Algebra liniowa, I rok Informatyki, 2012/2013

1. Struktury algebraiczne. Półgrupy, grupy, pierścienie, ciała. Liczby zespolone.
2. Macierze. Działania i własności. Równania liniowe.
3. Przestrzenie liniowe. Podprzestrzenie, wymiar, baza. Przekształcenia liniowe.
4. Obraz i jądro przekształcenia liniowego i macierzy. Rząd macierzy. Funkcjonały liniowe. Przestrzeń sprzężona. Baza dualna.
5. Układy równań liniowych. Twierdzenie Kroneckera-Capellego. Metoda eliminacji Gaussa. Rozkład macierzy na czynniki trójkątne.
6. Warstwy. Przestrzeń ilorazowa. Układy współrzędnych w przestrzeni liniowej.
7. Normy wektorów i macierzy. Nierówności norm.
8. Wyznaczniki. Wzory Cramera.
9. Formy dwuliniowe. Macierz formy. Kongruencje. Twierdzenie Sylvestra. Przestrzenie ortogonalne.
10. Równania i zbiory II stopnia. Afiniczna postać kanoniczna równania.
11. Przestrzenie euklidesowe i unitarne. Izometrie. Rzuty prostopadłe.
12. Bazy ortogonalne. Ortogonalizacja Grama-Schmidta. Objętość równoległościanu.
13. Bazy ortonormalne. Macierze ortogonalne.
14. Liniowe zadania najmniejszych kwadratów.

Zasady zaliczania przedmiotu

Blady atoli strach padał na wszystkich, kiedy król, dla niespodzianego kaprysu, ogłaszał zgadywanki. Z dawien dawna lubował się w nich i jeszcze wielkiego kanclerza podczas koronacji zaskoczył pytaniem, jak sądzi, czy pacierz i macierz różnią się czymś między sobą, a jeśli tak, to czym?

Stanisław Lem: Cyberiada.

Na zaliczenie przedmiotu składają się:

- Prace domowe zadawane co tydzień, a także kartkówki (i inne szykany ustalane indywidualnie przez prowadzących ćwiczenia). Na końcu semestru liczba otrzymanych punktów będzie podzielona przez maksymalną liczbę punktów możliwych do zdobycia w danej grupie ćwiczeniowej, pomnożona przez 50 i dodana do liczby punktów zdobytych na kolokwium.
- Kolokwium (pisemne, wspólne dla całego roku). Za kolokwium można dostać maksymalnie 50 punktów.
- Aby zaliczyć ćwiczenia, trzeba na nie chodzić i zdobyć (łącznie, z prac domowych i kolokwium) co najmniej 50 punktów.
- Egzamin pisemny na końcu semestru.
- Osoby, u których wystąpi niezgodność między oceną z egzaminu pisemnego i liczbą punktów z ćwiczeń, albo niezgodność proponowanej oceny końcowej z ambicjami, zdają egzamin ustny.
- Osoby, które otrzymały 88 lub więcej punktów z ćwiczeń mają prawo do zdawania tylko egzaminu ustnego.

Kolokwium z algebry liniowej, I rok Inf.

(*Ściśle tajne* przed godz. 14^{00} 6 grudnia 2012.)

Proszę bardzo uważnie przeczytać treść zadań. Podczas oceniania *nie mniej ważne niż rachunki* będą poprawne uzasadnienia wszystkich odpowiedzi (z powołaniem się na właściwe twierdzenia).
Za każde zadanie można otrzymać od 0 do 10 punktów.

1. Zbadaj, czy każdy ze zbiorów:

$$A_1 = \{(x, y) : x > 0,\ y < 0\}, \quad A_2 = \{(x, y) : xy \geqslant 0\}$$

jest (a) podgrupą, (b) podpółgrupą grupy $\mathbb{R}^2$ ze „zwykłym" działaniem dodawania (tj. $(x_1, y_1) + (x_2, y_2) = (x_1 + x_2, y_1 + y_2)$)?

2. Środkiem ciężkości pewnego trójkąta równobocznego w płaszczyźnie zespolonej jest punkt $(1, 1)$ (tzn. $1 + i$), a jednym z jego wierzchołków jest $(0, 0)$. Znajdź pozostałe dwa wierzchołki tego trójkąta.

3. Niech $\mathbf{a} = [1, 2, 3, 4, 5]^T$, $\mathbf{b} = [-5, 4, -3, 2, 1]^T$ i niech $A = \mathbf{a}\mathbf{b}^T$. Oblicz iloczyn $A\mathbf{x}$, gdzie $\mathbf{x} = [3, 0, -5, -1, 1]^T$.

4. Niech

$$\mathbf{x}_1 = \begin{bmatrix} 1 \\ 2 \\ 3 \\ 2 \end{bmatrix}, \ \mathbf{x}_2 = \begin{bmatrix} -2 \\ 3 \\ -1 \\ -1 \end{bmatrix}, \ \mathbf{x}_3 = \begin{bmatrix} -1 \\ 5 \\ 2 \\ 1 \end{bmatrix}, \ \mathbf{x}_4 = \begin{bmatrix} 0 \\ 5 \\ 5 \\ 3 \end{bmatrix}, \ \mathbf{x}_5 = \begin{bmatrix} 0 \\ 5 \\ -5 \\ 3 \end{bmatrix}$$

i niech $V_1 = \operatorname{lin}\{\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3\}$, $V_2 = \{\mathbf{x}_4, \mathbf{x}_5\}$. Znajdź wymiar sumy algebraicznej przestrzeni V_1 i V_2. Czy jest to suma prosta? Odpowiedź uzadasnij.

5. Symbol $\mathbb{R}[x]_n$ oznacza przestrzeń wielomianów rzeczywistych jednej zmiennej stopnia co najwyżej n. Przekształcenie liniowe $f : \mathbb{R}[x]_5 \to \mathbb{R}[x]_5$ określone jest wzorem $\big(f(w)\big)(x) = xw'(x) - 5w(x)$. Znajdź macierz przekształcenia f w bazie potęgowej przestrzeni $R[x]_5$ i wyznacz wymiar obrazu tego przekształcenia.

Kolokwium poprawkowe z algebry liniowej, I rok Inf.

(*Ściśle tajne* przed godz. 14^{00} 18 stycznia 2013.)

Proszę bardzo uważnie przeczytać treść zadań. Podczas oceniania *nie mniej ważne niż rachunki* będą poprawne uzasadnienia wszystkich odpowiedzi (z powołaniem się na właściwe twierdzenia).
Za każde zadanie można otrzymać od 0 do 10 punktów.

1. Zbadaj, czy zbiór liczb całkowitych z działaniem „$\diamond$" określonym wzorem

$$a \diamond b \overset{\text{def}}{=} a - ab + b$$

jest a) grupą, b) półgrupą. Czy działanie to jest przemienne?

2. Zbadaj, czy zbiór *zespolonych* macierzy hermitowskich $n \times n$ (dla ustalonego $n > 0$) jest przestrzenią liniową nad

a) ciałem liczb rzeczywistych,

b) ciałem liczb zespolonych.

W każdym przypadku, jeśli odpowiedź jest twierdząca, podaj wymiar przestrzeni.

3. Czy istnieją liczby całkowite dodatnie n, takie że

$$\operatorname{Im}\left[\left(\frac{\sqrt{2} + i\sqrt{2}}{1 - i\sqrt{3}}\right)^n\right] = 0\ ?$$

Odpowiedź uzasadnij i jeśli jest twierdząca, to znajdź najmniejsze takie n.

4. Niech $\mathbf{a} = [1, -2, 3, -4, 5]^T$, $\mathbf{b} = [5, -4, 3, -2, 1]^T$ i niech $A = \mathbf{a}\mathbf{b}^T$. Znajdź dowolną bazę przestrzeni rozwiązań układu równań że $A\mathbf{x} = \mathbf{0}$.

5. Symbol $\mathbb{R}[x]_n$ oznacza przestrzeń wielomianów rzeczywistych jednej zmiennej stopnia co najwyżej n. Przekształcenie liniowe $f : \mathbb{R}[x]_5 \to \mathbb{R}[x]_5$ określone jest wzorem $\big(f(w)\big)(x) = \frac{1}{2}\big(w(x) - w(-x)\big)$. Znajdź macierz przekształcenia f w bazie potęgowej przestrzeni $R[x]_5$ i wyznacz wymiar obrazu tego przekształcenia.

Działania

Def. Funkcję $f: X_1 \times \dots \times X_n \to X$ nazywamy działaniem n-argumentowym.

Def. Działaniem wewnętrznym n-argumentowym w zbiorze X nazywamy dowolną funkcję

$$f: \underbrace{X \times \dots \times X}_{n} \to X.$$

Zbiory $X_1, \dots, X_n, X$ w definicji działania mogą być różne, ale nie muszą.
Tak więc działanie wewnętrzne jest to szczególny przypadek działania w sensie tej ogólniejszej definicji, dla $X_1 = \dots = X_n = X$. Słowo „działanie" bywa też stosowane (niezbyt ściśle) na określenie funkcji nie objętych tymi definicjami.

Przykład: Dzielenie liczb *nie jest* działaniem w zbiorze liczb całkowitych dodatnich, $\mathbb{Z}_+$, bo nie dla każdej pary takich liczb istnieje liczba całkowita dodatnia, która jest ich ilorazem. Dlatego poprawniej byłoby mówić o „operacji" odwrotnej do mnożenia.
Dzielenie nie jest też działaniem wewnętrznym w zbiorze liczb wymiernych $\mathbb{Q}$, bo nie możemy dzielić przez 0. Jest to jednak działanie $\mathbb{Q} \times (\mathbb{Q} \setminus \{0\}) \to \mathbb{Q}$.

Notacja: Można pisać $f(x_1, \dots, x_n)$, $a \cdot b$, $\frac{a}{b}$, $\sin x$, $\sqrt{x}$ itd.
Sposób zapisu jest kwestią wygody.

Równości i równania

Wyrażeniem nazywamy dowolny symbol obiektu (elementu ustalonego zbioru albo zmiennej) lub poprawny (tj. zgodny z definicjami działań) opis obiektu za pomocą innych wyrażeń i działań na nich. Jeśli w wyrażeniu nie występują zmienne, to mamy jednoznacznie określoną wartość wyrażenia, czyli element odpowiedniego zbioru. W przeciwnym razie wyrażenie opisuje nam pewną funkcję — podstawiając w miejsce wszystkich zmiennych konkretne obiekty (argumenty) otrzymamy wyrażenie, które opisuje jeden konkretny obiekt (będący wartością funkcji). Z każdą zmienną jest związany zbiór (jeden ze zbiorów $X_1, \dots, X_n$), którego elementy wolno podstawiać w jej miejsce.

Dwa wyrażenia mogą opisywać ten sam obiekt lub różne obiekty. Zdanie „wyrażenie$_1$ = wyrażenie$_2$" oznacza stwierdzenie, że wartością obu wyrażeń jest ten sam obiekt. Takie zdanie nazywamy równością wyrażeń.

Uwaga: Nie należy mieszać symbolu równości, „=", z symbolem równości przybliżonej, „≈", który oznacza, że wartości wyrażeń różnią się dostatecznie mało (cokolwiek to znaczy). Pierwszy symbol należy do algebry, a drugi do analizy.
Do algebry należy natomiast symbol nierówności „≠".

Jeśli wyrażenia zawierają zmienne, to ich równość jest stwierdzeniem, że niezależnie od tego, jakie obiekty podstawimy w miejsce zmiennych (można też mówić „jakie wartości nadamy zmiennym"), oba wyrażenia będą opisywać ten sam obiekt. Równością tego typu jest np. zdanie

$$(a+b)^2 = a^2 + 2ab + b^2 \quad \text{dla } a, b \in \mathbb{R}.$$

Jeśli natomiast mamy za zadanie znalezienie obiektów, które podstawione w miejsce zmiennych dadzą nam zdanie prawdziwe (np. „znajdź wszystkie liczby x, takie że $x^2 + 1 = 0$"), to mówimy o równaniu. Zbiór rozwiązań równania (czyli obiektów, które podstawiwszy otrzymamy zdanie prawdziwe stwierdzające równość otrzymanych wyrażeń) może być pusty i wtedy mówimy, że jest ono sprzeczne (cokolwiek podstawimy w miejsce zmiennych, zdanie stwierdzające równość wyrażeń będzie fałszywe), ale uwaga: sprzeczność równania zależy od zbioru, w którym szukamy rozwiązań. Zbiór liczb dodatnich x spełniających równanie $x + 1 = 0$ jest pusty, ale w zbiorze liczb całkowitych rozwiązanie istnieje.

Struktury algebraiczne

Def. Struktura algebraiczna jest to pewien układ zbiorów $X_1, \dots, X_k$, w których są określone pewne działania $f_1, \dots, f_m$ ($f_j: X_{j_1} \times \dots \times X_{j_l} \to X_{j_{l+1}}$) oraz wyróżnione pewne elementy tych zbiorów.

W szczególności struktura algebraiczna może być jednym zbiorem X, w którym są określone pewne działania $f_1, \dots, f_m$ i wyróżnione pewne elementy $x_1, \dots, x_k \in X$.

Konkretną klasę struktur algebraicznych (np. grupy, ciała) definiuje się podając odpowiednie aksjomaty opisujące warunki spełnione przez działania.
Na podstawie aksjomatów można dowodzić twierdzenia orzekające o każdej takiej strukturze.
Mając konkretny układ zbiorów z działaniami możemy sprawdzić, czy są spełnione aksjomaty (i w szczególności, czy istnieją opisane przez nie elementy wyróżnione). Jeśli to się uda, to możemy do tego zbioru z działaniami stosować wszystkie twierdzenia, które pasują.

Grupy

Def. Grupą nazywamy niepusty zbiór X, w którym jest określone działanie dwuargumentowe „$\diamond$" i element wyróżniony e, takie że

G.1: Działanie „$\diamond$" jest łączne, tj.

$$\forall_{x,y,z\in X}\ (x\diamond y)\diamond z = x\diamond(y\diamond z)$$

Możemy więc pisać $x\diamond y\diamond z$, co pozostaje jednoznaczne.

G.2: Element e jest neutralnym elementem działania „$\diamond$":

$$\forall_{x\in X}\ e\diamond x = x\diamond e = x.$$

G.3: Dla każdego $x\in X$ istnieje element prawostronnie odwrotny (albo przeciwny):

$$\forall_{x\in X}\ \exists_{y\in X}\ x\diamond y = e.$$

Jeśli *dodatkowo* spełniony jest warunek

G.4: $\forall_{x,y\in X}\ x\diamond y = y\diamond x$, czyli działanie „$\diamond$" jest przemienne,

to grupa $(X,\diamond,e)$ nazywa się grupą abelową.

Twierdzenie: *Element prawostronnie odwrotny x' dowolnego elementu x w grupie jest tylko jeden. Element ten jest także elementem lewostronnie odwrotnym elementu x.*

Dowód: Niech $x\in X$. Na mocy G.3 istnieją $y,z\in X$, takie że $x\diamond y = y\diamond z = e$. Wtedy

$$z \underset{\substack{\uparrow\\ \text{G.2}}}{=} e\diamond z \underset{\substack{\uparrow\\ \text{G.3}}}{=} (x\diamond y)\diamond z \underset{\substack{\uparrow\\ \text{G.1}}}{=} x\diamond(y\diamond z) \underset{\substack{\uparrow\\ \text{G.3}}}{=} x\diamond e \underset{\substack{\uparrow\\ \text{G.2}}}{=} x. \qquad \square$$

Możemy więc mówić o elemencie odwrotnym do x, bez precyzowania strony.

Twierdzenie (prawostronne prawo skreśleń): *Jeśli zbiór X z działaniem „$\diamond$" jest grupą, to*

$$\forall_{x,y,z\in X}\ (x\diamond z = y\diamond z) \Rightarrow (x = y).$$

Dowód: Załóżmy, że $x\diamond z = y\diamond z$. Oznaczmy element odwrotny do z symbolem z'. Wtedy

$$x \underset{\substack{\uparrow\\ \text{G.2}}}{=} x\diamond e \underset{\substack{\uparrow\\ \text{G.3}}}{=} x\diamond(z\diamond z') \underset{\substack{\uparrow\\ \text{G.1}}}{=} (x\diamond z)\diamond z' \underset{\substack{\uparrow\\ \text{założenie}}}{=} (y\diamond z)\diamond z' \underset{\substack{\uparrow\\ \text{G.1}}}{=} y\diamond(z\diamond z') \underset{\substack{\uparrow\\ \text{G.3}}}{=} y\diamond e \underset{\substack{\uparrow\\ \text{G.2}}}{=} y. \qquad \square$$

Łatwo jest też dowieść lewostronnego prawa skreśleń,

$$\forall_{x,y,z\in X}\ (x\diamond y = x\diamond z) \Rightarrow (y = z),$$

ale tymczasem się powstrzymamy.

Uwaga: Dla tej samej klasy struktur algebraicznych można przyjmować różne układy aksjomatów. Na przykład moglibyśmy sformułować G.3 w bardziej symetryczny sposób:

G.3′: $\forall_{x\in X}\ \exists_{y\in X}\ x\diamond y = y\diamond x = e.$

Wszystko zależy od tego, kto to robi i w jakim celu (można dążyć do symetrii, albo poszukiwać najsłabszych warunków, z których wynika cała reszta).

O działaniu w grupie najczęściej myślimy jak o *mnożeniu* (mówimy wtedy: grupa multiplikatywna) i piszemy $a\cdot b$ lub ab zamiast $a\diamond b$. Element odwrotny do x oznaczamy wtedy symbolem x^{-1}, co pasuje do notacji potęgowej:

$$\underbrace{x\cdots x}_{n} \overset{\text{ozn.}}{=} x^n \qquad \text{(w szczególności } x^0 = e.)$$

Możemy też uznać działanie za *dodawanie* (wtedy jest grupa addytywna — piszemy $a+b$). Element przeciwny (czyli odwrotny w poprzedniej terminologii) do x oznaczamy symbolem $-x$. Mamy wtedy też notację mnożenia elementu przez liczbę całkowitą

$$\underbrace{x+\cdots+x}_{n} \overset{\text{ozn.}}{=} n\cdot x \qquad \text{(w szczególności } 0\cdot x = e.)$$

Przykłady grup:

- Zbiór liczb całkowitych $\mathbb{Z}$, z działaniem „$+$" (dodawaniem) i elementem neutralnym 0.
- Zbiór różnych od 0 liczb rzeczywistych, $\mathbb{R}\setminus\{0\}$, z działaniem „$\cdot$" (mnożeniem) i elementem neutralnym 1.

- Zbiór obrotów płaszczyzny wokół ustalonego punktu, o wielokrotności ustalonego kąta α. Działaniem jest *złożenie obrotów*, a elementem neutralnym obrót o kąt 0.
- Zbiór permutacji; permutacją zbioru n-elementowego nazywamy dowolną funkcję różnowartościową, której dziedziną i zbiorem wartości jest zbiór $\{1, 2, \ldots, n\}$. Działaniem w tej grupie jest złożenie funkcji, a elementem neutralnym jest funkcja identycznościowa. Grupę tę oznaczamy symbolem S_n.

Półgrupy

Def. Półgrupą nazywamy zbiór X, w którym jest określone działanie dwuargumentowe „$\diamond$", które jest łączne (zobacz aksjomat G.1).

Przykłady półgrup:

- Każda grupa jest półgrupą.
- Zbiór liczb naturalnych, $\mathbb{N} = \{1, 2, 3, \ldots\}$ (a także całkowitych nieujemnych, $\mathbb{N}_0 = \{0, 1, 2, \ldots\}$) z działaniem mnożenia.
- Zbiór napisów, czyli skończonych ciągów symboli ustalonego alfabetu $V = \{a_1, \ldots, a_n\}$. Półgrupę otrzymamy przyjmując za działanie konkatenację, tj. połączenie dwóch napisów w jeden. W półgrupie tej mamy element neutralny — napis pusty.

Pierścienie

Def. Pierścieniem nazywamy zbiór X, w którym są określone dwa działania: dodawanie („$+$") i mnożenie („$\cdot$"), spełniające następujące warunki:

P.1: Zbiór X z działaniem dodawania jest grupą abelową (jej element neutralny oznaczamy symbolem „0", czyli zero).

P.2: Zbiór X z działaniem mnożenia jest półgrupą.

Element neutralny tej półgrupy, *jeśli istnieje*, nazywamy jedynką i oznaczamy symbolem „1". Mamy wtedy pierścień z jedynką.

P.3: Mnożenie jest rozdzielne względem dodawania:

$$\forall_{a,b,c \in X}\ (a + b)c = ac + bc, \quad a(b + c) = ab + ac.$$

Twierdzenie: *Jeśli* $(X, +, \cdot, 0)$ *jest pierścieniem, to* $\forall_{a \in X}\ a \cdot 0 = 0 = 0 \cdot a$.

Dowód: $a \cdot 0 + 0 = a \cdot 0 = a \cdot (0 + 0) = a \cdot 0 + a \cdot 0$.
Na podstawie lewostronnego prawa skreśleń $0 = a \cdot 0$.
Dowód, że $0 \cdot a = 0$ jest podobny. □

Przykłady pierścieni:

- Zbiór liczb całkowitych $\mathbb{Z}$, ze „zwykłym" dodawaniem i mnożeniem, oraz elementami wyróżnionymi — liczbami 0 i 1.
- Zbiór $\mathbb{R}[x]$ wielomianów jednej zmiennej (oznaczonej symbolem x). Element neutralny dodawania to wielomian zerowy.

Ciała

Def. Ciałem nazywamy zbiór X z dwoma działaniami (dodawaniem i mnożeniem), taki że

C.1: Zbiór X z działaniem dodawania jest grupą abelową (element neutralny — 0),

C.2: Zbiór $X \setminus \{0\}$ z działaniem mnożenia jest grupą (element neutralny — 1),

C.3: Mnożenie jest rozdzielne względem dodawania (zobacz aksjomat P.3).

Jak widać, każde ciało ma co najmniej dwa elementy, 0 i 1, i jest pierścieniem z jedynką (uwaga: w pierścieniu jest możliwe $0 = 1$; w ciele jest to wykluczone). Jeśli grupa multiplikatywna ciała jest abelowa, to mamy ciało przemienne, a w przeciwnym razie ciało nieprzemienne (w tym wykładzie nie będzie o nich mowy).

Przykłady ciał:

- Zbiór liczb wymiernych $\mathbb{Q}$, rzeczywistych $\mathbb{R}$, albo zespolonych $\mathbb{C}$.
- Zbiór funkcji wymiernych; mają one postać $f(x) = \frac{w_1(x)}{w_2(x)}$, gdzie w_1 i w_2 to wielomiany zmiennej x.

Podstruktury

Def. Niech $(X, \diamond, e)$ będzie grupą. Zbiór $A \subset X$, taki że $e \in A$ oraz $(A, \diamond, e)$ jest grupą, nazywamy podgrupą grupy X.

Twierdzenie: *Warunek* $a, b \in A \Rightarrow ab^{-1} \in A$ *(stosujemy notację multiplikatywną) jest warunkiem koniecznym i dostatecznym tego, aby niepusty zbiór* $A \subset X$ *był podgrupą grupy* X.
Dowód: Sprawdzamy aksjomat G.2:

$$a \in A \Rightarrow aa^{-1} = e \in A.$$

Sprawdzamy G.3:

$$a \in A \Rightarrow e, a \in A \Rightarrow ea^{-1} = a^{-1} \in A.$$

Pozostaje sprawdzić, czy zbiór A jest zamknięty ze względu na działanie grupy X, ponieważ działanie to po obcięciu do zbioru A pozostaje oczywiście łączne. Zatem,

$$a, b \in A \Rightarrow a, b, b^{-1} \in A \Rightarrow a(b^{-1})^{-1} = ab \in A \quad \square$$

Pojęcia podpierścienia i podciała definiuje się analogicznie — jest to odpowiednio podzbiór ustalonego pierścienia lub ciała, który jest pierścieniem lub ciałem z tymi samymi działaniami i elementami wyróżnionymi. Na przykład, ciało liczb wymiernych $\mathbb{Q}$ jest podciałem ciała liczb rzeczywistych $\mathbb{R}$.

Liczby zespolone

Def. Liczbą zespoloną nazywamy parę uporządkowaną (a, b) liczb rzeczywistych. Liczba a nazywa się częścią rzeczywistą, a liczba b — częścią urojoną liczby zespolonej $z = (a, b)$.
Oznaczamy je symbolami $a = \operatorname{Re} z$, $b = \operatorname{Im} z$.
Zbiór liczb zespolonych oznaczamy literą $\mathbb{C}$.

Def. Dodawanie i mnożenie liczb zespolonych określa się według wzorów

$$(a_1, b_1) + (a_2, b_2) = (a_1 + a_2, b_1 + b_2),$$
$$(a_1, b_1) \cdot (a_2, b_2) = (a_1a_2 - b_1b_2, a_1b_2 + b_1a_2).$$

Twierdzenie: *Zbiór* $\mathbb{C}$ *z określonymi wyżej działaniami jest ciałem. Element neutralny dodawania to* $(0, 0)$*, a mnożenia* $(1, 0)$.
Szkic dowodu: 1. Należy sprawdzić, że $(\mathbb{C}, +, (0, 0))$ jest grupą abelową, na podstawie własności dodawania liczb rzeczywistych,
2. Należy sprawdzić, że $(\mathbb{C}\backslash\{(0, 0)\}, \cdot, (1, 0))$ jest grupą abelową; w szczególności

$$(a, b)^{-1} = \Big(\frac{a}{a^2 + b^2}, \frac{-b}{a^2 + b^2}\Big).$$

3. Dowód kończy sprawdzenie, że mnożenie jest rozdzielne względem dodawania. Rachunki — na ćwiczeniach.

Formalnie należałoby ciało liczb zespolonych oznaczać symbolem $(\mathbb{C}, +, \cdot, (0, 0), (1, 0))$, ale w skrócie pisze się po prostu $\mathbb{C}$. Ta sama uwaga dotyczy ciał liczb wymiernych i rzeczywistych, $\mathbb{Q}$ i $\mathbb{R}$, a także innych.

Liczbie rzeczywistej a przyporządkujemy liczbę zespoloną $(a, 0)$. Możemy sprawdzić, że zbiór liczb zespolonych o części urojonej równej 0 zawiera zero i jedynkę (tj. liczby $(0, 0)$ i $(1, 0)$), oraz jest zamknięty ze względu na dodawanie i mnożenie, oraz operacje przeciwne (tj. odejmowanie i dzielenie, z wyjątkiem dzielenia przez $(0, 0)$).
Możemy też sprawdzić, że dodawanie i mnożenie zespolone liczb o zerowej części urojonej zgadza się z dodawaniem i mnożeniem liczb rzeczywistych.
Utożsamiwszy w ten sposób liczby rzeczywiste z pewnymi liczbami zespolonymi możemy powiedzieć, że ciało liczb rzeczywistych jest podciałem ciała liczb zespolonych. Dlatego możemy (i będziemy) pisać a zamiast $(a, 0)$, i w szczególności 0 zamiast $(0, 0)$ i 1 zamiast $(1, 0)$.

Def. Liczbę $i = (0, 1)$ nazywamy jednostką urojoną.

Zachodzi równość $i^2 = (-1, 0) = -1$. Korzystając z tego symbolu możemy zapisywać liczby zespolone w postaci $a + bi$ zamiast (a, b).

Uwaga: Właśnie w tym napisie utożsamiliśmy liczbę rzeczywistą a z $(a, 0)$, b z $(b, 0)$; symbol „$+$” oznacza dodawanie *zespolone*!

Def. Wartością bezwzględną liczby zespolonej $z = (a, b)$ nazywamy liczbę rzeczywistą $\sqrt{a^2 + b^2}$. Jest ona zawsze nieujemna, oznacza się ją symbolem $|z|$.

Na wykładzie analizy są definiowane funkcje trygonometryczne $\sin x$, $\cos x$ i funkcja wykładnicza e^x (tu symbol e oznacza liczbę rzeczywistą $\sum_{k=0}^{\infty} \frac{1}{k!} = 2{,}71828\ldots$). Dowodzi się, że

1. Dla dowolnych liczb $c, s \in \mathbb{R}$, takich że $c^2 + s^2 = 1$ istnieje liczba $\varphi \in \mathbb{R}$, taka że $s = \sin\varphi$, $c = \cos\varphi$.
2. Dla dowolnej liczby $\varphi \in \mathbb{R}$ zachodzi równość liczb zespolonych:

$$e^{i\varphi} = (\cos\varphi, \sin\varphi).$$

W związku z tym dowolną liczbę zespoloną można przedstawić w postaci

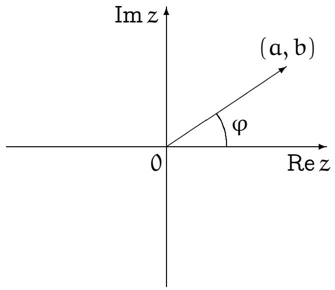
$$z = |z| \cdot (\cos\varphi, \sin\varphi) = |z| \cdot e^{i\varphi}.$$

Liczbę φ, która występuje w powyższym przedstawieniu, nazywamy argumentem liczby zespolonej z i oznaczamy symbolem $\arg z$.

Uwaga: Równość $z = |z| \cdot e^{i \arg z}$ jest uogólnieniem równości $x = |x| \cdot \operatorname{sgn} x$ dla dowolnej liczby rzeczywistej x. Czynnik $e^{i \arg z}$ pełni rolę znaku liczby zespolonej.

Jest oczywiste, że jeśli liczba φ jest argumentem pewnej liczby zespolonej z, to jest nim również każda liczba równa $\varphi + 2k\pi$ dla $k \in \mathbb{Z}$. Tylko jeden argument liczby z jest elementem przedziału $[0, 2\pi)$. Nazwiemy go argumentem głównym liczby zespolonej z i oznaczymy symbolem $\operatorname{Arg} z$.

Interpretacja geometryczna:
Utożsamiamy zbiór liczb zespolonych z płaszczyzną, w której jest określony kartezjański układ współrzędnych. Wartość bezwzględna liczby $z = (a, b)$ jest odległością odpowiedniego punktu od początku układu (tj. punktu, czyli liczby 0). Argument liczby z jest miarą kąta nachylenia odcinka $\overline{0z}$ względem osi rzeczywistej (poziomej). Argument liczby 0 jest nieokreślony.



Jeśli liczby z_1 i z_2 przedstawimy w postaci trygonometrycznej, $z_1 = |z_1| \cdot (c_1, s_1)$ i $z_2 = |z_2| \cdot (c_2, s_2)$, to możemy obliczyć

$$z_1 z_2 = |z_1| \cdot |z_2| \cdot (c_1 c_2 - s_1 s_2, c_1 s_2 + s_1 c_2) = |z| \cdot (c, s).$$

Rozpoznajemy wyżej wyrażenia opisujące cosinus i sinus sumy kątów i wyciągamy wniosek, że:

1. Wartość bezwzględna iloczynu liczb zespolonych jest iloczynem ich wartości bezwzględnych.
2. Argument iloczynu jest sumą argumentów mnożonych liczb.

Prostą konsekwencją tego spostrzeżenia jest poniższy wzór.

Wzór de Moivre'a: jeśli $z = |z| \cdot (\cos\varphi, \sin\varphi)$, to

$$z^n = |z|^n \cdot (\cos n\varphi, \sin n\varphi).$$

Def. Sprzężeniem nazywamy przyporządkowanie liczbie (a, b) liczby $(a, -b)$. Liczbę zespoloną sprzężoną z z oznaczamy symbolem $\overline{z}$.

Możemy zauważyć, że

1. $\overline{\overline{z}} = z$,
2. $\operatorname{Re} z = (z + \overline{z})/2$, $\operatorname{Im} z = (z - \overline{z})/(2i)$,
3. $z \cdot \overline{z} = |z|^2$, $z^{-1} = \dfrac{1}{z} = \dfrac{\overline{z}}{|z|^2}$,
4. $|\overline{z}| = |z|$, $\arg \overline{z} = -\arg z = \arg \dfrac{1}{z}$,
5. jeśli symbol „$\diamond$" oznacza dowolne z czterech działań arytmetycznych (dodawanie, odejmowanie, mnożenie lub dzielenie), to $\overline{z_1 \diamond z_2} = \overline{z}_1 \diamond \overline{z}_2$.
   Wynika stąd, że jeśli $f(z_1, \ldots, z_n)$ jest dowolnym wyrażeniem wymiernym, w którym oprócz $z_1, \ldots, z_n$ nie występują żadne inne liczby zespolone, to $f(\overline{z}_1, \ldots, \overline{z}_n) = \overline{f(z_1, \ldots, z_n)}$.

Wielomiany

Def. Wielomianem zmiennej x nazywamy wyrażenie, w którym oprócz symbolu x pojawiają się dowolne stałe i działania określone w przyjętej strukturze algebraicznej.
Jeśli rozpatrujemy ciało, to możemy mnożyć i dodawać, a zatem każdy wielomian (nad ciałem — o innych nie mówimy) możemy przedstawić w postaci

$$w(x) = a_0 + a_1 x + \cdots + a_n x^n \quad \left(= \sum_{k=0}^{n} a_k x^k\right),$$

dla pewnego $n \in \mathbb{N}$. Jeśli $a_n \neq 0$, to liczba n nazywa się stopniem wielomianu w (umawiamy się, że stopień wielomianu zerowego, $w(x) = 0$, jest równy $-\infty$). Liczby $a_0, \ldots, a_n$ to współczynniki wielomianu w.

Jednym z podstawowych historycznych zadań algebry jest rozwiązywanie równań, tj. znajdowanie liczb x (pierwiastków wielomianu), takich że $w(x) = 0$.

Twierdzenie zasadnicze algebry (Gauss, 1799): *Ciało liczb zespolonych jest algebraicznie domknięte, tj. każdy wielomian stopnia większego niż* 0 *o współczynnikach zespolonych ma pierwiastek zespolony.*
Dowód pomijamy. Nie można go przeprowadzić na gruncie samej algebry, bo gdyby można było, to ten sam dowód można by zastosować do dowolnego ciała, a nie wszystkie ciała są algebraicznie domknięte.

Konsekwencją zasadniczego twierdzenia algebry jest istnienie rozkładu wielomianu stopnia n o współczynnikach zespolonych $a_0, \dots, a_n$ na czynniki pierwszego stopnia:

$$\sum_{k=0}^{n} a_k x^k = a_n (x - x_1) \cdots (x - x_n).$$

Zbiór liczb zespolonych $x_1, \dots, x_n$ — pierwiastków wielomianu — jest określony przez współczynniki $a_0, \dots, a_n$ jednoznacznie.

Na przykład, pierwiastkami wielomianu $w(x) = x^n - 1$ są liczby $x_k = (\cos \frac{2\pi k}{n}, \sin \frac{2\pi k}{n})$ dla $k = 0, \dots, n-1$, co łatwo jest uzasadnić na podstawie wzoru de Moivre'a. Liczby te nazywamy pierwiastkami stopnia n z jedynki.
W ogólnym przypadku znalezienie pierwiastków wielomianu jest trudne.

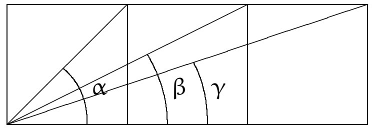
Zadania i problemy na ćwiczenia i do domu

1. Udowodnij, że z założenia, że w zbiorze X istnieje element prawostronnie neutralny e i element lewostronnie neutralny e' działania dwuargumentowego „$\diamond$" wynika, że $e = e'$ i jest tylko jeden element neutralny tego działania.
2. Zbadaj, czy zbiór liczb całkowitych $\mathbb{Z}$ z działaniem „$\diamond$" określonym wzorem $a \diamond b = a + ab + b$ jest grupą.
3. Zbadaj, czy zbiór liczb wymiernych $\mathbb{Q}$, z działaniem „$\diamond$" określonym wzorem takim jak w poprzednim zadaniu, jest grupą.
4. Udowodnij, że zbiór $\mathbb{Z}_n = \{0, 1, \dots, n-1\}$, dla $n > 1$ z działaniami dodawania i mnożenia modulo n jest pierścieniem z jedynką.
   Dla jakich n pierścień ten jest ciałem?
5. Udowodnij, że zbiór liczb rzeczywistych o postaci $a + b\sqrt{2}$, gdzie $a, b \in \mathbb{Q}$, jest podciałem ciała $\mathbb{R}$. Skorzystaj w dowodzie z twierdzenia opisującego warunek konieczny i dostateczny tego, aby podzbiór grupy był podgrupą.
6. Czy można koszulkę (tzw. *T-shirt*) założyć jednocześnie tył na przód, prawa ręka w lewym rękawie a lewa w prawym, tak, aby wewnętrzna strona była na zewnątrz? (zakładamy, że talia nie mieści się w dekolcie). Należy znaleźć ścisłe uzasadnienie odpowiedzi.
7. Def. Niech $i, j \in \{1, \dots, n\}$, $i \neq j$. Transpozycją nazywamy taką permutację T_{ij} zbioru n-elementowego, że $T_{ij}(i) = j$, $T_{ij}(j) = i$ oraz $T_{ij}(k) = k$ dla $k \notin \{i, j\}$.

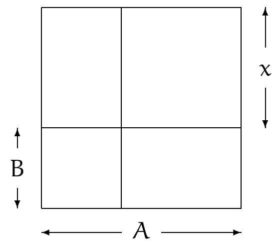
   Def. Nieporządkiem (albo inwersją) permutacji σ nazywamy parę liczb i, j, taką że $i < j$ i $\sigma(i) > \sigma(j)$.

   Udowodnij, że liczba nieporządków permutacji σ jest parzysta wtedy i tylko wtedy, gdy liczba nieporządków permutacji $T_{ij} \circ \sigma$ jest nieparzysta.
8. Def. Zbiór generatorów grupy jest to taki podzbiór grupy, że każdy element grupy można przedstawić w postaci wyrażenia, w którym występują tylko elementy tego zbioru (generatory) i ich elementów odwrotnych.

   Rozważmy dwa zbiory generatorów grupy S_n.
   Pierwszy to $\{T_{12}, T_{23}, \dots, T_{n-1,n}\}$, a drugi to $\{T_{12}, T_{13}, \dots, T_{1n}\}$.
   Ile transpozycji z pierwszego zbioru wystarczy złożyć, aby przedstawić dowolną permutację? A ile transpozycji z drugiego zbioru?
9. Oblicz liczbę $z = (1, -2)^7$ wykonując jak najmniej mnożeń.
10. Figura geometryczna na rysunku obok jest skonstruowana w oparciu o trzy kwadraty. Udowodnij, że suma kątów α, β i γ jest kątem prostym.



11. Napisz wzór wyrażający $\sin 4\alpha$ w zależności od $\sin \alpha$ i $\cos \alpha$.
12. Udowodnij, że zbiór wielomianów trygonometrycznych, tj. funkcji o postaci $f(x) = a_0 + \sum_{k=1}^{n} a_k \cos kx + b_k \sin kx$, $a_k, b_k \in \mathbb{R}$, z działaniami dodawania i mnożenia funkcji, jest pierścieniem.
13. Wskaż wszystkie pierwiastki (zespolone) wielomianu $x^6 + x^5 + x^4 + x^3 + x^2 + x$.
14. Wyprowadź wzory, pozwalające obliczyć część rzeczywistą i urojoną pierwiastka kwadratowego z liczby zespolonej, za pomocą działań arytmetycznych na liczbach rzeczywistych i pierwiastka kwadratowego z liczby rzeczywistej.
15. Geometryczne wyprowadzenie wzorów opisujących rozwiązanie równania algebraicznego drugiego stopnia $x^2 + ax = b$ (wg. *Wykładów z historii matematyki* M. Kordosa): Rozważmy kwadrat podzielony na kwadraty i prostokąty jak na rysunku. Mamy $A^2 = x^2 + 2Bx + B^2$, czyli $x^2 + 2Bx = A^2 - B^2$. Jeśli przyjmiemy $a = 2B$ oraz $b = A^2 - B^2$, to dostaniemy równanie wyjściowe. Aby obliczyć x wystarczy znaleźć liczby A i B. Oczywiście $B = \frac{a}{2}$, zaś $A^2 = b + B^2$, czyli $A = \sqrt{b + B^2}$.



   Opisany wyżej sposób rozwiązywania jest arabski. Średniowieczni Arabowie nie znali liczb ujemnych (ani tym bardziej zespolonych), więc nie wiedzieli, że można wziąć też $A = -\sqrt{b + B^2}$.
16. Geometryczne rozwiązanie równania trzeciego stopnia $x^3 + ax = b$ (Tartaglia, 1535, *op. cit.*): Sześcian o boku A dzielimy na sześcian o boku B, sześcian o boku x, takim że $A = B + x$ i trzy jednakowe prostopadłościany o bokach A, B, x.

Zatem $A^3 = x^3 + 3ABx + B^3$, skąd $x^3 + 3ABx = A^3 - B^3$. Równanie rozwiążemy, jeśli znajdziemy takie liczby A i B, że $3AB = a$ i $A^3 - B^3 = b$.

Oznaczamy $p = A^3$, $q = B^3$ i stąd $pq = \frac{a^3}{27}$, $p - q = b$. Podstawiając $p = b + q$ do pierwszego równania dostajemy równanie kwadratowe $q^2 + bq = \frac{a^3}{27}$ i dalej każdy umie.

17. Udowodnij, że podana niżej procedura dla dowolnej liczby naturalnej b oblicza $z = a^b$. Działanie div to dzielenie całkowite z odrzuceniem reszty.

```
x := a;   e := b;   z := 1;
repeat
  if nieparzyste(e) then z := z*x
  e := e div 2;
  if e = 0 then goto 1;
  x := x^2;
until false;
1:
```

Określ liczbę mnożeń wykonywaną przez tę procedurę (podnoszenie do kwadratu to też jedno mnożenie).

Macierze

Def. Macierzą liczbową o wymiarach $m \times n$ nazywamy układ liczb a_{ij} dla $i = 1, \dots, m$, $j = 1, \dots, n$. Liczby te nazywamy współczynnikami macierzy.

Często macierz przedstawia się w postaci prostokątnej tabeli:

$$\begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix},$$

w związku z czym liczba m jest liczbą wierszy, a n jest liczbą kolumn macierzy.

Krótszy zapis, którego będziemy używali gdy skądinąd znamy wymiary macierzy, to $[a_{ij}]_{i,j}$ (po nawiasie pierwsza litera określa indeks wiersza, a druga — kolumny). Macierze będziemy też oznaczać dużymi literami, np. A, B.

Macierze można tworzyć nie tylko z liczb, ale także z funkcji, zbiorów i innych obiektów. Będziemy wykonywać na współczynnikach macierzy działania — dodawanie, odejmowanie, mnożenie (czasem także dzielenie), a zatem założymy co najmniej, że współczynniki należą do pewnego pierścienia lub ciała. Na razie rozpatrujemy tylko macierze liczbowe.

Jeśli ciało, do którego należą współczynniki, oznaczymy literą $\mathbb{K}$, to zbiór wszystkich macierzy o wymiarach $m \times n$ oznaczymy symbolem $\mathbb{K}^{m,n}$.

Jeśli $n = 1$, czyli macierz ma tylko jedną kolumnę, to będziemy ją nazywać macierzą kolumnową, albo wektorem (pojęcie wektora później uogólnimy). Zbiór wszystkich macierzy kolumnowych o m wierszach oznaczymy symbolem $\mathbb{K}^m$. Wektory będą też symbolizowane przez małe tłuste litery, np. $\mathbf{a}$, $\mathbf{b}$.

Dla $m = 1$ mamy macierz wierszową.

Jeśli $m = n = 1$, czyli macierz ma tylko 1 współczynnik, to utożsamiamy ją z tym współczynnikiem. Zatem, $\mathbb{K}^{1,1} = \mathbb{K}^1 = \mathbb{K}$.

Def. Transpozycja macierzy jest to przekształcenie $\mathbb{K}^{m,n} \to \mathbb{K}^{n,m}$, które macierzy $A = [a_{ij}]_{i,j}$ przyporządkowuje macierz $A^T = [a_{ij}]_{j,i}$. Macierz A^T nazywa się macierzą transponowaną do A.

Na przykład

$$\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \\ a_{31} & a_{32} \end{bmatrix}^T = \begin{bmatrix} a_{11} & a_{21} & a_{31} \\ a_{12} & a_{22} & a_{32} \end{bmatrix}.$$

Def. Sprzężenie hermitowskie jest przekształceniem $\mathbb{C}^{m,n} \to \mathbb{C}^{n,m}$, które macierzy $A = [a_{ij}]_{i,j}$ przyporządkowuje macierz $A^H = [\overline{a}_{ij}]_{j,i}$.

Oczywiste są tożsamości $(A^T)^T = A$, $(A^H)^H = A$.
Jeśli zbiory $\mathbb{Q}$ i $\mathbb{R}$ traktujemy jak podzbiory $\mathbb{C}$, czyli zbiory $\mathbb{Q}^{m,n}$ i $\mathbb{R}^{m,n}$ jak podzbiory $\mathbb{C}^{m,n}$, to sprzężenie hermitowskie w tych podzbiorach jest identyczne z transpozycją.

Działania dwuargumentowe w $\mathbb{K}^{m,n}$: dodawanie i odejmowanie macierzy, są określone wzorami

$$[a_{ij}]_{i,j} + [b_{ij}]_{i,j} \stackrel{\text{def}}{=} [a_{ij} + b_{ij}]_{i,j},$$
$$[a_{ij}]_{i,j} - [b_{ij}]_{i,j} \stackrel{\text{def}}{=} [a_{ij} - b_{ij}]_{i,j}.$$

Zbiór $\mathbb{K}^{m,n}$ z działaniem dodawania jest grupą abelową, której elementem neutralnym jest macierz zerowa $0 = [0]_{i,j}$.
Element przeciwny do $A = [a_{ij}]_{i,j}$ to macierz $-A = [-a_{ij}]_{i,j}$.
Ponadto, mamy $(A \pm B)^T = A^T \pm B^T$, $(A \pm B)^H = A^H \pm B^H$.

Operacja mnożenia macierzy $A = [a_{ij}]_{i,j} \in \mathbb{K}^{m,n}$ i $B = [b_{ij}]_{i,j} \in \mathbb{K}^{n,l}$ jest określona w następujący sposób: Iloczynem macierzy A i B jest macierz $C = [c_{ij}]_{i,j} \in \mathbb{K}^{m,l}$, której współczynniki są równe

$$c_{ij} = a_{i1}b_{1j} + \dots + a_{in}b_{nj} = \sum_{k=1}^{n} a_{ik}b_{kj}.$$

W przyswojeniu sobie tego wzoru może pomóc taki schemat:

$$\begin{array}{ccc} \begin{array}{c} B \\ \downarrow \\ A \;\to\; C \end{array} & & \begin{bmatrix} b_{1j} \\ \vdots \\ b_{nj} \end{bmatrix} \\ & \begin{bmatrix} \\ a_{i1} \dots a_{in} \\ \\ \end{bmatrix} & \begin{bmatrix} \vdots \\ \dots \; c_{ij} \; \dots \\ \vdots \end{bmatrix} \end{array}$$

Własności algebraiczne mnożenia macierzy

1. Jeśli $m = n = l$, a zatem rozpatrujemy macierze kwadratowe, to mnożenie macierzy jest działaniem dwuargumentowym wewnętrznym w zbiorze $\mathbb{K}^{n,n}$.
W ogólności jest to działanie nieprzemienne, np.
$$\begin{bmatrix}1&2\\3&0\end{bmatrix}\cdot\begin{bmatrix}1&0\\0&2\end{bmatrix}=\begin{bmatrix}1&4\\3&0\end{bmatrix},\qquad \begin{bmatrix}1&0\\0&2\end{bmatrix}\cdot\begin{bmatrix}1&2\\3&0\end{bmatrix}=\begin{bmatrix}1&2\\6&0\end{bmatrix}.$$

2. Mnożenie macierzy jest łączne.
Dowód: Dodawanie współczynników jest łączne i przemienne, mnożenie współczynników jest łączne i rozdzielne względem dodawania. Możemy zatem przyjąć $A = [a_{ij}]_{i,j} \in \mathbb{K}^{m,n}$, $B = [b_{ij}]_{i,j} \in \mathbb{K}^{n,l}$ i $C = [c_{ij}]_{i,j} \in \mathbb{K}^{l,s}$ i obliczyć
$$(AB)C = [d_{ij}]_{i,j} \in \mathbb{K}^{m,s}\colon\ d_{ij} = \sum_{k=1}^{l}\Big(\sum_{t=1}^{n} a_{it}\cdot b_{tk}\Big)\cdot c_{kj} = \sum_{k=1}^{l}\sum_{t=1}^{n} a_{it}\cdot b_{tk}\cdot c_{kj},$$
$$A(BC) = [e_{ij}]_{i,j} \in \mathbb{K}^{m,s}\colon\ e_{ij} = \sum_{t=1}^{n} a_{it}\cdot\Big(\sum_{k=1}^{l} b_{tk}\cdot c_{kj}\Big) = \sum_{t=1}^{n}\sum_{k=1}^{l} a_{it}\cdot b_{tk}\cdot c_{kj}.$$
Ponieważ $d_{ij} = e_{ij}$ dla każdego i, j, więc $(AB)C = A(BC)$. Zauważmy, że w dowodzie nie trzeba zakładać, że mnożenie współczynników jest przemienne. □
Uwaga: Wyrażenie ABC, dla ustalonych macierzy A, B i C, opisuje ten sam wynik niezależnie od kolejności wykonywania działań. W każdym przypadku inne są wyniki pośrednie i może być ogromna różnica w ilości działań, które trzeba wykonać. Na przykład, jeśli $A \in \mathbb{R}^n$, $B \in \mathbb{R}^{1,n}$ i $C \in \mathbb{R}^n$, to obliczanie $(AB)C$ wymaga wykonania $2n^2$ mnożeń, podczas gdy obliczając $A(BC)$ wykonamy tylko $2n$ mnożeń.

3. Niech I_n oznacza macierz jednostkową $n \times n$; $I_n = [\delta_{ij}]_{i,j}$.
Użyty w tym określeniu symbol Kroneckera δ_{ij} oznacza 0 dla $i \neq j$ i 1 jeśli $i = j$.

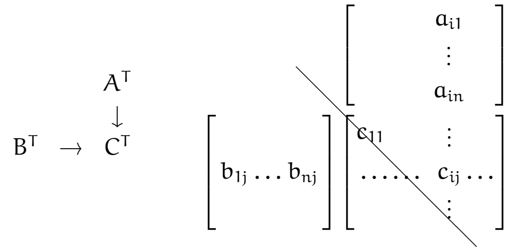
   Macierz I_n jest prawostronnie neutralnym argumentem mnożenia, tj. dla dowolnej macierzy $A \in \mathbb{K}^{m,n}$ zachodzi równość $AI_n = A$.
Podobnie, macierz I_m jest argumentem lewostronnie neutralnym, tj. $I_mA = A$.
Żadna inna macierz w $\mathbb{K}^{n,n}$ ani w $\mathbb{K}^{m,m}$ nie ma tej własności (ale istnieją macierze A oraz B i C inne niż jednostkowe, takie że $AB = A$ i $CA = A$).

