21}} & \color{blue}{a_{22}} & \color{teal}{a_{23}} \\ \color{red}{a_{31}} & \color{blue}{a_{32}} & \color{teal}{a_{33}} \end{bmatrix}.$$

Oczywiście zgodnie z definicją symetrii mamy:

$$\phi((10, 14, -4)) = (10, 14, -4) = 1 \cdot (10, 14, -4) + 0 \cdot (0, 1, 0) + 0 \cdot (0, 1, 1)$$

$$\phi((0, 1, 0)) = -(0, 1, 0) = 0 \cdot (10, 14, -4) + (-1) \cdot (0, 1, 0) + 0 \cdot (0, 1, 1)$$

$$\phi((0, 1, 1)) = -(0, 1, 1) = 0 \cdot (10, 14, -4) + 0 \cdot (0, 1, 0) + (-1) \cdot (0, 1, 1)$$

Nietrudno z tych warunków wywnioskować, że

$$\phi((0, 0, 1)) = \phi((0, 1, 1)) - \phi((0, 1, 0)), \quad \phi(1, 0, 0) = \frac{1}{10}\phi((10, 14, -4)) - \frac{7}{5}\phi((0, 1, 0)) + \frac{2}{5}\phi((0, 0, 1)),$$

ale naszym celem jest zaprezentowanie metody wykorzystującej formułę (*).

Weźmy bazę $\mathcal{A} = ((10, 14, -4), (0, 0, 1), (0, 1, 1))$. Z warunków zapisanych wyżej wynika, że:

$$M(\phi)^{\mathcal{A}}_{\mathcal{A}} = \begin{bmatrix} \color{red}{1} & \color{blue}{0} & \color{teal}{0} \\ \color{red}{0} & \color{blue}{-1} & \color{teal}{0} \\ \color{red}{0} & \color{blue}{0} & \color{teal}{-1} \end{bmatrix}.$$

Możemy teraz skorzystać z formuły $M(\phi)^{st}_{st} = M(\text{id})^{st}_{\mathcal{A}} \cdot M(\phi)^{\mathcal{A}}_{\mathcal{A}} \cdot M(\text{id})^{st}_{\mathcal{A}}$.

Wyznaczmy macierze $M(\text{id})^{st}_{\mathcal{A}}$ oraz $M(\text{id})^{st}_{\mathcal{A}}$. Dla każdego $v \in \mathbb{R}^3$ mamy $\text{id}(v) = v$, czyli

$$\text{id}((10, 14, -4)) = (10, 14, -4) = 10 \cdot (1, 0, 0) + 14 \cdot (0, 1, 0) - 4 \cdot (0, 0, 1)$$

$$\text{id}((0, 1, 0)) = (0, 1, 0) = 0 \cdot (1, 0, 0) + 1 \cdot (0, 1, 0) + 0 \cdot (0, 0, 1)$$

$$\text{id}((0, 1, 1)) = (0, 1, 1) = 0 \cdot (1, 0, 0) + 1 \cdot (0, 1, 0) + 1 \cdot (0, 0, 1).$$

Niedużo trudniej jest wyznaczyć współrzędne wektorów bazy standardowej w bazie \mathcal{A} :

$$\text{id}((1, 0, 0)) = (1, 0, 0) = \frac{1}{10} \cdot (10, 14, -4) - \frac{9}{5} \cdot (0, 1, 0) + \frac{2}{5} \cdot (0, 1, 1)$$

$$\text{id}((0, 1, 0)) = (0, 1, 0) = 0 \cdot (10, 14, -4) + 1 \cdot (0, 1, 0) + 0 \cdot (0, 1, 1)$$

$$\text{id}((0, 0, 1)) = (0, 0, 1) = 0 \cdot (10, 14, -4) + (-1) \cdot (0, 1, 0) + 1 \cdot (0, 1, 1).$$

$$\text{W szczególności } M(\text{id})^{st}_{\mathcal{A}} = \begin{bmatrix} 10 & 0 & 0 \\ 14 & 1 & 1 \\ -4 & 0 & 1 \end{bmatrix}, \quad M(\text{id})^{st}_{\mathcal{A}} = \begin{bmatrix} \frac{1}{10} & 0 & 0 \\ -\frac{9}{5} & 1 & -1 \\ \frac{2}{5} & 0 & 1 \end{bmatrix}.$$

W rezultacie

$$M(\phi)^{st}_{st} = M(\text{id})^{st}_{\mathcal{A}} \cdot M(\phi)^{\mathcal{A}}_{\mathcal{A}} \cdot M(\text{id})^{st}_{\mathcal{A}} = \begin{bmatrix} 10 & 0 & 0 \\ 14 & 1 & 1 \\ -4 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{bmatrix} \cdot \begin{bmatrix} \frac{1}{10} & 0 & 0 \\ -\frac{9}{5} & 1 & -1 \\ \frac{2}{5} & 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ \frac{14}{5} & -1 & 0 \\ -\frac{4}{5} & 0 & -1 \end{bmatrix}.$$

Otrzymaliśmy więc szukany wzór $\phi((x_1, x_2, x_3)) = (x_1, \frac{14}{5}x_1 - x_2, -\frac{4}{5}x_1 - x_3)$.

Następnym razem przyjrzymy się macierzom izomorfizmów w kontekście istnienia przekształcenia odwrotnego oraz mnożenia macierzy. Omówimy przy tym dokładniej ważny obiekt wprowadzony ostatnio – macierze odwracalne, wiążąc je z rozważanymi dziś krótko macierzami zmiany bazy.

Na koniec zostawiamy dowód znanego nam już twierdzenia, które wypowiadamy teraz za pomocą macierzy przekształcenia liniowego. Pozwala on nie tylko na utożsamianie skończenie wymiarowych przestrzeni liniowych i przestrzeni ich współrzędnych, ale także przekształceń liniowych i ich macierzy.

Wniosek 19.1.6

Niech V, W będą przestrzeniami liniowymi wymiaru odpowiednio n i m nad ciałem K . Dla dowolnej bazy \mathcal{A} przestrzeni V oraz \mathcal{B} przestrzeni W możemy sformułować izomorfizm przestrzeni liniowych $\Psi_{\mathcal{A}, \mathcal{B}} : L(V, W) \longrightarrow M_{m \times n}(K)$ dany wzorem

$$\Psi_{\mathcal{A}, \mathcal{B}}(\phi) = M(\phi)_{\mathcal{A}}^{\mathcal{B}}.$$

Dowód. Zauważmy najpierw, że $\Psi_{\mathcal{A}, \mathcal{B}}$ jest przekształceniem liniowym. Weźmy dwa przekształcenia liniowe $\phi, \psi : V \rightarrow W$. Niech $\mathcal{A} = (\alpha_1, \dots, \alpha_n)$ oraz $\mathcal{B} = (\beta_1, \dots, \beta_m)$. Założymy, że istnieją a_{1j}, \dots, a_{mj} oraz b_{1j}, \dots, b_{mj} takie, że:

$$\phi(\alpha_j) = a_{1j}\beta_1 + \dots + a_{mj}\beta_m, \quad \psi(\alpha_j) = b_{1j}\beta_1 + \dots + b_{mj}\beta_m.$$

Wówczas

$$(\phi + \psi)(\alpha_j) = \phi(\alpha_j) + \psi(\alpha_j) = a_{1j}\beta_1 + \dots + a_{mj}\beta_m + b_{1j}\beta_1 + \dots + b_{mj}\beta_m = (a_{1j} + b_{1j})\beta_1 + \dots + (a_{mj} + b_{mj})\beta_m.$$

A zatem współrzędna wektora $(\phi + \psi)(\alpha_j)$ w bazie \mathcal{B} przy wektorze β_i to $a_{ij} + b_{ij}$. Krótko mówiąc

$$\Psi_{\mathcal{A}, \mathcal{B}}(\phi + \psi) = M(\phi + \psi)_{\mathcal{A}}^{\mathcal{B}} = M(\phi)_{\mathcal{A}}^{\mathcal{B}} + M(\psi)_{\mathcal{A}}^{\mathcal{B}} = \Psi_{\mathcal{A}, \mathcal{B}}(\phi) + \Psi_{\mathcal{A}, \mathcal{B}}(\psi).$$

Analogicznie dowodzimy, że $M(\lambda\phi)_{\mathcal{A}}^{\mathcal{B}} = \lambda \cdot M(\phi)_{\mathcal{A}}^{\mathcal{B}}$, dla $\lambda \in K$. A zatem $\Psi_{\mathcal{A}, \mathcal{B}}$ jest przekształceniem liniowym.

Pozostaje pokazać, że $\Psi_{\mathcal{A}, \mathcal{B}}$ jest izomorfizmem. Oczywiście $\Psi_{\mathcal{A}, \mathcal{B}}$ jest różnicowartościowe, bo każde przekształcenie liniowe jest jednoznacznie określone na bazie, a kolumny $M(\phi)_{\mathcal{A}}^{\mathcal{B}}$ to obrazy wektorów z bazy \mathcal{A} .

Odwzorowanie $\Psi_{\mathcal{A}, \mathcal{B}}$ to również suriekcja: dla dowolnej macierzy $X = [x_{ij}] \in M_{m \times n}(K)$ można określić przekształcenie liniowe $\phi \in L(V, W)$, zadane na bazie \mathcal{A} (w sposób jednoznaczny) warunkiem

$$\alpha_j \mapsto x_{1j}\beta_1 + \dots + x_{mj}\beta_m.$$

Stąd $\Psi_{\mathcal{A}, \mathcal{B}}$ jest liniową bijekcją, czyli izomorfizmem przestrzeni liniowych. □

* * *

Czytelnik może teraz wrócić do wyjściowego diagramu wiążącego przekształcenie liniowe z jego macierzą i dodać do niego strzałkę utożsamiającą przekształcenie i jego macierz:

$$\begin{array}{ccc} V & \xrightarrow{\phi} & W \\ \downarrow \Psi_{\mathcal{A}} & \downarrow \Psi_{\mathcal{A}, \mathcal{B}}(\phi) & \uparrow \Psi_{\mathcal{B}}^{-1} \\ K^n & \xrightarrow{M(\phi)_{\mathcal{A}}^{\mathcal{B}}} & K^m \end{array}$$

Jeżeli kogoś z Państwa przy pierwszym czytaniu nieco zniechęcają powracające diagramy, warto podkreślić, że mają one na tym etapie wartość ilustracyjną. Za prostym w zasadzie wzorem na zmianę bazy, sprowadzającym się do przedstawienia macierzy za pomocą iloczynu trzech innych, stoi w istocie głęboki zamysł, będący u źródeł utożsamienia wszystkich przestrzeni ustalonego (skończonego) wymiaru, który pozwala przełożyć składanie przekształceń liniowych na odpowiednie mnożenie macierzy. Odsyłamy Czytelnika do zadań, których staranne przerobienie dać może niezbędne poczucie pewności w operowaniu obiektytami, które wracać będą w naszych rozważaniach przez cały kurs.

19.2 Wybór przykładowych pytań

Proszę się nie uczyć na pamięć odpowiedzi czy wyjaśnienie! Takich pytań można ułożyć znacznie więcej i służą one jedynie sprawdzeniu podstawowego zrozumienia treści zajęć. Jeśli odpowiedź na dowolne z poniższych pytań sprawia Państwu problem, proszę pomyśleć który fragment notatek dotyczy tego pytania, wrócić do ich lektury, a w razie dalszych trudności poprosić prowadzącego o wskazówkę.

- Przekształcenie liniowe $\phi : \mathbb{R}^2 \rightarrow \mathbb{R}^3$ dane jest wzorem

$$\phi((x_1, x_2)) = (x_1 + x_2, -x_1, x_2).$$

Jaki jest rozmiar macierzy tego przekształcenia liniowego? Wyznacz macierze tego przekształcenia w bazach \mathcal{A} przestrzeni \mathbb{R}^2 oraz \mathcal{B} przestrzeni \mathbb{R}^3 , gdzie:

- $\mathcal{A} = ((1, 0), (0, 1)), \mathcal{B} = ((1, 0, 0), (0, 1, 0), (0, 0, 1)),$
- $\mathcal{A} = ((0, 1), (1, 0)), \mathcal{B} = ((1, 0, 0), (0, 1, 0), (0, 0, 1)),$
- $\mathcal{A} = ((0, 1), (1, 0)), \mathcal{B} = ((0, 0, 1), (1, 0, 0), (0, 1, 0)).$

- Zapisz macierz w bazach standardowych rzutu ϕ przestrzeni \mathbb{R}^2 na podprzestrzeń $\text{lin}((1, 1))$ wzdłuż podprzestrzeni $\text{lin}((1, -1))$ oraz rzutu ψ przestrzeni \mathbb{R}^2 na podprzestrzeń $\text{lin}((1, -1))$ wzdłuż podprzestrzeni $\text{lin}((1, 1))$.
- Niech $\mathcal{A} = (\alpha, \beta, \gamma)$ będzie bazą przestrzeni liniowej V . Przekształcenie liniowe $\phi : V \rightarrow V$ spełnia

$$\phi(\alpha) = 2\alpha, \quad \phi(\beta) = -\beta, \quad \phi(\gamma) = 0.$$

Znajdź macierz $A = M(\phi)_{\mathcal{A}}^{\mathcal{A}}$ oraz macierze $A \cdot A$. Czy istnieje macierz A^{-1} ?

- Niech $\mathcal{A} = (\alpha, \beta)$ będzie bazą przestrzeni liniowej V . Przekształcenie liniowe $\phi : V \rightarrow V$ spełnia

$$\phi(\alpha) = \beta, \quad \phi(\beta) = \alpha.$$

Znajdź macierz $A = M(\phi)_{\mathcal{A}}^{\mathcal{A}}$ oraz macierze $A \cdot A$ i A^{-1} .

- Niech $\mathcal{A} = (\alpha, \beta)$ będzie bazą przestrzeni liniowej V . Przekształcenie liniowe $\phi : V \rightarrow V$ spełnia

$$\phi(\alpha) = \alpha, \quad \phi(\beta) = \alpha + \beta.$$

Znajdź macierz $M(\phi)_{\mathcal{A}}^{\mathcal{A}}$ oraz macierze $M(\phi \circ \phi \circ \phi)_{\mathcal{A}}^{\mathcal{A}}$ i $M(\phi^{-1})_{\mathcal{A}}^{\mathcal{A}}$.

- Niech $\mathcal{A} = (\alpha, \beta, \gamma)$ będzie bazą przestrzeni liniowej V . Przekształcenie liniowe $\phi : V \rightarrow V$ spełnia

$$\phi(\alpha) = \beta, \quad \phi(\beta) = \gamma, \quad \phi(\gamma) = 0.$$

Znajdź macierz $M(\phi)_{\mathcal{A}}^{\mathcal{A}}$ oraz macierze $M(\phi \circ \phi)_{\mathcal{A}}^{\mathcal{A}}$ i $M(\phi \circ \phi \circ \phi)_{\mathcal{A}}^{\mathcal{A}}$.

- Niech $\phi : V \rightarrow W$ będzie izomorfizmem przestrzeni liniowych, niech $\mathcal{A} = (\alpha_1, \dots, \alpha_n)$ będzie bazą przestrzeni W . Czy $\mathcal{B} = (\phi(\alpha_1), \dots, \phi(\alpha_n))$ jest bazą W ? Jeśli tak, to znajdź $M(\phi)_{\mathcal{A}}^{\mathcal{B}}$ oraz macierz $M(\phi^{-1})_{\mathcal{B}}^{\mathcal{A}}$.

- Niech \mathcal{A}, \mathcal{B} będą bazami przestrzeni liniowej V wymiaru n . Pokaż, że $M(\text{id})_{\mathcal{B}}^{\mathcal{A}} \circ M(\text{id})_{\mathcal{A}}^{\mathcal{B}} = I_n$.

- Niezerowe macierze A, B są macierzami pewnych przekształceń liniowych $\phi, \psi : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ w bazach standardowych. Wiadomo, że $A \cdot B = 0$. Jakie są wymiary $\ker \phi, \ker \psi, \text{im } \phi, \text{im } \psi$? Jakie są rzędy macierzy A i B ?

- Dane są przekształcenia liniowe $\phi : V \rightarrow W$ i $\psi : W \rightarrow U$ oraz bazy $\mathcal{A}, \mathcal{B}, \mathcal{C}$ przestrzeni V, W, U odpowiednio. Wiadomo, że dla $\alpha_1, \alpha_2 \in \mathcal{A}$ zachodzi

$$\phi(\alpha_1) = \phi(\alpha_2).$$

Czy pierwsza i druga kolumna macierzy $M(\phi)_{\mathcal{A}}^{\mathcal{B}}$ są takie same? Czy pierwsza i druga kolumna macierzy będącej iloczynem $M(\psi)_{\mathcal{B}}^{\mathcal{C}} \cdot M(\phi)_{\mathcal{A}}^{\mathcal{B}}$ są takie same?

- Przekształcenie liniowe $\phi : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ jest symetrią względem pewnej niezerowej podprzestrzeni oraz $\phi \neq \text{id}$. Niech A będzie macierzą przekształcenia ϕ w bazach standardowych. Udowodnij, że układ równań

$$A \cdot \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = - \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

ma niezerowe rozwiązanie.

19.3 Zadania do samodzielnej pracy

Zadania ilustrujące umiejętności niezbędne do zaliczenia przedmiotu oznaczone są symbolem ♠.

Przy tych zadaniach dodany jest opis umiejętności, które sprawdzają.

1. (♠ Wyznaczanie macierzy przekształcenia w bazie)

Znajdź macierz przekształcenia liniowego ϕ w bazach \mathcal{A}, \mathcal{B} :

- $\phi: \mathbb{R}^2 \rightarrow \mathbb{R}^4$, $\phi((x_1, x_2)) = (3x_1 + x_2, x_1 + 5x_2, -x_1 + 4x_2, 2x_1 + x_2)$,
 $\mathcal{A} = \{(3, 1), (4, 2)\}$, $\mathcal{B} = \{(1, 0, 1, 0), (0, 1, 1, 1), (0, 1, 2, 3), (0, 0, 0, 1)\}$,
- $\phi: \mathbb{R}^3 \rightarrow \mathbb{R}^3$, $\phi((x_1, x_2, x_3)) = (4x_1 + x_2 + x_3, 3x_1 + 2x_2 + x_3, 3x_1 + 2x_2 + x_3)$,
 $\mathcal{A} = \{(3, 1, 1), (1, 0, 0), (5, 1, 0)\}$, $\mathcal{B} = \{(3, 4, 5), (4, 1, 1), (2, 0, 1)\}$.
- $\phi: \mathbb{R}^4 \rightarrow \mathbb{R}^2$, $\phi((x_1, x_2, x_3, x_4)) = (5x_1 - 2x_2 + 3x_3 - x_4, 3x_1 + 4x_2 + 6x_4)$,
 $\mathcal{A} = \{(2, 1, 0, 1), (1, 0, 3, 1), (2, 1, 1, 3), (3, 1, 2, 1)\}$, $\mathcal{B} = \{(5, 2), (3, 1)\}$.

2. (♠ Wyznaczanie wzoru przekształcenia liniowego mając macierz w bazach)

Dane są bazy \mathcal{A}, \mathcal{B} przestrzeni \mathbb{R}^3 : $\mathcal{A} = \{(3, 1, 1), (1, 0, 0), (5, 1, 0)\}$, $\mathcal{B} = \{(3, 4, 5), (4, 1, 1), (2, 0, 1)\}$.

Znajdź wzór przekształcenia $\psi: \mathbb{R}^3 \rightarrow \mathbb{R}^3$ wiedząc, że: $M(\psi)_{\mathcal{A}}^{\mathcal{B}} = \begin{bmatrix} 1 & 1 & 4 \\ 2 & 1 & 3 \\ 0 & 1 & 1 \end{bmatrix}$.

3. Dla $\phi \in L(\mathbb{R}^3, \mathbb{R}^2)$ danego wzorem $\phi((x_1, x_2, x_3)) = (x_1 - x_2 + 2x_3, 3x_1 + x_2 + x_3)$ znajdź takie bazy

\mathcal{A} przestrzeni \mathbb{R}^3 oraz \mathcal{B} przestrzeni \mathbb{R}^2 , że $M(\phi)_{\mathcal{A}}^{\mathcal{B}} = \begin{bmatrix} 1 & 1 & 1 \\ 2 & 0 & 1 \end{bmatrix}$.

4. (♠ Baza jądra i obrazu przekształcenia liniowego danego macierzą w bazach)

Niech $\mathcal{A} = ((1, 1, 0), (1, 0, 1), (0, 1, 0))$ oraz $A = \begin{bmatrix} 1 & 2 & 3 \\ 1 & -1 & 1 \end{bmatrix}$. Znajdź bazy jądra i obrazu przekształceń ϕ, ψ , gdzie $M(\phi)_{st}^{\mathcal{A}} = A^T$ oraz $M(\psi)_{st}^{\mathcal{A}} = A$.

5. (♠ Wyznaczanie macierzy złożenia w bazach) Dane są następujące bazy przestrzeni $\mathbb{R}^3, \mathbb{R}^4, \mathbb{R}^2$:

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

Przy tym dane jest również przekształcenie $\phi: \mathbb{R}^3 \rightarrow \mathbb{R}^4$ o wzorze

$$\phi((x_1, x_2, x_3)) = (3x_1 - x_2 - 2x_3, 3x_1 + 4x_2 + x_3, 5x_1 + 2x_3, x_1 + x_2 + x_3)$$

oraz przekształcenie ψ takie, że $M(\psi)_{\mathcal{B}}^{\mathcal{C}} = \begin{bmatrix} 0 & 2 & 1 & 1 \\ 2 & 5 & 0 & 3 \end{bmatrix}$. Znajdź $M(\psi \circ \phi)_{\mathcal{A}}^{\mathcal{C}}$ oraz $M(\psi \circ \phi)_{st}^{\mathcal{A}}$

6. (♠ Zmianianie współrzędnych wektorów przy zmianie bazy)

Niech $\phi: V \rightarrow W$ oraz $\psi: W \rightarrow Z$ będą przekształceniami liniowymi i niech

$$M(\phi)_{\mathcal{A}}^{\mathcal{B}} = \begin{bmatrix} 2 & 1 & 4 & 5 \\ 1 & 0 & 4 & 3 \end{bmatrix}, \quad M(\psi)_{\mathcal{B}}^{\mathcal{C}} = \begin{bmatrix} 3 & 1 \\ 2 & 5 \\ 0 & 1 \end{bmatrix}$$

w pewnych bazach $\mathcal{A}, \mathcal{B}, \mathcal{C}$ przestrzeni V, W, Z . Niech $\alpha \in V$ ma w bazie \mathcal{A} współrzędne $1, -1, 3, -2$.

Znajdź współrzędne wektora $\phi(\alpha)$ w bazie \mathcal{B} oraz współrzędne wektora $(\psi \circ \phi)(\alpha)$ w bazie \mathcal{C} .

7. Wykaż, że dla dowolnych $A \in M_{m \times n}(K)$ oraz $B, C \in M_{n \times s}(K)$ mamy $A(B + C) = AB + AC$.

8. Niech $A, B \in M_{n \times n}(K)$. Udowodnij, że jeśli $AB = 0$, to $r(A) + r(B) \leq n$.

9. Niech $f: M_{2 \times 2}(\mathbb{C}) \rightarrow M_{2 \times 2}(\mathbb{C})$ będzie przekształceniem danym wzorem

$$f(X) = AX - XA, \quad \text{gdzie } A = \begin{bmatrix} 1 & i \\ 1 & 1 \end{bmatrix} \in M_2(\mathbb{C}).$$

Wykaż, że f jest przekształceniem liniowym. Wyznacz $M_{\mathcal{B}}^{\mathcal{B}}(f)$, gdzie $\mathcal{B} = \{E_{11}, E_{12}, E_{21}, E_{22}\}$, jest bazą złożoną z jedynek macierzowych. Znajdź bazy i wymiary przestrzeni $\ker f$ oraz $\text{im } f$.

10. Dane są przestrzenie liniowe V, W nad ciałem K , przy czym $\dim V = n$, $\dim W = 2$. Dane jest również przekształcenie liniowe $\phi \in L(V, V)$. Definiujemy funkcję $\Phi: L(V, W) \rightarrow L(V, W)$ wzorem $\Phi(\psi) = \psi \circ \phi$. Wykaż, że jest to przekształcenie liniowe. Niech \mathcal{A}, \mathcal{B} — bazy V, W . Wyznacz macierz tego przekształcenia liniowego w bazie ϕ_{ij} przestrzeni $L(V, W)$ określonej w dowodzie Twierdzenia 17.1.3. Jaki jest związek tej macierzy z macierzą $M(\phi)_{\mathcal{A}}^{\mathcal{B}}$?

Rozdział 20

Macierze odwracalne i izomorfizmy Macierze operacji elementarnych

20.1 Wykład 20

Od kilku wykładów mówimy o przekształceniach liniowych, czyli funkcjach pomiędzy przestrzeniami liniowymi (nad ustalonym ciałem) zachowujących kombinacje liniowe. W tym rozdziale powiążemy pojęcie izomorfizmu przestrzeni skończonego wymiaru z pojęciem macierzy odwracalnej. Przypomnijmy definicję.

Definicja 20.1.1: Macierz odwrotna

Powiemy, że macierz $B \in M_{n \times n}(K)$ jest ODWROTNĄ do macierzy $A \in M_{n \times n}(K)$, jeśli

$$AB = BA = I_n.$$

Macierz odwrotną do macierzy A oznaczamy, o ile istnieje, jako A^{-1} . Macierz, która ma odwrotną nazywamy MACIERZĄ ODWRACALNĄ.

Zauważmy, że jeśli dla macierzy A istnieją macierze B, C takie, że $AB = BA = I_n$ oraz $AC = CA = I_n$, to $AB = AC$. Mnożąc tę równość z obydwu stron przez B dostajemy $B(AB) = B(AC)$. Skoro mnożenie macierzy jest łączne, to mamy też $(BA)B = (BA)C$, czyli $I_nB = I_nC$, a stąd $B = C$. Macierz odwrotna do A jest więc wyznaczona jednoznacznie, o ile istnieje. Stąd oznaczenie A^{-1} .

Obserwacja 20.1.2: Odwrotność iloczynu

Jeśli $A, B \in M_{n \times n}(K)$ są macierzami odwracalnymi, to macierz AB jest odwracalna i

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

Dowód. Mnożenie macierzy jest łączne, więc

$$(AB) \cdot (B^{-1}A^{-1}) = A(BB^{-1})A^{-1} = AI_nA^{-1} = AA^{-1} = I_n.$$

□

Definicja 20.1.3: Macierz transponowana

MACIERZĄ TRANSPONOWANĄ do macierzy $A = [a_{ij}] \in M_{m \times n}(K)$ nazywać będziemy taką macierz $A^T = [b_{ij}] \in M_{n \times m}(K)$, której kolejne kolumny są kolejnymi wierszami macierzy A , czyli $b_{ij} = a_{ji}$, dla każdych i, j . Jeśli macierz A spełnia warunek $A = A^T$, to nazywamy ją macierzą SYMETRYCZNĄ.

Przykłady:

$$\begin{bmatrix} 1 & 2 & 3 & 4 \\ 5 & 6 & 7 & 8 \end{bmatrix}^T = \begin{bmatrix} 1 & 5 \\ 2 & 6 \\ 3 & 7 \\ 4 & 8 \end{bmatrix}, \quad \begin{bmatrix} 1 & 2 & 0 \\ 2 & 1 & 0 \\ 0 & 0 & 7 \end{bmatrix}^T = \begin{bmatrix} 1 & 2 & 0 \\ 2 & 1 & 0 \\ 0 & 0 & 7 \end{bmatrix}, \quad \begin{bmatrix} 1 & 2 & 0 \\ 2 & 1 & 1 \\ 0 & 1 & 4 \end{bmatrix}^T = \begin{bmatrix} 1 & 2 & 0 \\ 2 & 1 & 1 \\ 0 & 1 & 4 \end{bmatrix}.$$

Uwaga 20.1.4

Dla macierzy $A \in M_{m \times n}(K)$ mamy $(A^T)^T = A$, $r(A) = r(A^T)$ oraz $(AB)^T = B^T A^T$. Jeśli zaś macierz $A \in M_{n \times n}(K)$ jest odwracalna, to $(A^T)^{-1} = (A^{-1})^T$.

Dowód. Pierwsza i druga własność wynika bezpośrednio z definicji transponowania i rzędu. Aby uzasadnić trzecią, można zauważyc, że wyraz ij -ty macierzy $(AB)^T$ to wyraz ji -ty macierzy AB , powstający przez pomnożenie j -tego wiersza macierzy A przez i -tą kolumnę macierzy B . Z drugiej strony wyraz ij -ty macierzy $B^T A^T$ powstaje przez pomnożenie i -tego wiersza macierzy B^T przez j -tą kolumnę macierzy A , czyli przez pomnożenie i -tej kolumny macierzy B przez j -ty wiersz A . Wyniki tych mnożeń są identyczne.

Teza ostatniego punktu wynika stąd, że na mocy punktu trzeciego $A^T \cdot (A^{-1})^T = (A^{-1}A)^T = I^T = I$. \square

Przejdziemy teraz do powiązania pojęcia macierzy odwracalnej z izomorfizmami przestrzeni liniowych. Będziemy przy tym korzystać wielokrotnie z wyników z poprzedniego wykładu o postaci macierzy złożenia oraz z twierdzenia o zmianie bazy.

Zaczniemy od podstawowego rezultatu mówiącego, że macierz przekształcenia liniowego jest odwracalna wtedy i tylko wtedy, gdy jest macierzą izomorfizmu.

Twierdzenie 20.1.5

Niech $\phi : K^n \rightarrow K^n$ będzie przekształceniem liniowym. Następujące warunki są równoważne:

- (i) ϕ jest izomorfizmem,
- (ii) macierz $M(\phi)_{st}^{st}$ jest odwracalna,
- (iii) dla dowolnych baz \mathcal{A}, \mathcal{B} przestrzeni K^n macierz $M(\phi)_{\mathcal{A}}^{\mathcal{B}}$ jest odwracalna.

Dowód. Jeśli ϕ jest izomorfizmem oraz $\psi = \phi^{-1}$, to biorąc $M(\psi)_{st}^{st}$ mamy:

$$M(\phi)_{st}^{st} \cdot M(\psi)_{st}^{st} = M(\phi \circ \psi)_{st}^{st} = M(\phi \circ \phi^{-1})_{st}^{st} = M(\text{id})_{st}^{st} = I_n.$$

Analogicznie $M(\psi)_{st}^{st} \cdot M(\phi)_{st}^{st} = I_n$, co daje (i) \Rightarrow (ii). Jeśli $A = M(\phi)_{st}^{st}$ jest odwracalna i $AB = I_n$, to niech $\psi : K^n \rightarrow K^n$ będzie zadane warunkiem $M(\psi)_{st}^{st} = B$. Wówczas

$$M(\phi \circ \psi)_{st}^{st} = M(\phi)_{st}^{st} \cdot M(\psi)_{st}^{st} = A \cdot B = I_n = M(\text{id})_{st}^{st}.$$

Zatem $\phi \circ \psi = \text{id}$. Analogicznie z $BA = I_n$ mamy $\psi \circ \phi = \text{id}$. Zatem ϕ to izomorfizm i mamy (ii) \Rightarrow (i). Jest jasne, że (iii) \Rightarrow (ii) jest. Implikacja odwrotna wynika z rozkładu $M(\phi)_{\mathcal{A}}^{\mathcal{B}} = M(\text{id})_{st}^{\mathcal{B}} \cdot M(\phi)_{st}^{st} \cdot M(\text{id})_{st}^{\mathcal{A}}$. Macierz $M(\phi)_{\mathcal{A}}^{\mathcal{B}}$ jest więc odwracalna, jako iloczyn macierzy odwracalnych. Stąd (ii) \Rightarrow (iii). \square

Wniosek 20.1.6

Kolumny macierzy odwracalnej $A \in M_{n \times n}(K)$ tworzą bazę przestrzeni liniowej K^n .

Dowód. Rozważmy takie przekształcenie liniowe $\phi : K^n \rightarrow K^n$, że $A = M(\phi)_{st}^{st}$. Skoro macierz A jest odwracalna, to ϕ jest izomorfizmem, zgodnie z poprzednim rezultatem. Kolumny macierzy A to wektory $\phi(\epsilon_1), \dots, \phi(\epsilon_n)$, gdzie $\text{st} = \{\epsilon_1, \dots, \epsilon_n\}$. Izomorfizm przeprowadza jednak bazę na bazę, co kończy dowód. \square

Uwaga 20.1.7

Niech $\phi : K^n \rightarrow K^n$ będzie izomorfizmem. Wówczas każda macierz odwracalna w $M_{n \times n}(K)$ jest macierzą ϕ w pewnych bazach.

Dowód. Niech $A \in M_n(K)$ będzie macierzą odwracalną. Wykażemy, że jest to macierz izomorfizmu ϕ w pewnych bazach. Niech \mathcal{B} będzie taką bazą K^n , że $M(\phi)_{\text{st}}^{\mathcal{B}} = I_n$ (baza \mathcal{B} składa się z obrazów wektorów bazy standardowej przy ϕ). Wystarczy zatem pokazać, że istnieje taka baza \mathcal{A} przestrzeni K^n , że $G = M(\text{id})_{\mathcal{A}}^{\text{st}}$, bo wówczas

$$M(\phi)_{\mathcal{A}}^{\mathcal{B}} = M(\phi)_{\text{st}}^{\mathcal{B}} \cdot M(\text{id})_{\mathcal{A}}^{\text{st}} = I_n A = A.$$

Szukaną bazę \mathcal{A} można odczytać z kolumn macierzy odwracalnej A , gdzie $A = M(\text{id})_{\mathcal{A}}^{\text{st}}$. \square

A zatem dowolna macierz odwracalna rozmiaru n nad ciałem K jest macierzą dowolnego izomorfizmu przestrzeni wymiaru n nad ciałem K .

Definicja 20.1.8

Niech \mathcal{A}, \mathcal{B} będą bazami przestrzeni V . Macierz $M(\text{id}_V)_{\mathcal{A}}^{\mathcal{B}}$ nazywamy MACIERZĄ ZAMIANY (TRANSFORMACJI) WSPÓŁRZĘDNYCH z \mathcal{A} do \mathcal{B} .

Przykład. Macierz zmiany współrzędnych bazy przestrzeni \mathbb{R}^2 z bazy $\mathcal{A} = ((1, 1), (2, 1))$ do bazy $\mathcal{B} = ((2, 1), (-1, -1))$ jest macierz

$$M(\text{id})_{\mathcal{A}}^{\mathcal{B}} = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}.$$

Wniosek 20.1.9

Niech $A \in M_{n \times n}(K)$. Następujące warunki są równoważne:

- (i) A jest macierzą zamiany współrzędnych w K^n ,
- (ii) A jest macierzą odwracalną,
- (iii) przekształcenie liniowe $\phi : K^n \rightarrow K^n$ zadane warunkiem $M(\phi)_{\text{st}}^{\text{st}} = A$ jest izomorfizmem,
- (iv) $r(A) = n$.

Dowód. Równoważność warunków (ii) oraz (iii) pokazaliśmy już w Twierdzeniu 20.1.5. Warunek (iii) jest oczywiście równoważny (iv), gdyż macierz izomorfizmu jest odwracalna, a więc jej kolumny tworzą bazę przestrzeni K^n . Macierz ta więc ma rzad n . Odwrotnie zaś — gdy macierz A ma rzad n , to jej kolumny tworzą bazę \mathcal{A} przestrzeni K^n , czyli $A = M(\text{id})_{\mathcal{A}}^{\text{st}}$. Stąd A jest odwracalna, czyli jest macierzą izomorfizmu. Warunki (ii), (iii) i (iv) są więc równoważne, a ostatni argument pokazuje także, że (iv) \Rightarrow (i).

Jednak implikacja (i) \Rightarrow (ii) jest jasna, bo jeśli $A = M(\text{id})_{\mathcal{A}}^{\mathcal{B}}$, to macierz odwrotną do niej jest $M(\text{id})_{\mathcal{B}}^{\mathcal{A}}$. \square

Wniosek 20.1.10

Niech $A \in M_{n \times n}(K)$. Jeśli macierz $B \in M_{n \times n}(K)$ spełnia warunek $AB = I_n$, to $BA = I_n$.

Dowód. Możemy przyjąć, że A, B są macierzami w bazach standardowych pewnych przekształceń liniowych $\phi, \psi : K^n \rightarrow K^n$. Z Twierdzenia 20.1.5 wynika, że $\phi \circ \psi$ jest identycznością. Stąd $\phi : K^n \rightarrow K^n$ jest monomorfizmem, a $\psi : K^n \rightarrow K^n$ jest epimorfizmem (Wniosek 17.1.8). Wiemy jednak, że monomorfizm pomiędzy przestrzeniami tego samego skończonego wymiaru jest izomorfizmem, podobnie dla epimorfizmu (Wniosek 16.1.10). Stąd przekształcenia liniowe ϕ, ψ są wzajemnie odwrotnymi izomorfizmami, a stąd macierz B jest odwrotna do A . \square

Alternatywny, choć być może nieco bardziej pomysłowy dowód powyższego wniosku opiera się o Wniosek 20.1.9. Skoro $AB = I_n$, to $r(AB) = n$. Wiemy, że $r(AB) \leq \min\{r(A), r(B)\}$, więc $r(A) = r(B) = n$. Stąd macierze A, B są odwracalne i na mocy Obserwacji 20.1.2, również macierz BA jest odwracalna. Jeśli zatem przemnożymy równość $AB = I_n$ z lewej przez B , dostając $BAB = B$, a dalej z prawej przez A , dostając $BABA = BA$. Skoro istnieje $(BA)^{-1}$, mamy $(BABA)(BA)^{-1} = (BA)(BA)^{-1}$, a stąd $BA = I_n$.

* * *

Wniosek 20.1.9 oznacza, że macierz A jest odwracalna wtedy i tylko wtedy, gdy macierz A' powstała z A przez wykonanie operacji elementarnej (na wierszach lub kolumnach) również jest odwracalna.

Nadszedł czas, aby uzyskać świadomość, że wykonywanie operacji elementarnych również wiąże się z mnożeniem z odpowiedniej strony przez pewną macierz odwracalną. Zrozumienie tego podejścia, wyposaży nas w nowe narzędzia do badania własności macierzy (także ich rozkładów) oraz pozwoli uzyskać nowe kryterium istnienia macierzy odwrotnej (i sposób jej wyznaczania).

