Uniwersytet Wrocławski
Wydział Matematyki i Informatyki
Instytut Matematyczny
specjalność: matematyka nauczycielska

Barbara Szczepańska

Skrypt z Algebry Liniowej 2

Praca magisterska
napisana pod kierunkiem
dr. hab. prof. Jacka Świątkowskiego

Wrocław 2007

Podziękowania

Dziękuję Panu Profesorowi Jackowi Świątkowskiemu, mojemu promotorowi, za poświęcony mi czas, oraz liczne rady i wskazówki, które pomogły mi w napisaniu tej pracy.

Wstęp

Skrypt ten powstał na podstawie notatek do wykładu „Algbra liniowa 2” prowadzonego przez prof. dr. hab. Jacka Świątkowskiego. Przeznaczony jest dla studentów pierwszego roku matematyki, w szczególności dla słuchaczy kursu z algebry liniowej 2 w łatwiejszym nurcie A.

Założeniem skryptu była pomoc słuchaczom w lepszym zrozumieniu treści omawianych w trakcie wykładu, dlatego został on napisany prostym językiem i wzbogacony o liczne rysunki i przykłady.

Treści omawiane w trakcie wykładu, a tym samym zawarte w niniejszej pracy, obejmują zakres klasycznej algebry liniowej połączonej z geometrią analityczną, dla przestrzeni R^3, R^n oraz dowolnych przestrzeni wektorowych. Skrypt ten w wielu miejscach będzie zawierał odwołania do analogicznych pojęć omawianych w trakcie wykładu „Algebra liniowa 1” i zawartych w pierwszej części skryptu napisanej przez Patrycję Piechaczek, dlatego wskazane jest, aby czytelnik chcący zapoznać się z niniejszą pracą opanował najpierw materiał omawiany w pierwszej części.

Pierwsze cztery rozdziały poświęcone zostały na omówienie podstawowych pojęć używanych przy badaniu przestrzeni R^3, przy czym szczególny nacisk położono na geometryczną interpretację tych pojęć, a dopiero w oparciu o nie wprowadza się teorię z algebry liniowej. Następnie (rozdziały 5 - 11) omówiono przekształcenia przestrzeni R^3, szczególny nacisk położono na poznanie teorii dotyczących przekształceń liniowych, poczynając od geometrycznych przykładów takich przekształceń, poprzez omówienie najważniejszych własności, a kończąc na wprowadzeniu pewnych klas przekształceń liniowych. W odniesieniu do nich wprowadza się również pojęcie wartości i wektorów własnych oraz macierzy przekształcenia. Kolejne rozdziały (12-13) dotyczą pewnych rodzajów macierzy rozmiaru 3×3, wraz z odniesieniem do przekształceń zadanych tymi macierzami. W dalszej części (rozdział 14) wprowadzono nowe układy współrzędnych, zarówno na płaszczyźnie jak i w przestrzeni, które będą wykorzystywane do badania powierzchni stopnia drugiego, omawianych w rozdziale 15. W skrypcie dość dużo uwagi poświęca się różnym metodom rozwiązywania układów równań liniowych z wieloma niewiadomymi (rozdziały 16,18,21). Dalsza część zawiera teorię dotyczącą przestrzeni wyższych wymiarów, przy czym zrezygnowano tu z odwoływania się do intuicji geometrycznych. Pozostawiono jedynie analogie do przestrzeni R^2 i R^3. Następnie (rozdział 18) wprowadzono teorię dotyczącą macierzy dowolnego rozmiaru. Omówiono tam działania na macierzach, jak również ich zastosowanie. Na koniec (rozdział 24) podano przykłady pojęć występujących w przestrzeniach R^n oraz w dowolnych przestrzeniach wektorowych, które możemy rozpatrywać jako uogólnienia pojęć pojawiających się w R^3 i R^n.

Mam nadzieję, że skrypt ten przyczyni się do lepszego zrozumienia treści omawianych w ramach wykładu. Życzę Czytelnikowi czerpania jak największej przyjemności z odkrywania algebry liniowej.

Barbara Szczepańska

Spis treści

Rozdział 1

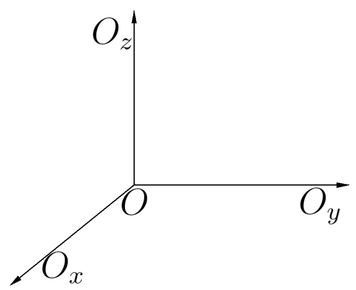
Punkty i wektory w R^3

W rozdziale tym zajmować będziemy się podstawowymi własnościami przestrzeni R^3, w której wprowadzony został układ współrzędnych. Zdefiniujemy pojęcia, które będziemy wykorzystywać w dalszych badaniach tej przestrzeni. Nauczymy się wykonywać podstawowe działania na wektorach z R^3, określimy również własności tych działań.

1.1 Współrzędne punktów w R^3.

➢ Układ współrzędnych O_{xyz} w przestrzeni

Przypomnijmy, że układ współrzędnych w R^3 składa się z trzech wzajemnie prostopadłych osi liczbowych o wspólnym początku w punkcie O, zwanym początkiem układu współrzędnych. Współrzędne na osiach określać będziemy zgodnie ze skalą jednostkową. Osie te oznaczamy odpowiednio O_x, O_y, O_z. Tradycyjnie przyjęło się oznaczać pionową oś układu współrzędnych przez O_z, natomiast poziome przez O_x i O_y, tak jak to pokazuje poniższy rysunek.



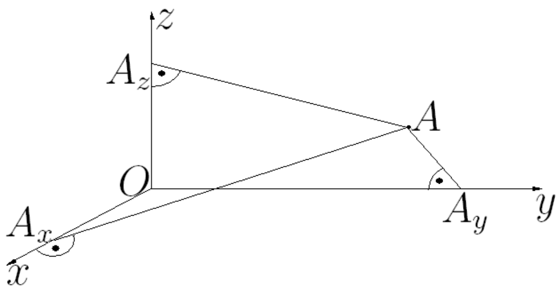
Często, na rysunkach, zamiast pisać O_x, O_y, O_z będziemy opisywać osie przez x, y, z.

Układ utworzony przez te osie nazywamy układem O_{xyz}, natomiast płaszczyzny, oznaczane O_{xy}, O_{xz}, O_{yz}, z których każda zawiera dwie osie układu, nazywamy płaszczyznami współrzędnych układu.

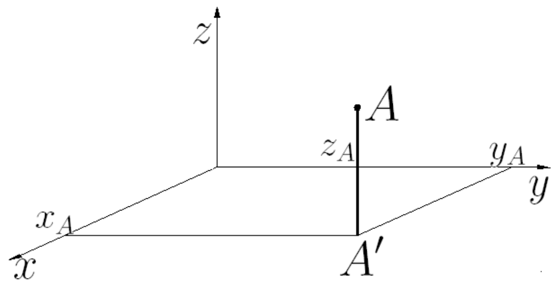
➢Określenie współrzędnych punktów

Niech A_x, A_y, A_z będą rzutami prostokątnymi punktu A na osie O_x, O_y, O_z odpowiednio. **Współrzędnymi** x_A, y_A, z_A punktu A nazywamy współrzędne rzutów

A_x, A_y, A_z na poszczególnych osiach. Współrzędne punktu A będziemy zapisywać jako (x_A, y_A, z_A).



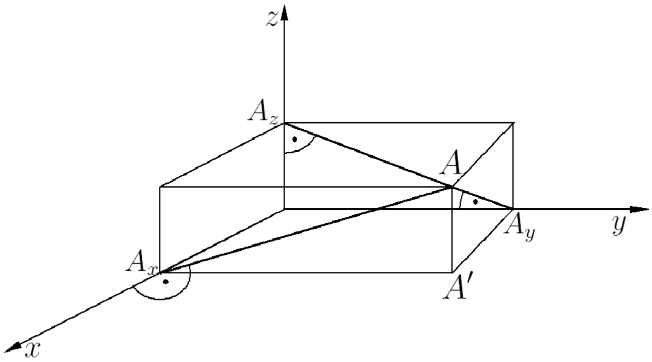
Rozważmy inny, bardziej poglądowy, sposób określenia współrzędnych punktu A. Niech A' będzie rzutem prostokątnym punktu A na płaszczyznę O_{xy}.



Współrzędne punktu A' w układzie O_{xy} to (x_A, y_A). Określmy z_A jako odległość punktu A od swojego rzutu A', gdy A leży powyżej płaszczyzny O_{xy}, lub z_A to ta odległość ze znakiem minus, gdy A leży poniżej płaszczyzny O_{xy}. Współrzędnymi punktu A będziemy wtedy nazywać trójkę liczb (x_A, y_A, z_A).

➢Równoważność określeń współrzędnych punktów

Uzasadnijmy teraz, że te dwa sposoby określenia współrzędnych są rzeczywiście równoważne. Weźmy dowolny punkt A. Poprowadźmy przez ten punkt płaszczyzny prostopadłe do osi układu. Nazwijmy te płaszczyzny A_{xy}, A_{yz}, A_{xz}. Zauważmy, że punkty przecięcia tych płaszczyzn z osiami układu będą rzutami A na poszczególne osie. Płaszczyzny $O_{xy}, O_{yz}, O_{xz}, A_{xy}, A_{yz}, A_{xz}$ ograniczają pewien prostopadłościan, którego wierzchołkami są A_x, A_y, A_z, O, A, A'.