   Zbiór macierzy kwadratowych, $\mathbb{K}^{n,n}$, z działaniem mnożenia, jest półgrupą, której elementem neutralnym jest macierz jednostkowa I_n.

4. Mnożenie macierzy jest rozdzielne względem dodawania, tj.
$$(A+B)C = AC+BC,\qquad A(B+C) = AB+AC.$$

Powyższe własności działań na macierzach oznaczają, że zbiór macierzy kwadratowych $\mathbb{K}^{n,n}$ z działaniami dodawania i mnożenia jest pierścieniem. Jego elementy wyróżnione to macierz zerowa (zero) i macierz jednostkowa (jedynka).

5. Transpozycja i sprzężenie hermitowskie iloczynu macierzy wyrażają się wzorami
$$(AB)^T = B^TA^T,\qquad (AB)^H = B^HA^H,$$
Trzeba przy tym założyć przemienność mnożenia współczynników macierzy.
Aby to uzasadnić, zamiast wykonywać rachunki, „odbijemy” schemat mnożenia macierzy względem diagonali wyniku, tj. linii ukośnej, na której leżą współczynniki c_{ii}.



Aby dokończyć dowód drugiego wzoru wystarczy przypomnieć, że jeśli $f(x_1,\ldots,x_n)$ jest dowolnym wyrażeniem wymiernym, w którym oprócz zmiennych $x_1,\ldots,x_n$ nie występują żadne inne liczby zespolone, to $f(\overline{x}_1,\ldots,\overline{x}_n) = \overline{f(x_1,\ldots,x_n)}$.

6. Jeśli argumenty są macierzami 1×1, to dodawanie i mnożenie macierzy jest zgodne z dodawaniem i mnożeniem odpowiadających im liczb, co usprawiedliwia utożsamienie $\mathbb{K}^{1,1} = \mathbb{K}^1 = \mathbb{K}$.
Transpozycja macierzy 1×1 jest przekształceniem identycznościowym, a sprzężenie hermitowskie jest sprężeniem zespolonym.

Def. Mnożenie macierzy przez liczbę jest działaniem $\mathbb{K} \times \mathbb{K}^{m,n} \to \mathbb{K}^{m,n}$, określonym wzorem
$$a \cdot [a_{ij}]_{i,j} \stackrel{\text{def}}{=} [a \cdot a_{ij}]_{i,j}.$$

Jeśli ciało $\mathbb{K}$ jest przemienne, to czynniki, tj. liczbę a i macierz A możemy pisać w dowolnej kolejności, a więc $aA = Aa$. Ponadto prawdziwe są równości:
$$(a+b)A = aA+bA,\qquad a(A+B) = aA+aB,$$
$$(a\cdot b)\cdot A = a\cdot(b\cdot A),$$
$$(aA)^T = aA^T,\qquad (aA)^H = \overline{a}A^H.$$

Podział blokowy macierzy

W wielu sytuacjach wygodnie jest wyróżnić w macierzy <u>bloki</u>, czyli macierze, które powstają przez odrzucenie pewnej liczby początkowych i końcowych wierszy i kolumn.

<u>Macierz blokowa</u> jest to macierz, której współczynnikami są macierze (bloki), takie że

- w każdym wierszu wszystkie bloki mają tyle samo wierszy,
- w każdej kolumnie wszystkie bloki mają tyle samo kolumn.

Będziemy pisać $A = [A_{ij}]_{i,j}$. Na przykład

$$\left[\begin{array}{cc|cc} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ \hline a_{31} & a_{32} & a_{33} & a_{34} \end{array}\right] = \left[\begin{array}{c|c} A_{11} & A_{12} \\ \hline A_{21} & A_{22} \end{array}\right], \qquad A_{11} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \quad \text{itd.}$$

Wszystkie bloki w szczególności mogą, ale nie muszą, mieć te same wymiary. „Zwykłe" macierze to macierze blokowe z wszystkimi blokami 1×1.

Dowolną macierz A możemy przedstawić w postaci <u>blokowo-kolumnowej</u>, $A = \begin{bmatrix} A_1 \\ \vdots \\ A_m \end{bmatrix}$, albo <u>blokowo-wierszowej</u>, $A = [A_1, \dots, A_n]$.

Jeśli macierze A i B o takich samych wymiarach podzielimy w identyczny sposób na bloki, to możemy napisać wzór na dodawanie macierzy w postaci blokowej:

$$[A_{ij}]_{i,j} + [B_{ij}]_{i,j} = [A_{ij} + B_{ij}]_{i,j}.$$

Poniższy wzór opisujący mnożenie macierzy obowiązuje wtedy, gdy wymiary odpowiednich bloków umożliwiają ich mnożenie:

$$[A_{ij}]_{i,j} \cdot [B_{ij}]_{i,j} = [C_{ij}]_{i,j}, \qquad C_{ij} = \sum_{k=1}^{p} A_{ik} \cdot B_{kj}.$$

Liczba p jest liczbą bloków w wierszu macierzy A i w kolumnie B. Nietrudno jest udowodnić, że powyższy wzór jest równoważny podanemu wcześniej wzorowi na „pojedynczy" współczynnik iloczynu macierzy.

Możemy na nowo zinterpretować „zwykłe" mnożenie macierzy. Macierz A przedstawimy jako blokowo-kolumnową (z blokami o jednym wierszu), a macierz B

jako blokowo-wierszową (z blokami o jednej kolumnie). Wtedy

$$\begin{array}{ccc} & & B \\ & & \downarrow \\ A & \rightarrow & C \end{array} \qquad \begin{array}{cc} & \begin{bmatrix} & B_j & \end{bmatrix} \\ \begin{bmatrix} \\ A_i \\ \\ \end{bmatrix} & \begin{bmatrix} & \vdots & \\ \dots & c_{ij} & \dots \\ & \vdots & \end{bmatrix} \end{array} \qquad c_{ij} = A_i B_j.$$

Szczególne klasy macierzy

<u>Def.</u> <u>Macierz trójkątna górna</u> jest to macierz $[a_{ij}]_{i,j}$ spełniająca warunek $a_{ij} = 0$ dla $i > j$.

<u>Def.</u> <u>Macierz trójkątna dolna</u> jest to macierz $[a_{ij}]_{i,j}$, której współczynniki spełniają warunek $a_{ij} = 0$ dla $i < j$.

<u>Def.</u> <u>Macierz diagonalna</u> jest to macierz $[a_{ij}]_{i,j}$, której współczynniki a_{ij} dla $i \neq j$ są równe 0.
Macierz diagonalna jest więc zarówno macierzą trójkątną górną, jak i dolną.
Będziemy ją oznaczać symbolem $\mathrm{diag}(a_{11}, \dots, a_{kk})$ albo $\mathrm{diag}(a_1, \dots, a_k)$, gdzie $k = \min(m, n)$.

Łatwo jest dowieść (na ćwiczeniach), że suma i iloczyn dwóch macierzy trójkątnych górnych jest macierzą trójkątną górną. Takie samo stwierdzenie dotyczy macierzy trójkątnych dolnych, a także macierzy diagonalnych.

Przyjmijmy oznaczenia $\boldsymbol{e}_1, \dots, \boldsymbol{e}_n$ dla wektorów (macierzy kolumnowych), takich że $\boldsymbol{e}_k = [0, \dots, 0, \underset{\substack{\uparrow \\ \text{pozycja } k}}{1}, 0, \dots, 0]^T$.

Macierz diagonalną możemy przedstawić w postaci blokowo-wierszowej, $\mathrm{diag}(a_1, \dots, a_n) = [a_1\boldsymbol{e}_1, \dots, a_n\boldsymbol{e}_n]$, albo blokowo-kolumnowej,

$$\mathrm{diag}(a_1, \dots, a_m) = \begin{bmatrix} a_1\boldsymbol{e}_1^T \\ \vdots \\ a_m\boldsymbol{e}_m^T \end{bmatrix}.$$

W szczególności wektory $\boldsymbol{e}_1, \dots, \boldsymbol{e}_n$ są kolumnami macierzy jednostkowej I_n.

<u>Def.</u> <u>Macierz permutacji</u> jest to macierz kwadratowa $n \times n$, która w każdym wierszu, a także w każdej kolumnie, ma jeden współczynnik równy 1,

a pozostałe 0. Nazwa bierze się stąd, że jeśli pomnożymy przez taką macierz wektor, np.

$$\begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} x_2 \\ x_1 \\ x_3 \end{bmatrix}$$

to otrzymamy wektor, który ma te same współczynniki, ustawione w innej kolejności.

Macierz P_σ, taką że dla dowolnej macierzy $\boldsymbol{x} = [x_1, \dots, x_n]^T$ iloczyn $\boldsymbol{y} = P_\sigma \boldsymbol{x} = [y_1, \dots, y_n]^T$ spełnia warunek $y_{\sigma(i)} = x_i$ dla każdego $i \in \{1, \dots, n\}$, nazywamy macierzą permutacji σ.
Macierz P_σ możemy przedstawić w postaci blokowo-wierszowej, a także blokowo-kolumnowej:

$$P_\sigma = [\boldsymbol{e}_{\sigma(1)}, \dots, \boldsymbol{e}_{\sigma(n)}] = \begin{bmatrix} \boldsymbol{e}^T_{\sigma^{-1}(1)} \\ \vdots \\ \boldsymbol{e}^T_{\sigma^{-1}(n)} \end{bmatrix}.$$

Łatwo jest sprawdzić, że dla dowolnej permutacji σ macierz P_σ^T jest odwrotnością macierzy P_σ; jest to macierz permutacji odwrotnej do σ, $P_{\sigma^{-1}}$.

Def. Macierz nieosobliwa jest to macierz kwadratowa $A \in \mathbb{K}^{n,n}$, dla której istnieje macierz odwrotna, A^{-1}, taka że $AA^{-1} = I_n$.
Macierz kwadratowa, która nie ma odwrotności nazywa się macierzą osobliwą.

Zbiór macierzy nieosobliwych z działaniem mnożenia spełnia podane na poprzednim wykładzie aksjomaty grupy, a zatem jest grupą (nieabelową).
Zauważmy, że dla dowolnej macierzy nieosobliwej A mamy $A^{-1}A = I_n$.
Warunki, które zapewniają nieosobliwość macierzy będą podane później, a tymczasem możemy wskazać dwie klasy macierzy nieosobliwych: macierze diagonalne, których współczynniki na diagonali są różne od 0 i macierze permutacji. Obie te klasy macierzy stanowią podgrupy grupy macierzy nieosobliwych $n \times n$.

Def. Macierz symetryczna jest to macierz spełniająca warunek $A^T = A$.
Macierz hermitowska jest to macierz spełniająca warunek $A^H = A$.

Obie te klasy macierzy są zamknięte ze względu na dodawanie, ale nie ze względu na mnożenie.

Def. Macierz antysymetryczna jest to macierz spełniająca warunek $A^T = -A$.

Odnotujmy jeszcze trzy łatwe do udowodnienia wzory (zakładamy, że występujące w nich macierze są nieosobliwe):

$$(A \cdot B)^{-1} = B^{-1} \cdot A^{-1},$$
$$(A^{-1})^T = (A^T)^{-1},$$
$$(A^{-1})^H = (A^H)^{-1}.$$

Zamiast pisać $(A^{-1})^T$ lub $(A^{-1})^H$, czasem w skrócie pisze się A^{-T} albo A^{-H}.

Układ równań liniowych

Def. Algebraicznym równaniem liniowym nazywamy równanie o postaci

$$a_1x_1 + a_2x_2 + \cdots + a_nx_n = b,$$

w którym dane są liczby (elementy ustalonego ciała $\mathbb{K}$) $a_1, \dots, a_n$ i b.
Symbole $x_1, \dots, x_n$ nazywamy niewiadomymi.
Rozwiązaniem równania jest dowolny ciąg liczb $x_1, \dots, x_n$, który spełnia to równanie.

Def. Układem równań liniowych nazywamy pewien zbiór równań liniowych, w których występują te same niewiadome.
Zbiorem rozwiązań układu równań jest część wspólna zbiorów rozwiązań poszczególnych równań.
Układ jest sprzeczny, jeśli nie ma rozwiązania, albo niesprzeczny, jeśli zbiór jego rozwiązań jest niepusty.
Układ niesprzeczny jest nieokreślony jeśli ma więcej niż jedno rozwiązanie i określony, jeśli ma tylko jedno.

Układ równań liniowych możemy przedstawić w postaci

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1, \\ \dots\dots\dots\dots\dots\dots\dots\dots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n = b_m. \end{cases}$$

Jeśli oznaczymy macierz współczynników

$$A = \begin{bmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \dots & a_{mn} \end{bmatrix},$$

oraz wektor prawej strony i wektor niewiadomych

$$\mathbf{b} = \begin{bmatrix} b_1 \\ \vdots \\ b_m \end{bmatrix}, \qquad \mathbf{x} = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix},$$

to układ równań liniowych możemy przedstawić w postaci macierzowej:

$$A\mathbf{x} = \mathbf{b}.$$

Zanim udowodnimy ogólne twierdzenia dotyczące rozwiązań układów równań liniowych, będziemy takie układy rozwiązywali „na piechotę". Zauważmy, że jeśli macierz A jest kwadratowa $n \times n$ i nieosobliwa, to mnożąc stronami równanie $A\mathbf{x} = \mathbf{b}$ przez A^{-1} otrzymujemy $A^{-1}\mathbf{b} = A^{-1}A\mathbf{x} = I_n\mathbf{x} = \mathbf{x}$. To przekształcenie często stosuje się w rachunkach symbolicznych (i w dowodach twierdzeń), ale praktyczne metody numeryczne rozwiązywania układów równań *nie polegają* na wyznaczaniu odwrotności macierzy A.

Zadania i problemy na ćwiczenia i do domu

1. Udowodnij, że jeśli P_1 i P_2 są macierzami permutacji σ_1 i σ_2, to macierz P_2P_1 jest macierzą permutacji $\sigma_2 \circ \sigma_1$.

   Uwaga: Przyjmujemy tu umowę notacyjną, że $(\sigma_2 \circ \sigma_1)(i) = \sigma_2(\sigma_1(i))$ dla każdego i.
2. Udowodnij, że jeśli P jest macierzą permutacji, to $P^{-1} = P^T$.
3. Udowodnij, że odwrotność kwadratowej macierzy trójkątnej górnej (jeśli istnieje) jest macierzą trójkątną górną i podaj warunek konieczny i dostateczny istnienia odwrotności.

   Zrób to samo zadanie po zmienieniu słów „górnej" i „górną" na „dolnej" i „dolną".
4. Def. Śladem macierzy kwadratowej A nazywamy liczbę, która jest sumą współczynników na diagonali tej macierzy. Ślad macierzy oznaczamy symbolem $\operatorname{tr} A$.

   Niech $A, B \in \mathbb{K}^{m,n}$. Udowodnij, że $\operatorname{tr}(A^TB) = \operatorname{tr}(BA^T)$.
5. Zbadaj, jaki skutek ma pomnożenie dowolnej macierzy A przez macierz diagonalną, a także przez macierz permutacji.

   Wskazówka: Jeśli macierz diagonalna lub macierz permutacji jest czynnikiem z lewej strony, to przedstaw macierz A w postaci blokowo-kolumnowej. W przeciwnym razie przedstaw ją w postaci blokowo-wierszowej.

6. Udowodnij, że zbiór macierzy rzeczywistych 2×2, o postaci $\begin{bmatrix} a & -b \\ b & a \end{bmatrix}$, z działaniami dodawania i mnożenia, jest ciałem.

   Def. Homomorfizmem struktur algebraicznych X_1 i X_2 nazywamy funkcję φ, której dziedziną jest zbiór elementów struktury X_1, której wartości są elementami struktury X_2 i która dla każdego działania f określonego w X_1 spełnia warunek $\varphi(f(a_1, \ldots, a_k)) = g(\varphi(a_1), \ldots, \varphi(a_1))$, gdzie g oznacza działanie w X_2 odpowiadające działaniu f w X_1.
   Na przykład homomorfizm pierścieni X_1 i X_2 oznacza spełnienie warunków $\forall_{a,b \in X_1}\ \varphi(a+b) = \varphi(a) + \varphi(b),\ \varphi(a \cdot b) = \varphi(a) \cdot \varphi(b)$.

   Def. Izomorfizm struktur algebraicznych X_1 i X_2 jest to bijekcja (odwzorowanie różnowartościowe i „na"), która jest homomorfizmem.

   Wskaż inne ciało izomorficzne z ciałem macierzy 2×2 o podanej wyżej strukturze. Wskaż jakieś jego podciała.
7. Oblicz macierz $A^2 - 5A - 2I_2$, gdzie $A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$.

   Użyj w tym celu schematu Hornera: $(A - 5I_2)A - 2I_2$.
8. Znajdź zbiory rozwiązań układów równań liniowych

$$\begin{bmatrix} 1 & 0 & 2 \\ 0 & -2 & 3 \\ 0 & 0 & 4 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 7 \\ 5 \\ 12 \end{bmatrix}, \qquad \begin{bmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 2 & 0 & 0 \\ 1 & -1 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 2 \\ -1 \end{bmatrix}.$$

9. Rozwiąż układ równań

$$\begin{bmatrix} (-1,1) & (0,2) \\ (0,0) & (2,1) \end{bmatrix} \begin{bmatrix} z_1 \\ z_2 \end{bmatrix} = \begin{bmatrix} (-5,-1) \\ (-1,2) \end{bmatrix},$$

10. Udowodnij, że dowolną macierz kwadratową A można jednoznacznie przedstawić w postaci sumy macierzy symetrycznej i antysymetrycznej (uwaga: zakładamy, że współczynniki macierzy są elementami pewnego ciała, które nie może mieć tzw. charakterystyki równej 2, czyli nie może w tym ciele być $1 + 1 = 0$; charakterystyka ciał liczbowych $\mathbb{Q}$, $\mathbb{R}$ i $\mathbb{C}$ jest równa 0).
    Przedstaw macierz A^T za pomocą tych macierzy.
11. Def. Macierz kwadratowa A nazywa się macierzą stochastyczną, jeśli jej współczynniki są nieujemne i suma współczynników w każdym wierszu jest równa 1. Wykaż, że jeśli macierze $A = [a_{ij}]_{i,j}$ i $B = [b_{ij}]_{i,j} \in \mathbb{K}^{n,n}$ są stochastyczne, to ich iloczyn też.

Wskazówka: Zbadaj, czym jest wektor Az, jeśli macierz A jest stochastyczna, a $z = [1, \dots, 1]^T$.

Czy zbiór macierzy stochastycznych $n \times n$ z działaniem mnożenia jest grupą?

12. Niech A będzie macierzą nieosobliwą $n \times n$. Wzór Shermana-Morrisona stwierdza, że jeśli $B = A + \boldsymbol{u}\boldsymbol{v}^T$ dla $\boldsymbol{u}, \boldsymbol{v} \in \mathbb{R}^n$ i macierz B jest nieosobliwa, to

$$B^{-1} = A^{-1} - aA^{-1}\boldsymbol{u}\boldsymbol{v}^T A^{-1},$$

dla pewnego a. Znajdź a.

Rozwiązanie: Jeśli wzór jest prawdziwy, to

$$I_n = (A + \boldsymbol{u}\boldsymbol{v}^T)(A^{-1} - aA^{-1}\boldsymbol{u}\boldsymbol{v}^T A^{-1}) = I_n + \big((1-a)I_n - a\boldsymbol{u}\boldsymbol{v}^T A^{-1}\big)\boldsymbol{u}\boldsymbol{v}^T A^{-1}$$

i równanie to powinno dać się rozwiązać ze względu na a. Możemy przyjąć, że $\boldsymbol{u} \neq \boldsymbol{0}$ i $\boldsymbol{v} \neq \boldsymbol{0}$, bo w przeciwnym razie $A = B$ i wzór jest oczywiście prawdziwy. Jeśli wzór jest prawdziwy, to wyrażenie

$$\big((1-a)I_n - a\boldsymbol{u}\boldsymbol{v}^T A^{-1}\big)\boldsymbol{u}\boldsymbol{v}^T A^{-1} = \boldsymbol{u}\big((1-a) - a\boldsymbol{v}^T A^{-1}\boldsymbol{u}\big)\boldsymbol{v}^T A^{-1} \qquad (*)$$

musi być macierzą zerową w $\mathbb{R}^{n,n}$. Wartościami wyrażeń $\boldsymbol{u}^T\boldsymbol{u}$ i $\boldsymbol{v}^T\boldsymbol{v}$ są liczby dodatnie. Możemy zatem pomnożyć wyrażenie (*) z lewej strony przez $\frac{1}{u^T u}\boldsymbol{u}^T$ i z prawej strony przez $A\frac{1}{v^T v}\boldsymbol{v}$. Po uproszczeniu otrzymamy wyrażenie, którego wartością musi być liczba 0:

$$(1-a) - a\boldsymbol{v}^T A^{-1}\boldsymbol{u} = 0.$$

Stąd łatwo wynika, że $a = 1/(1 + \boldsymbol{v}^T A^{-1}\boldsymbol{u})$.

Wzór S.-M. opisuje odwrotność macierzy B, jeśli $\boldsymbol{v}^T A^{-1}\boldsymbol{u} \neq -1$.

13. Oblicz macierz B i na podstawie wzoru Shermana-Morrisona macierz B^{-1}, jeśli

$$B = A + \boldsymbol{u}\boldsymbol{v}^T, \quad A = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}, \quad \boldsymbol{u} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \quad \boldsymbol{v} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}.$$

14. Udowodnij podany niżej wzór Shermana-Morrisona-Woodbury'ego: jeśli $U, V \in \mathbb{R}^{n,m}$ i macierze $A \in \mathbb{K}^{n,n}$ oraz $(I_m + V^T A^{-1} U) \in \mathbb{R}^{m,m}$ są nieosobliwe, to

$$(A + UV^T)^{-1} = A^{-1} - A^{-1}U(I_m + V^T A^{-1}U)^{-1}V^T A^{-1}.$$

Przestrzenie liniowe

Określenie przestrzeni liniowej

Def. Ustalmy dowolne ciało $\mathbb{K}$; jego elementy będziemy nazywać skalarami. Wektorami nazywamy obiekty, które możemy dodawać (chodzi o dwuargumentowe działanie wewnętrzne w ustalonym zbiorze wektorów, które jest określone w sposób zależny od tych obiektów) i mnożyć przez skalary.

Def. Zbiór V wektorów nazywamy przestrzenią liniową nad ciałem $\mathbb{K}$, jeśli spełnia on następujące warunki:

S.1: Zbiór V z działaniem dodawania wektorów jest grupą abelową. Element neutralny tej grupy nazywamy wektorem zerowym i oznaczamy symbolem $\mathbf{0}$.

S.2: Zbiór V jest zamknięty ze względu na mnożenie wektorów przez skalary.

S.3: Działanie mnożenia skalarów i wektorów jest rozdzielne względem dodawania wektorów, a także dodawania skalarów:

$$\forall_{a\in\mathbb{K},\, x,y\in V}\ a\cdot(x+y) = a\cdot x + a\cdot y,$$
$$\forall_{a,b\in\mathbb{K},\, x\in V}\ (a+b)\cdot x = a\cdot x + b\cdot x.$$

S.4: Zachodzi łączność mnożenia skalarów i wektorów:

$$\forall_{a,b\in\mathbb{K},\, x\in V}\ (ab)\cdot x = a\cdot(b\cdot x),$$

S.5: $\forall_{x\in V}\ 1\cdot x = x$.

Formalnie przestrzeń liniowa jest piątką spełniających powyższe aksjomaty elementów, $(V,\mathbb{K},+,\cdot,\mathbf{0})$, którymi są zbiór wektorów V, ciało skalarów $\mathbb{K}$, działanie dodawania wektorów „+”, działanie mnożenia skalara i wektora „·” i wektor zerowy $\mathbf{0}$.
Mówiąc o elemencie ustalonej przestrzeni liniowej mamy na myśli wektor (element zbioru V).

Przykłady:

- Dowolne ciało $\mathbb{K}$ jest przestrzenią liniową nad ciałem $\mathbb{K}$, a także nad dowolnym swoim podciałem (uwaga: dla każdego podciała otrzymujemy *inną* przestrzeń liniową!).
- Zbiór macierzy liczbowych o ustalonych wymiarach, $\mathbb{K}^{m,n}$, z działaniami dodawania i mnożenia przez liczbę określonymi w poprzednim wykładzie. W szczególności zbiór macierzy kolumnowych $\mathbb{K}^m$ jest przestrzenią liniową, o na tyle ważnym znaczeniu, że przez słowo „wektor” bywa rozumiany domyślnie element tej przestrzeni.

- Zbiór przesunięć płaszczyzny (wektorem jest operacja przesunięcia w danym kierunku na określoną odległość, a dodawaniem wektorów jest operacja złożenia takich przekształceń).
- Zbiór wszystkich równań liniowych wiążących określone niewiadome.
- Zbiór funkcji rzeczywistych określonych w ustalonej dziedzinie, z działaniami określonymi w „naturalny” sposób.
- Zbiór $\mathbb{R}[x]_n$ wielomianów stopnia nie większego niż ustalone n z działaniami j.w. (uwaga: zbiór wielomianów ustalonego stopnia $n\in\mathbb{N}$ *nie jest* przestrzenią liniową).

Def. Podzbiór X przestrzeni liniowej V nazywamy podprzestrzenią liniową, jeśli jest on przestrzenią liniową (nad tym samym ciałem, z działaniami otrzymanymi z obcięcia działań w przestrzeni V).

W dowolnej przestrzeni liniowej V możemy natychmiast wskazać dwie podprzestrzenie: podprzestrzeń zerową, której jedynym elementem jest wektor zerowy, i podprzestrzeń niewłaściwą, czyli całą przestrzeń V.

Def. Kombinacją liniową wektorów $x_1,\ldots,x_k\in V$ o współczynnikach $a_1,\ldots,a_k\in\mathbb{K}$ nazywamy wektor $a_1x_1+\cdots+a_kx_k$.

Def. Przestrzenią rozpiętą przez zbiór wektorów $\{x_1,\ldots,x_k\}$ nazywamy zbiór wszystkich kombinacji liniowych tych wektorów.
Oznaczamy ją symbolem $\mathrm{lin}\{x_1,\ldots,x_k\}$ albo $\mathrm{span}\{x_1,\ldots,x_k\}$.

Łatwo jest sprawdzić, że przestrzeń rozpięta przez dowolny zbiór wektorów w ustalonej przestrzeni V jest jej podprzestrzenią liniową. Jest też oczywiste, że jeśli $X=\mathrm{lin}\{x_1,\ldots,x_k\}$ i $Y=\mathrm{lin}\{x_1,\ldots,x_k,x_{k+1}\}$, to zbiór X jest podprzestrzenią (nie wiadomo, czy właściwą) przestrzeni Y.

Baza i wymiar przestrzeni liniowej

Def. Ustalony układ wektorów $x_1,\ldots,x_k$ jest liniowo niezależny, jeśli z równości $a_1x_1+\cdots+a_kx_k=\mathbf{0}$ wynika $a_1=\cdots=a_k=0$.
Jeśli istnieją skalary $a_1,\ldots,a_k$, nie wszystkie równe 0, takie że $a_1x_1+\cdots+a_kx_k=\mathbf{0}$, to układ $\{x_1,\ldots,x_k\}$ jest liniowo zależny.

Przykład: Układ wektorów $x_1=\begin{bmatrix}1\\-2\end{bmatrix}$ i $x_2=\begin{bmatrix}-2\\4\end{bmatrix}$ w przestrzeni $\mathbb{R}^2$ jest liniowo zależny, bo np. $2x_1+1x_2=\mathbf{0}$.

Układ wektorów $y_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ i $y_2 = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$ jest liniowo niezależny, bo
$a_1 y_1 + a_2 y_2 = \begin{bmatrix} a_1 + a_2 \\ 2a_2 \end{bmatrix}$, i dalej, z równości $a_1 y_1 + a_2 y_2 = 0$ wynika
$a_1 + a_2 = 0$, $2a_2 = 0$, czyli $a_1 = a_2 = 0$.

Uwaga: Słowo „układ” w powyższej definicji oznacza ciąg — w wielu sytuacjach istotna jest kolejność jego elementów, kiedy indziej interesuje nas tylko, czy jest to ciąg różnowartościowy (i uporządkowanie jest nieistotne) lub tylko zbiór elementów tego ciągu. Uporządkowanie jest potrzebne, aby wiązać wektory z odpowiednimi współczynnikami kombinacji liniowych. Jest oczywiste, że jeśli w danym układzie pewien wektor występuje więcej niż raz (tj. ciąg *nie jest* różnowartościowy), to układ ten jest liniowo zależny. Ustalony zbiór wektorów (niezależnie od uporządkowania) określa przestrzeń, którą rozpina.

Stwierdzenie: *Dowolny podzbiór zbioru liniowo niezależnego jest liniowo niezależny (w szczególności zbiór pusty jest liniowo niezależnym zbiorem wektorów).*
Dowolny zbiór, który zawiera zbiór liniowo zależny, jest liniowo zależny.
W szczególności dowolny zbiór, który zawiera wektor zerowy, jest liniowo zależny.

Def. Układ wektorów w przestrzeni V, który spełnia następujące warunki:

1. jest liniowo niezależny,
2. rozszerzenie go o dowolny wektor daje w wyniku układ liniowo zależny,

nazywa się bazą przestrzeni V.

Twierdzenie: *Jeśli układ wektorów $x_1, \ldots, x_n$ jest bazą przestrzeni V, to każdy wektor $y \in V$ można jednoznacznie przedstawić w postaci kombinacji liniowej*

$$y = a_1 x_1 + \cdots + a_n x_n.$$

Dowód: Układ wektorów $x_1, \ldots, x_n, y$ jest liniowo zależny, a zatem istnieją skalary $b_1, \ldots, b_n, c \in \mathbb{K}$, nie wszystkie równe 0, takie że

$$b_1 x_1 + \cdots + b_n x_n + c y = 0.$$

Z liniowej niezależności wektorów $x_1, \ldots, x_n$ wynika, że $c \neq 0$. Mnożąc powyższą równość przez c^{-1} i stosując prawo skreśleń otrzymujemy

$$y = \frac{-b_1}{c} x_1 + \cdots + \frac{-b_n}{c} x_n.$$

Gdyby istniały dwie różne kombinacje liniowe elementów bazy równe y, np.

$$y = a_1 x_1 + \cdots + a_n x_n = a_1' x_1 + \cdots + a_n' x_n,$$

to otrzymalibyśmy równość

$$(a_1 - a_1') x_1 + \cdots + (a_n - a_n') x_n = 0.$$

Warunek $a_i \neq a_i'$ jest dla każdego i sprzeczny z liniową niezależnością bazy. □

Twierdzenie o istnieniu bazy: *Dowolny liniowo niezależny układ wektorów w przestrzeni V można rozszerzyć tak, aby otrzymać bazę tej przestrzeni. W szczególności, każda przestrzeń liniowa ma bazę.*

Dowód tego twierdzenia opiera się na pewniku wyboru, pod postacią lematu Kuratowskiego-Zorna. Pominiemy go, ograniczając się do stwierdzenia, że użycie tego pewnika w dowodzie umożliwia wykazanie istnienia pewnego obiektu (w naszym przypadku bazy), ale nie daje żadnej efektywnej metody znalezienia tego obiektu. Na przykład nie umiemy wskazać żadnej bazy przestrzeni $\mathbb{R}$ nad ciałem $\mathbb{Q}$. Dowody oparte na pewniku wyboru noszą miano niekonstruktywnych. Dla wielu przestrzeni liniowych umiemy jednak wskazać bazy. Umiejętność ta umożliwia sprowadzenie do działań na liczbach (skalarach) wielu rachunków na wektorach (które mogą być bardzo „dziwnymi” obiektami) i ich zbiorach.

Przykłady baz:

- Układ wektorów (macierzy) $e_1, \ldots, e_n$ jest bazą przestrzeni $\mathbb{K}^n$.
- Układ wektorów (ciągów nieskończonych) $e_1, e_2, \ldots$, takich że k-ty wyraz ciągu e_k jest równy 1, a wszystkie pozostałe — 0 jest bazą przestrzeni $\mathbb{K}^\infty$ (której elementami są nieskończone ciągi elementów ciała $\mathbb{K}$, z których co najwyżej skończenie wiele jest różnych od 0).
- Układ wektorów (wielomianów) $1, x, x^2, \ldots, x^n$ jest bazą przestrzeni $\mathbb{R}[x]_n$ wielomianów jednej zmiennej stopnia co najwyżej n (jest to tak zwana baza potęgowa).
- Układ wektorów (równań) $x_1 = 0$, $x_2 = 0$, ..., $x_n = 0$, $0 = 1$ jest bazą przestrzeni równań liniowych z niewiadomymi $x_1, \ldots, x_n$.

Twierdzenie Steinitza o wymianie: *Niech $X = \operatorname{lin}\{x_1, \ldots, x_n\}$, przy czym układ wektorów $x_1, \ldots x_n$ jest liniowo niezależny. Niech $X \subset Y = \operatorname{lin}\{y_1, \ldots, y_m\}$. Wówczas $n \leqslant m$ i n spośród elementów zbioru $\{y_1, \ldots, y_m\}$ można zastąpić przez wektory $x_1, \ldots, x_n$, otrzymując zbiór wektorów, który rozpina tę samą przestrzeń Y.*

Dowód: Twierdzenie jest oczywiście prawdziwe dla $n = 0$. Przyjmiemy założenie indukcyjne, że twierdzenie jest prawdziwe dla $n-1$ (gdzie $n > 0$). Mamy więc $n - 1 \leqslant m$, a przy odpowiednim uporządkowaniu wektorów $y_1, \ldots, y_m$ jest

$$x_n \in Y = \operatorname{lin}\{x_1, \ldots, x_{n-1}, y_n, \ldots, y_m\} \supset \operatorname{lin}\{x_1, \ldots, x_n, y_{n+1}, \ldots, y_m\}.$$

Wektor x_n jest kombinacją liniową wektorów $x_1, \ldots, x_{n-1}, y_n, \ldots, y_m$, ale nie jest on kombinacją liniową wektorów $x_1, \ldots, x_{n-1}$ (bo zbiór $\{x_1, \ldots, x_n\}$ jest liniowo niezależny). Musi więc być $n \leqslant m$, a ponadto istnieją skalary $a_1, \ldots, a_m$, takie że

$$x_n = a_1 x_1 + \cdots + a_{n-1}x_{n-1} + a_n y_n + \cdots + a_m y_m,$$

przy czym $a_k \neq 0$ dla pewnego $k \in \{n, \ldots, m\}$. Przyjmijmy bez straty ogólności, że $a_n \neq 0$. Wtedy

$$y_n = \frac{-a_1}{a_n}x_1 + \cdots + \frac{-a_{n-1}}{a_n}x_{n-1} + \frac{1}{a_n}x_n + \frac{-a_{n+1}}{a_n}y_{n+1} + \cdots + \frac{-a_m}{a_n}y_m.$$

Dla dowolnego wektora $z \in Y$, istnieją skalary $b_1, \ldots, b_m$, takie że

$$z = b_1 x_1 + \cdots + b_{n-1}x_{n-1} + b_n y_n + \cdots + b_m y_m.$$

Podstawiając w miejsce y_n otrzymane wcześniej wyrażenie, otrzymujemy

$$\begin{aligned} z = \Big(b_1 - \frac{b_n a_1}{a_n}\Big)x_1 + \cdots + \Big(b_{n-1} - \frac{b_n a_{n-1}}{a_n}\Big)x_{n-1} + \frac{b_n}{a_n}x_n + \\ \Big(b_{n+1} - \frac{b_n a_{n+1}}{a_n}\Big)y_{n+1} + \cdots + \Big(b_m - \frac{b_n a_m}{a_n}\Big)y_m, \end{aligned}$$

a zatem $z \in \operatorname{lin}\{x_1, \ldots, x_n, y_{n+1}, \ldots, y_m\}$, czyli $Y \subset \operatorname{lin}\{x_1, \ldots, x_n, y_{n+1}, \ldots, y_m\}$. □

Wniosek: *Jeśli n-elementowy układ wektorów jest bazą przestrzeni V, to każda baza tej przestrzeni składa się z n wektorów.*
Dowód: Niech $\{x_1, \ldots, x_n\}$ i $\{y_1, \ldots, y_m\}$ będą dwiema bazami przestrzeni V. Na podstawie twierdzenia Steinitza, $n \leqslant m$, ale także $m \leqslant n$, czyli $n = m$. □

Można udowodnić mocniejsze twierdzenie: wszystkie bazy dowolnej przestrzeni liniowej V są równoliczne (a zatem dotyczy to także przestrzeni, które mają bazy złożone z nieskończenie wielu wektorów).

Def. Wymiarem przestrzeni liniowej V nazywamy liczbę elementów jej bazy (ogólniej — moc zbioru elementów bazy).
Wymiar przestrzeni V oznaczamy symbolem $\dim V$ (z angielskiego *dimension* — wymiar).

Wnioskiem z twierdzenia o istnieniu bazy i twierdzenia o równoliczności wszystkich baz danej przestrzeni jest
Stwierdzenie: *Jeśli przestrzeń X jest podprzestrzenią przestrzeni Y, to $\dim X \leqslant \dim Y$, przy czym jeśli wymiary są skończone i równe, to $X = Y$.*

Suma algebraiczna podprzestrzeni liniowych

Def. Niech Y i Z będą podprzestrzeniami liniowymi przestrzeni liniowej X. Sumą algebraiczną podprzestrzeni Y i Z nazywamy zbiór

$$Y + Z \stackrel{\text{def}}{=} \{\, y + z \colon y \in Y,\ z \in Z \,\}.$$

Przecięciem (albo częścią wspólną) podprzestrzeni Y i Z nazywamy zbiór

$$Y \cap Z = \{\, x \colon x \in Y,\ x \in Z \,\}.$$

Jeśli przecięcie podprzestrzeni Y i Z ma tylko jeden element — wektor zerowy, to sumę algebraiczną tych podprzestrzeni nazywamy sumą prostą i oznaczamy symbolem $Y \oplus Z$.

Łatwo jest udowodnić, że zarówno przecięcie jak i suma algebraiczna (w tym prosta) podprzestrzeni przestrzeni liniowej jest podprzestrzenią liniową. Zachodzą następujące, oczywiste lub łatwe do udowodnienia relacje:

$$\begin{aligned} &0 \leqslant \dim(Y \cap Z) \leqslant \min(\dim Y, \dim Z), \\ &\max(\dim Y, \dim Z) \leqslant \dim(Y + Z) \leqslant \min(\dim X, \dim Y + \dim Z), \\ &\dim Y + \dim Z = \dim(Y + Z) + \dim(Y \cap Z), \end{aligned}$$

a stąd w szczególności $\dim(Y \oplus Z) = \dim Y + \dim Z$ (bo $\dim\{0\} = 0$).

Pojęcia sumy algebraicznej i sumy prostej możemy uogólnić na przypadek większej liczby składników. Sumą algebraiczną s podprzestrzeni $Y_1, \ldots, Y_s$ przestrzeni X jest zbiór

$$\{\, y_1 + \cdots + y_s \colon \forall_j\ y_j \in Y_j \,\} = Y_1 + \cdots + Y_s \stackrel{\text{ozn.}}{=} \sum_{j=1}^{s} Y_j.$$

Jeśli $Y_1, \ldots, Y_s$ są podprzestrzeniami przestrzeni X, takimi że

$$\begin{aligned} &X = \{\, y_1 + \cdots + y_s \colon y_j \in Y_j \,\}, \quad \text{oraz} \\ &Y_i \cap \Big(\sum_{\substack{j \in \{1,\ldots,s\} \\ j \neq i}} Y_j \Big) = \{0\} \quad \text{dla } i = 1, \ldots, s, \end{aligned}$$

to suma algebraiczna tych podprzestrzeni jest sumą prostą. W zapisie symbolicznym

$$X = Y_1 \oplus \cdots \oplus Y_s = \bigoplus_{j=1}^{s} Y_j.$$

Wtedy przedstawienie dowolnego wektora $x \in X$ w postaci sumy $y_1 + \cdots + y_s$ wektorów takich że $\forall_{j \in \{1,\dots,s\}}\ y_j \in Y_j$ jest jednoznaczne.
Istotnie, dla $s = 2$, biorąc $x = y_1 + y_2 = y_1' + y_2'$, otrzymujemy $y_1 - y_1' = y_2' - y_2$, ale $y_1 - y_1' \in Y_1$, $y_2' - y_2 \in Y_2$, a jedynym elementem przecięcia przestrzeni Y_1 i Y_2 jest 0. Zatem, $y_1 = y_1'$, $y_2 = y_2'$. Przez indukcję przenosi się to na przypadek dowolnego s. □

Def. Wektory $y_1 \in Y_1, \dots, y_s \in Y_s$, przypisane jednoznacznie wektorowi $x \in X = \bigoplus_{j=1}^{s} Y_j$, nazywamy składowymi wektora x względem wskazanego rozkładu przestrzeni X na sumę prostą.

Przykład. Przestrzeń $\mathbb{R}[x]_n$ wielomianów rzeczywistych jednej zmiennej stopnia co najwyżej n możemy przedstawić w postaci sumy prostej podprzestrzeni wielomianów parzystych i nieparzystych. Dla n parzystego

$$\mathbb{R}[x]_n = \text{lin}\{1, x, x^2, x^3, x^4, \dots, x^n\} = \text{lin}\{1, x^2, x^4 \dots, x^n\} \oplus \text{lin}\{x, x^3, \dots, x^{n-1}\}.$$

Na przykład wielomian $1 - 2x - 3x^2 + 4x^5 + x^6$ jest sumą $1 - 3x^2 + x^6$ (składowa parzysta) i $-2x + 4x^5$ (składowa nieparzysta).

W szczególności, jeśli układ wektorów $x_1, \dots, x_n$ jest bazą przestrzeni X, to przestrzeń tę możemy przedstawić jako sumę prostą podprzestrzeni jednowymiarowych:

$$X = \text{lin}\{x_1\} \oplus \cdots \oplus \text{lin}\{x_n\},$$

a zatem możemy przedstawić dowolny wektor $z \in X$ w postaci kombinacji liniowej:

$$z = a_1 x_1 + \cdots + a_n x_n.$$

W ten sposób (prawie) dowiedliśmy ponownie, że współczynniki $a_1, \dots, a_n \in \mathbb{K}$ tej kombinacji są określone jednoznacznie.

Przekształcenia liniowe

Def. Niech V_1 i V_2 będą przestrzeniami liniowymi nad ciałem $\mathbb{K}$. Funkcja $f\colon V_1 \to V_2$ jest przekształceniem liniowym, jeśli spełnia warunek

$$\forall_{x_1, x_2 \in V_1} \forall_{a_1, a_2 \in \mathbb{K}}\ f(a_1 x_1 + a_2 x_2) = a_1 f(x_1) + a_2 f(x_2).$$

Łatwy dowód indukcyjny uzasadnia stwierdzenie, że przekształcenie liniowe zachowuje dowolne kombinacje liniowe. Wynikają stąd następujące własności takich przekształceń:

- Obrazem wektora zerowego $0_1 \in V_1$ jest wektor zerowy $0_2 \in V_2$.
- Jeśli układ wektorów $y_1, \dots, y_k \in V_1$ jest liniowo zależny, to układ $f(y_1), \dots, f(y_k) \in V_2$ jest liniowo zależny.
  Jeśli więc układ wektorów $f(y_1), \dots, f(y_k)$ jest liniowo niezależny, to układ $y_1, \dots, y_k$ też jest liniowo niezależny.
- Jeśli określimy sumę przekształceń f_1 i f_2 ustalonej przestrzeni V_1 w V_2 wzorem

  $$\forall_{z \in V_1}\ (f_1 + f_2)(z) = f_1(z) + f_2(z)$$

  oraz iloczyn przekształcenia f i liczby a wzorem

  $$\forall_{z \in V_1}\ (a \cdot f)(z) = a \cdot f(z),$$

  to możemy łatwo przekonać się, że
  zbiór przekształceń liniowych ustalonych przestrzeni liniowych nad ciałem $\mathbb{K}$ z określonymi wyżej działaniami jest przestrzenią liniową nad $\mathbb{K}$.
  Przestrzeń tę oznaczymy symbolem $L(V_1; V_2)$. Jej wektorem zerowym jest przekształcenie zerowe, które każdemu wektorowi $z \in V_1$ przyporządkowuje wektor $0_2 \in V_2$.
- Jeśli funkcje $f\colon V_1 \to V_2$ i $g\colon V_2 \to V_3$ są przekształceniami liniowymi, to przekształcenie złożone $g \circ f\colon V_1 \to V_3$ jest przekształceniem liniowym. Ogólnie, złożenie dowolnego skończonego ciągu przekształceń liniowych jest przekształceniem liniowym.

Def. Izomorfizmem przestrzeni liniowych V_1 i V_2 (nad tym samym ciałem $\mathbb{K}$) nazywamy przekształcenie liniowe $f\colon V_1 \to V_2$, które jest bijekcją.

Izomorfizm f przestrzeni liniowych ma przekształcenie odwrotne, $f^{-1}\colon V_2 \to V_1$, które też jest izomorfizmem. O przestrzeniach V_1 i V_2 mówimy wtedy, że są izomorficzne. Z definicji izomorfizmu wynika, że

- Izomorfizm zachowuje liniową zależność albo liniową niezależność dowolnego układu wektorów. W szczególności izomorfizm przekształca bazę przestrzeni V_1 na bazę przestrzeni V_2, a więc zachowuje wymiar przestrzeni.
- Warunek istnienia izomorfizmu między przestrzeniami określa relację, która jest zwrotna (każda przestrzeń jest izomorficzna ze sobą),
  symetryczna (jeśli V_1 jest izomorficzna z V_2, to V_2 jest izomorficzna z V_1)
  i przechodnia (jeśli V_1 jest izomorficzna z V_2, a V_2 z V_3, to V_1 jest izomorficzna z V_3). Jest to więc relacja równoważności; klasa przestrzeni liniowych nad ustalonym ciałem dzieli się na klasy przestrzeni izomorficznych.

Wspólną cechą przestrzeni izomorficznych jest ich wymiar; dowolna przestrzeń V o wymiarze n nad ciałem $\mathbb{K}$ jest więc izomorficzna z przestrzenią $\mathbb{K}^n$.
Mając ustaloną bazę $x_1, \ldots, x_n$ przestrzeni V możemy określić izomorfizm, który wektorowi $z = c_1x_1 + \cdots + c_nx_n$ przyporządkowuje wektor (macierz kolumnową) $f(z) = [c_1, \ldots, c_n]^T$. Izomorfizm odwrotny jest to przyporządkowanie układowi współczynników $c_1, \ldots, c_n$ wektora z, który jest kombinacją liniową wektorów bazy z tymi współczynnikami. Na dalszych wykładach będzie podany bardziej systematyczny sposób konstruowania tych izomorfizmów.

Wiemy, że jeśli $A \in \mathbb{K}^{m,n}$ oraz $x \in \mathbb{K}^n$, to $Ax \in \mathbb{K}^m$. Na podstawie znanych własności mnożenia macierzy możemy stwierdzić, że przekształcenie przestrzeni $\mathbb{K}^n$ w $\mathbb{K}^m$, polegające na mnożeniu ustalonej macierzy i wektora, jest liniowe. Macierz liczbową $m \times n$ możemy więc utożsamić z przekształceniem $\mathbb{K}^n \to \mathbb{K}^m$.

Ustalmy bazę $x_1, \ldots, x_n$ przestrzeni V_1 i bazę $y_1, \ldots, y_m$ przestrzeni V_2. Obrazem wektora x_j w dowolnym przekształceniu liniowym f jest pewna kombinacja liniowa $a_{1j}y_1 + \cdots + a_{mj}y_m$ wektorów bazy przestrzeni V_2. Przypuśćmy teraz, że $z = c_1x_1 + \cdots + c_nx_n$. Wtedy

$$f(z) = \sum_{j=1}^{n} c_j f(x_j) = \sum_{j=1}^{n} c_j \Big(\sum_{i=1}^{m} a_{ij} y_i\Big) = \sum_{i=1}^{m} \Big(\sum_{j=1}^{n} a_{ij} c_j\Big) y_i.$$

Współczynnikami w bazie $y_1, \ldots, y_m$ obrazu wektora z w przekształceniu f są więc liczby $d_i = \sum_{j=1}^{n} a_{ij}c_j$. Możemy je obliczyć wykonując mnożenie macierzy

$$\begin{bmatrix} a_{11} & \ldots\ldots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \ldots\ldots & a_{mn} \end{bmatrix} \begin{bmatrix} c_1 \\ \vdots \\ \vdots \\ c_n \end{bmatrix} = \begin{bmatrix} d_1 \\ \vdots \\ d_m \end{bmatrix}.$$

Z tego spostrzeżenia wnioskujemy, że mając dowolne przekształcenie liniowe $f\colon V_1 \to V_2$, dla ustalonych baz przestrzeni V_1 o wymiarze n i V_2 o wymiarze m możemy wskazać jednoznacznie określoną macierz $[a_{ij}]_{i,j} \in \mathbb{K}^{m,n}$, która reprezentuje to przekształcenie (tj. pozwala obliczyć współczynniki wektora $f(z)$ na podstawie współczynników wektora z, za pomocą mnożenia macierzy).

Z drugiej strony, dowolnej macierzy $[a_{ij}]_{i,j} \in \mathbb{K}^{m,n}$ odpowiada przekształcenie liniowe, które „przekształca współczynniki w bazach” w opisany sposób. Dlatego przestrzeń przekształceń liniowych przestrzeni n-wymiarowej V_1 w przestrzeń V_2 o wymiarze m jest izomorficzna z przestrzenią macierzy $\mathbb{K}^{m,n}$. Wymiar tej przestrzeni (a więc także przestrzeni $L(V_1; V_2)$) jest równy mn.

Dzięki istnieniu tego izomorfizmu (a właściwie całej klasy izomorfizmów, ponieważ możemy dowolnie wybierać bazy), badanie przekształceń liniowych możemy sprowadzić do badania własności macierzy. Co też uczynimy niebawem.