Definicja 20.1.11: Macierz dodawania wiersza/kolumny

Niech $a \in K$ oraz niech $n \in \mathbb{N}$. Definiujemy macierz

$$E_{ij}^n(a) = \begin{bmatrix} 1 & \cdots & a & \cdots & 1 \end{bmatrix},$$

to znaczy $E_{ij}^n(a) = [a_{st}] \in M_{n \times n}(K)$, gdzie

$$a_{st} = \begin{cases} a & \text{gdy } s = i, t = j \\ 1 & \text{gdy } s = t \\ 0 & \text{w pozostałych przypadkach.} \end{cases}$$

Przykład. Poniżej przedstawiony jest efekt mnożenia macierzy o wyrazach a, b, c, d, e, f odpowiednio przez macierze $E_{21}^2(2)$ oraz $E_{13}^3(2)$:

$$\begin{bmatrix} 1 & 0 \\ 2 & 1 \end{bmatrix} \cdot \begin{bmatrix} a & b & c \\ d & e & f \end{bmatrix} = \begin{bmatrix} a & b & c \\ d+2a & e+2b & f+2c \end{bmatrix}$$

$$\begin{bmatrix} a & b & c \\ d & e & f \end{bmatrix} \cdot \begin{bmatrix} 1 & 0 & 2 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} a & b & c+2a \\ d & e & f+2d \end{bmatrix}.$$

Definicja 20.1.12: Macierz zamiany wierszy/kolumn

Niech $n \in \mathbb{N}$. Definiujemy macierz

$$T_{ij}^n = \begin{bmatrix} 1 & \cdots & 1 & 0 & 1 & \cdots & 1 \\ & \ddots & & 1 & 0 & 1 & \cdots & 1 \\ & & \vdots & & \vdots & & \vdots & \\ & & & 1 & \cdots & 1 & 0 & 1 \\ & & & & \ddots & & & 1 \end{bmatrix},$$

to znaczy $T_{ij}^n = [a_{st}] \in M_{n \times n}(K)$, gdzie

$$a_{st} = \begin{cases} 1 & \text{gdy } s = t \neq i, j \\ 1 & \text{gdy } s = i, t = j \text{ lub } s = j, t = i \\ 0 & \text{w pozostałych przypadkach.} \end{cases}$$

Przykład. Poniżej przedstawiony jest efekt mnożenia macierzy o wyrazach a, b, c, d, e, f odpowiednio przez macierze T_{12}^2 oraz T_{12}^3 :

$$\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \cdot \begin{bmatrix} a & b & c \\ d & e & f \end{bmatrix} = \begin{bmatrix} d & e & f \\ a & b & c \end{bmatrix},$$

$$\begin{bmatrix} a & b & c \\ d & e & f \end{bmatrix} \cdot \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} b & a & c \\ e & d & f \end{bmatrix}$$

Definicja 20.1.13: Macierz mnożenia wiersza/kolumny przez stałą

Niech $c \in K$, $c \neq 0$ oraz niech $n \in \mathbb{N}$. Definiujemy macierz

$$I_i^n(c) = \begin{bmatrix} 1 & \cdots & 1 & c & 1 & \cdots & \cdots & 1 \end{bmatrix},$$

to znaczy $I_i^n(c) = [a_{st}] \in M_{n \times n}(K)$, gdzie

$$a_{st} = \begin{cases} c & \text{gdy } s = t = i \\ 1 & \text{gdy } s = t \neq i \\ 0 & \text{w pozostałych przypadkach.} \end{cases}$$

Przykład. Poniżej przedstawiony jest efekt mnożenia macierzy o wyrazach a, b, c, d, e, f odpowiednio przez macierze $I_2^2(x)$ oraz $I_2^3(x)$:

$$\begin{bmatrix} 1 & 0 \\ 0 & x \end{bmatrix} \cdot \begin{bmatrix} a & b & c \\ d & e & f \end{bmatrix} = \begin{bmatrix} a & b & c \\ xd & xe & xf \end{bmatrix}$$

$$\begin{bmatrix} a & b & c \\ d & e & f \end{bmatrix} \cdot \begin{bmatrix} 1 & 0 & 0 \\ 0 & x & 0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} a & xb & c \\ d & xe & f \end{bmatrix}.$$

Definicja 20.1.14

Macierze $E_{ij}(a)$, T_{ij} , $I_i(c)$ nazywamy *macierzami operacji elementarnych*.

Przykład.

$$E_{24}^5(a) = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & a & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}, \quad T_{35}^5 = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 \end{bmatrix}, \quad I_1^5(c) = \begin{bmatrix} c & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}.$$

Uwaga 20.1.15

Dla każdej macierzy $A \in M_{m \times n}(K)$ zachodzi:

- a) $E_{ij}^m(a) \cdot A$ powstaje z A przez dodanie do i -tego wiersza j -tego wiersza pomnożonego przez a ,
- b) $A \cdot E_{ij}^n(a)$ powstaje z A przez dodanie do j -tej kolumny i -tej kolumny pomnożonej przez a ,
- c) $T_{ij}^m \cdot A$ powstaje z A przez przestawienie i -tego i j -tego wiersza,
- d) $A \cdot T_{ij}^n$ powstaje z A przez przestawienie i -tej i j -tej kolumny,
- e) $I_i^m(c) \cdot A$ powstaje z A przez pomnożenie i -tego wiersza przez c ,
- f) $A \cdot I_i^n(c)$ powstaje z A przez pomnożenie i -tej kolumny przez c .

W szczególności, biorąc za A macierz jednostkową $I \in M_{n \times n}(K)$, otrzymujemy:

- (i) $E_{ij}^m(x) \cdot I$ powstaje z I przez dodanie do i -tego wiersza j -tego wiersza pomnożonego przez x ,
- (ii) T_{ij}^m powstaje z I przez przestawienie i -tego i j -tego wiersza,
- (iii) $I_i^m(c)$ powstaje z I przez pomnożenie i -tego wiersza przez c ,

Przyjrzymy się teraz odwracalności macierzy operacji elementarnych.

Uwaga 20.1.16

Dla każdej macierzy $S \in M_{n \times n}(K)$ jednego z typów $E_{ij}^n(a), T_{ij}^n, I_i^n(c)$ istnieje macierz S' tego samego typu taka, że:

$$S'S = SS' = I.$$

Dowód. Oczywiście S' ma być odwrotnością S . Łatwo sprawdzić, że

- jeśli $S = E_{ij}^n(a)$, to $S^{-1} = E_{ij}^n(-a)$,
- jeśli $S = T_{ij}^n$, to $S^{-1} = T_{ij}^n$,
- jeśli $S = I_i^n(c)$, to $S^{-1} = I_i^n(c^{-1})$

□

Wniosek 20.1.17

Niech $A' \in M_{n \times n}(K)$ będzie macierzą otrzymaną z A przez sprowadzenie do zredukowanej postaci schodkowej za pomocą elementarnych operacji na wierszach. Następujące warunki są równoważne:

- macierz A jest odwracalna
- $A' = I$.

W szczególności następujące warunki są równoważne:

- macierz A jest odwracalna,
- A jest iloczynem macierzy typu $E_{ij}(x), T_{ij}, I_i(y)$, gdzie $y \neq 0$.

Dowód. Teza wynika natychmiast z tego, że każdą macierz można sprowadzić do postaci schodkowej i schodkowej zredukowanej A' operacjami odpowiedniego typu. Niech $A' = W_r \cdot \dots \cdot W_1 \cdot A$, gdzie W_i – macierze operacji elementarnych. Wówczas zgodnie ze wzorem na odwrotność iloczynu mamy

$$A = (W_r \cdot \dots \cdot W_1)^{-1} A' = W_1^{-1} \cdot \dots \cdot W_r^{-1} A'.$$

Teza wynika zatem z faktu, że macierze W_i^{-1} są również macierzami operacji elementarnych. □

Przykład. Postacią zredukowaną macierzy

$$A = \begin{bmatrix} 6 & 6 & -2 \\ -1 & 0 & 0 \\ -1 & 1 & 0 \end{bmatrix}$$

jest macierz jednostkowa I_3 , dokładniej:

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & \frac{1}{3} \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & \frac{1}{6} & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ -6 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} A = I_3.$$

* * *

Na koniec warto odnotować jeszcze jedną własność, która będzie przydatna w drugim semestrze. Zauważmy, że transponowanie macierzy operacji elementarnych zmienia jedynie macierz operacji typu (1).

$$(E_{ij}(x))^T = E_{ji}(x), \quad (T_{ij})^T = T_{ij}, \quad (I_i(y))^T = I_i(y).$$

Innymi słowy, macierze operacji typu (2) i (3) są symetryczne. Wniosek jest następujący: jeśli P jest macierzą operacji elementarnej, to macierz P^TAP powstaje z macierzy A :

- albo przez dodanie do i -tego wiersza j -tej kolumny i dodanie do i -tej kolumny j -tej kolumny,
- albo przez zamianę i -tego wiersza i j -tego wiersza oraz zamianę i -tej kolumny i j -tej kolumny,
- albo przez przemnożenie i -tego wiersza i j -tej kolumny przez y .

□

20.2 Wybór przykładowych pytań

Proszę się nie uczyć na pamięć odpowiedzi czy wyjaśnienie! Takich pytań można ułożyć znacznie więcej i służą one jedynie sprawdzeniu podstawowego zrozumienia treści zajęć. Jeśli odpowiedź na dowolne z powyższych pytań sprawia Państwu problem, proszę pomyśleć który fragment notatek dotyczy tego pytania, wrócić do ich lektury, a w razie dalszych trudności poprosić prowadzącego o wskazówkę.

Uwaga. Przydatne będą formuły (zakładamy, że wszystkie działania są wykonalne):

$$A(BC) = (AB)C, \quad A(B + C) = AB + AC, \quad (AB)^{-1} = B^{-1}A^{-1}, \quad (AB)^T = B^T A^T$$

1. Niech $\mathcal{A} = ((1, 3), (2, 4))$ będzie bazą przestrzeni liniowej \mathbb{R}^2 . Wypisz macierze $M(\text{id})_{\mathcal{A}}^{st}$ oraz $M(\text{id})_{st}^{\mathcal{A}}$.
2. Założmy, że macierze $A, B \in M_{n \times n}(K)$ są odwracalne. Czy wynika stąd, że macierz $A + B$ jest odwracalna?
3. Założmy, że macierz $A \in M_{n \times n}(K)$ spełnia równość $A \cdot A^T = I_n$. Czy wynika stąd, że A jest macierzą odwracalną?
4. Macierz odwracalna $A \in M_{n \times n}(K)$ oraz pewna macierz $B \in M_{n \times n}(K)$ spełniają warunek $AB = 0$. Czy wynika stąd, że $A = 0$?
5. Macierz odwracalna $A \in M_{n \times n}(K)$ spełnia $A = A^3$. Czy wynika stąd, że $A = A^{-1}$? Jeśli tak, podaj przykład takiej macierzy.
6. Niech $A \in M_{n \times n}(K)$ będzie odwracalna. Rozstrzygnij, czy dla każdego n i każdego ciała K poniższa macierz jest odwracalna:
 - A^{-1} ,
 - A^2 ,
 - $A + I_n$,
 - $A + A^{-1}$,
 - A^T ,
 - $A^T A$ oraz AA^T ,
 - $A + (A^T)^{-1}$.
7. Macierz A jest odwracalna. Jak zmieni się macierz A^{-1} , jeśli w macierzy A
 - zamienimy i -ty i j -ty wiersz macierzy A ,
 - dodamy j -ty wiersz przemnożony przez a do i -tego wiersza,
 - przemnożymy i -ty wiersz przez $c \neq 0$?
 - wykonamy w analogiczny sposób którąś z operacji powyższego typu na kolumnach?
8. Niech $A \in M_{n \times n}(K)$ spełnia warunek $A^2 + 2A + I_n = 0$. Czy wynika stąd, że A jest odwracalna?
9. Niech $A \in M_{n \times n}(K)$ spełnia warunek $A^2 - 2A + 2I_n = 0$. Czy wynika stąd, że A jest odwracalna?
10. Uzasadnij, że jeśli $A \in M_{n \times m}(K)$, to macierze AA^T oraz $A^T A$ są symetryczne.
11. W $M_{2 \times 2}(K)$ rozważmy podzbiór X złożony ze wszystkich macierzy nieodwracalnych. Czy X jest podprzestrzenią $M_{2 \times 2}(K)$?
12. Czy każda macierz rozmiaru $n \times n$ jest iloczynem macierzy operacji elementarnych?
13. Niech $A \in M_{3 \times 4}(K)$. Przez jaką macierz należy pomnożyć macierz A , i z której strony, aby otrzymać macierz powstającą z A przez
 - dodanie czwartego wiersza do wiersza drugiego,
 - przemnożenie czwartej kolumny przez 2,
 - zamianę pierwszej i czwartej kolumny,
 - odjęcie od pierwszej kolumny drugiej kolumny przemnożonej przez 2.
14. Niech $A \in M_{4 \times 3}(K)$. Wyznacz taką macierz P , że macierz PA powstaje z A przez dodanie do czwartego wiersza sumy pozostałych wierszy.

20.3 Zadania do samodzielnej pracy

Zadania ilustrujące umiejętności niezbędne do zaliczenia przedmiotu oznaczone są symbolem ♠. Przy tych zadaniach dodany jest opis umiejętności, które sprawdzają.

Uwaga. Przydatne będą formuły (zakładamy, że wszystkie działania są wykonalne):

$$A(BC) = (AB)C, \quad A(B + C) = AB + AC, \quad (AB)^{-1} = B^{-1}A^{-1}, \quad (AB)^T = B^T A^T$$

1. (♠) Przedstaw macierz

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

jako iloczyn macierzy operacji elementarnych.

2. (♠) Przedstaw macierz

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

jako iloczyn macierzy $E_{ij}(a)$ dla pewnych i, j, a .

3. (♠) Znajdź taką macierz odwracalną $A \in M_{4 \times 4}(K)$, że

$$A^T \cdot \begin{bmatrix} 1 & 1 & 0 & 0 \\ 1 & 1 & 1 & 0 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 \end{bmatrix} \cdot A = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

4. Macierz $A \in M_{n \times n}(\mathbb{R})$ spełnia $A^3 = 2I_n$. Uzasadnij, że macierz $A^2 - 2A + 2I_n$ jest odwracalna.
5. Macierz $A \in M_{n \times n}(K)$ spełnia, dla pewnej liczby całkowitej k , warunek $A^k = 0$. Uzasadnij, że macierz $I_n - A$ jest odwracalna.
6. Niech $A \in M_{m \times n}(K)$, gdzie $m > n$ Uzasadnij, że macierz $A \cdot A^T$ nie jest odwracalna.
7. Niech $A \in M_{m \times n}(K)$ oraz $B \in M_{m \times n}$, przy czym $AB = I_n$ oraz $BA = I_m$. Uzasadnij, że $m = n$.
8. Udowodnij, że dowolna macierz $A \in M_{n \times n}(K)$ jest albo odwracalna, albo istnieją takie macierze $B, C \in M_{n \times n}(K)$, że $AB = 0$ oraz $CA = 0$.
9. Wykaż, że każdą macierz rzędu 1 rozmiaru $n \times n$ można zapisać jako iloczyn postaci $A \cdot B^T$, gdzie $A, B \in M_{1 \times n}(K)$.
10. Założmy, że macierze $A, B, A - B$ oraz $B^{-1} - A^{-1}$ są odwracalne w $M_{n \times n}(K)$. Uzasadnij, że

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

11. Niech $A \in M_{n \times n}, B \in M_{n \times n}$ będą takie, że $I_n - AB$ jest macierzą odwracalną i jej odwrotność jest macierzą C . Uzasadnij, że macierz $I + BCA$ jest macierzą odwrotną do macierzy $I - BA$.

12. Macierze $A, B \in M_{n \times n}(K)$ spełniają warunki:

$$A^2 = A = A^T, \quad B^2 = B = B^T, \quad A + B = I_n.$$

Uzasadnij, że

$$(A - B)(A - B)^T = I_n.$$

13. Wykaż, że jeśli $A, B \in M_{m \times n}(K)$ są macierzami schodkowymi zredukowanymi i B jest otrzymana z A elementarnymi operacjami na wierszach, to $A = B$. Wywnioskuj stąd, że dla każdej macierzy $A \in M_{m \times n}(K)$ istnieje dokładnie jedna macierz schodkowa zredukowana otrzymana z A operacjami elementarnymi na wierszach.

Rozdział 21

Macierze blokowe Równoważność macierzy

21.1 Wykład 21

Na tym wykładzie wprowadzimy pewne wygodne narzędzie przydatne do patrzenia zarówno na iloczyn macierzy, jak i na macierze przekształceń liniowych.

Definicja 21.1.1: Postać blokowa macierzy

Niech $M = [m_{ij}]$ będzie macierzą o n wierszach i m kolumnach. Dla dodatnich liczb całkowitych n_1, m_1 oraz nieujemnych liczb całkowitych n_2, m_2 takich, że $n_1 + n_2 = n$ oraz $m_1 + m_2 = m$ określamy macierze:

- M_{11} — o wyrazach $[m_{ij}]$, dla $1 \leq i \leq n_1$ oraz $1 \leq j \leq m_1$,
- M_{21} — o wyrazach $[m_{ij}]$, dla $n_1 + 1 \leq i \leq n$ oraz $1 \leq j \leq m_1$,
- M_{12} — o wyrazach $[m_{ij}]$, dla $1 \leq i \leq n_1$ oraz $m_1 + 1 \leq j \leq m$,
- M_{22} — o wyrazach $[m_{ij}]$, dla $n_1 + 1 \leq i \leq n$ oraz $m_1 + 1 \leq j \leq m$.

Mówimy wtedy, że macierz M jest w POSTACI BLOKOWEJ O BLOKACH M_{ij} (blok ij -ty)

$$M = \begin{bmatrix} M_{11} & M_{12} \\ M_{21} & M_{22} \end{bmatrix}$$

Dla $n_2 = 0$ lub $m_2 = 0$ pisać będziemy odpowiednio $M = [M_{11} \ M_{12}]$ oraz $M = \begin{bmatrix} M_{11} \\ M_{21} \end{bmatrix}$.

Czasem dla podkreślenia podziału na bloki używamy w notacji linii do ich separowania:

$$M = \left[\begin{array}{c|c} M_{11} & M_{12} \\ \hline M_{21} & M_{22} \end{array} \right], \quad M = [\ M_{11} \ | \ M_{12} \], \quad M = \left[\begin{array}{c} M_{11} \\ \hline M_{21} \end{array} \right]$$

Oczywiście macierz może mieć wiele postaci blokowych, a jej bloki można dzielić na mniejsze bloki. Najczęściej interesuje nas identyfikowanie bloków o prostych własnościach. Dla przykładu, biorąc macierz

$$A = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & \cos \theta & -\sin \theta \\ 0 & 0 & \sin \theta & \cos \theta \end{bmatrix}$$

wskazać można różne podziały na bloki, ale szczególnie przydatny dla zastosowań jest podział

$$A = \begin{bmatrix} I_2 & 0 \\ 0 & O_\theta \end{bmatrix},$$

o bloku I_2 , dwóch blokach zerowych, oraz bloku będącym macierzą obrotu.

Obserwacja 21.1.2: Mnożenie macierzy blokowych

Niech $A \in M_{n \times p}(K)$ oraz $B \in M_{p \times m}$ będą macierzami w postaciach blokowych

$$A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}, \quad B = \begin{bmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{bmatrix}.$$

Wówczas jeśli wszystkie iloczyny $A_{il} \cdot B_{lj}$ istnieją, to AB jest macierzą blokową postaci:

$$AB = \begin{bmatrix} A_{11}B_{11} + A_{12}B_{21} & A_{11}B_{12} + A_{12}B_{22} \\ A_{21}B_{11} + A_{22}B_{21} & A_{21}B_{12} + A_{22}B_{22} \end{bmatrix}.$$

Dowód. Uzasadnienie przeprowadzić można oczywiście bezpośrednio, ale oprzemy je na prostej własności rozdzielności mnożenia macierzy względem dodawania. Mówiąc ona, że dla macierzy X, Y, Z takich, że iloczyny XY oraz XZ istnieją, mamy $X(Y + Z) = XY + XZ$. Oczywiście składników może być więcej.

Każda macierz blokowa możemy zapisać w postaci

$$M = \begin{bmatrix} M_{11} & M_{12} \\ M_{21} & M_{22} \end{bmatrix} = \begin{bmatrix} M_{11} & 0 \\ 0 & 0 \end{bmatrix} + \begin{bmatrix} 0 & M_{12} \\ 0 & 0 \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ M_{21} & 0 \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ 0 & M_{22} \end{bmatrix}.$$

Nietrudno sprawdzić¹, że wymnożenie takich sum zapisanych dla A, B , daje macierz jak w tezie. \square

Przykłady prostych zastosowań formuły wyżej.

- Niech $A \in M_{n \times n}(K)$ oraz $B \in M_{m \times m}(K)$. Wówczas dla dowolnej liczby naturalnej k :

$$\begin{bmatrix} A & 0 \\ 0 & B \end{bmatrix}^k = \begin{bmatrix} A^k & 0 \\ 0 & B^k \end{bmatrix}.$$

- Niech $X \in M_{n \times n}(K)$. Wówczas macierz blokowa $\begin{bmatrix} I_n & X \\ 0 & I_n \end{bmatrix}$ jest odwracalna, gdyż:

$$\begin{bmatrix} I_n & X \\ 0 & I_n \end{bmatrix} \cdot \begin{bmatrix} I_n & -X \\ 0 & I_n \end{bmatrix} = \begin{bmatrix} I_n + 0 & -X + X \\ 0 + 0 & 0 + I_n \end{bmatrix} = I_{2n}.$$

- Niech $A \in M_{m \times p}(K)$, $B \in M_{m \times q}(K)$, $C \in M_{n \times p}(K)$, $D \in M_{n \times q}(K)$, gdzie $m+n = p+q$. Wówczas zachodzi równość rzędów macierzy blokowych

$$r \begin{bmatrix} A & B \\ C & D \end{bmatrix} = r \begin{bmatrix} D & C \\ B & A \end{bmatrix}.$$

Istotnie, mamy następującą równosć:

$$\begin{bmatrix} 0 & I_n \\ I_m & 0 \end{bmatrix} \cdot \begin{bmatrix} A & B \\ C & D \end{bmatrix} \cdot \begin{bmatrix} 0 & I_p \\ I_q & 0 \end{bmatrix} = \begin{bmatrix} D & C \\ B & A \end{bmatrix}.$$

Mnożenie przez macierz odwracalną nie zmienia rzędu (jest to macierz izomorfizmu), a macierze typu

$$\begin{bmatrix} 0 & I_m \\ I_n & 0 \end{bmatrix}$$

są oczywiście rzędu $m+n$, gdyż ich wiersze to parami różne wektory bazy standardowej, czyli są odwracalne. Można się o tym zresztą przekonać na wiele innych sposobów, w tym za pomocą podejścia funkcyjnego. Macierz ta jest w istocie macierzą standardową przekształcenia liniowego $\phi : K^{n+m} \rightarrow K^{n+m}$ postaci:

$$\phi((x_1, \dots, x_n, y_1, \dots, y_m)) = (y_1, \dots, y_m, x_1, \dots, x_n),$$

które zamienia pierwsze n współrzędnych wektora z ostatnimi m współrzędnymi. Oczywiście przekształceniem odwrotnym do ϕ jest $\psi : K^{n+m} \rightarrow K^{n+m}$, które zamienia pierwsze m współrzędnych z ostatnimi n . Mamy (oznaczając dla ułatwienia macierz zerową rozmiaru $k \times l$ jako $0_{k \times l}$):

$$\begin{bmatrix} 0_{n \times m} & I_n \\ I_m & 0_{m \times n} \end{bmatrix} \cdot \begin{bmatrix} 0_{m \times n} & I_m \\ I_n & 0_{n \times m} \end{bmatrix} = \begin{bmatrix} 0_{n \times m} \cdot 0_{m \times n} + I_n \cdot I_n & 0_{n \times m} \cdot I_m + I_n \cdot 0_{n \times m} \\ I_m \cdot 0_{m \times n} + 0_{m \times n} \cdot I_n & I_m \cdot I_m + 0_{m \times n} \cdot 0_{n \times m} \end{bmatrix} = \begin{bmatrix} I_n & 0 \\ 0 & I_m \end{bmatrix} = I_{n+m}.$$

¹Ten nietrudny i raczej evidentny argument można nieco sformalizować — piszę o tym pod koniec wykładu.

Poniższa uwaga jest prostym zastosowaniem naszych rozważań.

Uwaga 21.1.3

Macierz $A \in M_{n \times n}(K)$ jest odwracalna wtedy i tylko wtedy, gdy macierz blokowa $[A | I_n]$ po sprowadzeniu do zredukowanej postaci schodkowej ma postać $[I_n | B]$. Wówczas $B = A^{-1}$.

Dowód. Jeśli A jest odwracalna, to na mocy Wniosku 20.1.7 istnieje macierz P , będąca iloczynem macierzy typu $E_{ij}(a)$, T_{ij} oraz $I_i(c)$ spełniająca $PA = I_n$. Mamy zatem

$$P \cdot [A | I_n] = [PA | I_n P] = [I_n | P]$$

przy czym macierz $[I_n | P]$ jest w zredukowanej postaci schodkowej. Ponadto $P = A^{-1}$.

Na odwrót: założmy, że macierz $[A | I_n]$ po sprowadzeniu do zredukowanej postaci schodkowej ma postać $[I_n | B]$ dla pewnej macierzy $B \in M_{n \times n}(K)$. Wówczas

$$[I_n | B] = Q \cdot [A | I_n],$$

dla pewnej macierzy odwracalnej Q (będącej iloczynem macierzy typu $E_{ij}(a)$, T_{ij} oraz $I_i(c)$). Stąd dostajemy $QA = I_n$ oraz $B = QI = Q$, więc A jest odwracalna i $A^{-1} = Q = B$. \square

Przykład. Mamy

$$A = \begin{bmatrix} 1 & 0 & 1 \\ -1 & 1 & 0 \\ 1 & 1 & 3 \end{bmatrix}, \quad [A | I_3] = \left[\begin{array}{ccc|ccc} 1 & 0 & 1 & 1 & 0 & 0 \\ -1 & 1 & 0 & 0 & 1 & 0 \\ 1 & 1 & 3 & 0 & 0 & 1 \end{array} \right].$$

Sprowadzając $[A | I_3]$ do zredukowanej postaci schodkowej dostajemy:

$$\begin{aligned} \left[\begin{array}{ccc|ccc} 1 & 0 & 1 & 1 & 0 & 0 \\ -1 & 1 & 0 & 0 & 1 & 0 \\ 1 & 1 & 3 & 0 & 0 & 1 \end{array} \right] &\rightarrow \left[\begin{array}{ccc|ccc} 1 & 0 & 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 1 & 1 & 0 \\ 0 & 1 & 2 & -1 & 0 & 1 \end{array} \right] \rightarrow \left[\begin{array}{ccc|ccc} 1 & 0 & 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 1 & 1 & 0 \\ 0 & 0 & 1 & -2 & -1 & 1 \end{array} \right] \rightarrow \\ &\rightarrow \left[\begin{array}{ccc|ccc} 1 & 0 & 0 & 3 & 1 & -1 \\ 0 & 1 & 0 & 3 & 2 & -1 \\ 0 & 0 & 1 & -2 & -1 & 1 \end{array} \right]. \quad \text{Stąd } A^{-1} = \begin{bmatrix} 3 & 1 & -1 \\ 3 & 2 & -1 \\ -2 & -1 & 1 \end{bmatrix}. \end{aligned}$$

* * *

Wróćmy teraz do przekształceń liniowych i przyjrzymy się następującemu pytaniu: jakie macierze może mieć (w różnych bazach) dane przekształcenie liniowe $\phi : K^n \rightarrow K^m$? Jak się okazuje, zależy to jedynie od rzędu ϕ , czyli od wymiaru obrazu ϕ .

Definicja 21.1.4: Macierze równoważne

Macierze $A, B \in M_{m \times n}(K)$ nazywamy RÓWNOWAŻNYMI, jeśli istnieją takie macierze odwracalne $C \in M_m(K)$ oraz $D \in M_n(K)$ takie, że $B = CAD$.

Uwaga 21.1.5

Dla macierzy $A, B \in M_{m \times n}(K)$ następujące warunki są równoważne:

- (i) macierze A oraz B są równoważne,
- (ii) istnieje przekształcenie liniowe $\phi : K^n \rightarrow K^m$ oraz bazy $\mathcal{A}, \mathcal{A}'$ przestrzeni K^n i bazy $\mathcal{B}, \mathcal{B}'$ przestrzeni K^m , że $A = M(\phi)_{\mathcal{A}}^{\mathcal{B}}$ oraz $B = M(\phi)_{\mathcal{A}'}^{\mathcal{B}'}$,
- (iii) B może być otrzymana z A ciągiem operacji elementarnych na wierszach i kolumnach,
- (iv) $r(A) = r(B)$.

Dowód. Dowodzimy (i) \Rightarrow (ii). Niech $M(\phi)_{st}^{st} = A$ oraz niech $B = CAD$. Przez \mathcal{C}, \mathcal{D} określamy bazy złożone odpowiednio z kolumn macierzy odwracalnych C^{-1} i D . Wówczas $C = M(\text{id})_{st}^{\mathcal{C}}$ oraz $D = M(\text{id})_{\mathcal{D}}^{st}$. Stąd $B = CAD = M(\phi)_{\mathcal{D}}^{\mathcal{C}}$.

Implikacja (ii) \Rightarrow (iii) wynika z tego, że macierz jest odwracalna wtedy i tylko wtedy, gdy jest iloczynem macierzy operacji elementarnych (I semestr). W szczególności skoro $B = M(\text{id})_{\mathcal{B}}^{B'} \cdot A \cdot M(\text{id})_{\mathcal{A}'}^{\mathcal{A}}$, gdzie $M(\text{id})_{\mathcal{B}}^{B'}$ oraz $M(\text{id})_{\mathcal{A}'}^{\mathcal{A}}$ są odwracalne, to mamy (iii).

Implikacja (iii) \Rightarrow (iv) jest jasna, ponieważ wykonanie operacji elementarnej nie zmienia rzędu.

Natomiast jeśli zachodzi (iv) to znaczy, że za pomocą operacji elementarnych na wierszach i kolumnach każdej z macierzy A i B można sprowadzić za pomocą operacji elementarnych do macierzy $X = [x_{ij}]$, gdzie $r = r(A) = r(B)$ oraz

$$x_{ij} = \begin{cases} 1, & \text{dla } i = j = 1, \dots, r, \\ 0, & \text{dla pozostałych } i, j, \end{cases}$$

czyli do macierzy rozmiaru $m \times n$ o postaci blokowej

$$X = \begin{bmatrix} I_r & 0 \\ 0 & 0 \end{bmatrix}.$$

Rzeczywiście, wiemy że macierze tego samego rzędu można sprowadzić za pomocą operacji elementarnych na wierszach do postaci A' , B' , które są w zredukowanej postaci schodkowej i mają $r = r(A) = r(B)$ niezerowych wierszy. Innymi słowy istnieją macierze odwracalne $E, G \in M_m(K)$, że $A' = EA$ oraz $B' = GB$, gdzie E, G — iloczyny odpowiednich macierzy operacji elementarnych na wierszach. Teraz zaczynamy wykonywać operacje kolumnowe. Po pierwsze porządkujemy kolumny tak, by wyrazy wiodące stojące w pierwszych r wierszach stały w pierwszych r kolumnach. Innymi słowy za pomocą operacji kolumnowych sprowadzamy macierze A' oraz B' do postaci A'' oraz B'' postaci blokowej:

$$A'' = \begin{bmatrix} I_r & * \\ 0 & 0 \end{bmatrix}, \quad B'' = \begin{bmatrix} I_r & ** \\ 0 & 0 \end{bmatrix}.$$

Widzimy zatem, że za pomocą pierwszych r kolumn można za pomocą operacji kolumnowych doprowadzić te macierze do postaci X . Innymi słowy istnieją macierze odwracalne E, F, G, H , że

$$X = EAF \quad \text{oraz} \quad X = GBH.$$

Zatem $EAF = GBH$, czyli

$$A = E^{-1}GBH^{-1}.$$

Macierze $E^{-1}G$ oraz H^{-1} są odwracalne, więc uzyskujemy (i). \square

W świetle przytoczonego twierdzenia powyższa definicja mówi, że macierze są równoważne wtedy i tylko wtedy, gdy są macierzami tego samego przekształcenia liniowego. Równoważność macierzy nie jest na pożółk pojęciem mówiącym za wiele o geometrii przekształcenia, choć jest ono przydatne z punktu widzenia samej teorii macierzy.

Przedstawimy teraz wniosek, uogólniający Uwagę 20.1.7.

Wniosek 21.1.6

Dla każdego przekształcenia liniowego przestrzeni skończonego wymiaru $\phi : V \rightarrow W$ rzędu r istnieją bazy \mathcal{A}, \mathcal{B} przestrzeni V, W takie, że

$$M(\phi)_{\mathcal{A}}^{\mathcal{B}} = \begin{bmatrix} I_r & 0 \\ 0 & 0 \end{bmatrix}.$$

Czytelnik udowodni oczywiście powyższy fakt bez trudu także w oparciu o pojęcie macierzy przekształcenia liniowego. Jeśli $\alpha_{r+1}, \dots, \alpha_n$ jest bazą jądra ϕ oraz $\alpha_1, \dots, \alpha_r$ dowolnym układem dopełniającym ją do bazy V , wówczas biorąc bazę $\mathcal{A} = (\alpha_1, \dots, \alpha_n)$ przestrzeni V oraz dowolną bazę \mathcal{B} przestrzeni W , której pierwsze r wektorów to $\phi(\alpha_1), \dots, \phi(\alpha_r)$ (ten układ r wektorów jest bazą obrazu ϕ), otrzymujemy

$$M(\phi)_{\mathcal{A}}^{\mathcal{B}} = \begin{bmatrix} I_r & 0 \\ 0 & 0 \end{bmatrix}.$$

Wykład zakończymy uogólnieniem konstrukcji macierzy blokowej, dopuszczającym większą liczbę bloków. Ważny będzie dla nas szczególnie przypadek takiej konstrukcji dotyczącej macierzy rozmiaru $n \times n$.

Definicja 21.1.7: Postać blokowa związana z partycją

Rozważmy rozkład^a $n = n_1 + \dots + n_k$, gdzie n, n_1, \dots, n_k, k są dodatnimi liczbami całkowitymi. Dla macierzy $A = [a_{st}] \in M_{n \times n}(K)$ oraz dla dowolnych $1 \leq p, q \leq k$ rozważmy macierze

$$A_{pq} \in M_{n_p \times n_q}(K),$$

że w i -tym wierszu i j -tej kolumnie A_{pq} stoi wyraz a_{st} macierzy A , gdzie dla $n_0 = 0$ mamy

$$s = n_0 + \dots + n_{p-1} + i \quad \text{oraz} \quad t = n_0 + \dots + n_{q-1} + j.$$

Mówimy wtedy, że A jest w POSTACI BLOKOWEJ odpowiadającej rozkładowi $n = n_1 + \dots + n_k$ o blokach A_{pq} postaci:

$$A = \begin{bmatrix} A_{11} & A_{12} & \dots & A_{1k} \\ A_{21} & A_{22} & \dots & A_{2k} \\ \vdots & \vdots & \ddots & \vdots \\ A_{k1} & A_{k2} & \dots & A_{kk} \end{bmatrix}. \quad (*)$$

Bloki $A_{pp} \in M_{n_p \times n_p}(K)$ nazywamy BLOKAMI DIAGONALNYMI.

^aZwany czasem partycją liczby n .

Czytelnik z pewnością widzi, że definicję powyższą można uogólnić również na przypadek macierzy prostokątnych rozmiaru $m \times n$, gdzie każdą z liczb m, n przedstawiamy w postaci sumy $m = m_1 + \dots + m_k$ oraz $n = n_1 + \dots + n_s$ i tworzymy bloki A_{pq} rozmiarów $m_p \times n_q$. Jak wspomniałem wcześniej — macierze tego typu można dostać z macierzy blokowych z Definicji 21.1.1 przez traktowanie każdego bloku jako macierzy blokowej, tak jednak by rozmiary mniejszych bloków były zgodne, czyli aby bloki A_{pq} były rozmiarów $m_p \times n_q$ (czasem nawet tej zgodności rozmiarów się nie zakłada — choć zawsze można ją uzyskać).

Uogólnienie formuły na mnożenie macierzy w postaci blokowej przedstawionej wyżej nie przedstawia trudności, a jej dowód to prosta indukcja po liczbie bloków, oparta w swojej istocie o Obserwację 21.1.2. W przyszłości wiele twierdzeń formułowanych będzie właśnie w języku macierzy tego typu.

Na koniec naszkicujemy krótko rolę, jaką dla mnożenia macierzy i zrozumienia zarówno działań na operacjach elementarnych, jak i macierzach blokowych, odgrywają jedynki macierzowe. Podejście przedstawione niżej pozwala sformalizować dowody rezultatów omawianych w niniejszym i poprzednim rozdziale.

Przypomnijmy, że jedynki macierzowe to takie macierze $E_{ij} \in M_{m \times n}(K)$, których jedyny niezerowy wyraz równy jest 1 i znajduje się on w i -tym wierszu i j -tej kolumnie. Gdy chcemy podkreślić, że jedynki są danego rozmiaru piszemy $E_{ij}^{m,n}$, stosując notację podobną, jak przy macierzach operacji elementarnych. Gdy $m = n$ piszemy po prostu E_{ij}^n (nie należy tego mylić z macierzą operacji elementarnej $E_{ij}^n(a)$).

Jak wiemy, przestrzeń liniowa macierzy $M_{m \times n}(K)$ jest rozpięta przez jedynki macierzowe tak, że macierz $A = [a_{ij}] \in M_{m \times n}(K)$ można zapisać w postaci

$$A = \sum_{i,j} a_{ij} E_{ij}^{m,n}.$$

Widzimy zatem, że do wykonywania iloczynu macierzy $A = [a_{ij}] \in M_{m \times n}(K)$ oraz $B = [b_{kl}] \in M_{p \times q}(K)$ wystarczy w istocie zrozumienie w jaki sposób mnożą się jedynki macierzowe różnych rozmiarów. Mamy

$$AB = \left(\sum_{i,j} a_{ij} E_{ij}^{m,n} \right) \cdot \left(\sum_{i,j} b_{kl} E_{ij}^{p,q} \right) = \sum_{i,j,k,l} a_{ij} b_{kl} \cdot E_{ij}^{m,n} \cdot E_{kl}^{p,q},$$

Oczywiście iloczyn elementów postaci

$$E_{ij}^{m,n} \cdot E_{kl}^{p,q}$$

ma sens tylko, gdy $n = p$ (macierze mają odpowiednie rozmiary, by je pomnożyć). Wtedy jedyny niezerowy wyraz takiej macierzy rozmiaru $m \times q$ może stać w i -tym wierszu i l -tej kolumnie, o ile tylko $j = k$.