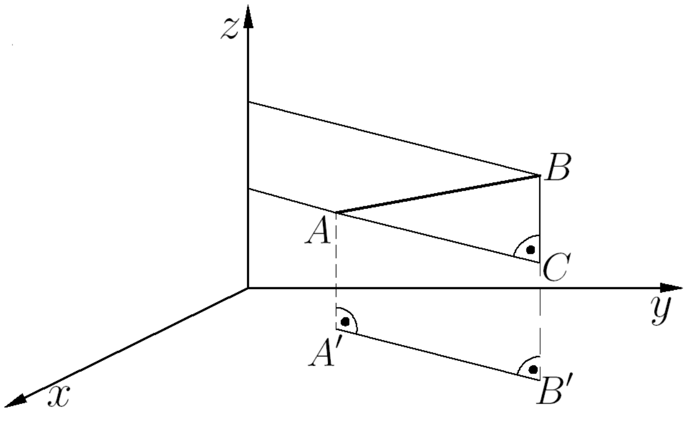
Zauważmy, że punkt A' jest rzutem punktu A na płaszczyznę O_{xy}. Współrzędne tego punktu na płaszczyźnie O_{xy} wyznaczone są przez rzuty punktu A' na osie O_x, O_y. Nietrudno przekonać się, że są to odpowiednio punkty A_x oraz A_y. Wystarczy teraz pokazać, że trzecia współrzędna również jest wyznaczona jednoznacznie. Z równoległości płaszczyzn O_{xy} oraz A_{xy} wynika, że odległość A od O_{xy} jest równa odległości A_z od O_{xy}, czyli współrzędnej A_z. Uzasadniliśmy zatem, że przedstawione powyżej dwa sposoby określania współrzędnych są równoważne.
Przyjmijmy, że zapis $A(x, y, z)$ oznaczać będzie punkt A o współrzędnych (x, y, z).

Uwaga 1.1.1 (jednoznaczność wyznaczenia punktu przez współrzędne). *Dla każdej trójki liczb rzeczywistych (x, y, z) istnieje dokładnie jeden punkt A o takich współrzędnych.*

Jako uzasadnienie powyższej uwagi spróbujmy, dla współrzędnych (x, y, z), znaleźć odpowiadający im punkt. Dwie pierwsze współrzędne określają położenie punktu na płaszczyźnie O_{xy}. Jak wiemy współrzędne punktu na płaszczyźnie są wyznaczone jednoznacznie, istnieje więc dokładnie jeden punkt A', który na płaszczyźnie O_{xy} ma współrzędne (x, y). Trzecia współrzędna określa odległość szukanego punktu od swojego rzutu na płaszczyznę O_{xy}. Jeżeli z jest liczbą dodatnią to szukany punkt A będzie znajdował się nad płaszczyzną O_{xy}, jeżeli natomiast jest liczbą ujemną, to pod. Zauważmy, że po ustalonej stronie płaszczyzny, istnieje dokładnie jeden punkt znajdujący się w odległości $|z|$ od A' taki, że A' jest jego rzutem na płaszczyznę O_{xy}. Punkt A jest zatem jednoznacznie wyznaczony przez swoje współrzędne.

1.2 Odległość punktów w R^3.

Aby obliczyć odległość punktu A o współrzędnych (x, y, z) od punktu B o współrzędnych (x', y', z') rozważmy rzuty tych punktów na płaszczyznę O_{xy}.



Na tej płaszczyźnie punkty A' i B' mają współrzędne $A' = (x, y), B' = (x', y')$. Zatem ich odległość wynosi

$$|A'B'| = \sqrt{(x - x')^2 + (y - y')^2}.$$

Figura $A'B'CA$ jest prostokątem, więc $|AC| = |A'B'|$. Ponadto $|BC| = |z - z'|$. Stosując twierdzenie Pitagorasa dla trójkąta ABC, otrzymujemy

$$|AB|^2 = |AC|^2 + |CB|^2 = \left(\sqrt{(x-x')^2 + (y-y')^2}\right)^2 + (z-z')^2.$$

Stąd

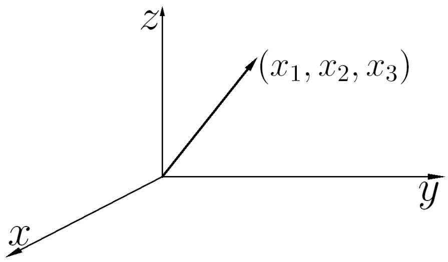
$$|AB|^2 = (x-x')^2 + (y-y')^2 + (z-z')^2.$$

Odległość tych punktów wyraża się zatem wzorem

$$\boxed{|AB| = \sqrt{(x-x')^2 + (y-y')^2 + (z-z')^2}.} \tag{1.1}$$

1.3 Wektory w R^3.

Podobnie jak w przypadku płaszczyzny punkt $X(x_1, x_2, x_3)$ będziemy utożsamiać z jego **wektorem wodzącym**, czyli wektorem o początku w początku układu współrzędnych i końcu w punkcie X. Wektory takie oznaczać będziemy $\overrightarrow{X}$ lub X.



➢Współrzędne wektora

Rozważać będziemy również wektory zaczepione w punkcie nie będącym początkiem układu współrzędnych. Niech

$$A = (a_1, a_2, a_3), \quad B = (b_1, b_2, b_3),$$

gdzie A jest początkiem wektora, a B jego końcem. Za współrzędne wektora $\overrightarrow{AB}$ przyjmuje się

$$\boxed{\overrightarrow{AB} = [b_1 - a_1, b_2 - a_2, b_3 - a_3].} \tag{1.2}$$

✔ **Przykład.** Wektor o początku w punkcie $A = (2, 1, -3)$ i końcu w punkcie $B = (-2, 6, 8)$ ma współrzędne

$$\overrightarrow{AB} = [-2-2, 6-1, 8-(-3)] = [-4, 5, 11].$$

Obserwacja 1.3.1. *Zauważmy, że jeśli początek wektora, punkt A, pokrywa się z początkiem układu, czyli $A = (0, 0, 0)$, a koniec wektora ma współrzędne $B = (b_1, b_2, b_3)$, to współrzędne wektora $\overrightarrow{AB}$ pokrywają się ze współrzędnymi punktu B. Tak więc wektory wodzące punktów mają takie same współrzędne jak te punkty.*

Uwaga 1.3.2 (kolumnowy zapis współrzędnych). *Współrzędne wektora będziemy często oznaczać*

$$X = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}.$$

Oznaczenie takie nazywane jest kolumnowym zapisem współrzędnych.

Definicja 1.3.3 (równość wektorów). Wektory U i V nazywamy równymi, jeśli mają jednakowe długości, kierunki i zwroty.

Zauważmy, że wektory równe wektorowi V to wszystkie wektory uzyskane przez przesunięcie wektora V do innych punktów zaczepienia. Równość wektorów możemy również rozpoznawać przy użyciu ich współrzędnych, o czym mowa jest w następującym twierdzeniu:

Twierdzenie 1.3.4 (współrzędne a równość wektorów). *Wektory w przestrzeni są równe wtedy i tylko wtedy gdy mają równe współrzędne.*

Dowód twierdzenia pomijamy.

✔ **Przykład.** Rozważmy dwa wektory $\overrightarrow{AB}$ oraz $\overrightarrow{CD}$, gdzie:

$$A = (1, 1, 1),\ B = (3, -2, 5),\ C = (5, -1, 7),\ D = (7, -4, 11).$$

Współrzędne tych wektorów to

$$\overrightarrow{AB} = \begin{pmatrix} 2 \\ -3 \\ 4 \end{pmatrix},\quad \overrightarrow{CD} = \begin{pmatrix} 2 \\ -3 \\ 4 \end{pmatrix}.$$

Zatem, jak wynika z powyższgo twierdzenia wektory $\overrightarrow{AB}$ i $\overrightarrow{CD}$ są równe.

Zauważmy, że każdy wektor zaczepiony jest równy wektorowi wodzącemu, mającemu takie same współrzędne.

1.4 Działania na wektorach.

Poniżej w czysto algebraiczny sposób zdefiniujemy podstawowe działania na wektorach.

➢Dodawanie wektorów

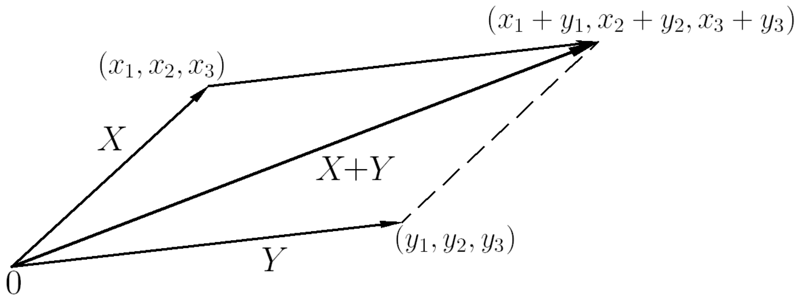
Niech dane będą wektory

$$X = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix},\quad Y = \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix}.$$

Ich sumą będziemy nazywać wektor

$$\boxed{X + Y = \begin{pmatrix} x_1 + y_1 \\ x_2 + y_2 \\ x_3 + y_3 \end{pmatrix}.} \tag{1.3}$$

Aby przekonać się o zgodności powyższego określenia z tradycyjną geometryczną definicją dodawania wektorów, rozważmy w przestrzeni dwa wektory X i Y o wspólnym początku w punkcie $(0, 0, 0)$.



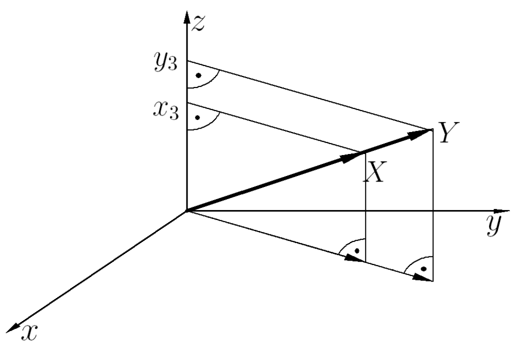
Geometrycznie sumę dwóch wektorów określa się jako przekątną równoległoboku zbudowanego na tych wektorach. Przenieśmy zatem równolegle wektor Y tak, aby jego początek znalazł się w punkcie (x_1, x_2, x_3). Nietrudno przekonać się, że koniec tak przesuniętego wektora wypada w punkcie $(x_1 + y_1, x_2 + y_2, x_3 + y_3)$. Wektor wodzący o końcu w tym punkcie, czyli o współrzędnych $\begin{pmatrix} x_1 + y_1 \\ x_2 + y_2 \\ x_3 + y_3 \end{pmatrix}$, jest sumą wektorów X i Y.

➢Mnożenie wektora przez skalar

Niech dany będzie wektor $X = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$ oraz liczba rzeczywista k. Iloczynem liczby k i wektora X nazywamy wektor

$$\boxed{k \cdot X = \begin{pmatrix} k \cdot x_1 \\ k \cdot x_2 \\ k \cdot x_3 \end{pmatrix}.} \tag{1.4}$$

Pokażemy teraz zgodność powyższego określenia z geometryczną definicją mnożenia wektora przez liczbę. Ograniczymy się do przypadku gdy $k > 0$. Iloczyn danego wektora przez k to wektor o tym samym kierunku i zwrocie, ale k razy dłuższy. Niech dany będzie wektor Y, będący iloczynem wektora X oraz liczby k.