Zadania i problemy na ćwiczenia i do domu

1. Udowodnij, że jeśli V jest przestrzenią liniową, to $\forall_{x \in V}\ 0 \cdot x = 0$.
2. Wskaż wymiar i dowolną bazę przestrzeni $\mathbb{K}^{m,n}$.
3. Wskaż bazę przestrzeni $\mathbb{C}[x_1, x_2]_2$ (zbioru wielomianów zespolonych dwóch zmiennych stopnia co najwyżej 2) nad ciałem $\mathbb{C}$.
4. Wskaż bazę przestrzeni $\mathbb{C}[x_1, x_2]_2$ nad ciałem $\mathbb{R}$.
5. Znajdź wymiar przestrzeni równań algebraicznych n zmiennych m-tego stopnia nad ciałem $\mathbb{C}$ (na przykład dla $m = n = 2$ to są równania o postaci $ax^2 + bxy + cy^2 + dx + ey + f = 0$, ze współczynnikami zespolonymi $a, \ldots, f$). Wskaż bazę tej przestrzeni.
6. Czy podane niżej zbiory macierzy $n \times n$ są podprzestrzeniami liniowymi przestrzeni $\mathbb{K}^{n,n}$?

   a) Zbiór macierzy permutacji.
   b) Zbiór macierzy trójkątnych.
   c) Zbiór macierzy trójkątnych górnych.
   d) Zbiór macierzy stochastycznych (których suma współczynników w każdym wierszu jest równa 1).
   e) Zbiór macierzy nieosobliwych.
   f) Zbiór macierzy osobliwych.
   g) Zbiór macierzy hermitowskich (zbadaj osobno przypadki gdy $\mathbb{K} = \mathbb{R}$ i $\mathbb{K} = \mathbb{C}$).
   h) Zbiór macierzy antysymetrycznych.
   i) Zbiór macierzy diagonalnie dominujących (takich, że dla każdego $i \in \{1, \ldots, n\}$ zachodzi nierówność $|a_{ii}| > \sum_{j \in \{1,\ldots,n\} \setminus \{i\}} |a_{ij}|$).

   W każdym przypadku podaj uzasadnienie i jeśli odpowiedź jest twierdząca, to wskaż wymiar i bazę.
7. Udowodnij, że zbiór $\mathbb{R}_+$ (zbiór liczb rzeczywistych dodatnich) z działaniem dodawania „$+'$” określonym wzorem $x +' y = xy$ (rozważane dodawanie jest „zwykłym” mnożeniem liczb rzeczywistych) i działaniem mnożenia przez liczbę rzeczywistą a „$\cdot'$” określonym wzorem $a \cdot' x = x^a$, jest rzeczywistą przestrzenią liniową. Znajdź wektor zerowy, wymiar i bazę tej przestrzeni.

   Czy zbiór $\mathbb{Q}_+$ z działaniami określonymi za pomocą tych samych wzorów jest przestrzenią liniową?

8. Zbadaj, czy zbiór $\mathbb{Q}[\sqrt{2}] = \{\, a + b\sqrt{2} \colon a, b \in \mathbb{Q} \,\}$ (z działaniem dodawania wektorów tożsamym z dodawaniem liczb) jest przestrzenią liniową nad ciałem $\mathbb{Q}$. Jeśli tak, to znajdź jakąś bazę i wymiar tej przestrzeni.
9. Które z podanych niżej podzbiorów przestrzeni $\mathbb{R}^3$ są podprzestrzeniami liniowymi?

$$\begin{array}{ll} \{\,(x,y,z) \colon x + y - z = a\,\} & \text{dla ustalonego } a \in \mathbb{R}, \\ \{\,(x,y,z) \colon x^n = 0\,\} & \text{dla ustalonego } n \in \mathbb{N}, \\ \{\,(x,y,z) \colon x^2 + y^2 > s\,\} & \text{dla ustalonego } s \in \mathbb{Q}, \\ \{\,(x,y,z) \colon (x+y)^n = sz^n\,\} & \text{dla ustalonego } n \in \mathbb{N},\ s \in \mathbb{Q}, \\ \{\,(x,y,z) \colon xy = 0\,\}. & \end{array}$$

10. Def. Trójkątną rodziną wielomianów nazywamy układ wielomianów $w_0, \ldots, w_n$, takich że dla każdego k stopień wielomianu w_k jest równy k.

    Udowodnij, że dowolna trójkątna rodzina wielomianów $w_0, \ldots, w_n$ o współczynnikach rzeczywistych jest bazą przestrzeni $\mathbb{R}[x]_n$.
11. Udowodnij podane na stronie 3.6 relacje między wymiarami podprzestrzeni Y, Z,$Y + Z$ i $Y \cap Z$ przestrzeni X.
12. Niech przestrzenie X, Y i Z będą podprzestrzeniami przestrzeni V i niech $\dim X = 3$, $\dim Y = 4$, $\dim Z = 2$ i $\dim V = 7$. Jakie mogą być wymiary przestrzeni $(Y + Z) \cap X$ i $(X \cap Y) + Z$?
13. Udowodnij, że dwie przestrzenie, V_1 i V_2 nad ciałem $\mathbb{K}$ są izomorficzne wtedy i tylko wtedy, gdy $\dim V_1 = \dim V_2$.
14. Zbadaj, czy układ wektorów

$$\begin{bmatrix} 1 \\ \sqrt{2} \\ 0 \end{bmatrix}, \quad \begin{bmatrix} \sqrt{2}/2 \\ 0 \\ 1 \end{bmatrix}, \quad \begin{bmatrix} 0 \\ -1 \\ 1 \end{bmatrix}$$

    jest liniowo zależny
    a) nad ciałem $\mathbb{R}$,
    b) nad ciałem $\mathbb{Q}$.
15. Znajdź macierz przekształcenia liniowego $d \colon \mathbb{R}[x]_3 \to \mathbb{R}[x]_3$, określonego wzorem $d(w) = w'$, odpowiadającą bazie potęgowej $1, x, x^2, x^3$ (w tej bazie reprezentujemy zarówno argument przekształcenia d, jak i wynik).
16. Udowodnij, że jeśli przekształcenia liniowe $f \colon V_1 \to V_2$ i $g \colon V_2 \to V_3$ są reprezentowane w bazach $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n$ przestrzeni V_1, $\boldsymbol{y}_1, \ldots, \boldsymbol{y}_m$ przestrzeni V_2 i $\boldsymbol{z}_1, \ldots, \boldsymbol{z}_k$ przestrzeni V_3 odpowiednio przez macierze A i B, to macierzą złożenia $g \circ f \colon V_1 \to V_3$ tych przekształceń w bazach $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n$ i $\boldsymbol{z}_1, \ldots, \boldsymbol{z}_k$ jest macierz BA.

17. Znajdź macierze przekształceń $d_n \colon \mathbb{R}[x]_3 \to \mathbb{R}[x]_3$, określonych wzorem $d_n(w) = \frac{d^n}{dx^n} w$ dla $n = 1, 2, 3, \ldots$, w bazie potęgowej $1, x, x^2, x^3$. Skorzystaj w tym celu z rozwiązań poprzednich dwóch zadań.

Obraz i jądro przekształcenia liniowego i macierzy

Równoważne układy równań liniowych

Def. Dwa układy równań liniowych o tych samych niewiadomych są równoważne, jeśli każde równanie z pierwszego układu równań jest kombinacją liniową równań drugiego układu i na odwrót.

Twierdzenie: *Zbiory rozwiązań układów równoważnych są identyczne.*
Dowód: Rozważmy dwa układy równań liniowych:

$$a_{i1}x_1 + \cdots + a_{in}x_n = b_i, \quad i = 1, \ldots, m,$$

oraz

$$c_{i1}x_1 + \cdots + c_{in}x_n = d_i, \quad i = 1, \ldots, l.$$

Wystarczy dowieść, że jeśli układy te są równoważne, to każde rozwiązanie pierwszego układu jest także rozwiązaniem drugiego.
Jeśli układy są równoważne, to i-te równanie drugiego układu jest kombinacją liniową równań pierwszego układu o pewnych współczynnikach $s_{i1}, \ldots, s_{im}$.
Zatem, i-te równanie drugiego układu można przedstawić w postaci

$$\sum_{k=1}^{m} s_{ik}(a_{k1}x_1 + \cdots + a_{kn}x_n) = \sum_{k=1}^{m} s_{ik}b_k,$$

przy czym $c_{ij} = \sum_{k=1}^{m} s_{ik}a_{kj}$ oraz $d_i = \sum_{k=1}^{m} s_{ik}b_k$. Podstawiając dowolne rozwiązanie $x_1, \ldots, x_n$ pierwszego układu otrzymujemy równość $\sum_{k=1}^{m} s_{ik}b_k = \sum_{k=1}^{m} s_{ik}b_k$, z której wynika, że jest to także rozwiązanie i-tego (czyli każdego, bo i może być dowolne) równania drugiego układu. □

Uwaga: Układy możemy przedstawić w postaci macierzowej, $A\mathbf{x} = \mathbf{b}$ i $C\mathbf{x} = \mathbf{d}$, gdzie $A = [a_{ij}]_{i,j} \in \mathbb{K}^{m,n}$, $\mathbf{b} \in \mathbb{K}^m$, $C = [c_{ij}]_{i,j} \in \mathbb{K}^{l,n}$ i $\mathbf{d} \in \mathbb{K}^l$. Drugi układ otrzymamy z pierwszego przez pomnożenie jego stron z lewej strony przez macierz $S = [s_{ij}]_{i,j} \in \mathbb{K}^{l,m}$. Mamy więc $C = SA$, $\mathbf{d} = S\mathbf{b}$.
Rozwiązania układu równań liniowych są rozwiązaniami dowolnego układu, który możemy otrzymać mnożąc strony przez macierz, ale taki układ może mieć więcej rozwiązań. Na przykład, mnożąc strony przez macierz zerową otrzymamy układ, którego rozwiązaniem jest każdy wektor $\mathbf{x} \in \mathbb{K}^n$.

Układ równań liniowych $A\mathbf{x} = \mathbf{b}$ możemy przedstawić w postaci

$$\mathbf{a}_1x_1 + \cdots + \mathbf{a}_nx_n = \mathbf{b},$$

gdzie wektory $\mathbf{a}_1, \ldots, \mathbf{a}_n \in \mathbb{K}^m$ są kolumnami macierzy A. Odpowiada to przedstawieniu macierzy A w postaci blokowo-wierszowej. Interpretację rozwiązania $x_1, \ldots, x_n$ jako współczynników kombinacji liniowej kolumn macierzy A, która to kombinacja jest równa wektorowi prawej strony, będziemy eksploatować dalej. Tymczasem możemy zauważyć, że jednorodny układ równań, który otrzymamy biorąc $\mathbf{b} = \mathbf{0}$, ma rozwiązanie różne od $\mathbf{0}$ wtedy i tylko wtedy, gdy wektory $\mathbf{a}_1, \ldots, \mathbf{a}_n$ są liniowo zależne. Jest tak zawsze dla $n > m$.

Rząd macierzy

Def. Rząd kolumnowy macierzy $A \in \mathbb{K}^{m,n}$ jest to wymiar rozpiętej przez kolumny tej macierzy podprzestrzeni przestrzeni $\mathbb{K}^m$.
Rząd wierszowy macierzy A jest to wymiar podprzestrzeni przestrzeni $\mathbb{K}^{1,n}$ rozpiętej przez wiersze macierzy A.
Inaczej mówiąc, rząd kolumnowy macierzy A jest maksymalną liczbą niezależnych liniowo kolumn tej macierzy, a rząd wierszowy macierzy jest maksymalną liczbą jej niezależnych liniowo wierszy.

Twierdzenie: *Rząd kolumnowy dowolnej macierzy jest równy jej rzędowi wierszowemu.*
Dowód: Oznaczmy rząd kolumnowy macierzy $A \in \mathbb{K}^{m,n}$ literą k, a jej rząd wierszowy literą w. Symbolem A_1 oznaczymy macierz $m \times k$, której kolumny są liniowo niezależnymi kolumnami macierzy A. Rząd wierszowy macierzy A_1 oznaczymy w_1; jest jasne, że rząd kolumnowy macierzy A_1 jest równy k, a $w_1 \leqslant w$.
Niech A_2 oznacza macierz $w_1 \times k$, otrzymaną z A_1 przez wybranie w_1 niezależnych liniowo wierszy.
Przypuśćmy, że $k > w_1$. Wtedy układ równań $A_2\mathbf{x} = \mathbf{0}$ ma rozwiązanie $\mathbf{x} \neq \mathbf{0}$ (bo kolumny macierzy A_2 są wektorami w $\mathbb{K}^{w_1}$, więc gdy jest ich więcej niż w_1, to są liniowo zależne). Ale ponieważ wektor $\mathbf{x}$ jest także rozwiązaniem układu $A_1\mathbf{x} = \mathbf{0}$ (oba układy równań są równoważne), więc to oznacza liniową zależność kolumn macierzy A_1. Otrzymana sprzeczność dowodzi, że $k \leqslant w_1$.
Mamy zatem $k \leqslant w$. Aby dowieść, że również $w \leqslant k$ wystarczy rozważyć macierz A^T. Jej rząd kolumnowy jest równy w, a wierszowy — k. □

Def. Rzędem macierzy A nazywamy liczbę, która jest jej rzędem kolumnowym i jednocześnie wierszowym. Będziemy oznaczać go symbolem $\operatorname{rank} A$
Macierz $A \in \mathbb{K}^{m,n}$, której kolumny są liniowo niezależne (tj. $\operatorname{rank} A = n$) nazywamy macierzą kolumnowo regularną.
Macierz $A \in \mathbb{K}^{m,n}$, której wiersze są liniowo niezależne (tj. $\operatorname{rank} A = m$) nazywamy macierzą wierszowo regularną.

Macierze kolumnowo regularne i wierszowo regularne będziemy określać wspólnym mianem macierzy pełnego rzędu, a pozostałe to macierze niepełnego rzędu.

Jest oczywiste, że $\operatorname{rank} A^T = \operatorname{rank} A$. Łatwo jest również pokazać (zostawiam to na ćwiczenia), że dla macierzy $A \in \mathbb{C}^{m,n}$ jest $\operatorname{rank} A^H = \operatorname{rank} A$.

Obraz i jądro macierzy

Def. Obrazem macierzy $A \in \mathbb{K}^{m,n}$ jest przestrzeń rozpięta przez kolumny tej macierzy. Jej wymiar jest równy rzędowi macierzy A. Do oznaczenia tej przestrzeni użyjemy symbolu $\operatorname{im} A$. (ang. *image* — obraz)
Jądrem (albo przestrzenią zerową) macierzy $A \in \mathbb{K}^{m,n}$ nazywamy zbiór wektorów $\{ x \in \mathbb{K}^n \colon Ax = 0 \}$. Zbiór ten jest podprzestrzenią liniową przestrzeni $\mathbb{K}^n$; będziemy go oznaczać symbolem $\ker A$ (ang. *kernel* — jądro).

Mając dowolną macierz A możemy łatwo znaleźć bazę jej obrazu; wystarczy wybrać liniowo niezależne kolumny. Nieco trudniejsze jest znalezienie bazy jądra; w tym celu zauważmy, że wymiar jądra nie zmieni się, jeśli poprzestawiamy kolumny.
Możemy tak poprzestawiać kolumny, aby pierwsze $r = \operatorname{rank} A$ kolumn otrzymanej macierzy $\hat{A}$ było niezależne liniowo. Macierz $\hat{A}$ przedstawimy w postaci blokowej, $[A_1, A_2]$, z blokiem A_1 pełnego rzędu o wymiarach $m \times r$ i blokiem A_2, którego kolumny są kombinacjami liniowymi kolumn bloku A_1. Wobec tego istnieje macierz $B \in \mathbb{K}^{r,n-r}$ (utworzona ze współczynników tych kombinacji), taka że $A_2 = A_1 B$. Niech $\hat{x} \in \mathbb{K}^n$. Wektor ten możemy przedstawić w postaci blokowej, $\hat{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$, dla $x_1 \in \mathbb{K}^r$, $x_2 \in \mathbb{K}^{n-r}$. Wektor $\hat{x}$ jest elementem przestrzeni $\ker \hat{A}$ wtedy, gdy

$$\hat{A}\hat{x} = \begin{bmatrix} A_1 & A_1 B \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = A_1(x_1 + Bx_2) = 0.$$

Macierz A_1 jest kolumnowo regularna, z czego natychmiast wynika, że jej jądro jest podprzestrzenią zerową. Wobec tego musi być $x_1 + Bx_2 = 0$, a zatem $x_1 = -Bx_2$. Mamy stąd

$$\ker \hat{A} = \left\{ \begin{bmatrix} -B \\ I_{n-r} \end{bmatrix} x_2 \colon x_2 \in \mathbb{K}^{n-r} \right\}.$$

Ponieważ wektor $x_2 \in \mathbb{K}^{n-r}$ można wybrać dowolnie, więc bazę przestrzeni $\ker \hat{A}$ otrzymamy biorąc np. wektory $e_1, \ldots, e_{n-r} \in \mathbb{K}^{n-r}$. Wtedy otrzymamy bazę

złożoną z wektorów

$$\hat{x}_1 = \begin{bmatrix} -b_1 \\ e_1 \end{bmatrix}, \ldots, \hat{x}_{n-r} = \begin{bmatrix} -b_{n-r} \\ e_{n-r} \end{bmatrix},$$

gdzie b_i jest i-tą kolumną macierzy B. Bazę przestrzeni $\ker A$ otrzymamy przestawiając odpowiednio współczynniki wektorów $\hat{x}_i$. Wymiar tej przestrzeni, jak się przekonaliśmy, jest równy $n - r = n - \operatorname{rank} A$. Natychmiast wynika stąd

Stwierdzenie: *Dla dowolnej macierzy $A \in \mathbb{K}^{m,n}$ zachodzi równość*

$$n = \dim \ker A + \dim \operatorname{im} A.$$

Twierdzenie: *Dla dowolnej macierzy $A \in \mathbb{C}^{m,n}$ zachodzą równości*

$$\mathbb{C}^m = \operatorname{im} A \oplus \ker A^H, \qquad \mathbb{C}^n = \operatorname{im} A^H \oplus \ker A$$

(dla $A \in \mathbb{R}^{m,n}$ lub $A \in \mathbb{Q}^{m,n}$ sprzężenie hermitowskie macierzy zastępujemy przez transpozycję).
Dowód: Wystarczy udowodnić pierwszą z tych równości. Obie przestrzenie, $\operatorname{im} A$ i $\ker A^H$, są podprzestrzeniami przestrzeni $\mathbb{C}^m$. Jeśli $y \in \operatorname{im} A$, to istnieje wektor $x \in \mathbb{C}^n$, taki że $y = Ax$. Jeśli równocześnie $y \in \ker A^H$, czyli $A^H y = 0$, to $y^H y = x^H A^H y = 0$, ale $y = [y_1, \ldots, y_m]^T$ (dla pewnych liczb zespolonych $y_1, \ldots, y_m$), a zatem $y^H y = \sum_{k=1}^m |y_k|^2 = 0$ wtedy i tylko wtedy gdy $y_1 = \cdots = y_m = 0$, czyli $y = 0$. Tak więc $\operatorname{im} A \cap \ker A^H = \{0\}$, czyli suma algebraiczna tych przestrzeni jest sumą prostą. Wymiar tej sumy jest więc równy

$$\dim(\operatorname{im} A \oplus \ker A^H) = \dim \operatorname{im} A + \dim \ker A^H = \operatorname{rank} A + (m - \operatorname{rank} A^H) = m,$$

co kończy dowód, bo jeśli wymiar podprzestrzeni jest skończony i równy wymiarowi przestrzeni, to ta podprzestrzeń jest niewłaściwa. □

Uwaga: Twierdzenie to dotyczy macierzy o współczynnikach zespolonych, rzeczywistych lub wymiernych, ale nie można go uogólnić na przypadek współczynników należących do dowolnego ciała. Istotnie, w dowodzie skorzystaliśmy z pojęcia wartości bezwzględnej i uporządkowania liczb rzeczywistych przez relację „$\leqslant$”.

Interpretacja iloczynu macierzy

Niech $A = [a_1, \ldots, a_n] \in \mathbb{K}^{m,n}$ i $B = [b_{ij}]_{i,j} \in \mathbb{K}^{n,l}$. Mamy

$$AB = \left[\sum_{k=1}^n a_k b_{k1}, \ldots, \sum_{k=1}^n a_k b_{kl} \right].$$

Kolumny iloczynu są więc kombinacjami liniowymi kolumn macierzy A, a zatem $\operatorname{im}(AB) \subset \operatorname{im} A$, skąd wynika

$$\operatorname{rank}(AB) = \dim \operatorname{im}(AB) \leqslant \operatorname{rank} A.$$

Podobnie, wiersze iloczynu macierzy są kombinacjami wierszy drugiego czynnika (macierzy B); możemy stąd wywnioskować, że $\operatorname{im}(AB)^T \subset \operatorname{im} B^T$, a stąd $\operatorname{rank}(AB) \leqslant \operatorname{rank} B$. Ostatecznie, $\operatorname{rank}(AB) \leqslant \min(\operatorname{rank} A, \operatorname{rank} B)$.

Przykład: Rząd obu macierzy,

$$A = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & -1 & -1 \end{bmatrix} \quad \text{i} \quad B = \begin{bmatrix} 1 & 1 \\ -1 & -1 \\ 1 & -1 \\ -1 & 1 \end{bmatrix},$$

jest równy 2. Ich iloczyn jest macierzą zerową 2×2, a zatem jego rząd jest równy 0. Z drugiej strony, rząd iloczynu kwadratowych macierzy $n \times n$ pełnego rzędu jest równy n.

Dla dowolnego wektora $\boldsymbol{x} \in \mathbb{K}^l$ mamy $(AB)\boldsymbol{x} = A(B\boldsymbol{x})$, jako że mnożenie macierzy jest łączne. Ponieważ $AB\boldsymbol{x} \in \mathbb{K}^m$, więc macierz $C = AB$ opisuje przekształcenie liniowe, które jest złożeniem przekształcenia $\mathbb{K}^l \to \mathbb{K}^n$ reprezentowanego przez macierz B i przekształcenia $\mathbb{K}^n \to \mathbb{K}^m$ reprezentowanego przez A.
Stwierdzenie, że rząd iloczynu macierzy jest nie większy niż rząd któregokolwiek czynnika możemy też uzasadnić geometrycznie, zauważając, że obraz dowolnej podprzestrzeni $X \in \mathbb{K}^n$ w przekształceniu liniowym $\mathbb{K}^n \to \mathbb{K}^m$ jest podprzestrzenią $Y \in \mathbb{K}^m$, której wymiar nie może być większy niż $\dim X$.

Macierze nieosobliwe

Dowolna macierz $A \in \mathbb{K}^{n,n}$ określa przekształcenie przestrzeni $\mathbb{K}^n$ w $\mathbb{K}^n$.

Twierdzenie: *Macierz jest nieosobliwa (czyli istnieje jej odwrotność) wtedy i tylko wtedy, gdy jest pełnego rzędu.*
Dowód. Kolumny $\boldsymbol{a}_1, \ldots, \boldsymbol{a}_n$ macierzy pełnego rzędu są liniowo niezależne i jest ich n, czyli rozpięta przez nie podprzestrzeń $\operatorname{im} A$ przestrzeni $\mathbb{K}^n$ jest niewłaściwa. Kolumny macierzy A stanowią bazę przestrzeni $\mathbb{K}^n$, a więc dowolny wektor $\boldsymbol{y} \in \mathbb{K}^n$ można przedstawić w postaci $\boldsymbol{y} = x_1\boldsymbol{a}_1 + \cdots + x_n\boldsymbol{a}_n$, przy czym układ współczynników $x_1, \ldots, x_n$ spełniających tę równość jest tylko jeden. Istnieje zatem przekształcenie odwrotne do A, które dowolnemu wektorowi $\boldsymbol{y} \in \mathbb{K}^n$ przyporządkowuje wektor $\boldsymbol{x} = [x_1, \ldots, x_n]^T \in \mathbb{K}^n$. Przekształcenie to jest

oczywiście liniowe, a więc istnieje macierz $B \in \mathbb{K}^{n,n}$, która reprezentuje to przekształcenie. Iloczyn BA tych macierzy reprezentuje złożenie obu tych przekształceń; przekształcenie to jest przekształceniem tożsamościowym przestrzeni $\mathbb{K}^n$ i jego macierz jest jednostkowa.
Z drugiej strony, jeśli kolumny macierzy A są liniowo zależne, to reprezentowane przez nią przekształcenie nie jest różnowartościowe. Gdyby istniała odwrotność macierzy A, to reprezentowałaby przekształcenie odwrotne, które nie istnieje. □

Funkcjonały liniowe i przestrzeń sprzężona

Def. Funkcjonałem liniowym w przestrzeni liniowej V nad ciałem $\mathbb{K}$ nazywamy każde przekształcenie liniowe $V \to \mathbb{K}$
(przypominam, że zbiór $\mathbb{K}$ jest przestrzenią liniową nad ciałem $\mathbb{K}$).

Def. Zbiór wszystkich funkcjonałów liniowych w ustalonej przestrzeni V nad ciałem $\mathbb{K}$ (z odpowiednio określonymi działaniami) jest przestrzenią liniową nad $\mathbb{K}$. Przestrzeń tę nazywamy przestrzenią sprzężoną (albo dualną) z V i oznaczamy symbolem V^* (który jest równoważny oznaczeniu $L(V; \mathbb{K})$, ale krótszy).

Def. Funkcjonał zerowy jest to funkcjonał, który każdemu wektorowi $\boldsymbol{x} \in V$ przyporządkowuje 0. Funkcjonał ten jest wektorem zerowym przestrzeni V^*.

Wymiary przestrzeni V i V^* są równe (co wynika z wcześniejszych ustaleń na temat wymiaru przestrzeni $L(V_1; V_2)$). Zauważmy, że zarówno elementy przestrzeni V^* są funkcjonałami liniowymi w przestrzeni V, jak i odwrotnie — dowolny element $\boldsymbol{x}$ przestrzeni V określa pewien funkcjonał liniowy $\boldsymbol{x}$ w przestrzeni V^*, za pomocą wzoru

$$\forall_{\varphi \in V^*}\ \boldsymbol{x}(\varphi) = \varphi(\boldsymbol{x}).$$

Przestrzeń $L(V^*; \mathbb{K}) = V^{**}$ funkcjonałów liniowych na V^* możemy więc utożsamić z przestrzenią V (na jednym z dalszych wykładów wrócimy do tego). Cecha pary przestrzeni V i V^*, która pozwala na „symetryczne” ich traktowanie, nazywa się dualnością.

Def. Niech układ wektorów $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n$ będzie bazą przestrzeni V. Układ wektorów $\varphi_1, \ldots, \varphi_n \in V^*$ nazywamy bazą dualną do $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n$ jeśli spełnione są równania

$$\varphi_i(\boldsymbol{x}_j) = \delta_{ij}$$

(δ_{ij}, przypominam, to symbol Kroneckera, 1 dla $i = j$ i 0 w przeciwnym razie).

Natychmiast zauważamy, że wybór bazy przestrzeni V określa bazę dualną jednoznacznie.
Ponadto, jeśli układ $\varphi_1, \ldots, \varphi_n \in V^*$ jest bazą dualną bazy $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n$ przestrzeni V, to baza $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n$ jest dualna względem bazy $\varphi_1, \ldots, \varphi_n$.

Przykłady

1. Przestrzeń sprzężona z $\mathbb{K}^n$ to $\mathbb{K}^{1,n}$. Baza sprzężona z bazą $\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n$ to układ wektorów (macierzy wierszowych) $\boldsymbol{e}_1^T, \ldots, \boldsymbol{e}_n^T$.
2. Elementy przestrzeni sprzężonej z $\mathbb{R}[x]_n$ (przestrzenią wielomianów rzeczywistych stopnia co najwyżej n; wymiar tej przestrzeni jest równy $n+1$) są funkcjami przyporządkowującymi dowolnemu wielomianowi stopnia co najwyżej n pewną liczbę rzeczywistą.

   Znajdziemy bazę sprzężoną z bazą potęgową, której elementami są wektory (wielomiany) $1, x, x^2, \ldots, x^n$. Łatwo jest przekonać się, że jeśli przyjmiemy funkcjonał φ_0 określony wzorem $\varphi_0(w) = w(0)$, oraz funkcjonały $\varphi_1, \ldots, \varphi_n$, takie że $\varphi_k(w) = \frac{1}{k!}\frac{d^k}{dx^k}w(x)|_{x=0}$, to otrzymamy bazę dualną.

   Niech $x_0, \ldots, x_n$ będą ustalonymi, parami różnymi liczbami. Układ funkcjonałów $\psi_0, \ldots, \psi_n$, określonych wzorami $\psi_i(w) = w(x_i)$ jest bazą przestrzeni sprzężonej z $\mathbb{R}[x]_n$. Łatwo jest znaleźć dualną do niej bazę przestrzeni wielomianów.
   Jej elementami są wielomiany $w_0, \ldots, w_n$, takie że

$$w_j(x) = \prod_{\substack{k=0,\ldots,n\\ k\neq j}} \frac{x - x_k}{x_j - x_k}.$$

   Istotnie, wszystkie te funkcje są wielomianami stopnia n, oraz $w_j(x_i) = \delta_{ij}$.

Obraz i jądro przekształcenia liniowego

Rozszerzymy pojęcie macierzy, dopuszczając, aby jej elementami były wektory i funkcjonały (czyli też wektory, z przestrzeni sprzężonej). W zasadzie jest to uogólnienie wcześniejszych rozważań na temat macierzy blokowych. Dodawanie takich macierzy polega na dodaniu odpowiednich wektorów (albo funkcjonałów). Mnożenie takiej macierzy przez skalar, a także przez macierz, której współczynnikami są skalary, jest określone za pomocą tych samych wzorów co mnożenie macierzy liczbowych (tylko, że symbole mnożenia i dodawania współczynników mają inne znaczenie).

Macierz, której współczynniki są funkcjonałami możemy pomnożyć przez macierz, której współczynniki są wektorami, przy czym „mnożenie” funkcjonału φ i wektora $\boldsymbol{x}$ polega na obliczeniu skalara — wartości wyrażenia $\varphi(\boldsymbol{x})$.

Uwaga: to mnożenie *nie jest* przemienne, tj. iloczyn $\boldsymbol{x}\varphi$, w którym taka kolejność czynników wynika z kolejności mnożonych macierzy, opisuje pewne przekształcenie liniowe; możemy je „pomnożyć” z prawej strony przez wektor $\boldsymbol{y}$ (współczynnik kolejnej macierzy, jeśli taka występuje w rozpatrywanym wyrażeniu) i otrzymać wektor $\boldsymbol{x}\varphi(\boldsymbol{y})$ (iloczyn wektora $\boldsymbol{x}$ i liczby $\varphi(\boldsymbol{y})$), który jest obrazem wektora $\boldsymbol{y}$ w tym przekształceniu liniowym.

Łatwo jest sprawdzić, że wymienione działania na macierzach, których współczynnikami są wektorami i funkcjonałami, mają podobne własności algebraiczne jak dodawanie i mnożenie macierzy liczbowych (w tym łączność, rozdzielność), możemy zatem w podobny sposób przekształcać wzory z takimi macierzami.

Niech V_1 i V_2 będą ustalonymi przestrzeniami liniowymi nad ciałem $\mathbb{K}$, o wymiarach odpowiednio n i m. Przestrzeń V_1 jest więc izomorficzna z $\mathbb{K}^n$, a V_2 jest izomorficzna z $\mathbb{K}^m$. Ustalmy bazę $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n$ przestrzeni V_1 i ustawmy jej elementy w macierz wierszową $X = [\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n]$. Podobnie, z elementów $\varphi_1, \ldots, \varphi_n$ bazy dualnej ustawimy macierz kolumnową $\Phi = \begin{bmatrix} \varphi_1 \\ \vdots \\ \varphi_n \end{bmatrix}$.
Podobnie, biorąc elementy $\boldsymbol{y}_1, \ldots, \boldsymbol{y}_m$ bazy przestrzeni V_2 i elementy jej bazy dualnej $\psi_1, \ldots, \psi_m$, określimy macierze $Y = [\boldsymbol{y}_1, \ldots, \boldsymbol{y}_m]$ i $\Psi = \begin{bmatrix} \psi_1 \\ \vdots \\ \psi_m \end{bmatrix}$.

Możemy teraz zapisać formalnie wiele operacji, o których dotychczas mogliśmy tylko gestykulować. Weźmy dowolny wektor $\boldsymbol{z} \in V_1$. Wyrażenie

$$\Phi\boldsymbol{z} = \begin{bmatrix} \varphi_1(\boldsymbol{z}) \\ \vdots \\ \varphi_n(\boldsymbol{z}) \end{bmatrix}$$

jest wektorem w przestrzeni $\mathbb{K}^n$ (tj. liczbową macierzą kolumnową), której współczynniki odpowiadają przedstawieniu wektora $\boldsymbol{z}$ w bazie $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n$. Istotnie,

$$\varphi_i(c_1\boldsymbol{x}_1 + \cdots + c_n\boldsymbol{x}_n) = c_1\varphi_i(\boldsymbol{x}_1) + \cdots + c_n\varphi_i(\boldsymbol{x}_n) = c_i.$$

Podobnie, mając dowolny wektor $\boldsymbol{c} = [c_1, \ldots, c_n]^T \in \mathbb{K}^n$ możemy wyrazić kombinację liniową wektorów bazy $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n$ o współczynnikach $c_1, \ldots, c_n$ w postaci iloczynu macierzy $X\boldsymbol{c}$.

Macierze X i Φ są swoimi odwrotnościami, w tym sensie, że iloczyn $X\Phi$ reprezentuje przekształcenie tożsamościowe przestrzeni V_1; dla dowolnego wektora $z \in V_1$ wyrażenie $X\Phi z$ jest równe z.
Podobnie, iloczyn ΦX jest przekształceniem tożsamościowym przestrzeni $\mathbb{K}^n$. Dla dowolnego wektora (macierzy kolumnowej) $c \in \mathbb{K}^n$ zachodzi równość $\Phi X c = c$, a zatem możemy stwierdzić, że iloczyn ΦX jest macierzą jednostkową I_n (która reprezentuje tożsamościowe przekształcenie przestrzeni $\mathbb{K}^n$). Możemy więc napisać $\Phi = X^{-1}$, albo $X = \Phi^{-1}$, rozumiejąc to w ten sposób, że jeśli np. macierz X jest przekształceniem liniowym $\mathbb{K}^n \to V_1$, to Φ opisuje przekształcenie odwrotne.

Tak więc mnożenie przez opisane wyżej macierze, które utożsamimy z bazami odpowiednich przestrzeni, określa izomorfizmy przestrzeni V_1 z $\mathbb{K}^n$ i V_2 z $\mathbb{K}^m$. Weźmy teraz dowolną macierz $A \in \mathbb{K}^{m,n}$. Określa ona pewne przekształcenie liniowe przestrzeni $\mathbb{K}^n$ w $\mathbb{K}^m$ (polega ono na pomnożeniu macierzy A i przekształcanego wektora $c \in \mathbb{K}^n$). Wyrażenie $YA\Phi$ opisuje pewne przekształcenie liniowe g przestrzeni V_1 w V_2. Jeśli obliczamy obraz w tym przekształceniu wektora $z \in V_1$, to Φz jest macierzą współczynników wektora z w bazie $x_1, \ldots, x_n$, zaś $A\Phi z$ jest macierzą współczynników wektora $g(z)$ w bazie $y_1, \ldots, y_m$. Wreszcie, $g(z) = YA\Phi z$.

Z wcześniejszych rozważań wynika, że jeśli ustalimy bazy $x_1, \ldots, x_n$ i $y_1, \ldots, y_m$, to każdemu przekształceniu liniowemu $g\colon V_1 \to V_2$ odpowiada jednoznacznie określona macierz $A \in \mathbb{K}^{m,n}$. Wybierając bazy przestrzeni V_1 i V_2 ustaliliśmy więc izomorfizm przestrzeni przekształceń, $L(V_1; V_2)$ i przestrzeni macierzy, $\mathbb{K}^{m,n}$.

Def. Jądrem przekształcenia liniowego $g\colon V_1 \to V_2$ nazywamy zbiór $\{x \in V_1 \colon g(x) = 0\}$.
Jest ono podprzestrzenią liniową przestrzeni V_1, oznaczamy je symbolem $\ker g$.
Obrazem przekształcenia g jest zbiór $\{y \in V_2 \colon \exists_{x \in V_1}\ y = g(x)\}$.
Jest to podprzestrzeń liniowa przestrzeni V_2, którą oznaczamy symbolem $\operatorname{im} g$.

Możemy teraz poszukiwać obrazu i jądra dowolnego przekształcenia liniowego $g\colon V_1 \to V_2$. Jest jasne, że wymiary obrazu i jądra przekształcenia g są takie same jak wymiary obrazu i jądra macierzy A, reprezentującej to przekształcenie w jakichkolwiek bazach X i Y. Mając te bazy i macierz A, możemy wyznaczyć bazę przestrzeni $\operatorname{im} A$ (przez wybranie niezależnych liniowo kolumn, oznaczmy je symbolami $a_{i_1}, \ldots, a_{i_r}$, liczba r jest rzędem macierzy A) i mamy

$$\operatorname{im} g = \operatorname{lin}\{Ya_{i_1}\} \oplus \cdots \oplus \operatorname{lin}\{Ya_{i_r}\}$$

i podobnie, mając bazę przestrzeni $\ker A$ złożoną z wektorów $c_1, \ldots, c_{n-r}$ (sposób ich znalezienia znamy, choć nie do końca; polega on na rozwiązaniu układów

równań liniowych, o czym będzie mowa później), mamy

$$\ker g = \operatorname{lin}\{Xc_1\} \oplus \cdots \oplus \operatorname{lin}\{Xc_{n-r}\}.$$

Niech układ wektorów $x_1, \ldots, x_n$ będzie bazą pewnej przestrzeni V nad $\mathbb{K}$. Niech macierz X reprezentuje tę bazę, a macierz Φ jej bazę dualną. Wtedy dowolne przekształcenie $g\colon V \to V$ możemy przedstawić w postaci $g = XA\Phi$, gdzie A jest pewną macierzą kwadratową. Macierz tę (inaczej: macierz przekształcenia g w bazie $x_1, \ldots, x_n$) możemy przedstawić w postaci $\Phi g X$.

Możemy również przyjąć *dwie* bazy przestrzeni V, reprezentowane odpowiednio przez macierze X i Y. Bazy dualne będą reprezentowane odpowiednio przez macierze Φ i Ψ. Przypuśćmy, że znamy współczynniki dowolnego wektora $z \in V$ w bazie X. Współczynniki tego wektora w bazie Y możemy obliczyć przez pomnożenie wektora współczynników w bazie X przez macierz zmiany bazy B. Wyraża się ona wzorem $B = \Psi X$.

Zadania i problemy

1. Udowodnij, że rząd macierzy nie zmieni się wskutek wykonania żadnej z wymienionych niżej operacji:

- pomnożenie kolumny albo wiersza przez dowolny skalar różny od 0,
- dodanie do kolumny (albo wiersza) innej kolumny (albo wiersza) pomnożonej (pomnożonego) przez dowolny skalar,
- przestawienie kolumn (albo wierszy).

2. Udowodnij, że macierze A i A^H mają ten sam rząd.
3. Niech

$$A = \begin{bmatrix} 1 \\ 2 \\ 3 \\ 4 \end{bmatrix}, \quad B = \begin{bmatrix} 1 & -1 & 1 & -1 \end{bmatrix}, \quad C = \begin{bmatrix} 1 \\ 2 \\ 2 \\ 1 \end{bmatrix},$$

$$D = \begin{bmatrix} 0 & -1 & 2 & -3 \end{bmatrix}, \quad E = \begin{bmatrix} 1 \\ -2 \\ 2 \\ -1 \end{bmatrix}.$$

Oblicz macierz $F = ABCDE$.

4. Zbadaj, jaki jest rząd macierzy

$$\begin{bmatrix} 1 & 2 & 3 & 4 & 5 \\ 0 & -1 & 2 & -3 & 4 \\ 1 & 1 & 5 & 1 & 9 \\ 1 & 0 & 7 & -2 & 13 \end{bmatrix}, \qquad \begin{bmatrix} 0 & 1 & -2 & 3 & -4 \\ -1 & 0 & 2 & -3 & 4 \\ 2 & -2 & 0 & 3 & -4 \\ -3 & 3 & -3 & 0 & 4 \\ 4 & -4 & 4 & -4 & 0 \end{bmatrix}.$$

5. Niech $n \geqslant 4$ i niech wektory $x, y \in \mathbb{R}^n$ będą niezależne liniowo. Znajdź jądro macierzy $A = [x, x+y, x-y, y]$ (macierz jest tu przedstawiona w postaci blokowej, tak aby uwidocznić kolumny).
6. Znajdź jądro i obraz przekształcenia liniowego $f\colon \mathbb{R}[x]_n \to \mathbb{R}[x]_n$ określonego wzorem $f(w) = x^2 w''$.
7. Wyznacz współczynniki macierzy przekształcenia f z poprzedniego zadania, w bazie potęgowej $1, x, x^2, \ldots, x^n$.
8. Niech $n > 1$ i niech przekształcenie $g\colon \mathbb{R}[x]_n \to \mathbb{R}[x]_{n-2}$ będzie określone wzorem $g(w) = w''$, a przekształcenie $h\colon \mathbb{R}[x]_{n-2} \to \mathbb{R}[x]_n$ wzorem $h(w) = x^2 w$. Wyznacz macierze tych przekształceń w bazach potęgowych przestrzeni $\mathbb{R}[x]_n$ i $\mathbb{R}[x]_{n-2}$ i oblicz iloczyn tych macierzy (to jest alternatywny sposób rozwiązania poprzedniego zadania).
9. Udowodnij, że w przestrzeni $\mathbb{R}[x]_n$ każdy funkcjonał liniowy jest kombinacją liniową wartości funkcji w ustalonych punktach.
10. Niech $V = \mathbb{R}[x]_2$. Ponieważ $\dim V^* = \dim V = 3$, więc dowolny funkcjonał liniowy φ w przestrzeni V można przedstawić w postaci $\varphi(f) = a_0 f(x_0) + a_1 f(x_1) + a_2 f(x_2)$, czyli $\varphi = a_0\varphi_0 + a_1\varphi_1 + a_2\varphi_2$, gdzie $\varphi_i(f) = f(x_i)$. Znajdź współczynniki w bazie $\varphi_0, \varphi_1, \varphi_2$ funkcjonału $\varphi(f) = \int_0^1 f(x)\,dx$.
11. Niech U będzie podprzestrzenią liniową przestrzeni V, przy czym $n = \dim U < m = \dim V$. Dowolny funkcjonał liniowy $\varphi \in V^*$ jest funkcjonałem liniowym określonym w podprzestrzeni U, czyli $\varphi \in U^*$, a zatem $V^* \subset U^*$. Ale wtedy $m = \dim V^* \leqslant n = \dim U^*$, a więc mamy sprzeczność. Wyjaśnij, gdzie tu jest błąd i na czym on polega.

Ćwiczenia potwórkowe

1. Niech $z = (1, \sqrt{3})$. Oblicz z^{-12}.
2. Zbadaj, czy zbiór macierzy zespolonych 2×2, o postaci

$$\begin{bmatrix} a & -b \\ \overline{b} & \overline{a} \end{bmatrix}$$

z działaniem mnożenia macierzy jest grupą. Jeśli tak, to czy jest to grupa abelowa?

3. Oblicz macierz

$$\begin{bmatrix} 1 & 2 & 0 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{bmatrix}^{12}.$$

4. Oblicz macierz $B = A^3 - 3A^2 + 3A - I_4$, gdzie

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

5. Znajdź zbiór rozwiązań układu równań liniowych

$$\begin{bmatrix} (1,2) & (2,-1) & (3,4) & (5,3) \\ (0,1) & (1,0) & (0,5) & (1,5) \\ (1,0) & (0,-1) & (4,3) & (4,2) \end{bmatrix} \begin{bmatrix} z_1 \\ z_2 \\ z_3 \\ z_4 \end{bmatrix} = \begin{bmatrix} (13,11) \\ (5,13) \\ (14,2) \end{bmatrix}.$$

6. Znajdź rząd oraz bazę jądra i obrazu macierzy

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

7. Wskaż wymiar i bazę podprzestrzeni przestrzeni $\mathbb{R}[x]_4$, składającej się z wielomianów w spełniających warunki $w(-1) = w(1)$ i $w'(0) = 0$.
8. Oblicz współczynniki funkcjonału φ w przestrzeni $\mathbb{R}[x]_2$, danego wzorem

$$\varphi(w) = \int_0^1 w(x)\,dx,$$

w bazie dualnej do bazy w_0, w_1, w_2, gdzie $w_0(x) = x(x-a)$, $w_1(x) = x(x+a)$, $w_2(x) = x^2 - a^2$, $a = \frac{1}{\sqrt{3}}$.
Zaproponuj inny (niż z użyciem podanego wyżej wzoru) sposób obliczania wartości $\varphi(w)$ dla ustalonego wielomianu w.

Układy równań liniowych

Układ m równań liniowych z n niewiadomymi możemy zapisać w czterech równoważnych postaciach. Pierwsza z nich jest „pełna”:

$$\begin{cases} a_{11}x_1 + \ a_{12}x_2 + \cdots + \ a_{1n}x_n = b_1, \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n = b_m. \end{cases}$$

Druga to postać macierzowa,

$$Ax = b,$$

z <u>macierzą współczynników</u> $A = [a_{ij}]_{i,j} \in \mathbb{K}^{m,n}$, <u>wektorem niewiadomym</u> $x \in \mathbb{K}^n$ i <u>wektorem prawej strony</u> $b \in \mathbb{K}^m$. Rozwiązanie układu w tej postaci interpretujemy jak znalezienie wektora (lub zbioru wektorów) x, którego obrazem w przekształceniu A jest wektor b.
Trzecia postać polega na przedstawieniu wektora b jako kombinacji liniowej kolumn macierzy $A = [a_1, \ldots, a_n]$ ze współczynnikami $x_1, \ldots, x_n$:

$$a_1x_1 + \cdots + a_nx_n = b.$$

Rozwiązanie układu polega na znalezieniu tych współczynników.
Wreszcie, czwarta postać wygląda następująco:

$$\begin{cases} \varphi_1(x) = b_1, \\ \cdots\cdots\cdots \\ \varphi_m(x) = b_m. \end{cases}$$

Funkcjonał liniowy $\varphi_i \in (\mathbb{K}^n)^*$ dla $i \in \{1, \ldots, m\}$ jest tu określony wzorem $\varphi_i([x_1, \ldots, x_n]^T) = a_{i1}x_1 + \cdots + a_{in}x_n$.
Każde równanie liniowe jest kombinacją liniową równań $x_1 = 0, \ldots, x_n = 0$ i $0 = 1$. Pierwsze n z nich możemy zapisać w postaci $e_1^H x = 0, \ldots, e_n^H x = 0$ (macierze $e_1^H, \ldots, e_n^H$ reprezentują elementy bazy dualnej do $e_1, \ldots, e_n$), a ostatnie z nich w postaci $0x = 1$ (macierz zerowa, $0 \in \mathbb{K}^{1,n}$ reprezentuje <u>funkcjonał zerowy</u>, czyli wektor zerowy przestrzeni sprzężonej). Tak więc w i-tym równaniu występuje funkcjonał $\varphi_i = a_{i1}e_1^H + \cdots + a_{in}e_n^H$. Rozwiązanie układu polega na znalezieniu takiego wektora x, dla którego funkcjonały te przyjmują podane wartości.

Na podstawie poznanych dotychczas faktów możemy podać warunek istnienia rozwiązań dowolnego układu równań liniowych oraz zbadać, czym jest zbiór rozwiązań. Macierz $[A, b] = [a_1, \ldots, a_n, b]$, otrzymaną przez dołączenie wektora b jako dodatkowej kolumny do macierzy współczynników nazwiemy <u>macierzą uzupełnioną</u>.

<u>Twierdzenie Kroneckera-Capellego:</u> *Układ m równań liniowych z n niewiadomymi, $Ax = b$, ma rozwiązanie wtedy i tylko wtedy, gdy rząd macierzy A jest równy rzędowi macierzy uzupełnionej $[A, b]$.*
Ponadto, jeśli wektor x jest rozwiązaniem układu $Ax = b$, a $z \in \ker A$, to wektor $x + z$ też jest rozwiązaniem. Każde rozwiązanie można przedstawić w postaci sumy dowolnego innego rozwiązania i pewnego wektora należącego do jądra macierzy A.

<u>Dowód:</u> Lewa strona układu (przedstawiona w drugiej z wymienionych wyżej postaci), po podstawieniu dowolnych liczb $x_1, \ldots, x_n$, reprezentuje pewną kombinację liniową kolumn macierzy A. Wektor prawej strony, równy takiej kombinacji, musi zatem być elementem przestrzeni rozpiętej przez kolumny macierzy A, tj. $\operatorname{im} A$.
Jeśli rząd macierzy uzupełnionej jest większy niż rząd macierzy A, to $b \notin \operatorname{im} A$, a zatem wektor b nie jest kombinacją liniową kolumn macierzy A.
Spośród kolumn macierzy A można wybrać bazę przestrzeni $\operatorname{im} A$.
Wektor $b \in \operatorname{im} A$ ma w tej bazie jednoznacznie określone współczynniki.
Rozwiązanie układu otrzymujemy wstawiając do tego ciągu współczynników zera w miejscach odpowiadających kolumnom nie należącym do bazy.
Jesli wektor x jest rozwiązaniem, a $z \in \ker A$, to $A(x + z) = Ax + Az = b + 0 = b$.
Przypuśćmy, że wektory x i x' są rozwiązaniami.
Wtedy $0 = b - b = Ax' - Ax = A(x' - x)$, a zatem $x' - x \in \ker A$. $\square$

Znając pewne rozwiązanie x układu $Ax = b$, możemy przedstawić zbiór wszystkich rozwiązań w postaci sumy algebraicznej zbioru jednoelementowego i podprzestrzeni:

$$\{x\} + \ker A = \{\, y\colon y = x + z,\ Az = 0\}.$$

O wektorze x (i każdym innym elemencie tego zbioru) mówimy, że jest to <u>rozwiązanie szczególne</u> układu równań.

<u>Rozwiązanie ogólne</u> układu jest to wyrażenie o postaci

$$x + \sum_{k=1}^{n-r} c_k y_k,$$

w którym wektor x jest dowolnym rozwiązaniem szczególnym układu, a układ wektorów $y_1, \ldots, y_{n-r}$ jest bazą przestrzeni $\ker A$. Wartość tego wyrażenia dla dowolnych liczb $c_1, \ldots, c_{n-r}$ jest rozwiązaniem szczególnym, każdemu rozwiązaniu szczególnemu odpowiada inny układ liczb $c_1, \ldots, c_{n-r}$ i nie ma żadnych innych rozwiązań.

Metoda rozwiązywania układów równań liniowych

W wyznaczaniu zbioru rozwiązań można wyróżnić dwa cele: sprawdzenie, czy układ jest niesprzeczny oraz znalezienie dowolnego rozwiązania szczególnego i przestrzeni $\ker A$. Ogólna metoda rozwiązywania układu równań polega na takim jego przekształcaniu, aby otrzymać układ *równoważny* (z dokładnością do uporządkowania niewiadomych), łatwiejszy do rozwiązania. Celem jest otrzymanie układu dostatecznie prostego, aby łatwo było wskazać zbiór jego rozwiązań (a w szczególności przekonać się, czy jest on niepusty).