Uwaga 21.1.8

Mamy

$$E_{ij}^{m,n} \cdot E_{kl}^{n,q} = \begin{cases} E_{il}^{m,q}, & \text{gdy } j = k, \\ 0, & \text{w przeciwnym przypadku.} \end{cases}$$

W szczególności dla $m = n = p = q$ mamy stąd:

$$E_{ij}^n \cdot E_{kl}^n = \begin{cases} E_{il}^n, & \text{gdy } j = k \\ 0, & \text{w przeciwnym przypadku.} \end{cases}$$

Oczywiście mnożenie macierzy w sposób powyższy nie wygląda na bardziej atrakcyjne od zwykłej definicji, ale pozwala sformalizować pewne rozumowania (na przykład bez trudu uzyskamy stąd łączność mnożenia macierzy). Zauważmy na przykład, że zachodzi następująca oczywista obserwacja.

Uwaga 21.1.9

Niech $A = [a_{ij}] \in M_{m \times n}(K)$ będzie macierzą o m wierszach i n kolumnach. Niech E_{ij}^k będą jedynkami macierzowymi w $M_{k \times k}(K)$. Wówczas zachodzą następujące fakty.

- (i) Jedyne niezerowe wyrazy macierzy $E_{ij}^m A$ są w i -tym wierszu i są to wyrazy j -tego wiersza A .
- (ii) Jedyne niezerowe wyrazy macierzy $A E_{ij}^n$ są w j -tej kolumnie i są to wyrazy i -tej kolumny A .

Dowód. Uzasadnimy tylko (i). Jedyne niezerowe iloczyny $E_{ij}^m \cdot E_{kl}^{m,n}$ to $E_{il}^{m,n}$, dla $j = k$. Zatem mamy

$$E_{ij}^m \cdot \sum_{k,l} a_{k,l} E_{kl}^{m,n} = \sum_l a_{jl} E_{il}^{m,n}$$

Uzyskana macierz ma jedynie i -ty wiersz niezerowy i jego kolejne wyrazy to elementy j -tego wiersza A . \square

Przykłady. Mnożenie z lewej strony przez E_{12}^3 oraz z prawej strony przez E_{31}^4

$$\begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \cdot \begin{bmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{bmatrix} = \begin{bmatrix} 3 & 4 \\ 0 & 0 \\ 0 & 0 \end{bmatrix}, \quad \begin{bmatrix} 1 & 2 & 3 & 4 \\ 5 & 6 & 7 & 8 \end{bmatrix} \cdot \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} = \begin{bmatrix} 3 & 0 & 0 & 0 \\ 7 & 0 & 0 & 0 \end{bmatrix}.$$

Obserwacja 21.1.10: Operacje elementarne jako kombinacje liniowe

Niech $E_{ij}^n(a), T_{ij}^n, I_i^n(c)$ będą macierzami operacji elementarnych rozmiaru $n \times n$. Wówczas:

$$E_{ij}^n(a) = I_n + aE_{ij}^n, \quad T_{ij}^n = I_n + E_{ji}^n + E_{ij}^n - E_{ii}^n - E_{jj}^n, \quad I_i^n(c) = I_n + (c-1) \cdot E_{ii}^n.$$

Powyższa interpretacja wraz z Uwagą 21.1.9 pozwala na łatwe formalne uzasadnienie Uwagi 20.1.5 mówiącej o działaniu operacji elementarnych. Dla przykładu: mamy

$$E_{ij}^n(a) \cdot A = (I_n + aE_{ij}^n) \cdot A = A + aE_{ij}^n A.$$

Wynik po prawej w istocie polega na dodaniu do macierzy A macierzy o jednym niezerowym i -tym wierszu, w którym znajdują się wyrazy j -tego macierzy A przemnożone przez a .

Podobnie sformalizować można dowód reguły mnożenia macierzy blokowych z Obserwacją 21.1.2. Dla

$$M = \begin{bmatrix} M_{11} & M_{12} \\ M_{21} & M_{22} \end{bmatrix} = \begin{bmatrix} M_{11} & 0 \\ 0 & 0 \end{bmatrix} + \begin{bmatrix} 0 & M_{12} \\ 0 & 0 \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ M_{21} & 0 \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ 0 & M_{22} \end{bmatrix},$$

gdzie $M \in M_{m \times n}(K)$, możemy traktować każdy ze składników wyżej jako macierz postaci $E_i^n \cdot M \cdot E_j^m$, gdzie E_i^n oraz E_j^m są odpowiednimi sumami jedynek macierzowych rozmiaru n oraz m . Mnożenie takich macierzy sprowadza się wówczas do problemu mnożenia jedynek macierzowych. Argument ten można uogólnić także na przypadek, gdy mnożymy macierze o większej liczbie bloków, na przykład postaci (*).

21.2 Wybór przykładowych pytań

Proszę się nie uczyć na pamięć odpowiedzi czy wyjaśnienie! Takich pytań można ułożyć znacznie więcej i służą one jedynie sprawdzeniu podstawowego zrozumienia treści zajęć. Jeśli odpowiedź na dowolne z poniższych pytań sprawia Państwu problem, proszę pomyśleć który fragment notatek dotyczy tego pytania, wrócić do ich lektury, a w razie dalszych trudności poprosić prowadzącego o wskazówkę.

1. Wykorzystując podział na bloki 2×2 , oblicz iloczyn

$$\begin{bmatrix} 0 & 0 & 1 & 2 \\ 0 & 0 & 3 & 4 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \cdot \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 2 \\ 0 & 0 & 3 & 4 \end{bmatrix}.$$

2. Wykorzystując podział na bloki 2×2 , oblicz iloczyn

$$\begin{bmatrix} 1 & 2 & 1 & 0 \\ 3 & 4 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix} \cdot \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 2 & 1 & 2 \\ 3 & 4 & 3 & 4 \end{bmatrix}.$$

3. Wykorzystując podział na bloki 2×2 , oblicz iloczyn

$$\begin{bmatrix} 0 & -1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & -1 \\ 0 & 0 & 0 & 0 & 1 & 0 \end{bmatrix}^{100}.$$

4. Niech $A \in M_{n \times n}(K)$. Wykonaj iloczyn

$$\begin{bmatrix} I_n & -A \\ 0 & I_n \end{bmatrix} \cdot \begin{bmatrix} A & -I_n \\ I_n & 0 \end{bmatrix}.$$

5. Znajdź takie macierze blokowe $X, Y \in M_{2n \times 2n}(K)$, że dla dowolnych macierzy $A, B, C \in M_{n \times n}(K)$ mamy

$$X \cdot \begin{bmatrix} A & B \\ 0 & C \end{bmatrix} \cdot Y = \begin{bmatrix} C & B \\ 0 & A \end{bmatrix}.$$

Czy istnieje macierz $X \in M_{2n \times 2n}(K)$, że dla dowolnych różnych macierzy $A, B \in M_{n \times n}(K)$ mamy

$$\begin{bmatrix} A & A \\ B & B \end{bmatrix} \cdot X = \begin{bmatrix} B & B \\ A & A \end{bmatrix}?$$

6. Rozstrzygnij, czy prawdziwe są następujące równości. Jeśli nie, podaj kontrprzykład.

$$r \begin{bmatrix} A & 0 \\ 0 & B \end{bmatrix} = r(A) + r(B), \quad r \begin{bmatrix} A & C \\ 0 & B \end{bmatrix} = r(A) + r(B), \quad r \begin{bmatrix} A & A \\ B & B \end{bmatrix} = r(A) + r(B).$$

7. Niech $A, B \in M_{n \times n}(K)$. Uzasadnij, że

$$r \begin{bmatrix} I_n & A \\ 0 & B \end{bmatrix} = n + r(B), \quad r \begin{bmatrix} A & A \\ A & B \end{bmatrix} = r(A) + r(B - A).$$

8. Niech A, B będą macierzami odwracalnymi. Znajdź macierze odwrotne do macierzy blokowych

$$\begin{bmatrix} A & 0 \\ 0 & -B \end{bmatrix}, \quad \begin{bmatrix} 0 & A \\ B & 0 \end{bmatrix}.$$

9. Uzasadnij, że poniższe macierze blokowe są przemienne, czyli $XY = YX$.

$$X = \begin{bmatrix} I_m & A \\ 0 & I_n \end{bmatrix}, \quad Y = \begin{bmatrix} I_m & B \\ 0 & I_n \end{bmatrix}$$

10. Uzasadnij, że macierz blokowa

$$X = \begin{bmatrix} -I_m & A \\ B & I_n \end{bmatrix}$$

spełnia $X^2 = I_{m+n}$, o ile $A = 0$ lub $B = 0$. Czy jest to warunek konieczny?

21.3 Zadania do samodzielnej pracy

Zadania ilustrujące umiejętności niezbędne do zaliczenia przedmiotu oznaczone są symbolem ♠. Przy tych zadaniach dodany jest opis umiejętności, które sprawdzają.

1. (♠) Rozważmy macierz

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

Zapisz macierz w postaci blokowej. Policz A^{100} . Podaj przykład macierzy B takiej, że $B^2 = A$.

2. Uzasadnij, że

$$r \begin{bmatrix} A & B \\ 2A & -5B \end{bmatrix} = r(A) + r(B).$$

3. Niech A, B będą macierzami rozmiaru $n \times n$. Znajdź taką macierz X , że

$$X^2 = \begin{bmatrix} AB & 0 \\ 0 & BA \end{bmatrix}.$$

4. Niech $A = XBY$, dla pewnych $A, B \in M_{m \times n}(K)$ oraz $X \in M_{m \times m}(K)$, $Y \in M_{n \times n}(K)$. Niech

$$A' = \begin{bmatrix} I_n & 0 \\ 0 & A \end{bmatrix}, \quad B' = \begin{bmatrix} I_n & 0 \\ 0 & B \end{bmatrix}.$$

Znajdź takie macierze X', Y' , że $A' = X'B'Y'$.

5. Niech $\phi \in L(V, V)$. Dla pewnej bazy $\mathcal{A} = (\alpha_1, \dots, \alpha_n)$ przestrzeni V macierz $M(\phi)_{\mathcal{A}}^{\mathcal{A}}$ ma postać blokową $\begin{bmatrix} A & B \\ C & D \end{bmatrix}$, gdzie $A \in M_{k \times k}(K)$. Wskaż taką bazę \mathcal{A}' , że $M(\phi)_{\mathcal{A}}^{\mathcal{A}'}$ ma postać $\begin{bmatrix} D & C \\ B & A \end{bmatrix}$.

6. Uzasadnij, że $r[A | B] \leq r(A) + r(B)$.

7. Dana jest macierz odwracalna A w postaci blokowej $X = \begin{bmatrix} A & B \\ C & D \end{bmatrix}$, przy czym $A \in M_{n \times n}(K)$ jest odwracalna. Wykaż, że macierz X ma rząd n wtedy i tylko wtedy, gdy $D = CA^{-1}B$. Wskazówka: przedstaw macierz X jako iloczyn macierzy blokowych.

8. Macierze $A, B, C, D, E, F, G, H \in M_{n \times n}(K)$ spełniają warunki:

$$AE + BG = I_n, \quad AF + BH = O, \quad CE + DG = 0, \quad CF + DH = I_n.$$

Uzasadnij, że zachodzą także równości:

$$EA + FC = I_n, \quad EB + FD = 0, \quad GA + HC = 0, \quad GB + HD = I_n.$$

9. Korzystając z tożsamości

$$\begin{bmatrix} I_n & 0 \\ -A & I_m \end{bmatrix} \cdot \begin{bmatrix} I_n & B \\ A & 0 \end{bmatrix} \cdot \begin{bmatrix} I_n & -B \\ 0 & I_m \end{bmatrix} = \begin{bmatrix} I_n & 0 \\ 0 & -AB \end{bmatrix},$$

dla macierzy $A \in M_{m \times n}(K)$, $B \in M_{n \times m}(K)$, uzasadnij, że $r(A) + r(B) \leq r(AB) + n$.

10. Wykaż, znajdując odpowiedni rozkład, że $A, B \in M_{n \times n}(K)$, to

$$r \begin{bmatrix} A & AB \\ B & B + B^2 \end{bmatrix} = r(A) + r(B).$$

11. Rozważmy macierz blokową $A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}$, gdzie $A_{11} \in M_{r \times r}(K)$ oraz $A_{22} \in M_{n-r \times n-r}(K)$.

Wykaż, że jeśli

$$r(A) = r[A_{11} | A_{12}] = r \begin{bmatrix} A_{11} \\ A_{21} \end{bmatrix},$$

to macierz A_{11} jest odwracalna.

Rozdział 22

Przestrzeń sprzężona

22.1 Wykład 22

W tym rozdziale wprowadzimy szczególnie ważny typ przekształceń liniowych, mający duże znaczenie i w samej matematyce i w naukach ją wykorzystujących. Damy bowiem podstawowe algebraiczne intuicje dla ważnego zjawiska dualności, znajdującego odpowiedni język w stworzonej w połowie XX wieku teorii kategorii, a mający podstawy choćby w klasycznej geometrii (zwłaszcza rzutowej).

Definicja 22.1.1: Funkcjonał liniowy, przestrzeń sprzężona (dualna)

FUNKCJONALEM LINIOWYM (albo FORMĄ LINIOWĄ) na przestrzeni liniowej V nad ciałem K nazywamy przekształcenie liniowe $\phi : V \rightarrow K$. Zbiór $V^* = L(V, K)$ funkcji liniowych na przestrzeni liniowej V nazywamy PRZESTRZENIĄ SPRZĘŻONĄ (DUALNĄ) do V .

Przykłady

- Odwzorowanie $\phi : \mathbb{R}^3 \rightarrow \mathbb{R}$ dane wzorem

$$\phi((x_1, x_2, x_3)) = 5x_1 + 7x_2 - 4x_3$$

jest funkcjonałem liniowym na przestrzeni \mathbb{R}^3 .

- Wiemy już z wcześniejszych rozdziałów, że dla każdego elementu $\phi \in (K^n)^*$ istnieją $a_1, \dots, a_n \in K$ takie, że ϕ zadana jest wzorem

$$\phi((x_1, \dots, x_n)) = a_1x_1 + \dots + a_nx_n.$$

- Przekształcenie $\text{tr} \in (M_{n \times n}(\mathbb{R}))^*$ zwane ŚLADEM, zadane wzorem

$$\text{tr}([a_{ij}]) = a_{11} + \dots + a_{nn}.$$

- Dla $t \in K$ przekształcenie $\phi_t : K[x] \rightarrow K$ dane wzorem $\phi_t(w) = w(t)$ jest funkcjonałem liniowym.
- Niech $C \subset \mathbb{R}^\infty$ będzie podprzestrzenią złożoną ze wszystkich ciągów zbieżnych. Dla każdego szeregu bezwzględnie zbieżnego $\sum_i a_i$ określić można funkcjonał $\phi : C \rightarrow \mathbb{R}$ wzorem $\phi((x_i)) = \sum_i a_i x_i$.

Uwaga 22.1.2

Jeśli V jest przestrzenią skończenie wymiarową, to $V \simeq V^*$.

Dowód. Niech $\dim V = n$. Wówczas $V^* = L(V, K)$ jest również wymiaru n , jako przestrzeń izomorficzna z $M_{1 \times n}(K)$. Dwie przestrzenie liniowe tego samego skończonego wymiaru są izomorficzne. \square

Uwaga 22.1.3

Niech $\mathcal{A} = (\alpha_1, \dots, \alpha_n)$ będzie bazą przestrzeni V i niech $\phi_i : V \rightarrow K$ będzie jedynym funkcjonałem liniowym takim, że:

$$\phi_i(\alpha_j) = \begin{cases} 1, & \text{jeśli } i = j \\ 0, & \text{jeśli } i \neq j. \end{cases} \quad (*)$$

Wówczas:

- (a) dla dowolnego $\alpha \in V$ mamy $\alpha = \phi_1(\alpha)\alpha_1 + \phi_2(\alpha)\alpha_2 + \dots + \phi_n(\alpha)\alpha_n$, czyli ϕ_i przyporządkowuje wektorowi jego i -tą współrzędną w bazie \mathcal{A} ,
- (b) dla dowolnego $\phi \in V^*$ mamy $\phi = \phi(\alpha_1)\phi_1 + \phi(\alpha_2)\phi_2 + \dots + \phi(\alpha_n)\phi_n$, i jest to przedstawienie jednoznaczne,
- (c) układ funkcjonałów $\mathcal{A}^* = (\phi_1, \dots, \phi_n)$ jest bazą V^* i wartość funkcjonału $\phi \in V^*$ na wektorze α_j jest j -tą współrzędną tego funkcjonału w bazie \mathcal{A}^* .

Przykłady.

- Dla bazy \mathcal{A} przestrzeni \mathbb{R}^3 postaci

$$\alpha_1 = (1, 1, 1), \quad \alpha_2 = (1, 1, 0), \quad \alpha_3 = (1, 0, 0)$$

układ funkcjonałów f_i określony warunkami wyżej istnieje i ma postać:

$$\phi_1((x_1, x_2, x_3)) = x_3, \quad \phi_2((x_1, x_2, x_3)) = x_2 - x_3, \quad \phi_3((x_1, x_2, x_3)) = x_1 - x_2.$$

Na przykład dla $\alpha = (10, 5, 2)$ otrzymujemy:

$$\phi_1(\alpha) = 2, \quad \phi_2(\alpha) = 3, \quad \phi_3(\alpha) = 5 \quad \text{oraz} \quad (10, 5, 2) = 2(1, 1, 1) + 3(1, 1, 0) + 5(1, 0, 0).$$

- Współrzędne funkcjonału $\phi \in (\mathbb{R}^3)^*$, gdzie

$$\phi((x_1, x_2, x_3)) = 2x_1 - 9x_2 + 5x_3$$

w bazie sprzężonej do $\mathcal{A} = ((1, 1, 1), (1, 1, 0), (1, 0, 0))$ wynoszą:

$$\phi((1, 1, 1)) = -2, \quad \phi((1, 1, 0)) = -7, \quad \phi((1, 0, 0)) = 2.$$

Dowód. Pierwszy punkt jest oczywisty. Istotnie, jeśli $\alpha = a_1\alpha_1 + \dots + a_n\alpha_n$, to obkładając tę równość z obydwu stron funkcjonałem ϕ_i dostajemy: $\phi_i(\alpha) = a_1\phi_i(\alpha_1) + a_2\phi_i(\alpha_2) + \dots + a_n\phi_i(\alpha_n)$. Tylko element $\phi_i(\alpha_i)$ sumy po prawej jest niezerowy, z definicji ϕ_i . A zatem $\phi_i(\alpha) = a_i$.

Punkty (b) i (c) postulają, że (ϕ_1, \dots, ϕ_n) jest bazą V^* . Wykażalismy to już w dowodzie Twierdzenia 17.1.3 (opisując pewną bazę $L(V, W)$), ale powtórzmy te argumenty. Sprawdźmy najpierw, że układ ten jest liniowo niezależny. Założymy, że istnieją takie $a_1, \dots, a_n \in K$, że

$$a_1\phi_1 + \dots + a_n\phi_n = 0,$$

przy czym 0 po prawej stronie interpretujemy jako funkcjonał zerowy! A zatem $a_1\phi_1 + \dots + a_n\phi_n$ jest funkcjonałem, który dowolny wektor posyła na zero. Z drugiej strony biorąc element α_i bazy \mathcal{A} mamy:

$$(a_1\phi_1 + \dots + a_n\phi_n)(\alpha_i) = a_1\phi_1(\alpha_i) + \dots + a_n\phi_n(\alpha_i) = a_i.$$

A zatem $a_i = 0$, dla każdego $1 \leq i \leq n$. A zatem (ϕ_1, \dots, ϕ_n) jest układem liniowo niezależnym.

Zobaczmy teraz, że (ϕ_1, \dots, ϕ_n) rozpinia V^* . Niech $\phi \in V^*$. Twierdzimy, że:

$$\phi = \phi(\alpha_1)\phi_1 + \phi(\alpha_2)\phi_2 + \dots + \phi(\alpha_n)\phi_n.$$

Aby stwierdzić czy dwa przekształcenia są identyczne wystarczy to sprawdzić na dowolnej bazie V , na przykład na $(\alpha_1, \dots, \alpha_n)$. Wówczas rzeczywiście:

$$(\phi(\alpha_1)\phi_1 + \phi(\alpha_2)\phi_2 + \dots + \phi(\alpha_n)\phi_n)(\alpha_i) = \phi(\alpha_1)\phi_1(\alpha_i) + \dots + \phi(\alpha_n)\phi_n(\alpha_i) = \phi(\alpha_i) \cdot 1.$$

□

Definicja 22.1.4: Baza sprzężona (dualna)

Bazę \mathcal{A}^* zdefiniowaną wyżej wzorem (\star) nazywamy BAZĄ SPRZĘŻONĄ (DUALNA) do bazy \mathcal{A} . Jej elementy ϕ_1, \dots, ϕ_n ozaczamy czasem jako $\alpha_1^*, \dots, \alpha_n^*$.

Przykład. Niech $\alpha_1 = (1, 3)$, $\alpha_2 = (2, 7)$ będzie bazą przestrzeni \mathbb{R}^2 . Weźmy

$$\alpha_1^*(x_1, x_2) = 7x_1 - 2x_2 \quad \text{oraz} \quad \alpha_2^*(x_1, x_2) = -3x_1 + 1x_2.$$

Wówczas, jak we wzorze (\star) mamy: $\alpha_1^*(\alpha_1) = 1$, $\alpha_1^*(\alpha_2) = 0$, $\alpha_2^*(\alpha_1) = 0$, $\alpha_2^*(\alpha_2) = 1$. Zauważmy, że jeśli wpiszemy współczynniki funkcjonalów w macierz (w kolumny), a wektory z wyjściowej bazy wpiszemy w wiersze macierzy, dostaniemy zależność:

$$\begin{bmatrix} 1 & 3 \\ 2 & 7 \end{bmatrix} \cdot \begin{bmatrix} 7 & -3 \\ -2 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = I_2.$$

Problem wyznaczania bazy sprzężonej jest problemem rozwiązywania układu równań danego warunkami z (\star) . Macierze: mająca w wierszach wektory z \mathcal{A} oraz: mająca w kolumnach wektory z \mathcal{A}^* są **odwrotne**. Można to też uzasadnić pisząc odpowiedni układ równań liniowych.

W przyszłości przekonają się Państwo, że przestrzeń sprzężona ma kluczowe znaczenie zarówno dla zrozumienia pewnych własności geometrycznych, jak i analitycznych, wyrażanych w języku dualności. Będzie to szczególnie widoczne w przestrzeniach liniowych rzeczywistych mających dodatkową strukturę wyznaczoną przez tzw. iloczyny skalarne. W tym miejscu podamy pewną podstawową intuicję, która pomoże nam zinterpretować wyniki wyżej. Wzmocnimy ją na końcu mówiąc o tzw. anihilatorze podprzestrzeni.

Do tej pory mogłyby się wydawać, że w skończeniu wymiarowym przypadku przestrzenie V oraz V^* są w zasadzie identyczne – w końcu to przestrzenie izomorficzne. Symetria między tymi przestrzeniami jest jednak dość ciekawa. Odwołajmy się tu do intuicji, jakie daje twierdzenie Kroneckera-Capellego.

Układ k równań liniowych od n zmiennych o współczynnikach w ciele K wyznacza jednoznacznie podprzestrzeń $(K^n)^*$ rozpiętą przez odpowiednie funkcjonały, np. układ

$$x_1 + x_2 + x_3 = 0, x_2 - x_3 = 0$$

wyznacza podprzestrzeń $(K^3)^*$ rozpiętą przez

$$\phi_1((x_1, x_2, x_3)) = x_1 + x_2 + x_3, \quad \phi_2((x_1, x_2, x_3)) = x_2 - x_3.$$

Zauważmy też, że dla dowolnego układu wektorów z $\text{lin}(\phi_1, \phi_2)$ zbiór rozwiązań tego układu zawiera przestrzeń rozwiązań układu $\phi_1 = 0, \phi_2 = 0$. A zatem możemy powiedzieć, że dwuwymiarowej przestrzeni $\text{lin}(\phi_1, \phi_2) \subseteq (K^3)^*$ odpowiada podprzestrzeń rozwiązań układu $\phi_1 = 0, \phi_2 = 0$, która ma wymiar 1. Innymi słowy, podprzestrzeniom „dużego” wymiaru w K^n odpowiadają pewne podprzestrzenie „małego” wymiaru w $(K^n)^*$. Nadajmy temu pewną intuicję geometryczną.

Powiedzieliśmy, że każdy funkcjonał ϕ w K^n jest postaci $a_1x_1 + \dots + a_nx_n$. Innymi słowy jądro tego funkcjonału składa się ze wszystkich takich wektorów (x_1, \dots, x_n) , które dla ustalonego wektora (a_1, \dots, a_n) spełniają $a_1x_1 + \dots + a_nx_n = 0$. Nazwijmy na użytek intuicji „wektorami prostopadłymi” do wektora $\alpha = (a_1, \dots, a_n)$. Zbiór wszystkich takich wektorów określmy przez α^\perp . Oczywiście jest to podprzestrzeń. Podajmy dwa powody dla mówienia o niej. Po pierwsze — dualność. Należenie wektora (x_1, \dots, x_n) do $(a_1, \dots, a_n)^\perp$ jest równoważne należeniu wektora (a_1, \dots, a_n) do $(x_1, \dots, x_n)^\perp$. Początkowa zróżnicowanie ról x_1, \dots, x_n jako niewiadomych i a_1, \dots, a_n jako współczynników równania znikło, a rola tych wektorów stała się symetryczna, choć symetria ta — jak wspomnialiśmy wyżej — dokonuje się mimo wszystko pomiędzy przestrzeniami potencjalnie różnych wymiarów.

Po drugie — rzutowania. Zauważmy, że jeśli $K = \mathbb{R}$, to zgodnie z Uwagą 22.1.3, biorąc funkcjonał ϕ opisany wzorem $\phi((x_1, \dots, x_n)) = a_1x_1 + \dots + a_nx_n$, i dowolny wektor β w K^n , mamy

$$\beta = b \cdot (a_1, \dots, a_n) + \alpha',$$

gdzie $\alpha' \in \ker \phi$ oraz $b = \phi(\beta)$. Istotnie, $\phi((a_1, \dots, a_n)) = a_1^2 + \dots + a_n^2$, więc o ile funkcjonał nie jest zerowy, to wektor (a_1, \dots, a_n) dopełnić można do bazy \mathbb{R}^n wektorami z $\ker \phi$.

O wartości $\phi(\beta)$ można zatem myśleć jako o współczynniku powstającym przez „zrzutowanie” wektora β na $\text{lin}((a_1, \dots, a_n))$, „wzdłuż” ϕ . W ten sposób Uwagę 22.1.3(a) można rozumieć w następujący sposób: jeśli $\alpha_1, \dots, \alpha_n$ jest bazą V to wektor $\alpha \in V$ jest sumą swoich „rzutów” na podprzestrzeni $\text{lin}(\alpha_i)$. Co więcej, również punkt (b) można rozumieć tak, że dowolna funkcja jest sumą swoich „rzutów” na podprzestrzenie wyznaczone przez pewne funkcje bazowe. Kiedyś przekonają się Państwo, że w podobny sposób rozumieć można oznaczenie dx przy całkowaniu.

Zdecydowanie trudniejsza niż dla przestrzeni skończonego wymiaru jest ogólna teoria przestrzeni sprzężonych nieskończonym wymiaru, mających jednakże kluczową rolę choćby w analizie. Zobaczmy, że nawet przestrzeń sprzężona do przestrzeni wielomianów nie jest zupełnie banalna.

Przykład. Pokażemy, że $K[x]^* \simeq K^\infty$.

Dowód. Niech $\phi : K[x]^* \rightarrow K^\infty$ będzie określone wzorem:

$$\phi(f) = (f(1), f(x), f(x^2), \dots)$$

Innymi słowy, i -ty wyraz ciągu $\phi(f)$ to wartość funkcjonału f na x^i . Oczywiście $\phi(af+bg) = a\phi(f)+b\phi(g)$, dla dowolnych $f, g \in K[x]^*$ oraz $a, b \in K$. Jest to więc przekształcenie liniowe.

Nietrudno widzieć, że ϕ ma trywialne jądro. Tylko przekształcenie zerowe ma tę własność, że $f(x^i) = 0$, dla każdego i (jednoznaczne określenie na bazie). Co więcej, ϕ jest surjekcją, bo dla $c = (a_0, a_1, \dots) \in K^\infty$ bierzemy $f \in K[x]^*$ taki, że $f(x^i) = a_i$ (określamy funkcjonal na bazie $K[x]$, więc taki f istnieje). Oczywiście $c = \phi(f)$, więc ϕ jest izomorfizmem. \square

Nie mamy do dyspozycji teorii liczb kardynalnych, ale w jej języku dowodzi się, że jeśli: $\dim K[x] = \omega$, to $\dim K^\infty \geq 2^\omega$, zależnie od mocy ciała K . Zatem $K[x] \not\simeq K[x]^*$ (jeśli nie ma bijekcji, nie ma izomorfizmu). Dowodzi się też, że jeżeli V jest przestrzenią liniową nad ciałem K i $\dim(V) = \infty$, to

$$\dim V^* = |K|^{\dim V}.$$

* * *

Czym jest przestrzeń sprzężona do przestrzeni sprzężonej, czyli V^{**} ? Jest to przestrzeń złożona z funkcjonałów $\phi : V^* \rightarrow K$. A zatem element $\phi \in V^{**}$ każdemu funkcjonalowi f z V^* przypisuje element z K .

Definicja 22.1.5: Ewaluacja

Niech $\phi : V \rightarrow K$ będzie funkcjonałem liniowym. EWALUACJĄ funkcjonału ϕ w wektorze $\alpha \in V$ nazywamy przekształcenie $\text{ev}_\alpha \in V^{**}$ zadane wzorem

$$\text{ev}_\alpha(\phi) = \phi(\alpha).$$

Jest zupełnie jasne, że ev_α jest przekształceniem liniowym, dla każdego $\alpha \in V$. Co więcej, każda ewaluacja jest elementem V^{**} . A zatem każdemu wektorowi z V przypisaliśmy w naturalny sposób element V^{**} . Ma to ważne skutki w przypadku skończenia wymiarowego.

Twierdzenie 22.1.6

Niech V będzie przestrzenią skończonego wymiaru. Przekształcenie $e_V : V \rightarrow V^{**}$ zadane wzorem

$$e_V(\alpha) = \text{ev}_\alpha$$

jest izomorfizmem przestrzeni liniowych.

Dowód. Oczywiście e_V jest przekształceniem liniowym. Skoro V oraz V^{**} są tego samego wymiaru wystarczy pokazać, że e_V jest monomorfizmem. Założmy, że $\alpha \in \ker(e_V)$. Wówczas ev_α jest elementem zerowym w V^{**} , czyli dla każdego $\phi : V \rightarrow K$ mamy $\text{ev}_\alpha(\phi) = 0$. Z definicji ev_α oznacza to, że $\phi(\alpha) = 0$, dla każdego $\phi \in V^*$. A zatem wektor α ma tę własność, że ewaluowany na każdym funkcjonalu liniowym jest zerem. Jedyny element z V o tej własności to 0, a więc $\ker(e_V) = \{0\}$. \square

Na koniec pokażemy przykład związku między przestrzenią sprzężoną i przestrzenią ilorazową.

Definicja 22.1.7

Niech U będzie podprzestrzenią V . ANIHILATOREM podprzestrzeni U w V^* , oznaczanym przez $\text{Ann}(U)$, nazywamy zbiór wszystkich funkcjonałów na V , które zerują się na całym U , czyli:

$$\text{Ann}(U) = \{f \in V^* \mid f(u) = 0, \text{ dla każdego } u \in U\}.$$

Nietrudno widzieć, że $\text{Ann}(U)$ to podprzestrzeń liniowa. Oczywiście $\text{Ann}(\{0\}) = V^*$ oraz $\text{Ann}(V) = \{0\}$. Im większa jest podprzestrzeń U , tym mniejszy jest jej anihilator. W istocie, anihilator jest jednym ze sposobów ścisłego wyrażenia intuicji dotyczącej odpowiedniości pomiędzy „małymi” podprzestrzeniami V i „dużymi” podprzestrzeniami V^* , gdy wychodzimy poza kontekst przestrzeni współrzędnych K^n .

Przykład. Niech $U = \text{lin}((1, 0, 0, 0, 0), (1, 1, 0, 0, 0)) \subseteq \mathbb{R}^5$. Interesują nas wszystkie funkcjonały

$$f((x_1, x_2, x_3, x_4, x_5)) = a_1x_1 + a_2x_2 + a_3x_3 + a_4x_4 + a_5x_5 \in (\mathbb{R}^5)^*,$$

że

$$f((1, 0, 0, 0, 0)) = f((1, 1, 0, 0, 0)) = 0 \Leftrightarrow a_1 = 0 \wedge a_1 + a_2 = 0.$$

Nietrudno zatem widzieć, że $\text{Ann}(U)$ jest podprzestrzenią dwuwymiarową. Być może Czytelnik zauważy, że jeśli weźmiemy bazę U i dopełnimy ją do bazy \mathcal{A} całego V , to baza $\text{Ann}(U)$ składa się z tych elementów bazy sprzężonej do \mathcal{A} , które zerują się na U .

Takie spojrzenie bardzo już przypomina język przestrzeni ilorazowej. Rzeczywiście, zachodzi następujący, nietrudny rezultat, wiążący anihilator U z przestrzenią ilorazową V/U , mający bardzo czytelną interpretację w języku twierdzenia Kroneckera-Capellego, gdzie anihilator $U \subseteq K^n$ utożsamiać można ze wszystkimi takimi równaniami, które spełnione są na wszystkich elementach z U .

Twierdzenie 22.1.8

Ma miejsce izomorfizm

$$\text{Ann}(U) \simeq (V/U)^*.$$

A zatem możemy identyfikować funkcjonały na V/U z elementami $\text{Ann}(U)$.

Dowód. Niech $f \in \text{Ann}(U)$. Jest to liniowy funkcjonal na V , który znika na U . A zatem możemy określić funkcjonał liniowy f' na V/U dany wzorem:

$$f'(v + U) = f(v).$$

Innymi słowy, f' posyła warstwę $v + U$ na skalar $f(v)$. Zobaczmy, że jest to przekształcenie dobrze określone. Założmy, że $v + U = v' + U$. Musimy sprawdzić, że $f'(v + U) = f'(v' + U)$. Istotnie, skoro $v + U = v' + U$, to $v - v' \in U$, a zatem:

$$0 = f(v - v') = f(v) - f(v').$$

Pokażmy, że przyporządkowanie $\Psi : \text{Ann}(U) \rightarrow (V/U)^*$ dane wzorem $\Psi(f) = f'$ jest izomorfizmem. Weźmy najpierw $f \in \ker(\Psi)$. A zatem $f' = \Psi(f)$ jest funkcjonałem zerowym na V/U . A zatem:

$$0 = f'(v + U) = f(v), \text{ dla każdego } v \in V.$$

Stąd f jest funkcjonałem zerowym. W szczególności jest to element zerowy $\text{Ann}(U)$. Zatem $\ker(\Psi) = \{0\}$. Aby pokazać, że Ψ to surjekcja, weźmy $g \in (V/U)^*$. Określamy element $f \in V^*$ jako:

$$f(v) = g(v + U), \text{ dla każdego } v \in V.$$

Twierdzimy, że f należy tak naprawdę do $\text{Ann}(U)$. Istotnie, jeśli $u \in U$, to $g(u + U) = g(U) = 0$, skoro U jest elementem zerowym V/U oraz g jest liniowe. Zatem $f(u) = 0$, czyli $f \in \text{Ann}(U)$. Z definicji Ψ wynika, że $\Psi(f) = g$, więc Ψ jest surjekcją. Oczywiście jest to przekształcenie liniowe. \square

22.2 Wybór przykładowych pytań

Proszę się nie uczyć na pamięć odpowiedzi czy wyjaśnienie! Takich pytań można ułożyć znacznie więcej i służą one jedynie sprawdzeniu podstawowego zrozumienia treści zajęć. Jeśli odpowiedź na dowolne z poniższych pytań sprawia Państwu problem, proszę pomyśleć który fragment notatek dotyczy tego pytania, wrócić do ich lektury, a w razie dalszych trudności poprosić prowadzącego o wskazówkę.

1. Dane są funkcjonały liniowe $f, g : V \rightarrow \mathbb{R}$. Wiedząc, że dla pewnych $v, w \in V$ mamy $f(v) = f(w) = 1$ oraz $g(v) = 2, g(w) = 3$:
 - wyznacz wartość $f + g$ na wektorze $w - v$,
 - wyznacz wartość $2f - g$ na wektorze $3v$,
 - rozstrzygnij, czy funkcjonały f, g są liniowo niezależne.
2. Czy każdy funkcjonał liniowy jest albo przekształceniem zerowym, albo epimorfizmem?
3. Niech $\dim V = 1$ oraz $f, g \in V^*$. Czy f, g są liniowo zależne?
4. Niech $\dim V = 2$ oraz $f, g \in V^*$ spełniają $\ker f = \ker g = \text{lin}(v)$, dla pewnego $v \neq 0$. Czy f, g są liniowo zależne?
5. Niech $f, g \in V^*$, gdzie V jest przestrzenią liniową nad K . Uzasadnij, że $\ker f \subseteq \ker g$ wtedy i tylko wtedy, gdy istnieje $\lambda \in K$, że $g = \lambda \cdot f$.
6. Czy dla każdego niezerowego $\alpha = (x_1, x_2, x_3) \in \mathbb{R}^3$ istnieje taki funkcjonał $f : \mathbb{R}^3 \rightarrow \mathbb{R}$, że $f(\alpha) = 1$?
7. Niech $U \neq V$ będzie podprzestrzenią przestrzeni liniowej V . Czy wynika stąd, że istnieje taki niezerowy funkcjonał $f \in V^*$, że $f(u) = 0$, dla każdego $u \in U$?
8. Czy odwzorowanie $F : K[x] \rightarrow K$ dane wzorem

$$F(p(x)) = p''(a) \quad \text{dla pewnego } a \in K$$

jest funkcjonałem liniowym?