Nietrudno uzasadnić, np. stosując tw. Talesa, że jeżeli rzutem wektora X na płaszczyznę O_{xy} jest wektor X', to rzutem wektora Y jest wektor Y', który ma ten sam kierunek i zwrot co wektor X', ale jest od niego k razy dłuższy, tak jak to zostało przedstawione na rysunku powyżej. Z określenia mnożenia wektora przez liczbę na płaszczyźnie otrzymujemy więc, że $y_1 = kx_1$ oraz $y_2 = kx_2$. Podobnie z własności rzutu na oś O_z otrzymujemy $y_3 = kx_3$. Tak więc mamy

$$Y = k \cdot X = \begin{pmatrix} k \cdot x_1 \\ k \cdot x_2 \\ k \cdot x_3 \end{pmatrix}.$$

W przypadku gdy $k < 0$ rozumowanie bedzie wyglądać analogicznie.

Dla $k = -1$ otrzymujemy $k \cdot X = (-1) \cdot \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} -x_1 \\ -x_2 \\ -x_3 \end{pmatrix}$. Wektor ten jest nazywany **wektorem przeciwnym** do X, tzn. wektorem o tym samym kierunku i długości, ale o przeciwnym zwrocie. Oznaczać go będziemy $-X$.

➢Własności działań na wektorach

Określimy teraz, podobnie jak na płaszczyźnie, podstawowe własności powyżej określonych działań.

(1) $X + U = U + X$

(2) $(X + U) + V = X + (U + V)$

(3) Istnieje wektor $O = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$ taki, że dla każdego X mamy

$$X + O = O + X = X.$$

(4) Dla każdego wektora X istnieje wektor $-X$, taki że

$$X + (-X) = O.$$

(5) $t \cdot (X + U) = t \cdot X + t \cdot U$

(6) $(t + s) \cdot X = t \cdot X + s \cdot X$

(7) $t \cdot (s \cdot X) = (t \cdot s) \cdot X$

(8) $1 \cdot X = X, \quad (-1) \cdot X = -X$

✍ **Ćwiczenie.** Uzasadnienie powyższych własności, na podstawie algebraicznych definicji działań na wektorach, pozostawimy do wykonania czytelnikowi.

➢Różnica wektorów

Niech dane będą wektory

$$X = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}, \quad Y = \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix}.$$

Wtedy ich różnicą będziemy nazywać wektor

$$X - Y = X + (-Y) = \begin{pmatrix} x_1 - y_1 \\ x_2 - y_2 \\ x_3 - y_3 \end{pmatrix}. \tag{1.5}$$

➢Długość wektora

Przypomnijmy, że długość wektora $\overrightarrow{AB}$ to z definicji długość odcinka AB. Z równania (1.1) dla punktów

$$A = \begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix}, \quad B = \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix}$$

mamy więc

$$|\overrightarrow{AB}| = \sqrt{(b_1 - a_1)^2 + (b_2 - a_2)^2 + (b_3 - a_3)^2}. \tag{1.6}$$

1.5 Rozkład wektora względem ustalonych trzech wektorów.

Definicja 1.5.1 (niewspółpłaszczyznowość wektorów). Wektory X, Y, Z w R^3 nazywamy niewspółpłaszczyznowymi, jeśli nie istnieje płaszczyzna, w której wszystkie są zawarte.

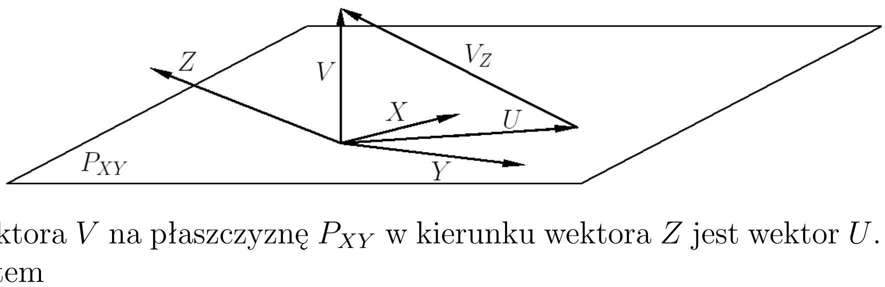
Definicja 1.5.2 (rozkład wektora). Rozkładem dowolnego wektora V względem ustalonych trzech wektorów niewspółpłaszczyznowych X, Y, Z, nazywamy przedstawienie wektora V w postaci

$$V = V_X + V_Y + V_Z,$$

gdzie $V_X || X, \quad V_Y || Y, \quad V_Z || Z$.

Fakt 1.5.3 (istnienie rozkładu). *Niech X, Y, Z będą niewspółpłaszczyznowymi wektorami w R^3. Wówczas każdy wektor V posiada rozkład względem X, Y, Z.*

Dowód. Niech P_{XY} będzie płaszczyzną zawierającą wektory X i Y.



Rzutem wektora V na płaszczyznę P_{XY} w kierunku wektora Z jest wektor U. Otrzymujemy zatem

$$V = U + V_Z,$$

gdzie $V_Z||Z$. Wektor U zawiera się w płaszczyźnie P_{XY}, więc możemy go na tej płaszczyźnie rozłożyć względem X i Y. Otrzymujemy

$$U = V_X + V_Y,$$

gdzie $V_X||X,\ V_Y||Y$.
Ostatecznie

$$V = U + V_Z = V_X + V_Y + V_Z.$$

□

1.6 Przedstawienie wektora w postaci kombinacji liniowej trzech wektorów.

Definicja 1.6.1 (kombinacja liniowa wektorów). Kombinacją liniową wektorów X, Y, Z nazywamy wektor postaci

$$s \cdot X + t \cdot Y + r \cdot Z,$$

gdzie s, t, r są liczbami rzeczywistymi, nazywanymi współczynnikami kombinacji.

Jeżeli wektor V daje się przedstawić w postaci kombinacji liniowej wektorów X, Y, Z, czyli $V = s \cdot X + t \cdot Y + r \cdot Z$, dla pewnych s, t, r, to mówimy wtedy, że wektor V rozkłada się w kombinację liniową tych wektorów.

✔ **Przykład.** Wektor $V = \begin{pmatrix} 7 \\ -12 \\ 6 \end{pmatrix}$ jest kombinacją liniową wektorów $X = \begin{pmatrix} 2 \\ 0 \\ 3 \end{pmatrix}$, $Y = \begin{pmatrix} 1 \\ -2 \\ -5 \end{pmatrix}$, $Z = \begin{pmatrix} -1 \\ 4 \\ 0 \end{pmatrix}$, ponieważ

$$V = 2 \cdot X + 0 \cdot Y + (-3) \cdot Z.$$

➢Wyznaczanie rozkładu wektorów w kombinację liniową

Dla $V = \begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix}$, $X = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$, $Y = \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix}$, $Z = \begin{pmatrix} z_1 \\ z_2 \\ z_3 \end{pmatrix}$ podaną w definicji równość $V = s \cdot X + t \cdot Y + r \cdot Z$ możemy zapisać jako

$$\begin{cases} v_1 = s \cdot x_1 + t \cdot y_1 + r \cdot z_1 \\ v_2 = s \cdot x_2 + t \cdot y_2 + r \cdot z_2 \\ v_2 = s \cdot x_3 + t \cdot y_3 + r \cdot z_3 \end{cases}$$

Szukanie rozkładu danego wektora w kombinację liniową zadanych wektorów sprowadza się zatem do rozwiązania pewnego układu równań liniowych.

✔ **Przykład.** Rozważmy wektory

$$V = \begin{pmatrix} 6 \\ 3 \\ -5 \end{pmatrix}, \quad X = \begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix}, \quad Y = \begin{pmatrix} 2 \\ 1 \\ -2 \end{pmatrix}, \quad Z = \begin{pmatrix} 2 \\ -2 \\ 4 \end{pmatrix}.$$

Szukanie rozkładu wektora V w kombinację liniową wektorów X, Y, Z, czyli znalezienie takich x, y, z, aby

$$x \cdot \begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix} + y \cdot \begin{pmatrix} 2 \\ 1 \\ -2 \end{pmatrix} + z \cdot \begin{pmatrix} 2 \\ -2 \\ 4 \end{pmatrix} = \begin{pmatrix} 6 \\ 3 \\ -5 \end{pmatrix},$$

sprowadza się do rozwiązania układu równań

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

Rozwiązaniem tego układu są

$$x = -1, \quad y = 3, \quad z = \frac{1}{2}.$$

Znaleźliśmy zatem rozkład wektora V w kombinację liniową wektorów X, Y, Z. Wygląda on następująco

$$V = (-1) \cdot X + 3 \cdot Y + \frac{1}{2} \cdot Z.$$

1.7 Równanie parametryczne prostej w przestrzeni.

Na początek zajmijmy się przypadkiem, gdy dana prosta L przechodzi przez początek układu współrzędnych. Weźmy dowolny wektor U zawarty w prostej L. Wtedy, podobnie jak na płaszczyźnie, prostą L możemy opisać równaniem

$$X = t \cdot U,$$

gdzie $t \in R$.

Zauważmy, że dla różnych wartości parametru t dostajemy kolejno współrzędne różnych punktów tej prostej, przy czym każdy punkt, przy odpowiedniej wartości t, da się w ten sposób przedstawić. Podstawiając za t kolejno wszystkie liczby rzeczywiste, dostaniemy współrzędne wszystkich punktów tej prostej.

W ogólnym przypadku, tzn. gdy prosta nie zawiera początku układu współrzędnych, wybierzmy punkt początkowy $V \in L$. Wtedy prostą można zapisać równaniem

$$X = V + t \cdot U,$$

gdzie $t \in R$, a U jest wektorem zawartym w L.
Przy użyciu współrzędnych i przy oznaczeniach

$$V = \begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix}, \quad U = \begin{pmatrix} u_1 \\ u_2 \\ u_3 \end{pmatrix},$$

równość powyższą możemy zapisać jako

$$\boxed{\begin{cases} x_1 = v_1 + t \cdot u_1 \\ x_2 = v_2 + t \cdot u_2 \\ x_2 = v_3 + t \cdot u_3 \end{cases}}$$

Takie przedstawienie nazywamy **równaniem parametrycznym prostej** we współrzędnych.