Przedstawiona niżej metoda rozwiązywania układów równań liniowych znana jest pod nazwą eliminacji Gaussa, choć była używana znacznie wcześniej niż Gauss mógł się nią zajmować (ponoć jedna z chińskich ksiąg z czasów dynastii Han zawiera opis rozwiązywania układów dwóch i trzech równań).

Układ możemy reprezentować za pomocą macierzy A i wektora $\mathbf{b}$, albo za pomocą macierzy uzupełnionej, $[A, \mathbf{b}]$, o wymiarach $m \times n + 1$. Przekształcając układ będziemy wykonywać następujące czynności:

- zmianę kolejności równań układu (odpowiada to przestawianiu wierszy macierzy $[A, \mathbf{b}]$),
- zmianę kolejności niewiadomych (odpowiada to przestawianiu kolumn macierzy A; kolejno wykonane przestawienia trzeba zapamiętać, aby po obliczeniu niewiadomych odpowiednio je uporządkować),
- dodanie do wiersza macierzy $[A, \mathbf{b}]$ innego wiersza pomnożonego przez dowolną liczbę.

Przekształcenie układu wymaga wykonania $r = \operatorname{rank} A \leqslant \min(m, n)$ kroków. Przed wykonaniem k-tego kroku mamy macierz $A^{(k-1)}$ i wektor prawej strony $\mathbf{b}^{(k-1)}$, które przedstawiamy w postaci

$$\left[\begin{array}{c|c|ccc} R^{(k-1)} & \multicolumn{4}{c}{T^{(k-1)}} \\ \hline & a_{kk}^{(k-1)} & a_{k,k+1}^{(k-1)} & \ldots & a_{kn}^{(k-1)} \\ \cline{2-5} & a_{k+1,k}^{(k-1)} & a_{k+1,k+1}^{(k-1)} & \ldots & a_{k+1,n}^{(k-1)} \\ 0 & \vdots & \vdots & & \vdots \\ & a_{mk}^{(k-1)} & a_{m,k+1}^{(k-1)} & \ldots & a_{mn}^{(k-1)} \end{array}\right], \quad \left[\begin{array}{c} \mathbf{b}_1^{(k-1)} \\ \hline b_k^{(k-1)} \\ \hline b_{k+1}^{(k-1)} \\ \vdots \\ b_m^{(k-1)} \end{array}\right].$$

Blok $R^{(k-1)}$ o wymiarach $k-1 \times k-1$ jest macierzą trójkątną górną, a pod nim znajduje się (zaznaczony symbolem 0) blok złożony z samych zer. Przed pierwszym krokiem ($k = 1$) mamy macierz $A^{(0)} = A$ z pustym blokiem $R^{(0)}$ i wektor $\mathbf{b}^{(0)} = \mathbf{b}$ (na początku którego mamy pusty blok $\mathbf{b}_1^{(0)}$).

Jeśli wszystkie współczynniki macierzy $A^{(k-1)}$ w wierszach $k, \ldots, m$ są równe 0, to $r = k - 1$ i praca polegająca na przekształcaniu układu została zakończona.
W przeciwnym razie, jeśli $a_{kk}^{(k-1)} = a_{k+1,k}^{(k-1)} = \cdots = a_{mk}^{(k-1)} = 0$, to musimy k-tą kolumnę macierzy $A^{(k)}$ zamienić miejscami z inną kolumną, która w k-tym wierszu lub niżej ma różny od zera współczynnik (a zatem zmieniamy kolejność niewiadomych).

Jeśli po ewentualnym przestawieniu kolumn współczynnik na diagonali w k-tej kolumnie jest równy 0, to wyszukujemy w tej kolumnie poniżej diagonali współczynnik różny od zera. Jeśli znajdziemy go w wierszu j-tym, to przestawiamy wiersze k i j i zamieniamy odpowiednie współrzędne wektora prawej strony.

W ten sposób zapewniliśmy, że $a_{kk}^{(k-1)} \neq 0$ (uwaga: powinienem ten współczynnik oznaczyć inaczej, ale nadmierna liczba różnych symboli może być groźniejsza niż drobna nieścisłość). Możemy więc wykonać następujące przekształcenie:
dla $i = k+1, \ldots, m$ obliczamy $l_{ik} = a_{ik}^{(k-1)}/a_{kk}^{(k-1)}$ i przypisujemy $a_{ij}^{(k)} := a_{ij}^{(k-1)} - l_{ik}a_{kj}^{(k-1)}$ dla $j = 1, \ldots, n$ (w szczególności wynika stąd $a_{ij}^{(k)} = 0$ dla $j \leqslant k$, nie musimy tego obliczać) oraz $b_i^{(k)} := b_i^{(k-1)} - l_{ik}b_k^{(k-1)}$.
W ten sposób w k-tej kolumnie pod diagonalą pojawiają się zera, a przekształcona macierz $A^{(k)}$ i wektor prawej strony $\mathbf{b}^{(k)}$ reprezentują równoważny układ równań liniowych.

Po wykonaniu r takich kroków otrzymamy macierz uzupełnioną $[A^{(r)}, \mathbf{b}^{(r)}]$, która reprezentuje układ równań liniowych równoważny układowi wyjściowemu.
Macierz ta ma następującą postać blokową:

$$\left[\begin{array}{ccc} R^{(r)} & T^{(r)} & \mathbf{b}_1^{(r)} \\ 0 & 0 & \mathbf{b}_2^{(r)} \end{array}\right].$$

Blok $R^{(r)}$ o wymiarach $r \times r$ jest macierzą trójkątną górną. Ponieważ wszystkie współczynniki na diagonali tego bloku są różne od 0, więc jest to macierz nieosobliwa.

Blok $T^{(r)}$ ma wymiary $n - r \times r$ i w ogólności może mieć dowolne współczynniki. Poniżej bloków $R^{(r)}$ i $T^{(r)}$ mamy bloki z wszystkimi współczynnikami równymi 0. Bloki $\mathbf{b}_1^{(r)} \in \mathbb{K}^r$ i $\mathbf{b}_2^{(r)} \in \mathbb{K}^{m-r}$ powstają z podziału przekształconego wektora prawej strony. Otrzymany układ równań można przedstawić w następującej postaci:

$$\left\{\begin{array}{r} R^{(r)}\mathbf{x}_1 + T^{(r)}\mathbf{x}_2 = \mathbf{b}_1^{(r)}, \\ 0\mathbf{x}_1 + \quad 0\mathbf{x}_2 = \mathbf{b}_2^{(r)}. \end{array}\right.$$

Przeprowadzimy dyskusję możliwych rozwiązań układu o tej szczególnej postaci.

Układ jest sprzeczny jeśli $\mathbf{b}_2^{(r)} \neq 0$. Przyjmijmy zatem, że $\mathbf{b}_2^{(r)} = 0$. Jeśli blok $T^{(r)}$ ma 0 kolumn (czyli $r = n$), to macierz współczynników układu jest kolumnowo regularna i układ jest określony, tj. ma jednoznaczne rozwiązanie. W przeciwnym razie możemy przyjąć dowolny wektor $x_2 \in \mathbb{K}^{n-r}$, a następnie obliczyć wektor x_1, który jest jedynym rozwiązaniem układu $R^{(r)}x_1 = \mathbf{b} - T^{(r)}x_2$. Rozwiązaniem całego układu jest wektor $\hat{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$. Jeśli nie przestawialiśmy kolumn macierzy A, to jest to również rozwiązanie wyjściowego układu równań. W przeciwnym razie trzeba odpowiednio poprzestawiać obliczone niewiadome, tj. współczynniki wektora $\hat{x}$.

Rozwiązanie układu równań z nieosobliwą macierzą trójkątną $R^{(r)}$ jest łatwe. Możemy niezbyt ściśle powiedzieć, że niewiadoma x_i nie występuje w danym równaniu, jeśli jest w nim pomnożona przez współczynnik równy 0. Rozpatrując równania „od końca” zauważamy, że w każdym z nich występuje tylko jedna niewiadoma nieobecna w poprzednich równaniach. Sposób jej obliczenia, oparty na tym spostrzeżeniu, jest oczywisty.

Rozkład trójkątny macierzy

Opisana wyżej metoda rozwiązywania układu m równań liniowych z n niewiadomymi, $Ax = \mathbf{b}$, składa się z dwóch etapów. Etap pierwszy polega na przekształceniu układu w taki sposób, aby otrzymać układ równoważny z macierzą trójkątną. Dokonamy nowej interpretacji tego procesu.

Przypuśćmy, że nie zachodzi konieczność przestawiania równań (wierszy macierzy uzupełnionej $[A, \mathbf{b}]$), ani niewiadomych (które odpowiada przestawianiu kolumn macierzy A). W kroku k dla $i = k+1, \ldots, m$ obliczamy liczbę l_{ik} i do i-tego wiersza macierzy układu dodajemy wiersz k-ty pomnożony przez $-l_{ik}$. Działanie to jest równoważne obliczeniu macierzy $[A^{(k)}, \mathbf{b}^{(k)}] = L_k^{-1}[A^{(k-1)}, \mathbf{b}^{(k-1)}]$, gdzie

$$L_k = \begin{bmatrix} 1 & & & & & \\ & \ddots & & & & \\ & & 1 & & & \\ & & l_{k+1,k} & 1 & & \\ & & \vdots & & \ddots & \\ & & l_{m,k} & & & 1 \end{bmatrix}, \quad L_k^{-1} = \begin{bmatrix} 1 & & & & & \\ & \ddots & & & & \\ & & 1 & & & \\ & & -l_{k+1,k} & 1 & & \\ & & \vdots & & \ddots & \\ & & -l_{m,k} & & & 1 \end{bmatrix}$$

(macierze te mają wymiary $n \times n$; nie zaznaczone współczynniki są równe 0). Macierze te możemy przedstawić w następującej postaci: $L_k = I_m + \mathbf{l}_k \boldsymbol{e}_k^T$ oraz $L_k^{-1} = I_m - \mathbf{l}_k \boldsymbol{e}_k^T$, gdzie $\mathbf{l}_k = [0, \ldots, 0, l_{k+1,k}, \ldots, l_{m,k}]^T$.

Po wykonaniu $r = \operatorname{rank} A$ kroków otrzymujemy macierz trójkątną górną $R = A^{(r)}$. Możemy napisać

$$R = L_r^{-1} \ldots L_1^{-1} A = L^{-1}A, \quad \text{czyli} \quad A = L_1 \ldots L_r R = LR.$$

Macierz L jest iloczynem macierzy trójkątnych dolnych, a zatem jest również macierzą trójkątną dolną. Co więcej, możemy się przekonać, że dla $i > k \leqslant r$ współczynnik macierzy L w i-tym wierszu i k-tej kolumnie jest równy l_{ik}. Współczynniki na diagonali macierzy L są równe 1, a pozostałe są równe 0.

Procedura przekształcania układu doprowadziła do znalezienia macierzy trójkątnej dolnej $L \in \mathbb{K}^{m,m}$ i trójkątnej górnej $R \in \mathbb{K}^{m,n}$, takich że $A = LR$. Przedstawienie macierzy A w postaci iloczynu tych macierzy nazywa się rozkładem trójkątno-trójkątnym (w skrócie — rozkładem trójkątnym). Zbadamy, czego rozkład otrzymamy, jeśli w trakcie przekształcania układu przestawiamy wiersze i kolumny.

Jeśli przed wykonaniem k-tego kroku eliminacji przestawiamy kolumny o indeksach k i $q(k)$ oraz wiersze k i $p(k)$, to jest to równoważne pomnożeniu przekształcanej macierzy z prawej strony przez macierz transpozycji $T_{k,q(k)} \in \mathbb{K}^{n,n}$ i z lewej strony przez $T_{k,p(t)} \in \mathbb{K}^{m,m}$. Zatem, po wykonaniu całego przekształcenia otrzymamy macierz

$$R = L_r^{-1} T_{r,p(r)} \ldots L_1^{-1} T_{1,p(1)} A T_{1,q(1)} \ldots T_{r,q(r)} = UAQ^{-1}.$$

Macierz $Q^{-1} = T_{1,q(1)} \ldots T_{r,q(r)}$ jest macierzą permutacji; ponieważ odwrotnością dowolnej transpozycji jest ta sama transpozycja, a ponadto macierze transpozycji są symetryczne, więc $Q = T_{r,q(r)} \ldots T_{1,q(1)}$ oraz $Q^{-1} = Q^T$.

Przekonamy się, że macierz U możemy przedstawić w postaci

$$U = L_r^{-1} T_{r,p(r)} \ldots L_1^{-1} T_{1,p(1)} = \hat{L}_r^{-1} \ldots \hat{L}_1^{-1} T_{r,p(r)} \ldots T_{1,p(1)} = \hat{L}^{-1} P.$$

Niech $l > k$; weźmy macierz transpozycji T_{jl} takiej że $k < j \leqslant l$. Przekształcimy wyrażenie

$$T_{jl} L_k^{-1} T_{jl} = T_{jl}(I_m - \mathbf{l}_k \boldsymbol{e}_k^T) T_{jl} = I_m - T_{jl} \mathbf{l}_k \boldsymbol{e}_k^T T_{jl}.$$

Mamy $\boldsymbol{e}_k^T T_{jl} = \boldsymbol{e}_k^T$ (transpozycja T_{jl} przestawia w $\boldsymbol{e}_k$ współczynniki na pozycjach większych od k, czyli zera), a ponadto istnieje wektor $\tilde{\mathbf{l}}_k = [0, \ldots, 0, \tilde{l}_{k+1,k}, \ldots, \tilde{l}_{m,k}]^T = T_{jl}[0, \ldots, 0, l_{k+1,k}, \ldots, l_{m,k}]^T$ (oba wektory, $\mathbf{l}_k$ i $\tilde{\mathbf{l}}_k$, mają k początkowych współczynników równych 0). Stąd wynika istnienie wektora $\hat{\mathbf{l}}_k$, takiego że

$$T_{r,p(r)} \ldots T_{k+1,p(k+1)} (I_m - \mathbf{l}_k \boldsymbol{e}_k^T) = (I_m - \hat{\mathbf{l}}_k \boldsymbol{e}_k^T) T_{r,p(r)} \ldots T_{k+1,p(k+1)}.$$

Macierz $\widehat{L}_k^{-1} = I_m - \hat{l}_k e_k^T$ jest więc macierzą trójkątną dolną, która poza diagonalą i k-tą kolumną ma wszystkie współczynniki równe 0. Zatem, macierz $L^{-1} = \widehat{L}_r^{-1} \dots \widehat{L}_1^{-1}$ jest trójkątna dolna. Mamy więc

$$R = L^{-1}PAQ^{-1}, \quad \text{czyli} \quad PAQ^{-1} = LR.$$

Macierze P i Q są macierzami permutacji. W wyniku eliminacji otrzymaliśmy więc czynniki L i R rozkładu trójkątnego macierzy powstałej z A przez przestawienie wierszy i kolumn.

Twierdzenie: *Dla dowolnej macierzy $A \in \mathbb{K}^{m,n}$ rzędu r istnieją macierze permutacji $P \in \mathbb{K}^{m,m}$ i $Q \in \mathbb{K}^{n,n}$, oraz macierz trójkątna dolna $L \in \mathbb{K}^{m,m}$ i macierz trójkątna górna $R \in \mathbb{K}^{m,n}$, takie że*

$$PAQ^{-1} = LR$$

przy czym współczynniki na diagonali macierzy L są równe 1 i w kolumnach $r+1, \dots, n$ tej macierzy wszystkie inne współczynniki są równe 0.
Dla ustalonych macierzy P i Q rozkład ten (o ile istnieje) jest jednoznaczny.

Dowód: Istnienie opisanych w twierdzeniu macierzy wynika z wykonalności procedury eliminacji i z rachunków przeprowadzonych przed chwilą. Pozostało zatem udowodnienie jednoznaczności rozkładu, jeśli macierze P i Q są ustalone. Oznaczmy $\widehat{A} = PAQ^{-1}$. Macierz tę przedstawimy w postaci blokowej,

$$\widehat{A} = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix},$$

z blokiem A_{11} o wymiarach $r \times r$ ($r = \operatorname{rank} A$), A_{12} $r \times n - r$, A_{21} $m - r \times r$ i A_{22} $m - r \times n - r$. Jeśli $r = m$ lub $r = n$, to odpowiednie bloki będą puste. Blok A_{11} jest macierzą nieosobliwą.
Macierze L i R również przedstawimy w postaci

$$L = \begin{bmatrix} L_{11} & 0 \\ L_{21} & L_{22} \end{bmatrix}, R = \begin{bmatrix} R_{11} & R_{12} \\ 0 & 0 \end{bmatrix},$$

z blokami L_{11} i R_{11} nieosobliwymi o wymiarach $r \times r$. Mamy wtedy $A_{11} = L_{11}R_{11}$. Gdyby istniały również inne macierze $r \times r$, takie że $A_{11} = \widehat{L}_{11}\widehat{R}_{11}$, to mielibyśmy równości

$$L_{11}R_{11} = \widehat{L}_{11}\widehat{R}_{11}, \quad \text{czyli} \quad \widehat{L}_{11}^{-1}L_{11} = \widehat{R}_{11}R_{11}^{-1},$$

przy czym macierz $\widehat{L}_{11}^{-1}L_{11}$ jest trójkątna dolna, macierz $\widehat{R}_{11}R_{11}^{-1}$ jest trójkątna górna, a więc obie te macierze są diagonalne. Ponieważ wszystkie współczynniki

na diagonali macierzy $\widehat{L}_{11}^{-1}$ i L_{11} są równe 1, więc iloczyn tych macierzy jest macierzą jednostkową $r \times r$. Stąd

$$\widehat{L}_{11}^{-1}L_{11} = I_r = \widehat{R}_{11}R_{11}^{-1}, \quad \text{a zatem} \quad \widehat{L}_{11} = L_{11}, \quad \widehat{R}_{11} = R_{11}.$$

Pozostałe bloki macierzy L i R znajdujemy tak: $A_{21} = L_{21}R_{11}$, czyli $L_{21} = A_{21}R_{11}^{-1}$. Z założeń twierdzenia wynika, że $L_{22} = I_{m-r}$. Musi też być $R_{12} = L_{11}^{-1}A_{21}$.
Tak więc wszystkie bloki macierzy L i R są określone jednoznacznie. □

Uwaga: Jeśli macierz A nie jest kwadratowa nieosobliwa, to macierz P lub Q w rozkładzie $PAQ^{-1} = LR$ może nie być określona jednoznacznie. Istotnie, jeśli po zakończeniu eliminacji zostaną wiersze lub kolumny z wszystkimi współczynnikami równymi 0, to możemy je przestawiać dowolnie.
Jeśli $r < m$, to blok L_{22} macierzy L jest niepusty. Możemy zauważyć, że wybierając dowolne współczynniki w tym bloku zawsze będziemy mieli równość $\widehat{A} = LR$.

Metoda eliminacji Gaussa jest nie tylko algebraicznym algorytmem rozwiązywania układów równań liniowych, ale (dla układów z macierzami kwadratowymi nieosobliwymi) także algorytmem numerycznym. Znalezienie trójkątnych czynników rozkładu macierzy $n \times n$ wymaga wykonania ok. $\frac{1}{3}n^3$ działań (mnożeń i odejmowań albo dzieleń). Jeśli wiele współczynników macierzy to 0, to koszt eliminacji może być istotnie mniejszy. W algorytmie numerycznym wiersze (i ewentualnie kolumny) przestawia się nie tylko po to, aby unikać dzielenia przez zero, ale także, aby zmniejszać skutki błędów zaokrągleń — przestawiamy tak, aby kolejne dzielniki miały jak największe wartości bezwzględne.

Konstrukcja obrazu i jądra macierzy

Niech macierze $L \in \mathbb{K}^{m,m}$ i $R \in \mathbb{K}^{m,n}$ będą czynnikami rozkładu trójkątnego macierzy $\widehat{A} = PAQ^{-1}$. Znając te czynniki oraz permutacje P i Q możemy łatwo znaleźć bazę obrazu i jądra macierzy A.

Działanie przekształcenia opisanego przez macierz Q^{-1} polega na takim przestawieniu kolumn macierzy A, że pierwsze $r = \operatorname{rank} A$ kolumn macierzy AQ^{-1} jest liniowo niezależnych. Kolumny te stanowią bazę przestrzeni macierzy $\operatorname{im} A$.

Ponieważ macierz L jest nieosobliwa, więc jądro macierzy $\widehat{A} = LR$ jest jądrem macierzy R. Macierz ta ma ostatnie $m - r$ wierszy z samymi zerami, a więc jej jądro jest również jądrem macierzy $[R_{11}, R_{12}]$ otrzymanej po odrzuceniu tych wierszy (bloki R_{11} i R_{12} są te same, co wcześniej). Jądro przestrzeni $\ker R$ składa się z wektorów $\hat{x} = \begin{bmatrix} y \\ z \end{bmatrix}$, takich że

$$R_{11}y + R_{12}z = 0, \quad \text{czyli} \quad R_{11}y = -R_{12}z.$$

Macierz R_{11} jest nieosobliwa, a zatem dla każdego wektora z powyższy układ równań ma jednoznaczne rozwiązanie. Podstawiając zamiast z wektor $e_i \in \mathbb{K}^{n-r}$ otrzymujemy po prawej stronie wektor równy i-tej kolumnie macierzy R_{12} pomnożonej przez -1. Rozwiązując takie układy dla $i = 1, \ldots, n-r$ otrzymamy $n-r$ liniowo niezależnych wektorów $\hat{x}_1, \ldots, \hat{x}_{n-r}$, które stanowią bazę przestrzeni $\ker R$. Mając tę bazę, możemy wskazać bazę przestrzeni $\ker A$; składa się ona z wektorów $x_1 = Q^{-1}\hat{x}_1, \ldots, x_{n-r} = Q^{-1}\hat{x}_{n-r}$. Istotnie,

$$Ax_i = P^{-1}LRQx_i = P^{-1}LRQQ^{-1}\hat{x}_i = P^{-1}LR\hat{x}_i = 0.$$

Zadania i problemy

1. Udowodnij, że

- układ równań jest niesprzeczny wtedy i tylko wtedy, gdy równanie $0x = 1$ nie jest kombinacją liniową równań układu,
- dwa niesprzeczne układy równań liniowych są równoważne wtedy i tylko wtedy, gdy mają ten sam zbiór rozwiązań (na wykładzie był dowód wynikania tylko w jedną stronę).

<u>Wskazówka:</u> Wybierz z każdego układu maksymalny podukład, który składa się z równań liniowo niezależnych, i udowodnij, że te podukłady są równoważne. Spróbuj użyć w dowodzie pojęcia przestrzeni sprzężonej do $\mathbb{K}^n$.

2. Znajdź rozwiązanie ogólne układu równań $Ax = b$, gdzie

$$A = \begin{bmatrix} 0 & 1 & 2 & 3 \\ 1 & 2 & 3 & 4 \\ 1 & 1 & 1 & 1 \end{bmatrix}, \qquad b = \begin{bmatrix} -3 \\ -3 \\ 0 \end{bmatrix}.$$

Sprawdź wynik.

3. Znajdź rozwiązanie ogólne układu równań $Ax = b$, gdzie

$$A = \begin{bmatrix} (0,0) & (1,0) & (2,1) \\ (1,0) & (0,1) & (-1,2) \\ (1,0) & (0,2) & (-2,4) \end{bmatrix}, \qquad b = \begin{bmatrix} (1,1) \\ (0,1) \\ (-1,2) \end{bmatrix}.$$

Sprawdź wynik.

4. Udowodnij, że jeśli macierz $A \in \mathbb{K}^{m,n}$ jest kolumnowo regularna, to podczas rozwiązywania układu równań liniowych $Ax = b$ nie trzeba przestawiać kolumn, a co najwyżej tylko wiersze.
5. Znajdź rozwiązanie ogólne układu równań liniowych $Ax = b$, gdzie

$$A = \begin{bmatrix} 1 & 1 & 1 \\ 1 & a & 1 \\ 1 & 1 & a^2 \end{bmatrix}, \qquad b = \begin{bmatrix} 1 \\ -a \\ a^2 \end{bmatrix},$$

w zależności od parametru $a \in \mathbb{R}$. Sprawdź wynik.

6. Udowodnij, że jeśli macierz $A \in \mathbb{K}^{n,n}$ jest nieosobliwa, to następujące postępowanie:

- Przyjmij macierz $M_0 = [A, I_n] \in \mathbb{K}^{n,2n}$,
- Niech dla $k > 0$ macierz M_k powstaje przez wykonanie jednej z następujących operacji: pomnożenie wiersza przez stałą, przestawienie wierszy lub dodanie do wiersza innego wiersza pomnożonego przez dowolną stałą (z takich operacji składa się eliminacja Gaussa bez przestawiania kolumn, ale dopuszczamy dowolne takie działania w dowolnej kolejności),

spowoduje, że jeśli w macierzy M_k w miejscu bloku A pojawi się macierz jednostkowa $n \times n$, to w miejscu bloku I_n pojawi się macierz A^{-1}.

7. Znajdź czynniki trójkątne macierzy

$$A = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 6 & 9 \\ 3 & 10 & 54 \end{bmatrix}.$$

Sprawdź wynik.

8. Dokonaj rozkładu macierzy

$$A = \begin{bmatrix} 0 & 4 & 4 \\ 1 & 2 & 1 \\ 2 & 6 & 3 \end{bmatrix}$$

na czynniki trójkątne oraz odpowiednie macierze permutacji. Sprawdź wynik.

9. Oblicz koszt rozwiązywania układu równań liniowych $Rx = b$ z nieosobliwą macierzą trójkątną górną R.

Podprzestrzenie, warstwy i przestrzeń ilorazowa

Zasada abstrakcji

Def. Niech $\sim$ oznacza relację dwuargumentową w pewnym zbiorze X. Relację tę nazywamy równoważnością, jeśli ma ona następujące własności:

- zwrotność, tj. $\forall_{x \in X}\, x \sim x$,
- symetrię, tj. $\forall_{x,y \in X}\, x \sim y \Rightarrow y \sim x$,
- przechodniość, tj. $\forall_{x,y,z \in X}\, x \sim y$ oraz $y \sim z \Rightarrow x \sim z$.

Twierdzenie (zasada abstrakcji): *Relacja równoważności dzieli zbiór X na rozłączne zbiory X_j, takie że jeśli $x \in X_j$ to $y \in X_j \Leftrightarrow x \sim y$.*
Dowód jest prostym ćwiczeniem, więc go na wykładzie pominiemy.

Def. Niech X oznacza zbiór, w którym jest określona relacja równoważności $\sim$.
Klasą abstrakcji elementu $x \in X$ nazywamy zbiór X_j wszystkich elementów y zbioru X, takich że $x \sim y$.
Dowolny element x klasy abstrakcji X_j nazywamy reprezentantem tej klasy.

Przykłady:

- W zbiorze $\mathbb{Z}$ dla dowolnego $n > 0$ określamy następującą relację $\sim$: $x \sim y \Leftrightarrow (x - y) \bmod n = 0$. Jest to relacja równoważności, która dzieli zbiór $\mathbb{Z}$ na n klas abstrakcji (jakich?).
- W zbiorze wielomianów zespolonych stopnia co najwyżej n określamy relację $\sim$ warunkiem: $w_1 \sim w_2 \Leftrightarrow (w_1(x) = 0 \Leftrightarrow w_2(x) = 0)$.

Wiele różnych konstrukcji w matematyce (a zwłaszcza w algebrze) polega na rozpatrywaniu klas abstrakcji jako indywidualnych obiektów. Przypomnijmy na przykład, że aby formalnie zdefiniować liczby całkowite, określa się w zbiorze par liczb naturalnych następującą relację „$\sim$”: $(a, b) \sim (c, d) \Leftrightarrow a + d = b + c$ (w tej definicji występuje tylko dodawanie liczb naturalnych). Liczby całkowite to klasy abstrakcji tej relacji. Podanie innych przykładów to problem do zastanowienia.

Warstwy przestrzeni liniowej

Niech X będzie ustaloną podprzestrzenią liniową przestrzeni liniowej V nad ciałem $\mathbb{K}$. Wprowadzimy następującą relację w przestrzeni V:

$$\boldsymbol{x} \sim \boldsymbol{y} \Leftrightarrow \boldsymbol{x} - \boldsymbol{y} \in X.$$

Możemy sprawdzić, że jest to relacja równoważności. Istotnie, $\boldsymbol{x} - \boldsymbol{x} = \boldsymbol{0} \in X$, a więc relacja jest zwrotna. Jeśli $\boldsymbol{x} - \boldsymbol{y} = \boldsymbol{z} \in X$, to $\boldsymbol{y} - \boldsymbol{x} = -\boldsymbol{z} \in X$, czyli relacja $\sim$ jest symetryczna. Wreszcie, jeśli $\boldsymbol{x} - \boldsymbol{y} \in X$ oraz $\boldsymbol{y} - \boldsymbol{z} \in X$, to $\boldsymbol{x} - \boldsymbol{z} = (\boldsymbol{x} - \boldsymbol{y}) + (\boldsymbol{y} - \boldsymbol{z}) \in X$, a zatem jest to również relacja przechodnia.

Def. Klasy abstrakcji relacji $\sim$ określonej wyżej nazwiemy warstwami przestrzeni V równoległymi do podprzestrzeni X. Klasę abstrakcji wektora $\boldsymbol{x} \in V$ oznaczymy symbolem $[\boldsymbol{x}]$ (trudno, będziemy starać się nie mylić tego symbolu z zapisem macierzy). W szczególności warstwa zerowa, $[\boldsymbol{0}]$, jest zbiorem elementów podprzestrzeni X.

Jeśli podprzestrzeń X jest podprzestrzenią zerową, to wszystkie warstwy do niej równoległe są zbiorami jednoelementowymi. Podobnie łatwo jest sprawdzić, że jeśli podprzestrzeń X jest niewłaściwa, tj. $X = V$, to jest tylko jedna warstwa — zbiór wszystkich elementów przestrzeni V. Między tymi przypadkami skrajnymi może być wiele innych możliwości.

Wszystkie warstwy przestrzeni V równoległe do podprzestrzeni X są równoliczne. Aby tego dowieść, wystarczy pokazać, że każda warstwa jest równoliczna z podprzestrzenią X. Jest to łatwe — jeśli ustalimy dowolny wektor $\boldsymbol{x} \in [\boldsymbol{x}]$, to możemy określić przekształcenie $f\colon X \to [\boldsymbol{x}]$ wzorem $\forall_{\boldsymbol{z} \in X}\, f(\boldsymbol{z}) = \boldsymbol{x} + \boldsymbol{z}$ i sprawdzić, że jest to bijekcja.

Mając wektor $\boldsymbol{x} \in V$ oraz bazę $\boldsymbol{y}_1, \ldots, \boldsymbol{y}_l$ podprzestrzeni X, możemy warstwę $[\boldsymbol{x}]$ przedstawić w następującej postaci:

$$[\boldsymbol{x}] = \{\, \boldsymbol{y} \colon \boldsymbol{y} = \boldsymbol{x} + \sum_{k=1}^{l} a_k \boldsymbol{y}_k,\ a_1, \ldots, a_l \in \mathbb{K} \,\}.$$

Możemy też napisać $[\boldsymbol{x}] = \boldsymbol{x} + X$ (albo ściślej $[\boldsymbol{x}] = \{\boldsymbol{x}\} + X$), rozumiejąc, że symbol „$+$” oznacza tu sumę algebraiczną.
Wymiar warstwy równoległej do podprzestrzeni X jest to liczba $\dim X$.

Możemy zauważyć następujący fakt: w takiej samej postaci jak warstwę powyżej przedstawiliśmy rozwiązanie ogólne dowolnego niesprzecznego układu równań liniowych. Mamy stąd wniosek, że zbiór rozwiązań niesprzecznego układu m równań liniowych z n niewiadomymi, $A\boldsymbol{x} = \boldsymbol{b}$, jest warstwą równoległą do podprzestrzeni $\ker A \subset \mathbb{K}^n$.

Jest też prawdziwe stwierdzenie odwrotne: każda warstwa przestrzeni $\mathbb{K}^n$ jest zbiorem rozwiązań pewnego układu równań liniowych. Teraz to udowodnimy.

Def. Hiperpłaszczyzną w przestrzeni liniowej V o wymiarze n nazywamy jej dowolną podprzestrzeń liniową o wymiarze $n-1$.

Lemat 1: *Każda hiperpłaszczyzna w przestrzeni $\mathbb{K}^n$ jest zbiorem rozwiązań pewnego jednorodnego równania liniowego $\varphi(x)=0$.*
Dowód: Niech $z_1,\ldots,z_{n-1}$ będzie bazą danej hiperpłaszczyzny. Można tę bazę rozszerzyć o 1 wektor, otrzymując bazę $z_1,\ldots,z_n$ przestrzeni $\mathbb{K}^n$. Funkcjonał, o którym mowa w stwierdzeniu, to element φ_n bazy $\varphi_1,\ldots,\varphi_n$ dualnej do bazy $z_1,\ldots,z_n$ (funkcjonał ten zależy od z_n, ale to nie szkodzi). $\square$

Lemat 2: *Dowolna podprzestrzeń k-wymiarowa X przestrzeni n wymiarowej V jest przecięciem $n-k$ lub większej liczby hiperpłaszczyzn.*
Dowód: Wystarczy dowieść tezy dla przestrzeni $\mathbb{K}^n$. Podobnie jak w lemacie 1, bazę $z_1,\ldots,z_k$ podprzestrzeni X możemy rozszerzyć do bazy $z_1,\ldots,z_n$ przestrzeni $\mathbb{K}^n$, a następnie znaleźć bazę dualną $\varphi_1,\ldots,\varphi_n$.
Zbiorem rozwiązań równania $\varphi_i(x)=0$ jest hiperpłaszczyzna, którą oznaczymy symbolem π_i. Zbiór rozwiązań układu $\varphi_{k+1}(x)=0,\ldots,\varphi_n(x)=0$ jest przecięciem hiperpłaszczyzn $\pi_{k+1},\ldots,\pi_n$. Zbiorem tym jest podprzestrzeń X o wymiarze k, którą w ten sposób przedstawiliśmy w postaci przecięcia $n-k$ hiperpłaszczyzn.
Pokażemy, że wymiar przestrzeni, która jest przecięciem $n-k$ hiperpłaszczyzn, jest nie mniejszy niż k. Przestrzeń ta jest zbiorem rozwiązań pewnego układu $n-k$ równań (każde równanie określa jedną z hiperpłaszczyzn). Rząd macierzy współczynników tego układu jest nie większy niż $n-k$, a zatem wymiar jej jądra jest nie mniejszy niż $n-(n-k)=k$. $\square$

Twierdzenie: *Dowolna warstwa k-wymiarowa przestrzeni $\mathbb{K}^n$ jest zbiorem rozwiązań pewnego układu równań liniowych, który składa się co najmniej z $n-k$ równań.*
Dowód: Ustalmy dowolną warstwę $[x]\subset\mathbb{K}^n$ równoległą do podprzestrzeni k-wymiarowej X. Na podstawie lematu 2 możemy skonstruować macierz A o wymiarach $n-k\times n$, której jądrem jest podprzestrzeń X (i nie umiemy skonstruować macierzy o mniejszej liczbie wierszy i tym samym jądrze).
Jeśli wektor y jest reprezentantem warstwy $[x]$ (można przyjąć $y=x$), to układ równań liniowych, którego zbiorem rozwiązań jest warstwa $[x]$ otrzymamy biorąc macierz współczynników A i wektor prawej strony $b=Ay$. $\square$

Przestrzenie ilorazowe

Niech X oznacza ustaloną podprzestrzeń liniową przestrzeni V nad $\mathbb{K}$. Zbiór elementów przestrzeni V jest sumą warstw równoległych do X. Rozważmy zbiór

tych warstw. Możemy w nim określić dodawanie warstw, wzorem

$$[x]+[y]\stackrel{\text{def}}{=}[x+y],$$

oraz mnożenie warstwy przez skalar $a\in\mathbb{K}$, wzorem

$$a[x]\stackrel{\text{def}}{=}[ax].$$

Musimy jednak przekonać się, że powyższe działania są dobrze określone, co oznacza, że wynik każdego z tych działań jest warstwą równoległą do X i nie zależy od wyboru reprezentantów. To pierwsze jest oczywiste: dla dowolnego wektora $z\in V$ możemy skonstruować warstwę równoległą do X, której z jest reprezentantem, a suma wektorów $x+y$ i iloczyn ax są wektorami.
Dowód, że wyniki działań nie zależą od wyboru reprezentantów jest też prosty. Niech $x'\in[x]$, tj. $x-x'=z\in X$, oraz $y'\in[y]$, czyli $y-y'=w\in X$.
Jeśli $a\in[x+y]$, czyli $(x+y)-a=b\in X$, to

$$(x'+y')-a=(x-z)+(y-w)-a=b-z-w\in X,$$

a zatem $a\in[x'+y']$, co dowodzi, że $[x'+y']=[x+y]$.
Dowód, że jeśli $x'\in[x]$ to $a[x']=a[x]$ pozostawiam na ćwiczenia.

Zbiór warstw przestrzeni V równoległych do podprzestrzeni X oznaczymy symbolem V/X. Zbiór ten z działaniem dodawania warstw jest grupą abelową, której elementem neutralnym jest warstwa $[0]$. Mnożenie warstwy przez skalar jest rozdzielne względem dodawania warstw, a także dodawania skalarów:

$$\forall_{a\in\mathbb{K},\,[x],[y]\in V/X}\ a([x]+[y])=a[x]+a[y],$$
$$\forall_{a,b\in\mathbb{K},\,[x]\in V/X}\ (a+b)[x]=a[x]+b[x].$$

Wreszcie, prawdziwe są warunki

$$\forall_{a,b\in\mathbb{K},\,[x]\in V/X}\ (ab)[x]=a(b[x]),\qquad \forall_{[x]\in V/X}\ 1[x]=[x],$$

a zatem zbiór V/X z działaniem dodawania warstw i mnożenia warstw przez skalar jest przestrzenią liniową nad ciałem $\mathbb{K}$. Przestrzeń tę nazywamy przestrzenią ilorazową.

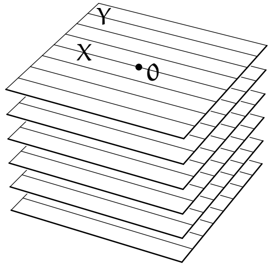
Podstawowe pytania, jakie należy zadać mając do czynienia z dowolną przestrzenią liniową, to jaki jest jej wymiar i jak znaleźć jej bazę (jeśli można to zrobić).

Aby odpowiedzieć na te pytania, załóżmy, że $\dim X=k\leqslant\dim V=n$ i weźmy bazę $x_1,\ldots,x_k$ podprzestrzeni X. Dołączając do niej pewne wektory $x_{k+1},\ldots,x_n$

możemy otrzymać bazę przestrzeni V. Udowodnimy, że układ warstw $[x_{k+1}], \dots, [x_n]$ jest bazą przestrzeni ilorazowej V/X, której wymiar jest w związku z tym równy $n-k$.
Przypuśćmy, że $a_{k+1}[x_{k+1}] + \dots + a_n[x_n] = [0]$, a zatem istnieje wektor $z \in X$, taki że $a_{k+1}x_{k+1} + \dots + a_n x_n + z = 0$. Wektor z jest kombinacją liniową wektorów $x_1, \dots, x_k$, a zatem wyrażenie po lewej stronie powyższej równości można przedstawić w postaci $a_1x_1 + \dots + a_n x_n$. Ponieważ układ wektorów $x_1, \dots, x_n$ jest bazą, więc wszystkie współczynniki tej kombinacji liniowej, w tym $a_{k+1}, \dots, a_n$, są równe 0.
Układ warstw $[x_{k+1}], \dots, [x_n]$ jest więc liniowo niezależny. Pozostaje dowieść, że jest to maksymalny układ liniowo niezależny, tj. dołączenie do niego dowolnej warstwy $[y]$ spowoduje powstanie układu liniowo zależnego. Weźmy dowolny wektor $y \in V$, $y \notin X$; istnieją liczby $b_1, \dots, b_n \in \mathbb{K}$, takie że $y = b_1x_1 + \dots + b_nx_n$ i co najmniej jedna z liczb $b_{k+1}, \dots, b_n$ jest różna od 0. Rozważmy równanie $a_{k+1}[x_{k+1}] + \dots + a_n[x_n] - [y] = [0]$. Jeśli jest ono spełnione przez pewne liczby $a_{k+1}, \dots, a_n$, to istnieje wektor $z \in X$, taki że $a_{k+1}x_{k+1} + \dots + a_nx_n - y + z = 0$. Ale równanie to możemy przedstawić w postaci $(a_1 - b_1)x_1 + \dots + (a_n - b_n)x_n = 0$ i spełniają je współczynniki $a_i = b_i$ dla $i = 1, \dots, n$. Co najmniej jedna z liczb $a_{k+1}, \dots, a_n$ jest różna od 0, tak więc układ warstw $[x_{k+1}], \dots, [x_n], [y]$ jest liniowo zależny. □

Rozważmy podprzestrzenie liniowe X i Y przestrzeni V, takie że $X \subset Y \subset V$. Możemy określić przestrzenie ilorazowe V/X i Y/X. Przestrzeń Y/X jest zbiorem warstw przestrzeni V, równoległych do X i zawartych w Y. Jest to więc podzbiór przestrzeni V/X, a ponieważ jest to przestrzeń liniowa, jest ona więc podprzestrzenią liniową przestrzeni V/X. Możemy zatem określić przestrzeń ilorazową $(V/X)/(Y/X)$. Zbadamy jej związek z przestrzenią V/Y.
Przede wszystkim, łatwo jest sprawdzić, że $\dim V/Y = \dim(V/X)/(Y/X)$, a więc przestrzenie te są izomorficzne. Elementami przestrzeni V/Y są warstwy przestrzeni V równoległe do Y. Elementami przestrzeni $(V/X)/(Y/X)$ są warstwy przestrzeni V/X, równoległe do podprzestrzeni Y/X, która jest zbiorem warstw przestrzeni Y równoległych do X. Ponieważ przestrzeń Y jest sumą (mnogościową) swoich warstw równoległych do X, więc każdy element $[x] = x + Y$ (tu suma algebraiczna) przestrzeni V/Y jest również sumą (mnogościową) pewnych warstw przestrzeni V równoległych do X (czyli elementów przestrzeni V/X).

Przykład. Rozważmy przestrzeń trójwymiarową V, jej podprzestrzeń dwuwymiarową Y (łatwo ją sobie wyobrazić jako płaszczyznę zawierającą punkt przyjęty za wektor 0) i jednowymiarową podprzestrzeń X zawartą w Y (wyobrażamy ją sobie jako prostą w płaszczyźnie Y, przechodzącą przez 0).



Przestrzeń Y/X jest zbiorem prostych równoległych do X, które leżą w płaszczyźnie Y. Przestrzeń V/X jest zbiorem wszystkich prostych w przestrzeni V równoległych do X. Przestrzeń V/Y jest zbiorem płaszczyzn równoległych do Y. Każdą taką płaszczyznę można przedstawić w postaci sumy położonych w niej prostych równoległych do X. Tak więc elementy przestrzeni V/Y są płaszczyznami (które możemy interpretować jako zbiory punktów), a elementy przestrzeni $(V/X)/(Y/X)$ to te same płaszczyzny, które interpretujemy jako zbiory prostych równoległych do X.

Układy współrzędnych w przestrzeni liniowej

Def. Układem współrzędnych w zbiorze X nazywamy dowolną funkcję różnowartościową $f\colon X \to \mathbb{K}^n$. Przy ustalonym f przestrzeń $\mathbb{K}^n$ nazywa się przestrzenią współrzędnych zbioru X.

Znając współrzędne dowolnego elementu $a \in X$ w zasadzie możemy znaleźć ten element; definicja nie wyklucza jednak możliwości, że pewien wektor $v \in \mathbb{K}^n$ nie jest wartością funkcji f dla żadnego $a \in X$, ani nie podpowiada sposobu znajdowania $a = f^{-1}(v)$. W przypadku n-wymiarowej przestrzeni liniowej V nad ciałem $\mathbb{K}$ szczególnie wygodne jest posługiwanie się układem współrzędnych, który jest izomorfizmem przestrzeni V i $\mathbb{K}^n$.

Dowolny izomorfizm przestrzeni V i $\mathbb{K}^n$ możemy skonstruować biorąc pewną bazę $x_1, \dots, x_n$ przestrzeni X i przyjmując, że jeśli $z = a_1x_1 + \dots + a_nx_n$, to $f(z) = [a_1, \dots, a_n]^T$. Liczby $a_1, \dots, a_n$ nazywamy współrzędnymi wektora x w bazie $x_1, \dots, x_n$ i możemy je otrzymać obliczając wyrażenie $\begin{bmatrix} \varphi_1(z) \\ \vdots \\ \varphi_n(z) \end{bmatrix}$, które zapisujemy krócej Φz. Macierz Φ reprezentuje bazę sprzężoną z bazą $x_1, \dots, x_n$, której elementy ustawimy w macierz X.

Podobnie, możemy ustalić inną (albo i tę samą) bazę $y_1, \dots, y_n$ i jej bazę sprzężoną $\psi_1, \dots, \psi_n$ i przedstawić je w postaci macierzy Y i Ψ. Wtedy macierz $C = (c_{ij})_{i,j} = \Phi Y$, o współczynnikach $c_{ij} = \varphi_i(y_j)$, jest macierzą zmiany bazy (to już było, tu jest tylko powtórzenie). Macierz ta opisuje zmianę układu współrzędnych, tj. umożliwia obliczenie współrzędnych dowolnego wektora z w układzie, którego układem odniesienia jest baza $x_1, \dots, x_n$, na podstawie współrzędnych tego wektora w bazie $y_1, \dots, y_n$.

Macierze zmiany bazy mają następujące własności:

- Są jednoznacznie określone przez odpowiednie bazy.
- Jeśli A jest macierzą przejścia od bazy $x_1, \dots, x_n$ do $y_1, \dots, y_n$, a B jest macierzą przejścia od $y_1, \dots, y_n$ do $z_1, \dots, z_n$, to macierz przejścia od $x_1, \dots, x_n$ do $z_1, \dots, z_n$ jest równa BA.
- Są nieosobliwe; mając dane dowolne bazy $x_1, \dots, x_n$ i $y_1, \dots, y_n$ możemy znaleźć macierz A przejścia od pierwszej z nich do drugiej, a także od drugiej do pierwszej, czyli A^{-1}.
- Macierzą przejścia od bazy $x_1, \dots x_n$ do $x_1, \dots x_n$ jest macierz jednostkowa I_n.
- Dowolna macierz nieosobliwa jest macierzą zmiany bazy. Mianowicie, macierz A jest macierzą przejścia od bazy $x_1, \dots x_n$ do bazy, której elementami są współczynniki macierzy $Y = [x_1, \dots, x_n]A^{-1}$.

  Udowodnimy to. Oznaczmy współczynniki macierzy Y symbolami $y_1, \dots, y_n$. Są to wektory liniowo niezależne, bo $YA = X = [x_1, \dots, x_n]$, a więc możemy otrzymać bazę wyjściową wybierając odpowiednie kombinacje liniowe wektorów $y_1, \dots, y_n$, zatem wektory te rozpinają przestrzeń V. Dowolnemu wektorowi $a \in \mathbb{K}^n$ odpowiada wektor $x = Xa \in V$. Podstawiając $X = YA$ mamy $x = YAa = Yb$, gdzie $b = Aa \in \mathbb{K}^n$ jest wektorem współrzędnych wektora x w bazie $y_1, \dots, y_n$.

Dowolny funkcjonał liniowy ξ możemy reprezentować za pomocą pewnego wektora $g \in \mathbb{K}^n$ w ten sposób, że jeśli $a \in \mathbb{K}^n$ jest wektorem współrzędnych wektora $z \in V$ w bazie $x_1, \dots, x_n$, to $\xi(z) = g^H a$.

Uwaga: Jeśli $\mathbb{K} = \mathbb{C}$, to we wzorach powyżej i poniżej występuje sprzężenie hermitowskie, a w innych przypadkach zwykła transpozycja macierzy. Powód sięgania po liczby sprzężone ze współczynnikami zespolonymi wyjaśni się na jednym z dalszych wykładów (było go widać w dowodzie twierdzenia na str. 4.4).

Zbadajmy, jak przy zmianie bazy zmienia się baza sprzężona. Niech C oznacza macierz przejścia od bazy $y_1, \dots, y_n$ do $x_1, \dots, x_n$. Niech teraz $b \in \mathbb{K}^n$ będzie wektorem współrzędnych wektora x w bazie $y_1, \dots, y_n$. Mamy zatem $a = Cb$,

czyli $b = C^{-1}a$. Ten sam funkcjonał ξ w bazie sprzężonej z bazą $y_1, \dots, y_n$ jest reprezentowany przez wektor $h = C^H g$; istotnie, $h^H b = g^H C C^{-1} a$.

Jeśli więc dokonujemy zmiany bazy przestrzeni V i macierzą zmiany bazy jest C^{-1}, to baza sprzężona podlega zmianie opisanej za pomocą macierzy C.
W związku z tym elementy przestrzeni V^*, czyli funkcjonały liniowe na V określa się mianem wektorów kowariantnych (bardziej po polsku: współzmienniczych), a elementy przestrzeni V (czyli „zwykłe” wektory) noszą miano wektorów kontrawariantnych (przeciwzmienniczych). Ponieważ role przestrzeni V i V^* można zamienić, więc ta terminologia ma charakter umowny.

Z powyższych uwag na temat przestrzeni V i V^* wynika, że choć są to przestrzenie izomorficzne (bo mają ten sam wymiar), nie istnieje żaden „naturalny” izomorfizm, który byłby niezmienniczy względem wyboru baz. Istotnie, gdybyśmy utożsamili elementy pewnej bazy $x_1, \dots, x_n$ przestrzeni V z elementami $\varphi_1, \dots, \varphi_n$ bazy sprzężonej, to utożsamienie to nie dotyczy innych baz sprzężonych ze sobą; zauważmy, że w $\mathbb{K}^n$ na ogół $A^H x \neq A^{-1}x$.

Przestrzeń V^{**}, tj. przestrzeń funkcjonałów określonych w przestrzeni V^* jest jednak łatwo utożsamić w rozpatrywanym sensie z przestrzenią V. Wystarczy wziąć dowolną bazę $x_1, \dots, x_n$ przestrzeni V, bazę z nią sprzężoną $\varphi_1, \dots, \varphi_n$ przestrzeni V^* i bazę $f_1, \dots, f_n$ przestrzeni V^{**} sprzężoną z tą ostatnią bazą. Utożsamiamy funkcjonał f_i z wektorem x_i i sprawdzamy, że dowolna zmiana bazy przestrzeni V i odpowiadająca jej zmiana bazy przestrzeni V^{**} są opisane za pomocą tej samej macierzy.