9. Niech $U = \text{lin}((1, 0)) \subseteq (\mathbb{R}^2)^*$. Znajdź bazę i wymiar $\text{Ann}(U)$.
 10. Założmy, że V jest przestrzenią liniową skończonego wymiaru oraz U, W są jej podprzestrzeniami. Uzasadnij, że
- $$\text{Ann}(W) \subseteq \text{Ann}(U) \iff U \subseteq W.$$
11. Znajdź bazę sprzężoną do bazy $\mathcal{B} = \{(1, 1), (1, 0)\}$ w $(\mathbb{R}^2)^*$.
 12. Niech $\mathcal{A} = (\alpha, \beta)$ będzie bazą przestrzeni liniowej V . Dla $a, b \neq 0$ rozważmy bazę $\mathcal{B} = (a\alpha, b\beta)$. Jaki jest związek pomiędzy bazami dualnymi \mathcal{A}^* oraz \mathcal{B}^* ?
 13. Znajdź bazę sprzężoną do bazy $1, x - 1$ w przestrzeni $\mathbb{R}[x]_{\leqslant 1}$.
 14. Niech $f, g : \mathbb{C}^2 \rightarrow \mathbb{C}$ będą funkcjonalami zadanimi wzorami

$$f((a, b)) = a + bi \quad \text{oraz} \quad g((a, b)) = a - bi.$$

Wykaż, że układ $\{f, g\}$ stanowi bazę $(\mathbb{C}^2)^*$. Znajdź taką bazę $\alpha_1, \alpha_2 \in \mathbb{C}^2$, że $\{f, g\}$ jest bazą sprzężoną do $\{\alpha_1, \alpha_2\}$.

15. Niech $w^{(k)}$ będzie k -tą pochodną wielomianu w . Uzasadnij, że baza dualna do bazy $1, x, \dots, x^m$ przestrzeni $\mathbb{R}[x]_{\leqslant m}$ złożona jest z funkcjonalów f_1, \dots, f_m danych wzorami $f_k(w) = \frac{w^{(k)}(0)}{k!}$.
16. Niech $\alpha_1, \dots, \alpha_n$ będzie bazą przestrzeni liniowej V nad ciałem K oraz niech $f_1, \dots, f_n \in V^*$ będzie bazą sprzężoną do niej. Rozważmy przekształcenia $\Phi : V \rightarrow K^n$ oraz $\Psi : K^n \rightarrow V$ dane wzorami:

$$\Phi(v) = (f_1(v), \dots, f_n(v)), \quad \Psi((a_1, \dots, a_n)) = a_1 v_1 + \dots + a_n v_n.$$

Uzasadnij, że Φ oraz Ψ są wzajemnie odwrotnymi izomorfizmami przestrzeni liniowych.

17. Uzasadnij, że przekształcenie $e_V : V \rightarrow V^{**}$ określone wzorem $e_V(\alpha) = e_V \alpha$ jest liniowe. Dla $V = \mathbb{R}^2$ znajdź wartości $e_V((1, -1))$ na bazie st^* .
18. Uzasadnij, że jeśli V jest skończonym wymiarową przestrzenią liniową, W jest podprzestrzenią V , zaś $e_V : V \rightarrow V^{**}$ jest izomorfizmem ewaluacji, to

$$e_V(\text{Ann}(W)) = \text{Ann}(e_V(W)).$$

22.3 Zadania do samodzielnej pracy

Zadania ilustrujące umiejętności niezbędne do zaliczenia przedmiotu oznaczone są symbolem ♠.

Przy tych zadaniach dodany jest opis umiejętności, które sprawdzają.

1. (♠) Niech $W = \{(x_1, x_2, x_3, x_4) \mid 2x_1 + x_2 - 4x_3 + 5x_4 = 0\}$. Znaleźć wzór na funkcjonał $f : \mathbb{R}^4 \rightarrow \mathbb{R}$ spełniający: $f(\alpha) = 0$ dla każdego $\alpha \in W$ oraz $f((1, 0, 1, 0)) = 1$.
2. (♠) Znajdź bazę $\{\alpha_1^*, \alpha_2^*, \alpha_3^*\}$ przestrzeni $(\mathbb{R}^3)^*$ sprzężoną do bazy $\{\alpha_1, \alpha_2, \alpha_3\}$ przestrzeni \mathbb{R}^3 , gdzie $\alpha_1 = (1, 2, 2)$, $\alpha_2 = (2, 5, 5)$, $\alpha_3 = (1, 3, 4)$.
3. (♠) Niech $f(x_1, x_2, x_3) = 2x_1 + x_2 - 3x_3$ będzie funkcjonałem na \mathbb{R}^3 . Znajdź współrzędne f w bazie sprzężonej do bazy standardowej \mathbb{R}^3 oraz w bazie sprzężonej do bazy $\{(2, 0, 0), (1, 2, 0), (0, 1, 2)\}$.
4. (♠) W \mathbb{R}^3 dany jest układ wektorów $\mathcal{A} = (\alpha_1, \alpha_2, \alpha_3)$, gdzie

$$\alpha_1 = (1, 1, -1), \alpha_2 = (1, 1, 0), \alpha_3 = (1, 2, 1).$$

Znaleźć współrzędne funkcjonału $f(x_1, x_2, x_3) = 2x_1 - x_2 + 3x_3$ w bazie \mathcal{A}^* sprzężonej do \mathcal{A} . Znaleźć wzór funkcjonału $g \in (\mathbb{R}^3)^*$ mającego w bazie \mathcal{A}^* współrzędne 1, -1, 1 i wzory funkcjonałów f_1, f_2, f_3 takich, że $\mathcal{A}^* = (f_1, f_2, f_3)$.

5. (♠) Znajdź bazę sprzężoną do bazy $1, x - 5, (x - 5)^2, (x - 5)^3$ w przestrzeni $\mathbb{R}[x]_{\leq 3}$.

6. (♠) Sprawdź, że funkcjonały

$$f_1(x_1, x_2, x_3) = 5x_1 - 2x_2, \quad f_2(x_1, x_2, x_3) = -x_1 - x_3, \quad f_3(x_1, x_2, x_3) = -x_1 + x_2 + x_3$$

tworzą bazę przestrzeni $(\mathbb{R}^3)^*$ i znajdź taką bazę $\alpha_1, \alpha_2, \alpha_3$ przestrzeni \mathbb{R}^3 , że $\{f_1, f_2, f_3\}$ jest bazą sprzężoną do tej bazy.

7. Niech $V = M_{2 \times 2}(\mathbb{C})$. Zdefiniujmy funkcjonały $f_1, f_2, f_3, f_4 \in V^*$ wzorami:

$$\begin{aligned} f_1 \left(\begin{bmatrix} x_{11} & x_{12} \\ x_{21} & x_{22} \end{bmatrix} \right) &= x_{11} + x_{12}, & f_2 \left(\begin{bmatrix} x_{11} & x_{12} \\ x_{21} & x_{22} \end{bmatrix} \right) &= -x_{12} + 2x_{22}, \\ f_3 \left(\begin{bmatrix} x_{11} & x_{12} \\ x_{21} & x_{22} \end{bmatrix} \right) &= -x_{22}, & f_4 \left(\begin{bmatrix} x_{11} & x_{12} \\ x_{21} & x_{22} \end{bmatrix} \right) &= x_{21} + 3x_{22}. \end{aligned}$$

Sprawdź, że zbiór $\{f_1, f_2, f_3, f_4\}$ jest bazą przestrzeni V^* , a następnie znajdź taką bazę $\{e_1, e_2, e_3, e_4\}$ przestrzeni V , aby $e_j^* = f_j$, dla $1 \leq j \leq 4$.

8. W przestrzeni liniowej V rozważamy bazę $\mathcal{A} = \{\alpha_1, \alpha_2, \dots\}$ zaś w przestrzeni sprzężonej V^* dualny do niej układ funkcjonałów $\mathcal{A}^* = \{\alpha_1^*, \alpha_2^*, \dots\}$. Udowodnij, że jeśli $\dim(V) = \infty$, to układ \mathcal{A}^* jest liniowo niezależny, ale nie rozpina V^* .

9. Niech $\dim(V) = n$ oraz niech $f_1, f_2, \dots, f_n \in V^*$. Wykaż, że f_1, f_2, \dots, f_n są liniowo niezależne wtedy i tylko wtedy, gdy $\ker(f_1) \cap \ker(f_2) \cap \dots \cap \ker(f_n) = \{0\}$. Uzasadnij też, że dla każdego $m < n$ mamy: $\dim(\ker f_1 \cap \ker f_2 \cap \dots \cap \ker f_m) = n - m$.

10. Opisz $\text{Ann}(V)$ równaniami, gdzie

- a) $V \subseteq \mathbb{R}^4$ jest opisana równaniami $x_1 = x_2 = x_3 = x_4$.
- b) $V = \text{lin}((1, 2, 0, -3), (-2, 3, 2, -3), (-3, 1, 2, 0)) \subseteq \mathbb{R}^4$.

11. Niech $U \subseteq W$ będą podprzestrzeniami skończenie wymiarowej przestrzeni liniowej V , to:

$$\text{Ann}(U + W) = \text{Ann}(U) \cap \text{Ann}(W), \quad \text{Ann}(U \cap W) = \text{Ann}(U) + \text{Ann}(W).$$

12. Niech $f : M_{n \times n}(K) \rightarrow K$ będzie funkcjonałem takim, że $f(AB) = f(BA)$, dla dowolnych macierzy $A, B \in M_{n \times n}(\mathbb{R})$. Wykaż, że istnieje $c \in K$, że $f(A) = c \cdot \text{tr}(A)$.

13. Niech $X = \{(x_n)_{n=1}^\infty \in \mathbb{R}^\mathbb{N} : \sum_{n=1}^\infty x_n^2 < \infty\}$. Wykaż, że:

- (a) gdy $x = (x_n)_{n=1}^\infty \in X$, to funkcja $f_x : X \rightarrow \mathbb{R}$ zadana wzorem $f_x(y) = \sum_{n=1}^\infty x_n y_n$ dla $y = (y_n)_{n=1}^\infty \in X$ jest dobrze określona i liniowa (czyli $f_x \in X^*$).
- (b) odwzorowanie $f : X \rightarrow X^*$ zdefiniowane jako $f(x) = f_x$ dla $x \in X$ jest liniowe oraz injektywne (czyli jest monomorfizmem).

Czy odwzorowanie f jest epimorfizmem? Jeśli nie, to wskaż przykład funkcjonału z X^* , który nie leży w obrazie $\text{im } f$ odwzorowania f .

Rozdział 23

Przekształcenie sprzężone

23.1 Wykład 23

Na ostatnim wykładzie wprowadziliśmy pojęcie przestrzeni sprzężonej i w oparciu o wcześniejsze rezultaty uzasadniliśmy, że w skończenie wymiarowym przypadku przestrzeń V oraz V^* są w zasadzie identyczne – są to przestrzenie izomorficzne. Wprowadziliśmy również pojęcie bazy sprzężonej \mathcal{A}^* do bazy \mathcal{A} przestrzeni V .

Definicja 23.1.1: Przekształcenie sprzężone

Niech $\phi : V \rightarrow W$ będzie przekształceniem liniowym. PRZEKSZTAŁCENIEM SPRZĘŻONYM do ϕ nazywamy przekształcenie $\phi^* : W^* \rightarrow V^*$ określone wzorem

$$\phi^*(g) = g \circ \phi.$$

Innymi słowy jest to takie przekształcenie, które bierze funkcjonał g z W^* i przeprowadza go na funkcjonał $\phi^*(g) : V \rightarrow K$ tak, że następujący diagram jest przemienny:

$$\begin{array}{ccc} V & \xrightarrow{\phi^*(g)} & K \\ \phi \searrow & & \swarrow g \\ & W & \end{array}$$

Pozostawiamy Czytelnikowi nietrudne sprawdzenie, że ϕ^* w istocie jest przekształceniem liniowym. Z definicji łatwo sprawdzić, że przekształcenie sprzężone do zerowego jest zerowe oraz $(\text{id}_V)^* = \text{id}_{V^*}$.

Przykład przekształcenia sprzężonego. Rozważmy $\psi : \mathbb{R}^3 \rightarrow \mathbb{R}^2$ dane wzorem:

$$\psi((x_1, x_2, x_3)) = (2x_1 + 3x_2 + x_3, 5x_1 - x_2 - 2x_3).$$

Wówczas dla funkcjonału $g : \mathbb{R}^2 \rightarrow \mathbb{R}$ zadanego wzorem:

$$g((y_1, y_2)) = 3y_1 - 2y_2$$

funkcjonał $\psi^*(g) : \mathbb{R}^3 \rightarrow \mathbb{R}$ jest zadany wzorem:

$$\begin{aligned} \psi^*(g)((x_1, x_2, x_3)) &= (g \circ \psi)((x_1, x_2, x_3)) = g(\psi((x_1, x_2, x_3))) = \\ &= g((2x_1 + 3x_2 + x_3, 5x_1 - x_2 - 2x_3)) = \\ &= 3(2x_1 + 3x_2 + x_3) - 2(5x_1 - x_2 - 2x_3) = \\ &= \color{blue}{-4x_1 + 11x_2 + 7x_3}. \end{aligned}$$

A jak wygląda wzór przekształcenia ψ^* ? Jak je zapisać? Przypomnijmy, że bazą sprzężoną do bazy standardowej $\epsilon_1, \dots, \epsilon_n$ przestrzeni K^n jest baza złożona z funkcjonałów postaci $\epsilon_i^*((x_1, \dots, x_n)) = x_i$, dla $1 \leq i \leq n$. Bazę tę oznaczać będziemy przez st^* . Aby zapisać wzór ϕ^* , dla dowolnych $y_1, y_2 \in K$ musimy mieć takie (zależne od nich) a_1, a_1, a_3 , aby $\psi^*(\color{red}{y_1}\epsilon_1^* + \color{red}{y_2}\epsilon_2^*) = \color{blue}{a_1}\epsilon_1^* + \color{blue}{a_2}\epsilon_2^* + \color{blue}{a_3}\epsilon_3^*$.

Twierdzenie 23.1.2

Niech $\psi : V \rightarrow W$ będzie przekształceniem liniowym, $\mathcal{A} = \{\alpha_1, \dots, \alpha_n\}$ będzie bazą przestrzeni V i niech $\mathcal{B} = \{\beta_1, \dots, \beta_m\}$ będzie bazą przestrzeni W . Wówczas dla przekształcenia sprzężonego $\psi^* : W^* \rightarrow V^*$ oraz baz sprzężonych $\mathcal{A}^*, \mathcal{B}^*$ mamy

$$M(\psi^*)_{\mathcal{B}^*}^{\mathcal{A}^*} = (M(\psi)_{\mathcal{A}}^{\mathcal{B}})^T.$$

Dowód. Oznaczmy $\mathcal{B}^* = \{\phi_1, \dots, \phi_m\}$ oraz niech $M(\psi)_{\mathcal{A}}^{\mathcal{B}} = [a_{ij}]$, czyli dla każdego $j = 1, \dots, n$ mamy

$$\psi(\alpha_j) = \sum_{i=1}^m a_{ij} \beta_i.$$

W j -tej kolumnie macierzy $M(\psi^*)_{\mathcal{B}^*}^{\mathcal{A}^*}$ stoją współrzędne funkcjonału $\psi^*(\phi_j)$ w bazie \mathcal{A}^* , czyli (na mocy Uwagi 22.1.3(b)) wartości tego funkcjonału na kolejnych wektorach bazy \mathcal{A} . Dostajemy:

$$\begin{aligned} (\psi^*(\phi_j))(\alpha_1) &= (\phi_j \circ \psi)(\alpha_1) = \phi_j(\sum_{i=1}^m a_{i1} \beta_i) = a_{j1} \\ (\psi^*(\phi_j))(\alpha_2) &= (\phi_j \circ \psi)(\alpha_2) = \phi_j(\sum_{i=1}^m a_{i2} \beta_i) = a_{j2} \\ \vdots &\quad \vdots \quad \vdots \quad \vdots \\ (\psi^*(\phi_j))(\alpha_n) &= (\phi_j \circ \psi)(\alpha_n) = \phi_j(\sum_{i=1}^m a_{in} \beta_i) = a_{jn} \end{aligned}$$

Zatem j -ta kolumna macierzy $M(\psi^*)_{\mathcal{B}^*}^{\mathcal{A}^*}$ jest równa j -temu wierszowi macierzy $M(\psi)_{\mathcal{A}}^{\mathcal{B}}$. Stąd mamy tezę. \square

Twierdzenie wyżej będzie dla nas miało duże znaczenie w kontekście badania macierzy symetrycznych i tzw. przekształceń samosprzężonych, ale możemy też o nim myśleć jako o narzędziu pozwalającym na przeformułowanie pewnych znanych nam już faktów. Czytelnik bez trudu sprawdzi (zostawiamy to jako zadanie w zestawie niżej), że jeśli $\phi : V \rightarrow W$ oraz $\psi : W \rightarrow Z$, to przekształcenie sprzężone do przekształcenia $\psi \circ \phi : V \rightarrow Z$ spełnia następującą zależność:

$$(\psi \circ \phi)^* = \phi^* \circ \psi^*.$$

Formuła ta, widziana w języku Twierdzenia 23.1.2 jest przypomnieniem, że transpozycja iloczynu macierzy jest iloczynem transpozycji w odwrotnej kolejności, czyli $(BA)^T = A^T B^T$.

Podobnie, bez trudu wykazać można, że jeśli $\phi : V \rightarrow W$ jest izomorfizmem, to również $\phi^* : W^* \rightarrow V^*$ jest izomorfizmem oraz:

$$(\phi^*)^{-1} = (\phi^{-1})^*,$$

co na poziomie macierzy bezpośrednio nawiązuje do formuły $(A^T)^{-1} = (A^{-1})^T$.

Na podstawie tych obserwacji sformułować możemy wniosek, który również zostawimy jako ważne ćwiczenie: jeśli $\phi : V \rightarrow W$ jest przekształceniem liniowym przestrzeni liniowych V, W wymiarów odpowiednio n i m , to przyporządkowanie $\phi \mapsto \phi^*$ zadaje izomorfizm $*$ pomiędzy $L(V, W)$ oraz $L(W^*, V^*)$. Jak?

Jeśli wybierzemy bazy \mathcal{A}, \mathcal{B} przestrzeni V oraz W , to na mocy Wniosku 19.1.6 mamy następujące izomorfizmy: $\Psi_{\mathcal{A}, \mathcal{B}} : L(V, W) \rightarrow M_{m \times n}(K)$ oraz $\Psi_{\mathcal{B}^*, \mathcal{A}^*} : L(W^*, V^*) \rightarrow M_{n \times m}(K)$. Okazuje się, że jeśli $T : M_{m \times n}(K) \rightarrow M_{n \times m}(K)$ jest izomorfizmem polegającym na braniu transpozycji, to złożenie

$$* = \Psi_{\mathcal{B}^*, \mathcal{A}^*} \circ T \circ \Psi_{\mathcal{A}, \mathcal{B}}$$

jest szukanym izomorfizmem, czyli dla dowolnych baz \mathcal{A}, \mathcal{B} poniższy diagram jest przemienny

$$\begin{array}{ccc} L(V, W) & \xrightarrow{*} & L(W^*, V^*) \\ \downarrow \Psi_{\mathcal{A}, \mathcal{B}} & & \downarrow \Psi_{\mathcal{B}^*, \mathcal{A}^*} \\ M_{m \times n}(K) & \xrightarrow{T} & M_{n \times m}(K) \end{array}$$

Czytelnika może irytować fakt, że przemienność tego naturalnego diagramu wymaga wyboru baz przestrzeni liniowych V, W . Problem ten zniką przy drugim braniu sprzężenia. W jaki sposób to uzasadnimy?

Wiemy, że dla dowolnej macierzy $(A^T)^T = A$. W Twierdzeniu 22.1.6 wykazaliśmy, że dla przestrzeni skończenie wymiarowej V istnieje izomorfizm ewaluacji $e_V : V \rightarrow V^{**}$, przypisujący wektorowi $\alpha \in V$ funkcjonał $ev_\alpha : V^* \rightarrow K$, którego wartość na funkcjonale $f : V \rightarrow K$ równa jest $f(\alpha)$. Posłużymy się tym izomorfizmem do uzasadnienia następującego rezultatu, utożsamiającegogo ϕ z ϕ^{**} .

Twierdzenie 23.1.3

Jeśli V, W są skończenie wymiarowymi przestrzeniami liniowymi oraz $\phi \in L(V, W)$, wówczas dla każdego $\alpha \in V$ mamy

$$\phi^{**}(e_V(\alpha)) = e_W(\phi(\alpha)),$$

czyli następujący diagram przekształceń jest przemienny

$$\begin{array}{ccc} V & \xrightarrow{\phi} & W \\ e_V \downarrow & & \downarrow e_W \\ V^{**} & \xrightarrow{\phi^{**}} & W^{**} \end{array}$$

W języku teorii kategorii, który kiedyś Państwo poznają, okaże się powyższy diagram uzasadnia, że rodzina $\{e_V\}_{V \in Ob(Vect)}$ jest tzw. *naturalną transformacją* między funktorem identyczności, a funktorem brania drugiego sprzężenia. Biorąc pod uwagę, że ograniczamy się do przestrzeni skończonego wymiaru jest to nawet równoważność naturalna. Jest ważne, gdyż był to przykład, od którego Eilenberg i MacLane rozpoczęli wyjaśnianie teorii kategorii w artykule „General theory of equivalences” z roku 1945.

Dowód. Weźmy $\alpha \in V$. Oczywiście mamy:

$$\phi^{**}(e_V(\alpha)) = (\phi^*)^*(e_V(\alpha)) = ev_\alpha \circ \phi^*.$$

Idąc natomiast w drugą stronę diagramu mamy

$$e_W(\phi(\alpha)) = ev_{\phi(\alpha)}.$$

Należy więc wykazać, że funkcjonały $ev_\alpha \circ \phi^*$ oraz $ev_{\phi(\alpha)}$ są równe jako elementy W^{**} . Weźmy więc dowolny $g \in W^*$. Wówczas:

$$(ev_\alpha \circ \phi^*)(g) = ev_\alpha(\phi^*(g)) = ev_\alpha(g \circ \phi) = (g \circ \phi)(\alpha) = g(\phi(\alpha)) = ev_{\phi(\alpha)}(g).$$

□

Dla zarysowania pewnych podstaw omawianej teorii, pozostało rozstrzygnąć jeszcze jedno naturalne pytanie: co sprząganie robi z monomorfizmami i epimorfizmami? W dowodzie rezultatu odpowiadającego na to pytanie skorzystamy z następującej naturalnej obserwacji.

Uwaga 23.1.4

Dla każdego niezerowego wektora $\alpha \in V$ istnieje funkcjonal liniowy $\phi : V \rightarrow K$ taki, że $\phi(\alpha) = 1$.

Dowód. Dopełniamy wektor α do bazy przestrzeni V i definiujemy funkcjonal ϕ zadając go na otrzymanej bazie następująco: 1 na α , 0 na pozostałych wektorach bazy. □

Warto odnotować, że choć powyższy fakt wygląda bardzo naturalnie, w istocie korzystamy tu z twierdzenia mówiącego, że dowolna przestrzeń liniowa ma bazę — nie tylko skończenie wymiarowa. Korzystamy też z uogólnienia twierdzenia o jednoznaczny definiowaniu na bazie na przypadek dowolnych — niekiedy skończenie wymiarowych przestrzeni liniowych.

Jesteśmy gotowi do sformułowania ostatniego istotnego dla nas rezultatu dotyczącego przestrzeń sprzężonych. Uzasadnimy je osobno w przypadku przekształceń pomiędzy przestrzeniami skończonymi wymiaru, i osobno w przypadku ogólnym, korzystając z obserwacji wyżej.

Twierdzenie 23.1.5

Niech $\psi : V \rightarrow W$ będzie przekształceniem liniowym. Wówczas

- a) ψ jest monomorfizmem $\iff \psi^*$ jest epimorfizmem.
- b) ψ jest epimorfizmem $\iff \psi^*$ jest monomorfizmem.

Dowód. Przypadek skończenie wymiarowy.

Niech $\mathcal{A} = \{\alpha_1, \dots, \alpha_n\}$ będzie bazą V i niech $\mathcal{B} = \{\beta_1, \dots, \beta_m\}$ będzie bazą W . Wiemy już, że $r(\phi) = r(\phi^*)$, a zatem na mocy twierdzenia o sumie wymiarów jądra i obrazu przekształcenia mamy równoważności:

$$\begin{aligned}\phi \text{ jest monomorfizmem} &\iff r(\phi) = \dim V \iff \\ &\iff r(\phi^*) = \dim V^* \iff \\ &\iff \phi^* \text{ jest epimorfizmem.}\end{aligned}$$

oraz równoważności:

$$\begin{aligned}\phi \text{ jest epimorfizmem} &\iff r(\phi) = \dim W \iff \\ &\iff r(\phi^*) = \dim W^* \iff \\ &\iff \phi^* \text{ jest monomorfizmem.}\end{aligned}$$

Przypadek ogólny.

a) Założmy, że ψ jest monomorfizmem. Pokażemy, że ψ^* jest epimorfizmem. Niech $\phi \in V^*$ będzie dowolnym funkcjonałem. Znajdziemy funkcjonał $\phi_1 \in W^*$ spełniający warunek $\phi = \psi^*(\phi_1)$.

Niech $\{\alpha_t\}_{t \in T}$ będzie bazą przestrzeni V . Wtedy $\{\psi(\alpha_t)\}_{t \in T}$ jest układem liniowo niezależnym w W , bo ψ jest monomorfizmem. Uzupełniamy ten układ do bazy \mathcal{B} przestrzeni W . Definiujemy funkcjonał ϕ_1 przez zadanie go na bazie \mathcal{B} następująco: $\phi_1(\psi(\alpha_t)) = \phi(\alpha_t)$ dla każdego $t \in T$ oraz $\phi_1(\beta) = 0$ dla tych $\beta \in \mathcal{B}$, które nie należą do $\{\psi(\alpha_t)\}_{t \in T}$. Otrzymujemy $\phi = \psi^*(\phi_1)$.

Założmy, że ψ^* jest epimorfizmem. Pokażemy, że ψ jest monomorfizmem. Niech α będzie dowolnym wektorem w $\ker(\psi)$. Przypuśćmy że $\alpha \neq 0$. Niech $\phi \in V^*$ będzie dowolnym funkcjonałem spełniającym $\phi(\alpha) = 1$. Taki funkcjonał istnieje na mocy Uwagi 23.1.4. Zauważmy, że dla każdego funkcjonału $\phi_1 \in W^*$ funkcjonał $\psi^*(\phi_1)$ zeruje się na wektorze α , bo

$$(\psi^*(\phi_1))(\alpha) = (\phi_1 \circ \psi)(\alpha) = \phi_1(\psi(\alpha)) = \phi_1(0) = 0.$$

Stąd $\phi \notin \text{im}(\psi^*)$, co przeczy temu, że ψ^* jest epimorfizmem. Stąd musi zachodzić $\alpha = 0$, a więc $\ker(\psi) = \{0\}$, czyli ψ jest monomorfizmem.

b) Założmy, że ψ jest epimorfizmem. Pokażemy, że ψ^* jest monomorfizmem. Niech $\phi \in \ker(\psi^*)$. Oznacza to, że $\phi \circ \psi(\alpha) = 0$ dla każdego wektora $\alpha \in V$. Skoro jednak ψ jest epimorfizmem, więc każdy wektor $\beta \in W$ jest postaci $\beta = \psi(\alpha)$ dla pewnego $\alpha \in V$. Otrzymujemy więc, że

$$\phi(\beta) = \phi(\psi(\alpha)) = \phi \circ \psi(\alpha) = 0,$$

dla każdego $\beta \in W$, czyli ϕ jest funkcjonałem zerowym. Wykaźaliśmy więc, że $\ker(\psi^*) = 0$, czyli ψ^* jest monomorfizmem.

Założmy, że ψ^* jest monomorfizmem. Pokażemy, że ψ jest epimorfizmem. Przypuśćmy, że tak nie jest, to znaczy istnieje wektor $\beta \in W$, który nie należy do $\text{im}(\psi)$. Niech $\{\beta_t\}_{t \in T}$ będzie bazą przestrzeni $\text{im}(\psi)$. Wtedy układ $\{\beta\} \cup \{\beta_t\}_{t \in T}$ jest liniowo niezależny. Dopełniamy go do bazy \mathcal{B} przestrzeni W . Definiujemy funkcjonał $\phi \in W^*$ przez zadanie jego wartości na bazie \mathcal{B} następująco: 1 na wektorze β oraz 0 na pozostałych wektorach bazy \mathcal{B} . Tak otrzymany funkcjonał ϕ jest niezerowy, ale $\psi^*(\phi) = \phi \circ \psi$ jest funkcjonałem zerowym. Przeczy to założeniu, że ψ^* jest monomorfizmem. Zatem $W = \text{im}(\psi)$, czyli ψ jest epimorfizmem. \square

23.2 Wybór przykładowych pytań

Proszę się nie uczyć na pamięć odpowiedzi czy wyjaśnienie! Takich pytań można ułożyć znacznie więcej i służą one jedynie sprawdzeniu podstawowego zrozumienia treści zajęć. Jeśli odpowiedź na dowolne z poniższych pytań sprawia Państwu problem, proszę pomyśleć który fragment notatek dotyczy tego pytania, wrócić do ich lektury, a w razie dalszych trudności poprosić prowadzącego o wskazówkę.

- Niech $\phi : \mathbb{R}^2 \rightarrow \mathbb{R}^3$ dane będzie wzorem

$$\phi((x_1, x_2)) = (x_1 + x_2, x_1, x_2).$$

- Znajdź wzory funkcjonalów $\phi^*(\epsilon_1^*), \phi^*(\epsilon_2^*), \phi^*(\epsilon_3^*)$, gdzie $\text{st}^* = \{\epsilon_1^*, \epsilon_2^*, \epsilon_3^*\}$.
- Rozważmy funkcjonał $f : \mathbb{R}^3 \rightarrow \mathbb{R}$ dany wzorem $f((x_1, x_2, x_3)) = x_1 - x_2 + x_3$. Co ma wspólnego wyznaczenie wzoru $\phi^*(f)$ z mnożeniem macierzy postaci:

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

- Przekształcenie liniowe $\phi : \mathbb{R}^3 \rightarrow \mathbb{R}^2$ dane jest wzorem

$$\phi((x_1, x_2, x_3)) = (ax_1 + bx_2 + cx_3, dx_1 + ex_2 + fx_3)$$

oraz $g : \mathbb{R}^2 \rightarrow \mathbb{R}$ dane jest wzorem $g((y_1, y_2)) = y_1 + y_2$. Znajdź $\phi^*(g)$.

- Niech V będzie przestrzenią liniową nad \mathbb{R} . Niech $\mathcal{A} = (\alpha, \beta), \mathcal{B} = (\alpha + \beta, \alpha - \beta)$. będą bazami V . Znajdź macierze

$$M(\text{id}_V)_{\mathcal{B}}^{\mathcal{A}}, \quad M(\text{id}_{V^*})_{\mathcal{A}^*}^{\mathcal{B}^*}.$$

- Niech $\mathcal{A} = (\alpha, \beta)$ oraz $\mathcal{B} = (\alpha, \alpha + \beta)$ będą bazami przestrzeni liniowej V . Znajdź macierze przekształcenia id_{V^*} postaci

$$M(\text{id}_{V^*})_{\mathcal{A}^*}^{\mathcal{B}^*}, \quad M(\text{id}_{V^*})_{\mathcal{B}^*}^{\mathcal{A}^*}$$

- Uzasadnij, w oparciu o definicję, że przekształcenie sprzężone do przekształcenia liniowego jest przekształceniem liniowym.

- Uzasadnij, że przekształcenie sprzężone do id_V jest równe id_{V^*} .

- Niech W będzie przestrzenią skończenie wymiarową oraz niech $\phi \in L(V, W)$. Uzasadnij, że

$$\phi^* = 0 \iff \phi = 0.$$

- Uzasadnij, że przekształcenie $\Psi : L(V, W) \rightarrow L(W^*, V^*)$ dane wzorem $\Psi(\phi) = \phi^*$ jest liniowe.

- Uzasadnij, że jeśli $\phi \in L(V, W)$ oraz $\psi : L(W, Z)$, to

$$(\psi \circ \phi)^* = \phi^* \circ \psi^*.$$

Uzasadnij również, że jeśli ϕ jest izomorfizmem, to

$$(\phi^*)^{-1} = (\phi^{-1})^*.$$

- Niech $\phi \in L(V, W)$ będzie izomorfizmem. Rozstrzygnij, czy również ϕ^* jest izomorfizmem?
- Niech L będzie podprzestrzenią przestrzeni liniowej V . Rozważmy przekształcenie liniowe $\phi : V \rightarrow W$. Uzasadnij, że $L \subseteq \ker \phi$ wtedy i tylko wtedy, gdy $\text{im } \phi^* \subseteq \text{Ann}(L)$.
- Niech $\phi : V \rightarrow W$ oraz $\psi : W \rightarrow Z$ będą takimi przekształceniami liniowymi, że $\phi \circ \psi$ jest izomorfizmem. Rozstrzygnij, czy $\psi^* \circ \phi^*$ jest izomorfizmem?
- Niech U będzie podprzestrzenią przestrzeni liniowej V . Niech $i : U \rightarrow V$ będzie inkluzją określona wzorem $i(u) = i$.
 - Wykaż, że $\ker i^* = \text{Ann}(U)$.
 - Wykaż, że jeśli V jest skończenie wymiarowa, to $\text{im } i^* = U^*$.
 - Wskaż izomorfizm przestrzeni liniowych $V^*/\text{Ann}(U)$ oraz U^* .

23.3 Zadania do samodzielnej pracy

Zadania ilustrujące umiejętności niezbędne do zaliczenia przedmiotu oznaczone są symbolem ♠. Przy tych zadaniach dodany jest opis umiejętności, które sprawdzają.

- Niech $\mathcal{A} = (\alpha_1, \dots, \alpha_n)$ będzie bazą przestrzeni liniowej V nad ciałem K . Niech $\mathcal{A}^* = (\alpha_1^*, \dots, \alpha_n^*)$ będzie bazą sprzężoną do \mathcal{A} w V^* . Niech \mathcal{B} powstaje z \mathcal{A} przez:

- (a) zamianę i -tego i j -tego wektora miejscami,
- (b) zamianę wektora α_i na $\alpha_i + k\alpha_j$, dla pewnego $i \neq j$ oraz $k \in K \neq 0$.

Opisz w każdym z tych dwóch przypadków postać macierzy $M(\text{id})_{\mathcal{B}^*}^{\mathcal{A}^*}$.

- (♣) Dane są dwie bazy $\mathcal{A} = \{(1, 1, 1), (1, 1, 2), (2, 1, 1)\}$, $\mathcal{B} = \{(2, 1, 1), (2, 1, 2), (2, 2, 2)\}$ przestrzeni \mathbb{R}^3 , a także przekształcenie liniowe $\phi : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ takie, że

$$M(\phi)_{\mathcal{A}}^{\mathcal{B}} = \begin{bmatrix} 1 & 1 & 2 \\ 2 & 1 & 1 \\ 3 & 1 & 0 \end{bmatrix}$$

i funkcjonał $f \in (\mathbb{R}^3)^*$ mający w bazie \mathcal{A}^* współrzędne $-1, 2, 1$. Znaleź wzór na funkcjonał $\phi^*(f)$.

- (♣) Niech $V = \mathbb{R}^3$ oraz

$$A = \{(2, 1, 4), (1, 1, 1), (3, 2, 1)\}, \quad B = \{(1, 2, 3), (1, 1, 2), (3, 5, 3)\}.$$

Rozważmy przekształcenie liniowe $\phi : V \rightarrow V$ zadane warunkiem

$$M_{\mathcal{A}}^{\mathcal{B}}(\phi) = \begin{bmatrix} 1 & 2 & 3 \\ 1 & 1 & 2 \\ 1 & 4 & 5 \end{bmatrix}.$$

Dla jakich $t \in \mathbb{R}$ funkcjonał $f \in V^*$, dany wzorem $f(x, y, z) = 2x - y + tz$, należy do $\ker \phi^*$?

- (♣) Znaleźć bazy przestrzeni $\ker(\psi^*)$ oraz $\text{im}(\psi^*)$, gdzie

- $\psi : \mathbb{R}^2 \rightarrow \mathbb{R}^3$ dane jest wzorem

$$\psi((x_1, x_2)) = (x_1 + 2x_2, 3x_1 + 6x_2, 2x_1 + 4x_2).$$

- $\psi : \mathbb{R}^4 \rightarrow \mathbb{R}^3$ dane jest wzorem

$$\psi((x_1, x_2, x_3, x_4)) = (x_1 + 2x_2 + 3x_3 + 4x_4, 2x_1 + 5x_2 + 8x_3 + 9x_4, x_1 + 3x_2 + 5x_3 + 5x_4),$$

- $\psi : \mathbb{R}^3 \rightarrow \mathbb{R}^4$ dane jest wzorem

$$\psi((x_1, x_2, x_3)) = x_1 + 2x_2, x_1 + 3x_2 + 2x_3, x_2 + 2x_3, x_1 + x_2 - 2x_3).$$

- (♣) Przekształcenie liniowe $\psi : \mathbb{R}^2 \rightarrow \mathbb{R}^3$ dane jest wzorem

$$\psi((x_1, x_2)) = (x_1 + 3x_2, 2x_1 + tx_2, -2x_1 - 6x_2).$$

Dla jakich $t \in \mathbb{R}$ przekształcenie ψ^* jest epimorfizmem?

- Niech $\phi : V \rightarrow V$ będzie rzutem. Wykaż, że $\pi^* : V^* \rightarrow V^*$ również jest rzutem. Uzasadnij również, że

$$\text{im } \pi^* = \text{Ann}(\ker \pi), \quad \text{im } \pi = \text{Ann}(\ker \pi^*).$$

- Niech U będzie podprzestrzenią V i niech $\pi : V \rightarrow V/U$ będzie naturalnym rzutowaniem danym wzorem $\pi(v) = v + U$.