1.8 Równanie parametryczne płaszczyzny w przestrzeni.

Dla płaszczyzny Π zawierającej początek układu współrzędnych rozważmy dwa niewspółliniowe wektory U, W zawarte w tej płaszczyźnie. Zauważmy, że dowolny wektor z płaszczyzny Π można zapisać jako kombinację liniową U, W i że każda taka kombinacja wyznacza wektor z płaszczyzny Π. Daną płaszczyznę możemy zatem opisać rówaniem

$$X = t \cdot U + s \cdot W,$$

gdzie $t \in R, s \in R$ (parametry).

Dla dowolnej płaszczyzny niech V będzie punktem początkowym. Wtedy

$$X = V + t \cdot U + s \cdot W,$$

gdzie t i s są parametrami rzeczywistymi, a U, W są wektorami z tej płaszczyzny. Używając współrzędnych

$$V = \begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix}, \quad U = \begin{pmatrix} u_1 \\ u_2 \\ u_3 \end{pmatrix}, \quad W = \begin{pmatrix} w_1 \\ w_2 \\ w_3 \end{pmatrix},$$

równość powyższa przedstawia się jako

$$\boxed{\begin{cases} x_1 = v_1 + t \cdot u_1 + s \cdot w_1 \\ x_2 = v_2 + t \cdot u_2 + s \cdot w_2 \\ x_2 = v_3 + t \cdot u_3 + s \cdot w_3 \end{cases}}$$

Taki sposób zapisu nazywamy **równaniem parametrycznym płaszczyzny** we współrzędnych.

1.9 Liniowa zależność i niezależność wektorów w przestrzeni.

Dwa wektory X, Y nazywamy liniowo zależnymi, jeżeli

$$X = t \cdot Y \ \text{ lub } Y = t \cdot X.$$

Geometrycznie liniowa zależność oznacza współliniowość lub równoległość wektorów. Trzy wektory X, Y, Z nazywamy liniowo zależnymi, jeżeli

$$X = t \cdot Y + s \cdot Z \quad \text{lub} \quad Y = t \cdot X + s \cdot Z \quad \text{lub} \quad Z = t \cdot X + s \cdot Y,$$

tzn. jeden z nich daje się przedstawić jako kombinacja liniowa dwóch pozostałych. Geometrycznie równość $X = t \cdot Y + s \cdot Z$ oznacza, że wektory X, Y, Z leżą na jednej płaszczyźnie. Zauważmy, że jeżeli Y i Z są niewspółliniowe, to wyznaczają płaszczyznę $t \cdot Y + s \cdot Z$ i X należy do tej płaszczyzny. Jeżeli natomiast Y i Z są współliniowe, to kombinacje $t \cdot Y + s \cdot Z$ wyznaczają prostą lub w szczególnym przypadku, gdy $Y = 0$ i $Z = 0$, punkt.

✔ **Przykład.**
(1) Rozważmy wektory

$$A = \begin{pmatrix} 4 \\ 7 \\ 6 \end{pmatrix}, \quad B = \begin{pmatrix} 2 \\ 3 \\ -2 \end{pmatrix}, \quad C = \begin{pmatrix} 1 \\ 2 \\ 4 \end{pmatrix}.$$

Wektory A, B, C są liniowo zależne, bo

$$A = 1 \cdot B + 2 \cdot C.$$

(2) Wektory $\vec{O}$, X są liniowo zależne, bo

$$\vec{O} = 0 \cdot X.$$

Definicja 1.9.1 (liniowa niezależność trójki wektorów). Wektory X, Y, Z są liniowo niezależne, jeśli żaden z nich nie wyraża się jako kombinacja liniowa dwóch pozostałych.

Podamy teraz inną, bardziej użyteczną, charakteryzację liniowej niezależności dla trzech wektorów.

Lemat 1.9.2 (liniowa niezależność a kombinacja wektorów). *Wektory X, Y, Z są liniowo niezależne wtedy i tylko wtedy, gdy jedyna ich kombinacja liniowa $t \cdot X + s \cdot Y + r \cdot Z$ dająca wektor zerowy to kombinacja ze współczynnikami*

$$t = s = r = 0$$

(tzw. kombinacja trywialna).

Dowód. Zakładamy, że wektory X, Y, Z są liniowo niezależne. Gdyby $t \cdot X + s \cdot Y + r \cdot Z = O$ i jeden ze współczynników np. $t \neq 0$, wówczas mielibyśmy

$$t \cdot X = -s \cdot Y - r \cdot Z \qquad / : t$$

$$X = -\frac{s}{t} \cdot Y - \frac{r}{t} \cdot Z.$$

Zatem X jest kombinacją liniową wektorów Y i Z, co przeczy założeniu o liniowej niezależności. Dowiedliśmy zatem, że z faktu liniowej niezależności wektorów

X, Y, Z wynika trywialność kombinacji liniowej dającej wektor zerowy. Spróbujmy teraz dowieść implikacji w przeciwną stronę. Zakładamy, że jedyną kombinacją liniową wektorów X, Y, Z dającą $\vec{O}$ jest kombinacja trywialna. Gdyby wektory X, Y, Z były liniowo zależne, czyli np.

$$X = t \cdot Y + s \cdot Z,$$

wtedy mielibyśmy

$$1 \cdot X - t \cdot Y - s \cdot Z = \vec{O}$$

co przeczy założeniu o trywialności kombinacji dającej $\vec{O}$. □

✔ **Przykład.** Rozważmy wektory

$$A = \begin{pmatrix} 3 \\ 0 \\ 0 \end{pmatrix}, \quad B = \begin{pmatrix} 0 \\ 2 \\ 0 \end{pmatrix}, \quad C = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}.$$

Szukamy takich $t,\ s,\ r$ aby kombinacja liniowa $t \cdot A + s \cdot B + r \cdot C$ dawała wektor zerowy.

$$t \cdot A + s \cdot B + r \cdot C = t \begin{pmatrix} 3 \\ 0 \\ 0 \end{pmatrix} + s \begin{pmatrix} 0 \\ 2 \\ 0 \end{pmatrix} + r \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 3t \\ 2s \\ r \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}.$$

Otrzymujemy zatem, że $t = s = r = 0$. Czyli wektory A, B, C są liniowo niezależne.

Twierdzenie 1.9.3 (jednoznaczność rozkładu w kombinację liniową). *Niech U, V, W będą liniowo niezależnymi wektorami, zaś X niech będzie dowolnym innym wektorem w przestrzeni. Wówczas X rozkłada się w kombinację liniową*

$$X = t \cdot U + s \cdot V + r \cdot W,$$

i to dokładnie na jeden sposób.

Dowód.
Istnienie rozkładu:
Jeśli U, V, W są liniowo niezależne, to w szczególności są niezerowe, żadne dwa z nich nie są współliniowe, a wszystkie trzy nie leżą na jednej płaszczyźnie. Wektor X, jak wynika z Faktu 1.5.3, posiada rozkład względem U, V, W, możemy go zatem zapisać w postaci $X = X_U + X_V + X_W$, gdzie $X_U || U,\ X_V || V,\ X_W || W$. Z równoległości wektorów X_U oraz U otrzymujemy, że istnieje takie t, że $X_U = t \cdot U$. Podobnie znajdziemy takie s, że $X_V = s \cdot V$, oraz takie r, że $X_W = r \cdot W$. Podstawiając te iloczyny za X_U, X_V, X_W otrzymujemy szukany rozkład

$$X = t \cdot U + s \cdot V + r \cdot W.$$

Jednoznaczność rozkładu:
Załóżmy, że rozkład ten nie jest jednoznaczny.
Rozważmy dwa rozkłady

$$X = t_1 \cdot U + s_1 \cdot V + r_1 \cdot W = t_2 \cdot U + s_2 \cdot V + r_2 \cdot W.$$

Pokażemy, że rozkłady te są identyczne, tzn. mają jednakowe współczynniki przy poszczególnych wektorach. Zauważmy, że możemy napisać następujące równości

$$\begin{aligned}\vec{O} = X - X = t_1 \cdot U + s_1 \cdot V + r_1 \cdot W - t_2 \cdot U - s_2 \cdot V - r_2 \cdot W = \\ = (t_1 - t_2) \cdot U + (s_1 - s_2) \cdot V + (r_1 - r_2) \cdot W.\end{aligned}$$

Z liniowej niezależności wektorów U, V, W wynika, że

$$t_1 - t_2 = 0, \quad s_1 - s_2 = 0, \quad r_1 - r_2 = 0,$$

czyli

$$t_1 = t_2, \quad s_1 = s_2, \quad r_1 = r_2.$$

Zatem te dwa rozkłady są identyczne. □

Rozdział 2

Iloczyn skalarny w R^3

W tym rozdziale zajmiemy się pojęciem iloczynu skalarnego. Ma ono bardzo wiele zastosowań, z których część tutaj podamy. Na początek w czysto algebraiczny sposób zdefiniujemy to pojęcie i wyprowadzimy jego najważniejsze własności. W kolejnych podrozdziałach zapoznamy się z geometrycznymi konsekwencjami iloczynu skalarnego. W dalszych rozdziałach tego skryptu będziemy bardzo często wracać do tego pojęcia. Okaże się, że jest ono jednym z ważniejszych narzędzi, wykorzystywanych do badania własności przestrzeni.

2.1 Definicja i podstawowe własności iloczynu skalarnego.

Definicja, którą teraz podamy ma charakter czysto algebraiczny, geometrycznym określeniem iloczynu skalarnego zajmiemy się w dalszej części.

Definicja 2.1.1 (iloczyn skalarny). Iloczynem skalarnym wektorów

$$X = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}, \quad Y = \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix},$$

oznaczanym przez $X \circ Y$, nazywamy liczbę

$$\boxed{X \circ Y = x_1 \cdot y_1 + x_2 \cdot y_2 + x_3 \cdot y_3.} \tag{2.1}$$

Zauważmy, że iloczyn skalarny to funkcja, która dowolnej parze wektorów przyporządkowuje liczbę rzeczywistą.