Transformacje macierzy przekształceń liniowych

Weźmy teraz dwie bazy $x_1, \dots, x_n$ i $y_1, \dots, y_n$ przestrzeni V, ich bazy sprzężone $\varphi_1, \dots, \varphi_n$ i $\psi_1, \dots, \psi_n$, a także dwie bazy $u_1, \dots, u_m$ i $v_1 \dots, v_m$ przestrzeni W i bazy z nimi sprzężone $\omega_1, \dots, \omega_m$ i $\gamma_1, \dots, \gamma_m$. Bazy te będziemy reprezentowali przez macierze odpowiednio X, Y, Φ, Ψ, oraz U, V, Ω i Γ. Interpretując te macierze jako przekształcenia, otrzymujemy zależności $\Phi = X^{-1}$, $\Psi = Y^{-1}$, $U = \Omega^{-1}$ i $V = \Gamma^{-1}$. Dodatkowo oznaczmy macierz $C = \Phi Y$, która jest macierzą przejścia od bazy $y_1, \dots, y_n$ do $x_1, \dots, x_n$ i macierz $D = \Omega V$, która jest macierzą przejścia od bazy $v_1, \dots, v_m$ do $u_1, \dots, u_m$.

Niech A oznacza macierz pewnego przekształcenia liniowego $f\colon V \to W$ w bazach $x_1, \dots, x_n$ i $u_1, \dots, u_m$. Macierz B tego przekształcenia w bazach $y_1, \dots, y_n$ i $v_1 \dots, v_m$ jest równa $D^{-1}AC$. Istotnie, jeśli współczynniki wektora $b \in \mathbb{K}^n$ są współrzędnymi wektora $\boldsymbol{x}$ w bazie $\boldsymbol{y}_1, \ldots, \boldsymbol{y}_n$, to współczynniki wektora $C\boldsymbol{b}$ są współrzędnymi wektora $\boldsymbol{x}$ w bazie $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n$, współczynniki wektora $AC\boldsymbol{b} \in \mathbb{K}^m$ są współrzędnymi wektora $f(\boldsymbol{x})$ w bazie $\boldsymbol{u}_1, \ldots, \boldsymbol{u}_m$ i wreszcie wektor $B\boldsymbol{b} = D^{-1}AC\boldsymbol{b} \in \mathbb{K}^m$ składa się ze współczynników $f(\boldsymbol{x})$ w bazie $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_m$.

Jeśli macierz A reprezentuje przekształcenie liniowe $f\colon V \to W$ w bazach $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n$ i $\boldsymbol{u}_1, \ldots, \boldsymbol{u}_m$, to reprezentuje ona również pewne przekształcenie liniowe $W^* \to V^*$ w bazach $\omega_1, \ldots, \omega_m$ i $\varphi_1, \ldots, \varphi_n$. Przekształcenie to nazwiemy przekształceniem sprzężonym z f i oznaczymy symbolem f^*.
Przekształcenie to będziemy zapisywać nieco innym wzorem niż f; mając wektor $\boldsymbol{a} \in \mathbb{K}^n$ współrzędnych wektora $\boldsymbol{x} \in V$ w bazie $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n$, możemy obliczyć wektor $\boldsymbol{b} \in \mathbb{K}^m$ współrzędnych wektora $f(\boldsymbol{x})$ ze wzoru $\boldsymbol{b} = A\boldsymbol{a}$. Przypuśćmy, że dany funkcjonał $\beta \in W^*$ możemy przedstawić w postaci $\beta = \bar{c}_1\omega_1 + \cdots + \bar{c}_m\omega_m$ (czyli współczynniki macierzy $\boldsymbol{c}^H = [\bar{c}_1, \ldots, \bar{c}_m] \in \mathbb{K}^{1,m}$ są współrzędnymi funkcjonału β w bazie $\omega_1, \ldots, \omega_m$). Wtedy macierz $\boldsymbol{d}^H = [\bar{d}_1, \ldots, \bar{d}_n] \in \mathbb{K}^{1,n}$ współrzędnych funkcjonału $f^*(\beta)$ otrzymamy ze wzoru $\boldsymbol{d}^H = \boldsymbol{c}^H A$. Równoważnie możemy pisać $\boldsymbol{d} = A^H\boldsymbol{c}$.

Wcześniej już rozpatrywaliśmy przekształcenia sprzężone (str. 4.4); dowolnemu przekształceniu liniowemu $f\colon \mathbb{K}^n \to \mathbb{K}^m$ odpowiada pewna macierz $A \in \mathbb{K}^{m,n}$. Macierz $A^H \in \mathbb{K}^{n,m}$ reprezentuje przekształcenie $f^*\colon \mathbb{K}^m \to \mathbb{K}^n$ sprzężone z f, jeśli przyjmiemy, że przestrzeń sprzężona z $\mathbb{K}^n$ to $\mathbb{K}^n$ (uwaga: przestrzeń sprzężona z $\mathbb{K}^n$ to $\mathbb{K}^{1,n}$, jeśli obliczenie wartości funkcjonału na wektorze polega na zwykłym mnożeniu macierzy, ale zapis, którego tu używam jest usprawiedliwiony tym, że hermitowskie sprzężenie jest bijekcją), a przestrzeń sprzężona z $\mathbb{K}^m$ to (analogicznie) $\mathbb{K}^m$. Jeśli wektor $\boldsymbol{y} \in \mathbb{K}^n$ uznamy za funkcjonał, to jego wartość dla wektora $\boldsymbol{x} \in \mathbb{K}^n$ jest równa $\boldsymbol{y}^H\boldsymbol{x}$.

Podobnie, jak w przypadku przekształcenia f, znajdziemy macierz reprezentującą przekształcenie f^* w bazach $\gamma_1, \ldots, \gamma_m$ i $\psi_1, \ldots, \psi_n$. Ponieważ to są bazy sprzężone z bazami $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_m$ i $\boldsymbol{y}_1, \ldots, \boldsymbol{y}_n$, a macierzą przekształcenia f w tych bazach jest $B = D^{-1}AC$, więc jest to również poszukiwana macierz przekształcenia f^*. Jeśli chcemy reprezentować funkcjonały za pomocą macierzy kolumnowych, to macierz przekształcenia f^* jest równa $B^H = C^H A^H D^{-H}$.

W szczególnym przypadku weźmy $V = W$ i $\boldsymbol{u}_i = \boldsymbol{x}_i$, $\boldsymbol{v}_i = \boldsymbol{y}_i$, $\omega_i = \varphi_i$ oraz $\gamma_i = \psi_i$ dla $i = 1, \ldots, n$. Macierz przejścia od bazy $\boldsymbol{y}_1, \ldots, \boldsymbol{y}_n$ do $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n$ oznaczymy literą C. Jeśli macierz A reprezentuje przekształcenie f w bazie $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n$ (tj. zarówno argument przekształcenia jak i jego obraz reprezentujemy za pomocą współczynników w tej samej bazie), to macierzą przekształcenia f w bazie

$\boldsymbol{y}_1, \ldots, \boldsymbol{y}_n$ jest macierz $B = C^{-1}AC$, a jeśli chcemy funkcjonały reprezentować za pomocą macierzy kolumnowych, to macierzą przekształcenia f^* jest macierz $B^H = C^H A^H C^{-H}$.

Jeśli dla ustalonych dwóch macierzy, A i B, istnieje macierz C, taka że $B = C^{-1}AC$, to macierze te opisują to samo przekształcenie liniowe w dwóch różnych bazach. Macierze o takiej własności nazywamy macierzami podobnymi.

Zadania i problemy

1. Udowodnij, że działanie na liczbach całkowitych, tj. klasach abstrakcji opisanej w wykładzie relacji równoważności w zbiorze par liczb naturalnych, zdefiniowane wzorem

   $$[(a,b)] \diamond [(c,d)] \stackrel{\text{def}}{=} [(a+c, b+d)]$$

   nie zależy od wyboru reprezentantów (to działanie jest dodawaniem liczb całkowitych).

   Określ mnożenie liczb całkowitych przez znalezienie wzoru, który umożliwia znalezienie reprezentanta wyniku na podstawie danych reprezentantów argumentów. Udowodnij, że to działanie nie zależy od wyboru reprezentantów.
2. Użyj zasady abstrakcji w zbiorze par liczb całkowitych do zdefiniowania liczb wymiernych. Określ dodawanie i mnożenie tych liczb za pomocą odpowiednich działań na reprezentantach.
3. Wskaż wszystkie klasy abstrakcji określone przez relację $\sim$ w zbiorze $\mathbb{C}[x]_n$, zdefiniowaną wzorem $w_1 \sim w_2 \Leftrightarrow (w_1(x) = 0 \Leftrightarrow w_2(x) = 0)$. Czy te klasy abstrakcji są równoliczne?
4. Relacja „$\sim$” w zbiorze $\mathbb{R}[x]_n$ jest zdefiniowana warunkiem $w_1 \sim w_2 \Leftrightarrow (w_1(x) = 0 \Leftrightarrow w_2(x) = 0)$, natomiast relacja „$\sim'$” warunkiem $w_1 \sim' w_2 \Leftrightarrow (w_1 = a w_2(x)$ dla pewnego $a \neq 0)$. Czy relacje te są identyczne?
5. Pierścień wielomianów rzeczywistych jednej zmiennej, $\mathbb{R}[x]$, zawiera podpierścień wielomianów podzielnych przez wielomian $x^2 + 1$. Wprowadzimy w $\mathbb{R}[x]$ relację równoważności „$\sim$”, spełnioną przez dwa wielomiany, a i b, jeśli wielomian $a - b$ jest podzielny przez $x^2 + 1$.
   Zbiór klas abstrakcji takiej relacji jest tzw. pierścieniem ilorazowym. Dowolną klasę abstrakcji można reprezentować za pomocą reszty z dzielenia przez $x^2 + 1$ dowolnego wielomianu należącego do tej klasy. Reszta jest wielomianem stopnia co najwyżej 1, a więc ten pierścień ilorazowy można utożsamić ze zbiorem (pierścieniem) wielomianów, w którym mnożenie odbywa się modulo $x^2 + 1$. Dodawanie klas abstrakcji jest równoważne dodawaniu reprezentantów tych klas (czyli jest realizowane przez „zwykłe” dodawanie wielomianów).

a) Udowodnij, że reprezentantem stopnia co najwyżej 1 klasy abstrakcji wielomianu $w(x) = a_0 + a_1x + a_2x^2 + \cdots + a_nx^n$ jest wielomian $p(x) = (a_0 - a_2 + a_4 - \cdots) + (a_1 - a_3 + a_5 - \cdots)x$.

b) Sprawdź, że $((a_1 + b_1x) + (a_2 + b_2x)) \bmod (x^2 + 1) = (a_1 + a_2) + (b_1 + b_2)x$ oraz $((a_1 + b_1x)(a_1 + b_2x)) \bmod (x^2 + 1) = (a_1a_2 - b_1b_2) + (a_1b_2 + b_1a_2)x$.

Jeśli zamiast x napiszesz i, to otrzymasz wzory definiujące mnożenie liczb zespolonych. Jest to więc inny sposób ich określenia. Zwróć uwagę, że ten pierścień ilorazowy jest ciałem.

6. Udowodnij, że jeśli $\boldsymbol{x}' \in [\boldsymbol{x}]$ to $a[\boldsymbol{x}'] = a[\boldsymbol{x}]$.
7. Udowodnij, że jeśli macierze A i B są podobne, to mają ten sam rząd.
8. Jaki jest rząd macierzy przekształcenia przekształcenia $f\colon \mathbb{R}[x]_2 \to \mathbb{R}[x]_2$, danego wzorem $f(w) = aw''x^2 + bw'x + w$, w zależności od parametrów a i b?
9. Zbadaj, czy jeśli macierze A i B są podobne, oraz macierz A jest hermitowska (tj. $A^H = A$) i macierz C podobieństwa macierzy A i B (taka, że $B = C^{-1}AC$) jest hermitowska, to macierz B jest hermitowska.
10. Niech f oznacza przekształcenie przestrzeni $\mathbb{R}[x]_2$, określone wzorem $f(w) = w'' + w' + w$.

    Znajdź macierz tego przekształcenia w bazie potęgowej $1, x, x^2$.

    Znajdź macierz C przejścia do bazy Newtona, $1, x-1, (x-1)(x-2)$.

    Oblicz macierz B przekształcenia f w tej bazie Newtona.

Normy wektorów

Pojęcie normy wektora

Def. Niech V oznacza przestrzeń liniową nad ciałem $\mathbb{K}$, które jest jednym z ciał liczbowych, $\mathbb{Q}$, $\mathbb{R}$ lub $\mathbb{C}$. Normą w tej przestrzeni nazywamy dowolną funkcję $\|\cdot\|\colon V \to \mathbb{R}$, o następujących własnościach:

N.1: $\forall_{x\in V}\, \|x\| \geqslant 0$, oraz $\|x\| = 0 \;\Leftrightarrow\; x = 0$.
Norma jest więc funkcją nieujemną, a jedynym jej miejscem zerowym jest wektor 0.

N.2: $\forall_{a\in\mathbb{K}, x\in V}\, \|ax\| = |a|\|x\|$.
Własność ta bywa nazywana półliniowością, przy czym „pół” związane jest z ignorowaniem znaku liczby a.

N.3: $\forall_{x,y\in V}\, \|x + y\| \leqslant \|x\| + \|y\|$.
Powyższa nierówność nazywa się nierównością trójkąta.

Def. Normą wektora x nazywamy liczbę $\|x\|$.

Def. Przestrzeń unormowana jest to przestrzeń liniowa V, w której jest określona ustalona norma $\|\cdot\|$.

Warto podkreślić, że w przestrzeni, której wymiar jest większy od 0, można określić nieskończenie wiele różnych norm. Wybierając różne normy w tej samej przestrzeni liniowej otrzymamy za każdym razem *inną* przestrzeń unormowaną.

Stwierdzenie: *Jeśli funkcja $\|\cdot\|_a\colon V \to \mathbb{R}$ jest normą, a przekształcenie liniowe $f\colon V \to V$ jest ograniczone i różnowartościowe, to funkcja $\|\cdot\|_b\colon V \to \mathbb{R}$, określona wzorem $\|x\|_b = \|f(x)\|_a$, też jest normą w przestrzeni V.*
Dowód polega na sprawdzeniu, że funkcja $\|\cdot\|_b$ ma wszystkie podane wyżej własności normy. Mając pewną normę w przestrzeni możemy więc definiować inne normy. Trzeba tylko umieć określić choć jedną normę.

Powód, dla którego wprowadza się normy, jest następujący: mając dwie liczby (np. rzeczywiste) x i y, możemy zmierzyć *odległość* tych liczb, albo błąd, z jakim x przybliża y, obliczając $|x - y|$. Mając dwa wektory w dowolnej przestrzeni liniowej V, chcielibyśmy móc zrobić to samo, tj. wyrazić błąd, z jakim wektor x przybliża wektor y, *za pomocą jednej liczby.* W tym celu potrzebna jest metryka, tj. funkcja, która opisuje odległość dowolnych dwóch wektorów

w przestrzeni V. Jednym z najprostszych sposobów określenia metryki w przestrzeni liniowej jest właśnie użycie normy; za odległość wektorów x i y przyjmujemy liczbę równą $\|x - y\|$.

Normy Höldera

Jeden z najczęściej stosowanych sposobów określenia normy w przestrzeni $\mathbb{K}^n$ polega na użyciu wzoru (dla $x = [x_1, \dots, x_n]^T$)

$$\|x\| = \Big(\sum_{i=1}^{n} |x_i|^p\Big)^{1/p},$$

w którym występuje ustalona liczba rzeczywista $p \geqslant 1$. Łatwo jest udowodnić, że tak określona funkcja ma dwie spośród trzech podanych wyżej własności, tj. jest dodatnia dla $x \neq 0$ i półliniowa. Dowód, że funkcja ta spełnia nierówność trójkąta, tj. że

$$\Big(\sum_{i=1}^{n} |x_i + y_i|^p\Big)^{1/p} \leqslant \Big(\sum_{i=1}^{n} |x_i|^p\Big)^{1/p} + \Big(\sum_{i=1}^{n} |y_i|^p\Big)^{1/p}$$

dla dowolnych liczb $x_1, \dots, x_n$ i $y_1, \dots, y_n \in \mathbb{K}$, jest znacznie trudniejszy i pominiemy go (dalej zbadamy tylko pewne przypadki szczególnie łatwe). Powyższa nierówność nosi nazwę nierówności Minkowskiego, a normy przestrzeni $\mathbb{K}^n$ określone w opisany tu sposób są nazywane normami Höldera. Normę p-tą oznacza się symbolem $\|\cdot\|_p$.

Przypadki szczególne norm Höldera, często spotykane w praktycznych rachunkach, odpowiadają $p = 1$ i $p = 2$. Mamy też przypadek graniczny, dla $p \to \infty$. Te trzy szczególne normy można obliczać na podstawie łatwych do udowodnienia wzorów

$$\|x\|_1 = \sum_{i=1}^{n} |x_i|, \qquad \|x\|_2 = \sqrt{\sum_{i=1}^{n} |x_i|^2} = \sqrt{x^H x}, \qquad \|x\|_\infty = \max_{i\in\{1,\dots,n\}} |x_i|.$$

Norma druga, $\|\cdot\|_2$, bywa określana mianem normy euklidesowej. O normie $\|\cdot\|_\infty$ często mówi się „norma maksimum”.

Możemy określić normę w dowolnej przestrzeni n-wymiarowej V, ustalając w niej układ współrzędnych (czyli izomorfizm $V \to \mathbb{K}^n$), a następnie przyjmując za normę wektora $y \in V$ wybraną normę Höldera jego wektora współrzędnych.

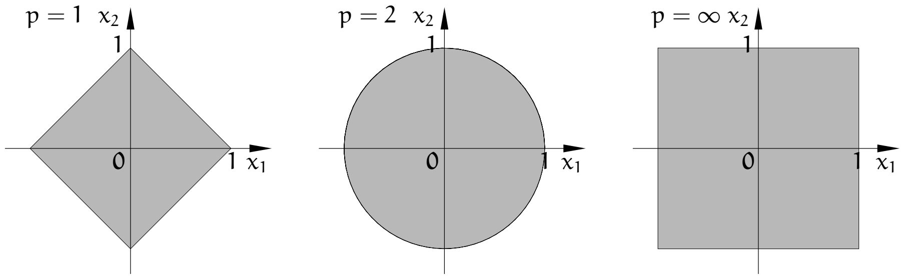
Wspomnijmy w tym miejscu o normach Höldera w przestrzeniach funkcji; dla rzeczywistej lub zespolonej funkcji f o dziedzinie A można napisać

$$\|f\|_p = \Big(\int_A |f(x)|^p \,dx\Big)^{1/p}.$$

Jeśli całka występująca w tym wzorze istnieje i jest skończona, to wyrażenie po prawej stronie wzoru określa normę funkcji f. Zdanie to sugeruje, że nie każda funkcja ma normę. Dlatego rozpatruje się przestrzenie liniowe złożone z wszystkich tych funkcji (o dziedzinie A), których norma jest określona (tj. całka istnieje i jest skończona), a zatem norma jest elementem określenia przestrzeni. W powyższym określeniu normy funkcji jest pewna nieścisłość, wyjaśniona w pierwszym zadaniu po tym wykładzie.

Def. Kulą jednostkową nazywamy zbiór wszystkich tych wektorów x, których norma jest mniejsza lub równa 1.
Sfera jednostkowa jest zbiorem wektorów x, takich że $\|x\| = 1$.

Określenie metryki w przestrzeni V, poprzez ustalenie normy, umożliwia określenie otoczenia dowolnego punktu (wektora) x; może nim być zbiór tych punktów y, dla których liczba $\|y-x\|$ jest mniejsza niż ustalone $\varepsilon > 0$. Mając dowolny podzbiór X przestrzeni V możemy wyróżnić wnętrze tego zbioru, czyli zbiór tych elementów zbioru X, które należą do niego wraz z pewnym otoczeniem, a także brzeg zbioru X, czyli zbiór tych punktów, których wszystkie otoczenia zawierają punkty należące do X i do $V \setminus X$.
Możemy teraz powiedzieć, że sfera jednostkowa jest brzegiem kuli jednostkowej.



Na rysunku są przedstawione kule jednostkowe w przestrzeni $\mathbb{R}^2$ dla $p = 1$, $p = 2$ i $p = \infty$, czyli dla przypadków, z którymi będziemy się stykać najczęściej. Zauważmy, że wszystkie te kule są zbiorami wypukłymi i symetrycznymi względem punktu 0. W ogólności w przestrzeni liniowej V nad $\mathbb{R}$ można wybrać dowolny zbiór K, który ma te dwie własności, a ponadto ma niepuste wnętrze

i jest domknięty i ograniczony (te własności badamy przy użyciu dowolnej normy), i określić normę, dla której zbiór K jest kulą jednostkową.

Udowodnimy nierówność Minkowskiego w szczególnym przypadku $p = 2$.
W tym celu dowiedziemy najpierw nierówności Schwarza: jeśli $x, y \in \mathbb{C}^n$, to

$$|x^H y| \leqslant \|x\|_2 \|y\|_2.$$

Nierówność ta jest oczywista, jeśli któryś z wektorów jest równy 0; załóżmy zatem, że $y \neq 0$ i przyjmując parametr zespolony t obliczmy

$$\|x + yt\|_2^2 = (x+yt)^H(x+yt) = x^Hx + x^Hyt + y^Hx\bar{t} + y^Hyt\bar{t}.$$

Wartością tego wyrażenia dla każdego $t \in \mathbb{C}$ jest nieujemna liczba rzeczywista. Możemy oznaczyć $x^Hy = a = |a|e^{i\varphi}$. Podstawiając $t = |t|e^{i\psi}$ otrzymujemy

$$0 \leqslant \|x\|_2^2 + |a|\,|e^{i(\varphi+\psi)} + e^{-i(\varphi+\psi)}|\,|t| + \|y\|_2^2|t|^2.$$

Przyjmując kolejno $\psi = -\varphi$ i $\psi = \pi - \varphi$ otrzymujemy nierówności

$$0 \leqslant \|x\|_2^2 + 2|x^Hy|\,|t| + \|y\|_2^2|t|^2 \quad \text{oraz} \quad 0 \leqslant \|x\|_2^2 - 2|x^Hy|\,|t| + \|y\|_2^2|t|^2.$$

Biorąc następnie liczbę rzeczywistą $\tau = |t|$ w pierwszym przypadku i $\tau = -|t|$ w drugim, otrzymujemy za każdym razem nierówność

$$0 \leqslant \|x\|_2^2 + 2|x^Hy|\tau + \|y\|_2^2\tau^2.$$

Nierówność ta obowiązuje dla każdego $\tau \in \mathbb{R}$. Trójmian kwadratowy po prawej stronie nie może mieć dwóch różnych rzeczywistych miejsc zerowych, a stąd wyróżnik tego trójmianu (dla $a + 2b\tau + c\tau^2$ jest to $\Delta = b^2 - ac$) nie może być dodatni. Mamy zatem

$$|x^Hy|^2 - \|x\|_2^2\|y\|_2^2 \leqslant 0,$$

skąd nierówność Schwarza wynika natychmiast.
Możemy napisać

$$\|x+y\|_2^2 = (x+y)^H(x+y) = \|x\|_2^2 + x^Hy + y^Hx + \|y\|_2^2 = \\ \|x\|_2^2 + 2\mathrm{Re}(x^Hy) + \|y\|_2^2,$$

oraz $(\|x\|_2 + \|y\|_2)^2 = \|x\|_2^2 + 2\|x\|_2\|y\|_2 + \|y\|_2^2$.
Na podstawie nierówności Schwarza $2\mathrm{Re}(x^Hy) \leqslant 2|x^Hy| \leqslant 2\|x\|_2\|y\|_2$, zatem $\|x+y\|_2^2 \leqslant (\|x\|_2 + \|y\|_2)^2$ i dowód nierówności Minkowskiego dla $p = 2$ na tym się kończy. □

Uwaga 1: Nierówność Schwarza jest ostra (tj. równość nie zachodzi), jeśli wektory x i y są niezależne liniowo. Nierówność Minkowskiego jest ostra wtedy, gdy nie istnieje liczba nieujemna a, taka że $x = ay$ lub $y = ax$.
Norma, dla której nierówność trójkąta jest ostra, jeśli podstawić do niej wektory spełniające podany wyżej warunek, nazywa się normą ostrą. Wszystkie normy Höldera oprócz $\|\cdot\|_1$ i $\|\cdot\|_\infty$ są ostre.

Uwaga 2: Dowolna norma Höldera wektora $x \in \mathbb{K}^n$ nie zmieni się, jeśli poprzestawiamy jego współczynniki, ani jeśli pozmieniamy ich znaki.

Normy przekształceń liniowych

Niech f będzie przekształceniem liniowym przestrzeni unormowanej X w przestrzeń unormowaną Y. Możemy określić normę przekształcenia f, wzorem

$$\|f\| = \sup_{x\in X\setminus\{0\}} \frac{\|f(x)\|_Y}{\|x\|_X}.$$

Tak określona funkcja określa, „ile razy dłuższy" od dowolnego wektora x może być jego obraz w przekształceniu f. Sprawdzenie, że funkcja ta w przestrzeni $L(X;Y)$ ma wszystkie własności normy, pozostawiam na ćwiczenia. Określona w ten sposób norma nazywa się normą indukowaną przez normy przestrzeni X i Y.

Można zauważyć, że dla dowolnych norm $\|\cdot\|_X$ i $\|\cdot\|_Y$ wyrażenie z prawej strony powyższego wzoru jest równe

$$\sup_{x\in X:\ \|x\|_X\leqslant 1} \|f(x)\|_Y = \sup_{x\in X:\ \|x\|_X=1} \|f(x)\|_Y.$$

Chcąc znaleźć normę danego przekształcenia f można badać te wyrażenia, choć w ogólności zadanie to nadal może być trudne. W przestrzeniach skończenie wymiarowych kule jednostkowe są zbiorami zwartymi (zbiór skończenie wymiarowy jest zwarty wtedy i tylko wtedy gdy jest domknięty i ograniczony). Zamiast supremum możemy wtedy poszukiwać maksimum.

Niech $f\colon X \to Y$ i $g\colon Y \to Z$ będą przekształceniami liniowymi przestrzeni unormowanych. Wtedy norma indukowana przez $\|\cdot\|_X$ i $\|\cdot\|_Z$ przekształcenia złożonego $f\circ g$, spełnia nierówność

$$\|f\circ g\| \leqslant \|g\|\|f\|,$$

w której występują normy przekształceń f i g, indukowane odpowiednio przez $\|\cdot\|_X$ i $\|\cdot\|_Y$ oraz $\|\cdot\|_Y$ i $\|\cdot\|_Z$.

Ważny przypadek szczególny otrzymamy rozpatrując przekształcenie $f\colon X \to X$, dla przestrzeni X, w której rozpatrujemy tylko jedną normę.

Normy indukowane macierzy

Macierz $A \in \mathbb{K}^{m,n}$ interpretujemy jako przekształcenie liniowe przestrzeni $\mathbb{K}^n$ w $\mathbb{K}^m$. Wybrawszy normy w tych przestrzeniach możemy określić normę indukowaną w przestrzeni $\mathbb{K}^{m,n}$. W praktyce ograniczymy się do norm indukowanych przez normę p-tą Höldera w jednej i drugiej przestrzeni.

Normę macierzy A, indukowaną przez normy p-te w przestrzeniach $\mathbb{K}^n$ i $\mathbb{K}^m$, będziemy oznaczać symbolem $\|A\|_p$ (który wygląda tak samo jak symbol normy p-tej macierzy kolumnowej, ale ma ogólniejsze znaczenie). Zauważmy, że macierz $x \in \mathbb{K}^{m,1} = \mathbb{K}^m$ opisuje przekształcenie przestrzeni $\mathbb{K}^1$ w $\mathbb{K}^m$; przekształcając dowolny element sfery jednostkowej w $\mathbb{K}^1$, tj. liczbę a, taką że $|a| = 1$, otrzymujemy wektor ax, którego norma jest równa $\|x\|$. Tak więc, jeśli $n = 1$, to obie interpretacje symbolu $\|x\|_p$, tj. norma p-ta wektora $x \in \mathbb{K}^m$ i norma przekształcenia $x\colon \mathbb{K}^1 \to \mathbb{K}^m$ indukowana przez normę p-tą, są ze sobą zgodne.

Interpretując normy macierzy jako normy przekształceń indukowane przez normy wektorów, możemy uzasadnić następującą nierówność:

$$\|AB\| \leqslant \|A\|\|B\|,$$

dla wszystkich macierzy, których iloczyn jest określony. Dla pewnego wektora z, takiego że $\|z\| = 1$, mamy bowiem

$$\|AB\| = \|(AB)z\| = \|A(Bz)\| \leqslant \|A\|\|Bz\| \leqslant \|A\|\|B\|.$$

Ta własność norm macierzy nazywa się submultiplikatywnością.

Indukowana norma pierwsza i norma maksimum macierzy $A = [a_{ij}]_{i,j} \in \mathbb{K}^{m,n}$ są szczególnie łatwe do obliczenia, na podstawie wzorów

$$\|A\|_1 = \max_{j\in\{1,\dots,n\}} \sum_{i=1}^{m} |a_{ij}|,$$
$$\|A\|_\infty = \max_{i\in\{1,\dots,m\}} \sum_{j=1}^{n} |a_{ij}|.$$

Udowodnimy drugi z tych wzorów, dowód pierwszego zostawiając na ćwiczenia. Weźmy wektor $z = [z_1,\dots,z_n]^T$, taki że $\|z\|_\infty \leqslant 1$. Zatem, dla $j = 1,\dots,n$ mamy $|z_j| \leqslant 1$. Współczynniki wektora $Az = x = [x_1,\dots,x_m]^T$ można oszacować następująco:

$$|x_i| = \Bigl|\sum_{j=1}^{n} a_{ij}z_j\Bigr| \leqslant \sum_{j=1}^{n} |a_{ij}z_j| = \sum_{j=1}^{n} |a_{ij}|\,|z_j| \leqslant \sum_{j=1}^{n} |a_{ij}|.$$

Ponieważ $\|x\|_\infty = \max_{i\in\{1,\dots,m\}} |x_i|$, więc mamy górne oszacowanie normy, $\|A\|_\infty \leqslant \max_{i\in\{1,\dots,m\}} \sum_{i=1}^{m} |a_{ij}|$.
Aby oszacować normę macierzy A z dołu, możemy obliczyć normę ustalonego wektora z. Niech i oznacza numer wiersza, w którym suma wartości bezwzględnych współczynników macierzy A jest największa. Współczynniki w tym wierszu są liczbami zespolonymi $a_{i1} = |a_{i1}|s_1, \dots, a_{in} = |a_{in}|s_n$ (jeśli macierz A jest rzeczywista, to $s_j = \pm 1$). Ponieważ $s_j\bar{s}_j = 1$, więc biorąc wektor $z = [\bar{s}_1, \dots, \bar{s}_n]^T$ i obliczając współczynnik x_i wektora $x = Az$ otrzymamy

$$x_i = \sum_{j=1}^{n} |a_{ij}|,$$

a ponieważ $\|x\| \geqslant |x_i|$, więc stąd wynika nierówność $\|A\|_\infty \geqslant \max_{i\in\{1,\dots,m\}} \sum_{j=1}^{m} |a_{ij}|$. Z otrzymanych dwóch nierówności wynika wzór, którego mieliśmy dowieść. □

Z podanych dwóch wzorów, z których jeden udowodniliśmy wynika, że $\|A^T\|_1 = \|A\|_\infty$, oraz $\|A^T\|_\infty = \|A\|_1$, a ponadto możemy przepisać te wzory zamieniając transpozycję na hermitowskie sprzężenie.

Często stosuje się normę drugą wektorów. Indukowana przez nią norma druga macierzy jest trudniejsza do obliczenia, a przy tym we wzorze, który ją opisuje, występują pojęcia spoza zakresu tego wykładu[1]. Ale możemy udowodnić (ćwiczenia), że

$$\|A\|_2 = \max_{\substack{x\in\mathbb{K}^m, y\in\mathbb{K}^n\\ \|x\|_2=\|y\|_2=1}} |x^H A y|.$$

Na podstawie tego wzoru łatwo jest pokazać, że

$$\|A\|_2 = \|A^T\|_2 = \|A^H\|_2 = \|\bar{A}\|_2.$$

<u>Stwierdzenie:</u> *Jeśli normy $\|\cdot\|_a$ i $\|\cdot\|_b$ w $\mathbb{K}^n$ spełniają warunek $\forall_x \|x\|_b = \|Bx\|_a$, dla ustalonej nieosobliwej macierzy B, to normy w przestrzeni $\mathbb{K}^{n,n}$, indukowane przez te dwie normy, są związane zależnością $\|A\|_b = \|BAB^{-1}\|_a$.*
<u>Dowód:</u> Mamy

$$\|A\|_b = \max_{\|y\|_b=1} \|Ay\|_b = \max_{\|y\|_b=1} \|AB^{-1}By\|_b = \max_{\|x\|_a=1} \|BAB^{-1}x\|_a = \|BAB^{-1}\|_a.$$

□

[1] Jest $\|A\|_2 = \sqrt{\lambda_{\max}(A^H A)}$, gdzie $\lambda_{\max}(A^H A)$ oznacza największą tzw. wartość własną macierzy hermitowskiej $A^H A$, tj. największą liczbę λ, taką że macierz $A^H A - \lambda I$ jest osobliwa (wszystkie takie liczby są rzeczywiste i nieujemne).

Nierówności norm wektorów

<u>Def.</u> Dwie normy, $\|\cdot\|_a$ i $\|\cdot\|_b$, określone w tej samej przestrzeni V, nazywamy <u>równoważnymi</u>, jeśli istnieją liczby dodatnie c_1 i c_2, takie że dla każdego $x \in V$ spełnione są nierówności

$$c_1\|x\|_a \leqslant \|x\|_b \leqslant c_2\|x\|_a.$$

Wszystkie normy w przestrzeni o skończonym wymiarze są równoważne, zatem kwestia polega na znalezieniu odpowiednich liczb c_1 i c_2. Jeśli je znamy, to wykonując rachunki mające na celu oszacowanie normy $\|\cdot\|_a$ pewnego wektora x możemy posłużyć się oszacowaniami $\|x\|_b$, jeśli tylko umiemy je znaleźć.

Najważniejsze w praktyce normy w przestrzeni $\mathbb{K}^n$ spełniają następujące nierówności (dowody na ćwiczeniach):

$$\|x\|_\infty \leqslant \|x\|_1 \leqslant n\|x\|_\infty,$$
$$\|x\|_\infty \leqslant \|x\|_2 \leqslant \sqrt{n}\|x\|_\infty,$$
$$\|x\|_2 \leqslant \|x\|_1 \leqslant \sqrt{n}\|x\|_2.$$

Nierówności norm macierzy

Rozważmy p-tą normę wektorową i normę macierzy przez nią indukowaną.
Jeśli macierz $A \in \mathbb{K}^{m,n}$ przedstawimy w postaci blokowo-wierszowej, $A = [A_1, \dots, A_r]$, w której bloki $A_1, \dots, A_r$ mogą mieć różną liczbę kolumn, to dla $j = 1, \dots, r$ mamy $\|A_j\|_p \leqslant \|A\|_p$.

<u>Dowód:</u> Dowolny wektor w $\mathbb{K}^n$ możemy przedstawić w postaci $z = \begin{bmatrix} z_1 \\ \vdots \\ z_r \end{bmatrix}$. Wtedy

$$\|A_j\|_p = \max_{z_j\colon \|z_j\|_p=1} \|A_j z_j\|_p \leqslant \max_{z\colon \|z\|_p=1} \Bigl\|\sum_{j=1}^{r} A_j z_j\Bigr\| = \|A\|_p. \quad \square$$

Podobnie, jeśli macierz A przedstawimy w postaci blokowo-kolumnowej, $A = \begin{bmatrix} A_1 \\ \vdots \\ A_q \end{bmatrix}$, to $\|A_i\|_p \leqslant \|A\|_p$ dla $i = 1, \dots, q$. Łatwy dowód tego faktu pozostawiam na ćwiczenia. Tymczasem zauważmy, że stąd i z poprzedniego stwierdzenia wynika, że dla dowolnego podziału blokowego macierzy $A = [A_{ij}]_{i,j}$, nierówność $\|A_{ij}\|_p \leqslant \|A\|_p$ zachodzi dla każdego bloku A_{ij}.

W wielu zastosowaniach naturalne jest poszukiwanie wartości normy macierzy, indukowanej przez normę drugą w $\mathbb{K}^n$. Ponieważ norma druga macierzy nie jest bardzo łatwa do obliczenia, więc zamiast niej bywa stosowana norma Frobeniusa, określona wzorem

$$\|A\|_F = \sqrt{\sum_{i=1}^{m}\sum_{j=1}^{n}|a_{ij}|^2}.$$

Zwróćmy uwagę, że (poza przypadkiem $m = 1$ lub $n = 1$) norma ta *nie jest* normą indukowaną przez żadną normę wektorową. Aby tego dowieść, rozważmy macierz jednostkową, I_n. Reprezentuje ona przekształcenie tożsamościowe przestrzeni $\mathbb{K}^n$, oczywiste jest więc, że każda norma indukowana tej macierzy jest równa 1. Tymczasem $\|I_n\|_F = \sqrt{n}$

Stwierdzenie: *Norma Frobeniusa i norma druga macierzy (tj. norma w $\mathbb{K}^{m,n}$ indukowana przez normę drugą w $\mathbb{K}^n$) są równoważne i zachodzą następujące nierówności:*

$$\|A\|_2 \leqslant \|A\|_F \leqslant \sqrt{\min\{m,n\}}\,\|A\|_2.$$

Dowód: Przedstawmy macierz A w postaci blokowo-kolumnowej: $A = \begin{bmatrix} \mathbf{a}_1 \\ \vdots \\ \mathbf{a}_m \end{bmatrix}$.

Weźmy $z \in \mathbb{K}^n$, takie że $\|z\|_2 = 1$ i $\|A\|_2 = \|Az\|_2$. Wtedy (na podstawie nierówności Schwarza)

$$\|A\|_2^2 = \|Az\|_2^2 = \sum_{i=1}^{m}|\mathbf{a}_i z|^2 \leqslant \sum_{i=1}^{m}\|\mathbf{a}_i\|_2^2 = \|A\|_F^2.$$

Na podstawie wcześniejszych stwierdzeń dotyczących podziału blokowego macierzy mamy

$$\|A\|_F^2 = \sum_{j=1}^{m}\|\mathbf{a}_j\|_2^2 \leqslant m\|A\|_2^2,$$

a także $\|A\|_F^2 = \|A^T\|_F^2 \leqslant n\|A^T\|_2^2 = n\|A\|_2^2$, co kończy dowód. □

Dla wygody określa się wartość bezwzględną macierzy $A = [a_{ij}]_{i,j}$, jako macierz macierz $|A| \stackrel{\text{def}}{=} [|a_{ij}|]_{i,j}$.

W zbiorach $\mathbb{Q}^{m,n}$ i $\mathbb{R}^{m,n}$ (ale nie w $\mathbb{C}^{m,n}$) istnieje relacja „nie większy niż” (ozn. symbolem „$\leqslant$”), która jest zwrotna, antysymetryczna i przechodnia, czyli

jest częściowym porządkiem:

$$[a_{ij}]_{i,j} \leqslant [b_{ij}]_{i,j} \stackrel{\text{def}}{\Leftrightarrow} \forall_{i,j}\ a_{ij} \leqslant b_{ij}.$$

Na przykład, dla dowolnej macierzy A zachodzi relacja $A \leqslant |A|$.
Możemy też rozpatrywać relacje „mniejszy niż” ($<$), „nie mniejszy niż” ($\geqslant$) i „większy niż” ($>$), których definicje łatwo jest zgadnąć.

Oprócz szacowania jednej normy danej macierzy przez inną, stosuje się oszacowania normy macierzy przez normę innej macierzy. Norma w przestrzeni $\mathbb{K}^{m,n}$ nazywa się monotoniczną, jeśli z nierówności $|A| \leqslant |B|$ wynika $\|A\| \leqslant \|B\|$. Nie wszystkie normy są monotoniczne; na przykład nie jest monotoniczna norma druga macierzy.

Bez dowodu podam następujące związki między normami macierzy (niektóre z nich są oczywiste, inne mniej):

$$\|A\|_F = \|A^T\|_F = \|A^H\|_F = \|\overline{A}\|_F,$$
$$\||A|\|_1 = \|A\|_1, \quad \||A|\|_\infty = \|A\|_\infty, \quad \||A|\|_F = \|A\|_F,$$
$$\|A\|_2 \leqslant \||A|\|_2, \quad \text{(nie zawsze zachodzi równość)},$$
$$\|A^HA\|_2 = \|A\|_2^2 \leqslant \|A\|_1\|A\|_\infty,$$

Zadania i problemy

1. Def. Seminormą (albo półnormą) nazywamy funkcję $\|\cdot\|'\colon V \to \mathbb{K}$, która spełnia wszystkie warunki określające normę, z wyjątkiem tego, że jedynym jej miejscem zerowym jest wektor $\mathbf{0}$.

   Udowodnij, że zbiór wektorów, których seminorma jest równa 0, jest podprzestrzenią liniową przestrzeni V.

   Udowodnij, że jeśli wszystkie wektory $\mathbf{x}$, takie że $\|\mathbf{x}\|' = 0$ należą do podprzestrzeni X przestrzeni V, to wzór $\|[\mathbf{x}]\| = \|\mathbf{x}\|'$ określa normę w przestrzeni ilorazowej V/X.

   Uzupełnienie informacji na temat norm w przestrzeniach funkcji: istnieją funkcje f różne od 0, takie że $\int_A |f(x)|^p\,dx = 0$; na przykład funkcja f może przyjmować wartości różne od 0 tylko w skończonej liczbie punktów. Dlatego wzór podany na wykładzie określa funkcję nie spełniającą warunku dodatniości, czyli seminormę. Normę otrzymamy, jeśli utożsamimy każde dwie funkcje, których całka różnicy podniesionej do potęgi p jest równa 0. Dlatego np. określenie „przestrzeń funkcji całkowalnych w potędze p” w rzeczywistości oznacza unormowaną przestrzeń ilorazową (o czym często przy różnych okazjach się nie wspomina, bo wszyscy wiedzą).

2. Jaka jest średnica kuli jednostkowej?
3. Udowodnij, że jeśli X i Y są podprzestrzeniami liniowymi przestrzeni V, w których są określone normy odpowiednio $\|\cdot\|_X$ i $\|\cdot\|_Y$, oraz $V = X \oplus Y$ (to jest rozkład wektora $\boldsymbol{v}$ na składowe względem rozkładu przestrzeni V na sumę prostą podprzestrzeni X i Y; rozkład taki jest, przypominam, jednoznaczny), to wzory $\|\boldsymbol{v}\|_a = \|\boldsymbol{x}\|_X + \|\boldsymbol{y}\|_Y$ i $\|\boldsymbol{v}\|_b = \max(\|\boldsymbol{x}\|_X, \|\boldsymbol{y}\|_Y)$, gdzie $\boldsymbol{v} = \boldsymbol{x} + \boldsymbol{y}$, $\boldsymbol{x} \in X$, $\boldsymbol{y} \in Y$, określają normy w przestrzeni V.
4.* Udowodnij, że przy założeniach takich jak w poprzednim zadaniu wzór $\|\boldsymbol{v}\| = (\|\boldsymbol{x}\|_X^p + \|\boldsymbol{y}\|_Y^p)^{1/p}$ dla $p \geqslant 1$ określa normę w przestrzeni V.
5. Sprawdź, że norma indukowana w przestrzeni $L(X; Y)$ rzeczywiście spełnia wszystkie warunki potrzebne do tego, aby być normą.
6. Udowodnij wzór umożliwiający obliczenie normy pierwszej macierzy $A \in \mathbb{K}^{m,n}$.
7. Udowodnij, że norma p-ta macierzy $A \in \mathbb{K}^{m,n}$ jest nie mniejsza niż norma każdej macierzy, która powstaje z A przez odrzucenie pewnych wierszy.
8. Czy norma macierzy określona wzorem $\|[a_{ij}]_{i,j}\| = \max_{i,j} |a_{ij}|$ jest submultiplikatywna? Czy jest ona normą indukowaną przez jakąś normę wektorową? Wskazówka: Spróbuj znaleźć kontrprzykład.
9. Oblicz normy $\|A\|_1$, $\|A\|_\infty$ i $\|A\|_F$ macierzy

$$A = \begin{bmatrix} (1,1) & (2,-1) & (3,0) \\ (0,2) & (3,-2) & (2,1) \end{bmatrix}.$$

10. Udowodnij podane na wykładzie związki między normami $\|\cdot\|_1$, $\|\cdot\|_2$ i $\|\cdot\|_\infty$ w przestrzeni $\mathbb{K}^n$.
11. Niech $\boldsymbol{u} \in \mathbb{K}^m$, $\boldsymbol{v} \in \mathbb{K}^n$. Udowodnij, że

$$\|\boldsymbol{u}\boldsymbol{v}^H\|_p = \|\boldsymbol{u}\|_p \|\boldsymbol{v}\|_q,$$

gdzie $p = 1$ i $q = \infty$, albo $p = q = 2$ lub $p = \infty$, $q = 1$.

Wskazówka: Oszacuj z góry i z dołu normę p-tą wektora $\boldsymbol{u}\boldsymbol{v}^H\boldsymbol{z}$ dla $\boldsymbol{z} \in \mathbb{K}^n$, $\|\boldsymbol{z}\|_p = 1$.

Powyższa nierówność jest prawdziwa dla dowolnych liczb p i q, takich że $1/p + 1/q = 1$ oraz $p \geqslant 1$, $q \geqslant 1$. Do jej dowodu można zastosować następującą nierówność Höldera:

$$|\boldsymbol{y}^H\boldsymbol{z}| \leqslant \|\boldsymbol{y}\|_q \|\boldsymbol{z}\|_p.$$

Szczególnym przypadkiem tej nierówności, dla $p = q = 2$, jest udowodniona na wykładzie nierówność Schwarza.

12. Udowodnij, że indukowana norma druga macierzy $A \in \mathbb{K}^{m,n}$ wyraża się wzorem

$$\|A\|_2 = \max_{\substack{\boldsymbol{x}\in\mathbb{K}^m, \boldsymbol{y}\in\mathbb{K}^n \\ \|\boldsymbol{x}\|_2=\|\boldsymbol{y}\|_2=1}} |\boldsymbol{x}^H A \boldsymbol{y}|.$$

13. Znajdź wzór umożliwiający obliczenie dowolnej normy indukowanej przez p-tą normę Höldera macierzy *diagonalnej* $n \times n$ i podaj dowód jego poprawności.

Wyznaczniki

Def. Wyznacznikiem stopnia n nazywamy funkcję $\det_n\colon \mathbb{K}^{n,n} \to \mathbb{K}$, taką że

- jeśli określimy funkcję $f\colon \mathbb{K}^n \to \mathbb{K}$ wzorem $f(\boldsymbol{x}) = \det_n[\boldsymbol{a}_1, \dots, \boldsymbol{a}_{n-1}, \boldsymbol{x}]$, to dla dowolnych ustalonych wektorów $\boldsymbol{a}_1, \dots, \boldsymbol{a}_{n-1} \in \mathbb{K}^n$ funkcja ta jest funkcjonałem liniowym,
- jeśli macierz A' otrzymamy przez przestawienie dwóch sąsiednich kolumn macierzy A, to $\det_n A' = -\det_n A$,
- $\det_n I_n = 1$.

Def. Wyznacznikiem macierzy $A \in \mathbb{K}^{n,n}$ nazywamy skalar (liczbę) $\det_n A$.

Dalej zamiast $\det_n$ na ogół będziemy pisać det. Mając definicję sformułowaną przez podanie pewnych warunków musimy upewnić się, czy definiowany obiekt jest *dobrze określony*, a zatem czy istnieje funkcja mająca wymienione własności i czy jest tylko jedna taka funkcja. To pierwsze możemy sprawdzić jeśli zgadniemy dowolny wzór, który określa funkcję spełniającą te warunki. To drugie jest prawdą, jeśli nałożone warunki (nie wystarczy, jeśli odgadnięty wzór) umożliwiają *jednoznaczne* obliczenie wartości takiej funkcji dla ustalonej macierzy.

Łatwo jest zauważyć, że z podanych warunków wynikają następujące własności funkcji det: zamiana miejscami dowolnych dwóch kolumn macierzy powoduje zmianę znaku wyznacznika na przeciwny. Ogólniej, poddanie kolumn macierzy dowolnej permutacji powoduje zmianę znaku na przeciwny wtedy, gdy permutacja ta jest nieparzysta (czyli zawiera nieparzystą liczbę *nieporządków*, str. 1.11, zad. 7), a więc można ją przedstawić jako złożenie nieparzystej liczby transpozycji. W przeciwnym razie przestawienie kolumn nie zmienia wartości wyznacznika. Ponadto wyznacznik jest funkcjonałem liniowym ze względu na każdą (nie tylko ostatnią) kolumnę macierzy (bo wystarczy zamienić tę kolumnę z ostatnią, sprawdzić liniowość i przestawić z powrotem).
Wreszcie, wyznacznik macierzy osobliwej jest równy 0; przypuśćmy bowiem, że ostatnia kolumna macierzy jest kombinacją liniową pozostałych. Wtedy

$$\det_n\Big[\boldsymbol{a}_1, \dots, \boldsymbol{a}_{n-1}, \sum_{i=1}^{n-1} c_i\boldsymbol{a}_i\Big] = \sum_{i=1}^{n-1} c_i \det_n[\boldsymbol{a}_1, \dots, \boldsymbol{a}_{n-1}, \boldsymbol{a}_i],$$

ale jeśli macierz ma dwie kolumny identyczne, to jej wyznacznik jest równy 0 bo

$$\det_n[\boldsymbol{a}_1, \dots, \boldsymbol{a}_i, \dots, \boldsymbol{a}_{n-1}, \boldsymbol{a}_i] = -\det_n[\boldsymbol{a}_1, \dots, \boldsymbol{a}_i, \dots, \boldsymbol{a}_{n-1}, \boldsymbol{a}_i].$$

(zakładamy tu, że $\mathbb{K}$ nie jest ciałem charakterystyki 2).

Jawny wzór opisujący wyznacznik jest taki:

$$\det_n[a_{ij}]_{i,j} = \sum_{\sigma\in S_n} \operatorname{sgn}\sigma \prod_{i=1}^{n} a_{i,\sigma(i)}.$$

Suma w tym wzorze przebiega po wszystkich permutacjach σ zbioru n-elementowego, których jest $n!$. Znak permutacji, $\operatorname{sgn}\sigma$, jest równy $+1$ jeśli permutacja σ jest parzysta, albo -1 jeśli nieparzysta. Współczynniki macierzy, których iloczyn (opatrzony odpowiednim znakiem) jest składnikiem sumy, są brane po jednym z każdego wiersza (dla $i = 1, \dots, n$), a także po jednym z każdej kolumny (bo dla $i = 1, \dots, n$ wartość $\sigma(i)$ przebiega zbiór $\{1, \dots, n\}$).