- Wykaż, że π^* jest monomorfizmem.
- Wykaż, że $\text{im } \pi^* = \text{Ann}(U)$.
- Wykaż, że $\pi^* : (V/U)^* \rightarrow \text{Ann}(U)$ jest izomorfizmem odwrotnym do przekształcenia Ψ określonego w dowodzie Twierdzenia 22.1.8.

- Uzasadnij, że przekształcenie $\Psi : L(V, W) \rightarrow L(W^*, V^*)$ dane wzorem $\Psi(\phi) = \phi^*$ jest monomorfizmem. Wykaż, że jeśli V, W są przestrzeniami skończonymi wymiarowymi, to Ψ jest również izomorfizmem. Podaj przykład wskazujący, że założenie o skończonym wymiarze jest potrzebne.

Rozdział 24

Wyznacznik — rozwinięcie Laplace'a

24.1 Wykład 24

Zakończyliśmy podstawową część wykładu dotyczącą przestrzeni i przekształceń liniowych. Kolejne dwa wykłady poświęcimy najbardziej zapewne znanemu (z tych niebanalnych) pojęciu algebraicznemu — wyznacznikowi. Czytelnik mógł o nim słyszeć w kontekście rozwiązywania układów równań. Pojęcie wyznacznika można określić i badać w sposób czysto algebraiczny, co jest naszym celem na ten wykład. Niektóre istotne motywacje wygodnie jest wysławiać także w języku geometrii, o czym powiemy później.

Definicja 24.1.1: Macierze kwadratowe

Dla każdego całkowitego $n \geq 1$ zbiór $M_{n \times n}(K)$ macierzy o n wierszach i n kolumnach nazywamy zbiorem MACIERZY KWADRATOWYCH ROZMIARU n i oznaczamy przez $M_n(K)$.

Z punktu widzenia przekształceń liniowych macierze kwadratowe są macierzami przekształceń liniowych pomiędzy przestrzeniami tego samego (skończonego) wymiaru. W szczególności, są to macierze przekształceń liniowych przestrzeni liniowej do niej samej, którymi zajmiemy się w drugim semestrze.

Definicja 24.1.2

Niech $A = [a_{ij}] \in M_n(K)$, gdzie $n > 1$. Dla każdej pary $1 \leq i, j \leq n$ określamy macierz postaci

$$A_{ij} \in M_{n-1}(K)$$

otrzymaną z macierzy A przez skreślenie odpowiednio i -tego wiersza i j -tej kolumny.

Przykład: dla macierzy $A \in M_3(\mathbb{R})$ postaci

$$A = \begin{bmatrix} 1 & 2 & 0 \\ 0 & 1 & 3 \\ 1 & 0 & 5 \end{bmatrix}$$

mamy:

$$A_{11} = \begin{bmatrix} 1 & 3 \\ 0 & 5 \end{bmatrix}, \quad A_{23} = \begin{bmatrix} 1 & 2 \\ 1 & 0 \end{bmatrix}, \quad A_{31} = \begin{bmatrix} 2 & 0 \\ 1 & 3 \end{bmatrix}.$$

Intuicja jest następująca: chcemy określić funkcję $\det : M_n(K) \rightarrow K$, przypominającą „funkcję objętości”. Objętość często liczymy według formuł postaci: „długość/pole podstawy razy wysokość”. Ogólnie objętość obiektu w przestrzeni n wymiarowej wyznaczać chcemy poprzez znajomość objętości $n - 1$ oraz 1-wymiarowej. Okazuje się, że aby określić wyznacznik macierzy $A \in M_n(K)$ potrzebna jest znajomość:

- wyrazów macierzy w pierwszej kolumnie: $a_{11}, a_{21}, \dots, a_{n1}$,
- wyznaczników macierzy rozmiaru $n - 1$ postaci $A_{11}, A_{21}, \dots, A_{n1}$.

Definicja 24.1.3: Wyznacznik — rozwinięcie Laplace'a względem pierwszej kolumny

Definiujemy funkcję $\det : M_n(K) \rightarrow K$ w sposób rekurencyjny

- Dla $n = 1$ kładziemy $\det : M_1(K) \rightarrow K$, gdzie $\det(A) = a$, dla $A = [a]$.
- Dla $A = [a_{ij}] \in M_n(K)$ określamy $\det : M_n(K) \rightarrow K$ znając $\det : M_{n-1}(K) \rightarrow K$ wzorem:

$$\det A = a_{11} \det A_{11} - a_{21} \det A_{21} + a_{31} \det A_{31} + \dots + (-1)^{n+1} a_{n1} \det A_{n1} = \sum_{i=1}^n (-1)^{i+1} a_{i1} \det A_{i1}.$$

Funkcję $\det : M_n(K) \rightarrow K$ nazywamy **WYZNACZNIKIEM**. Czasem zamiast pisać $\det A$, piszemy $|A|$.

Od razu warto dodać pewne doprecyzowanie. W zasadzie definiujemy ciąg funkcji – formalnie należałoby być może pisać (ale nikt tego nie robi) $\det_n : M_n(K) \rightarrow K$. Przy takiej konwencji mielibyśmy

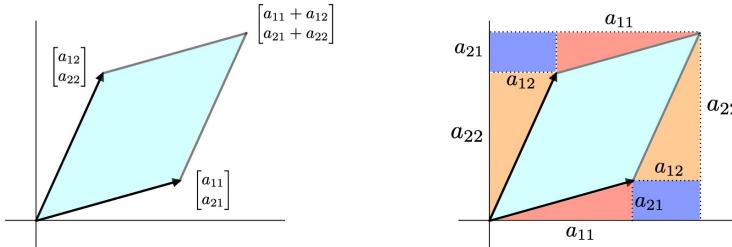
$$\det_n A = \sum_{i=1}^n (-1)^{i+1} a_{i1} \det_{n-1} A_{i1}.$$

Zaczniemy od kilku przykładów dla małych n .

- Dla $A = \begin{bmatrix} 1 & 4 \\ 3 & 2 \end{bmatrix} \in M_2(\mathbb{R})$ mamy $A_{11} = [2], A_{21} = [4]$, czyli:

$$\det A = (-1)^{1+1} \cdot 1 \cdot \det A_{11} + (-1)^{2+1} \cdot 3 \cdot \det A_{21} = 1 \cdot 2 - 3 \cdot 4 = -10.$$

Ogólnie dla macierzy $A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$ mamy $|A| = a_{11}a_{22} - a_{12}a_{21}$. Aby było dla Czytelnika już w tym momencie jasne, że formula ta ma rzeczywiście coś wspólnego z polem czy ogólnie rzecz biorąc „ n -wymiarową objętością”, spójrzmy na poniższy rysunek.



Jeśli o kolumnach macierzy A pomyślimy jak o wektorach na płaszczyźnie \mathbb{R}^2 , które rozpinają pewien równoległobok, to prosty rachunek pokazuje, że wartość bezwzględna wyznacznika tej macierzy równa jest polu równoległoboku. Analogicznie wartość bezwzględna wyznacznika macierzy 1×1 można interpretować jako długość odcinka, a wartość bezwzględna wyznacznika macierzy 3×3 — jako objętość pewnego równoległościanu. Dodajemy wartość bezwzględną, gdyż wyznacznik może być ujemny. Można myśleć o wyznaczniku jako o „objętości zorientowanej”.

Nie będziemy na razie dokładnie wyjaśniać związku wyznacznika z objętością (zrobimy to precyzyjnie w drugim semestrze, dysponując pojęciem prostopadłości), choć do interpretacji tej wróćmy jeszcze kilkukrotnie. Chodzi raczej o ilustrację tego, że z pozoru skomplikowane formuły na wyznacznik mają pochodzenie geometryczne oraz o podkreślenie, że tak jak do policzenia objętości równoległościanu przydatna jest znajomość pola równoległoboku, tak do policzenia wyznacznika macierzy 3×3 przydatna jest znajomość wyznaczników macierzy 2×2 . Warto samodzielnie zastanawiać się jak interpretować geometrycznie własności wyznacznika.

- Dla $A = \begin{bmatrix} 4 & 0 & 0 \\ 0 & 2 & -1 \\ 0 & 5 & 3 \end{bmatrix} \in M_3(\mathbb{R})$ mamy $a_{11} = 4, a_{21} = 0, a_{31} = 0, |A_{11}| = 11, |A_{21}| = 0, |A_{31}| = 0$, czyli

$$\det A = 4 \cdot 11 - 0 \cdot 0 + 0 \cdot 0 = 44.$$

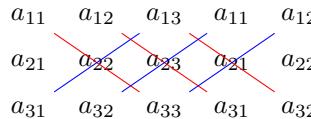
Ogólnie dla macierzy

$$A = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}$$

mamy:

$$\begin{aligned} |A| &= (-1)^{1+1} a_{11} \begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix} + (-1)^{2+1} a_{21} \begin{vmatrix} a_{12} & a_{13} \\ a_{32} & a_{33} \end{vmatrix} + (-1)^{3+1} a_{31} \begin{vmatrix} a_{12} & a_{13} \\ a_{22} & a_{23} \end{vmatrix} = \\ &= a_{11}a_{22}a_{33} - a_{11}a_{23}a_{32} - a_{12}a_{21}a_{33} + a_{13}a_{21}a_{32} + a_{12}a_{23}a_{31} - a_{13}a_{22}a_{31}. \end{aligned}$$

Formułę tą można uzyskać korzystając z tzw. metody Sarrusa, polegającej na wypisaniu obok macierzy A dwóch pierwszych jej kolumn. Wówczas trzy składniki powyższej sumy występujące ze znakiem + uzyskujemy przez iloczyn wyrazów połączonych na czerwono (rys. niżej), a trzy składniki ze znakiem - uzyskujemy przez wymnożenie wyrazów połączonych na niebiesko.



Wyróżnimy teraz macierze, których wyznacznik ma szczególnie elegancką postać.

Definicja 24.1.4: Macierze trójkątne

Niech $A = [a_{ij}] \in M_n(K)$. Zbiór wyrazów $\{a_{11}, \dots, a_{nn}\}$ nazywamy PRZEKĄTNĄ macierzy A . Powiemy, że macierz A jest:

- GÓRNOTRÓJKĄTNA, jeśli $a_{ij} = 0$, dla $i > j$
(pod przekątną macierzy A stoją wyrazy zerowe),
- DOLNOTRÓJKĄTNA, jeśli $a_{ij} = 0$, dla $j > i$
(czyli nad przekątną macierzy A stoją wyrazy zerowe).

Przykłady macierzy odpowiednio górnoprójkątnej i dolnotrójkątnej w $M_2(K)$:

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

Uwaga 24.1.5

Jeśli $A \in M_n(K)$ jest macierzą górnoprójkątną (na przykład: macierzą w postaci schodkowej), to jej wyznacznik równy jest iloczynowi wyrazów na przekątnej.

Dowód. Dowodzimy tezę przez indukcję. Dla $n = 1$ wynika ona wprost z definicji. Weźmy zatem macierz $A = [a_{ij}]$ rozmiaru $n \times n$ oraz zauważmy, że $a_{21} = \dots = a_{n1} = 0$, a zatem z definicji wyznacznika mamy

$$\det A = (-1)^{1+1} \cdot a_{11} \cdot \det A_{11}.$$

Macierz A_{11} powstaje z A przez usunięcie pierwszego wiersza i kolumny. Skoro A jest górnoprójkątna, to również A_{11} jest górnoprójkątna. A zatem z założenia indukcyjnego $\det A_{11} = a_{22} \cdot \dots \cdot a_{nn}$. \square

Przykład.

$$\begin{aligned} \begin{vmatrix} 1 & 2 & 3 & 4 \\ 0 & 5 & 6 & 7 \\ 0 & 0 & 8 & 9 \\ 0 & 0 & 0 & 10 \end{vmatrix} &= (-1)^{1+1} \cdot \begin{vmatrix} 5 & 6 & 7 \\ 0 & 8 & 9 \\ 0 & 0 & 10 \end{vmatrix} \\ &= (-1)^{1+1} \cdot (-1)^{1+1} \cdot 5 \cdot \begin{vmatrix} 8 & 9 \\ 0 & 10 \end{vmatrix} = (-1)^{-1+1} \cdot (-1)^{1+1} \cdot 5 \cdot (-1)^{1+1} \cdot 8 \cdot |10| = 1 \cdot 5 \cdot 8 \cdot 10. \end{aligned}$$

Czy za pomocą rozumowania jak wyżej policzyć możemy wyznacznik macierzy dolnotrójkątnej? Niestety za pomocą rozwinięcia względem pierwszej kolumny nie łatwo dostrzec, że uzyskamy w istocie wynik analogiczny, jak dla macierzy górnoprójkątnych — wyznacznik macierzy dolnotrójkątnej równy jest iloczynowi wyrazów na przekątnej.

Nawet rozważenie prostego przykładu pokazuje, że nie jest możliwe powtórzenie rozumowania z dowodu Uwagi 24.1.5. Dla macierzy

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

wyznacznik liczony za pomocą pierwszej kolumny to suma trzech wyrażeń postaci:

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

Znacznie wygodniej liczyłoby się powyższy wyznacznik, gdyby można było stosować „rozwinięciem względem pierwszego wiersza” postaci:

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

Czy możliwe jest obliczenie wyznacznika za pomocą rozwinięcia względem innych kolumn, albo nawet wierszy? Innymi słowy, czy mając macierz $A \in M_n(K)$ oraz znając:

- wyrazy macierzy w i -tym wierszu (odp. j -tej kolumnie): $a_{i1}, a_{i2}, \dots, a_{in}$ (odp. $a_{1j}, a_{2j}, \dots, a_{nj}$)
- wyznaczniki macierzy rozmiaru $n - 1$ postaci $A_{i1}, A_{i2}, \dots, A_{in}$ (odp. $A_{1j}, A_{2j}, \dots, A_{nj}$)

prawdziwe są formuły zwane ROZWINIĘCIAMI LAPLACE'A względem i -tego wiersza (odp. j -tej kolumny):

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

Okazuje się, że tak jest i pokażemy to pod koniec kolejnego wykładu. Nie jest to zupełnie natychmiastowe.

Oto przykłady rozwinięć Laplace'a wyznacznika macierzy

$$A = \begin{bmatrix} 1 & 2 & 3 & 4 \\ 5 & 6 & 7 & 8 \\ 9 & 10 & 11 & 12 \\ 13 & 14 & 15 & 16 \end{bmatrix} :$$

- względem pierwszego wiersza:

$$(-1)^{1+1} \cdot 1 \cdot \begin{vmatrix} 6 & 7 & 8 \\ 10 & 11 & 12 \\ 14 & 15 & 16 \end{vmatrix} + (-1)^{1+2} \cdot 2 \cdot \begin{vmatrix} 5 & 7 & 8 \\ 9 & 11 & 12 \\ 13 & 15 & 16 \end{vmatrix} + (-1)^{1+3} \cdot 3 \cdot \begin{vmatrix} 5 & 6 & 8 \\ 9 & 10 & 12 \\ 13 & 14 & 16 \end{vmatrix} + (-1)^{1+4} \cdot 4 \cdot \begin{vmatrix} 5 & 6 & 7 \\ 9 & 10 & 11 \\ 13 & 14 & 15 \end{vmatrix}.$$

- względem drugiej kolumny

$$(-1)^{1+2} \cdot 2 \cdot \begin{vmatrix} 5 & 7 & 8 \\ 9 & 11 & 12 \\ 13 & 15 & 16 \end{vmatrix} + (-1)^{2+2} \cdot 6 \cdot \begin{vmatrix} 1 & 3 & 4 \\ 9 & 11 & 12 \\ 13 & 15 & 16 \end{vmatrix} + (-1)^{3+2} \cdot 10 \cdot \begin{vmatrix} 1 & 3 & 4 \\ 5 & 7 & 8 \\ 13 & 15 & 16 \end{vmatrix} + (-1)^{4+2} \cdot 14 \cdot \begin{vmatrix} 1 & 3 & 4 \\ 5 & 7 & 8 \\ 9 & 11 & 12 \end{vmatrix}.$$

Aby zobaczyć, że istotnie zachodzą równości (\dagger) , kluczowe będzie zrozumienie w jaki sposób operacje elementarne zmieniają wyznacznik i w rezultacie wykazanie następującego twierdzenia.

Twierdzenie 24.1.6: Wyznacznik macierzy transponowanej

Dla dowolnej macierzy $A \in M_n(K)$ mamy $\det A^T = \det A$.

Transponowanie pozwoli oczywiście na przechodzenie od rozwinięć względem wierszy do rozwinięć względem kolumn. Zrozumiemy to dokładnie na kolejnym wykładzie, ale już teraz odnotować możemy naturalny, od razu widoczny wniosek: liczenie wyznacznika macierzy A za pomocą rozwinięcia Laplace'a względem 1-wszego wiersza jest tym samym, co liczenie wyznacznika, co liczenie wyznacznika A^T za pomocą rozwinięcia względem 1-wszej kolumny. Rzeczywiście, jeśli $A^T = [b_{ij}]$, to $a_{ij} = b_{ji}$ oraz $A_{ij} = (A^T)_{ji}$, a stąd

$$\sum_{k=1}^n (-1)^{1+k} a_{1k} \det A_{1k} = \sum_{k=1}^n (-1)^{k+1} b_{k1} \det(A^T)_{k1}.$$

Po prawej stronie znajduje się, zgodnie z powyższym twierdzeniem, wyznacznik macierzy A , więc za pomocą rozwinięcia po lewej można wyliczyć wyznacznik. Szczególnym wnioskiem z tego rozumowania jest stwierdzenie, że wyznacznik macierzy dolnotrójkątnej jest iloczynem wyrazów z przekątnej.

Do uzasadnienia (\dagger) potrzebny będzie również poniższy rezultat.

Twierdzenie 24.1.7: Wyznacznik, a operacje elementarne

- (1) Jeśli macierz A' została otrzymana z macierzy A przez dodanie do pewnego wiersza skalarnej wielokrotności innego wiersza, wówczas $\det(A') = \det(A)$.
- (2) Przetawienie wierszy zmienia znak wyznacznika, tzn. jeśli macierz A' została otrzymana z macierzy A przez zamianę miejscami dwóch wierszy, to $\det(A') = -\det(A)$.
- (3) Jeśli macierz A' została otrzymana z macierzy A przez pomnożenie pewnego wiersza przez c , to $\det(A') = c \cdot \det(A)$.

Oto przykład prostego wyliczenia wyznacznika, który nie korzysta z rozwinięcia Laplace'a, ale z faktu, że wyznacznik nie zmienia się przy wykonywaniu operacji dodawania do wiersza innego wiersza pomnożonego przez stałą oraz ze znajomości wyznacznika macierzy górnoprójkątnej.

$$\det \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 10 \end{bmatrix} = \det \begin{bmatrix} 1 & 2 & 3 \\ 3 & 3 & 3 \\ 6 & 6 & 7 \end{bmatrix} = \det \begin{bmatrix} 1 & 2 & 3 \\ 3 & 3 & 3 \\ 0 & 0 & 1 \end{bmatrix} = \det \begin{bmatrix} 1 & 2 & 3 \\ 0 & -4 & -6 \\ 0 & 0 & 1 \end{bmatrix} = 1 \cdot (-4) \cdot 1 = -4.$$

Łącząc zatem powyższe twierdzenie z obserwacją mówiącą, że wyznacznik macierzy górnoprójkątnej jest iloczynem wyrazów na przekątnej dostajemy prosty algorytm liczenia wyznacznika (inaczej niż z rozwinięcia Laplace'a) poprzez „schodkowanie”: aby policzyć $|A|$ sprowadzamy macierz A do postaci schodkowej A' operacjami typu (1) i (2). Wyznacznik A' to po prostu iloczyn wyrazów na przekątnej. Mamy też

$$|A| = (-1)^k \cdot |A'|,$$

gdzie k oznacza liczbę operacji typu (2) użytych przy sprowadzaniu A do A' .

Kluczowy wniosek z powyższych rezultatów, który formułujemy już teraz, jest następujący.

Twierdzenie 24.1.8: macierz jest odwracalna \iff macierz ma niezerowy wyznacznik

Dla każdej macierzy $A \in M_n(K)$ równoważne są warunki:

- $\det(A) \neq 0$,
- A jest odwracalna.

Dowód. Zgodnie z Twierdzeniem 24.1.7 macierz A' powstająca z A poprzez operację elementarną ma niezerowy wyznacznik wtedy i tylko wtedy, gdy A ma niezerowy wyznacznik. Skoro każdą macierz można sprowadzić za pomocą operacji (1)-(3) do postaci zredukowanej, która jest macierzą górnoprójkątną, to wyznacznik A jest równy zero wtedy i tylko wtedy, gdy postać zredukowana tej macierzy ma mniej niż n schodków — wtedy bowiem na jej przekątnej znajduje się zero. Tak jest wtedy i tylko wtedy, gdy macierz A jest nieodwracalna. Macierz jest zaś odwracalna wtedy i tylko wtedy, gdy jej postać zredukowana jest macierzą identycznościową (Wniosek 20.1.17), mającą oczywiście wyznacznik 1. \square

Kluczowe dowody Twierdzeń 24.1.7, 24.1.6 oraz (†) przedstawimy na kolejnym wykładzie. Nie bez przyczyny są one wymienione w odwrotnej kolejności — w istocie to będzie właśnie kolejność ich uzasadniania.

* * *

Na koniec tych wstępnych rozważań warto dodać, że będziemy liczyć wyznaczniki wielu typów macierzy, ale szczególnie istotne będą dla nas macierze stanowiące analog macierzy trójkątnych w kontekście blokowym. Im poświęcimy ostatni fragment wykładu.

Definicja 24.1.9

Niech A będzie macierzą blokową o blokach $[D_{ij}]$, przy czym zakładamy, że macierze D_{ii} są kwadratowe, jak w definicji 21.1.7 (nazywamy je blokami diagonalnymi). Macierz A nazywamy:

- **BLOKOWO GÓRNOTRÓJKĄTNĄ**, jeśli istnieje rozbicie $n = n_1 + \dots + n_k$ na dodatnie składniki takie, że postać blokowa (D_{ij}) macierzy A względem tego rozbicia spełnia $D_{ij} = 0$, dla $i > j$,
- **BLOKOWO DOLNOTRÓJKĄTNĄ**, jeśli istnieje rozbicie $n = n_1 + \dots + n_k$ na dodatnie składniki takie, że postać blokowa (D_{ij}) macierzy A względem tego rozbicia spełnia $D_{ij} = 0$, dla $i < j$.

Oto przykłady macierzy zdefiniowanych wyżej:

$$\begin{bmatrix} 1 & 2 & 0 & 0 \\ 3 & 4 & 0 & 0 \\ 5 & 6 & 1 & 1 \\ 7 & 8 & 1 & 1 \end{bmatrix}, \quad \begin{bmatrix} -1 & 0 & 0 & 0 \\ 0 & 1 & 2 & 3 \\ 0 & 4 & 5 & 6 \\ 0 & 7 & 8 & 9 \end{bmatrix}, \quad \begin{bmatrix} 1 & 1 & 1 & 2 & 2 \\ 1 & 1 & 1 & 2 & 2 \\ 1 & 1 & 1 & 2 & 2 \\ 0 & 0 & 0 & 3 & 3 \\ 0 & 0 & 0 & 3 & 3 \end{bmatrix}.$$

Obserwacja 24.1.10: Wyznacznik macierzy blokowo-trójkątnej

Niech A będzie macierzą blokowo górnoprójkątną lub blokowo dolnotrójkątną o blokach diagonalnych D_{11}, \dots, D_{kk} . Wówczas

$$\det A = \det D_{11} \cdot \det D_{22} \cdot \dots \cdot \det D_{kk}.$$

Dowód. Pokażmy najpierw tezę dla macierzy blokowo górnoprójkątnej $X = [x_{ij}]$ rozmiaru $n \times n$ postaci:

$$X = \begin{bmatrix} A & B \\ 0 & D \end{bmatrix}.$$

Rozumowanie jest indukcją ze względu na rozmiar n macierzy X . Oczywiście teza zachodzi dla $n = 2$ i macierzy o czterech blokach rozmiarów 1×1 . Niech $n > 2$. Oznaczmy kolejne wyrazy pierwszej kolumny macierzy A przez $a_{11}, a_{21}, \dots, a_{s1}$. Liczymy wyznacznik przez rozwinięcie względem pierwszej kolumny, otrzymując

$$\det X = (-1)^{1+1} a_{11} \det X_{11} + \dots + (-1)^{s+1} a_{s1} \det X_{s1}.$$

Rzeczywiście, kolejne $n - s$ składników rozwinięcia zawiera czynnik x_{j1} , który dla $j > s$ równy jest zero. Zauważmy też, że X_{i1} są, dla $1 \leq i \leq s$ macierzami blokowo-górnoprójkątnymi postaci

$$X_{i1} = \begin{bmatrix} A_{i1} & * \\ 0 & D \end{bmatrix}.$$

A zatem zgodnie z założeniem indukcyjnym

$$\det X_{i1} = \det A_{i1} \cdot \det D.$$

W ten sposób uzyskujemy krok indukcyjny, bowiem:

$$\det X = (-1)^{1+1} a_{11} \det A_{11} \cdot \det D + \dots + (-1)^{s+1} a_{s1} \det A_{s1} \cdot \det D = \det A \cdot \det D.$$

Dla macierzy blokowo-górnoprójkątnej o więcej niż 2 blokach rozumowanie jest prostą indukcję ze względu na liczbę bloków. Zauważmy bowiem, że macierz o $k > 1$ blokach diagonalnych D_{11}, \dots, D_{kk} traktować można jako macierz o dwóch blokach diagonalnych: D_{11} oraz bloku, którego blokami diagonalnymi są D_{22}, \dots, D_{kk} . Rozumowanie dla macierzy blokowo dolnotrójkątnych wynika natomiast natychmiast z tego, że wyznacznik nie zmienia się przy transponowaniu, a transpozycja macierzy blokowo górnoprójkątnej jest macierzą blokowo dolnotrójkątną. \square

24.2 Wybór przykładowych pytań

Proszę się nie uczyć na pamięć odpowiedzi czy wyjaśnienie! Takich pytań można ułożyć znacznie więcej i służą one jedynie sprawdzeniu podstawowego zrozumienia treści zajęć. Jeśli odpowiedź na dowolne z poniższych pytań sprawia Państwu problem, proszę pomyśleć który fragment notatek dotyczy tego pytania, wrócić do ich lektury, a w razie dalszych trudności poprosić prowadzącego o wskazówkę.

- Oblicz w pamięci wyznaczniki macierzy

$$\begin{bmatrix} \sin x & \cos x \\ -\cos x & \sin x \end{bmatrix}, \quad \begin{bmatrix} a+b & a-b \\ a-b & a+b \end{bmatrix}, \quad \begin{bmatrix} 1 & 2 & 3 \\ 0 & -1 & 5 \\ 0 & 0 & 18 \end{bmatrix}, \quad \begin{bmatrix} -\frac{1}{2} & 0 & 0 \\ -\frac{3}{5} & -\frac{1}{2} & 0 \\ -\frac{4}{5} & -\frac{3}{5} & -\frac{1}{2} \end{bmatrix}.$$

- Uzasadnij, bez wyliczania, że poniższe wyznaczniki są liczbami całkowitymi podzielnymi przez 3

$$\begin{vmatrix} 0 & 1 & 2 \\ 3 & 4 & 5 \\ 6 & 7 & 8 \end{vmatrix}, \quad \begin{vmatrix} 48 & 60 & 99 \\ 47 & 46 & 50 \\ 32 & 50 & 81 \end{vmatrix}, \quad \begin{vmatrix} 2 & 5 & 7 \\ 4 & 3 & 2 \\ 1 & 4 & 2 \end{vmatrix}.$$

- Opisaliśmy, jak wyznacznik zachowuje się przy operacjach elementarnych na wierszach macierzy. Czy możemy sformułować analogiczne reguły dla operacji elementarnych na kolumnach macierzy?
- Wyjaśnij, bez obliczania odpowiednich wyznaczników, że zachodzą równości:

$$\begin{vmatrix} a-c & b-d \\ c & d \end{vmatrix} = \begin{vmatrix} a & b \\ c & d \end{vmatrix}, \quad \begin{vmatrix} a+bx & b \\ c+dx & d \end{vmatrix} = \begin{vmatrix} a & b \\ c & d \end{vmatrix}$$

- Oblicz wyznacznik poniższej macierzy $A \in M_3(\mathbb{Z}_5)$ i rozstrzygnij, czy jest ona odwracalna:

$$A = \begin{bmatrix} 1 & 2 & 2 \\ 4 & 3 & 3 \\ 2 & 2 & 0 \end{bmatrix}.$$

- Niech $\omega \in \mathbb{C}$ będzie nierzeczywistym pierwiastkiem stopnia 3 z 1. Zapisz w najprostszej możliwej postaci wyznaczniki postaci

$$\begin{vmatrix} 1 & \omega & \omega^2 \\ \omega & \omega^2 & 1 \\ \omega^2 & 1 & \omega \end{vmatrix}, \quad \begin{vmatrix} 1 & 1 & \omega \\ 1 & 1 & \omega^2 \\ \omega^2 & \omega & 1 \end{vmatrix}, \quad \begin{vmatrix} \omega^2 & \omega & 1 \\ \omega & \omega^2 & 1 \\ 1 & 1 & \omega^2 \end{vmatrix}.$$

- Niech $A = [x \ y]$. Oblicz $\det(A^T \cdot A)$ oraz $\det(A \cdot A^T)$.

- Dana jest macierz

$$A = \begin{bmatrix} a & b & x \\ c & d & y \end{bmatrix},$$

przy czym $ad - bc \neq 0$. Wyznacz rząd macierzy A .

- Uzasadnij, bez wykonywania rachunków, że poniższa macierz rzeczywista ma zerowy wyznacznik.

$$\begin{bmatrix} 2 & 3 & 0 & 9 & 0 & 1 & 0 & 1 & 1 & 2 & 1 \\ 1 & 1 & 0 & 3 & 0 & 0 & 0 & 9 & 2 & 3 & 1 \\ 1 & 4 & 0 & 2 & 8 & 5 & 0 & 3 & 6 & 1 & 9 \\ 0 & 0 & 0 & 0 & 0 & 5 & 0 & 0 & 0 & 0 & 0 \\ 2 & 2 & 4 & 1 & 1 & 2 & 1 & 6 & 9 & 0 & 7 \\ 0 & 0 & 0 & 6 & 0 & 7 & 0 & 1 & 0 & 0 & 0 \\ 2 & 5 & 0 & 7 & 0 & 4 & 6 & 8 & 5 & 1 & 3 \\ 0 & 0 & 0 & 1 & 0 & 4 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 8 & 0 & 2 & 0 & 0 & 0 & 0 & 0 \\ 2 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 2 & 1 & 1 \\ 2 & 6 & 0 & 1 & 0 & 30 & 0 & 2 & 3 & 2 & 1 \end{bmatrix}$$

- Podaj przykład macierzy kwadratowych A, B , że $\det A = \det B = 0$, zaś $\det(A + B) = 1$.
- Załóżmy, że $A \in M_2(\mathbb{R})$ spełnia warunek $A^2 = 0$ oraz, że macierz A jest górnopróbką. Uzasadnij, że $\det(A + I_2) = 1$.
- Niech $A \in M_n(K)$. Uzasadnij, że dla każdego $\lambda \in K$ zachodzi równość $\det(\lambda I_n + A) = \det(\lambda I_n + A^T)$.

24.3 Zadania do samodzielnej pracy

Zadania ilustrujące umiejętności niezbędne do zaliczenia przedmiotu oznaczone są symbolem ♠.

Przy tych zadaniach dodany jest opis umiejętności, które sprawdzają.

1. (♠) Obliczanie wyznacznika za pomocą rozwinięcia Laplace'a) Oblicz wyznaczniki macierzy:

$$A_1 = \begin{bmatrix} 1 & 2 & 1 & 3 \\ 0 & 1 & 1 & 0 \\ 2 & 1 & 3 & 1 \\ 1 & 0 & 2 & 0 \end{bmatrix}, \quad A_2 = \begin{bmatrix} 3 & 0 & 0 & 0 \\ 5 & 1 & 0 & 0 \\ 4 & 3 & 6 & 0 \\ 7 & 9 & 2 & 8 \end{bmatrix}, \quad A_3 = \begin{bmatrix} 2 & 0 & 0 & 1 \\ 0 & 3 & 1 & 0 \\ 2 & 1 & 1 & 2 \\ 3 & 4 & 5 & 4 \end{bmatrix}, \quad A_4 = \begin{bmatrix} 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}.$$

2. (♠) Obliczanie wyznacznika przez sprowadzenie do postaci trójkątnej) Oblicz wyznaczniki:

$$\begin{vmatrix} 3 & 4 & 2 & 2 \\ 4 & 5 & 6 & 5 \\ 2 & 3 & 6 & 0 \\ 8 & 7 & 7 & 8 \end{vmatrix}, \quad \begin{vmatrix} 36 & 60 & 72 & 37 \\ 43 & 71 & 78 & 34 \\ 44 & 69 & 73 & 32 \\ 30 & 50 & 65 & 38 \end{vmatrix}, \quad \begin{vmatrix} 35 & 59 & 71 & 52 \\ 42 & 70 & 77 & 54 \\ 43 & 68 & 72 & 52 \\ 29 & 49 & 65 & 50 \end{vmatrix}.$$

3. (♠) Oblicz wyznaczniki następujących macierzy blokowych

$$C_1 = \begin{bmatrix} 7 & 9 & 8 & 2 \\ 4 & 6 & 7 & 0 \\ 8 & 1 & 0 & 0 \\ 3 & 0 & 0 & 0 \end{bmatrix}, \quad C_2 = \begin{bmatrix} 3 & 1 & 0 & 0 \\ 5 & 1 & 0 & 0 \\ 0 & 3 & 6 & 7 \\ 7 & 9 & 2 & 8 \end{bmatrix}, \quad C_3 = \begin{bmatrix} 4 & 1 & 2 & 9 \\ 5 & 0 & 1 & 7 \\ 2 & 1 & 1 & 8 \\ 0 & 0 & 0 & 4 \end{bmatrix}, \quad C_4 = \begin{bmatrix} 5 & 4 & 7 & 1 \\ 9 & 7 & 8 & 3 \\ 0 & 0 & 8 & 7 \\ 0 & 0 & 6 & 5 \end{bmatrix}.$$

4. (Obliczanie wyznacznika za pomocą rozwinięcia Laplace'a i indukcji matematycznej)

Oblicz wyznaczniki następujących macierzy rozmiaru $n \times n$:

$$A_1 = \begin{bmatrix} 1 & 0 & 0 & \dots & 0 & 1 \\ 0 & 1 & 0 & \dots & 0 & 0 \\ 0 & 0 & 1 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & 1 & 0 \\ 1 & 0 & 0 & \dots & 0 & 1 \end{bmatrix}, \quad A_2 = \begin{bmatrix} 0 & 0 & 0 & \dots & 0 & 1 \\ 0 & 0 & 0 & \dots & 1 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 1 & \dots & 0 & 0 \\ 0 & 1 & 0 & \dots & 0 & 0 \\ 1 & 0 & 0 & \dots & 0 & 0 \end{bmatrix}, \quad A_3 = \begin{bmatrix} 1 & 1 & 1 & \dots & 1 & 1 \\ -1 & 1 & 0 & \dots & 0 & 0 \\ 0 & -1 & 1 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & 1 & 0 \\ 0 & 0 & 0 & \dots & -1 & 1 \end{bmatrix},$$

$$A_4 = \begin{bmatrix} 0 & 1 & 1 & \dots & 1 & 1 \\ 1 & 1 & 0 & \dots & 0 & 0 \\ 1 & 0 & 1 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 1 & 0 & 0 & \dots & 1 & 0 \\ 1 & 0 & 0 & \dots & 0 & 1 \end{bmatrix}, \quad A_5 = \begin{bmatrix} 1 & i & 0 & \dots & 0 & 0 \\ i & 1 & i & \dots & 0 & 0 \\ 0 & i & 1 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & 1 & i \\ 0 & 0 & 0 & \dots & i & 1 \end{bmatrix}, \quad A_6 = \begin{bmatrix} 1 & 1 & 1 & \dots & 1 & 1 \\ 1 & 2 & 2 & \dots & 2 & 2 \\ 1 & 2 & 3 & \dots & 3 & 3 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 1 & 2 & 3 & \dots & n-1 & n-1 \\ 1 & 2 & 3 & \dots & n-1 & n \end{bmatrix}.$$

5. Oblicz wyznaczniki następujących macierzy $n \times n$ w zależności od parametrów $s, t \in \mathbb{R}$

$$C_1 = \begin{bmatrix} 0 & 1 & 1 & 1 & \dots & 1 \\ 1 & 0 & s & s & \dots & s \\ 1 & s & 0 & s & \dots & s \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & s & s & s & \dots & s \\ 1 & s & s & s & \dots & 0 \end{bmatrix}, \quad C_2 = \begin{bmatrix} 2 & t & 0 & 0 & \dots & 0 \\ 0 & 2 & t & 0 & \dots & 0 \\ 0 & 0 & 2 & t & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & t \\ t & 0 & 0 & 0 & \dots & 2 \end{bmatrix}, \quad C_3 = \begin{bmatrix} s & t & t & t & \dots & t \\ t & s & t & t & \dots & t \\ t & t & s & t & \dots & t \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ t & t & t & t & \dots & t \\ t & t & t & t & \dots & s \end{bmatrix}.$$

6. Wyznacz $\det A$, gdzie $A = [a_{ij}] \in M_n(\mathbb{R})$ oraz $a_{ij} = \begin{cases} 2, & \text{dla } i = j, \\ (-1)^{|i-j|}, & \text{dla } i \neq j. \end{cases}$

7. Wyrazami macierzy kwadratowej A należącej do zbioru $M_4(\mathbb{R})$ są wyłącznie liczby -2 oraz 1 (dowolnie ustawione). Wykaż, że wyznacznik macierzy A jest liczbą całkowitą podzielną przez 27 .

8. Niech $A \in M_n(\mathbb{R})$ będzie macierzą o wyrazach nieujemnych taką, że suma wyrazów w każdym wierszu jest mniejsza niż 1 . Wykaż, że $|\det(A)| \leq 1$.

9. Wykaż, że dla dowolnych dodatnich liczb całkowitych m, n mamy

$$\det \begin{bmatrix} 0 & I_n \\ I_m & 0 \end{bmatrix} = (-1)^{mn}.$$

Rozdział 25

Wyznacznik — funkcja objętości

25.1 Wykład 25

Zdefiniowaliśmy na ostatnim wykładzie wyznacznik postulując, że funkcja ta ma pewne naturalne właściwości przy wykonywaniu operacji elementarnych. W tym rozdziale będziemy te właściwości uzasadniać. Dodamy najpierw pewien naturalny kontekst geometryczny, jedynie zasygnalizowany ostatnio, wiążący wyznacznik z objętością (wróćmy do niego w pełni, jak zapowiadaliśmy, w kolejnym semestrze).