✔ **Przykład.** Iloczyn skalarny wektorów

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

to

$$A \circ B = (-3) \cdot (-2) + 2 \cdot 4 + 4 \cdot (-1) = 10.$$

➢ Własności iloczynu skalarnego

Dla dowolnych wektorów $X, Y, Z \in R^3$ oraz dowolnej liczby rzeczywistej t zachodzą równości:

(1) $X \circ Y = Y \circ X$

(2) $X \circ (Y + Z) = X \circ Y + X \circ Z$

(3) $(t \cdot X) \circ Y = t(X \circ Y)$

(4) $X \circ X = |X|^2$

W dowodach tych własności korzystać będziemy z definicji iloczynu skalarnego, oraz z podstawowych własności operacji na zbiorze liczb rzeczywistych, takich jak przemienność i łączność mnożenia, oraz rozdzielność mnożenia względem dodawania.

Dowód (1).

$$X \circ Y = x_1 \cdot y_1 + x_2 \cdot y_2 + x_3 \cdot y_3 = y_1 \cdot x_1 + y_2 \cdot x_2 + y_3 \cdot x_3 = Y \circ X$$

□

Dowód (2).

$$\begin{aligned} X \circ (Y + Z) &= \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \circ \left(\begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} + \begin{pmatrix} z_1 \\ z_2 \\ z_3 \end{pmatrix} \right) = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \circ \begin{pmatrix} y_1 + z_1 \\ y_2 + z_2 \\ y_3 + z_3 \end{pmatrix} = \\ &= x_1(y_1 + z_1) + x_2(y_2 + z_2) + x_3(y_3 + z_3) = \\ &= x_1y_1 + x_2y_2 + x_3y_3 + x_1z_1 + x_2z_2 + x_3z_3 = \\ &= \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \circ \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} + \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \circ \begin{pmatrix} z_1 \\ z_2 \\ z_3 \end{pmatrix} = X \circ Y + X \circ Z \end{aligned}$$

□

Dowód (3).

$$\begin{aligned} (t \cdot X) \circ Y &= \left(t \cdot \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \right) \circ \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} = \begin{pmatrix} tx_1 \\ tx_2 \\ tx_3 \end{pmatrix} \circ \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} = \\ &= (tx_1) \cdot y_1 + (tx_2) \cdot y_2 + (tx_3) \cdot y_3 = x_1 \cdot (ty_1) + x_2 \cdot (ty_2) + x_3 \cdot (ty_3) = \\ &= t(x_1 \cdot y_1) + t(x_2 \cdot y_2) + t(x_3 \cdot y_3) = t(X \circ Y) \end{aligned}$$

□

Dowód (4).

$$\begin{aligned} X \circ X &= \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \circ \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = x_1x_1 + x_2x_2 + x_3x_3 = \\ &= x_1^2 + x_2^2 + x_3^2 = (\sqrt{x_1^2 + x_2^2 + x_3^2})^2 = |X|^2 \end{aligned}$$

□

2.2 Kąt między wektorami.

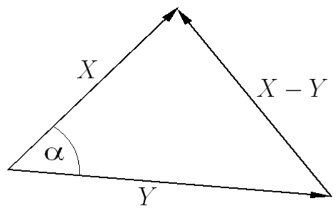
Równość, którą udowodnimy poniżej, bywa często przyjmowana za geometryczną definicję iloczynu skalarnego.

Fakt 2.2.1 (iloczyn skalarny a kąt między wektorami). *Niech X, Y będą wektorami z przestrzeni R^3, α niech będzie kątem pomiędzy X a Y. Wtedy*

$$\boxed{X \circ Y = |X| \cdot |Y| \cdot \cos\alpha.} \tag{2.2}$$

Dowód. Korzystając z twierdzenia cosinusów, dla wektorów $X, Y, X - Y$, mamy

$$|X - Y|^2 = |X|^2 + |Y|^2 - 2 \cdot |X| \cdot |Y| \cdot \cos\alpha.$$



Ale z własności iloczynu skalarnego wiemy również, że

$$|X - Y|^2 = (X - Y) \circ (X - Y) = X \circ X - X \circ Y - Y \circ X + Y \circ Y =$$
$$= |X|^2 + |Y|^2 - 2 \cdot |X| \circ |Y|.$$

Po podstawieniu do twierdzenia cosinusów otrzymujemy

$$|X|^2 + |Y|^2 - 2 \cdot |X| \circ |Y| = |X|^2 + |Y|^2 - 2 \cdot |X| \cdot |Y| \cdot \cos\alpha.$$

Zatem, redukując jednakowe składniki po obu stronach, dostajemy

$$X \circ Y = |X| \cdot |Y| \cdot \cos\alpha.$$

□

Prostą konsekwencją powyższego faktu jest zależność, która pozwoli nam obliczać kąt między wektorami bezpośrednio z ich współrzędnych.

Wniosek 2.2.2. *Jeśli α jest kątem pomiędzy wektorami X, Y, to*

$$\boxed{\cos\alpha = \frac{X \circ Y}{|X| \cdot |Y|} = \frac{x_1y_1 + x_2y_2 + x_3y_3}{\sqrt{x_1^2 + x_2^2 + x_3^2} \cdot \sqrt{y_1^2 + y_2^2 + y_3^2}}.} \tag{2.3}$$

✔ **Przykład.** Obliczmy kąt α pomiędzy wektorem $A = \begin{pmatrix} 3 \\ -4 \\ -1 \end{pmatrix}$ a wektorem $B = \begin{pmatrix} 2 \\ 1 \\ -3 \end{pmatrix}$. Korzystając ze wzoru (2.3) otrzymujemy

$$\cos\alpha = \frac{5}{\sqrt{26} \cdot \sqrt{14}} \approx 0,2621.$$

Z tablic możemy teraz odczytać przybliżoną wartość kąta α

$$\alpha \approx 75^0.$$

Ze wzoru (2.3) możemy wyprowadzić bardzo użyteczną zależność, która pozwoli nam bardzo szybko sprawdzić prostopadłość wektorów.

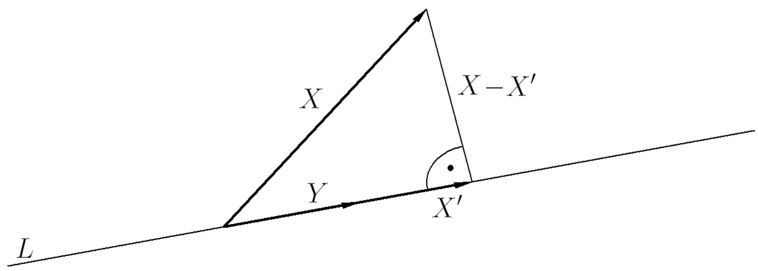
Wniosek 2.2.3 (**iloczyn skalarny a prostopadłość wektorów**). *Wektory* $X, Y \in R^3$ *w przestrzeni są prostopadłe wtedy i tylko wtedy, gdy* $\cos\alpha = 0$, *czyli* $X \circ Y = 0$.

✍ **Ćwiczenie.** Jako ćwiczenie czytelnik może przy użyciu powyższego wniosku sprawdzić prostopadłość wersorów układu współrzędnych.

2.3 Rzut prostokątny wektora X na prostą wzdłuż wektora Y.

W tym podrozdziale zastosujemy iloczyn skalarny do wyprowadzenia wzoru na rzut prostopadły wektora na prostą.

Dla pary wektorów X, Y możemy określić przekształcenie, które przyporządkowuje wektorowi X jego rzut na prostą wzdłuż wektora Y. Rzut ten również będzie wektorem. Spróbujmy znaleźć zależność jaką musi spełniać. Oznaczmy ten rzut przez X'.



Zauważmy, że wektor X' jest współliniowy z wektorem Y, więc $X' = t{\cdot}Y$ dla pewnego $t \in R$. Skoro X' jest rzutem prostokątnym X, to $(X - X')$ musi być prostopadły do wektora Y. Zatem, jak wynika z Wniosku 2.2.3

$$(X - X') \circ Y = 0.$$

Stąd, po podstawieniu i uproszczeniu otrzymujemy

$$0 = (X - t \cdot Y) \circ Y = X \circ Y - t \cdot Y \circ Y = X \circ Y - t \cdot |Y|^2.$$

Korzystając z tej zależności możemy wyznaczyć t

$$t \cdot |Y|^2 = X \circ Y$$

$$t = \frac{X \circ Y}{|Y|^2}.$$

Podstawiając do równości określającej X' mamy

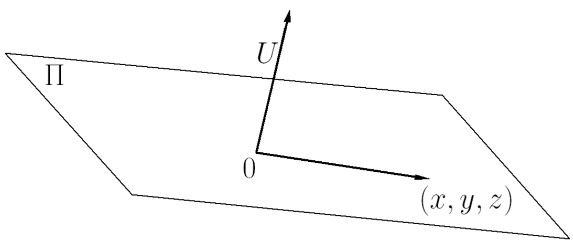
$$\boxed{X' = t \cdot Y = \tfrac{X \circ Y}{|Y|^2} \cdot Y.} \tag{2.4}$$

Zależność tę, we współrzędnych, dla $X = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$, $Y = \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix}$, możemy zapisać jako

$$X' = \frac{x_1y_1 + x_2y_2 + x_3y_3}{y_1^2 + y_2^2 + y_3^2} \cdot \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix}.$$

2.4 Równanie ogólne płaszczyzny przechodzącej przez początek układu współrzędnych.

Spróbujmy znaleźć równanie ogólne płaszczyzny przechodzącej przez $(0,0,0)$. Oznaczmy ją przez Π. Płaszczyzna ta jest wyznaczona przez wektor do niej prostopadły $U = \begin{pmatrix} a \\ b \\ c \end{pmatrix}$.



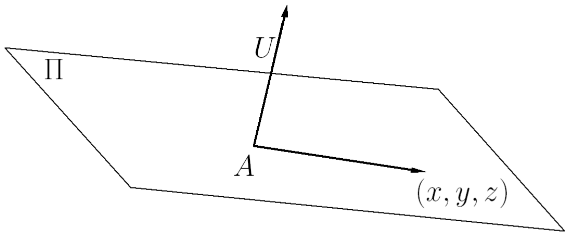
Weźmy dowolny punkt $X = \begin{pmatrix} x \\ y \\ z \end{pmatrix}$ z tej płaszczyzny, wtedy również wektor $\overrightarrow{OX}$ należy do tej płaszczyzny, tak więc jest prostopadły do wektora U, zatem jak wiemy z Wniosku 2.2.3, iloczyn skalarny $\overrightarrow{OX}$ i U jest równy 0, stąd

$$\boxed{ax + by + cz = 0.}$$

Zauważmy, że każdy wektor z płaszczyzny Π musi spełniać tę równość, może ona być przyjęta za ogólne równanie tej płaszczyzny.