Uwaga: Podany wyżej wzór przyda nam się do rozważań teoretycznych, ale nie należy na jego podstawie obliczać wyznacznika numerycznie (chyba, że macierz ma wymiary 3×3 lub mniejsze). Obliczenie takie byłoby bardzo czasochłonne i niedokładne. Poznamy dużo lepszą metodę obliczania wyznacznika.

Sprawdźmy, że funkcja określona powyższym wzorem spełnia podane warunki.
W każdym składniku występuje dokładnie 1 współczynnik z ostatniej kolumny. Jeśli w miejsce tej kolumny podstawimy dowolną kombinację liniową wektorów $\boldsymbol{x}_1, \dots, \boldsymbol{x}_k \in \mathbb{K}^n$, to w każdym składniku będziemy mieli czynnik, który jest kombinacją liniową (z tymi samymi współczynnikami) odpowiednich współczynników wektorów $\boldsymbol{x}_1, \dots, \boldsymbol{x}_k$.
Zamiana miejscami dwóch sąsiednich (a nawet dowolnych) kolumn powoduje złożenie każdej permutacji σ z transpozycją T opisującą tę zamianę. Złożenie takie ma znak przeciwny niż σ i wszystkie takie złożenia przebiegają cały zbiór S_n. Zatem zamiana kolumn powoduje zmianę na przeciwny znaku wszystkich składników.
Podstawiając macierz jednostkową I_n przekonamy się, że tylko jeden składnik sumy we wzorze jest różny od 0: ten, który odpowiada permutacji identycznościowej. Składnik ten jest iloczynem samych jedynek.
Tak więc funkcja spełniająca podane warunki istnieje.

Aby sprawdzić, że istnieje tylko jedna funkcja, która odpowiada definicji wyznacznika stopnia n, załóżmy, że macierz A jest nieosobliwa (dla osobliwej każda taka funkcja ma wartość 0, a więc funkcje takie obcięte do zbioru macierzy osobliwych są identyczne) i rozważmy ciąg przekształceń macierzy, który prowadzi do otrzymania macierzy jednostkowej. Proces ten jest ciągiem przekształceń, z których każde jest przestawianiem kolumn (znak wyznacznika się zmienia, zapamiętujemy czynnik -1), mnożeniem kolumny przez pewną liczbę $a \neq 0$ (zapamiętujemy czynnik $1/a$) i dodawaniem pewnej kolumny do innej (ta operacja nie zmienia wartości wyznacznika). Wykonując te operacje, możemy po skończenie wielu krokach otrzymać macierz jednostkową. Wyznacznik macierzy wyjściowej jest równy iloczynowi zapamiętanych czynników i nie może być inny (bo w którymś kroku nie zachowalibyśmy przyjętych własności wyznacznika).

Dalsze własności wyznaczników

Każda permutacja σ ma taki sam znak jak jej permutacja odwrotna. Ponadto, jeśli σ przebiega zbiór S_n wszystkich permutacji zbioru n-elementowego, to σ^{-1} też przebiega cały ten zbiór. Dlatego

$$\sum_{\sigma\in S_n} \operatorname{sgn}\sigma \prod_{i=1}^{n} a_{i,\sigma(i)} = \sum_{\sigma^{-1}\in S_n} \operatorname{sgn}\sigma^{-1} \prod_{j=1}^{n} a_{\sigma^{-1}(j),j},$$

skąd wynika, że $\det A = \det A^T$. Konsekwencją tego spostrzeżenia jest fakt, że wszystkie własności wyznacznika sformułowane dla kolumn obowiązują także dla wierszy.

Wyznaczniki można opisać (i czasem tak się je definiuje) za pomocą wzoru rekurencyjnego, zwanego rozwinięciem Laplace'a:

$$\begin{aligned}\det{}_1[a_{11}] &= a_{11},\\ \det{}_n[a_{ij}]_{i,j} &= \sum_{i=1}^{n}(-1)^{i+j}(\det{}_{n-1} A_{ij})a_{ij}.\end{aligned}$$

We wzorze tym symbol A_{ij} oznacza macierz $n-1\times n-1$ otrzymaną z $A=[a_{ij}]_{i,j}$ przez skreślenie i-tego wiersza i j-tej kolumny. Wyznacznik (stopnia $n-1$) tej macierzy nazywa się minorem macierzy A. We wzorze tym wyróżniona jest j-ta kolumna (j może być dowolne od 1 do n); żaden z minorów nie zależy od współczynników w tej kolumnie. Dlatego mówi się o rozwinięciu względem j-tej kolumny. Dowód, że wzór jawny i wzór rekurencyjny są równoważne, pozostawiam na ćwiczenia. Tymczasem zauważmy, że można napisać wzór rekurencyjny, który przedstawia rozwinięcie wyznacznika względem dowolnego wiersza.

Macierz A można przedstawić w postaci blokowej, $A=\begin{bmatrix} A_{11} & A_{12}\\ A_{21} & A_{22}\end{bmatrix}$, w której bloki A_{11} i A_{22} są macierzami kwadratowymi $k\times k$ oraz $l\times l$ (jest $n=l+k$). Jeśli blok A_{12} lub A_{21} ma wszystkie współczynniki równe 0, to $\det A = \det A_{11}\det A_{22}$. Istotnie, załóżmy, że blok A_{12} zawiera same zera. Rozważmy składnik sumy w jawnym wzorze na wyznacznik, w którym jednym z czynników jest

współczynnik a_{ij} z bloku A_{21}. Pozostałe czynniki pochodzą z wierszy o numerach innych niż i i kolumn innych niż j-ta.
Musimy wybrać jeszcze $k+l-1$ czynników. Wybierając współczynnik z bloku A_{11} lub A_{21} musimy odrzucić odpowiednią kolumnę każdego z tych bloków (a zatem wybierzemy najwyżej $k-1$ takich współczynników), a także odpowiedni wiersz całej macierzy. Odrzuciliśmy zatem wszystkie k kolumn bloków A_{11} i A_{21}, oraz k wierszy bloków A_{12} i A_{22}, w tym co najmniej 1 wiersz bloku A_{22}. Pozostałe czynniki trzeba znaleźć w macierzy kwadratowej utworzonej z pozostałych l wierszy bloków A_{12} i A_{22}. Ale co najmniej jeden z tych wierszy należy do bloku A_{12}, czyli zawiera same zera. Tak więc każdy składnik, który zawiera czynnik z bloku A_{21} zawiera również czynnik z bloku A_{12}, równy 0. Wszystkie te składniki są więc równe 0. Sumę pozostałych składników otrzymamy mnożąc sumy opisujące wyznaczniki bloków A_{11} i A_{22}.

Rozumowanie to możemy powtórzyć zakładając, że wszystkie współczynniki w bloku A_{21} są równe 0. Wnioski z obu tych rozważań są takie: jeśli macierz A jest blokowo-trójkątna (górna lub dolna), przy czym wszystkie bloki diagonalne są macierzami kwadratowymi, to wyznacznik macierzy A jest równy iloczynowi wyznaczników tych bloków. W szczególności, jeśli macierz A jest trójkątna, to jej wyznacznik jest iloczynem współczynników na diagonali. W szczególności stwierdzenia te dotyczą macierzy blokowo-diagonalnych i diagonalnych.

Twierdzenie Cauchy'ego: *Jeśli* $A, B\in\mathbb{K}^{n,n}$*, to* $\det(AB)=\det A\det B$.

Dowód: Rozważmy macierz C o wymiarach $2n\times 2n$, o następującej strukturze blokowej:

$$C=\begin{bmatrix} A & 0\\ -I_n & B\end{bmatrix}.$$

Na podstawie wcześniejszego stwierdzenia $\det C=\det A\det B$. Obliczmy iloczyn macierzy C i macierzy

$$D=\begin{bmatrix} I_n & B\\ 0 & I_n\end{bmatrix}.$$

Macierz $C'=CD$ ma następującą postać:

$$C'=\begin{bmatrix} A & AB\\ -I_n & 0\end{bmatrix}.$$

Jeśli kolumny macierzy C oznaczymy symbolami $\mathbf{c}_1,\dots,\mathbf{c}_{2n}$, to wykonane mnożenie jest równoważne dodaniu do $n+j$-tej kolumny macierzy C wektora

$\sum_{k=1}^{n} c_k b_{kj}$, dla $j = 1, \ldots, n$, a zatem wyznacznik macierzy C' jest równy $\det C$. Możemy jeszcze zamienić kolumnę j-tą z $n+$j-tą dla $j = 1, \ldots, n$, co spowoduje zmianę znaku wyznacznika jeśli n jest nieparzyste, a potem pomnożyć ostatnie n kolumn przez -1, (co n-krotnie zmieni znak na przeciwny). W wyniku tego postępowania otrzymamy macierz

$$C'' = \begin{bmatrix} AB & -A \\ 0 & I_n \end{bmatrix}.$$

Mamy $\det C'' = \det(AB)$, ale także $\det C'' = \det C$, co kończy dowód. □

Wniosek: *Jeśli macierz* A *jest nieosobliwa, to* $\det A \neq 0$.
Dowód: $1 = \det I_n = \det A \det A^{-1}$. □

Wyznaczniki, baza dualna i wzory Cramera

Rozważmy układ n równań liniowych z n niewiadomymi, $Ax = \mathbf{b}$, z nieosobliwą macierzą A, który zapiszemy w postaci

$$\mathbf{a}_1 x_1 + \cdots + \mathbf{a}_n x_n = \mathbf{b}.$$

Kolumny $\mathbf{a}_1, \ldots, \mathbf{a}_n$ macierzy A są liniowo niezależne, zatem stanowią bazę przestrzeni $\mathbb{K}^n$ i istnieje jej baza dualna $\varphi_1, \ldots, \varphi_n$. Znając funkcjonały wchodzące w skład tej bazy możemy obliczyć każdą z niewiadomych w taki sposób:

$$\varphi_i(\mathbf{b}) = \varphi_i(\mathbf{a}_1 x_1 + \cdots + \mathbf{a}_n x_n) = \varphi_i(\mathbf{a}_1)x_1 + \cdots + \varphi_i(\mathbf{a}_n)x_n = x_i.$$

Zobaczymy sposób znalezienia tych funkcjonałów. Rozważmy wyrażenie

$$\beta_i(\mathbf{y}) = \det[\mathbf{a}_1, \ldots, \mathbf{a}_{i-1}, \mathbf{y}, \mathbf{a}_{i+1}, \ldots, \mathbf{a}_n].$$

Definiuje ono pewien funkcjonał liniowy β_i, którego argumentem może być dowolny wektor $\mathbf{y} \in \mathbb{K}^n$. Dla $i \neq j$ mamy $\beta_i(\mathbf{a}_j) = 0$, natomiast $\beta_i(\mathbf{a}_i) = \det A \neq 0$. Dlatego funkcjonał φ_i, który jest i-tym elementem bazy dualnej do $\mathbf{a}_1, \ldots, \mathbf{a}_n$, jest równy

$$\varphi_i = \frac{\beta_i}{\beta_i(\mathbf{a}_i)}.$$

Stąd mamy dla $i = 1, \ldots, n$ wzory Cramera:

$$x_i = \varphi_i(\mathbf{b}) = \frac{\beta_i(\mathbf{b})}{\beta_i(\mathbf{a}_i)} = \frac{\det[\mathbf{a}_1, \ldots, \mathbf{a}_{i-1}, \mathbf{b}, \mathbf{a}_{i+1}, \ldots, \mathbf{a}_n]}{\det[\mathbf{a}_1, \ldots, \mathbf{a}_n]}.$$

Wzory Cramera w szkole bywają przedstawiane jako praktyczny sposób rozwiązywania układów równań liniowych, w związku z czym konieczny jest pewien komentarz. W pewnych sytuacjach rzeczywiście można ich użyć, głównie wtedy, gdy liczba równań i niewiadomych jest mała (np. 2 lub 3), albo gdy informacja o zadaniu umożliwia szczególnie łatwe obliczanie odpowiednich wyznaczników. Jednak już nawet dla $n = 2$ algorytm oparty na wzorach Cramera nie jest numerycznie poprawny (tylko stabilny), a zatem jest gorszy niż eliminacja Gaussa. Ponadto obliczanie wyznacznika zwykle jest kosztowne; stosunkowo tania i dokładna metoda polega na dokonaniu rozkładu macierzy A na czynniki trójkątne (metodą eliminacji Gaussa), a następnie obliczenie iloczynu wyznaczników tych czynników (tj. iloczynu ich współczynników na diagonalach). W świetle powyższych uwag stosowanie wzorów Cramera w obliczeniach numerycznych jest zwykle bezsensowne.

Wyznaczniki i odwrotność macierzy

Def. Niech $A = [a_{ij}]_{i,j}$ oznacza macierz $n \times n$. Dopełnieniem algebraicznym współczynnika a_{ij} macierzy A nazywamy liczbę $(-1)^{i+j} \det A_{ij}$, gdzie macierz A_{ij} powstaje przez skreślenie w macierzy A i-tego wiersza i j-tej kolumny.

Dopełnienie algebraiczne jest więc minorem macierzy, opatrzonym odpowiednim znakiem. Udowodnimy, że jeśli macierz A jest nieosobliwa, to macierz $A^{-1} = [b_{ij}]_{i,j}$ ma współczynniki

$$b_{ij} = \frac{(-1)^{i+j} \det A_{ji}}{\det A},$$

a zatem jest to transponowana macierz dopełnień algebraicznych poszczególnych współczynników macierzy A, pomnożona przez $\frac{1}{\det A}$. Istotnie, dla ustalonego $j \in \{1, \ldots, n\}$ niech $\mathbf{y}_j^H = [(-1)^{1+j} \det A_{1j}, \ldots, (-1)^{n+j} \det A_{nj}]$. Jeśli $\mathbf{a}_k$ oznacza k-tą kolumnę macierzy A, to liczba $\mathbf{y}_j^H \mathbf{a}_k$ jest równa $\det A$, jeśli $j = k$ (wyrażenie $\mathbf{y}_j^H \mathbf{a}_k$ opisuje rozwinięcie Laplace'a wyznacznika macierzy A względem j-tej kolumny), albo 0 jeśli $j \neq k$. Stąd wynika równość

$$\begin{bmatrix} \mathbf{y}_1^H \\ \vdots \\ \mathbf{y}_n^H \end{bmatrix} [\mathbf{a}_1, \ldots, \mathbf{a}_n] = \det A \cdot I_n$$

i na tym koniec dowodu. □

Uwaga: Macierz wierszowa $\frac{1}{\det A}\mathbf{y}_j^H$ użyta w powyższym dowodzie jest funkcjonałem φ_j, należącym do bazy przestrzeni $\mathbb{K}^{1,n} = (\mathbb{K}^n)^*$ dualnej do bazy $\mathbf{a}_1, \ldots, \mathbf{a}_n$.

Wyznacznik przekształcenia liniowego

Niech $f\colon V \to V$ będzie przekształceniem liniowym przestrzeni V, której wymiar jest skończony. Ustalmy dowolną bazę $x_1, \dots, x_n$, z której utworzymy macierz $X = [x_1, \dots, x_n]$. Wtedy wyrażenie $A = X^{-1} f X$ opisuje macierz A tego przekształcenia w ustalonej bazie (macierz X^{-1} reprezentuje bazę dualną do $x_1, \dots, x_n$). Udowodnimy, że wyznacznik macierzy przekształcenia nie zależy od wyboru bazy.

Istotnie, niech macierz $Y = [y_1, \dots, y_n]$ reprezentuje dowolną inną bazę przestrzeni V. Macierz przejścia od bazy Y do X jest równa $B = X^{-1}Y$. Zatem, macierz przekształcenia f w bazie Y jest równa $C = B^{-1}AB$ (mówimy, że macierze A i C są podobne). Na podstawie twierdzenia Cauchy'ego $\det C = \det B^{-1} \det A \det B = \frac{1}{\det B} \det A \det B = \det A$, co kończy dowód. □

Ponieważ wyznacznik każdej macierzy reprezentującej ustalone przekształcenie jest taki sam (nie zależy od wyboru bazy), więc możemy zdefiniować pojęcie wyznacznika przekształcenia liniowego. Dla ustalonego przekształcenia liniowego $f\colon V \to V$ jest to liczba, równa wyznacznikowi macierzy reprezentującej to przekształcenie w dowolnej bazie.

Orientacja bazy i układu współrzędnych

Niech $x_1, \dots, x_n$ oraz $y_1, \dots, y_n$ będą bazami pewnej rzeczywistej przestrzeni liniowej V (dowolnej, niekoniecznie $\mathbb{R}^n$). Bazy te mogą służyć do określenia różnych układów współrzędnych, o czym była mowa wcześniej. Istnieje macierz C przejścia od jednej bazy do drugiej. Macierz ta jest nieosobliwa, a więc jej wyznacznik jest liczbą rzeczywistą dodatnią albo ujemną. Możemy między bazami wprowadzić relację *zgodności orientacji*: dwie bazy są zgodnie zorientowane jeśli wyznacznik macierzy przejścia między nimi jest dodatni i przeciwnie zorientowane, jeśli jest ujemny. Oczywiście, orientacja bazy jest związana z kolejnością jej elementów; zmiana tej kolejności może zmienić orientację bazy.

Łatwo jest sprawdzić (na podstawie twierdzenia Cauchy'ego), że określona wyżej relacja jest równoważnością, która ma dwie klasy abstrakcji. Jest sprawą arbitralnej umowy, na podstawie której układy współrzędnych określone przez bazy z jednej z tych klas nazwiemy prawoskrętnymi i wtedy układy współrzędnych określone przez bazy z drugiej klasy będą lewoskrętne. Co ciekawe, nie mając wzorca (czyli pewnej bazy w rozpatrywanej przez nas przestrzeni, o której *ktoś powiedział*, że określa ona układ prawoskrętny), nie mamy żadnej

szansy na to, aby skonstruować układ np. prawoskrętny na podstawie informacji, która takiego wzorca nie definiuje.

Iloczyny wektorowe

Def. Iloczynem wektorowym w przestrzeni $\mathbb{K}^n$ nazywamy przekształcenie $f\colon \underbrace{\mathbb{K}^n \times \cdots \times \mathbb{K}^n}_{n-1} \to (\mathbb{K}^n)^*$ określone wzorem

$$\forall_{y \in \mathbb{K}^n}\ f(x_1, \dots, x_{n-1})(y) = \det[x_1, \dots, x_{n-1}, y].$$

Tak więc iloczynem wektorowym ustalonych $n-1$ wektorów $x_1, \dots, x_{n-1} \in \mathbb{K}^n$ jest pewien funkcjonał liniowy w przestrzeni $\mathbb{K}^n$. Możemy go utożsamić z pewnym wektorem $n \in \mathbb{K}^n$, takim że dla każdego $y \in \mathbb{K}^n$ jest $\det[x_1, \dots, x_{n-1}, y] = n^H y$. W związku z tym mówimy, że iloczyn wektorowy wektorów jest wektorem. Kolejne współczynniki wektora n możemy otrzymać następująco: tworzymy macierz kwadratową $A = [x_1, \dots, x_{n-1}, y]$ (dla dowolnego wektora $y \in \mathbb{K}^n$), a następnie obliczamy dopełnienia algebraiczne kolejnych współczynników w kolumnie y (i w przypadku zespolonym bierzemy liczbę sprzężoną).

Uwaga: Dla $n > 3$ to nie jest praktyczny algorytm obliczania iloczynu wektorowego, tylko fakt, na którym opierając się można różne praktyczne algorytmy znaleźć.

Iloczyn wektorowy wektorów $x_1, \dots, x_{n-1}$ oznaczamy symbolem

$$x_1 \wedge \cdots \wedge x_{n-1}, \quad \text{albo} \quad x_1 \times \cdots \times x_{n-1}$$

(drugie oznaczenie może mylić się z iloczynem kartezjańskim zbiorów, więc osobiście wolę to pierwsze).

Większość własności iloczynu wektorowego w łatwy sposób wynika z własności wyznaczników. Na przykład, iloczyn wektorowy jest przekształceniem liniowym ze względu na każdy argument. Przestawienie dwóch argumentów jest równoważne pomnożeniu iloczynu przez -1. Jeśli wektory $x_1, \dots, x_{n-1}$ są liniowo zależne, to ich iloczyn wektorowy jest funkcjonałem zerowym.

Powyższa definicja dla dwóch wektorów w $\mathbb{R}^3$ jest równoważna definicji iloczynu wektorowego znanej ze szkoły. Ale w szkolna definicja wymaga pojęcia kąta między wektorami, do określenia którego jest potrzebne pojęcie iloczynu skalarnego — poznamy je później i dopiero wtedy będziemy mogli udowodnić tę równoważność.

Zadania i problemy

1. Udowodnij, że wzór jawny na wyznacznik i wzór rekurencyjny (rozwinięcie Laplace'a) są równoważne. Zacznij od udowodnienia indukcyjnego równoważności wzoru jawnego i rozwinięcia względem ostatniej kolumny, a następnie udowodnij (rozpatrując odpowiednie przestawienia kolumn), że ten sam wynik daje rozwinięcie względem dowolnej kolumny.
2. Udowodnij reguły Sarrusa, tj. znane ze szkoły schematy pozwalające obliczać wyznaczniki macierzy 2×2 i 3×3.
3. Udowodnij, że macierz antysymetryczna (tj. taka, że $A^T = -A$) o wymiarach $n \times n$ jest osobliwa, jeśli n jest nieparzyste.

   Macierz $A \in \mathbb{C}^{n,n}$ jest antyhermitowska, jeśli $A^H = -A$. Co można powiedzieć o wyznaczniku takiej macierzy w zależności od jej wymiarów?
4. Metoda eliminacji Gaussa, przy użyciu której można rozwiązywać układy równań liniowych, znajduje macierze L, R, P i Q, takie że $LR = PAQ^T$, przy czym macierze P i Q to macierze permutacji, macierz L jest trójkątna dolna, a macierz R jest trójkątna górna. Jak można obliczyć wyznacznik macierzy A, znając macierze L i R oraz łączną liczbę przestawień kolumn i wierszy wykonanych podczas eliminacji (przestawienia te są reprezentowane odpowiednio przez macierze P i Q)? Odpowiedź uzasadnij, powołując się na stosowne twierdzenia.
5. Def. Macierz cykliczna dwudiagonalna jest to macierz $A = [a_{ij}]_{i,j}$ o wymiarach $n \times n$, której wszystkie współczynniki z wyjątkiem a_{ii} dla $i = 1, \ldots, n$ oraz $a_{i,i+1}$ dla $i = 1, \ldots, n-1$ i a_{n1} są równe 0.

   Znajdź wzór opisujący wyznacznik takiej macierzy i warunek (równanie, w którym występują współczynniki) konieczny i dostateczny nieosobliwości takiej macierzy.
6. Oblicz współczynniki macierzy odwrotnych do

$$A = \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix} \quad \text{oraz} \quad B = \begin{bmatrix} 2 & 1 & 0 \\ 1 & 2 & 1 \\ 0 & 1 & 2 \end{bmatrix}$$

   na podstawie jawnego wzoru.
7. Oblicz iloczyn wektorowy wektorów

$$\boldsymbol{x} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} \quad \text{i} \quad \boldsymbol{y} = \begin{bmatrix} 3 \\ 2 \\ 1 \end{bmatrix}.$$

8. Dla ustalonego wektora $\boldsymbol{x} \in \mathbb{R}^3$ przekształcenie $f\colon \mathbb{R}^3 \to \mathbb{R}^3$, dane wzorem $f(\boldsymbol{y}) = \boldsymbol{x} \wedge \boldsymbol{y}$, jest liniowe. Znajdź macierz jego przekształcenia.

9. Zbadaj, czy zbiór $\mathbb{R}^3$ z działaniem dwuargumentowym — iloczynem wektorowym — jest grupą (w tym i w następnym zadaniu utożsamiamy przestrzenie $\mathbb{R}^3$ i $\mathbb{R}^{1,3}$). Jeśli tak, to czy to jest grupa abelowa? Jeśli nie, to które własności działania w grupie nie są spełnione?
10. Oblicz wyznacznik przekształcenia liniowego $f\colon \mathbb{R}[x]_4 \to \mathbb{R}[x]_4$, danego wzorem $f(w)(x) = (x+1)w'(x) + w(x)$.
11. Znajdź ogólny wzór na wyznacznik macierzy

$$A_n = \begin{bmatrix} 2 & 1 & & & \\ 1 & 2 & 1 & & \\ & \ddots & \ddots & \ddots & \\ & & 1 & 2 & 1 \\ & & & 1 & 2 \end{bmatrix}.$$

12. Def. Macierz Vandermonde'a jest to macierz $n+1 \times n+1$ o postaci

$$\begin{bmatrix} 1 & x_0 & x_0^2 & \ldots & x_0^n \\ 1 & x_1 & x_1^2 & \ldots & x_1^n \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & x_n & x_n^2 & \ldots & x_n^n \end{bmatrix},$$

   dla ustalonych liczb $x_0, \ldots, x_n$. Udowodnij, że wyznacznik takiej macierzy jest równy

$$\prod_{0 \leqslant i < j \leqslant n} (x_j - x_i)$$

   (wynika stąd wniosek, że taka macierz jest osobliwa wtedy i tylko wtedy, gdy nie wszystkie liczby x_i są parami różne).

   Pierwszy sposób: Zastosuj jeden krok eliminacji Gaussa, a następnie przekształć blok $n \times n$ w prawym dolnym rogu (wykonując pewne działania na wierszach i kolumnach) tak, aby otrzymać macierz Vandermonde'a $n \times n$. Następnie użyj indukcji.

   Drugi sposób: Napisz macierz Vandermonde'a, zamieniając *liczbę* x_n na *zmienną* x. Wyznacznik tej macierzy jest wielomianem zmiennej x stopnia co najwyżej n, a jego miejscami zerowymi są liczby $x_0, \ldots, x_{n-1}$. Zatem wielomian ten można przedstawić w postaci

$$\varepsilon(x - x_0) \cdot \ldots \cdot (x - x_{n-1}).$$

   Liczba ε jest współczynnikiem przy x^n, a więc jest to również wyznacznik macierzy ... . Następnie użyj indukcji.

   Który sposób bardziej Ci się podoba? (odpowiedź na to pytanie nie jest punktowana).

Formy kwadratowe, dwuliniowe i półtoraliniowe

Def. Niech $V_1, \ldots, V_n$ oraz V będą przestrzeniami liniowymi nad ustalonym ciałem $\mathbb{K}$. Funkcję $f\colon V_1 \times \ldots \times V_n \to V$ nazywamy przekształceniem n-liniowym, jeśli dla każdego $k \in \{1, \ldots, n\}$, przy ustalonych wszystkich argumentach z wyjątkiem k-tego, funkcja ta jest przekształceniem liniowym przestrzeni V_k w V.

Na przykład, jeśli utożsamimy przestrzeń $\mathbb{K}^{n,n}$ z $(\mathbb{K}^n)^n$, to funkcja $\det_n$ (wyznacznik) jest przekształceniem n-liniowym (bo jest to przekształcenie liniowe ze względu na każdą kolumnę macierzy, przy ustalonych wszystkich pozostałych).

Odtąd przez najbliższe kilkanaście wykładów będziemy się zajmować przekształceniami dwuliniowymi i symetrycznymi, $\varphi\colon V \times V \to \mathbb{K}$. Przekształcenie dwuliniowe φ jest symetryczne, jeśli dla dowolnych wektorów $x, y \in V$ zachodzi równość $\varphi(x, y) = \varphi(y, x)$. Przekształcenia takie będziemy nazywać formami dwuliniowymi.

Def. Niech V oznacza przestrzeń liniową nad ciałem $\mathbb{K}$. Formą kwadratową w przestrzeni V nazywamy dowolną funkcję $\Phi\colon V \to \mathbb{K}$, określoną wzorem $\Phi(x) = \varphi(x, x)$, za pomocą pewnej formy dwuliniowej $\varphi\colon V \times V \to \mathbb{K}$.

Określone wyżej formy kwadratowe spełniają warunek $\Phi(ax) = a^2\Phi(x)$ dla każdego $x \in V$ oraz $a \in \mathbb{K}$. Własność funkcji Φ: $\Phi(ax) = h(a)\Phi(x)$ dla każdego x i a, gdzie h jest pewną ustaloną funkcją, nazywa się jednorodnością przekształcenia Φ. Przekształcenia liniowe są jednorodne, z funkcją $h(a) = a$. W poprzednim wykładzie była mowa o współrzędnych jednorodnych. W związku z tym zauważmy, że przekształcenie afiniczne nie jest jednorodne, ale reprezentowaliśmy je przez pewne przekształcenie liniowe; stąd nazwa „współrzędne jednorodne". Można uogólnić to postępowanie, np. na formy kwadratowe w przestrzeniach afinicznych — będzie jeszcze o tym mowa.

Obliczmy, korzystając z dwuliniowości i symetrii

$$\varphi(x + y, x + y) = \varphi(x, x) + 2\varphi(x, y) + \varphi(y, y)$$

skąd wynika

$$\varphi(x, y) = \frac{1}{2}\big(\Phi(x + y) - \Phi(x) - \Phi(y)\big).$$

Podstawiając do powyższego wzoru dowolną formę kwadratową Φ otrzymamy pewną jednoznacznie określoną formę dwuliniową φ. Jednoznaczność związku

symetrycznych przekształceń dwuliniowych i form kwadratowych nosi nazwę zasady polaryzacji, a właściwie jest szczególnym przypadkiem bardzo ogólnego twierdzenia o tej nazwie, dotyczącego symetrycznych przekształceń n-liniowych dla dowolnego n.

Jeśli $V = \mathbb{K}^n$, to każdą formę dwuliniową φ oraz formę kwadratową Φ możemy przedstawić wzorami

$$\varphi(x, y) = y^T A x, \qquad \Phi(x) = x^T A x,$$

dla pewnej jednoznacznie określonej macierzy symetrycznej A. Mając dowolną bazę $X = [x_1, \ldots, x_n]$ ustalonej przestrzeni liniowej V możemy odpowiednie formy w tej przestrzeni przedstawić w postaci

$$\varphi(x, y) = (X^{-1}y)^T A X^{-1}x, \qquad \Phi(x) = (X^{-1}x)^T A X^{-1}x,$$

za pomocą macierzy bazy dualnej X^{-1} oraz pewnej macierzy symetrycznej $A \in \mathbb{K}^{n,n}$. Zauważmy, że formę kwadratową otrzymamy biorąc niekoniecznie symetryczną macierz A; symetria macierzy jest konieczna do zapewnienia symetrii formy dwuliniowej. Założenie o symetrii zapewnia *jednoznaczność* związku form dwuliniowych i kwadratowych.

W przestrzeni nad ciałem liczb zespolonych możemy wykonać wszystkie konstrukcje opisane wyżej bez problemu. Chcielibyśmy jednak, aby wartości form kwadratowych w przestrzeni nad dowolnym ciałem liczbowym (tj. $\mathbb{Q}$, $\mathbb{R}$ i $\mathbb{C}$) były liczbami rzeczywistymi (a dokładniej, aby część urojona wartości formy kwadratowej w przestrzeni zespolonej była równa 0). Dlatego musimy nieco zmienić warunek jednorodności. Trzeba, aby dla każdego wektora $x \in V$ i liczby a była spełniona równość $\Phi(ax) = |a|^2\Phi(x)$. Określamy w ten sposób formy kwadratowe drugiego rodzaju; w przestrzeniach nad ciałem liczb rzeczywistych lub wymiernych dostajemy to samo, co przedtem.

Dowolną formę kwadratową (dalej dla skrótu pomijamy słowa „drugiego rodzaju", ale tylko takimi formami będziemy się zajmować) w przestrzeni $\mathbb{C}^n$ możemy przedstawić wzorem

$$\Phi(x) = x^H A x,$$

w którym występuje macierz hermitowska A (tj. $A^H = A$; sprawdzenie, że dla każdego wektora x część urojona wartości formy jest równa 0, jest prostym ćwiczeniem). Dla ustalonej bazy $X = [x_1, \ldots, x_n]$, dowolną formę kwadratową Φ w przestrzeni V możemy przedstawić wzorem

$$\Phi(x) = (X^{-1}x)^H A X^{-1}x,$$

z macierzą hermitowską A. Z tak określoną formą kwadratową możemy związać przekształcenie φ, określone wzorem

$$\varphi(x, y) = (X^{-1}y)^H A X^{-1} x.$$

Natychmiast możemy zauważyć, że przekształcenie to, z wyjątkiem przypadku gdy macierz A jest zerowa, *nie jest* ani symetryczne, ani dwuliniowe!

Przekształcenie $\varphi\colon \mathbb{C}^n \times \mathbb{C}^n \to \mathbb{C}$, określone wzorem

$$\varphi(x, y) = y^H A x,$$

który jest szczególnym, najłatwiejszym do zbadania przypadkiem wzoru poprzedniego, spełnia następujące warunki dla dowolnych wektorów x, y, x_1, x_2, y_1, y_2 i liczb a_1, a_2:

$$\begin{aligned}\varphi(a_1x_1 + a_2x_2, y) &= a_1\varphi(x_1, y) + a_2\varphi(x_2, y),\\ \varphi(x, a_1y_1 + a_2y_2) &= \overline{a_1}\varphi(x, y_1) + \overline{a_2}\varphi(x, y_2),\\ \varphi(x, y) &= \overline{\varphi(y, x)}.\end{aligned}$$

Przekształcenie spełniające powyższe warunki jest nazywane <u>formą półtoraliniową</u>. Przekształcenia dowolnej przestrzeni, opisane ogólniejszym wzorem podanym wcześniej, też spełniają te warunki.

Między formami półtoraliniowymi a formami kwadratowymi drugiego rodzaju również zachodzi jednoznaczna odpowiedniość. Sposób obliczania wartości formy kwadratowej za pomocą półtoraliniowej jest oczywisty, zaś w drugą stronę mamy

$$\begin{aligned}\varphi(x + y, x + y) &= \varphi(x, x) + 2\,\mathrm{Re}\,\varphi(x, y) + \varphi(y, y),\\ \varphi(x + iy, x + iy) &= \varphi(x, x) + \bar{\imath}\varphi(x, y) + i\varphi(y, x) + i\bar{\imath}\varphi(y, y) =\\ &\qquad \varphi(x, x) + 2\,\mathrm{Im}\,\varphi(x, y) + \varphi(y, y),\end{aligned}$$

skąd wynika następująca <u>tożsamość polaryzacyjna</u>:

$$\varphi(x, y) = \frac{1}{2}\big(\Phi(x + y) - \Phi(x) - \Phi(y)\big) + \frac{i}{2}\big(\Phi(x + iy) - \Phi(x) - \Phi(y)\big).$$

Znając formę kwadratową umiemy więc obliczyć odpowiadającą jej formę półtoraliniową, a z jej pomocą możemy obliczyć współczynniki macierzy $A = [a_{pq}]_{p,q}$, która reprezentuje obie te formy w ustalonej bazie $X = [x_1, \ldots, x_n]$:

$$a_{pq} = \varphi(x_q, x_p).$$

<u>Dygresja.</u> Formy kwadratowe i dwuliniowe lub półtoraliniowe określa się także w przestrzeniach nieskończenie wymiarowych. Niech np. $a, b \in \mathbb{R}$, $a < b$ i niech V oznacza zbiór funkcji rzeczywistych lub zespolonych na odcinku $[a, b]$, zaś ρ oznacza ustaloną nieujemną funkcję rzeczywistą na $[a, b]$. Możemy wtedy określić następujące formy:

$$\begin{aligned}\Phi(f) &= \int_a^b |f(x)|^2\rho(x)\,dx,\\ \varphi(f, g) &= \int_a^b f(x)\overline{g(x)}\rho(x)\,dx,\end{aligned}$$

przy czym bierzemy pod uwagę tylko te funkcje, dla których powyższe całki są określone i skończone. O funkcjach takich mówi się, że są <u>całkowalne w kwadracie</u> z wagą ρ na odcinku $[a, b]$; ich zbiór jest przestrzenią liniową (podprzestrzenią przestrzeni V). Zauważmy, że przyjęta forma kwadratowa jest istotnym elementem użytym do *zdefiniowania* przestrzeni liniowej[2]. Koniec dygresji.

Kongruencje macierzy

Jeśli macierz hermitowska A reprezentuje formę półtoraliniową φ w bazie $X = [x_1, \ldots, x_n]$, zaś $C = X^{-1}Y$, gdzie $Y = [y_1, \ldots, y_n]$ jest macierzą innej bazy, to macierz $B = C^H AC$ reprezentuje formę φ w bazie Y. Łatwo jest sprawdzić, że macierz B jest hermitowska. Ponieważ każda nieosobliwa macierz C jest macierzą zmiany bazy, więc każda macierz równa $C^H AC$ reprezentuje tę samą formę φ w pewnej bazie.

<u>Def.</u> Przekształcenie, które macierzy hermitowskiej A przyporządkowuje macierz $C^H AC$ (dla pewnej nieosobliwej macierzy C) nazywamy <u>przystawaniem</u> albo <u>kongruencją macierzy</u>.

Kongruencje stanowią grupę (nieabelową) z działaniem złożenia. Działanie to jest realizowane przez mnożenie macierzy kongruencji. Istotnie,

$$B = C^H AC, \quad D = E^H BE = E^H C^H ACE = (CE)^H A(CE).$$

Łączność złożenia kongruencji wynika z łączności mnożenia macierzy, elementem neutralnym w grupie kongruencji jest przekształcenie tożsamościowe (reprezentowane przez macierz jednostkową), a kongruencja odwrotna do realizowanej przez macierz C jest realizowana przez macierz C^{-1}.

[2] Jeśli funkcja ρ jest dodatnia, to wzór $\|f\| = \sqrt{\Phi(f)}$ definiuje pewną normę w przestrzeni funkcji określonych na odcinku $[a, b]$, zobacz też stronę 7.3.

W grupie kongruencji można wskazać różne podgrupy, realizowane przez macierze należące do odpowiednich podgrup grupy macierzy nieosobliwych $n \times n$ z działaniem mnożenia. Mamy np. kongruencje realizowane przez macierze permutacji, kongruencje trójkątne górne, kongruencje trójkątne dolne i kongruencje diagonalne — ta ostatnia podgrupa jest abelowa.

Mając dowolną grupę przekształceń warto znać jej niezmienniki. Oczywistym niezmiennikiem kongruencji jest rząd macierzy poddawanej przekształceniu. Inne niezmienniki poznamy dalej.

Przystawanie macierzy jest relacją w zbiorze macierzy hermitowskich $n \times n$: dwie macierze, A i B, są przystające, jeśli istnieje macierz nieosobliwa C, taka że $B = C^H AC$. Relacja przystawania jest równoważnością w zbiorze macierzy hermitowskich, a zatem dzieli ten zbiór na klasy abstrakcji. Przekonamy się, że klas tych jest skończenie wiele.

Uwaga: Mieliśmy już do czynienia z inną relacją równoważności w zbiorze macierzy $n \times n$, mianowicie z podobieństwem (s. 6.10). Przypomnijmy, że macierze A i B są podobne, jeśli istnieje macierz C taka że $B = C^{-1}AC$. Istnieją macierze C, takie że $C^{-1} = C^H$, a więc w tym przypadku (i tylko w tym) podobieństwo jest jednocześnie przystawaniem. Zwróćmy uwagę, że macierz podobieństwa C jest określona z dokładnością do czynnika stałego, tj. jeśli $B = C^{-1}AC$, to również $B = (aC)^{-1}A(aC)$ dla każdego $a \neq 0$. Dla przystawania tak nie jest.

Podsumowując: macierze A i B są podobne, jeśli reprezentują to samo przekształcenie liniowe $f\colon V \to V$ w dwóch różnych bazach. Macierze hermitowskie A i B są przystające, jeśli reprezentują tę samą formę półtoraliniową $\varphi\colon V \times V \to \mathbb{K}$ w dwóch różnych bazach. Proszę tych pojęć nie mylić.

Twierdzenie Sylvestra

Twierdzenie Sylvestra (macierzowe): *Dla każdej macierzy hermitowskiej A istnieje macierz nieosobliwa C, taka że macierz $C^H AC$ ma następującą strukturę blokowo-diagonalną:*

$$D = \begin{bmatrix} I_\pi & & \\ & -I_\nu & \\ & & 0_\zeta \end{bmatrix}.$$

Macierz C może nie być określona jednoznacznie, ale wymiary poszczególnych bloków, π, ν i ζ, są przez macierz A określone jednoznacznie.

Dowód: Przeprowadzony niżej dowód istnienia odpowiedniej macierzy C polega na wskazaniu sposobu jej znalezienia. Kongruencja realizowana przez macierz C jest złożeniem kongruencji, które kolejno

1. doprowadzają macierz A do postaci blokowo-diagonalnej z blokami 1×1 lub 2×2 na diagonali,
2. przekształcają bloki diagonalne 2×2 w macierze diagonalne,
3. doprowadzają do otrzymania macierzy diagonalnej o postaci podanej w twierdzeniu.

Lemat: *Dla macierzy hermitowskiej A o podziale blokowym*

$$A = \begin{bmatrix} M & G \\ G^H & W \end{bmatrix},$$

z blokami M i W o wymiarach odpowiednio $s \times s$ i $n-s \times n-s$, takiej że blok M jest macierzą nieosobliwą, kongruencja realizowana przez macierz

$$R = \begin{bmatrix} I_s & -M^{-1}G \\ 0 & I_{n-s} \end{bmatrix}$$

doprowadza macierz A do postaci blokowo-diagonalnej.

Dowód lematu: Wystarczy przeprowadzić odpowiednie rachunki, tj. obliczyć bloki macierzy $R^H AR$. Skorzystamy ze schematu

$$\begin{array}{ccccc} & & A & & R \\ & & \downarrow & & \downarrow \\ R^H & \to & R^H A & \to & R^H AR \end{array}$$

$$\begin{array}{ccc} & \begin{bmatrix} M & G \\ G^H & W \end{bmatrix} & \begin{bmatrix} I_s & -M^{-1}G \\ 0 & I_{n-s} \end{bmatrix} \\ \begin{bmatrix} I_s & 0 \\ -G^H M^{-1} & I_{n-s} \end{bmatrix} & \begin{bmatrix} M & G \\ 0 & -G^H M^{-1}G + W \end{bmatrix} & \begin{bmatrix} M & 0 \\ 0 & W - G^H M^{-1}G \end{bmatrix} \end{array}$$

□

Dowód twierdzenia:
Krok 1. Dowód istnienia kongruencji, której efektem jest otrzymanie macierzy blokowo-diagonalnej jest indukcyjny ze względu na wymiary macierzy. Dla macierzy 1×1 i 2×2 wystarczy kongruencja tożsamościowa, przypuśćmy więc, że istnieje odpowiednia kongruencja dla wszystkich macierzy o wymiarach mniejszych niż $n \times n$. Jeśli macierz A jest zerowa, to jest diagonalna, a więc nie trzeba jej przekształcać. Jeśli macierz A jest niezerowa, to mamy dwie możliwości:

1. Pewien współczynnik na diagonali, a_{pp}, jest różny od zera. Macierz $T_{1p}^H A T_{1p}$ ma „w górnym lewym rogu" nieosobliwy blok $M = [a_{pp}]$ o wymiarach 1×1 (Uwaga: Macierz transpozycji T_{pq} jest jednostkowa, jeśli $p = q$).
2. Wszystkie współczynniki na diagonali macierzy A są równe 0, ale istnieje niezerowy współczynnik a_{pq} poza diagonalą. Macierz $(T_{1p}T_{2q})^H A (T_{1p}T_{2q})$ ma „w górnym lewym rogu" blok M o wymiarach 2×2, o postaci $\begin{bmatrix} 0 & c \\ \bar{c} & 0 \end{bmatrix}$, nieosobliwy, bo $|c| = |a_{pq}| \neq 0$.

   Po odpowiednim przestawieniu wierszy i kolumn uzyskujemy nieosobliwy blok M o wymiarach 1×1 lub 2×2. Na podstawie lematu możemy wskazać macierz R, taką że macierz $R^H A R$ ma zera w pierwszym wierszu i kolumnie, albo w pierwszych dwóch wierszach i kolumnach, poza blokiem diagonalnym. Następnie możemy wskazać macierz blokowo-diagonalną, o postaci $\begin{bmatrix} I & 0 \\ 0 & C \end{bmatrix}$, z blokiem diagonalnym I — macierzą jednostkową 1×1 lub 2×2 i blokiem C, który jest macierzą kongruencji sprowadzającej drugi blok diagonalny naszej macierzy po przkształceniu do pożądanej postaci. Na mocy założenia indukcyjnego taka macierz istnieje.

Krok 2. Macierz kongruencji zastosowanej w kroku pierwszym jest złożeniem macierzy permutacji i macierzy trójkątnych górnych. Aby przekształcić otrzymane na diagonali bloki 2×2 to nie wystarczy. Mając blok $G = \begin{bmatrix} 0 & g \\ \bar{g} & 0 \end{bmatrix} = |g| \begin{bmatrix} 0 & e^{i\varphi} \\ e^{-i\varphi} & 0 \end{bmatrix}$ zastosujemy macierz $E_\varphi = \begin{bmatrix} e^{i\varphi} & e^{i\varphi} \\ 1 & -1 \end{bmatrix}$.

Łatwo możemy sprawdzić, że $E_\varphi^H G E_\varphi = 2|g| \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$. Aby otrzymać macierz diagonalną $n \times n$ z macierzy blokowo-diagonalnej otrzymanej w pierwszym kroku, stosujemy kongruencję reprezentowaną przez macierz blokowo-diagonalną z blokami o takich samych wymiarach; każdy z bloków 1×1 ma współczynnik równy 1, a bloki 2×2 są macierzami E_φ, dla odpowiednio wybranych φ.

Krok 3. Po wykonaniu kroku drugiego mamy hermitowską macierz diagonalną, zatem wszystkie jej współczynniki są liczbami rzeczywistymi. Teraz należy odpowiednio uporządkować i przeskalować współczynniki diagonalne, aby mieć na diagonali kolejno $+1$, -1 i 0. W tym celu stosujemy transpozycje T_{pq} tak, aby poprzestawiać na początek wiersze (i kolumny) zawierające dodatnie współczynniki diagonalne, następnie ujemne i na końcu zera. Po przestawieniu mamy macierz $D' = \mathrm{diag}(d_1, \ldots, d_\pi, -d_{\pi+1}, \ldots, -d_{\pi+\nu}, 0, \ldots, 0)$, gdzie wszystkie liczby $d_1, \ldots, d_{\pi+\nu}$ są dodatnie. Aby otrzymać macierz D, wystarczy zastosować kongruencję diagonalną realizowaną przez macierz $\mathrm{diag}(\frac{1}{\sqrt{d_1}}, \ldots, \frac{1}{\sqrt{d_{\pi+\nu}}}, 1, \ldots, 1)$.

Jest oczywiste, że powyższa konstrukcja jest w ogólnym przypadku niejednoznaczna (tj. że istnieje więcej niż jedna macierz C, taka że macierz $C^H A C$ ma postać opisaną w tezie twierdzenia). Pozostało dowieść, że niezależnie od sposobu przekształcania macierzy, zawsze otrzymamy te same liczby π, ν i ζ współczynników 1, -1 i 0 na diagonali.

Ponieważ kongruencje zachowują rząd macierzy, więc jasne jest, że zawsze $\zeta = n - \mathrm{rank}\, A$. Rozważmy nieosobliwe macierze C_1 i C_2, takie że $C_i^H A C_i = D_i = \mathrm{diag}(\underbrace{1, \ldots, 1}_{\pi_i}, \underbrace{-1, \ldots, -1}_{\nu_i}, \underbrace{0, \ldots, 0}_{\zeta})$. Podprzestrzeń liniowa przestrzeni K^n, której bazą jest zbiór pierwszych π_1 kolumn macierzy C_1 oznaczymy symbolem V_1. Symbolem V_2 oznaczymy podprzestrzeń rozpiętą przez ostatnie $\nu_2 + \zeta$ kolumn macierzy C_2.

Rozważmy formę kwadratową $\Phi\colon \Phi(\mathbf{x}) = \mathbf{x}^H A \mathbf{x}$. Dla każdego niezerowego wektora $\mathbf{x} \in V_1$ mamy $\Phi(\mathbf{x}) > 0$, podobnie dla każdego $\mathbf{x} \in V_2$ jest $\Phi(\mathbf{x}) \leqslant 0$. Przypuśćmy, że $\pi_1 > \pi_2$. Wtedy $\dim V_1 + \dim V_2 = \pi_1 + \nu_2 + \zeta > \pi_2 + \nu_2 + \zeta = n$, zatem wymiar podprzestrzeni $V_1 \cap V_2$ jest większy od zera (nie może być mniejszy niż $\pi_1 - \pi_2$). Stąd wynika istnienie niezerowego wektora $\mathbf{x} \in \mathbb{K}^n$ ($\mathbf{x} \in V_1 \cap V_2$), takiego że $\Phi(\mathbf{x}) > 0$ i $\Phi(\mathbf{x}) \leqslant 0$. Otrzymana sprzeczność dowodzi, że $\pi_1 \leqslant \pi_2$. W podobny sposób można dowieść, że $\pi_2 \leqslant \pi_1$, co kończy dowód. □

Twierdzenie Sylvestra (ogólne): *Niech Φ oznacza dowolną formę kwadratową (drugiego rodzaju) w przestrzeni V o wymiarze n i niech φ oznacza związaną z nią formę półtoraliniową. Istnieją podprzestrzenie V_+, V_- i V_0 przestrzeni V, takie że*

1. *dla każdego $\mathbf{x} \in V_+$, $\mathbf{x} \neq \mathbf{0}$, zachodzi równość $\Phi(\mathbf{x}) > 0$,*
2. *dla każdego $\mathbf{x} \in V_-$, $\mathbf{x} \neq \mathbf{0}$, zachodzi równość $\Phi(\mathbf{x}) < 0$,*
3. *dla każdego $\mathbf{x} \in V_0$, $\mathbf{y} \in V$, jest spełniona równość $\varphi(\mathbf{x}, \mathbf{y}) = 0$,*
4. *jeśli $\mathbf{x} \in V_+$ i $y \in V_-$ to $\varphi(\mathbf{x}, \mathbf{y}) = 0$.*

Wymiary podprzestrzeni V_+, V_- i V_0, odpowiednio π, ν i ζ, są określone jednoznacznie przez formę Φ. Przestrzeń V jest sumą prostą tych podprzestrzeni. Ponadto istnieje baza przestrzeni V, taka że macierz formy Φ w tej bazie ma postać $\mathrm{diag}(\underbrace{1, \ldots, 1}_{\pi}, \underbrace{-1, \ldots, -1}_{\nu}, \underbrace{0, \ldots, 0}_{\zeta})$.

Powyższe twierdzenie jest wnioskiem z twierdzenia macierzowego. Dodajmy jeszcze kilka uwag.