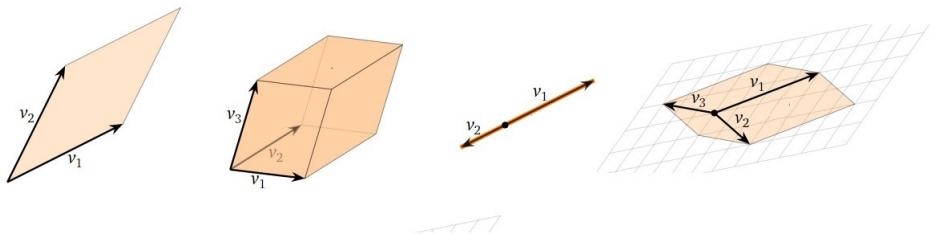
Definicja 25.1.1

RÓWNOLEGŁOŚCIANEM rozpiętym na wektorach $v_1, \dots, v_n \in \mathbb{R}^n$ nazywamy podzbiór

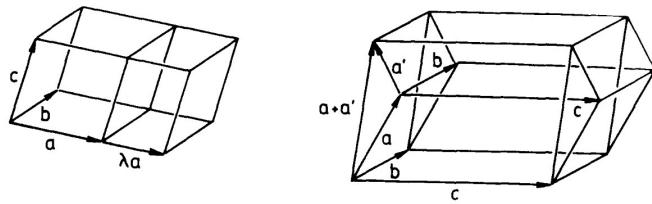
$$R(v_1, \dots, v_n) = \{a_1 v_1 + \dots + a_n v_n \mid 0 \leq a_1, \dots, a_n \leq 1\}.$$

Zauważmy, że powyższa definicja zakłada, że każdy równoległościan zawiera wektor zerowy. W drugim semestrze rozszerzymy tę definicję na odpowiednie podzbiory (euklidesowych) przestrzeni afnicznych.

Poniżej przedstawionych jest kilka ilustracji równoległościanów, nawiązujących do definicji szkolnych. Pierwsze dwa (od lewej) rozpięte są przez układy liniowo niezależne, a kolejne dwa (odpowiednio w przestrzeniach dwu- i trójwymiarowej) rozpięte są przez układy liniowo zależne. Czym różnią się te sytuacje?



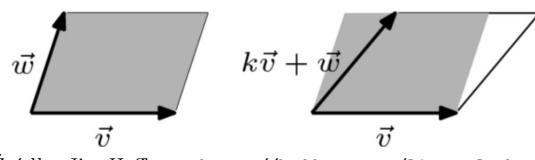
Dlaczego formułujemy taką definicję, zwłaszcza skoro dotyczy ona jedynie przestrzeni nad ciałem \mathbb{R} ? Spójrzmy na ilustrację równoległościanów $R(\lambda \cdot a, b, c)$ oraz $R(a + a', b, c)$, gdzie $a, b, c \in \mathbb{R}^3$ oraz $\lambda > 0$.



Źródło: K. Spindler, Abstract Algebra with Applications

Na gruncie geometrii elementarnej wiemy, że „objętość” równoległościanu $R(\lambda \cdot a, b, c)$ równa jest λ razy „objętość” $R(a, b, c)$, zaś „objętość” $R(a+a', b, c)$ równa jest sumie „objętości” $R(a, b, c)$ oraz $R(a', b, c)$.

Co się stanie z „objętością”, jeśli w miejsce wektora a wstawimy $a + \lambda b$? Dobre ilustruje to już przypadek dwuwymiarowy, gdzie widzimy, że „objętość” nie ulega po takiej operacji zmianie.



Źródło: Jim Hefferon, <https://hefferon.net/linearalgebra/>

Jak wyznacznik ma się do owej „objętości”? W istocie, traktując wektory rozpinające równoległościan jako wiersze pewnej macierzy (lub kolumny), można powiedzieć, że rozważana przez nas „objętość” zachowuje się w zasadzie tak samo jak wyznacznik, czyniąc wyznacznik jakby „funkcją objętości”.

Porządne powiązanie wyznacznika z objętością wykonamy w drugim semestrze. Teraz, dla uzasadnienia postulowanych na poprzednim wykładzie własności wyznacznika — w szczególności jego zachowania przy wykonywaniu operacji elementarnych — zajmiemy się badaniem funkcji, które mają takie własności jak opisana wyżej „objętość”. Są to jak się okazuje te same własności, jakie postulujemy dla wyznacznika. Przekonamy się, że owe własności wyznaczają spełniającą ją funkcję w zasadzie jednoznacznie, pozwalając tym samym udowodnić szereg interesujących nas rezultatów dotyczących wyznacznika. Oto te własności.

Definicja 25.1.2: Funkcje jednorodne i addytywne względem wierszy macierzy

Powiemy, że funkcja $\phi : M_n(K) \rightarrow K$, jest

- JEDNORODNA WZGLĘDEM k -TEGO WIERSZA, jeśli dla każdej $A \in M_n(K)$ oraz każdego $c \in K$ mamy $\phi(A') = c \cdot \phi(A)$, gdzie A' powstaje z A przez pomnożenie k -tego wiersza przez c .
- ADDYTYWNA WZGLĘDEM k -TEGO WIERSZA, jeśli $\phi(C) = \phi(A) + \phi(B)$, dla każdej trójki macierzy $A, B, C \in M_n(K)$ spełniającej poniższe dwa warunki
 - k -ty wiersz macierzy C to suma k -tego wiersza A oraz k -tego wiersza B ,
 - l -te macierzy A, B, C są identyczne, dla $l \neq k$.

Przykład. Niech $A, B, C \in M_3(K)$ mają takie same pierwsze i drugie wiersze oraz przyjmijmy, że trzeci wiersz macierzy C jest sumą trzeciego wiersza macierzy A oraz trzeciego wiersza macierzy B :

$$A = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ \textcolor{red}{1} & \textcolor{red}{1} & \textcolor{red}{2} \end{bmatrix}, \quad B = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ -1 & 3 & 0 \end{bmatrix}, \quad C = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ \textcolor{teal}{0} & \textcolor{teal}{4} & \textcolor{teal}{2} \end{bmatrix}.$$

Nietrudno sprawdzić (choćby na mocy wzoru Sarrusa), że $\det C = \det A + \det B$.

Przykład ten sugeruje addytywność wyznacznika macierzy rozmiaru 3×3 względem trzeciego wiersza.

Twierdzenie 25.1.3

Funkcja $\det : M_n(K) \rightarrow K$ jest jednorodna i addytywna względem każdego wiersza.

Zanim przedstawimy dowód, podkreślmy raz jeszcze — z geometrycznego punktu widzenia funkcja przypisująca objętość równoległościanowi (rozumianą na razie intuicyjnie) jest jednorodna i addytywna ze względu na każdą „współrzędną”. W języku algebry liniowej mówimy, że jest to tzw. funkcja *wieloliniovą*. Ogólnie dla przestrzeni liniowych V_1, \dots, V_n nad ciałem K przekształcenie $\varphi : V_1 \times \dots \times V_n \rightarrow V$ nazywamy *n-LINIOWYM*, gdy dla dowolnych $v_1 \in V_1, \dots, v_n \in V_n$ następujące przekształcenia są liniowe

$$V_i \ni x \mapsto \varphi(v_1, \dots, v_{i-1}, x, v_{i+1}, \dots, v_n) \in V \quad (1 \leq i \leq n).$$

Naszym celem będzie w istocie pokazanie, że wyznacznik macierzy $A \in M_n(K)$, interpretowany jako funkcja od n wierszy, jest funkcją *n-liniową* na $K^n \times \dots \times K^n \simeq M_n(K)$ — i to w pewnym sensie jedyną. Trzymać się będziemy jednak terminologii funkcji addytywnych i jednorodnych. Wykażemy również dalej, że wyznacznik jest funkcją addytywną i jednorodną ze względu na każdą kolumnę.

Dowód. Indukcja ze względu na n . Dla $n = 1$ – jasne. Weźmy $A = [a_{ij}] \in M_n(K)$. Mnożymy k -ty wiersz A przez $c \in K$ dostając B . Chcemy porównać wyznaczniki macierzy A i B poprzez przyjrzenie się ich rozwinięciom względem pierwszej kolumny (to na razie jedyna definicja wyznacznika, z której możemy korzystać). Wówczas:

- $B_{k1} = A_{k1}$,
- dla $j \neq k$ każda z macierzy B_{j1} powstaje z A_{j1} przez pomnożenie pewnego wiersza przez stałą. Za założenia indukcyjnego mamy więc $\det B_{j1} = c \det A_{j1}$ (są to macierze rozmiaru $n - 1$).

Zatem:

$$\begin{aligned}\det B &= (-1)^{1+1}a_{11} \det B_{11} + \dots + (-1)^{k+1}ca_{k1} \det B_{k1} + \dots + (-1)^{n+1}a_{n1} \det B_{n1} = \\ &= (-1)^{1+1}a_{11}c \det A_{11} + \dots + (-1)^{k+1}ca_{k1} \det A_{k1} + \dots + (-1)^{n+1}a_{n1}c \det A_{n1} = c \cdot \det A.\end{aligned}$$

Pokażmy teraz addytywność wyznacznika ze względu na k -ty wiersz. Ponownie argumentujemy przez indukcję ze względu na n . Dla $n > 1$ chcemy pokazać, że:

$$\det \underbrace{\begin{bmatrix} a_{11} & \dots & a_{1n} \\ \vdots & \ddots & \vdots \\ x_{k1} + y_{k1} & \dots & x_{kn} + y_{kn} \\ \vdots & \ddots & \vdots \\ a_{n1} & \dots & a_{nn} \end{bmatrix}}_Z = \det \underbrace{\begin{bmatrix} a_{11} & \dots & a_{1n} \\ \vdots & \ddots & \vdots \\ x_{k1} & \dots & x_{kn} \\ \vdots & \ddots & \vdots \\ a_{n1} & \dots & a_{nn} \end{bmatrix}}_X + \det \underbrace{\begin{bmatrix} a_{11} & \dots & a_{1n} \\ \vdots & \ddots & \vdots \\ y_{k1} & \dots & y_{kn} \\ \vdots & \ddots & \vdots \\ a_{n1} & \dots & a_{nn} \end{bmatrix}}_Y$$

gdzie $X, Y, Z \in M_n(K)$ różnią się tylko k -tymi wierszami — k -ty wiersz Z jest sumą k -tych wierszy X, Y .

Zauważmy, że:

- dla każdego $1 \leq k \leq n$ macierze powstające z Z, X, Y przez usunięcie k -tego wiersza i pierwszej kolumny są równe, tzn. $Z_{k1} = Y_{k1} = X_{k1}$,
- dla $j \neq k$ macierze Z_{j1}, Y_{j1}, X_{j1} różnią się tylko k -tym wierszem, przy czym k -ty wiersz Z_{j1} jest sumą k -tych wierszy Y_{j1} oraz X_{j1} . Za założenia indukcyjnego mamy zatem

$$\det Z_{j1} = \det X_{j1} + \det Y_{j1}, \quad \text{dla } j \neq k.$$

Stąd:

$$\begin{aligned}\det Z &= (-1)^{1+1}a_{11} \det Z_{11} + \dots + (-1)^{k+1}(x_{k1} + y_{k1}) \det Z_{k1} + \dots + (-1)^{n+1}a_{n1} \det Z_{n1} = \\ &= \sum_{j \neq k} (-1)^{j+1}a_{j1}(\det X_{j1} + \det Y_{j1}) + (-1)^{k+1}x_{k1} \det X_{k1} + y_{k1} \det Y_{k1} = \det X + \det Y\end{aligned}$$

□

Podkreślmy raz jeszcze kluczowy wątek tego dowodu — uzasadniliśmy pewną własność funkcji wyznacznik określonej za pomocą rozwinięcia względem pierwszej kolumny i nigdzie nie korzystaliśmy z rezultatów, które dopiero chcemy udowodnić. To za pomocą rozwinięcia względem pierwszej kolumny uzasadniliśmy, że wyznacznik na własność addytywności i jednorodności względem każdego wiersza!

Wniosek 25.1.4

Jeśli funkcja $\phi : M_n(K) \rightarrow K$ jest jednorodna i addytywna względem każdego wiersza, oraz macierz $A \in M_n(K)$ ma zerowy wiersz, to $\phi(A) = 0$. W szczególności dla macierzy A o zerowym wierszu mamy $\det(A) = 0$.

Dowód. Jeśli k -ty wiersz macierzy A jest zerowy, to z addytywności ϕ względem k -tego wiersza mamy $\phi(A) = \phi(A) + \phi(A)$ (w dowodzie wyżej przyjmujemy $X = Y = Z = A$). □

Uwaga 25.1.5

Jeśli dwa sąsiednie wiersze macierzy $A \in M_n(K)$ są identyczne, dla $n \geq 2$, wówczas $\det A = 0$.

Dowód. Dowód to indukcja ze względu na n . Dla $n = 2$ teza jest oczywiście prawdziwa. Założymy, że $n > 2$ oraz identyczne są i -ty oraz $i + 1$ -ty wiersz macierzy $A = [a_{ij}]$, czyli dla $1 \leq k \leq n$ mamy $a_{ik} = a_{i+1,k}$.

$$A = \begin{bmatrix} \vdots & \vdots & \ddots & \vdots \\ a_{i1} & a_{i2} & \dots & a_{in} \\ a_{i+1,1} & a_{i+1,2} & \dots & a_{i+1,n} \\ \vdots & \vdots & \ddots & \vdots \end{bmatrix}.$$

Zauważmy po pierwsze, że $A_{i1} = A_{i+1,1}$. Natomiast dla $k \neq i, i + 1$ macierze A_{k1} mają dwa identyczne wiersze. Z założenia indukcyjnego w tym drugim przypadku mamy zatem $\det A_{k1} = 0$. Zatem

$$\begin{aligned} \det A &= \sum_{k \neq i, i+1} (-1)^{k+1} a_{k1} \det A_{k1} + (-1)^{i+1} a_{i1} \det A_{i1} + (-1)^{i+1+1} a_{i+1,1} \det A_{i+1,1} \\ &= (-1)^{i+1} (1 - 1) a_{i1} \det A_{i1} = 0. \end{aligned}$$

□

Oto przykład ilustrujący krok indukcyjny w dowodzie wyżej. Rozważmy macierz rozmiaru 4×4 o dwóch identycznych wierszach i założymy, że obserwacja jest prawdziwa dla macierzy 3×3 . Weźmy macierz o dwóch identycznych wierszach i policzmy jej wyznacznik

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

Mamy zatem (stosujemy notację $|A|$ alternatywną do $\det A$):

$$\begin{aligned} |A| &= 1 \cdot |A_{11}| - 5 \cdot |A_{21}| + 5 \cdot |A_{31}| - 9 \cdot |A_{41}| \\ &= 1 \cdot \underbrace{\begin{vmatrix} 6 & 7 & 8 \\ 6 & 7 & 8 \\ 8 & 7 & 1 \end{vmatrix}}_0 - 5 \cdot \underbrace{\begin{vmatrix} 2 & 3 & 4 \\ 6 & 7 & 8 \\ 8 & 7 & 1 \end{vmatrix}}_{20} + 5 \cdot \underbrace{\begin{vmatrix} 2 & 3 & 4 \\ 6 & 7 & 8 \\ 8 & 7 & 1 \end{vmatrix}}_{20} - 9 \cdot \underbrace{\begin{vmatrix} 2 & 3 & 4 \\ 6 & 7 & 8 \\ 6 & 7 & 8 \end{vmatrix}}_0. \end{aligned}$$

Czytelnik widzi oczywiście, że uzasadnioną wyżej własność mają wszystkie funkcje z $M_n(K)$ do K będące jednorodne i addytywne względem każdego wiersza. Wykażemy kolejne własności \det , dając do uzasadnienia, że wszystkie rozwinięcia Laplace'a określają tę samą funkcję, co wyraża równość (†) z poprzedniego rozdziału. Kluczowy będzie następujący rezultat.

Twierdzenie 25.1.6

Dla każdego $n \geq 1$ istnieje dokładnie jedna funkcja $\phi : M_n(K) \rightarrow K$, taka, że:

- (1) Dla każdego $1 \leq k \leq n$ funkcja ϕ jest jednorodna względem k -tego wiersza.
- (2) Dla każdego $1 \leq k \leq n$ funkcja ϕ jest addytywna względem k -tego wiersza.
- (3) $\phi(A) = 0$, jeśli A ma identyczne dwa sąsiednie wiersze.
- (4) $\phi(I_n) = 1$.

Łatwym zadaniem jest na tym etapie rozważań wskazanie przykładu funkcji spełniającej (1)-(4), trudniejszym będzie natomiast dowód jej jedyności.

Uwaga 25.1.7

Funkcja $\phi = \det$ spełnia warunki (1)-(4).

Dowód. Spełnianie warunków (1), (2) zapewnia Twierdzenie 25.1.3. Warunek (3) wynika stąd, że wyznacznik macierzy A jest niezerowy wtedy i tylko przy zamianie wierszy jest dalej niezerowy, a przy dwóch identycznych sąsiednich wierszach jest zerowy (Uwaga 25.1.5). Oczywiście $\det(I_n) = 1$. \square

Twierdzimy, że żadnej innej funkcji niż \det spełniającej warunki (1)-(4) nie ma. Idea dowodu polega na pokazaniu, że funkcja spełniająca warunki (1)-(4) jest jednoznacznie określona na macierzach operacji elementarnych, co pozwoli nam wywnioskować, że dla każdej macierzy może ona przyjmować tylko jedną wartość. W szczególności dowodzić będziemy w kolejnych obserwacjach Twierdzenie 24.1.7 mówiącego o tym jak zachowuje się wyznacznik przy operacjach elementarnych na wierszach.

Uwaga 25.1.8

Załóżmy, że funkcja $\phi : M_n(K) \rightarrow K$ spełnia warunki (1)-(4) (np. $\phi = \det$, o czym już wiemy). Jeśli $C' \in M_n(K)$ powstaje z C przez zamianę dwóch sąsiednich wierszy, to $\phi(C) = -\phi(C')$.

Dowód. Niech wiersze macierzy C mają postać w_1, \dots, w_n . Niech C' powstaje z C przez zamianę wiersza k -tego i $k+1$ -wszego. Na mocy własności (2) i (3) funkcji ϕ :

$$\phi \underbrace{\begin{bmatrix} w_1 \\ \vdots \\ w_k + w_{k+1} \\ w_k + w_{k+1} \\ \vdots \\ w_n \end{bmatrix}}_0 = \phi \underbrace{\begin{bmatrix} w_1 \\ \vdots \\ w_k \\ w_k \\ \vdots \\ w_n \end{bmatrix}}_0 + \phi \underbrace{\begin{bmatrix} w_1 \\ \vdots \\ w_{k+1} \\ w_{k+1} \\ \vdots \\ w_n \end{bmatrix}}_0 + \phi \underbrace{\begin{bmatrix} w_1 \\ \vdots \\ w_k \\ w_{k+1} \\ \vdots \\ w_n \end{bmatrix}}_{\phi(C)} + \phi \underbrace{\begin{bmatrix} w_1 \\ \vdots \\ w_{k+1} \\ w_k \\ \vdots \\ w_n \end{bmatrix}}_{\phi(C')}$$

Uwaga 25.1.9

Załóżmy, że funkcja $\phi : M_n(K) \rightarrow K$ spełnia warunki (1)-(4) (np. $\phi = \det$, o czym już wiemy). Jeśli macierz $C \in M_n(K)$ ma dwa identyczne wiersze, to $\phi(C) = 0$.

Uzasadnienie: za pomocą skończenie wielu operacji zamiany wierszy możemy zamienić C w macierz C' o dwóch sąsiednich wierszach równych. Z poprzedniej obserwacji mamy $\phi(C) = \pm\phi(C') = 0$.

Uwaga 25.1.10

Załóżmy, że funkcja $\phi : M_n(K) \rightarrow K$ spełnia warunki (1)-(4) (np. $\phi = \det$, o czym już wiemy). Niech B będzie macierzą otrzymaną z macierzy A w wyniku dodania do wiersza l -tego wiersza k -tego pomnożonego przez $a \in K$. Wówczas: $\phi(B) = \phi(A)$.

Schemat uzasadnienia, korzystający z poprzednich wyników:

$$\phi \underbrace{\begin{bmatrix} w_1 \\ \vdots \\ w_l + aw_k \\ \vdots \\ w_k \\ \vdots \\ w_n \end{bmatrix}}_{\phi(B)} = \phi \underbrace{\begin{bmatrix} w_1 \\ \vdots \\ w_l \\ \vdots \\ w_k \\ \vdots \\ w_n \end{bmatrix}}_{\phi(A)} + \phi \underbrace{\begin{bmatrix} w_1 \\ \vdots \\ aw_k \\ \vdots \\ w_k \\ \vdots \\ w_n \end{bmatrix}}_{a \cdot \phi(A)} = \phi \underbrace{\begin{bmatrix} w_1 \\ \vdots \\ w_l \\ \vdots \\ w_k \\ \vdots \\ w_n \end{bmatrix}}_{\phi(A)} + a \cdot \phi \underbrace{\begin{bmatrix} w_1 \\ \vdots \\ w_k \\ \vdots \\ w_n \end{bmatrix}}_0.$$

Zauważmy, że powyższe uwagi domykając pełen opis zachowania się funkcji spełniających warunki (1)-(4) przy operacjach elementarnych, i w szczególności dają nam one dowód Twierdzenia 24.1.7. Sformułujemy je w języku mnożenia macierzy przez macierze operacji elementarnych.

Obserwacja 25.1.11: Wyznacznik macierzy operacji elementarnych

Załóżmy, że funkcja $\phi : M_n(K) \rightarrow K$ spełnia warunki (1)-(4) (np. \det , o czym już wiemy). Niech M będzie macierzą operacji elementarnej oraz $A \in M_n(K)$. Wówczas:

$$\phi(MA) = \begin{cases} \phi(A), & \text{dla } M \text{ dodającej do wiersza skalar razy inny wiersz,} \\ -\phi(A), & \text{dla } M \text{ zamieniającej dwa wiersze miejscami,} \\ c \cdot \phi(A), & \text{dla } M \text{ mnożącej pewien wiersz przez } c \neq 0. \end{cases}$$

W szczególności dla $A = I_n$ mamy

$$\phi(M) = \begin{cases} 1, & \text{dla } M \text{ dodającej do wiersza skalar razy inny wiersz,} \\ -1, & \text{dla } M \text{ zamieniającej dwa wiersze miejscami,} \\ c, & \text{dla } M \text{ mnożącej pewien wiersz przez } c \neq 0. \end{cases}$$

W każdym z opisanych przypadków zachodzi równość

$$\phi(MA) = \phi(M) \cdot \phi(A). \quad (\diamond)$$

Dowód. To, że $\phi(MA) = \phi(A)$, dla M dodającej do wiersza inny wiersz przemnożony przez skalar to wniosek z Uwagi 25.1.10. Z Uwagi 25.1.9 wiemy, że dla macierzy A o dwóch identycznych — niekoniecznie sąsiednich wierszach — mamy $\phi(A) = 0$. Stąd dowód tego, że $\phi(MA) = -\phi(A)$ dla macierzy M zamieniającej dwa wiersze miejscami jest analogiczny do dowodu Uwagi 25.1.8 (gdzie zakładaliśmy dodatkowo, że wiersze są sąsiednie). Wreszcie, dowód tego, że $\phi(MA) = c\phi(A)$, dla macierzy M mnożącej wiersz przez stałą c wynika z jednorodności funkcji ϕ względem każdego wiersza. \square

Przypomnijmy, że na poprzednim wykładzie z Twierdzeniem 24.1.7 (czyli z rezultatu wyżej dla $\phi = \det$) wywnioskowaliśmy, że wyznacznik macierzy kwadratowej jest niezerowy wtedy i tylko wtedy, gdy jest ona odwracalna. Dowód ten powtórzyć można bez żadnych zmian dla każdej funkcji ϕ spełniającej warunki (1)-(4) (gdyż wyznacznik macierzy identycznościowej jest niezerowy).

Wniosek 25.1.12

Niech $\phi : M_n(K) \rightarrow K$ będzie funkcją spełniającą warunki (1)-(4). Wówczas $\phi(A) \neq 0$ wtedy i tylko wtedy, gdy A jest macierzą odwracalną.

W kolejnym kroku wykażemy ważny rezultat, będący z jednej strony istotnym wynikiem o wyznaczniku, a z drugiej — narzędziem do uzasadnienia kolejnych rezultatów, w tym Twierdzenia 25.1.6 o jedyności funkcji spełniającej (1)-(4) oraz twierdzenia o wyznaczniku macierzy transponowanej.

Twierdzenie 25.1.13: Wzór Cauchy'ego

Załóżmy, że funkcja $\phi : M_n(K) \rightarrow K$ spełnia warunki (1)-(4) (np. $\phi = \det$, o czym już wiemy). Niech $A, B \in M_n(K)$. Wówczas:

$$\phi(AB) = \phi(A) \cdot \phi(B).$$

Dowód wzoru Cauchy'ego rozbija się na dwa przypadki.

- Przypadek 1. Macierz AB nie jest odwracalna. Zgodnie z Wnioskiem 25.1.2 mamy $\phi(AB) = 0$. Oznacza to, że $\phi(A) = 0$ lub $\phi(B) = 0$. Inaczej ponownie na mocy Wniosku 25.1.2 macierze A, B byłyby odwracalne, a z nimi i AB , bo jak wiadomo $(AB) \cdot B^{-1}A^{-1} = I$.

- Przypadek 2. Założymy, że AB jest odwracalna. Zgodnie z wynikiem wyżej $r(AB) = n$, a wiemy z wcześniejszych wykładów¹, że to oznacza, że $r(A) = n$ oraz $r(B) = n$. Z Twierdzenia 12.4 mamy $\phi(A) \neq 0$ oraz $\phi(B) \neq 0$, więc $\phi(AB) \neq 0$. W szczególności postacią schodkową zredukowaną A oraz B jest I . Mówiąc inaczej: istnieją macierze operacji elementarnych M_1, \dots, M_s oraz N_1, \dots, N_t takie, że

$$I = M_1 M_2 M_3 \dots M_s A, \quad I = N_1 N_2 N_3 \dots N_t B.$$

Zatem $A = M_s^{-1} M_{s-1}^{-1} \dots M_1^{-1}$, $B = N_t^{-1} N_{t-1}^{-1} \dots N_1^{-1}$. Ale M_i^{-1} oraz N_j^{-1} to macierze operacji elementarnych, więc z Obserwacji 25.1.11, a dokładniej formuły (\diamond):

$$\begin{aligned}\phi(AB) &= \phi(M_s^{-1} M_{s-1}^{-1} \dots M_1^{-1} N_t^{-1} N_{t-1}^{-1} \dots N_1^{-1}) = \\ &= \phi(M_s^{-1}) \cdot \phi(M_{s-1}^{-1}) \dots \phi(M_1^{-1}) \cdot \phi(N_t^{-1}) \cdot \phi(N_{t-1}^{-1}) \dots \phi(N_1^{-1}) = \\ &= \phi(M_s^{-1} M_{s-1}^{-1} \dots M_1^{-1}) \cdot \phi(N_t^{-1} N_{t-1}^{-1} \dots N_1^{-1}) = \phi(A)\phi(B).\end{aligned}$$

DOWÓD TWIERDZENIA 25.1.6. Twierdzimy, że wartość funkcji ϕ spełniającej (1)–(4) (w tym, jak wiemy, funkcji $\phi = \det$) jest jednoznacznie wyznaczona, dla każdej macierzy $A \in M_n(K)$. Rzeczywiście:

- Jeśli A nie jest odwracalna, to $\phi(A) = 0$, zgodnie z Wnioskiem 25.1.2.
- Jeśli A jest odwracalna to algorytm Gaussa podaje jednoznaczny, najkrótszy możliwy ciąg operacji elementarnych pozwalających na sprowadzenie A do postaci zredukowanej I . Niech macierze tych operacji to M_1, \dots, M_s . Na mocy wzoru Cauchy'ego:

$$1 = \phi(I) = \phi(M_s)\phi(M_{s-1}) \dots \phi(M_1)\phi(A).$$

Zatem gdy A jest odwracalna, to

$$\det \phi = (\phi(M_s)\phi(M_{s-1}) \dots \phi(M_1))^{-1},$$

gdzie M_1, \dots, M_s jest jednoznacznie wyznaczonym ciągiem macierzy operacji elementarnych. Skoro znamy $\phi(M_i)$, to $\phi(A)$ jest wyznaczona jednoznacznie, co kończy dowód.

Odnotujmy kolejny ważny wniosek ze wzoru Cauchy'ego, kluczowy do dowodu własności (\dagger).

Uwaga 25.1.14

Założymy, że funkcja $\phi : M_n(K) \rightarrow K$ spełnia warunki (1)-(4) (np. $\phi = \det$, o czym już wiemy). Dla każdej macierzy $A \in M_n(K)$ mamy $\phi(A) = \phi(A^T)$.

Dowód. Korzystamy z faktu, że $r(A) = r(A^T)$. Rozważamy dwa przypadki.

- Jeśli $r(A) < n$, to $r(A^T)$, czyli obie macierze nie są odwracalne i ich wartości na ϕ (w szczególności – wyznaczniki) są równe 0.
- Jeśli $r(A) = n$, to A rozkłada się na iloczyn macierzy operacji elementarnych

$$A = M_1 M_2 \dots M_s.$$

Zatem zgodnie ze wzorem $(XY)^T = Y^T X^T$ mamy:

$$A^T = M_s^T M_{s-1}^T \dots M_1^T.$$

Łatwo sprawdzić, że dla każdej macierzy operacji elementarnej M mamy

$$\phi(M) = \phi(M^T).$$

Rzeczywiście, dla macierzy operacji typu (2) i (3) po prostu mamy $M = M^T$. Co do macierzy operacji (1) to przecież M^T jest również macierzą operacji typu (1), a wszystkie te macierze mają wartość ϕ (a więc też wyznacznika) równą 1, zgodnie z Obserwacją 25.1.11. Zatem z twierdzenia Cauchy'ego:

$$\phi(A) = \phi(M_1)\phi(M_2) \dots \phi(M_s) = \phi(M_s^T)\phi(M_{s-1}^T) \dots \phi(M_1^T) = \phi(A^T).$$

¹Przykładowe argumenty: (1) AB to macierz izomorfizmu będącego złożeniem przekształceń o macierzach A oraz B , co oznacza, że A to macierz monomorfizmu, a B – macierz epimorfizmu. Ale te przekształcenia działają pomiędzy przestrzeniami wymiaru n , więc to izomorfizmy., że skoro $A, B \in M_n(K)$. Zatem A, B są odwracalne. (2) Mamy $r(AB) \leq \min\{r(A), r(B)\}$.

□

Pozostało uzasadnić wzór (\dagger) mówiący, że obliczanie wyznacznika za pomocą rozwinięcia względem dowolnego wiersza i dowolnej kolumny daje ten sam wynik. Argumenty są następujące:

- Skoro wyznaczniki macierzy i macierzy transponowanej są równe, to rozwinięcie względem pierwszego wiersza jest identyczną funkcją do rozwinięcia względem pierwszej kolumny.
- Dlaczego rozwinięcie względem i -tej kolumny jest tą samą funkcją, co rozwinięcie względem pierwszej kolumny? Dlatego, że przedstawione wyżej przeprowadzić można zupełnie analogicznie pokazując, że tak zdefiniowana funkcja spełnia (1)-(4). Skoro zaś funkcja spełniająca (1)-(4), jest jedyna, to jest to ta sama funkcja, co rozwinięcie względem pierwszej kolumny, o którym już wiemy, że spełnia (1)-(4).

Wzór (\dagger) jest zatem uzasadniony, z dokładnością do prostych powtórzeń dowodów. Wniosek jest następujący. Licząc wyznaczniki konkretnych macierzy możemy korzystać zamienne z różnych rozwinięć: jeśli na przykład sprowadzimy przez rozwinięcie względem drugiej kolumny obliczenie wyznacznika macierzy 4×4 do obliczenia czterech wyznaczników macierzy 3×3 , to każdy z tych czterech wyznaczników możemy liczyć za pomocą innego rozwinięcia – możemy korzystać zarówno z rozwinięć na wierszach i na kolumnach. Dla ścisłości – całe rozumowanie powyższe należy rozumieć indukcyjnie: najpierw stwierdzamy równość wszystkich (dostępnych) rozwinięć dla macierzy 2×2 , a dalej korzystając z niej i równości (\dagger), możemy postulować równość rozwinięć dla macierzy 3×3 rozumując, że licząc w ramach tych różnych rozwinięć wyznaczniki macierzy typu A_{ij} , możemy już korzystać z dowolnego (dostępnego) rozwinięcia.

* * *

Na koniec tej części wykładu warto wspomnieć, że wzór Cauchy'ego pozwala na uzasadnienie szeregu rezultatów dotyczących wyznaczników, w tym na bardzo eleganckie dowody związane z wyznacznikami macierzy blokowych. Oto dwa przykłady.

Przykład 1. Dla macierzy blokowych $A \in M_{n \times n}(K)$, $B = M_{n \times m}(K)$, $C = M_{m \times m}(K)$, $D = M_{m \times n}(K)$:

$$\begin{bmatrix} 0 & A \\ D & C \end{bmatrix} \cdot \begin{bmatrix} 0 & I_n \\ I_m & 0 \end{bmatrix} = \begin{bmatrix} A & 0 \\ D & C \end{bmatrix}.$$

A zatem korzystając z wzoru Cauchy'ego mamy

$$\begin{vmatrix} 0 & A \\ D & C \end{vmatrix} \cdot \begin{vmatrix} 0 & I_n \\ I_m & 0 \end{vmatrix} = \begin{vmatrix} A & 0 \\ D & C \end{vmatrix} \cdot (-1)^{m+n} = |A| \cdot |C| \Rightarrow \begin{bmatrix} 0 & A \\ D & C \end{bmatrix} = (-1)^{mn} \cdot |A| \cdot |C|.$$

Przykład 2. Niech $A, B, C, D \in M_n(\mathbb{K})$. Wówczas jeśli A jest macierzą odwracalną, to:

$$\begin{vmatrix} A & B \\ C & D \end{vmatrix} = |A| \cdot |D - CA^{-1}B|.$$

Gdy istnieje A^{-1} , wówczas mamy:

$$\begin{bmatrix} I & 0 \\ -CA^{-1} & I \end{bmatrix} \cdot \begin{bmatrix} A & B \\ C & D \end{bmatrix} = \begin{bmatrix} A & B \\ 0 & D - CA^{-1}B \end{bmatrix}.$$

Wyznacznik macierzy blokowo-dolnotrójkątnej o blokach diagonalnych I równy jest 1, natomiast wyznacznik macierzy blokowo-górnotrójkątnej o blokach A oraz $D - CA^{-1}B$ równy jest $|A| \cdot |D - CA^{-1}B|$.

Czytelnika może dziwić, że wychodzi tak skomplikowana równość, a nie na przykład wynik typu $|AD - BC|$ czy $|AD - CB|$. Okazuje się, że żaden z tych wzorów nie musi mieć miejsca. Wystarczy rozważyć choćby macierze A, B, C, D postaci:

$$A = \begin{bmatrix} 1 & -1 \\ 0 & 0 \end{bmatrix}, \quad B = \begin{bmatrix} 1 & 0 \\ -1 & 1 \end{bmatrix}, \quad C = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}, \quad D = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}.$$

Więcej efektywnych zastosowań tego typu znajdzie Czytelnik w zadaniach do niniejszego rozdziału.

25.2 Wybór przykładowych pytań

Proszę się nie uczyć na pamięć odpowiedzi czy wyjaśnienie! Takich pytań można ułożyć znacznie więcej i służą one jedynie sprawdzeniu podstawowego zrozumienia treści zajęć. Jeśli odpowiedź na dowolne z poniższych pytań sprawia Państwu problem, proszę pomyśleć który fragment notatek dotyczy tego pytania, wrócić do ich lektury, a w razie dalszych trudności poprosić prowadzącego o wskazówkę.

1. Czy wyznacznik macierzy jest funkcją jednorodną i addytywną względem każdej kolumny? Czy twierdzenia opisane w powyższym rozdziale są prawdziwe, jeśli zamiast wierszy będziemy rozważać kolumny? Uzasadnij, że wyznacznik macierzy o zerowej kolumnie jest zerowy.
2. Macierz $A \in M_{n \times n}(\mathbb{R})$ ma w każdym wierszu dokładnie jeden niezerowy wyraz, równy x i w każdej kolumnie dokładnie jeden niezerowy wyraz. Ile wynosi $|\det(A)|$?
3. Niech $A \in M_3(K)$ oraz $\det A = 1$. Macierz B powstaje z A przez zamianę pierwszego i ostatniego wiersza. Oblicz $\det(A + B)$ oraz $\det(3 \cdot B)$.
4. Uzasadnij, że jeśli $A \in M_n(K)$ oraz $\lambda \in K$, to $\det(\lambda \cdot A) = \lambda^n \cdot \det A$.
5. Uzasadnij, że jeśli A jest macierzą odwracalną, to
 - (a) $\det(A^{-1}) = (\det A)^{-1}$
 - (b) $\det(A \cdot A^T) > 0$,
 - (c) jeśli $B, C \in M_n(K)$ spełniają $AB = CA$, to $\det(B) = \det(C)$.
6. Macierz odwracalna $X \in M_4(\mathbb{R})$ spełnia równanie $X - 4X^{-1} = 0$. Wyznacz $|\det(X)|$.
7. Założmy, że $A \in M_n(\mathbb{Z})$ jest odwracalna i $A^{-1} \in M_n(\mathbb{Z})$. Jakie są możliwe wartości $\det A$?
8. Macierz B powstaje z macierzy $A \in M_n(K)$ w wyniku następujących operacji
 - (a) od pierwszego wiersza odejmujemy drugi, od drugiego trzeci, od trzeciego odejmujemy pierwotny pierwszy,
 - (b) pierwszą kolumnę przesuwamy na koniec, a pozostałe kolumny przesuwamy w lewo zachowując ich kolejność,
 - (c) pierwsze k wierszy piszemy w odwrotnym porządku i w odwrotnym porządku piszemy ostatnie $n - k$ wierszy,
 - (d) wszystkie wiersze zapisujemy w odwrotnym porządku.