2.5 Równanie ogólne płaszczyzny.

Zajmijmy się teraz płaszczyzną, która niekoniecznie przechodzi przez początek układu współrzędnych. Weźmy dowolny punkt $A(x_0, y_0, z_0)$ należący do Π. Płaszczyzna ta jest prostopadła do wektora $U = \begin{pmatrix} a \\ b \\ c \end{pmatrix}$.



Weźmy dowolny punkt $X(x, y, z)$ należący do tej płaszczyzny. Wektor $\overrightarrow{AX}$ również należy do tej płaszczyzny, więc jest prostopadły do wektora U, zatem ich iloczyn skalarny $\overrightarrow{AX} \circ U = 0$. Współrzędne wektora $\overrightarrow{AX} = \begin{pmatrix} x - x_0 \\ y - y_0 \\ z - z_0 \end{pmatrix}$, więc, po podstawieniu, mamy

$$\begin{pmatrix} x - x_0 \\ y - y_0 \\ z - z_0 \end{pmatrix} \circ \begin{pmatrix} a \\ b \\ c \end{pmatrix} = 0$$

$$a(x - x_0) + b(y - y_0) + c(z - z_0) = 0$$

$$ax + by + cz - ax_0 - by_0 - cz_0 = 0$$

Przyjmując za $d = -ax_0 - by_0 - cz_0$, otrzymujemy, że dowolny punkt z tej płaszczyzny spełnia

$$\boxed{ax + by + cz + d = 0,}$$

zatem równanie to możemy przyjąć za ogólne równanie płaszczyzny Π.

✔ **Przykład.** Równanie płaszczyzny przechodzącej przez punkt $\begin{pmatrix} -1 \\ 1 \\ 2 \end{pmatrix}$ i prostopadłej do wektora $\begin{pmatrix} 3 \\ 5 \\ 7 \end{pmatrix}$ ma postać

$$3x + 5y + 7z + (3 - 5 - 14) = 0,$$

czyli

$$3x + 5y + 7z - 16 = 0.$$

Wniosek 2.5.1. *Każde równanie postaci $ax+by+cz+d = 0$, w którym przynajmniej jeden ze współczynników a, b, c jest niezerowy, opisuje pewną płaszczyznę.*

Dowód. Aby się o tym przekonać, weźmy dowolną trójkę liczb x_0, y_0, z_0 spełniającą to równanie. Otrzymujemy wtedy $ax_0 + by_0 + cz_0 + d = 0$ czyli $d = -ax_0 - by_0 - cz_0$. Równanie $ax+by+cz+d = 0$ możemy zapisać jako $a(x-x_0)+b(y-y_0)+c(z-z_0) = 0$, czyli

$$\begin{pmatrix} x - x_0 \\ y - y_0 \\ z - z_0 \end{pmatrix} \circ \begin{pmatrix} a \\ b \\ c \end{pmatrix} = 0.$$

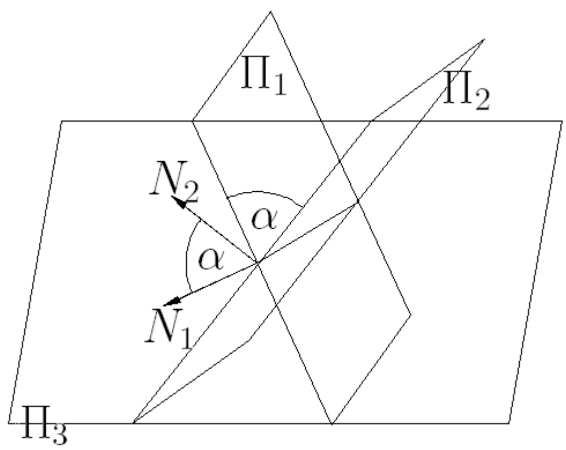
Zauważmy, że $(x - x_0, y - y_0, z - z_0)$ to współrzędne wektora łączącego dowolny punkt o współrzędnych $\begin{pmatrix} x \\ y \\ z \end{pmatrix}$, z punktem stałym $A_0(x_0, y_0, z_0)$. Równanie to opisuje zbiór punktów $A(x, y, z)$, dla których wektor $\overrightarrow{A_0A}$ jest prostopadły do wektora $\begin{pmatrix} a \\ b \\ c \end{pmatrix}$. Równanie to określa zatem płaszczyznę przechodzącą przez punkt (x_0, y_0, z_0) i prostopadłą do wektora $\begin{pmatrix} a \\ b \\ c \end{pmatrix}$. Wektor $\begin{pmatrix} a \\ b \\ c \end{pmatrix}$ nazywany jest **wektorem normalnym** płaszczyzny $ax + by + cz + d = 0$. □

Wniosek 2.5.2. *Jeżeli $d = 0$, to płaszczyzna przechodzi przez początek układu współrzędnych.*

Dowód. Jak wiemy z poprzedniego wniosku, równanie postaci $ax + by + cz + d = 0$ określa pewną płaszczyznę. Zauważmy, że jeżeli $d = 0$ to $x = 0, y = 0, z = 0$ spełniają to równanie, czyli punkt $(0, 0, 0)$ jest zawarty w płaszczyźnie określonej tym wzorem. □

2.6 Kąt pomiędzy płaszczyznami w R^3 zadanymi równaniami ogólnymi.

Dwie różne płaszczyzny w przestrzeni są równoległe lub się przecinają. W tym drugim przypadku naturalnym wydaje się pytanie pod jakim kątem się one przecinają. Rozważmy nierównoległe płaszczyzny Π_1, Π_2. Zauważmy, że przecinają się one wzdłuż prostej, nazwijmy ją L. Weźmy płaszczyzę Π_3, prostopadłą do rostej L. Zauważmy, że częścią wspólną płaszczyzn Π_3 i Π_1, oraz płaszczyzn Π_3 i Π_2, są proste. Kąt między tymi prostymi przyjmujemy za kąt między płaszczyznami Π_1, Π_2. Zauważmy, że kąt o takiej mierze tworzą również wektory normalne płaszczyzn Π_1, Π_2, zaczepione we wspólnym punkcie płaszczyzn Π_1, Π_2, Π_3.



Udowodnijmy, że te kąty rzeczywiście są sobie równe. Niech płaszczyzny Π_1, Π_2 będą dane równaniami

$$\Pi_1 : a_1x + b_1y + c_1z + d_1 = 0,$$

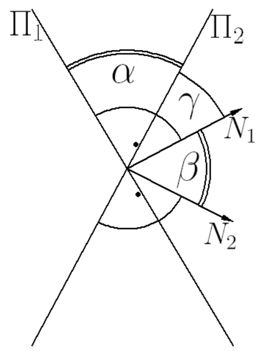
$$\Pi_2 : a_2x + b_2y + c_2z + d_2 = 0.$$

Wówczas wektory normalne tych płaszczyzn to odpowiednio

$$N_1 = \begin{pmatrix} a_1 \\ b_1 \\ c_1 \end{pmatrix}, \quad N_2 = \begin{pmatrix} a_2 \\ b_2 \\ c_2 \end{pmatrix}.$$

Oznaczmy kąt między wektorami normalnymi tych płaszczyzn przez β. Mogą wówczas zajść dwie sytuacje.

Przypadek 1 Poniższy rysunek przedstawia płaszczyznę Π_3, na której zaznaczone są jej przekroje z płaszczyznami Π_1, Π_2.

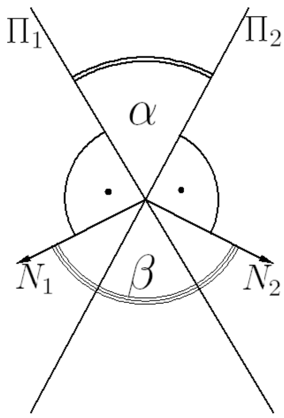


Zauważmy, że w tym wypadku $\alpha + \gamma = 90^0$, zatem $\alpha = 90^0 - \gamma$, ale również $\beta + \gamma = 90^0$, skąd mamy, że $\beta = 90^0 - \gamma$. Otrzymujemy zatem

$$\alpha = \beta.$$

Tak więc wybrany kąt pomiędzy płaszczyznami Π_1 i Π_2 jest taki sam jak kąt pomiędzy wektorami normalnymi tych płaszczyzn.

Przypadek 2 Podobnie jak poprzednio, rysunek przedstawia płaszczyznę Π_3 i jej przekroje z płaszczyznami Π_1, Π_2.



W tym wypadku mamy $\alpha+\beta+90^0+90^0=360^0$ a stąd $\alpha+\beta=180^0$. Otrzymujemy zatem $\alpha=180^0-\beta$.
Ostatecznie, po rozważeniu obu przypadków mamy

$$\alpha=\beta \text{ lub } \alpha=180^0-\beta.$$

Oznacza to, że jeden spośród kątów między płaszczyznami jest równy kątowi pomiędzy ich wektorami normalnymi, zaś drugi jest dopełnieniem tego kąta do 180^0.