- Niezależność liczb π, ν, ζ od konkretnych podprzestrzeni V_+ i V_- nazywa się bezwładnością formy Φ. Twierdzenie Sylvestra bywa w związku z tym nazywane twierdzeniem o bezwładności.
- Jeśli znamy macierz A formy kwadratowej Φ w bazie $X = [x_1, \dots, x_n]$ oraz macierz C kongruencji sprowadzającej macierz A do postaci D takiej jak w macierzowym twierdzeniu Sylvestra (tzw. postaci kanonicznej), to macierz D reprezentuje formę Φ w bazie $XC = [y_1, \dots, y_n] = Y$. Początkowe π elementów tej bazy rozpina podprzestrzeń V_+, kolejne ν wektorów tworzy bazę V_-, zaś bazę podprzestrzeni V_0 otrzymamy biorąc ζ ostatnich wektorów bazy Y.
- Z wyjątkiem sytuacji, gdy jedna z podprzestrzeni, V_+ lub V_-, jest całą przestrzenią V (czyli gdy $\pi = n$ albo $\nu = n$), jedynie podprzestrzeń V_0 jest przez formę Φ określona jednoznacznie. Przestrzeń tę nazywamy jądrem formy kwadratowej Φ. W przestrzeni $\mathbb{K}^n$ jest nią oczywiście podprzestrzeń $\ker A$, gdzie A jest macierzą formy Φ.
- Forma, dla której podprzestrzeń V_+ jest całą przestrzenią V, nazywa się formą dodatnio określoną. Jeśli $V_- = V$ to mamy formę ujemnie określoną. Oprócz tego możemy mieć do czynienia z formą nieujemnie określoną, jeśli $V_- = \{0\}$ albo z formą niedodatnio określoną, gdy $V_+ = \{0\}$.
  Forma kwadratowa Φ, dla której istnieją wektory x i y, takie że $\Phi(x) > 0$ i $\Phi(y) < 0$, nazywa się nieokreślona.

Przestrzenie ortogonalne

Def. Niech φ oznacza formę półtoraliniową w przestrzeni V. Jeśli dla pewnych wektorów $x, y \in V$ zachodzi równość $\varphi(x, y) = 0$, to mówimy, że wektory x i y są ortogonalne ze względu na formę φ, albo φ-ortogonalne.

Def. Parę (V, φ), w której V jest przestrzenią liniową nad ciałem liczbowym $\mathbb{K}$, a φ — formą półtoraliniową w tej przestrzeni, nazywamy przestrzenią ortogonalną.

Oczywiście, wybierając różne formy półtoraliniowe w ustalonej przestrzeni V dostaniemy różne przestrzenie ortogonalne. Z twierdzenia Sylvestra wynika, że dla ustalonego wymiaru n istnieje tylko skończenie wiele (mianowicie $\frac{1}{2}(n+1)(n+2)$) typów przestrzeni ortogonalnych, różniących się wymiarami podprzestrzeni V_+, V_- i V_0, które można w nich znaleźć. Najważniejsze znaczenie w zastosowaniach praktycznych mają przestrzenie, w których macierz formy jest dodatnio określona (tzw. przestrzenie euklidesowe, w przypadku zespolonym mówi się „unitarne”), ale nie tylko. Na przykład przestrzenie Minkowskiego, dla których $\pi = n - 1$, $\nu = 1$, $\zeta = 0$, są podstawą teorii względności.

Def. Bazę przestrzeni ortogonalnej, w której forma φ jest reprezentowana przez macierz diagonalną, nazywamy bazą ortogonalną. Jeśli w szczególności wszystkie współczynniki tej macierzy są równe 1, -1 lub 0, to mamy bazę ortonormalną.

Def. Jeśli przestrzeń ortogonalna V jest sumą prostą swoich podprzestrzeni $V_1, \dots, V_k$ i jest spełniony warunek, że jeśli $i \neq j$ to dowolne wektory $x \in V_i$ i $y \in V_j$ są ortogonalne ze względu na formę φ, to mówimy, że przestrzeń V jest sumą ortogonalną podprzestrzeni $V_1, \dots, V_k$.

Def. Niezerowe wektory x w przestrzeni ortogonalnej V, które są miejscami zerowymi formy kwadratowej Φ, nazywają się wektorami izotropowymi. Są to wszystkie elementy jądra tej formy, ale w przypadku, gdy jest ona nieokreślona, istnieją także inne wektory izotropowe. Przykład zobaczymy na ćwiczeniach.

Zadania i problemy

1. Udowodnij, że każdej formie kwadratowej w $\mathbb{K}^n$ odpowiada jednoznacznie określona macierz symetryczna (Uwaga: *Nie należy* wierzyć w zapewnienia na wykładzie, jeśli nie były udowodnione).

   Wskazówka: Udowodnij, że każdej formie dwuliniowej pierwszego rodzaju odpowiada jednoznacznie określona macierz symetryczna, a następnie powołaj się na jednoznaczną odpowiedniość między formami dwuliniowymi i kwadratowymi.
2. Znajdź macierz C taką, że jeśli
$$A = \begin{bmatrix} 4 & 2 & & \\ 2 & 5 & 2 & \\ & 2 & 5 & 2 \\ & & 2 & 5 \end{bmatrix}$$
   to macierz $C^H A C$ ma postać kanoniczną Sylvestra.
3. Udowodnij, że zbiór kongruencji ortogonalnych, realizowanych przez macierze C, takie że $C^{-1} = C^H$, jest podgrupą grupy kongruencji. Czy jest to podgrupa abelowa?

   Udowodnij, że zbiór kongruencji realizowanych przez macierze permutacji jest podgrupą grupy kongruencji ortogonalnych. Czy jest to podgrupa właściwa?
4. Forma kwadratowa Φ w przestrzeni $\mathbb{R}[x]_2$ jest określona wzorem
$$\Phi(f) = \int_{-1}^{1} f(x)^2 \, dx.$$
   Znajdź macierz tej formy w bazie $[1, x, x^2]$.

Szczególna teoria względności

5. Przestrzeń ortogonalna z formą kwadratową Φ, taką że $\nu = 1$, $\zeta = 0$, nazywa się przestrzenią Minkowskiego, albo czasoprzestrzenią. W fizyce bada się czasoprzestrzeń czterowymiarową, my ograniczymy się do dwuwymiarowej. Przestrzeń ta jest sumą ortogonalną dwóch podprzestrzeni jednowymiarowych, w których forma Φ jest odpowiednio dodatnio określona i ujemnie określona.

   Punkty przestrzeni nazywają się zdarzeniami.

   Wprowadźmy układ współrzędnych tak, aby macierz A formy Φ miała postać kanoniczną, tj. $\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$. Współrzędne wektora $\boldsymbol{v} = \begin{bmatrix} x \\ t \end{bmatrix}$ określają przemieszczenie w przestrzeni (na odległość x) i w czasie (na odległość t). Wartość formy $\Phi(\boldsymbol{v})$ nazywa się interwałem czasoprzestrzennym.

   Obserwator w czasoprzestrzeni może mierzyć przyrosty czasu (swoim zegarkiem) i przemieszczenia (swoją linijką). Szczególna teoria względności jest konsekwencją stwierdzenia, że jeśli mamy wielu obserwatorów poruszających się względem siebie, to każdy z nich między ustalonymi dwoma zdarzeniami zmierzy zawsze tę samą wartość interwału czasoprzestrzennego.

6. Pierwszy obserwator (dajmy na to, Asia) zmierzył między dwoma zdarzeniami $\boldsymbol{v} = \left[\begin{smallmatrix} x \\ t \end{smallmatrix}\right]$ i obliczy $\Phi(\boldsymbol{v}) = x^2 - t^2$. Drugi obserwator (niech będzie Basia) zmierzy między tymi samymi zdarzeniami wektor $\boldsymbol{v}' = \left[\begin{smallmatrix} x' \\ t' \end{smallmatrix}\right]$ i musi dostać

$$x'^2 - t'^2 = x^2 - t^2.$$

   Znajdźmy macierz kongruencji C, która nie zmienia macierzy A, tj. spełnia warunek $C^T A C = A$:

$$\begin{bmatrix} e & g \\ f & h \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} e & f \\ g & h \end{bmatrix} = \begin{bmatrix} e^2 - g^2 & ef - gh \\ ef - gh & f^2 - h^2 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}.$$

   Stąd wynika, że $e^2 - g^2 = 1$, $ef - gh = 0$ i $f^2 - h^2 = -1$.

   Przypuśćmy, że Asia stwierdziła, że $x = tv$, a Basia zaobserwowała $x' = 0$, $t' = ?$. A więc Basia porusza się względem Asi z prędkością v (w STW rozpatrujemy tylko ruchy ze stałą prędkością) i oba zdarzenia zaobserwowała „w tym samym miejscu”. Zatem,

$$\begin{bmatrix} e & f \\ g & h \end{bmatrix} \begin{bmatrix} 0 \\ t' \end{bmatrix} = \begin{bmatrix} tv \\ t \end{bmatrix}.$$

   Mamy więc $ft' = tv$ oraz $ht' = t$. Dzieląc to stronami dostajemy $\frac{f}{h} = v$. Podstawiając $f = hv$ do $f^2 - h^2 = -1$ dostajemy $h^2(v^2 - 1) = -1$, czyli $h^2 = \frac{1}{1-v^2}$. Dalej przyjmiemy $h = \frac{1}{\sqrt{1-v^2}}$, bo w przeciwnym razie zegarki Asi i Basi chodziłyby w przeciwne strony. Mamy zatem $f = \frac{v}{\sqrt{1-v^2}}$ i dalej podobnie możemy obliczyć (proszę zrobić to osobiście) $e = h$, $g = f$. Stąd macierz kongruencji, która zależy od parametru v (prędkości ruchu Basi względem Asi) jest symetryczna:

$$C_v = \frac{1}{\sqrt{1-v^2}} \begin{bmatrix} 1 & v \\ v & 1 \end{bmatrix}.$$

   Aby ta macierz była rzeczywista i nieosobliwa, musi być $|v| < 1$.

7. Kongruencje opisane przez macierze o powyższej postaci (dla $|v| < 1$) nazywają się przekształceniami Lorentza.

   Udowodnij, że zbiór przekształceń Lorentza (z działaniem złożenia) jest grupą abelową. Znajdź przy tym wzór opisujący składanie prędkości (trzeba w tym celu obliczyć współczynnik poza diagonalą iloczynu macierzy $C_w C_v$, który opisuje złożenie przekształceń opisujących ruch z prędkościami v i w).

   Przekonaj się przy tym, że jeśli $|v| < 1$ i $|w| < 1$, to prędkość ruchu powstałego ze złożenia ruchów z tymi prędkościami jest co do modułu mniejsza niż 1, czyli prędkość światła.

   Jeśli Basia ucieka od Asi z połową prędkości światła ($v = \frac{1}{2}$), zaś Cesia tak samo ucieka przed Basią, to jaką prędkością Cesia oddala się od Asi?

8. Wektory $\left[\begin{smallmatrix} 1 \\ 1 \end{smallmatrix}\right]$ i $\left[\begin{smallmatrix} -1 \\ 1 \end{smallmatrix}\right]$ są wektorami izotropowymi. Każdy z nich opisuje ruch z prędkością światła (w jedną lub drugą stronę). Udowodnij, że jeśli coś porusza się z prędkością światła względem pewnego obserwatora, to ma też tę samą prędkość ruchu względem każdego innego obserwatora.

9. Asia zauważyła pewne zdarzenia (punkty czasoprzestrzeni) odległe od siebie o wektor $\boldsymbol{v} = \left[\begin{smallmatrix} d \\ 0 \end{smallmatrix}\right]$ (a więc zaobserwowała je jednocześnie).

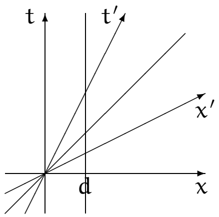
   Basia zobaczyła te same zdarzenia, które jej zdaniem różniły się o wektor

$$\boldsymbol{v}' = \frac{1}{\sqrt{1-v^2}} \begin{bmatrix} 1 & -v \\ -v & 1 \end{bmatrix} \begin{bmatrix} d \\ 0 \end{bmatrix} = \begin{bmatrix} d/\sqrt{1-v^2} \\ -dv/\sqrt{1-v^2} \end{bmatrix},$$

   a więc zaobserwowała je w różnych chwilach.

   Przypuśćmy, że obserwatorzy mają do czynienia ze zjawiskami trwającymi w czasie. Asia zwraca uwagę na proste $x = 0$ i $x = d$ (tj. Asia obserwuje „dwa nieruchome punkty”). Basia jednocześnie ze zdarzeniem $\left[\begin{smallmatrix} 0 \\ 0 \end{smallmatrix}\right]$ zaobserwuje zdarzenie, którego współrzędnymi w układzie Asi są liczby d i dv. Interwał czasoprzestrzenny tych zdarzeń jest równy $\Phi(\left[\begin{smallmatrix} d \\ dv \end{smallmatrix}\right]) = d^2(1 - v^2)$. Ale Basia „te dwa punkty” też obserwuje cały czas, widząc jak poruszają się względem niej, pozostając w stałej odległości od siebie. Tak więc odległość „tych samych” punktów zmierzona przez Basię jest równa $d\sqrt{1-v^2}$, czyli jest mniejsza.

   Zjawisko to nazywa się relatywistycznym skróceniem długości. Kiedy dwaj obserwatorzy mijają się, każdy z nich uważa, że to ten drugi jest krótszy.



10. Asia i Basia są bliźniaczkami w tym samym wieku. Asia obserwuje, jak Basia wyrusza w drogę, podróżuje w jedną stronę przez czas $t/2$, i natychmiast po osiągnięciu celu podróży zawraca i przez drugie tyle samo czasu wraca z tą samą prędkością v.

 Dla Asi zdarzenie, którym było zawrócenie Basi z drogi, ma współrzędne $vt/2$ i $t/2$. Basia zawróciła w punkcie (zdaniem Basi)

$$\frac{1}{\sqrt{1-v^2}}\begin{bmatrix} 1 & -v \\ -v & 1 \end{bmatrix}\begin{bmatrix} vt/2 \\ t/2 \end{bmatrix} = \frac{1}{2\sqrt{1-v^2}}\begin{bmatrix} 0 \\ -v^2t+t \end{bmatrix},$$

 a więc do tego momentu upłynęło jej $t\sqrt{1-v^2}/2$ czasu. Droga powrotna zajęła Basi tyle samo, a więc cała wyprawa trwała $t' = t\sqrt{1-v^2}$ czasu zmierzonego przez jej zegarek.

 Tak więc po powrocie Basia jest młodsza od siostry. Wniosek: Należy jak najwięcej podróżować, wtedy człowiek wolniej się starzeje. Można też szybko spacerować po (nawet niewielkim) pomieszczeniu w tę i z powrotem (to wyjaśnia dosyć rozpowszechniony zwyczaj spacerowania po celi).

Równania drugiego stopnia

Zbiory algebraiczne

Def. Podzbiór A przestrzeni $\mathbb{K}^n$ nazywamy zbiorem algebraicznym, jeśli istnieje układ wielomianów n zmiennych, $p_1, \ldots, p_k$, taki że $\mathbf{x} \in A$ wtedy i tylko wtedy gdy $p_i(\mathbf{x}) = 0$ dla $i = 1, \ldots, k$.

Zbiory algebraiczne możemy określić w dowolnej przestrzeni liniowej lub afinicznej, wybierając odpowiedni (kartezjański) układ współrzędnych i żądając, aby współrzędne punktów spełniały pewne równania. Jest oczywiste, że reprezentowanie zbiorów algebraicznych za pomocą wielomianów nie jest jednoznaczne. Nie każdy zbiór jest algebraiczny, np. kula (w sensie dowolnej normy) zbiorem algebraicznym nie jest.

Skończona suma (teoriomnogościowa), a także przecięcie zbiorów algebraicznych jest zbiorem algebraicznym. Zauważmy też, że każdy zbiór jednopunktowy jest algebraiczny. Tak więc każdy zbiór skończony jest algebraiczny.

Def. Stopniem zbioru algebraicznego A nazywamy najmniejszą liczbę, którą można otrzymać wybierając pewien układ wielomianów reprezentujących ten zbiór i obliczając iloczyn stopni tych wielomianów.

Jedynym zbiorem algebraicznym stopnia 0 jest zbiór pusty. Ponieważ miejsca zerowe wielomianów stopnia 1 spełniają równania liniowe, więc zbiory algebraiczne stopnia 1 to warstwy przestrzeni $\mathbb{K}^n$ (albo, jeśli traktujemy przestrzeń $\mathbb{K}^n$ jak afiniczną, to zbiorami algebraicznymi stopnia 1 są jej podprzestrzenie afiniczne).

Równania drugiego stopnia

Zbadamy zbiory algebraiczne drugiego stopnia w przestrzeni $\mathbb{R}^n$, traktując ją jak przestrzeń afiniczną, tj. rezerwując sobie możliwość dowolnego wybierania układu współrzędnych kartezjańskich (w szczególności o początku w dowolnym punkcie).

Wystarczy zająć się zbiorami, które są miejscami zerowymi jednego wielomianu drugiego stopnia. Jeśli mamy do tego jakieś równania liniowe, to z ich pomocą możemy wyeliminować niektóre zmienne, otrzymując równanie drugiego stopnia zbioru w przestrzeni o mniejszym wymiarze. Przekonamy się, że każde równanie drugiego stopnia w przestrzeni $\mathbb{R}^n$ możemy, przez odpowiednią zamianę zmiennych, przedstawić w „najprostszej” postaci.

Dowolny wielomian drugiego stopnia możemy przedstawić w postaci macierzowej:

$$p(\mathbf{x}) = \mathbf{x}^T A\mathbf{x} + 2\mathbf{b}^T\mathbf{x} + c.$$

Wielomian p jest tu sumą formy kwadratowej i wielomianu pierwszego stopnia. Bez straty ogólności, macierz formy jest symetryczna, przy czym założymy, że nie jest to macierz zerowa (bo wtedy mielibyśmy równanie liniowe). Ten sam wielomian możemy przedstawić w postaci jednorodnej:

$$p(\mathbf{x}) = \begin{bmatrix} \mathbf{x} \\ 1 \end{bmatrix}^T \begin{bmatrix} A & \mathbf{b} \\ \mathbf{b}^T & c \end{bmatrix} \begin{bmatrix} \mathbf{x} \\ 1 \end{bmatrix}.$$

Stosowanie jednego lub drugiego zapisu jest kwestią wygody. Zbadamy kolejne kroki (polegające na zamianie zmiennych), jakie prowadzą do otrzymania „najprostszej” postaci równania $p(\mathbf{x}) = 0$, którą nazwiemy postacią kanoniczną.

Pierwszy krok jest przekształceniem liniowym, dzięki któremu zamiast macierzy symetrycznej A otrzymamy macierz diagonalną D ze współczynnikami diagonalnymi 1, -1 i 0. Na podstawie twierdzenia Sylvestra istnieje taka macierz nieosobliwa C, że $C^T AC = D$. Mamy zatem

$$p(\mathbf{x}) = \mathbf{x}^T C^{-T} DC^{-1}\mathbf{x} + 2\mathbf{b}^T CC^{-1}\mathbf{x} + c = \mathbf{x}'^T D\mathbf{x}' + 2\mathbf{b}'^T\mathbf{x}' + c,$$

gdzie $\mathbf{b}' = C^T\mathbf{b}$ oraz $\mathbf{x}' = C^{-1}\mathbf{x}$. Uwaga: Symbol „$'$” *nie jest* w tym wykładzie symbolem pochodnej.

Możemy założyć, że liczba diagonalnych współczynników macierzy D równych 1, tj. π, jest nie mniejsza niż liczba współczynników -1. Jeśli bowiem $\pi < \nu$, to zajmiemy się równaniem $-p(\mathbf{x}) = 0$, które ma ten sam zbiór rozwiązań.

W kroku drugim przypuśćmy najpierw, że układ równań liniowych $D\mathbf{t} = -\mathbf{b}'$ z niewiadomym wektorem $\mathbf{t}$ jest niesprzeczny. Możemy zatem wziąć dowolne jego rozwiązanie $\mathbf{t}$ i podstawiając $\mathbf{x}' = \mathbf{x}'' + \mathbf{t}$, otrzymać

$$\begin{aligned} p(\mathbf{x}) &= (\mathbf{x}'' + \mathbf{t})^T D(\mathbf{x}'' + \mathbf{t}) + 2\mathbf{b}'^T(\mathbf{x}'' + \mathbf{t}) + c = \\ &\mathbf{x}''^T D\mathbf{x}'' + 2(D\mathbf{t})^T\mathbf{x}'' + 2\mathbf{b}'^T\mathbf{x}'' + \mathbf{t}^T D\mathbf{t} + 2\mathbf{b}'^T\mathbf{t} + c = \\ &\mathbf{x}''^T D\mathbf{x}'' + c'', \end{aligned}$$

gdzie $c'' = -\mathbf{t}^T D\mathbf{t} + c$. W ten sposób „pozbyliśmy się” liniowego składnika wielomianu p.

Zanim poddamy równanie dalszym przekształceniom, zauważmy, że jeśli dla pewnego wektora $\mathbf{y}$ zachodzi równość $\mathbf{y}^T D\mathbf{y} + c'' = 0$, to również $(-\mathbf{y})^T D(-\mathbf{y}) + c'' = 0$. Punkt $\mathbf{0}$ jest więc środkiem symetrii zbioru rozwiązań równania $y^T D y + c'' = 0$, skąd wnioskujemy, że punkt $\mathbf{t}$ jest środkiem symetrii zbioru rozwiązań równania $x'^T D x' + 2\mathbf{b}'^T x' + c = 0$ (z niewiadomym wektorem x'). Zbiór rozwiązań równania $p(x) = 0$ ma środek symetrii w punkcie $\mathbf{s} = C\mathbf{t}$. Dalej przekonamy się, że nie ma innych środków symetrii, a zatem zbiór środków symetrii dowolnego zbioru algebraicznego drugiego stopnia jest warstwą (albo podprzestrzenią afiniczną) przestrzeni $\mathbb{R}^n$, albo jest pusty.

Jeśli udało się przedstawić wielomian p w postaci $x''^T D x'' + c''$, to możliwe są dwa przypadki: może być $c'' = 0$ albo $c'' \neq 0$. Jeśli $c'' = 0$, to mamy już postać kanoniczną, która w rozwiniętej postaci wygląda tak:

$$y_1^2 + \cdots + y_\pi^2 - y_{\pi+1}^2 - \cdots - y_{\pi+\nu}^2 = 0$$

(symbole $y_1, \ldots, y_n$ oznaczają teraz zmienne, które są współrzędnymi w układzie, w którym ostatecznie przedstawiliśmy nasze równanie). Jeśli $c'' \neq 0$, to po podstawieniu $x'' = \sqrt{|c''|}x'''$ otrzymamy wielomian $|c''|(x'''^T D' x''' \pm 1)$, który możemy podzielić przez $|c''|$ i otrzymać wielomian o tych samych miejscach zerowych. Postać kanoniczna równania otrzymana w tym przypadku wygląda następująco:

$$y_1^2 + \cdots + y_\pi^2 - y_{\pi+1}^2 - \cdots - y_{\pi+\nu}^2 \pm 1 = 0.$$

Pozostał do zbadania przypadek, gdy układ równań $D\mathbf{t} = -\mathbf{b}'$ jest sprzeczny (czyli, jak się przekonamy, gdy zbiór rozwiązań nie ma środka symetrii). Wektor $\mathbf{b}'$ możemy przedstawić w postaci sumy dwóch wektorów, $\mathbf{b}_1 \neq 0$ i $\mathbf{b}_2$, takich że $\mathbf{b}_1 \in \ker D$ i $\mathbf{b}_2 \in \operatorname{im} D$. Wektory te spełniają warunek $\mathbf{b}_1^T \mathbf{b}_2 = 0$.

<u>Uwaga:</u> Na stronie 4.4 jest dowód takiego twierdzenia: *dla dowolnej macierzy* $A \in \mathbb{C}^{m,n}$ *jest* $\ker A \oplus \operatorname{im} A^H = \mathbb{C}^n$ *oraz* $\operatorname{im} A \oplus \ker A^H = \mathbb{C}^m$. W przypadku, gdy macierz A jest rzeczywista i symetryczna, wynika stąd $\mathbb{R}^n = \ker A \oplus \operatorname{im} A$. Szczególnie łatwo jest to widoczne dla macierzy diagonalnej, z jaką mamy teraz do czynienia.

Pierwsze $\pi + \nu = \operatorname{rank} D$ współczynników wektora $\mathbf{b}_1$ to zera, a pozostałe są odpowiednimi współczynnikami wektora $\mathbf{b}'$. Wektor $\mathbf{b}_2$ otrzymamy zamieniając na zera ostatnie $\zeta = n - \operatorname{rank} D$ współczynników wektora $\mathbf{b}'$. Układ równań $D\mathbf{t} = -\mathbf{b}_2$ jest niesprzeczny i istnieje (być może więcej niż jedno) jego rozwiązanie $\mathbf{t}$ spełniające warunek $\mathbf{b}_1^T \mathbf{t} = 0$. Za pomocą takiego rozwiązania określmy wektor x'', taki że $x' = x'' + \mathbf{t} + s\mathbf{b}_1$; parametr s dobierzemy za chwilę. Liczymy

$$\begin{aligned} p(x) = {} & (x'' + \mathbf{t} + s\mathbf{b}_1)^T D (x'' + \mathbf{t} + s\mathbf{b}_1) + 2(\mathbf{b}_1 + \mathbf{b}_2)^T (x'' + \mathbf{t} + s\mathbf{b}_1) + c = \\ & x''^T D x'' + 2\mathbf{t}^T D x'' + 2\mathbf{b}_2^T x'' + \mathbf{t}^T D \mathbf{t} + 2\mathbf{b}_1^T x'' + 2\mathbf{b}_2^T \mathbf{t} + 2s\mathbf{b}_1^T \mathbf{b}_1 + c = \\ & x''^T D x'' + 2\mathbf{b}_1^T x'' + (c - \mathbf{t}^T D \mathbf{t} + 2s\mathbf{b}_1^T \mathbf{b}_1). \end{aligned}$$

Podstawiając $s = (\mathbf{t}^T D \mathbf{t} - c)/(2\mathbf{b}_1^T \mathbf{b}_1)$ dostajemy

$$p(x) = x''^T D x'' + 2\mathbf{b}_1^T x'',$$

a zatem „pozbywamy się" wyrazu wolnego.

Niech E oznacza macierz $n \times n$, $E = [\mathbf{e}_1, \ldots, \mathbf{e}_{\pi+\nu}, 2\mathbf{b}_1, \mathbf{f}_2, \ldots, \mathbf{f}_\zeta]$, której początkowe kolumny są kolumnami macierzy jednostkowej, dalej mamy wektor $2\mathbf{b}_1$, a po nim wektory $\mathbf{f}_2, \ldots, \mathbf{f}_\zeta$ wybrane spośród wektorów $\mathbf{e}_{\pi+\nu+1}, \ldots, \mathbf{e}_n$ tak, aby macierz E była nieosobliwa. Podstawiając $x'' = E^{-T} x'''$ otrzymamy

$$p(x) = x'''^T E^{-1} D E^{-T} x''' + 2\mathbf{b}_1^T E^{-T} x'''.$$

Macierz E skonstruowana w opisany wyżej sposób spełnia następujące warunki: $E^{-1} D E^{-T} = D$ oraz $E^{-1}\mathbf{b}_1 = \frac{1}{2}\mathbf{e}_{\pi+\nu+1}$. Dlatego po tym podstawieniu mamy

$$p(x) = x'''^T D x''' + \mathbf{e}_{\pi+\nu+1}^T x'''.$$

To jest właśnie trzecia postać kanoniczna wielomianu drugiego stopnia:

$$y_1^2 + \cdots + y_\pi^2 - y_{\pi+1}^2 - \cdots - y_{\pi+\nu}^2 + y_{\pi+\nu+1} = 0.$$

(zmienne y_i są współrzędnymi wektora x'''). Zbiór rozwiązań powyższego równania jest niepusty (bo jego elementem jest wektor zerowy). Dowód, że zbiór rozwiązań równania $x^T A x + 2\mathbf{b}^T x + c = 0$, takiego że układ $A\mathbf{s} = -\mathbf{b}$ jest sprzeczny (czyli każdego równania drugiego stopnia, które można przez afiniczne przekształcenie wektora niewiadomych sprowadzić do tej postaci), nie ma środka symetrii jest umiarkowanie trudny i pozostawiam go na ćwiczenia.

Rachunki przeprowadzone wyżej są z jednej strony procedurą sprowadzania równań do postaci kanonicznej, z drugiej zaś są dowodem poniższego twierdzenia.

<u>Twierdzenie:</u> *Dowolne równanie drugiego stopnia n niewiadomych o współczynnikach rzeczywistych może być, poprzez odpowiednie afiniczne przekształcenie wektora niewiadomych, przedstawione w jednej (i tylko jednej) z trzech tzw. <u>afinicznych postaci kanonicznych</u>:*

$$\begin{aligned} & y_1^2 + \cdots + y_\pi^2 - y_{\pi+1}^2 - \cdots - y_{\pi+\nu}^2 = 0, \\ & y_1^2 + \cdots + y_\pi^2 - y_{\pi+1}^2 - \cdots - y_{\pi+\nu}^2 \pm 1 = 0, \\ & y_1^2 + \cdots + y_\pi^2 - y_{\pi+1}^2 - \cdots - y_{\pi+\nu}^2 + y_{\pi+\nu+1} = 0, \end{aligned}$$

przy czym $\pi \geqslant \nu$. W ogólności może istnieć więcej niż jedno przekształcenie sprowadzające dane równanie do którejś z tych postaci, ale liczby π i ν są zawsze takie same.

Klasyfikacja zbiorów stopnia 2 w płaszczyźnie

Ponieważ każde równanie dwóch zmiennych 2-go stopnia ma jedną ze skończenie wielu postaci kanonicznych, określenie czym jest zbiór rozwiązań sprowadza się do odpowiedniego nazwania zbioru rozwiązań równania w postaci kanonicznej. Dowolne dwa równania o tej samej postaci kanonicznej mają zbiory rozwiązań identyczne z dokładnością to podobieństwa afinicznego. Mamy zatem takie przypadki (symbole x i y oznaczają niewiadome):

$x^2 = 0$:	Zbiorem rozwiązań jest prosta $x = 0$.
$x^2 + y^2 = 0$:	Zbiór jednopunktowy, $x = y = 0$.
$x^2 - y^2 = 0$:	Dwie proste, o równaniach $x + y = 0$ i $x - y = 0$.
$x^2 - 1 = 0$:	Dwie proste równoległe, $x = 1$ i $x = -1$.
$x^2 + 1 = 0$:	Zbiór pusty.
$x^2 + y^2 + 1 = 0$:	Zbiór tym bardziej pusty.
$x^2 - y^2 + 1 = 0$:	Hiperbola.
$x^2 + y^2 - 1 = 0$:	Elipsa.
$x^2 + y = 0$:	Parabola.

Klasyfikacja zbiorów stopnia 2 w przestrzeni trójwymiarowej

$x^2 = 0$:	Płaszczyzna $x = 0$.
$x^2 + y^2 = 0$:	Prosta $x = y = 0$.
$x^2 - y^2 = 0$:	Dwie płaszczyzny, o równaniach $x + y = 0$ i $x - y = 0$.
$x^2 + y^2 + z^2 = 0$:	Zbiór jednopunktowy, $x = y = z = 0$.
$x^2 + y^2 - z^2 = 0$:	Stożek.
$x^2 - 1 = 0$:	Dwie płaszczyzny równoległe, $x = 1$ i $x = -1$.
$x^2 + 1 = 0$:	Zbiór pusty.
$x^2 + y^2 + 1 = 0$:	Zbiór jak wyżej pusty.
$x^2 - y^2 + 1 = 0$:	Walec hiperboliczny.
$x^2 + y^2 - 1 = 0$:	Walec eliptyczny.
$x^2 + y^2 + z^2 + 1 = 0$:	Zbir pusty.
$x^2 + y^2 - z^2 + 1 = 0$:	Hiperboloida dwupowłokowa.
$x^2 + y^2 - z^2 - 1 = 0$:	Hiperboloida jednopowłokowa.
$x^2 + y^2 + z^2 - 1 = 0$:	Elipsoida.
$x^2 + y = 0$:	Walec paraboliczny.
$x^2 + y^2 + z = 0$:	Paraboloida eliptyczna.
$x^2 - y^2 + z = 0$:	Paraboloida hiperboliczna.

Powierzchnie drugiego stopnia, tj. elipsoida, stożek, walce, paraboloidy i hiperboloidy są nazywane kwadrykami.

Zbiory prostokreślne

Def. Zbiór S rozwiązań równania drugiego stopnia nazywa się zbiorem prostokreślnym, jeśli każdy jego punkt leży na pewnej prostej afinicznej, która jest zawarta w zbiorze S.

Na przykład zbiór rozwiązań równania $x^2 - y^2 = 0$ na płaszczyźnie jest zbiorem prostokreślnym, ponieważ składa się on z dwóch prostych. Nie jest zbiorem prostokreślnym elipsa, tj. zbiór rozwiązań równania $x^2 + y^2 - 1 = 0$.

Znacznie ciekawiej to wygląda w przypadku zbiorów rozwiązań równań z trzema niewiadomymi. My spróbujmy rozwiązać w przypadku ogólnym (dla równań z dowolną liczbą niewiadomych) następujący problem: mając dane równanie, należy stwierdzić, czy jego zbiór rozwiązań S jest prostokreślny i jeśli tak, to znaleźć wszystkie proste przechodzące przez dowolny punkt $\boldsymbol{x}_0 \in S$, zawarte w zbiorze S.

Niech $p(\boldsymbol{x}) = \boldsymbol{x}^T A\boldsymbol{x} + 2\boldsymbol{b}^T\boldsymbol{x} + c$. Jeśli $p(\boldsymbol{x}_0) = 0$, to dowolny punkt $\boldsymbol{x}' = \boldsymbol{x} - \boldsymbol{x}_0$, taki że $p(\boldsymbol{x}) = 0$, jest rozwiązaniem równania

$$\boldsymbol{x}'^T A\boldsymbol{x}' + 2\boldsymbol{b}'^T\boldsymbol{x}' = 0.$$

Jeśli $\boldsymbol{b}' = \boldsymbol{0}$, to mamy równanie $\boldsymbol{x}'^T A\boldsymbol{x}' = 0$, którego rozwiązaniem jest każdy wektor izotropowy formy kwadratowej reprezentowanej przez macierz A.

Jeśli $\boldsymbol{b}' \neq \boldsymbol{0}$, to dowolny wektor $\boldsymbol{x}' \in \mathbb{R}^n$ możemy (jednoznacznie) przedstawić w postaci sumy: $\boldsymbol{x}' = d\boldsymbol{b}' + \boldsymbol{z}$, dla pewnego wektora $\boldsymbol{z}$ spełniającego warunek $\boldsymbol{b}'^T\boldsymbol{z} = 0$ (dowód na ćwiczeniach). Podstawiając tę sumę, otrzymujemy

$$p(\boldsymbol{x}) = p'(\boldsymbol{x}') = (d\boldsymbol{b}' + \boldsymbol{z})^T A(d\boldsymbol{b}' + \boldsymbol{z}) + 2\boldsymbol{b}'^T(d\boldsymbol{b}' + \boldsymbol{z}) = \\ d^2\boldsymbol{b}'^T A\boldsymbol{b}' + 2d\boldsymbol{b}'^T A\boldsymbol{z} + \boldsymbol{z}^T A\boldsymbol{z} + 2d\boldsymbol{b}'^T\boldsymbol{b}'.$$

Jeśli powyższe wyrażenie ma wartość 0 dla pewnego wektora $\boldsymbol{x}' = d\boldsymbol{b}' + \boldsymbol{z}$, to chcemy, aby było $p'(e\boldsymbol{x}') = 0$ dla każdego $e \in \mathbb{R}$. Ponieważ jednak $\boldsymbol{b}'^T\boldsymbol{b}' \neq 0$, więc musimy przyjąć $d = 0$. Istnieje niezależny liniowo układ wektorów $\boldsymbol{z}_1, \ldots, \boldsymbol{z}_{n-1}$, taki że $\boldsymbol{b}'^T\boldsymbol{z}_i = 0$ dla $i = 1, \ldots, n-1$ (sposób znalezienia takich wektorów będzie tematem dalszych wykładów). Możemy więc przyjąć $\boldsymbol{x}' = Z\boldsymbol{y}$, gdzie macierz $Z = [\boldsymbol{z}_1, \ldots, \boldsymbol{z}_{n-1}] \in \mathbb{R}^{n,n-1}$ jest kolumnowo regularna. Ponieważ $\boldsymbol{b}^T Z = \boldsymbol{0} \in \mathbb{R}^{n-1}$, więc

$$p'(Z\boldsymbol{y}) = \boldsymbol{y}^T(Z^T A Z)\boldsymbol{y}.$$

Otrzymujemy w ten sposób równanie jednorodne $p'(Z\boldsymbol{y}) = 0$, którego rozwiązania określają kierunki prostych zawartych w zbiorze rozwiązań równania $p(\boldsymbol{x}) = 0$, przechodzących przez punkt $\mathbf{x}_0$. Wektory $\mathbf{y}$ spełniające to równanie są wektorami izotropowymi formy kwadratowej reprezentowanej przez macierz $Z^T A Z \in \mathbb{R}^{n-1,n-1}$.

Zadania i problemy

1. Udowodnij, że punkt $\mathbf{s}$ spełniający układ równań $A\mathbf{s} = -\mathbf{b}$ jest środkiem symetrii zbioru rozwiązań równania $\mathbf{x}^T A\mathbf{x} + 2\mathbf{b}^T\mathbf{x} + c = 0$.

   Wykaż, że niesprzeczność układu równań $A\mathbf{s} = -\mathbf{b}$ jest równoważna niesprzeczności układu $D\mathbf{t} = -\mathbf{b}'$, gdzie $D = C^T AC$, $\mathbf{b}' = C^T\mathbf{b}$, dla dowolnej macierzy nieosobliwej C. Warunkiem koniecznym i dostatecznym istnienia środka symetrii zbioru rozwiązań równania drugiego stopnia jest więc niesprzeczność dowolnego z tych układów.
2. Udowodnij, że zbiór rozwiązań równania $\mathbf{x}^T A\mathbf{x} + 2\mathbf{b}^T\mathbf{x} + c = 0$, takiego że układ równań liniowych $A\mathbf{s} = -\mathbf{b}$ jest sprzeczny, nie ma środka symetrii.

   Wskazówka: Uzasadnij, że wystarczy zrobić to dla macierzy A w postaci kanonicznej. Udowodnij, że jeśli układ z taką macierzą jest sprzeczny, to punkt $\mathbf{0}$ nie jest środkiem symetrii zbioru rozwiązań równania $\mathbf{x}^T A\mathbf{x} + 2\mathbf{b}^T\mathbf{x} + c = 0$.
3. Oblicz, ile różnych rodzajów zbiorów drugiego stopnia (tj. zbiorów rozwiązań równań drugiego stopnia) w sensie klasyfikacji afinicznej jest w przestrzeni $\mathbb{R}^n$ dla dowolnego n.

   Podaj ile jest w tej klasyfikacji rodzajów zbiorów, które mają środki symetrii.

   Uwaga: Zbiór pusty jest zbiorem rozwiązań więcej niż jednego równania w postaci kanonicznej.
4. Znajdź wzory, które dla ustalonych wektorów $\mathbf{b} \neq \mathbf{0}$ i $\mathbf{x}$ umożliwiają obliczenie wektora $\mathbf{z}$ i liczby d, takich że $\mathbf{x} = d\mathbf{b} + \mathbf{z}$ oraz $\mathbf{b}^T\mathbf{z} = 0$.
5. Znajdź postać kanoniczną D macierzy trójdiagonalnej

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

   Napisz macierz kongruencji R, takiej że $D = R^T AR$.
6. Zapisz w postaci macierzowej, a następnie sprowadź do postaci kanonicznej równania

   a) $4x^2 - 12xy + 9y^2 + 4x - 6y + 1 = 0$,

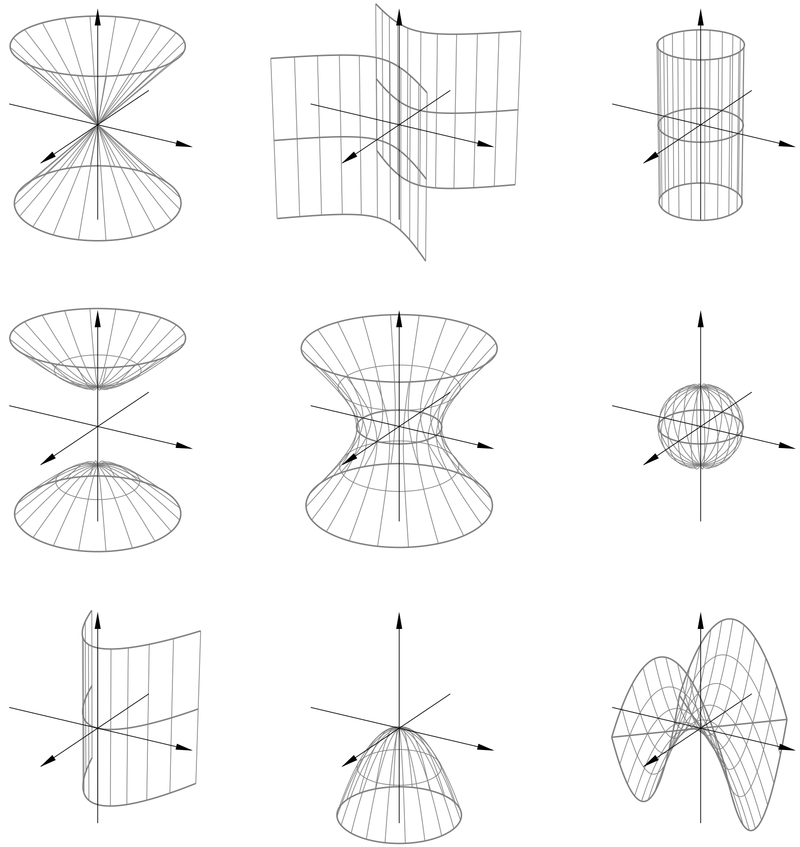
   b) $x^2 + 2xy - 2yz - z^2 + x - z = 0$.

   Zidentyfikuj zbiory rozwiązań tych równań.

7. Zidentyfikuj zbiory rozwiązań równania

$$x^2 + ay^2 + (a-2)z^2 + (1-a)^2 = 0$$

   w zależności od parametru $a \in \mathbb{R}$.
8. Zidentyfikuj (tj. dobierz odpowiednie nazwy i równania) powierzchnie drugiego stopnia, których fragmenty są przedstawione na rysunku.



9. Które spośród kwadryk (powierzchni drugiego stopnia w przestrzeni trójwymiarowej) są prostokreślne?

   Ile prostych zawartych w takiej powierzchni przechodzi przez każdy jej punkt?
10. Wskaż zbiór środków symetrii każdej z kwadryk.
11. Znajdź wszystkie proste zawarte w zbiorze rozwiązań równania

$$x^2 + y^2 - z^2 - 1 = 0$$

   przechodzące przez punkt $[1, 2, 2]^T$.

Przestrzenie euklidesowe i unitarne

Macierze dodatnio określone

Pojęcia macierzy dodatnio, nieujemnie, ujemnie i niedodatnio określonej, a także macierzy nieokreślonej, wiążą się z formami kwadratowymi reprezentowanymi przez te macierze. I tak np. macierz hermitowska (czyli symetryczna w przypadku rzeczywistym) $A \in \mathbb{K}^{n,n}$ jest <u>dodatnio określona</u>, jeśli dla każdego wektora $x \in \mathbb{K}^n$, $x \neq 0$, jest spełniona nierówność

$$x^H A x > 0.$$

Często symbolicznie zapisuje się to tak: $A > 0$, ale <u>uwaga:</u> niech to się nikomu nie myli z napisami takimi jak $A > B$, oznaczającymi wcześniej przedstawioną relację częściowego porządku między macierzami A i B.

<u>Twierdzenie</u> (kryterium Sylvestra): *Niech* $A^{(k)}$ *oznacza macierz o wymiarach* $k \times k$*, która jest blokiem macierzy hermitowskiej* A *o wymiarach* $n \times n$*, powstałym przez odrzucenie końcowych* $n-k$ *kolumn i wierszy. Macierz* A *jest dodatnio określona wtedy i tylko wtedy gdy* $\det A^{(k)} > 0$ *dla* $k = 1, \ldots, n$.

<u>Dowód:</u> Jest oczywiste, że warunek $A^{(k)} > 0$ dla $k = 1, \ldots, n-1$ jest konieczny dla dodatniej określoności macierzy $A = A^{(n)}$. Istotnie, gdyby było inaczej, tj. gdyby istniał wektor $y \in \mathbb{K}^k$, $y \neq 0$, taki że $y^H A^{(k)} y \leqslant 0$, to dla wektora x otrzymanego przez dopisanie $n-k$ zer do wektora y mielibyśmy $x^H A x = y^H A^{(k)} y \leqslant 0$.

Dla macierzy dodatnio określonej $A^{(k)}$ istnieje macierz nieosobliwa C_k, taka że $C_k^H A^{(k)} C_k = I_k$. Na podstawie twierdzenia Cauchy'ego mamy $|\det C_k|^2 \det A^{(k)} = 1$, skąd wynika nierówność $\det A^{(k)} > 0$. Zatem wyznaczniki wszystkich macierzy $A^{(k)}$ otrzymanych z macierzy dodatnio określonej A są dodatnie.

Należy jeszcze dowieść, że spełnienie nierówności $\det A^{(k)} > 0$ dla $k = 1, \ldots, n$ jest warunkiem dostatecznym dodatniej określoności macierzy A. Udowodnimy to indukcyjnie. Dla $n = 1$ twierdzenie jest oczywiste, załóżmy więc, że twierdzenie jest prawdziwe dla wszystkich macierzy hermitowskich o wymiarach mniejszych niż $n \times n$. Macierz A przedstawimy w postaci blokowej,

$$A = \begin{bmatrix} A^{(n-1)} & g \\ g^H & a_{nn} \end{bmatrix}$$

i założymy, że $\det A^{(k)} > 0$ dla $k = 1, \ldots, n$. Stąd i z założenia indukcyjnego

wynika, że macierz $A^{(n-1)}$ jest dodatnio określona. Kongruencja o macierzy

$$R = \begin{bmatrix} I_{n-1} & -\big(A^{(n-1)}\big)^{-1} g \\ 0^H & 1 \end{bmatrix}$$

sprowadza macierz A do postaci blokowo-diagonalnej (zobacz dowód macierzowego twierdzenia Sylvestra):

$$A' = R^H A R = \begin{bmatrix} A^{(n-1)} & 0 \\ 0^H & a'_{nn} \end{bmatrix} \quad \text{gdzie } a'_{nn} = a_{nn} - g^H \big(A^{(n-1)}\big)^{-1} g.$$

Wyznacznik macierzy R jest równy 1, a zatem $\det A' = \det A > 0$. Ponieważ $\det A^{(n-1)} > 0$, więc musi być $a'_{nn} > 0$. Dlatego dla dowolnego wektora $x = \left[\begin{smallmatrix} y \\ y_n \end{smallmatrix}\right] \neq 0$ mamy $x^H A' x = y^H A^{(n-1)} y + |y_n|^2 a'_{nn} > 0$, a zatem macierz A', czyli także macierz A, która jest macierzą przystającą do A', jest dodatnio określona. □

Iloczyny skalarne

<u>Def.</u> Niech V oznacza rzeczywistą lub zespoloną przestrzeń liniową. Forma dwuliniowa (albo półtoraliniowa) φ, taka że odpowiednia forma kwadratowa Φ jest dodatnio określona, nazywa się <u>iloczynem skalarnym</u> w przestrzeni V. Przestrzeń V z iloczynem skalarnym φ w przypadku rzeczywistym nazywa się <u>przestrzenią euklidesową</u>, a w przypadku zespolonym <u>przestrzenią unitarną</u>.

Dalej w przypadku gdy forma φ jest iloczynem skalarnym, zamiast pisać $\varphi(x, y)$ będziemy używać krótszej notacji $\langle x, y\rangle$. Liniowa przestrzeń euklidesowa lub unitarna jest więc szczególnym przypadkiem przestrzeni ortogonalnej. Przy okazji zdefiniujemy najważniejsze (niewątpliwie) pojęcie geometrii:

<u>Def.</u> Przestrzeń afiniczna E, której przestrzeń wektorów swobodnych V jest euklidesowa, nazywa się <u>przestrzenią afiniczną euklidesową</u>.

W poprzednich wykładach pojawiło się wiele faktów mających bezpośredni związek z iloczynami skalarnymi. Możemy je teraz zebrać i dokonać ponownej interpretacji.

- Dla dowolnego iloczynu skalarnego w przestrzeni V o wymiarze n istnieje baza $X = [x_1, \ldots, x_n]$, taka że macierz iloczynu skalarnego w tej bazie jest jednostkowa. Zatem jeśli wektory x i y reprezentujemy za pomocą kolumnowych macierzy współczynników w bazie X, odpowiednio a i $b \in \mathbb{K}^n$ (czyli $x = Xa$ i $y = Xb$), to

  $$\langle x, y\rangle = b^H a.$$

- Wcześniej udowodniliśmy następującą nierówność Schwarza: dla dowolnych wektorów $x, y \in \mathbb{K}^n$ (gdzie $\mathbb{K}$ oznacza ciało liczb rzeczywistych lub zespolonych) jest $|x^H y| \leqslant \|x\|_2 \|y\|_2$, gdzie $\|x\|_2 \stackrel{\text{def}}{=} \sqrt{x^H x}$. Udowodnimy ponownie tę nierówność dla wektorów w *dowolnej* przestrzeni euklidesowej lub unitarnej:

$$|\langle x, y\rangle| \leqslant \sqrt{\langle x, x\rangle\langle y, y\rangle}.$$

Dowód: Jeśli istnieje liczba a, taka że $x = ay$ lub $y = ax$, to oczywiście zachodzi równość $|\langle x, y\rangle| = \sqrt{\langle x, x\rangle\langle y, y\rangle}$. Wystarczy zatem zbadać przypadek, gdy wektory x i y są liniowo niezależne. Wtedy rozpinają one przestrzeń dwuwymiarową, w której naszemu iloczynowi skalarnemu odpowiada macierz

$$A = \begin{bmatrix} \langle x, x\rangle & \langle y, x\rangle \\ \langle x, y\rangle & \langle y, y\rangle \end{bmatrix}.$$

Na podstawie kryterium Sylvestra $\det A > 0$, czyli $\langle x, x\rangle\langle y, y\rangle > |\langle x, y\rangle|^2$. □

- Funkcja $f\colon V \to \mathbb{R}$, określona wzorem $f(x) = \sqrt{\langle x, x\rangle}$ dla dowolnego ustalonego iloczynu skalarnego ma następujące własności: jest dodatnia dla każdego wektora $x \in V$ z wyjątkiem 0, jest półliniowa, tj. $f(ax) = |a| f(x)$ dla każdego $a \in \mathbb{K}$, $x \in V$ i spełnia nierówność trójkąta: $f(x + y) \leqslant f(x) + f(y)$. Ten ostatni warunek wynika z nierówności Schwarza (dowód — ćwiczenie). Dlatego funkcja f jest normą i zamiast oznaczenia literowego będziemy pisać

$$\|x\| = \sqrt{\langle x, x\rangle}.$$

Przestrzeń liniowa euklidesowa lub unitarna jest więc przestrzenią unormowaną. Oczywiście, przyjmując różne iloczyny skalarne w ustalonej przestrzeni V (lub w innej przestrzeni o tym samym wymiarze) otrzymamy różne przestrzenie unormowane, ale wszystkie te przestrzenie są izomorficzne, a przy tym dla każdego iloczynu skalarnego istnieje baza (a nawet wiele baz), w której macierz tego iloczynu jest jednostkowa; jeśli pewien fakt wyrażalny za pomocą iloczynu skalarnego jest prawdziwy w jednej z tych przestrzeni, to ma on miejsce również we wszystkich pozostałych.