Wyraź $\det B$ w zależności od $\det A$.

9. Czy istnieją macierze $A \in M_{4 \times 3}(\mathbb{R})$ i $B = M_{3 \times 4}(\mathbb{R})$ takie, że $\det(AB) = 1$?
10. Niech suma kolumn macierzy $A \in M_n(K)$ będzie wektorem zerowym. Czy $\det(A) = 0$?
11. Uzasadnij, bez wyliczenia wyznaczników, że zachodzi równość

$$\begin{vmatrix} bc & a^2 & a^2 \\ b^2 & ac & b^2 \\ c^2 & c^2 & ab \end{vmatrix} = \begin{vmatrix} ac & bc & ab \\ bc & ab & ac \\ ab & ac & bc \end{vmatrix}$$

12. Uzasadnij, bez wyliczania, że poniższe wyznaczniki są zerowe

$$\begin{vmatrix} a & b & c \\ 2a & 2b & 2c \\ x & y & z \end{vmatrix}, \quad \begin{vmatrix} 1 & 1 & 1 \\ a & b & c \\ b+c & c+a & a+b \end{vmatrix}, \quad \begin{vmatrix} a^2 & b^2 & c^2 & d^2 \\ (a+1)^2 & (b+1)^2 & (c+1)^2 & (d+1)^2 \\ (a+2)^2 & (b+2)^2 & (c+2)^2 & (d+2)^2 \\ (a+3)^2 & (b+3)^2 & (c+3)^2 & (d+3)^2 \end{vmatrix}.$$

13. Wyrazami macierzy $A \in M_{5 \times 5}(\mathbb{R})$ są 0 i 1, rozmieszczone w taki sposób, że w każdym wierszu występują dokładnie trzy jedynki. Uzasadnij, że $\det(A)$ jest liczbą podzielną przez 3.
14. Niech $A \in M_n(K)$ będzie macierzą odwracalną. Dla jakich n zachodzi $\det(A) + \det(-A) = 0$?
15. Niech $A \in M_n(\mathbb{R})$.
 - (a) Niech n będzie liczbą nieparzystą oraz $A = -A^T$. Uzasadnij, że $\det(A) = 0$.
 - (b) Uzasadnij, że jeśli $A^2 + I = 0$, to n jest liczbą parzystą.
 - (c) Czy jeśli $A \in M_n(\mathbb{C})$, to teza (b) pozostaje prawdziwa?

25.3 Zadania do samodzielnej pracy

Zadania ilustrujące umiejętności niezbędne do zaliczenia przedmiotu oznaczone są symbolem ♠.

Przy tych zadaniach dodany jest opis umiejętności, które sprawdzają.

1. (♠ Wyznacznik iloczynu i wyznacznik macierzy odwrotnej)

Macierz $A \in M_5(\mathbb{R})$ ma wyznacznik równy 2. Oblicz wyznaczniki macierzy:

$$2A, \quad -3A, \quad A^2, \quad -A^3, \quad (A^T)^2.$$

2. (♠ Wyznacznik iloczynu i wyznacznik macierzy odwrotnej)

Oblicz $\det(A \cdot B)$, $\det(A^7)$, $\det(A^3 \cdot B^{-1})$ dla poniższych macierzy:

$$A = \begin{bmatrix} 6 & 1 & 0 & 4 \\ 2 & 0 & 0 & 0 \\ 7 & 0 & 0 & 1 \\ 6 & 0 & 1 & 0 \end{bmatrix}, \quad B = \begin{bmatrix} 0 & 0 & 0 & 2 \\ 0 & 0 & 1 & 2 \\ 0 & 5 & 6 & 6 \\ 1 & 5 & 9 & 7 \end{bmatrix}.$$

3. Policz poniższy wyznacznik mnożąc macierz przez transponowaną do niej.

$$\begin{vmatrix} a & b & c & d \\ -b & a & d & -c \\ -c & -d & a & b \\ -d & c & -b & a \end{vmatrix}$$

4. Korzystając z własności macierzy $\begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}$ uzasadnij, że wyrazy F_0, F_1, F_2, \dots ciągu Fibonacciego spełniają dla każdego $n \in \mathbb{N}$ równość

$$F_{n+1}F_{n-1} - F_n^2 = (-1)^n.$$

5. Wykaż, że jeśli w macierzy $n \times n$ na przecięciu k wierszy i l kolumn znajdują się same zera, przy czym $k + l > n$, to wyznacznik tej macierzy jest równy 0.

6. Dana jest macierz kwadratowa $X \in M_{155}(\mathbb{Q})$, taka że $X^2 = X^T$. Czy wynika z tego, że X jest macierzą identycznościową lub zerową? Jakią są możliwe wartości $\det X$?

7. Pewna macierz $A \in M_{n \times n}(\mathbb{R})$ spełnia równanie $A^2 + A = I_n$. Wykaż, że macierz A jest odwracalna.

8. Niech $n \geq 1$. Założmy, że macierz $A \in M_n(\mathbb{C})$ spełnia $A^T \cdot A = I$ oraz $\det A < 0$. Oblicz $\det(I + A)$.

9. Przypuśćmy, że dla pewnej macierzy A mamy $\det(A) = -3$, $\det(A + I) = 2$, $\det(A + 2I) = 5$, gdzie I jest macierzą identycznościową. Oblicz $\det(A^4 + 3A^3 + 2A^2)$.

10. Niech $A, B \in M_n(\mathbb{R})$ będą macierzami odwracalnymi, gdzie n jest liczbą nieparzystą. Korzystając z porównania wyznaczników AB i BA wykaż, że $AB + BA \neq 0$.

11. Niech $A, B \in M_n(\mathbb{R})$. Rozważmy macierze o wyrazach zespolonych postaci $A + iB, A - iB \in M_n(\mathbb{C})$.

- Wykaż, że $\det(A - iB) = \overline{\det(A + iB)}$.
- Wykaż, że jeśli $AB = BA$, to $\det(A^2 + B^2) \geq 0$.

12. Niech $X \in M_{m \times n}(K)$ oraz $Y \in M_{n \times m}(K)$. Oblicz iloczyny macierzy blokowych

$$\begin{bmatrix} I_n & -Y \\ X & I_m \end{bmatrix} \cdot \begin{bmatrix} I_n & Y \\ 0 & I_m \end{bmatrix}, \quad \text{oraz} \quad \begin{bmatrix} I_n & Y \\ 0 & I_m \end{bmatrix} \cdot \begin{bmatrix} I_n & -Y \\ X & I_m \end{bmatrix}.$$

Wywnioskuj stąd, że zachodzi tożsamość Sylvestera $\det(I_m + XY) = \det(I_n + YX)$.

13. Niech $A \in M_{p \times n}(K)$, $B \in M_{n \times q}(K)$. Oblicz iloczyn macierzy blokowych

$$\begin{bmatrix} I_n & 0 \\ -A & I_p \end{bmatrix} \cdot \begin{bmatrix} I_n & B \\ A & 0 \end{bmatrix} \cdot \begin{bmatrix} I_n & -B \\ 0 & I_p \end{bmatrix}.$$

Wywnioskuj stąd, że zachodzi nierówność Sylvestra $r(A) + r(B) \leq r(AB) + n$.

14. Dane są macierze blokowe $A \in M_{m \times p}(K)$, $B \in M_{m \times q}(K)$, $C \in M_{n \times p}(K)$, $D \in M_{n \times q}(K)$, przy czym spełniony jest warunek $m + n = p + q$. Wykaż, że

$$\det \begin{bmatrix} A & B \\ C & D \end{bmatrix} = (-1)^{mn+pq} \cdot \det \begin{bmatrix} D & C \\ B & A \end{bmatrix}.$$

Rozdział 26

Wyznacznik i układy równań

26.1 Wykład 26

Na tym wykładzie opowiemy o podstawowym zastosowaniu wyznacznika – historycznie rzecz biorąc – źródłowym dla jego powstania, a więc o rozwiązywaniu układów równań oraz zagadnieniach pokrewnych.

Rozważamy układ U złożony n równań liniowych z n niewiadomymi o współczynnikach w ciele K :

$$\begin{cases} a_{11}x_1 + \dots + a_{1n}x_n = b_1 \\ a_{21}x_1 + \dots + a_{2n}x_n = b_2 \\ \vdots \\ a_{n1}x_1 + \dots + a_{nn}x_n = b_n. \end{cases} \quad (\diamond).$$

Wiemy już, że możemy ten układ zapisać w postaci iloczynu macierzy współczynników oraz wektora o współrzędnych złożonych ze zmiennych tak, by wynikiem była macierz o kolumnie z wyrazami b_i :

$$A \cdot x = b \quad (\spadesuit),$$

gdzie

$$A = \begin{bmatrix} a_{11} & \dots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \dots & a_{nn} \end{bmatrix}, \quad x = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}, \quad b = \begin{bmatrix} b_1 \\ \vdots \\ b_n \end{bmatrix}.$$

Równanie typu (\spadesuit) jest przykładem **równania macierzowego**. Rozmaite problemy algebraiczne można formułować w języku tych równań, a ich rozwiązywanie bywa trudne z uwagi na nieprzemienność mnożenia macierzy oraz to, że nie są one zawsze odwracalne (mogą one dotyczyć też macierzy prostokątnych). Rozpoczniemy od fundamentalnej obserwacji, wynikającej z Twierdzenia 12.4.

Uwaga 26.1.1

Następujące warunki są równoważne:

- układ (\diamond) ma dokładnie jedno rozwiązanie,
- $\det A \neq 0$,
- macierz A jest odwracalna.

Gdy zachodzi dowolny z powyższych warunków, to $x = A^{-1}b$.

Obserwacja ta pozwala sformułować **metodę macierzową rozwiązywania układów równań**, których macierz współczynników ma niezerowy wyznacznik.

Zobaczmy przykład. Rozważmy układ równań o współczynnikach w \mathbb{R} postaci:

$$\begin{cases} x_1 + x_2 + x_3 = 6 \\ 2x_2 + 5x_3 = -4 \\ 2x_1 + 5x_2 - x_3 = 27 \end{cases}.$$

Równanie macierzowe równoważne powyższemu układowi ma postać:

$$\begin{bmatrix} 1 & 1 & 1 \\ 0 & 2 & 5 \\ 2 & 5 & -1 \end{bmatrix} \cdot \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 6 \\ -4 \\ 27 \end{bmatrix}.$$

Wyznaczamy teraz macierz odwrotną do macierzy A współczynników. Możemy to zrobić korzystając z algorytmu przedstawionego na poprzednich wykładach, to znaczy: za pomocą elementarnych operacji na wierszach sprowadzić macierz $[A | I]$ do macierzy $[I | A^{-1}]$, uzyskując:

$$\begin{bmatrix} 1 & 1 & 1 \\ 0 & 2 & 5 \\ 2 & 5 & -1 \end{bmatrix}^{-1} = \frac{1}{-21} \begin{bmatrix} -27 & 6 & 3 \\ 10 & -3 & -5 \\ -4 & -3 & 2 \end{bmatrix} \implies \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \frac{1}{-21} \begin{bmatrix} -27 & 6 & 3 \\ 10 & -3 & -5 \\ -4 & -3 & 2 \end{bmatrix} \cdot \begin{bmatrix} 6 \\ -4 \\ 27 \end{bmatrix} = \begin{bmatrix} 5 \\ 3 \\ 2 \end{bmatrix}.$$

Znamy metodę wyznaczania macierzy odwrotnej przy pomocy operacji elementarnych na wierszach. Pokażemy teraz metodę opartą o wyznacznik i tzw. macierz stwarzyszoną (inaczej: dołączoną).

Definicja 26.1.2: Macierz stwarzyszona

Załóżmy, że $A \in M_n(K)$. MACIERZĄ STOWARZYSZONĄ z A definiujemy następująco:

$$\text{adj}(A) = [(-1)^{i+j} \det(A_{ij})]^T = \begin{bmatrix} (-1)^{1+1} |A_{11}| & (-1)^{1+2} |A_{12}| & \dots & (-1)^{1+n} |A_{1n}| \\ (-1)^{2+1} |A_{21}| & (-1)^{2+2} |A_{22}| & \dots & (-1)^{2+n} |A_{2n}| \\ \vdots & \vdots & \ddots & \vdots \\ (-1)^{n+1} |A_{n1}| & (-1)^{n+2} |A_{n2}| & \dots & (-1)^{n+n} |A_{nn}| \end{bmatrix}^T.$$

Przykład.

$$A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}, \quad \text{adj}(A) = \begin{bmatrix} a_{22} & -a_{12} \\ -a_{21} & a_{11} \end{bmatrix}.$$

Zauważmy też, że w powyższym przypadku:

$$\text{adj}(A) \cdot A = \begin{bmatrix} a_{11}a_{22} - a_{12}a_{21} & 0 \\ 0 & a_{11}a_{22} - a_{12}a_{21} \end{bmatrix} = |A| \cdot \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}.$$

Twierdzenie 26.1.3

Zachodzi równość $\text{adj } A \cdot A = \det(A) \cdot I_n$. W szczególności, jeśli A jest macierzą odwracalną, to

$$A^{-1} = \frac{1}{\det A} \cdot \text{adj}(A).$$

Dowód. Niech $A = [a_{ij}]$, dla $1 \leq i, j \leq n$. Mnożymy $\text{adj}(A)$ przez A , czyli

$$\begin{bmatrix} (-1)^{1+1} \det A_{11} & (-1)^{2+1} \det A_{21} & \dots & (-1)^{n+1} \det A_{n1} \\ (-1)^{1+2} \det A_{12} & (-1)^{2+2} \det A_{22} & \dots & (-1)^{2+n} \det A_{n2} \\ \vdots & \vdots & \ddots & \vdots \\ (-1)^{n+1} \det A_{1n} & (-1)^{n+2} \det A_{2n} & \dots & (-1)^{n+n} \det A_{nn} \end{bmatrix} \cdot \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{bmatrix}$$

Wymnóżmy i -ty wiersz $\text{adj}(A)$ oraz j -tą kolumnę w A . Mamy:

$$(-1)^{i+1} \det A_{1i} \cdot a_{1j} + (-1)^{i+2} \det A_{2i} \cdot a_{2j} + \dots + (-1)^{n+i} \det A_{ni} \cdot a_{nj}, \quad (\dagger)$$

To wyrażenie wygląda prawie jak wzór na wyznacznik w rozwinięciu Laplace'a względem i -tej kolumny macierzy A z tym, że zamiast wyrazów z i -tej kolumny macierzy A w poszczególnych składnikach pojawiają się wyrazy z j -tej kolumny. Możemy jednak powiedzieć, że (\dagger) to wyznacznik macierzy D_{ij} powstającej z A przez zastąpienie i -tej kolumny kolumną j -tą (wystarczy policzyć wyznacznik D_{ij} rozwiązując względem i -tej kolumny). Zauważmy jednak, że jeśli $i \neq j$, to D_{ij} ma dwie identyczne kolumny, czyli:

$$\det D_{ij} = \begin{cases} \det A, & \text{dla } i = j \\ 0, & \text{dla } i \neq j. \end{cases}$$

W rezultacie:

$$\text{adj}(A) \cdot A = \begin{bmatrix} \det A & 0 & \dots & 0 \\ 0 & \det A & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \det A \end{bmatrix},$$

co kończy dowód. \square

Jesteśmy również gotowi do sformułowania wzorów pozwalających na uzyskanie rozwiązania równania (\diamond) w przypadku, gdy jest ono jedyne.

Twierdzenie 26.1.4: Wzory Cramera

Niech U będzie układem n równań liniowych z n niewiadomymi

$$\begin{cases} a_{11}x_1 + \dots + a_{1n}x_n = b_1 \\ a_{21}x_1 + \dots + a_{2n}x_n = b_2 \\ \vdots \\ a_{n1}x_1 + \dots + a_{nn}x_n = b_n. \end{cases} \quad (\diamond).$$

o macierzy współczynników $A \in M_n(K)$ i kolumnie wyrazów wolnych $B \in M_{n \times 1}(K)$. Założymy, że $\det A \neq 0$. Wówczas układ U ma dokładnie jedno rozwiązanie s_1, \dots, s_n , przy czym dla każdego i mamy

$$s_i = \frac{\det G_i}{\det A},$$

gdzie G_i jest macierzą powstałą z A przez zastąpienie i -tej kolumny kolumną B .

Zobaczmy, dla przykładu, układ równań nad \mathbb{Q} postaci

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

Jeśli A jest macierzą współczynników tego układu to to zgodnie z definicją G_i oraz wzorami Cramera:

$$|A| = \begin{vmatrix} 1 & 1 \\ 1 & -1 \end{vmatrix}, \quad |G_1| = \begin{vmatrix} 2 & 1 \\ 0 & -1 \end{vmatrix}, \quad |G_2| = \begin{vmatrix} 1 & 2 \\ 1 & 0 \end{vmatrix} \Rightarrow x = \frac{|G_1|}{|A|} = 1, \quad y = \frac{|G_2|}{|A|} = 1.$$

Dowód. Jak wiemy z metody macierzowej, aby rozwiązać równanie $AX = B$ powstałe z równania (\diamond) należy wykonać następujące mnożenie:

$$A^{-1}A \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = A^{-1} \begin{bmatrix} b_1 \\ \vdots \\ b_n \end{bmatrix} \Rightarrow \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = A^{-1} \begin{bmatrix} b_1 \\ \vdots \\ b_n \end{bmatrix}.$$

Ale macierz A^{-1} ma wyrazy c_{ij} postaci:

$$(-1)^{j+i} \frac{\det A_{ji}}{\det A},$$

czyli jeśli G_i to macierz powstała z A przez zamianę i -tej kolumny na B , to argumentując podobnie jak w poprzednim dowodzie widzimy, że iloczyn i -tego wiersza macierzy A^{-1} przez kolumnę B równy jest

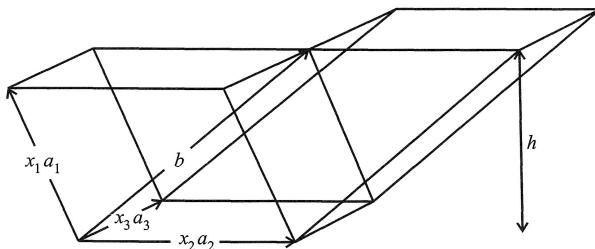
$$\begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = \frac{1}{\det A} \begin{bmatrix} (-1)^{1+1} \det A_{11} & (-1)^{2+1} \det A_{21} & \dots & (-1)^{n+1} \det A_{n1} \\ (-1)^{1+2} \det A_{12} & (-1)^{2+2} \det A_{22} & \dots & (-1)^{2+n} \det A_{n2} \\ \vdots & \vdots & \ddots & \vdots \\ (-1)^{1+n} \det A_{1n} & (-1)^{2+n} \det A_{2n} & \dots & (-1)^{n+n} \det A_{nn} \end{bmatrix} \begin{bmatrix} b_1 \\ \vdots \\ b_n \end{bmatrix},$$

czyli ostatnia równość ma postać:

$$\begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} \sum_{s=1}^n (-1)^{s+1} \det A_{s1} b_s \\ \det A \\ \vdots \\ \sum_{s=1}^n (-1)^{s+n} \det A_{sn} b_s \\ \det A \end{bmatrix} = \begin{bmatrix} \sum_{s=1}^n (-1)^{s+1} \det(G_1)_{s1} b_s \\ \det A \\ \vdots \\ \sum_{s=1}^n (-1)^{s+n} \det(G_n)_{sn} b_s \\ \det A \end{bmatrix} = \begin{bmatrix} \frac{\det G_1}{\det A} \\ \vdots \\ \frac{\det G_n}{\det A} \end{bmatrix}.$$

□

Wzory Cramera mają ciekawą interpretację geometryczną w przestrzeni trójwymiarowej. Niech $A \in M_3(\mathbb{R})$ będzie macierzą odwracalną o kolumnach a_1, a_2, a_3 i rozważmy wektor $b \in \text{lin}(a_1, a_2, a_3)$ tak, że układ $Ax = b$ ma dokładnie jedno rozwiązanie. Innymi słowy, istnieją jednoznacznie wyznaczone $x_1, x_2, x_3 \in \mathbb{R}$ takie, że $x_1 a_1 + x_2 a_2 + x_3 a_3 = b$. Przyjmijmy, dla uproszczenia, że $\det A > 0$, $x_1, x_2, x_3 > 0$ i rozważmy równoległościany $R = R(x_1 a_1, x_2 a_2, x_3 a_3)$ oraz $R_1 = R(b, x_2 a_2, x_3 a_3)$:



Rysunek. Źródło: A Geometric Interpretation of Cramer's Rule, Gregory Conner and Michael Lundquist

Równoległościan $R(x_2 a_2, x_3 a_3)$ traktować możemy jako wspólną podstawę obydwu tych równoległościanów. Mają one również wspólną wysokość opuszczoną na tę podstawę. Stąd, na mocy naszej (intuicyjnej na razie) interpretacji wyznacznika jako objętości (z dokładnością do wartości bezwzględnej, ale korzystamy z $x_1, x_2, x_3 > 0$):

$$\det[x_1 a_1 \ x_2 a_2 \ x_3 a_3] = \det[b \ x_2 a_2 \ x_3 a_3].$$

Nietrudno wywnioskować stąd wzór Cramera, korzystając z jednorodności i addytywności względem kolumn macierzy (patrz też Pytanie 9 w „Wyborze przykładowych pytań”).

Wzory Cramera można w łatwy sposób uogólniać na przypadek, układu m równań z n zmiennymi, gdzie $m < n$. Oto przykład takiego twierdzenia.

Twierdzenie 26.1.5: Wzory Cramera — przykładowe uogólnienie (ćwiczenie)

Niech U będzie układem m równań liniowych z n niewiadomymi, gdzie $m < n$

$$\begin{cases} a_{11}x_1 + \dots + a_{1n}x_n = b_1 \\ \vdots \\ a_{m1}x_1 + \dots + a_{mn}x_n = b_m. \end{cases}.$$

o macierzy współczynników $A \in M_{m \times n}(K)$ i kolumnie wyrazów wolnych $B \in M_{m \times 1}(K)$. Niech G będzie macierzą powstałą z A przez wykreślenie ostatnich $m - n$ kolumn. Założymy, że $\det G \neq 0$. Wówczas zbiór rozwiązań układu U składa się z n -tek $(s_1, \dots, s_m, s_{m+1}, \dots, s_n)$, gdzie s_{m+1}, \dots, s_n są dowolnymi elementami ciała K , zaś dla $1 \leq i \leq m$ mamy

$$s_i = \frac{\det G_i}{\det G} - \sum_{j=m+1}^n \frac{\det G'_{ij}}{\det G} s_j,$$

gdzie G_i jest macierzą powstałą z G przez zastąpienie i -tej kolumny kolumną B , zaś G'_{ij} jest macierzą powstałą z G przez zastąpienie i -tej kolumny j -tą kolumną macierzy A .

Na koniec omówimy jedno zagadnienie, mające szereg zastosowań, a związane również z teorią wielomianów. **Zadanie interpolacyjne Lagrange'a** polega na znalezieniu dla danej funkcji $f : K \rightarrow K$ wielomianu $P_n \in K[x]$ stopnia co najwyżej n , którego wartości dla $n+1$ z góry zadanych parami różnych elementów x_0, \dots, x_n ciała K są takie same, jak wartości interpolowanej funkcji, tzn.

$$P_n(x_i) = f(x_i), \quad \text{dla } i = 0, 1, \dots, n.$$

Twierdzenie 26.1.6

Zadanie interpolacyjne Lagrange'a ma dokładnie jedno rozwiązanie. Mianowicie konstruując funkcje:

$$p_i(x) = \prod_{\substack{j=0 \\ j \neq i}}^n \frac{x - x_j}{x_i - x_j}. \quad i = 0, 1, 2, \dots, n,$$

określamy rozwiązanie zadania interpolacyjnego wzorem:

$$P_n(x) = f(x_0)p_0(x) + f(x_1)p_1(x) + \dots + f(x_n)p_n(x) = \sum_{i=0}^n f(x_i) \prod_{\substack{j=0 \\ j \neq i}}^n \frac{x - x_j}{x_i - x_j}. \quad (\heartsuit)$$

Dowód. Nietrudno widzieć, że $p_i(x)$ to wielomian stopnia n takie, że¹:

$$p_i(x_j) = \begin{cases} 1, & \text{dla } i = j \\ 0, & \text{dla } i \neq j. \end{cases}$$

Stąd $P_n(x)$ jest wielomianem stopnia co najwyżej n przyjmującym w punktach x_i wartości $f(x_i)$, czyli jest rozwiązaniem problemu interpolacyjnego.

Jednoznaczność wynika natychmiast z Twierdzenia 5.4.1, które mówi, że wielomian stopnia n o współczynnikach w ciele K ma nie więcej niż n parami różnych pierwiastków. Gdyby jakiś wielomian P'_n stopnia nie większego od n również spełniał zadanie interpolacyjne, wówczas $P_n(x) - P'_n(x)$ byłby wielomianem stopnia n o $n+1$ pierwiastkach x_0, \dots, x_n , licząc krotności, co implikuje, że $P_n(x) = P'_n(x)$. \square

Wykażemy teraz, że jednoznaczność istnienia wielomianu interpolacyjnego można dostać również bez Twierdzenia 5.4.1. Istotnie, rozważmy układ $n+1$ równań, w którym niewiadomymi są współczynniki wielomianu

$$P_n = a_0 + a_1x + a_2x^2 + \dots + a_{n-1}x^{n-1} + a_nx^n,$$

spełniającego dla pewnych z góry zadanych $f(x_0), \dots, f(x_n)$ warunki:

$$\begin{cases} P_n(x_0) = a_0 + a_1x_0 + a_2x_0^2 + \dots + a_{n-1}x_0^{n-1} + a_nx_0^n & = f(x_0) \\ P_n(x_1) = a_0 + a_1x_1 + a_2x_1^2 + \dots + a_{n-1}x_1^{n-1} + a_nx_1^n & = f(x_1) \\ \vdots \\ P_n(x_n) = a_0 + a_1x_n + a_2x_n^2 + \dots + a_{n-1}x_n^{n-1} + a_nx_n^n & = f(x_n) \end{cases}.$$

Policzmy wyznacznik macierzy współczynników tego układu – zwany WYZNACZNIKIEM VANDERMONDE'A.

$$\Delta(x_0, \dots, x_n) = \begin{vmatrix} 1 & x_0 & x_0^2 & \dots & x_0^{n-1} & x_0^n \\ 1 & x_1 & x_1^2 & \dots & x_1^{n-1} & x_1^n \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 1 & x_n & x_n^2 & \dots & x_n^{n-1} & x_n^n \end{vmatrix}.$$

¹Mówiąc dokładniej, układ wielomianów p_i jest bazą przestrzeni wielomianów stopnia nie większego niż n , zaś współrzędne dowolnego innego wielomianu $f \in K_{\leq n}[X]$ w tej bazie to $f(x_1), f(x_2), \dots, f(x_n)$. To powinno się kojarzyć z bazą sprzężoną. Rzeczywiście, rozważając liniowo niezależne funkcjonały $f \xrightarrow{\mu_i} f(x_i)$ w $(K_n[X])^*$ widzimy, że stanowią one bazę sprzężoną do zaprezentowanego wyżej układu wielomianów $p_i \in K_{\leq n}[X]$.

Twierdzenie 26.1.7

Zachodzi równość $\Delta(x_0, \dots, x_n) = \prod_{0 \leq i < j \leq n} (x_j - x_i)$.

Czy Czytelnik widzi, że wyznacznik Vandermonde'a pojawia się w określeniu wielomianów $p_i(x)$? Formuła (\heartsuit) nie ujawnia współczynników wielomianu interpolacyjnego, ale teraz widzimy, że mogą być one wyznaczone z wzorów Cramera. Widzimy tu duże podobieństwo do iloczynu postaci $A^{-1}b$, rozważanego w dowodzie wzorów Cramera. Kluczowy wniosek jest tu taki: wyznaczenie $\Delta(x_0, \dots, x_n)$ zapewnia egzystencjalny dowód istnienia wielomianu interpolacyjnego, bez „zgadywania” wielomianów p_i .

Dowód. Indukcja ze względu na liczbę n . Dla $n = 1$ mamy: $\det \begin{bmatrix} 1 & x_0 \\ 1 & x_1 \end{bmatrix} = x_1 - x_0$. Niech $n > 1$. Idea jest taka, by rozbić macierz Vandermonde'a na iloczyn macierzy i skorzystać ze wzoru Cauchy'ego i założenia indukcyjnego. Dokładniej, wystarczy pokazać, że:

$$\Delta(x_0, \dots, x_n) = (x_1 - x_0)(x_2 - x_0) \cdots (x_n - x_0) \cdot \Delta(x_1, \dots, x_n). \quad (\spadesuit)$$

Bierzemy macierz Vandermonde'a i odejmujemy pierwszy wiersz od pozostałych. Zrobimy to za pomocą mnożenia macierzy, żeby Czytelnik mógł się przekonać, że nie tylko macierze operacji elementarnych wykonują pewne operacje na wierszach macierzy, przez które pomnożyliśmy je z prawej strony:

$$\begin{bmatrix} 1 & 0 & 0 & \dots & 0 \\ -1 & 1 & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ -1 & 0 & 0 & \dots & 1 \end{bmatrix} \cdot \begin{bmatrix} 1 & x_0 & x_0^2 & \dots & x_0^{n-1} & x_0^n \\ 1 & x_1 & x_1^2 & \dots & x_1^{n-1} & x_1^n \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 1 & x_n & x_n^2 & \dots & x_n^{n-1} & x_n^n \end{bmatrix} = \begin{bmatrix} 1 & x_0 & x_0^2 & \dots & x_0^{n-1} & x_0^n \\ 0 & x_1 - x_0 & x_1^2 - x_0^2 & \dots & x_1^{n-1} - x_0^{n-1} & x_1^n - x_0^n \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & x_n - x_0 & x_n^2 - x_0^2 & \dots & x_n^{n-1} - x_0^{n-1} & x_n^n - x_0^n \end{bmatrix}.$$

Macierz po prawej jest dolnotrójkątna i ma wyznacznik równy 1, co zgadza się z formułą Cauchy'ego i obserwacją mówiącą, że ciąg operacji typu (1) nie zmienia wyznacznika. Idźmy dalej. Uzyskana macierz jest blokowo górnortrójkątna, a zatem mamy (inaczej mówiąc: rozwijając względem pierwszej kolumny):

$$\Delta(x_0, \dots, x_n) = \det \begin{bmatrix} x_1 - x_0 & x_1^2 - x_0^2 & \dots & x_1^{n-1} - x_0^{n-1} & x_1^n - x_0^n \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ x_n - x_0 & x_n^2 - x_0^2 & \dots & x_n^{n-1} - x_0^{n-1} & x_n^n - x_0^n \end{bmatrix}$$

Teraz przedstawimy powyższą macierz w postaci iloczynu trzech macierzy. Po pierwsze wyciągamy wspólne czynniki $x_i - x_0$ z każdego wiersza i korzystając ze wzorów skróconego mnożenia mamy:

$$\begin{bmatrix} x_1 - x_0 & 0 & \dots & 0 \\ 0 & x_2 - x_0 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & x_n - x_0 \end{bmatrix} \cdot \begin{bmatrix} 1 & x_1 + x_0 & \dots & \sum_{i=0}^{n-1} x_1^{n-1-i} x_0^i \\ 1 & x_2 + x_0 & \dots & \sum_{i=0}^{n-1} x_2^{n-1-i} x_0^i \\ \vdots & \vdots & \ddots & \vdots \\ 1 & x_n + x_0 & \dots & \sum_{i=0}^{n-1} x_n^{n-1-i} x_0^i \end{bmatrix}.$$

Macierz po prawej wygląda nieco nieprzyjemnie, ale to się zmieni po rozbiciu jej na następujący iloczyn:

$$\begin{bmatrix} 1 & x_1 & x_1^2 & \dots & x_1^{n-1} \\ 1 & x_2 & x_2^2 & \dots & x_2^{n-1} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_n & x_n^2 & \dots & x_n^{n-1} \end{bmatrix} \cdot \begin{bmatrix} 1 & x_0 & x_0^2 & \dots & x_0^{n-1} \\ 0 & 1 & x_1 & \dots & x_0^{n-2} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & 1 \end{bmatrix}.$$

A zatem pokazaliśmy, że wyznacznik Vandermonde'a $\Delta(x_0, \dots, x_n)$ równy jest w istocie iloczynowi wyznaczników trzech macierzy:

- macierz diagonalnej o wyrazach $x_i - x_0$,
- macierz Vandermonde'a o wyznaczniku $\Delta(x_1, \dots, x_n)$,
- macierz górnortrójkątnej mającej jedynki na przekątnej.

A zatem z formuły Cauchy'ego mamy (\spadesuit). □

Widzimy zatem, że wyznacznik Vandermonde'a jest niezerowy wtedy i tylko wtedy, gdy liczby x_0, \dots, x_n są parami różne. Obserwacja ta jest przydatna w wielu rozważaniach algebraicznych.

26.2 Wybór przykładowych pytań

Proszę się nie uczyć na pamięć odpowiedzi czy wyjaśnienie! Takich pytań można ułożyć znacznie więcej i służą one jedynie sprawdzeniu podstawowego zrozumienia treści zajęć. Jeśli odpowiedź na dowolne z poniższych pytań sprawia Państwu problem, proszę pomyśleć który fragment notatek dotyczy tego pytania, wrócić do ich lektury, a w razie dalszych trudności poprosić prowadzącego o wskazówkę.

- Zapisz w postaci macierzowej układ równań liniowych:

$$\begin{cases} x_1 + 3x_2 - 2x_3 = 2 \\ 3x_1 + 9x_2 - 2x_3 = 2 \\ 2x_1 + 6x_2 + x_3 = 4 \end{cases} .$$

- Rozwiąż za pomocą metody macierzowej układ równań

$$\begin{cases} x_1 + 2x_2 = 4 \\ 3x_1 - 5x_2 = 1 \end{cases} .$$

- Znajdź $\text{adj}(A)$, gdzie A równa jest jednej z poniższych macierzy

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix}, \quad \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{bmatrix}, \quad \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix} .$$

- Rozważmy macierz

$$A = \begin{bmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{bmatrix}$$

Uzasadnij, że $b_1 \det A_{11} + b_2 \det A_{12} + b_3 \det A_{13} = 0$.

- Znajdź funkcję wielomianową taką, że jej wykres na płaszczyźnie kartezjańskiej przechodzi przez punkty

$$(1, 3), \quad (3, 4), \quad (5, 6), \quad (7, -10).$$

- Uzasadnij, naśladowując dowód z wykładu, że

$$\begin{vmatrix} 1 & a & a^2 \\ 1 & b & b^2 \\ 1 & c & c^2 \end{vmatrix} = (b-a)(c-a)(c-b).$$

- Czy istnieje taka macierz $A \in M_3(\mathbb{R})$, że $\text{adj}(A)$ ma tylko jeden niezerowy wyraz?

- Uzasadnij, że jeśli $A \in M_n(K)$ jest macierzą górnoprójkątną, to również macierz $\text{adj}(A)$ jest górnoprójkątna. Czy macierz odwrotna do górnoprójkątnej jest górnoprójkątna?

- (Inny dowód wzorów Cramera) Zapiszmy macierz $A \in M_n(K)$ w postaci $A = [A_1 \ A_2 \ \dots \ A_n]$, gdzie $A_i \in M_{n \times 1}(K)$ są kolumnami A . Niech $Ax = b$ będzie układem równań liniowych zmiennych x_1, \dots, x_n , gdzie $b \in M_{n \times 1}(K)$.

- Uzasadnij, że $Ax = x_1 A_1 + \dots + x_n A_n = b$.
- Oblicz wyznacznik

$$|b \ A_2 \ A_3 \ \dots \ A_n|$$

wstawiając za b kombinację $x_1 A_1 + \dots + x_n A_n$ i korzystając z addytywności i jednorodności wyznacznika względem pierwszej kolumny oraz korzystając z tego, że wyznacznik macierzy o dwóch identycznych kolumnach równy jest 0. Co widzisz?

- Uzasadnij, że

- jeśli macierz A jest odwracalna, to $\text{adj}(A^{-1}) = \frac{1}{\det A} \text{adj}(A)$,
- $\text{adj}(A^T) = (\text{adj}(A))^T$

- Niech $\phi, \psi : K^n \rightarrow K^n \in L(K^n, K^n)$, takimi że jeśli $A = M(\phi)^{st}$, to $M(\psi)^{st} = \text{adj}(A)$.

- Uzasadnij, że $\phi \circ \psi$ jest homotetią.
- Uzasadnij, że jeśli ϕ nie jest izomorfizmem, to $\ker \psi \subseteq \text{im } \phi$ oraz $\ker \phi \subseteq \text{im } \psi$.

26.3 Zadania do samodzielnej pracy

Zadania ilustrujące umiejętności niezbędne do zaliczenia przedmiotu oznaczone są symbolem ♠. Przy tych zadaniach dodany jest opis umiejętności, które sprawdzają.

- (♠ Wyznacznikowe kryterium odwracalności i wzór na macierz odwrotną)
Dla jakich wartości parametru $s \in \mathbb{R}$ poniższa macierz

$$A = \begin{bmatrix} 2 & 5 & 3 \\ 1 & s & 2 \\ 1 & 2 & 1 \end{bmatrix}$$

jest odwracalna? Dla każdego takiego s znajdź A^{-1} .

- (♠ Stosowanie wzorów Cramera)

Przedyskutuj rozwiązywalność następującego układu równań, w zależności od parametru $m \in \mathbb{R}$:

$$3x + z = 1, \quad mx - my + z = -m, \quad x + my + z = 3.$$

Dla tych m , dla których istnieją rozwiązania, wyznacz je.