Wniosek 2.6.1 (kąt między płaszczyznami). *Cosinus jednego z kątów pomiędzy płaszczyznami* Π_1, Π_2 *danymi równaniami*

$$\Pi_1 : a_1x+b_1y+c_1z+d_1=0,$$

$$\Pi_2 : a_2x+b_2y+c_2z+d_2=0$$

wyraża się wzorem

$$\boxed{\cos\alpha=\frac{N_1\circ N_2}{|N_1||N_2|}=\frac{a_1a_2+b_1b_2+c_1c_2}{\sqrt{a_1^2+b_1^2+c_1^2}\cdot\sqrt{a_2^2+b_2^2+c_2^2}}.} \tag{2.5}$$

✔ **Przykład.** Cosinus jednego z kątów pomiędzy płaszczyznami o równaniach $x+2y-3z+2=0$ oraz $2x-4y+z+1=0$, wynosi

$$\cos\alpha=\frac{2-8-3}{\sqrt{1+4+9}\cdot\sqrt{4+16+1}}=\frac{-9}{\sqrt{14}\cdot\sqrt{21}}=\frac{-9}{7\cdot\sqrt{6}}=-\frac{3\sqrt{6}}{14}\approx-0,52.$$

Zatem, jak łatwo możemy sprawdzić, ten kąt to w przybliżeniu

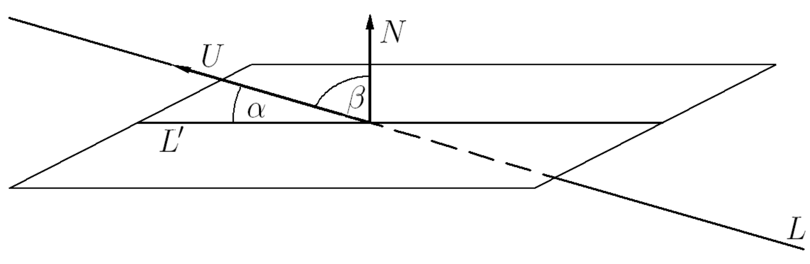
$$\alpha\approx122^0.$$

Otrzymaliśmy zatem, że jeden z kątów pomiędzy danymi płaszczyznami to kąt o mierze około 122^0.

2.7 Kąt pomiędzy płaszczyzną i prostą w R^3.

Postaramy się teraz wyznaczyć kąt pomiędzy płaszczyzną Π a prostą L. Oznaczmy przez L' rzut prostokątny prostej L na płaszczyznę Π. Kąt pomiędzy prostą L a płaszczyzną Π, to kąt pomiędzy prostą a jej rzutem na płaszczyznę.

Przyjmijmy, że płaszczyzna Π zadana jest równaniem $ax+by+cz+d=0$, natomiast L niech będzie prostą o wektorze kierunkowym U.

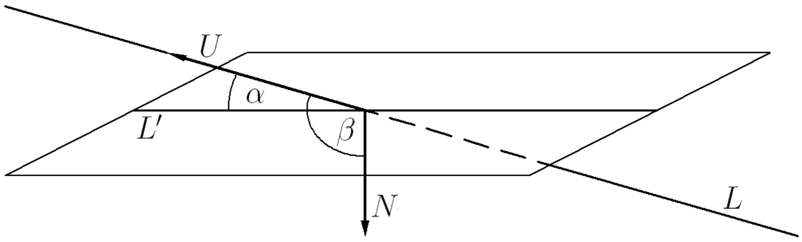


Niech β będzie kątem pomiędzy wektorem kierunkowym U a wektorem normalnym płaszczyzny Π, oznaczanym N. Zauważmy, że jeśli β jest kątem ostrym jak na rysunku powyżej, to mamy

$$\alpha = 90^0 - \beta.$$

Jeżeli natomiast β jest kątem rozwartym, jak na rysunku poniżej, to

$$\alpha = \beta - 90^0.$$



Mając dane równanie płaszczyzny Π oraz wektor kierunkowy U prostej L, możemy z tych współrzędnych wyliczyć kąt β, a następnie α, który jest jednym z kątów pomiędzy zadaną płaszczyzną a prostą.

Przykład. Znajdźmy kąt pomiędzy płaszczyzną Π o równaniu $2x - y + 2z - 4 = 0$ a prostą o wektorze kierunkowym $U = \begin{pmatrix} 1 \\ -2 \\ -1 \end{pmatrix}$. Wektor normalny płaszczyzny Π to $N = \begin{pmatrix} 2 \\ -1 \\ 2 \end{pmatrix}$. Cosinus kąta pomiędzy wektorami U i N, jak wynika z Wniosku 2.2.2, wynosi

$$\cos\beta = \frac{2}{\sqrt{9}\cdot\sqrt{6}} = \frac{\sqrt{6}}{9} \approx 0,2722.$$

Kąt β jest więc kątem ostrym, zatem α wyraża się wzorem $\alpha = 90^0 - \beta$, czyli $\sin\alpha = \cos\beta$. Stąd $\alpha \approx 15^0$. Tak więc kąt pomiędzy płaszczyzną Π a prostą o wektorze kierunkowym U, to kąt o mierze około 15^0.

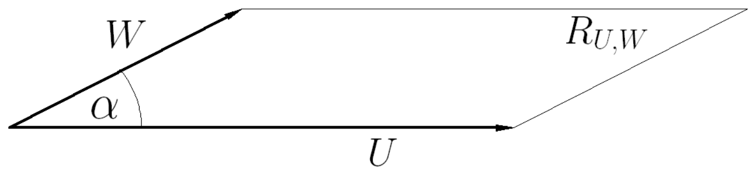
2.8 Pole równoległoboku zbudowanego na wektorach U, W w R^3.

W tym rozdziale zastosujemy iloczyn skalarny do obliczania pól równoległoboków w przestrzeni R^3.

Weźmy dwa dowolne wektory z przestrzeni R^3

$$U = \begin{pmatrix} u_1 \\ u_2 \\ u_3 \end{pmatrix}, \quad W = \begin{pmatrix} w_1 \\ w_2 \\ w_3 \end{pmatrix}.$$

Zauważmy, że wektory te wyznaczają pewien równoległobok. Równoległobok ten oznaczać będziemy $R_{U,W}$.



Z geometrii wiemy, że

$$P(R_{U,W}) = |U| \cdot |W| \cdot \sin\alpha,$$

gdzie α to kąt pomiędzy wektorami U i W. Mamy ponadto

$$\sin\alpha = \sqrt{1-\cos^2\alpha}, \quad \text{gdy} \ \ 0^0 < \alpha < 180^0.$$

Korzystając z tych zależności oraz ze wzoru (2.5), otrzymujemy

$$P(R_{U,W}) = |U| \cdot |W| \cdot \sin\alpha = |U||W|\sqrt{1-\cos^2\alpha} = |U||W|\sqrt{1-\left(\frac{U\circ W}{|U||W|}\right)^2} =$$

$$= \sqrt{|U|^2|W|^2 - |U|^2|W|^2\frac{(U\circ W)^2}{|U|^2|W|^2}} = \sqrt{|U|^2|W|^2-(U\circ W)^2}.$$

Podsumowując - pole równoległoboku rozpiętego przez wektory U, W wynosi

$$\boxed{P(R_{U,W}) = \sqrt{|U|^2|W|^2-(U\circ W)^2}.} \tag{2.6}$$

Rozdział 3

Iloczyn wektorowy

W tym rozdziale zajmiemy się iloczynem wektorowym. W pierwszym podrozdziale zdefiniujemy to pojęcie, w kolejnym podamy jego podstawowe własności, pokażemy jak możemy je wykorzystać. Ostatni podrozdział poświęcimy na podanie geometrycznej interpretacji iloczynu wektorowego, która pozwoli na wprowadzenie kolejnej bardzo użytecznej własności - związku iloczynu wektorowego z objętością równoległościanu.

3.1 Definicja iloczynu wektorowego.

Dla danego wektora $A \in R^2$ nietrudno podać przykład wektora $A' \in R^2$ prostopadłego do A. Dla $A = \begin{pmatrix} a_1 \\ a_2 \end{pmatrix}$ jest nim wektor o współrzędnych $A' = \begin{pmatrix} -a_2 \\ a_1 \end{pmatrix}$. Zastanówmy się teraz w jaki sposób dla danych dwóch wektorów $A, B \in R^3$ znaleźć wektor prostopadły do nich obu. Sprawdźmy, że dla $A = \begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix}$, $B = \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix}$ wektor $X = \begin{pmatrix} a_2b_3 - b_2a_3 \\ -a_1b_3 + b_1a_3 \\ a_1b_2 - b_1a_2 \end{pmatrix}$ jest właśnie takim wektorem, czyli jest prostopadły do A i do B. Jak wynika z Wniosku 2.2.3, dwa wektory są prostopadłe, jeżeli ich iloczyn skalarny jest równy 0. Sprawdźmy zatem prostopadłość X i A.

$$X \circ A = \begin{pmatrix} a_2b_3 - b_2a_3 \\ -a_1b_3 + b_1a_3 \\ a_1b_2 - b_1a_2 \end{pmatrix} \circ \begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix} =$$

$$= a_2b_3a_1 - b_2a_3a_1 - a_1b_3a_2 + b_1a_3a_2 + a_1b_2a_3 - b_1a_2a_3 =$$

$$= b_1(a_3a_2 - a_2a_3) + b_2(a_1a_3 - a_3a_1) + b_3(a_2a_1 - a_1a_2) = 0.$$

Podobnie sprawdźmy prostopadłość X i B.

$$X \circ B = \begin{pmatrix} a_2b_3 - b_2a_3 \\ -a_1b_3 + b_1a_3 \\ a_1b_2 - b_1a_2 \end{pmatrix} \circ \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix} =$$

$$= a_2b_3b_1 - b_2a_3b_1 - a_1b_3b_2 + b_1a_3b_2 + a_1b_2b_3 - b_1a_2b_3 =$$
$$= a_1(b_2b_3 - b_3b_2) + a_2(b_3b_1 - b_1b_3) + a_3(b_1b_2 - b_2b_1) = 0.$$

Tak określony wektor X jest zatem dobrym przykładem wektora prostopadłego do obu wektorów A oraz B.

Definicja 3.1.1 (iloczyn wektorowy). Wektor X określony powyżej, czyli

$$\boxed{X = \begin{pmatrix} a_2b_3 - b_2a_3 \\ -a_1b_3 + b_1a_3 \\ a_1b_2 - b_1a_2 \end{pmatrix},} \tag{3.1}$$

nazywany jest iloczynem wektorowym wektorów A i B, co oznaczamy przez

$$X = A \times B.$$

✔ **Przykład.** Iloczynem wektorowym wektorów $A = \begin{pmatrix} 3 \\ 0 \\ -2 \end{pmatrix}$, $B = \begin{pmatrix} -1 \\ -2 \\ 1 \end{pmatrix}$ jest wektor $X = \begin{pmatrix} -4 \\ -1 \\ -6 \end{pmatrix}$.