Norma przestrzeni $\mathbb{K}^n$ dana wzorem $\|x\|_2 = \sqrt{x^H x}$, o której wspomniałem, że bywa nazywana normą euklidesową, jest w szczególności związana z iloczynem skalarnym, którego macierz w bazie $e_1, \ldots, e_n$ jest jednostkowa.

- Dowolna norma w przestrzeni liniowej V określa metrykę w tej przestrzeni, za pomocą wzoru $\rho(x, y) = \|x - y\|$. W szczególności może to być metryka euklidesowa, jeśli zastosowana norma jest związana z iloczynem skalarnym. W związku z tą metryką będziemy używać określeń długość wektora x oraz długość odcinka $\overline{ab}$ (która jest odległością punktów a i b, czyli długością wektora $a - b$).

W pewnych zastosowaniach konieczne jest użycie różnych iloczynów skalarnych w tej samej przestrzeni liniowej (na przykład w celu szacowania normy pewnego wektora przez inną normę tego wektora) i wtedy spotyka się oznaczenia w rodzaju $\|x\|_a = \sqrt{\langle x, x\rangle_a}$, które mają na celu identyfikację użytych iloczynów skalarnych i norm. Zauważmy, że *pojęcie długości wektora jest związane z konkretnym iloczynem skalarnym*, i w szczególności w przestrzeni o wymiarze większym niż 1 istnieją wektory x i y oraz iloczyny skalarne $\langle\cdot,\cdot\rangle_a$ i $\langle\cdot,\cdot\rangle_b$, takie że $\|x\|_a = \sqrt{\langle x, x\rangle_a} < \|y\|_a = \sqrt{\langle y, y\rangle_a}$ oraz $\|x\|_b = \sqrt{\langle x, x\rangle_b} > \|y\|_b = \sqrt{\langle y, y\rangle_b}$.

Izometrie przestrzeni euklidesowych i unitarnych

Def. Dowolne przekształcenie f przestrzeni metrycznych $X_1 \to X_2$ nazywa się zanurzeniem izometrycznym, jeśli dla $a, b \in X_1$ jest $\rho_1(a, b) = \rho_2(f(a), f(b))$, tj. odległość każdych dwóch punktów jest niezmiennikiem przekształcenia f. Jeśli istnieje przekształcenie odwrotne do zanurzenia izometrycznego f, to przekształcenie to (a także jego odwrotność) nazywa się izometrią przestrzeni X_1 i X_2.

Twierdzenie: *Niech f oznacza przekształcenie liniowe przestrzeni euklidesowej lub unitarnej V_1 na przestrzeń euklidesową lub unitarną V_2. Przekształcenie f jest zanurzeniem izometrycznym wtedy i tylko wtedy, gdy zachowuje iloczyn skalarny, tj. dla dowolnych wektorów $x, y \in V_1$ zachodzi równość $\langle x, y\rangle_1 = \langle f(x), f(y)\rangle_2$.*

Dowód: Jeśli iloczyn skalarny wektorów jest niezmiennikiem przekształcenia f, to dla dowolnych wektorów $x, y \in V_1$ jest

$$\rho_1(x - y) = \|x - y\| = \sqrt{\langle x - y, x - y\rangle_1} = \sqrt{\langle f(x) - f(y), f(x) - f(y)\rangle_2} = \|f(x) - f(y)\| = \rho_2(f(x), f(y)),$$

a zatem przekształcenie f jest zanurzeniem izometrycznym. Z drugiej strony, jeśli f jest zanurzeniem izometrycznym, czyli zachowuje normę każdego wektora, to fakt, że zachowuje również iloczyn skalarny dowolnych dwóch wektorów wynika z tożsamości polaryzacyjnej. □

Jeśli wymiary przestrzeni V_1 i V_2 są skończone, to przekształcenie odwrotne do liniowego przekształcenia f istnieje wtedy i tylko wtedy, gdy wymiary przestrzeni są równe. W tym przypadku zanurzenie izometryczne f jest izometrią.

Pojęcia kąta

Def. Wektory x i y są prostopadłe (w sensie ustalonego iloczynu skalarnego $\langle\cdot,\cdot\rangle$) jeśli $\langle x, y\rangle = 0$. Prostopadłość wektorów x i y będziemy zapisywać symbolicznie $x \perp y$.

Twierdzenie Pitagorasa:

- *Jeśli wektory x i y są prostopadłe, to* $\|x+y\|^2 = \|x\|^2 + \|y\|^2$.
- *Jeśli wektory $x_1, \ldots, x_k$ są parami prostopadłe, to* $\left\|\sum_{i=1}^k x_i\right\|^2 = \sum_{i=1}^k \|x_i\|^2$.

Dowód tego twierdzenia, jako oczywisty, pominę (należy umieć go przeprowadzić).

Def. Podprzestrzenie V_1 i V_2 przestrzeni euklidesowej (lub unitarnej) V są prostopadłe jeśli $x \perp y$ dla każdego $x \in V_1$ oraz $y \in V_2$.

Def. Niech x i y będą niezerowymi wektorami w przestrzeni euklidesowej V. Na podstawie nierówności Schwarza liczba c, taka że $\langle x, y\rangle = c\|x\|\|y\|$, spełnia warunek $|c| \leqslant 1$. Dlatego istnieje liczba rzeczywista $\alpha \in [0, \pi]$, taka że $c = \cos\alpha$. Liczbę tę nazywamy miarą kąta niezorientowanego między wektorami x i y. Oznaczać ją będziemy symbolem $\operatorname{arc}(x, y)$.

Jest oczywiste, że zawsze $\operatorname{arc}(x, y) = \operatorname{arc}(y, x)$, a ponadto dla dowolnych liczb dodatnich a, b, $\operatorname{arc}(ax, by) = \operatorname{arc}(x, y)$. Łatwo jest dowieść, że miara kąta między wektorami jest niezmiennikiem każdej izometrii. Można też dowieść, że jeśli dla pewnych wektorów x, y, u, v zachodzi równość $\operatorname{arc}(x, y) = \operatorname{arc}(u, v)$, to istnieje taka izometria f oraz liczby dodatnie a i b, że $u = af(x)$ oraz $v = bf(y)$. Na tej podstawie można zdefiniować pojęcie kąta niezorientowanego przy użyciu zasady abstrakcji.

Def. Niech $\sim$ oznacza relację w zbiorze par (nieuporządkowanych) niezerowych wektorów w przestrzeni euklidesowej V, określoną przez warunek $\{x, y\} \sim \{u, v\}$ wtedy i tylko wtedy gdy istnieje izometria $f\colon V \to V$ i liczby dodatnie a, b, takie że $u = af(x)$ oraz $v = bf(y)$. Relacja $\sim$ jest równoważnością; każdą jej klasę abstrakcji nazywamy kątem niezorientowanym. Klasę, której reprezentantem jest para $\{x, y\}$, będziemy oznaczali symbolem $\measuredangle(x, y)$.

Na podstawie rozważań poprzedzających powyższą definicję możemy orzec, że miara kąta niezorientowanego jest funkcją stałą w każdej klasie abstrakcji relacji $\sim$ i jej wartość jest inna w każdej klasie abstrakcji. Ponadto miara kąta niezorientowanego jest niezmiennikiem wszystkich izometrii.

Pojęcie kąta niezorientowanego jest określone w przestrzeni euklidesowej o dowolnym wymiarze. W odróżnieniu od niego, pojęcie kąta zorientowanego możemy określić tylko w przestrzeni dwuwymiarowej. W tym celu wprowadzimy następującą relację „$\sim'$” w zbiorze *uporządkowanych* par niezerowych wektorów z takiej przestrzeni: $(x, y) \sim' (u, v)$ wtedy i tylko wtedy gdy istnieje izometria f, której wyznacznik jest dodatni (czyli równy $+1$) oraz dodatnie liczby a i b, takie że $u = af(x)$ i $v = bf(y)$.

Przypomnijmy, że dwie bazy dowolnej przestrzeni liniowej są zorientowane zgodnie, jeśli wyznacznik macierzy przejścia między nimi jest dodatni, oraz przeciwnie, jeśli jest ujemny. Każdą bazę zorientowaną zgodnie z bazą e_1, e_2 (lub z dowolną ustaloną bazą ortonormalną w przestrzeni innej niż $\mathbb{R}^2$) nazwiemy bazą dodatnio zorientowaną.

Def. Kątem zorientowanym nazywamy każdą klasę abstrakcji relacji $\sim'$ określonej wyżej. Jeśli para (x, y) jest reprezentantem pewnej takiej klasy abstrakcji, to miarą kąta zorientowanego nazywamy liczbę $\phi \in (-\pi, \pi]$, taką że $\cos\phi = \cos\operatorname{arc}(x, y)$ oraz jeśli wektory x, y stanowią bazę dodatnio zorientowaną, to $\sin\phi > 0$, a w przeciwnym razie $\sin\phi \leqslant 0$. Miarę kąta zorientowanego oznaczymy symbolem $\operatorname{Arc}(x, y)$.

W powyższej definicji zawarte jest twierdzenie, że miara kąta zorientowanego jest funkcją stałą w każdej klasie abstrakcji relacji $\sim'$, tj. liczba $\phi = \operatorname{Arc}(x, y)$ nie zależy od wyboru reprezentanta klasy, tj. pary (x, y). Często za miarę kąta zorientowanego wygodnie jest przyjąć każdą z liczb $\operatorname{Arc}(x, y) + 2k\pi$ dla $k \in \mathbb{Z}$.

Niewątpliwie przydadzą się bardziej praktyczne wzory. Zatem, miara kąta niezorientowanego między wektorami x i y jest to liczba α, taka że

$$\cos\alpha = \frac{\langle x, y\rangle}{\|x\|\|y\|}.$$

Miara kąta zorientowanego między wektorami $x = ax_1 + bx_2$ oraz $y = cx_1 + dx_2$ *w przestrzeni dwuwymiarowej*, której baza x_1, x_2 jest dodatnio zorientowana i ortonormalna, jest to liczba ϕ, taka że

$$\cos\phi = \frac{ac + bd}{\sqrt{(a^2+b^2)(c^2+d^2)}}, \qquad \sin\phi = \frac{ad - bc}{\sqrt{(a^2+b^2)(c^2+d^2)}}.$$

Miara kąta zorientowanego jest niezmiennikiem izometrii zachowującej orientację (tj. o dodatnim wyznaczniku). Miara kąta niezorientowanego jest niezmiennikiem każdej izometrii.

Macierze Grama

Niech φ oznacza pewną formę półtoraliniową w przestrzeni V. Dla dowolnych wektorów $x, y \in V$ możemy w rachunkach formalnych napis $y^H \cdot x$ interpretować jako $\varphi(x, y)$ (trzeba tylko pamiętać, jaką formę φ ustaliliśmy; jeśli ktoś nie czuje się z tą notacją zbyt pewnie, to może pisać $y^H \underset{\varphi}{\cdot} x$).

Weźmy dowolne dwie bazy przestrzeni V, mianowicie $X = [x_1, \ldots, x_n]$ i $Y = [y_1, \ldots, y_n]$. Możemy utworzyć macierz $A = [\varphi(x_j, y_i)]_{i,j}$ i wtedy jeśli $a \in \mathbb{K}^n$ jest wektorem współczynników pewnego wektora $x \in V$ w bazie X, zaś $b \in \mathbb{K}^n$ jest wektorem współczynników wektora $y \in V$ w bazie Y, to $\varphi(x, y) = b^H A a$.

W rachunkach symbolicznych powyższą macierz możemy oznaczać symbolem $\varphi(X, Y)$ albo $Y^H \cdot X$. Okazuje się, że wtedy możemy stosować zwykłe reguły przekształcania iloczynów macierzy i takie „mnożenie" jest w szczególności łączne. Na przykład dla $a, b \in \mathbb{K}^n$ mamy

$$\varphi(Xa, Yb) = b^H Y^H \cdot Xa = (Yb)^H \cdot (Xa).$$

Powyższe spostrzeżenia stosują się w szczególności do iloczynów skalarnych. Dla ustalonych baz X i Y będziemy pisać $\langle X, Y\rangle$ albo $Y^H \cdot X$. Na podstawie kryterium Sylvestra macierz $\langle X, X\rangle = X^H \cdot X$ jest dodatnio określona (i w szczególności nieosobliwa). Macierz ta nazywa się macierzą Grama bazy X.

Uwaga: Kropka odróżnia w tej notacji mnożenie polegające na obliczaniu wartości formy φ lub iloczynu skalarnego od zwykłego mnożenia macierzy, w celu uniknięcia niejednoznaczności. W szczególności równość $Y^H \cdot X = Y^H X$ odpowiada sytuacji, gdy macierze X i Y są liczbowe i $\varphi(x, y) = y^H x$.

Rzuty prostopadłe

Def. Rzutem na podprzestrzeń V_1 przestrzeni liniowej V nazywamy przekształcenie liniowe $p\colon V \to V_1 \subset V$, takie że dla każdego $x \in V_1$ $p(x) = x$.

Przestrzeń V_1 nazywamy podprzestrzenią niezmienniczą przekształcenia p. W szczególności przekształcenie tożsamościowe jest rzutem, którego podprzestrzenią niezmienniczą jest cała przestrzeń V; inny trywialny przykład rzutu to przekształcenie zerowe.

Rzuty są przekształceniami idempotentnymi, tj. $p = p^2$. Zauważmy, że w określeniu rzutu nie występuje pojęcie iloczynu skalarnego. Istotnie, możemy określić rzut w dowolnej przestrzeni liniowej. W tym celu wybieramy podprzestrzenie V_1 i V_2, takie że $V = V_1 \oplus V_2$. Dowolny wektor x możemy jednoznacznie przedstawić w postaci sumy wektorów $x_1 \in V_1$ i $x_2 \in V_2$. Przekształcenie p, które wektorowi x przyporządkowuje wektor x_1 jest rzutem na podprzestrzeń V_1. Przekształcenie $\mathrm{id} - p$ jest również rzutem, na podprzestrzeń V_2.

Jeśli wymiar podprzestrzeni V_1 jest równy r, to rząd macierzy A rzutu na tę podprzestrzeń jest równy r. Macierz ta musi spełniać warunek $A^2 = A$. W ogólności możemy wybrać dowolną bazę przestrzeni V_1 i uzupełnić ją (przez dołączenie wektorów rozpinających podprzestrzeń V_2) do bazy całej przestrzeni V. Macierz dowolnego rzutu na podprzestrzeń V_1 w takiej bazie składa się z r początkowych kolumn macierzy jednostkowej i $n - r$ kolumn będących dowolnymi kombinacjami liniowymi początkowych r kolumn.

Def. Rzutem prostopadłym na podprzestrzeń V_1 przestrzeni euklidesowej lub unitarnej V nazywamy rzut p, taki że dla każdego wektora $x \in V$ wektory $p(x)$ i $x - p(x)$ są prostopadłe.

Można udowodnić (ćwiczenie), że dla dowolnej podprzestrzeni V_1 istnieje dokładnie jeden rzut prostopadły na tę podprzestrzeń.

Twierdzenie: *Rzut prostopadły p przestrzeni V na podprzestrzeń V_1, której bazą jest $X = [x_1, \ldots, x_r]$, można zapisać w postaci wyrażenia macierzowego $p(x) = X(X^H \cdot X)^{-1} X^H \cdot x$, gdzie wyrażenie $X^H \cdot X$ oznacza macierz Grama bazy X.*

Dowód: Wystarczy dowieść, że jeśli $p = X(X^H \cdot X)^{-1} X^H$ to $\operatorname{im} p = V_1$, $p^2 = p$ i dla każdego wektora x prostopadłego do podprzestrzeni V_1 (czyli do wszystkich wektorów w tej podprzestrzeni) $p(x) = 0$. Obliczając

$$p(X) = X(X^H \cdot X)^{-1} X^H \cdot X = X$$

przekonujemy się, że obrazem bazy X podprzestrzeni V_1 jest baza X, skąd pierwsze dwa warunki wynikają natychmiast. Warunek $y \perp V_1$ jest równoważny warunkom $y \perp x_i$ dla $i = 1, \ldots, r$, a zatem $X^H \cdot y = 0$, skąd wynika równość $p(y) = 0$. $\square$

Jeśli $V = \mathbb{K}^n$ i stosujemy iloczyn skalarny $\langle x, y \rangle = y^H x$, to macierz rzutu prostopadłego na podprzestrzeń $\operatorname{im} A$, gdzie $A \in \mathbb{K}^{n,r}$, $\operatorname{rank} A = r$, jest równa $A(A^H A)^{-1} A^H$.

Ważny przypadek szczególny rzutu prostopadłego to rzut na podprzestrzeń jednowymiarową, rozpiętą przez ustalony wektor $v \neq 0$. W tym przypadku macierz Grama ma wymiary 1×1; jej jedyny współczynnik jest równy $\langle v, v \rangle$ i rozpatrywany rzut prostopadły możemy przedstawić wzorem

$$p(x) = v(v^H \cdot v)^{-1} v^H \cdot x = v \frac{\langle x, v \rangle}{\langle v, v \rangle}.$$

Jeśli $v \in \mathbb{K}^n$ i $\langle x, y \rangle = y^H x$, to macierz rzutu przestrzeni $\mathbb{K}^n$ na podprzestrzeń $\operatorname{lin}\{v\}$ jest równa $\frac{1}{v^H v} v v^H$. W przypadku rzeczywistym zamiast hermitowskiego sprzężenia we wszystkich podanych tu wzorach występuje transpozycja.

Zadania i problemy

1. Podaj (i udowodnij) warunek konieczny i dostateczny, aby macierz hermitowska A była <u>ujemnie określona</u>, tj. aby dla każdego wektora $x \neq 0$ było $x^H A x < 0$.
2. Udowodnij (na podstawie nierówności Schwarza) nierówność trójkąta dla norm generowanych przez iloczyny skalarne.
3. Na wszelki wypadek jednak udowodnij twierdzenie Pitagorasa.
4. Udowodnij, że macierz $\langle X, Y \rangle$, dla dowolnych baz $X = [x_1, \ldots, x_n]$ i $Y = [y_1, \ldots, y_n]$ i dowolnego iloczynu skalarnego jest nieosobliwa. Czy zawsze jest ona dodatnio określona?
5. Jaki warunek spełnia wartość wyznacznika dowolnego przekształcenia, które jest izometrią?
6. Udowodnij podane na wykładzie wzory, które pozwalają obliczyć miarę kąta zorientowanego między wektorami.

   <u>Wskazówka:</u> Trzeba udowodnić, że suma kwadratów wyrażeń opisujących sinus i cosinus jest równa 1, a ponadto że znak obliczonego w ten sposób sinusa zgadza się z definicją.
7. Oblicz miarę kąta niezorientowanego i miarę kąta zorientowanego między wektorami $[0, 1]^T$ i $[-1, -1]^T$, przy założeniu, że

   a) macierz iloczynu skalarnego w bazie e_1, e_2 jest jednostkowa,

   b) macierz iloczynu skalarnego w bazie e_1, e_2 jest równa $\begin{bmatrix} 1 & -1 \\ -1 & 2 \end{bmatrix}$.
8. Znajdź macierz Grama iloczynu skalarnego w $\mathbb{R}^3$, takiego że baza

   $$\left\{ \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \right\}$$

   jest ortonormalna (w sensie tego iloczynu).
9. Znajdź macierz rzutu prostopadłego przestrzeni $\mathbb{R}^3$ na podprzestrzeń

   a) rozpiętą przez wektor $[3, 4, 0]^T$,

   b) rozpiętą przez wektory $[3, 4, 0]^T$ i $[8, -6, 1]^T$,

   przy założeniu, że iloczyn skalarny jest dany wzorem $\langle x, y \rangle = y^T x$.
10. Niech

    $$\langle f, g \rangle = \int_{-1}^{1} f(x) g(x) \, dx$$

    Znajdź obraz wielomianu $w(x) = x^2 + x$ w rzucie prostopadłym na podprzestrzeń rozpiętą przez $a(x) = 1$.

11. Znajdź macierz rzutu prostopadłego przestrzeni $\mathbb{R}[x]_2$ na przestrzeń $\mathbb{R}[x]_1$ w bazie $[1, x, x^2]$ dla iloczynu skalarnego

$$\langle f, g \rangle = \int_{-1}^{1} f(x)g(x)(1 + x)\,dx$$

12. Napisz procedurę obliczania rzutu wektora $\boldsymbol{x} \in \mathbb{R}^n$ na podprzestrzeń jednowymiarową rozpiętą przez wektor $\boldsymbol{v}$, taki że $\boldsymbol{v}^T\boldsymbol{v} = 1$.
13. Udowodnij, że dla dowolnej podprzestrzeni V_1 istnieje dokładnie jeden rzut prostopadły przestrzeni euklidesowej lub unitarnej V na tę podprzestrzeń.
14. Niech $A, B \in \mathbb{C}^{n,n}$ będą macierzami hermitowskimi, nieujemnie określonymi. Udowodnij, że jeśli dla dowolnego $s \in [0, 1]$ macierz $(1 - s)A + sB$ jest dodatnio określona, to dla każdego $t \in (0, 1)$ macierz $(1 - t)A + tB$ jest dodatnio określona.

Bazy ortogonalne

Niech $(V, \langle\cdot,\cdot\rangle)$ będzie ustaloną przestrzenią euklidesową lub unitarną, a V_1 jej podprzestrzenią. Jak wiemy, istnieje dokładnie jeden rzut prostopadły przestrzeni V na podprzestrzeń V_1. Jeśli rzut ten oznaczymy symbolem p_1, to przekształcenie liniowe $p_2 = \mathrm{id} - p_1$ jest rzutem prostopadłym na podprzestrzeń $V_2 \subset V$, taką że

- $\dim V_1 + \dim V_2 = \dim V$, $\dim V_1 \cap V_2 = 0$, a zatem $V_1 \oplus V_2 = V$,
- $x \in V_1$, $y \in V_2 \Rightarrow \langle x, y\rangle = 0$.

W przypadku gdy dowolne podprzestrzenie V_1 i V_2 spełniają powyższe warunki mówimy, że przestrzeń V jest ich sumą ortogonalną.

Ponieważ dla dowolnej podprzestrzeni V_1 istnieje (dokładnie jeden) rzut prostopadły na tę podprzestrzeń, więc dla dowolnej przestrzeni V_1 istnieje (dokładnie jedna) przestrzeń prostopadła do niej, V_2, taka że przestrzeń V jest sumą ortogonalną podprzestrzeni V_1 i V_2.

Pojęcie sumy ortogonalnej można rozszerzyć na dowolną liczbę (a nawet nieskończenie wiele, ale nie jest to trywialne) podprzestrzeni parami prostopadłych. Pojęcie sumy ortogonalnej określa się też w przestrzeniach ortogonalnych innych niż euklidesowe i unitarne, tj. których forma półtoraliniowa nie jest iloczynem skalarnym (odpowiednia forma kwadratowa nie jest dodatnio określona) — postępowanie w dowodzie twierdzenia Sylvestra przedstawia pewien sposób znalezienia takich podprzestrzeni.

W szczególności przestrzeń V można przedstawić w postaci sumy ortogonalnej podprzestrzeni jednowymiarowych. Baza przestrzeni V, która jest sumą baz takich podprzestrzeni, nazywa się bazą ortogonalną. Każde dwa elementy takiej bazy są wektorami prostopadłymi.

Poniższe twierdzenie na temat baz w przestrzeniach z iloczynem skalarnym stanowi analogię (a właściwie wzmocnienie) ogólnego twierdzenia dotyczącego istnienia baz w przestrzeniach liniowych.

Twierdzenie: *W dowolnej przestrzeni euklidesowej lub unitarnej istnieją bazy ortogonalne. Dowolny ortogonalny układ wektorów różnych od wektora zerowego możemy rozszerzyć do bazy ortogonalnej.*

Uwaga: Twierdzenie to dla przestrzeni nieskończenie wymiarowych zawiera kilka subtelności, których nie można pominąć, nawet jeśli teraz nie będziemy się wgłębiali w ten temat. Po pierwsze, dotyczy ono innego pojęcia bazy.

Baza Hamela to układ wektorów taki, że dowolny element przestrzeni ma jednoznaczne przedstawienie w postaci kombinacji liniowej (skończonej) tych wektorów. Baza Schaudera to układ wektorów w przestrzeni *unormowanej*, taki że każdy element tej przestrzeni ma jednoznaczne przedstawienie w postaci skończonej kombinacji liniowej albo nieskończonego szeregu (który musi być zbieżny w sensie używanej normy). Myśmy zajmowali się dotąd bazami Hamela, ale pojęcie bazy ortogonalnej w przestrzeni nieskończenie wymiarowej jest związane z bazami Schaudera. Oba pojęcia w przestrzeni skończenie wymiarowej są identyczne. Jeśli układ wektorów, który chcemy rozszerzyć do bazy ortogonalnej, jest nieskończony, to aby to było możliwe, trzeba jeszcze założyć, że długości (tj. normy) tych wektorów są zawarte między pewnymi dwiema liczbami dodatnimi.

Uwaga: Przymiotniki „prostopadły” i „ortogonalny” można w zasadzie stosować zamiennie. Jest jeszcze jeden przymiotnik występujący w tym znaczeniu, mianowicie „normalny” (ale później jeszcze pojawi się pojęcie „baza ortonormalna”; druga część przymiotnika w tym terminie oznacza co innego).

Ortogonalizacja Grama–Schmidta

Zajmiemy się następującym zadaniem: Mając bazę $x_1, \dots, x_n$ przestrzeni euklidesowej (lub unitarnej) V, należy znaleźć bazę $y_1, \dots, y_n$, taką że

- $y_i \perp y_k$ dla $i \neq k$ (baza $y_1, \dots, y_n$ jest ortogonalna),
- $\mathrm{lin}\{x_1, \dots, x_k\} = \mathrm{lin}\{y_1, \dots, y_k\}$ dla $k = 1, \dots, n$, oraz
- $\langle x_k, y_k\rangle > 0$ dla $k = 1, \dots, n$.

W szczególności przekonamy się, że to zadanie ma jednoznaczne (z dokładnością do n stałych dodatnich) rozwiązanie.

Drugi z postawionych warunków oznacza, że dla każdego k wektor y_k jest kombinacją liniową wektorów $x_1, \dots, x_k$, albo, równoważnie, wektorów $y_1, \dots, y_{k-1}, x_k$. Aby spełnić trzeci warunek, wystarczy przyjąć, że współczynnik tej kombinacji liniowej przy x_k jest dodatni — najbardziej oczywisty wybór to 1.

Wektory $y_1, \dots, y_n$ będziemy konstruować kolejno. Wektor y_k otrzymamy dokonując rzutowania prostopadłego wektora x_k na podprzestrzeń prostopadłą do przestrzeni rozpiętej przez wektory $x_1, \dots, x_{k-1}$. Łatwo możemy się przekonać, że w ten sposób rzeczywiście dostaniemy wektory spełniające postawione warunki. Istotnie, mamy wtedy

$$x_k = y_k + z_k,$$

gdzie $z_k \in \mathrm{lin}\{x_1, \ldots, x_{k-1}\} = \mathrm{lin}\{y_1, \ldots, y_{k-1}\}$ oraz $y_k \perp z_k$. Jest oczywiste, że wektor y_k jest prostopadły do wektorów $y_1, \ldots, y_{k-1}$ (czyli każdy z nich jest prostopadły do y_k — relacja prostopadłości wektorów jest symetryczna), a ponadto

$$\langle x_k, y_k\rangle = \langle y_k + z_k, y_k\rangle = \langle y_k, y_k\rangle + \langle z_k, y_k\rangle > 0.$$

Opisany rzut możemy konstruować na wiele sposobów, które różnią się kosztem realizacji. W szczególności, niech $X = [x_1, \ldots, x_{k-1}]$ i $Y = [y_1, \ldots, y_{k-1}]$. Wtedy (zgodnie z wcześniej udowodnionym twierdzeniem)

$$z_k = X(X^H \cdot X)^{-1}X^H \cdot x_k = Y(Y^H \cdot Y)^{-1}Y^H \cdot x_k.$$

Jest jasne, że mniej pracy wymaga zastosowanie bazy Y, ponieważ jej macierz Grama jest diagonalna, co znakomicie ułatwia znalezienie jej odwrotności. Zatem, na podstawie ostatniego wyrażenia możemy napisać

$$y_k = x_k - z_k = x_k - \sum_{i=1}^{k-1} y_i \frac{\langle x_k, y_i\rangle}{\langle y_i, y_i\rangle}$$

Na podstawie powyższego wzoru możemy napisać następującą procedurę, która realizuje algorytm ortogonalizacji Grama–Schmidta:

```
for k := 1 to n do begin
  y := x_k;
  for i := 1 to k−1 do
    y := y − y_i ⟨x_k,y_i⟩/⟨y_i,y_i⟩;
  y_k := y
end;
```

Powyższa procedura może być zastosowana w celu znalezienia bazy ortogonalnej dowolnej przestrzeni euklidesowej lub unitarnej o wymiarze n. Stosując ją należy obliczyć $\frac{1}{2}n(n+1)$ iloczynów skalarnych.

Związek między bazami $X = [x_1, \ldots, x_n]$ i $Y = [y_1, \ldots, y_n]$ możemy przedstawić w postaci równania macierzowego

$$X = YR,$$

w którym macierz $R = (r_{ij})_{i,j} \in \mathbb{K}^{n,n}$ jest trójkątna górna: $r_{ij} = 0$ dla $i > j$ oraz $r_{ij} = \frac{\langle x_j, y_i\rangle}{\langle y_i, y_i\rangle}$ dla $i \leqslant j$. Współczynniki diagonalne r_{ii} macierzy R są równe 1, a zatem $\det R = 1$.

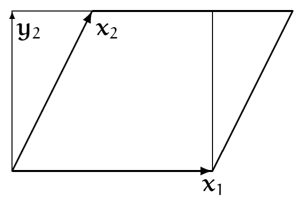
Objętość równoległościanu

Def. Niech $x_1, \ldots, x_k$ będą ustalonymi wektorami w pewnej przestrzeni euklidesowej. Równoległościanem k-wymiarowym rozpiętym przez te wektory nazywamy zbiór $\{\, x = \sum_{i=1}^{k} a_i x_i \colon a_i \in [0, 1]\,\}$.

Jednym z najważniejszych zadań geometrii i analizy jest badanie miary różnych zbiorów (w szczególności figur w przestrzeniach euklidesowych). Na przykład miarą odcinka może być długość tego odcinka. Innymi miarami są funkcje, które wyrażają pole figury płaskiej lub objętość bryły. Nie dotykając teorii miary powiedzmy, że:

- Miara w przestrzeni V jest rzeczywistą funkcją nieujemną, której dziedziną jest pewien podzbiór zbioru podzbiorów przestrzeni V (mogą istnieć tzw. zbiory niemierzalne).
- Znając miarę pewnego zbioru A możemy podzielić go na rozłączne, mierzalne podzbiory, a następnie poddać te podzbiory takim przekształceniom, które nie zmieniają miary. Jeśli ich obrazy są rozłączne, to ich suma jest zbiorem mierzalnym, którego miara jest równa mierze zbioru A. W ten sposób starożytni Grecy obliczali pola i objętości figur, tnąc je (w wyobraźni) na kawałki i układając z nich inne figury, których miary znali.
- W przestrzeni euklidesowej można określić pojęcie prostopadłościanu k-wymiarowego (jest to równoległościan rozpięty przez k wektorów parami prostopadłych) i określić miarę każdego takiego prostopadłościanu jako iloczyn długości rozpinających go wektorów. Tak określoną miarę można uogólnić na inne podzbiory tej przestrzeni (to jest bardzo nietrywialna konstrukcja, ponieważ mierzone zbiory mogą mieć bardzo skomplikowany opis), przez co powstaje tzw. k-wymiarowa miara Lebesgue'a, zgodna z intuicyjnym pojęciem „długości" (dla $k = 1$), „pola" (dla $k = 2$) „objętości" (dla $k = 3$) itd. Każda miara Lebesgue'a jest funkcją ciągłą (ale ścisłe wyjaśnienie co to oznacza jednak pominę) i na zbiorach mierzalnych (w sensie tej miary) identycznych z dokładnością do izometrii przyjmuje tę samą wartość.

Na obrazku obok mamy przykład równoległościanu dwuwymiarowego, rozpiętego przez wektory x_1 i x_2. Jest on równoważny przez rozkład (jak powiedzieliby starożytni Grecy, oczywiście w tłumaczeniu z języka polskiego) równoległościanowi rozpiętemu przez wektory x_1 i y_2; oba równoległościany mają więc tę samą miarę (objętość dwuwymiarową).



Wektor y_2 jest obrazem x_2 w rzucie prostopadłym na podprzestrzeń prostopadłą do x_1. Ponieważ z równoległoboku (równoległościanu dwuwymiarowego) przez rozkład otrzymaliśmy prostokąt, którego miara jest równa iloczynowi długości boków, więc taka też jest miara równoległoboku wyjściowego. Widzimy też, że takie samo pole będą miały wszystkie równoległoboki rozpięte przez wektory x_1 i $x_2 + ax_1$ dla dowolnego $a \in \mathbb{R}$.

Podobnie wygląda sytuacja dla równoległościanów o innym wymiarze. Jeśli pewien równoległościan jest rozpięty przez wektory $x_1, \ldots, x_k$, to dodanie do któregoś z tych wektorów dowolnej kombinacji liniowej pozostałych wektorów da nam równoległościan, którego k-wymiarowa objętość jest taka sama (równoważność przez rozkład obu równoległościanów można sprawdzić bezpośrednim rachunkiem).

Rozważmy ponownie proces ortogonalizacji Grama–Schmidta. Przypuśćmy, że wektory $x_1, \ldots, x_k$ są liniowo niezależne, a zatem stanowią bazę pewnej przestrzeni liniowej. W wyniku ortogonalizacji dostaniemy pewien układ wektorów $y_1, \ldots, y_k$ i łatwo jest przekonać się, że w świetle poczynionych przed chwilą uwag równoległościany rozpięte przez te dwa układy wektorów są równoważne przez rozkład. Mają one zatem tę samą k-wymiarową objętość. Drugi z tych równoległościanów jest k-wymiarowym prostopadłościanem i jego miara jest równa iloczynowi długości krawędzi (tj. iloczynowi długości parami prostopadłych wektorów $y_1, \ldots, y_k$).

Jeśli wektory $x_1, \ldots, x_k$ są liniowo zależne, tj. dla pewnego j wektor x_j jest kombinacją liniową wektorów $x_1, \ldots, x_{j-1}$, to procedura ortogonalizacji da w wyniku wektor $y_j = 0$; objętość równoległościanu rozpiętego przez te wektory jest równa 0.

Niech $X = [x_1, \ldots, x_k]$, $Y = [y_1, \ldots, y_k]$ (uwaga: nie zakładamy liniowej niezależności wektorów $x_1, \ldots, x_k$ ani otrzymanych z nich $y_1, \ldots, y_k$). Istnieje macierz trójkątna górna R o wymiarach $k \times k$, taka że $X = YR$ i $\det R = 1$. Macierz Grama układu $x_1, \ldots, x_k$ jest równa

$$\langle X, X\rangle = X^H \cdot X = (YR)^H \cdot YR = R^H Y^H \cdot YR = R^H \langle Y, Y\rangle R.$$

Zatem $\det\langle X, X\rangle = \det\langle Y, Y\rangle$. Ale macierz $\langle Y, Y\rangle$ jest diagonalna i jej współczynniki diagonalne są równe $\langle y_i, y_i\rangle = \|y_i\|^2$ (a zatem są to kwadraty długości wektorów y_i w sensie normy związanej z iloczynem skalarnym $\langle\cdot,\cdot\rangle$). Stąd miara (objętość k-wymiarowa) równoległościanu P rozpiętego przez wektory $x_1, \ldots, x_k$ jest równa

$$\mu_k(P) = \sqrt{\det\langle X, X\rangle}$$

Zauważmy (wynika to z twierdzenia Pitagorasa), że jeśli układ $y_1, \ldots, y_k$ powstał z $x_1, \ldots, x_k$ w drodze ortogonalizacji, to $\|x_i\| \geqslant \|y_i\|$ dla $i = 1, \ldots, k$, przy czym równość $\|x_i\| = \|y_i\|$ zachodzi tylko wtedy, gdy $x_i = y_i$. Na tej podstawie możemy napisać

$$\det\langle X, X\rangle = \det\langle Y, Y\rangle = \prod_{i=1}^{k}\langle y_i, y_i\rangle = \prod_{i=1}^{k}\|y_i\|^2 \leqslant \prod_{i=1}^{k}\|x_i\|^2.$$

Udowodniliśmy w ten sposób następującą nierówność Hadamarda: jeśli $X = [x_1, \ldots, x_k]$, to

$$\det\langle X, X\rangle \leqslant \prod_{i=1}^{k}\|x_i\|^2.$$

Nierówność ta jest ostra z wyjątkiem przypadku, gdy wektory $x_1, \ldots, x_k$ są parami prostopadłe, lub gdy któryś z nich jest wektorem zerowym.

Iloczyn skalarny i funkcjonały liniowe

Niech $(V, \langle\cdot,\cdot\rangle)$ oznacza skończenie wymiarową przestrzeń euklidesową lub unitarną. Weźmy pod uwagę wyrażenie $\langle x, y\rangle$. Przy ustalonym wektorze y opisuje ono funkcjonał liniowy; zatem dowolnemu wektorowi y odpowiada dokładnie jeden funkcjonał liniowy f określony wzorem $f(x) = \langle x, y\rangle$.

Udowodnimy, że przyporządkowanie elementom przestrzeni V (wektorom) elementów przestrzeni sprzężonej V^* (funkcjonałów), określone powyższym wzorem, jest wzajemnie jednoznaczne. Inaczej mówiąc, dla każdego funkcjonału $f \in V^*$ istnieje dokładnie jeden wektor y, taki że dla każdego $x \in V$ $f(x) = \langle x, y\rangle$.

W przestrzeni V istnieje baza ortogonalna $y_1, \ldots, y_n$. Niech $z_i = \frac{1}{\langle y_i, y_i\rangle} y_i$. Łatwo jest przekonać się, że układ funkcjonałów $f_1, \ldots, f_n$, takich że $f_i(x) = \langle x, z_i\rangle$ dla $i = 1, \ldots, n$, jest bazą przestrzeni V^*, sprzężoną z bazą $y_1, \ldots, y_n$. Wystarczy sprawdzić, że jeśli $Y = [y_1, \ldots, y_n]$ i $Z = [z_1, \ldots, z_n]$ to $\langle Y, Z\rangle = I_n$. Ponieważ elementy bazy przestrzeni V^* utożsamiliśmy z wektorami z przestrzeni V, więc możemy to samo zrobić z każdym elementem przestrzeni V^*. □

Tak więc funkcjonały liniowe na przestrzeni V możemy utożsamić z wektorami z tej przestrzeni. Izomorfizm przestrzeni V i V^*, w trochę węższym kontekście, rozpatrywaliśmy w pierwszym semestrze. Obecnie, znając pojęcie iloczynu skalarnego, możemy dokonać geometrycznej interpretacji funkcjonału liniowego. Zbiorem miejsc zerowych dowolnego niezerowego funkcjonału f jest pewna podprzestrzeń o wymiarze $n - 1$ (hiperpłaszczyzna). Jest ona zbiorem wektorów prostopadłych (w sensie iloczynu skalarnego przestrzeni V) do wektora y, który utożsamiliśmy z funkcjonałem f. Wektor y prostopadły do hiperpłaszczyzny $U \subset V$ zazwyczaj jest nazywany wektorem normalnym hiperpłaszczyzny U.

Uwaga: Nie w każdej nieskończenie wymiarowej przestrzeni z iloczynem skalarnym dowolny funkcjonał liniowy można utożsamić z pewnym wektorem; można to zrobić w tzw. przestrzeni Hilberta, czyli przestrzeni zupełnej (której elementami są granice wszystkich ciągów Cauchy'ego zawartych w tej przestrzeni) — orzeka o tym twierdzenie Riesza, ale to temat na wykład z innego przedmiotu.

Iloczyn wektorowy w $\mathbb{R}^3$

W wykładzie o wyznacznikach znajduje się definicja iloczynu wektorowego $n-1$ wektorów w przestrzeni $\mathbb{K}^n$. Zbadamy, jak ta definicja ma się do bardziej znanej definicji „geometrycznej" iloczynu wektorowego dwóch wektorów w przestrzeni euklidesowej $\mathbb{R}^3$ z iloczynem skalarnym $\langle x, y\rangle = y^T x$.

Iloczyn wektorowy dwóch wektorów, $x, y \in \mathbb{R}^3$, zdefiniowaliśmy jako funkcjonał liniowy f określony wzorem $f(z) = \det[x, y, z]$. Utożsamimy funkcjonał f z takim wektorem n, że dla każdego wektora z jest $f(z) = \langle z, n\rangle$. Wiemy, że wektor n jest prostopadły do wektorów x i y (kto nie wie, jak to udowodnić, niech powie to teraz, lub zamilknie na wieki).

Niech z oznacza wektor prostopadły do wektorów x i y (o których założymy, że są liniowo niezależne) i taki, że $\|z\|^2 = \langle z, z\rangle = 1$ (wektor taki istnieje, bo potrafimy go skonstruować; dokładniej, istnieją dwa takie wektory). Wektor z ma kierunek wektora n, skąd wynika, że $|\langle z, n\rangle| = \|z\|\|n\| = \|n\|$. Stąd $\|n\|^2 = \langle z, n\rangle^2 = \det G$, gdzie G oznacza macierz Grama układu wektorów x, y, z. Mamy

$$G = \langle [x, y, z], [x, y, z]\rangle = \begin{bmatrix} \langle x, x\rangle & \langle y, x\rangle & \langle z, x\rangle \\ \langle x, y\rangle & \langle y, y\rangle & \langle z, y\rangle \\ \langle x, z\rangle & \langle y, z\rangle & \langle z, z\rangle \end{bmatrix} = \begin{bmatrix} \langle x, x\rangle & \langle y, x\rangle & 0 \\ \langle x, y\rangle & \langle y, y\rangle & 0 \\ 0 & 0 & 1 \end{bmatrix},$$

skąd wynika, że wyznacznik tej macierzy jest równy

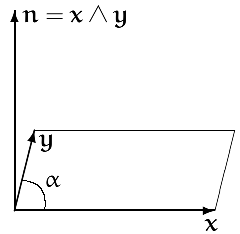
$$\|n\|^2 = \det G = \det \begin{bmatrix} \langle x, x\rangle & \langle y, x\rangle \\ \langle x, y\rangle & \langle y, y\rangle \end{bmatrix} = \langle x, x\rangle\langle y, y\rangle - \langle x, y\rangle^2$$

Jak widzimy, długość wektora n jest równa mierze (polu) równoległoboku rozpiętego przez wektory x i y. Korzystając w dalszych rachunkach z równości $\langle x, y\rangle = \|x\|\|y\| \cos\alpha$, gdzie $\alpha = \operatorname{arc}(x, y)$, możemy obliczyć

$$\|n\|^2 = \|x\|^2\|y\|^2(1 - \cos^2\alpha) = \|x\|^2\|y\|^2 \sin^2\alpha.$$

Zatem definicja iloczynu wektorowego podana wcześniej na wykładzie, w przypadku iloczynu wektorowego dwóch wektorów w $\mathbb{R}^3$, jest równoważna znanej definicji geometrycznej:

Def. Iloczynem wektorowym wektorów $x, y \in \mathbb{R}^3$ jest wektor $n \in \mathbb{R}^3$ prostopadły do x i y, którego długość jest równa $\|n\| = \|x\|\|y\| \sin\alpha$, gdzie $\alpha = \operatorname{arc}(x, y)$ (uwaga: zawsze $\sin\alpha \geqslant 0$) i jeśli wektory x i y są liniowo niezależne, to układ x, y, n jest zorientowany dodatnio.



Zadania i problemy

1. Udowodnij, że jeśli wektory $x_1, \ldots, x_k \in V$ są różne od wektora zerowego i parami prostopadłe, tj. $\langle x_i, x_j\rangle = 0$ dla $i \neq j$, to są liniowo niezależne.
2. Sprawdź bezpośrednim rachunkiem, że jeśli $y_i \perp y_j$ dla $i, j < k$, $i \neq j$, to dla każdego wektora x zachodzi równość
$$\langle y_j, x - \sum_{i=1}^{k-1} y_i \frac{\langle x, y_i\rangle}{\langle y_i, y_i\rangle}\rangle = 0.$$
Przy tym jeśli wektory $y_1, \ldots, y_{k-1}, x$ są liniowo niezależne, to drugi argument iloczynu skalarnego w wyrażeniu po lewej stronie nie jest wektorem zerowym.
3. Zastosuj ortogonalizację Grama–Schmidta do bazy $1, x, x^2$ przestrzeni $\mathbb{R}[x]_2$ z iloczynem skalarnym określonym za pomocą wzoru
$$\langle f, g\rangle = \int_0^1 f(x)g(x)\,dx.$$
4. Znajdź wyrażenie macierzowe, które opisuje bazę sprzężoną z bazą $X = [x_1, \ldots, x_n]$, wiedząc że $X = YR$, gdzie macierz $Y = [y_1, \ldots, y_n]$ reprezentuje bazę ortogonalną.
5. Udowodnij, że równoległościan jest zbiorem wypukłym.
6. Oblicz objętość (trójwymiarową) równoległościanu w przestrzeni euklidesowej $\mathbb{R}^4$ z iloczynem skalarnym $\langle x, y\rangle = y^T x$, rozpiętego przez wektory $[1, 2, 2, 0]^T$, $[-2, 2, 0, 1]$, $[1, -1, 1, -1]$. Oblicz długości krawędzi i pole powierzchni ścian tego równoległościanu.
7. Oblicz kąty między wektorami rozpinającymi równoległościan z poprzedniego zadania (wystarczy obliczyć cosinusy tych kątów).
8. Niech $x = [x_1, x_2, x_3]$ oznacza ustalony wektor w $\mathbb{R}^3$. Znajdź macierz X, taką że dla każdego wektora $y \in \mathbb{R}^3$ zachodzi równość $Xy = x \wedge y$.

Wskazówka: Oblicz wektory $x \wedge e_1$, $x \wedge e_2$ i $x \wedge e_3$ i skorzystaj z tego, że układ e_1, e_2, e_3 jest bazą.

9. Udowodnij równoważność przez rozkład równoległościanu trójwymiarowego rozpiętego przez dowolne wektory x_1, x_2, x_3 i równoległościanu rozpiętego przez x_1, $x_1 + ax_2$, $x_3 + bx_1 + cx_2$ dla dowolnych $a, b, c \in \mathbb{R}$.

   Wskazówka: Trzeba zacząć od przypadku $a \in [0, 1]$, $b = c = 0$, a następnie uogólnić wynik dla dowolnego a. Brzeg równoległościanu i jego kawałków należy przy tym zaniedbać (objętość brzegu jest równa 0).

10. Objętość k-wymiarowa sympleksu S, położonego w przestrzeni $\mathbb{R}^n$ z iloczynem skalarnym $\langle x, y\rangle = y^T x$, którego wierzchołkami są punkty $p_0, \dots, p_k$, jest równa

$$\mu_k(S) = \frac{1}{k!}\sqrt{\det X^T X},$$

gdzie $X = [x_1, \dots, x_k]$ oraz $x_i = p_i - p_0$ dla każdego i. Równoległościan k-wymiarowy można podzielić na k! sympleksów o jednakowej objętości, ale nie tędy droga, aby powyższy wzór wyprowadzić. Rzecz w tym, że dla $k > 2$ otrzymane sympleksy nie są równoważne przez rozkład, tj. nie można żadnego z nich podzielić na skończenie wiele części, z których można by złożyć każdy inny. Dlatego powyższy wzór otrzymuje się w wyniku przejścia granicznego (rozpatrując podziały sympleksu na nieskończenie wiele części), co oznacza obliczenie całki.

Udowodnij powyższy wzór, korzystając z indukcji. W tym celu rozważ sympleks k-wymiarowy. Każda jego ściana jest sympleksem $k-1$-wymiarowym; wybierz jedną ze ścian i podziel sympleks na „plasterki” o grubości dx za pomocą afinicznych przestrzeni $k-1$-wymiarowych równoległych do wybranej ściany. Objętość k-wymiarowa każdego plasterka jest w przybliżeniu (tym lepszym im mniejsze dx) równa $dx \cdot$ objętość $k-1$-wymiarowa „podstawy” plasterka. Wystarczy teraz znaleźć odpowiednią funkcję podcałkową i obliczyć całkę.

11. Udowodnij, że objętość k-wymiarowa sympleksu S o wierzchołkach w punktach $p_0, \dots, p_k \in \mathbb{R}^n$ można obliczyć na podstawie wzoru

$$\mu_k(S) = \frac{1}{k!}\sqrt{\det X^T X},$$

w którym występuje macierz $X = [x_0, \dots, x_k] \in \mathbb{R}^{n+1,k+1}$, której kolumny powstały przez dopisanie liczby 1 jako dodatkowej współrzędnej do punktów $p_0, \dots, p_k$. Udowodnij następnie, że jeśli $Y = [y_0, \dots y_k]$ gdzie $y_i = w_i x_i$ (liczby w_i są dowolne, różne od 0, czyli punkty $p_0, \dots, p_k$ reprezentujemy za pomocą dowolnych macierzy współrzędnych jednorodnych), to

$$\mu_k(S) = \frac{1}{k!}\frac{\sqrt{\det Y^T Y}}{|\prod_{i=0}^k w_i|}.$$

Wyjaśnij, jak można obliczyć objętość równoległościanu na podstawie współrzędnych jednorodnych jego wierzchołków.