- (♠ Stosowanie wzorów Cramera)

Dane są liczby rzeczywiste a, b, c . Rozwiąż układ równań liniowych postaci

$$\begin{cases} (b+c)x_1 + bx_2 + cx_3 = 1 \\ ax_1 + (c+a)x_2 + cx_3 = 1 \\ ax_1 + bx_2 + (a+b)x_3 = 1. \end{cases}$$

- Wyznacz macierz A^{-1} , gdzie $A = [a_{ij}] \in M_n(\mathbb{R})$, przy czym:

$$a_{ij} = \begin{cases} 1, & \text{dla } i \geq j \\ 0, & \text{dla } i < j. \end{cases}$$

- Znajdź macierze odwrotne do macierzy $n \times n$ postaci:

$$\begin{bmatrix} 1 & 1 & 1 & \dots & 1 \\ 1 & 2 & 2 & \dots & 2 \\ 1 & 2 & 3 & \dots & 3 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & 2 & 3 & \dots & n \end{bmatrix}, \quad \begin{bmatrix} 0 & 1 & 1 & \dots & 1 \\ 1 & 0 & 1 & \dots & 1 \\ 1 & 1 & 0 & \dots & 1 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & 1 & 1 & \dots & 0 \end{bmatrix}, \quad \begin{bmatrix} 0 & a_1 & 0 & \dots & 0 \\ 0 & 0 & a_2 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & a_{n-1} \\ a_n & 0 & 0 & \dots & 0 \end{bmatrix},$$

przy czym zakładamy, że $a_1, \dots, a_n \neq 0$.

- Załóżmy, że układ n równań liniowych U o n zmiennych i o współczynnikach całkowitych ma dla dowolnej liczby pierwszej p dokładnie jedno rozwiązanie w ciele \mathbb{Z}_p (współczynniki traktujemy modulo p). Czy układ równań U posiada rozwiązanie, którego wszystkie współrzędne są całkowite?

- Załóżmy, że $A \in M_n(\mathbb{Z})$ oraz $\det A = \pm 1$. Jaki może być $\det A^{-1}$? Wykaż, że $A^{-1} \in M_n(\mathbb{Z})$. Czy teza pozostaje prawdziwa, gdy $\det A = 2$?

- Wykaż następujące fakty o macierzy $\text{adj}(A)$, gdzie $A \in M_n(\mathbb{R})$.

- (a) $r(A) = n - 1 \iff r(\text{adj}(A)) = 1$,
- (b) $r(A) < n - 1 \iff r(\text{adj}(A)) = 0$,
- (c) $\det(\text{adj}(A)) = \det(A)^{n-1}$,
- (d) $\text{adj}(\text{adj}(A)) = \det(A)^{n-2}A$,
- (e) $\text{adj}(AB) = \text{adj}(B) \cdot \text{adj}(A)$.

- Niech $\lambda_1, \dots, \lambda_n$ będą liczbami zespolonymi o tej własności, że dla każdej liczby całkowitej $k > 0$ zachodzi równość $\sum_{i=1}^n \lambda_i^k = 0$. Wykaż, że $\lambda_i = 0$, dla $i = 1, 2, \dots, n$.

- Wykaż, że dla dowolnych liczb całkowitych a_1, \dots, a_n poniższa liczba jest całkowita

$$\prod_{1 \leq i < j \leq n} \frac{a_i - a_j}{i - j}.$$

Rozdział 27

Wzór permutacyjny. Minory

27.1 Wykład 27

Ostatni wykład pierwszego semestru poświęcimy alternatywnemu spojrzeniu na wyznacznik, opartemu na pojęciu permutacji. Jest ono rzeczywiście alternatywne, bowiem w wielu źródłach wyznacznik definiowany jest właśnie za pomocą formuły, którą za chwilę wprowadzimy – tzw. wzoru Leibniza, inaczej nazywanego wzorem permutacyjnym. Idea jest prosta – zamiast skomplikowanej rekurencyjnej definicji dostajemy otwarty wzór, pozwalający zresztą wykazać łatwo wiele podstawowych, znanych nam już własności wyznacznika. Kłopot stanowi dość skomplikowana postać wzoru, wymagająca wstępnych wyjaśnień i stosunkowo złożonej notacji. Zapewne nie będą Państwo często liczyć wyznacznika właśnie za pomocą tej formuły. Wielokrotnie jednak definiować będziemy w drugim semestrze funkcje, właśnie w oparciu o wyznaczniki. Będą to najpierw pewne niezmienniki przekształceń liniowych, a później też niezmienniki przestrzeni liniowych wyposażonych w dodatkową strukturę (i przekształceń liniowych zachowujących te dodatkowe struktury). Rozumienie własności tych funkcji ma duże znaczenie w algebrze linowej.

Nie odejdziemy więc zupełnie od kontekstu geometrycznego. Z jednej strony domykamy więc teorię klużcowego dla nas pojęcia algebraicznego (rozwiązywalność układów równań, odwrocalność macierzy, bycie macierzą izomorfizmu), z drugiej — przygotowujemy grunt do pracy, która dopiero przez nam.

Przypomnijmy najpierw rozważaną ostatnio notację kolumnową, którą będziemy dziś intensywnie stosować. Niech A_1, \dots, A_n będą kolumnami macierzy $A = [a_{ij}] \in M_{m \times n}(K)$. Wówczas pisać będziemy

$$A = [A_1, \dots, A_n].$$

Kolumny macierzy identycznościowej I_n oznaczamy (kolejno) przez E_1, \dots, E_n .

Formalnie rzecz biorąc, o kolumnach A_i myślimy jako o elementach K^m (czasem też jak o macierzach rozmiaru $m \times 1$), ponieważ jeśli weźmiemy $\alpha_i = (a_{1i}, \dots, a_{mi})$, to odwzorowanie przypisujące n -ce $(\alpha_1, \dots, \alpha_n)$ macierz $[a_{ij}]$ jest izomorfizmem przestrzeni liniowych $K^m \times \dots \times K^m$ (złożonej z n -tek wektorów w K^m z naturalnymi działaniami po współrzędnych) oraz $M_{m \times n}(K)$.

Będziemy korzystać z tego, że wyznacznik jest jedyną funkcją $M_n(K) \rightarrow K$ spełniającą warunki: (1) jednorodność ze względu na k -tą kolumnę, (2) addytywność względem k -tej kolumny, (3) funkcja zeruje się, jeśli macierz ma identyczne dwie (sąsiednie) kolumny, (4) przyjmuje wartość 1 na macierzy I_n . Innymi słowy, dla każdego $1 \leq k \leq n$ mamy równości (świadczące o "n-liniowości" wyznacznika jako funkcji na kolumnach macierzy)

$$\det[A_1, \dots, A_{k-1}, \mathbf{B} + \mathbf{C}, A_{k+1}, \dots, A_n] = \det[A_1, \dots, A_{k-1}, \mathbf{B}, A_{k+1}, \dots, A_n] + \det[A_1, \dots, A_{k-1}, \mathbf{C}, A_{k+1}, \dots, A_n],$$
$$\det[A_1, \dots, A_{k-1}, \mathbf{a}\mathbf{C}, A_{k+1}, \dots, A_n] = \mathbf{a} \cdot \det[A_1, \dots, A_{k-1}, \mathbf{C}, A_{k+1}, \dots, A_n].$$

Czytelnik zapyta — czy nie mówiliśmy dotąd raczej o jednorodności i addytywności na wierszach? Owszem, ale analogiczne własności zachodzą dla kolumn, gdyż $\det X = \det X^T$ (i w ogóle każda funkcja $\phi : M_n(K) \rightarrow K$ spełniająca warunki (1)-(3) dla wierszy, spełnia je też dla kolumn).

Do naszych rozważań niezbędna będzie definicja permutacji, podana zostały już w jednym z dodatków do wcześniejszych wykładów. W zasadzie na nasze potrzeby o permutacjach będziemy myśleć przede wszystkim przez pryzmat tzw. macierzy permutacji.

Definicja 27.1.1: Permutacje i macierze permutacji, znak permutacji

Przez S_n oznaczać będziemy zbiór wszystkich bijekcji (funkcji różnowartościowych i „na”) zbioru n -elementowego $\{1, 2, \dots, n\}$. Funkcje te nazywamy PERMUTACJAMI zbioru n -elementowego.

Dla dowolnej permutacji $\sigma \in S_n$ określamy macierz $P_\sigma \in M_n(K)$, zwaną MACIERZĄ PERMUTACJI σ , której i -ta kolumna stanowi $\sigma(i)$ -ty wektor bazy standardowej przestrzeni K^n . Podzbiór $M_n(K)$ złożony ze wszystkich macierzy permutacji zbioru n -elementowego również oznaczamy przez S_n .

Dla dowolnej permutacji σ , przez $\text{sgn}(\sigma) \in \{-1, 1\}$ oznaczamy wyznacznik macierzy permutacyjnej $\det P_\sigma$. Liczbę tą nazywamy ZNAKIEM PERMUTACJI σ .

Oczywiście macierz P_{id} permutacji identycznościowej zbioru n -elementowego, to macierz identycznościowa I_n . Innymi słowy o macierzy P_σ myśleć można jako o macierzy przekształcenia liniowego przeprowadzającego i -ty wektor ϵ_i bazy standardowej przestrzeni liniowej K^n na wektor $\epsilon_{\sigma(i)}$. Stąd jest jasne, że jeśli $\sigma, \tau \in S_n$, to

$$P_{\sigma \circ \tau} = P_\sigma P_\tau \quad \text{oraz} \quad \text{sgn}(\sigma \circ \tau) = \text{sgn}(\sigma) \cdot \text{sgn}(\tau).$$

Skoro zaś każda macierz permutacyjna powstaje przez pewną liczbę zamian kolumn (warto to uzasadnić!) macierzy identycznościowej, to istotnie $\text{sgn}(\sigma) \in \{-1, 1\}$. W zadaniach do tego rozdziału podajemy bardziej kombinatoryczną interpretację permutacji i znaku, związaną z rozkładem na tzw. transpozycje. Permutacje o znaku 1 nazywamy *parzystymi*, a permutacje o znaku -1 nazywamy *nieparzystymi*.

Przykład. Rozważmy permutację $\sigma \in S_4$, czyli bijekcję $\sigma : \{1, 2, 3, 4\} \rightarrow \{1, 2, 3, 4\}$, daną wzorem:

$$\sigma(1) = 3, \quad \sigma(2) = 2, \quad \sigma(3) = 4, \quad \sigma(4) = 1.$$

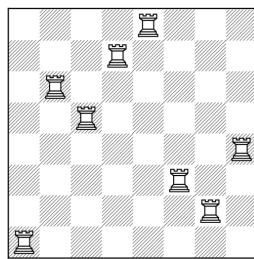
Notacja *tabelkowa* (chyba nie wymagająca wyjaśnienia — na górze argument, na dole wartość):

$$\sigma = \begin{bmatrix} 1 & 2 & 3 & 4 \\ 3 & 2 & 4 & 1 \end{bmatrix}$$

Macierz P_σ odpowiadająca tej permutacji ma postać:

$$P_\sigma = \begin{bmatrix} 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}.$$

Z punktu widzenia kombinatoryki interesująca może być dla Państwa następująca obrazowa intuicja. Rozważmy takie rozstawienie wież na szachownicy (a_{ij}) rozmiaru $n \times n$, by w każdym wierszu i kolumnie znajdowała się dokładnie jedna wieża. Permutacji $\sigma \in S_n$ odpowiada jedno z $n!$ różnych rozstawień.



W powyższym przykładzie dla macierzy $[a_{ij}]$ rozmiaru 8×8 wieże rozstawione są na miejscach:

$$a_{\sigma(1)1}, a_{\sigma(2)2}, a_{\sigma(3)3}, a_{\sigma(4)4}, a_{\sigma(5)5}, a_{\sigma(6)6}, a_{\sigma(7)7}, a_{\sigma(8)8},$$

czyli:

$$a_{81}, a_{32}, a_{43}, a_{24}, a_{15}, a_{66}, a_{77}, a_{58}.$$

Jesteśmy gotowi do wprowadzenia i dowodu tytułowego *wzoru permutacyjnego* na wyznacznik.

Twierdzenie 27.1.2

Dla macierzy $A = [a_{ij}] \in M_n(K)$ zachodzi wzór permutacyjny:

$$\det A = \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) a_{\sigma(1)1} a_{\sigma(2)2} \dots a_{\sigma(n)n}.$$

Przykład 1. Dla macierzy

$$A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$$

rozmiaru 2×2 mamy dwie możliwe permutacje kolumn reprezentowane przez następujące rozstawienia wież:



Dokładniej, $S_2 = \{\sigma_1, \sigma_2\}$, gdzie $\sigma_1(1) = 1, \sigma_1(2) = 2$ oraz $\sigma_2(1) = 2, \sigma_2(2) = 1$. Oczywiście widzimy, że $\operatorname{sgn}(\sigma_1) = 1, \operatorname{sgn}(\sigma_2) = -1$, skąd

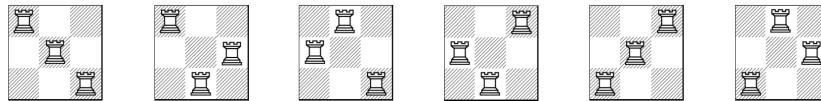
$$\det A = 1 \cdot a_{11}a_{22} + (-1)a_{21}a_{12}.$$

Przykład 2. Dla macierzy

$$A = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}$$

rozmiaru 3×3 mamy:

$$\det A = a_{11}a_{22}a_{33} - a_{11}a_{32}a_{23} - a_{21}a_{12}a_{33} + a_{21}a_{32}a_{13} - a_{31}a_{22}a_{13} + a_{31}a_{12}a_{23}.$$



Ważne: jeśli dla $\sigma \in S_n$ jakaś wieża stoi na zerze*, tzn. $a_{\sigma(i)i} = 0$, dla pewnego i , to odpowiedniego składnika ($= 0$) nie wliczamy do obliczania wyznacznika. Popatrzmy na dwa kolejne przykłady.

Przykład 3. Dla macierzy

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

mamy, podobnie jak wyżej,

$$a_{\sigma(1)1}a_{\sigma(2)2}a_{\sigma(3)3}a_{\sigma(4)4}a_{\sigma(5)5} \neq 0 \Leftrightarrow \sigma = \begin{bmatrix} 1 & 2 & 3 & 4 & 5 \\ 3 & 5 & 2 & 4 & 1 \end{bmatrix} \Rightarrow \det(A) = -1.$$

Przykład 4. Dla macierzy górnopróbkowej

$$\begin{bmatrix} a_{11} & * & * & \dots & * \\ 0 & a_{22} & * & \dots & * \\ 0 & 0 & a_{33} & \dots & * \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & a_{nn} \end{bmatrix}$$

$$a_{\sigma(1)1}a_{\sigma(2)2} \dots a_{\sigma(n)n} \neq 0 \Leftrightarrow \sigma(i) = i, \text{ dla } i = 1, 2, \dots, n. \Rightarrow \det A = a_{11}a_{22} \dots a_{nn}.$$

* * *

DOWÓD FORMUŁY PERMUTACYJNEJ. Korzystać będziemy wielokrotności z jednorodności i addytywności wyznacznika względem dowolnej kolumny. Mamy też, dla i -tej kolumny macierzy A :

$$\begin{bmatrix} a_{1i} \\ \vdots \\ a_{ni} \end{bmatrix} = a_{1i} \cdot \begin{bmatrix} 1 \\ \vdots \\ 0 \end{bmatrix} + \dots + a_{ni} \cdot \begin{bmatrix} 0 \\ \vdots \\ 1 \end{bmatrix} = a_{1i}E_1 + \dots + a_{ni}E_n.$$

Niech $A = [a_{ij}]$. Mamy:

$$\det[A_1, \dots, A_n] = \det[a_{11}E_1 + \dots + a_{n1}E_n, \dots, a_{n1}E_1 + \dots + a_{nn}E_n].$$

Korzystamy teraz z addytywności i jednorodności względem pierwszej kolumny dostając:

$$\begin{aligned} \det[A_1, \dots, A_n] &= \det[a_{11}E_1 + \dots + a_{n1}E_n, \dots, a_{n1}E_1 + \dots + a_{nn}E_n] = \\ &= a_{11} \det[E_1, a_{12}E_1 + \dots + a_{n2}E_n, \dots, a_{1n}E_1 + \dots + a_{nn}E_n] + \\ &\quad + a_{21} \det[E_2, a_{12}E_1 + \dots + a_{n2}E_n, \dots, a_{1n}E_1 + \dots + a_{nn}E_n] + \\ &\quad + \dots + \\ &\quad + a_{n1} \det[E_n, a_{12}E_1 + \dots + a_{n2}E_n, \dots, a_{1n}E_1 + \dots + a_{nn}E_n]. \end{aligned}$$

Teraz dla dwóch z otrzymanych n składników korzystamy z liniowości względem drugiej kolumny:

$$\begin{aligned} \det[A_1, \dots, A_n] &= \det[a_{11}E_1 + \dots + a_{n1}E_n, \dots, a_{n1}E_1 + \dots + a_{nn}E_n] = \\ &\quad + a_{11}a_{12} \det[E_1, E_1, \dots, a_{1n}E_1 + \dots + a_{nn}E_n] + \\ &\quad + a_{11}a_{22} \det[E_1, E_2, \dots, a_{1n}E_1 + \dots + a_{nn}E_n] + \\ &\quad + \dots + \\ &\quad + a_{11}a_{n2} \det[E_1, E_n, \dots, a_{1n}E_1 + \dots + a_{nn}E_n] + \\ &\quad + a_{21}a_{12} \det[E_1, E_1, \dots, a_{1n}E_1 + \dots + a_{nn}E_n] + \\ &\quad + a_{21}a_{22} \det[E_1, E_2, \dots, a_{1n}E_1 + \dots + a_{nn}E_n] + \\ &\quad + \dots + \\ &\quad + a_{21}a_{n2} \det[E_1, E_n, \dots, a_{1n}E_1 + \dots + a_{nn}E_n] + \\ &\quad + \dots + \\ &\quad + a_{n1} \det[E_n, a_{12}E_1 + \dots + a_{n2}E_n, \dots, a_{1n}E_1 + \dots + a_{nn}E_n]. \end{aligned}$$

Tą samą procedurę wykonujemy dla pozostałych $n - 2$ składników, dostając n^2 składników postaci:

$$\det[A_1, \dots, A_n] = \sum_{i_1=1}^n \sum_{i_2=1}^n a_{i_1 1} a_{i_2 2} \det[E_{i_1}, E_{i_2}, \dots, a_{1n}E_1 + \dots + a_{nn}E_n].$$

Teraz dla każdego z n^2 składników korzystamy z liniowości względem trzeciej kolumny, co da nam n^3 składników – i tak dalej aż otrzymamy przedstawienie w postaci n^n składników:

$$\det[A_1, \dots, A_n] = \sum_{i_1=1}^n \sum_{i_2=1}^n \dots \sum_{i_n=1}^n a_{i_1 1} a_{i_2 2} \dots a_{i_n n} \det[E_{i_1}, E_{i_2}, \dots, E_{i_n}].$$

Zauważmy, że z tych n^n składników tylko $n!$ może być niezerowych – te, dla których

$$\det[E_{i_1}, E_{i_2}, \dots, E_{i_n}] \neq 0.$$

Tymczasem z własności (3) dostajemy:

$$\det[E_{i_1}, E_{i_2}, \dots, E_{i_n}] = \begin{cases} 0, & \text{gdy } i_k \text{ nie są parami różne,} \\ \operatorname{sgn}(\sigma) & \text{dla } \sigma \in S_n : \sigma(k) = i_k. \end{cases}$$

Wzór permutacyjny pozwala, jak widzieliśmy wyżej, na wykazanie wielu własności wyznacznika. Można też, w oparciu o abstrakcyjną definicję znaku permutacji, uznać wzór permutacyjny za definicję wyznacznika. Takie podejście przyjmowane jest zwyczaczka w starszych podręcznikach.

* * *

Ostatnim pojęciem, któremu przyjrzymy się nieco bliżej, jest pojęcie minora stopnia k macierzy A . Jest to wyznacznik macierzy kwadratowej rozmiaru k powstały przez usunięcie odpowiedniej liczby wierszy i kolumn macierzy A — niekoniecznie kwadratowej.

Definicja 27.1.3: Minory

Dla macierzy $A \in M_{m \times n}(K)$ oraz dla każdego $k \leq \min(m, n)$

- przez $A_{j_1, \dots, j_k}^{i_1, \dots, i_k}$ oznaczamy macierz powstałą z A przez wykreślenie $m - k$ wierszy o indeksach różnych od i_1, \dots, i_k oraz wykreślenie $n - k$ kolumn o indeksach różnych od j_1, \dots, j_k ,
- MINORAMI STOPNIA k nazywamy wyznaczniki macierzy $A_{j_1, \dots, j_k}^{i_1, \dots, i_k}$.

Przykład. Dla macierzy

$$A = \begin{bmatrix} 1 & 2 & 3 & 4 \\ 5 & 6 & 7 & 8 \\ 9 & 10 & 11 & 12 \end{bmatrix}$$

mamy

$$A_{2,4}^{1,3} = \begin{bmatrix} 2 & 4 \\ 10 & 12 \end{bmatrix}, \quad A_{1,2,4}^{1,2,3} = \begin{bmatrix} 1 & 2 & 4 \\ 5 & 6 & 8 \\ 9 & 10 & 12 \end{bmatrix}.$$

Oczywiście minorami stopnia $n - 1$ posługiwały się przy wyliczaniu wyznacznika macierzy $n \times n$. Wyznacznik macierzy A_{ij} , nazywamy czasem *minorem odpowiadającym wyrazowi a_{ij}* macierzy A .

W wielu źródłach pojęcie minora jest podstawą do następującej wygodnej charakteryzacji (a nawet alternatywnej definicji) rzędu macierzy.

Twierdzenie 27.1.4

Niech $A \in M_{m \times n}(K)$. Następujące warunki są równoważne:

- (1) $r(A) = k$,
- (2) wszystkie minory stopnia $> k$ macierzy A są zerowe, o ile istnieją, oraz istnieje minor stopnia k macierzy A , który jest niezerowy.

Dowód. Rozpoczniemy od przypomnienia następującej elementarnej obserwacji. Jeśli dany jest układ wektorów $\alpha_1, \dots, \alpha_r \in K^m$ taki, że usunięcie pewnych ustalonych $m - r$ współrzędnych tych wektorów daje wektory $\alpha'_1, \dots, \alpha'_r \in K^r$ tworzące układ liniowo niezależny, to $\dim \text{lin}(\alpha_1, \dots, \alpha_r) \geq r$.

Uzasadnienie zostawiamy jako proste ćwiczenie. Dla przykładu, podprzestrzeń przestrzeni liniowej \mathbb{R}^4 rozpięta przez wektory $(1, 2, 3, 4), (0, 1, 0, 0)$ jest wymiaru co najmniej 2, gdyż usuwając trzecie i czwarte ich współrzędne dostajemy układ $(1, 2), (0, 1)$, który jest ewidentnie liniowo niezależny w \mathbb{R}^2 .

Przypuśćmy najpierw, że $r(A) = k$. Wówczas, skoro rząd jest wymiarem przestrzeni wierszowej macierzy A , to można wybrać k wierszy $\alpha_{i_1}, \dots, \alpha_{i_k}$, które tworzą układ liniowo niezależny. Jednakże te k wierszy tworzy macierz A' rzędu k . W takim razie można wybrać k kolumn o indeksach j_1, \dots, j_k macierzy A' , które tworzą układ liniowo niezależny (wymiar przestrzeni wierszowej równy jest wymiarowi przestrzeni kolumnowej). Stąd także macierz kwadratowa $A_{j_1, \dots, j_k}^{i_1, \dots, i_k}$ ma rząd k , czyli wyznacznik tej macierzy — minor rzędu k — jest niezerowy.

Z drugiej strony, istnienie niezerowego minora $\det A_{j_1, \dots, j_r}^{i_1, \dots, i_r}$ stopnia r macierzy A o wierszach $\alpha_1, \dots, \alpha_r$ oznacza, że rząd macierzy $A_{j_1, \dots, j_r}^{i_1, \dots, i_r}$ jest równy r , a stąd wiersze $\alpha_{i_1}, \dots, \alpha_{i_r}$ macierzy A tworzą układ liniowo niezależny. Zatem $r(A) \geq r$. \square

Przykład. Rozważmy macierz rzeczywistą postaci:

$$A = \begin{bmatrix} 1 & 2 & -1 & 4 \\ 3 & 1 & 2 & 2 \\ 4 & 3 & 1 & 6 \end{bmatrix}.$$

Nietrudno widzieć, że

$$\det A_{1,2}^{1,2} = \det \begin{bmatrix} 1 & 2 \\ 3 & 1 \end{bmatrix} = -5 \neq 0.$$

Analizując natomiast minory stopnia 3 mamy:

$$\begin{vmatrix} 1 & 2 & -1 \\ 3 & 1 & 2 \\ 4 & 3 & 1 \end{vmatrix} = \begin{vmatrix} 1 & 2 & 4 \\ 3 & 1 & 2 \\ 4 & 3 & 6 \end{vmatrix} = \begin{vmatrix} 1 & -1 & 4 \\ 3 & 2 & 2 \\ 4 & 1 & 6 \end{vmatrix} = \begin{vmatrix} 2 & -1 & 4 \\ 1 & 2 & 2 \\ 3 & 1 & 6 \end{vmatrix} = 0,$$

czyli wszystkie minory stopnia 3 są zerowe. Mamy więc $r(A) = 2$.

W drugim semestrze poznamy inny ważny przykład minora — główny minor wiodący stopnia k , powstający z macierzy kwadratowej rozmiaru n przez usunięcie ostatnich $n - k$ wierszy i kolumn. Będzie on służył do sformułowania bardzo istotnego faktu — tzw. kryterium Sylvester'a.

Ostatnie zastosowanie wyznacznika, jakie tu przedstawimy, ma duże znaczenie w geometrii analitycznej i analizie, a dotyczy sytuacji, gdy chcemy policzyć wyznacznik iloczynu macierzy, które nie są kwadratowe (ale iloczyn jest).

Twierdzenie 27.1.5: Cauchy-Binet

Jeśli $A \in M_{l \times m}(K)$, $B \in M_{m \times n}(K)$ oraz $r \leq \min(l, m, n)$, to

$$\det(AB)_{j_1, \dots, j_r}^{i_1, \dots, i_r} = \sum_{1 \leq k_1 < \dots < k_r \leq m} \left(\det A_{k_1, \dots, k_r}^{i_1, \dots, i_r} \right) \left(\det B_{j_1, \dots, j_r}^{k_1, \dots, k_r} \right).$$

Oto przykład rachunku używającego powyższego twierdzenia. Weźmy

$$A = \begin{bmatrix} 1 & 1 & 2 \\ 3 & 1 & -1 \end{bmatrix}, \quad B = \begin{bmatrix} 1 & 1 \\ 3 & 1 \\ 0 & 2 \end{bmatrix}.$$

Interesuje nas po prostu wyznacznik AB , czyli macierzy 2×2 . Mamy więc $l = n = 2$, $m = 3$, a wyznacznik jest sumą iloczynów odpowiednich minorów stopnia 2, czyli:

$$\begin{aligned} \det AB &= \sum_{1 \leq k_1 < k_2 \leq 3} \det A_{k_1, k_2}^{1, 2} \det B_{1, 2}^{k_1, k_2} = \det A_{1, 2}^{1, 2} \det B_{1, 2}^{1, 2} + \det A_{2, 3}^{1, 2} \det B_{1, 2}^{2, 3} + \det A_{1, 3}^{1, 2} \det B_{1, 2}^{1, 3} \\ &= \begin{vmatrix} 1 & 1 \\ 3 & 1 \end{vmatrix} \cdot \begin{vmatrix} 1 & 1 \\ 3 & 1 \end{vmatrix} + \begin{vmatrix} 1 & 2 \\ 1 & -1 \end{vmatrix} \cdot \begin{vmatrix} 3 & 1 \\ 0 & 2 \end{vmatrix} + \begin{vmatrix} 1 & 2 \\ 3 & -1 \end{vmatrix} \cdot \begin{vmatrix} 1 & 1 \\ 0 & 2 \end{vmatrix} = -28. \end{aligned}$$

Idea dowodu twierdzenia Cauchy'ego-Bineta jest następująca. Wystarczy pokazać, że jeśli $A \in M_{n \times N}(K)$ oraz $B \in M_{N \times n}(K)$ oraz jeśli \mathcal{S} jest rodziną n elementowych podzbiorów $\{1, \dots, N\}$ to

$$\det(AB) = \sum_{S \in \mathcal{S}} \det A_S \det B_S,$$

gdzie A_S powstaje z A przez usunięcie kolumn o indeksach spoza S , oraz B_S powstaje przez usunięcie wierszy o indeksach spoza S . Niech A_1, \dots, A_n będą wierszami A oraz B_1, \dots, B_n będą kolumnami B , traktowanymi jako wektory w K^N . Rozważamy funkcje

$$f(A, B) = \det(AB) = f(A_1, \dots, A_n, B_1, \dots, B_n) \text{ oraz } g(A, B) = \sum_{S \in \mathcal{S}} \det A_S \det B_S = g(A_1, \dots, A_n, B_1, \dots, B_n).$$

Jak się okazuje, funkcje f, g są jednorodne i addytywne względem każdej współrzędnej (sprawdzamy to poprzez własności wyznacznika). Aby pokazać, że są identyczne wystarczy zatem rozważyć przypadek, gdy A_i, B_j są wektorami bazy standardowej. I dalej działamy jak w dowodzie wzoru permutacyjnego.

Jednym z ładnych zastosowań twierdzenia wyżej jest zauważenie, że dla $A \in M_{n \times N}(K)$ oraz $B = A^T$ mamy $\det A_S = \det B_S$, dla każdego n -elementowego podzbioru $S \subseteq \{1, 2, \dots, N\}$, czyli uzyskujemy nierówność $\det(AA^T) = \sum(\det A_S)^2$. Wynik ten ma interpretację geometryczną, którą zrozumiemy w kolejnym semestrze, nazywaną *uogólnionym twierdzeniem Pitagorasa*. W elementarnym wydaniu mówi ono, że kwadrat pola równoległoboku R w przestrzeni \mathbb{R}^3 równy jest sumie kwadratów pól równoległoboków powstających przez rzutowanie R na płaszczyznę rozpięte przez pary prostopadłych osi układu współrzędnych. W kolejnym semestrze dowiemy się wreszcie co łączy wyznacznik i objętość (ale nie od razu).

27.2 Wybór przykładowych pytań

Proszę się nie uczyć na pamięć odpowiedzi czy wyjaśnienie! Takich pytań można ułożyć znacznie więcej i służą one jedynie sprawdzeniu podstawowego zrozumienia treści zajęć. Jeśli odpowiedź na dowolne z poniższych pytań sprawia Państwu problem, proszę pomyśleć który fragment notatek dotyczy tego pytania, wrócić do ich lektury, a w razie dalszych trudności poprosić prowadzącego o wskazówkę.

1. Określ znaki permutacji $\sigma \in S_6$ danych w notacji tabelkowej:

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

2. Wskaż takie indeksy i, j, k , że iloczyn

$$a_{51}a_{i6}a_{1j}a_{35}a_{44}a_{6k}$$

pojawia się jako składnik w formule permutacyjnej na wyznacznik macierzy $[a_{ij}] \in M_6(\mathbb{R})$ ze znakiem $-$.

3. Ile jest co najwyżej niezerowych składników, gdy wyznaczamy za pomocą formuły permutacyjnej wyznacznik macierzy górnopróbkątnej?
4. Znajdź wszystkie składniki występujące we wzorze permutacyjnym na wyznacznik

$$\begin{vmatrix} x & 1 & 2 & 3 \\ x & x & 1 & 2 \\ 1 & 2 & x & 3 \\ x & 1 & 2 & 2x \end{vmatrix}$$

zawierające x^4 oraz x^3 .

5. Niech $n \geq 3$. Korzystając ze wzoru permutacyjnego oblicz wyznacznik macierzy

$$A = \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 & 0 & 1 \\ 1 & 0 & 1 & \cdots & 0 & 0 & 0 \\ 0 & 1 & 0 & \cdots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 0 & 1 & 0 \\ 0 & 0 & 0 & \cdots & 1 & 0 & 1 \\ 1 & 0 & 0 & \cdots & 0 & 1 & 0 \end{bmatrix} \in M_n(\mathbb{R}).$$

6. Uzasadnij, że dla dowolnej macierzy $A \in M_n(K)$ wyznacznik dowolnej macierzy $\det(A + \lambda I)$ należy do $K[\lambda]$ i jest wielomianem stopnia $\leq n$.
7. Uzasadnij, że jeśli $A = [a_{ij}] \in M_n(\mathbb{Z})$ oraz ij -ty wyraz macierzy A_p powstaje przez policzenie reszty z a_{ij} modulo p , to wyznacznik macierzy A_p równy jest reszcie z dzielenia $\det A$ przez p .
8. Czy poniższy wyznacznik jest niezerowy?

$$\begin{vmatrix} 102495 & 550429 & 873298 & 660697 \\ 370628 & 909093 & 127450 & 925601 \\ 835044 & 601178 & 624655 & 263392 \\ 663780 & 487252 & 292276 & 593107 \end{vmatrix}$$

9. Uzasadnij, że jeśli rząd macierzy $A \in M_{m \times n}$ jest równy r , to macierz ta posiada niezerowy minor stopnia r .

10. Uzasadnij, że z Twierdzenia Cauchy'ego-Bineta wynika nierówność

$$\det(AA^T) \geq 0,$$

dla dowolnej macierzy $A \in M_{m \times n}(\mathbb{R})$. Zastosuj ten fakt dla macierzy

$$A = \begin{bmatrix} x_1 & x_2 & \dots & x_n \\ y_1 & y_2 & \dots & y_n \end{bmatrix}.$$

Czy znasz nierówność, powstającą w rezultacie?

27.3 Zadania do samodzielnej pracy

Zadania ilustrujące umiejętności niezbędne do zaliczenia przedmiotu oznaczone są symbolem ♠.

Przy tych zadaniach dodany jest opis umiejętności, które sprawdzają.

1. (♠) Korzystając z definicji permutacyjnej wyznacznika oblicz:

$$\begin{vmatrix} 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 0 \end{vmatrix}, \quad \begin{vmatrix} 0 & 0 & 1 & 2 & 1 \\ 0 & 2 & 3 & 4 & 0 \\ 5 & 0 & 5 & 6 & 0 \\ 0 & 0 & 1 & 2 & 0 \\ 0 & 0 & 3 & 4 & 0 \end{vmatrix}, \quad \begin{vmatrix} 1 & 2 & 3 & 4 & 1 \\ 0 & -1 & 2 & 4 & 2 \\ 0 & 0 & 4 & 0 & 0 \\ 0 & 0 & 0 & 0 & 7 \\ 0 & 0 & 1 & 1 & 1 \end{vmatrix}.$$

2. Niech $n \geq 3$. Korzystając ze wzoru permutacyjnego oblicz wyznaczniki macierzy

$$\begin{bmatrix} -t & 0 & 0 & \dots & 0 & a_1 \\ a_2 & -t & 0 & \dots & 0 & 0 \\ 0 & a_3 & -t & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & -t & 0 \\ 0 & 0 & 0 & \dots & a_n & -t \end{bmatrix}, \quad \begin{bmatrix} a_1 & 0 & 0 & \dots & 0 & b_n \\ b_1 & a_2 & 0 & \dots & 0 & 0 \\ 0 & b_2 & a_3 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & a_{n-1} & 0 \\ 0 & 0 & 0 & \dots & b_{n-1} & a_n \end{bmatrix}$$

3. Uzasadnij, że dla dowolnej macierzy $A \in M_n(K)$ mamy

$$\det(A - \lambda I) = (a_{11} - \lambda)(a_{22} - \lambda) \dots (a_{nn} - \lambda) + v(\lambda),$$

gdzie $v(\lambda) \in K[\lambda]$ jest wielomianem stopnia co najwyżej $n - 2$.

4. Uzasadnij, że jeśli liczba wyrazów zerowych w macierzy $n \times n$ jest większa od $n^2 - n$, to jej wyznacznik jest równy 0.

5. Uzasadnij, że $\text{sgn}(\sigma) = \text{sgn}(\sigma^{-1})$, dla dowolnej permutacji $\sigma \in S_n$.

6. Uzasadnij (ponownie), że jeśli w macierzy $n \times n$ na przecięciu k wierszy i l kolumn znajdują się same zera, przy czym $k + l > n$, to wyznacznik tej macierzy jest równy 0.

7. Uzasadnij, że każda macierz permutacji P_σ jest iloczynem pewnej liczby macierzy operacji elementarnych zamiany wierszy. Uzasadnij, że choć przedstawienie to nie musi być jednoznaczne, to parzystość liczby czynników tego rozkładu jest jednoznacznie wyznaczona przez permutację σ .

8. Założmy, że $n \geq 1$ jest liczbą nieparzystą.

- (i) Dowiedź, że gdy permutacja $\sigma \in S_n$ spełnia równość $\sigma \circ \sigma = \text{id}$, to σ ma punkt stały (tzn. istnieje takie $k \in \{1, \dots, n\}$, że $\sigma(k) = k$).
- (ii) Wywnioskuj z punktu (i), że gdy macierz symetryczna $A = [a_{ij}] \in M_n(\mathbb{Q})$ spełnia $a_{ij} \in \mathbb{Z}$ dla $i \neq j$ oraz $a_{11} = \dots = a_{nn} = 0$, to $\det A \in \mathbb{Z}$ jest liczbą parzystą.

9. *Cyklem* długości k nazwiemy taką permutację σ zbioru n -elementowego $\{1, 2, \dots, n\}$, że istnieje zbiór i_1, \dots, i_k zwany *nośnikiem* σ , że

- $\sigma(i_1) = i_2, \sigma(i_2) = i_3, \dots, \sigma(i_{k-1}) = i_k, \sigma(i_k) = i_1$,
- $\sigma(i) = i$, dla $i \notin \{i_1, \dots, i_k\}$.

- (a) Uzasadnij, że każda permutacja jest cyklem lub może być przedstawiona jako złożenie cykli o parami rozłącznych nośnikach (przyjmujemy tu, że identyczność jest cyklem długości 0). Czy rozkład ten jest jednoznaczny? Powiąż liczbę cykli w rozkładzie ze znakiem permutacji.
- (b) Niech P_σ będzie macierzą cyklu $\sigma \in S_n$. Uzasadnij, że $\det(I_n + P_\sigma) \in \{0, 2\}$.
- (c) Niech P będzie dowolną macierzą permutacyjną. Wykaż, że $\det(I_n + P)$ jest zerem lub potągą dwójki.

10. Dla jakich $n \geq 1$ istnieje taka macierz $A = [a_{ij}] \in M_n(\mathbb{R})$, że każdy składnik $(\text{sgn } \sigma)a_{\sigma(1)1} \dots a_{\sigma(n)n}$ sumy

$$\det A = \sum_{\sigma \in S_n} (\text{sgn } \sigma)a_{\sigma(1)1} \dots a_{\sigma(n)n}$$

jest dodatni?