➢ Sposób na zapamiętanie wzoru na iloczyn wektorowy

Wprowadzimy teraz interpretację iloczynu wektorowego, która okaże się być łatwym sposobem na zapamiętanie wzoru na ten iloczyn. Interpretacja ta korzysta z pojęcia wyznacznika macierzy 2×2. W tym celu przypomnijmy, że wyznacznikiem macierzy $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ nazywamy liczbę określoną wzorem $\det A = a \cdot d - b \cdot c$.

Uwaga 3.1.2 (wyznacznikowa interpretacja iloczynu wektorowego). *Jeśli* $\begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix} \times \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$, *to*

$$x_1 = \det \begin{pmatrix} a_2 & b_2 \\ a_3 & b_3 \end{pmatrix}, \ x_2 = -\det \begin{pmatrix} a_1 & b_1 \\ a_3 & b_3 \end{pmatrix}, \ x_3 = \det \begin{pmatrix} a_1 & b_1 \\ a_2 & b_2 \end{pmatrix}.$$

Zauważmy, że macierz, której wyznacznik liczymy w celu wyznaczenia x_1, powstaje przez wykreślenie z macierzy $\begin{pmatrix} a_1 & b_1 \\ a_2 & b_2 \\ a_3 & b_3 \end{pmatrix}$ pierwszego wiersza, podobnie macierz pojawiająca się we wzorze na x_2 powstaje przez wykreślenie z macierzy $\begin{pmatrix} a_1 & b_1 \\ a_2 & b_2 \\ a_3 & b_3 \end{pmatrix}$ drugiego wiersza. Analogicznie powstaje trzecia macierz.

3.2 Własności iloczynu wektorowego.

Podamy teraz najważniejsze algebraiczne własności iloczynu wektorowego. Wektory A, B, C, występujące poniżej, to dowolne wektory z przestrzeni R^3, zaś t jest dowolną liczbą rzeczywistą.

(1) $A \times B = -B \times A$. Własność ta nazywana jest **antyprzemiennością** iloczynu wektorowego.

(2) Jeśli $A = \overrightarrow{O} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$, to $A \times B = B \times A = \overrightarrow{O}$.

(3) Jeśli $A = t \cdot B$, to $A \times B = \overrightarrow{O}$.

(4) $A \times (B + C) = A \times B + A \times C$

(5) $(tA) \times B = t(A \times B)$

Dowód (1).
Własność ta wynika z antyprzemienności wyznacznika 2×2. Zauważmy, że

$$\det \begin{pmatrix} a & b \\ c & d \end{pmatrix} = ad - bc = -(bc - ad) = -\det \begin{pmatrix} b & a \\ d & c \end{pmatrix}.$$

Zatem, jeśli $A \times B = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$, a $B \times A = \begin{pmatrix} x'_1 \\ x'_2 \\ x'_3 \end{pmatrix}$, to korzystając z powyższej równości oraz z wyznacznikowej interpretacji iloczynu skalarnego, którą wprowadziliśmy w Uwadze 3.1.2, dla $A = \begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix}$, $B = \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix}$ mamy

$$x'_1 = \det \begin{pmatrix} b_2 & a_2 \\ b_3 & a_3 \end{pmatrix} = -\det \begin{pmatrix} a_2 & b_2 \\ a_3 & b_3 \end{pmatrix} = -x_1.$$

Podobnie $x'_2 = -x_2$ oraz $x'_3 = -x_3$. Otrzymujemy zatem $B \times A = -A \times B$. □

Dowód (2).
Zauważmy, że dla $A = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$, $B = \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix}$, jeżeli $A \times B = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$, to korzystając z wyznacznikowej interpretacji iloczynu skalarnego, otrzymujemy

$$x_1 = \det \begin{pmatrix} 0 & b_2 \\ 0 & b_3 \end{pmatrix} = 0.$$

Podobnie $x_2 = 0$ oraz $x_3 = 0$. □

Dowód (3).

Z wyznacznikowej interpretacji iloczynu skalarnego, dla $A \times B = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$, gdzie $A = \begin{pmatrix} tb_1 \\ tb_2 \\ tb_3 \end{pmatrix}, B = \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix}$, mamy

$$x_1 = \det \begin{pmatrix} tb_2 & b_2 \\ tb_3 & b_3 \end{pmatrix} = tb_2 b_3 - tb_2 b_3 = 0.$$

Podobnie $x_2 = 0,\ x_3 = 0$. □

Dowód (4).
Zauważmy, że

$$\det \begin{pmatrix} a_2 & b_2 + c_2 \\ a_3 & b_3 + c_3 \end{pmatrix} = \det \begin{pmatrix} a_2 & b_2 \\ a_3 & b_3 \end{pmatrix} + \det \begin{pmatrix} a_2 & c_2 \\ a_3 & c_3 \end{pmatrix}.$$

Zatem pierwsza współrzędna wektora $A \times (B + C)$ jest równa pierwszej współrzędnej wektora $A \times B + A \times C$. Podobnie druga i trzecia współrzędna. Otrzymujemy zatem $A \times (B + C) = A \times B + A \times C$. □

Dowód (5).
Dla $A = \begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix}$, $B = \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix}$ jeśli $A \times B = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$, $(tA) \times B = \begin{pmatrix} x_1' \\ x_2' \\ x_3' \end{pmatrix}$, to

$$x_1' = \det \begin{pmatrix} ta_2 & b_2 \\ ta_3 & b_3 \end{pmatrix} = t \cdot \det \begin{pmatrix} a_2 & b_2 \\ a_3 & b_3 \end{pmatrix} = t \cdot x_1.$$

Podobnie $x_2' = t \cdot x_2,\ x_3' = t \cdot x_3$. Otrzymaliśmy zatem $(tA) \times B = t(A \times B)$. □

Bezpośrednią konsekwencją własności (2) i (3) jest wniosek pozwalający na bardzo szybkie sprawdzenie liniowej zależności pary wektorów.

Wniosek 3.2.1 (iloczyn wektorowy a liniowa zależność wektorów). *Jeśli wektory A i B są liniowo zależne, to $A \times B = \overrightarrow{O}$.*

3.3 Geometryczna interpretacja iloczynu wektorowego.

Iloczyn wektorowy jest wektorem, którego kierunek, długość i zwrot możemy wyznaczyć, korzystając z jego geometrycznych własności.

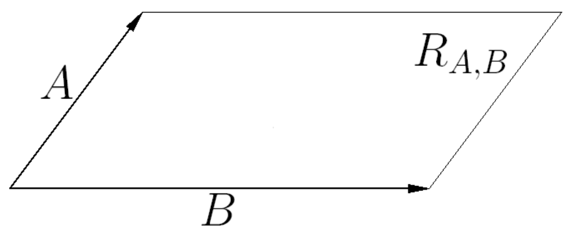
➢ Kierunek iloczynu wektorowego

Fakt 3.3.1. *(1) Jeśli wektory A i B są współliniowe, to $A \times B = \overrightarrow{O}$.*
(2) Jeśli A i B są niewspółliniowe, to wektor $A \times B$ jest prostopadły zarówno do A jak i do B, czyli wyznacza kierunek prostopadły do płaszczyzny rozpiętej przez wektory A i B.

Dowód. Pierwsza część powyższego faktu wynika z trzeciej własności udowodnionej w poprzednim podrozdziale. Precyzyjny dowód pomijamy. Prostopadłość wektorów $A \times B$ i A oraz $A \times B$ i B została pokazana na początku tego rozdziału. Zauważmy, że jeżeli A i B nie są współliniowe, to wyznaczają płaszczyznę, więc jeśli wektor $A \times B$ jest prostopadły do A i do B, to jest prostopadły również do płaszczyzny wyznaczonej przez te wektory, wyznacza więc kierunek do niej prostopadły. □

➢ Długość iloczynu wektorowego

Niech $R_{A,B}$ będzie równoległobokiem rozpiętym przez wektory A i B.



Związek iloczynu wektorowego $A \times B$ z tym równoległobokiem określa następujący fakt.

Fakt 3.3.2 (iloczyn wektorowy a pole równoległoboku). *Długość wektora $A \times B$ jest równa polu równoległoboku $R_{A,B}$.*

Dowód. Niech $A = \begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix}$, $B = \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix}$, i niech α oznacza miarę kąta pomiędzy wektorami A i B. Przypomnijmy, że wtedy pole równoległoboku $R_{A,B}$ można obliczyć korzystając ze wzoru $P(R_{A,B}) = |A| \cdot |B| \cdot \sin\alpha$. Mamy również

$$|A \times B|^2 = (a_2b_3 - b_2a_3)^2 + (-a_1b_3 + b_1a_3)^2 + (a_1b_2 - b_1a_2)^2.$$

Po łatwym przekształceniu, które pozostawimy do wykonania czytelnikowi, otrzymujemy

$$|A \times B|^2 = (a_1^2 + a_2^2 + a_3^2) \cdot (b_1^2 + b_2^2 + b_3^3) - (a_1b_1 + a_2b_2 + a_3b_3)^2 = |A|^2 \cdot |B|^2 - (A \circ B)^2 =$$
$$= |A|^2 \cdot |B|^2 - (|A| \cdot |B| \cdot \cos\alpha)^2 = |A|^2 \cdot |B|^2 \cdot (1 - \cos^2\alpha) =$$
$$= |A|^2 \cdot |B|^2 \cdot \sin^2\alpha = [P(R_{A,B})]^2.$$

Pierwiastkując obie strony, otrzymujemy $|A \times B| = P(R_{A,B})$, tak jak chcieliśmy. □

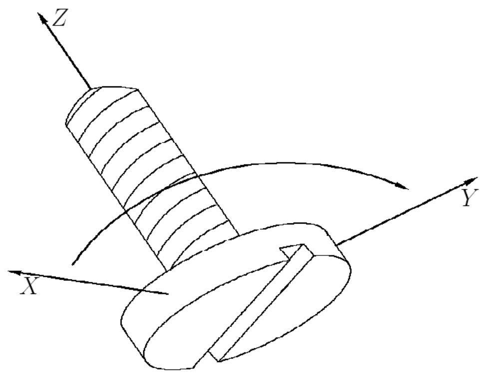
Wniosek 3.3.3. *$A \times B$ to wektor prostopadły do płaszczyzny wyznaczonej przez A i B, którego długość jest równa polu równoległoboku rozpiętego przez A i B.*

➢ Zwrot iloczynu wektorowego

Dla dowolnych dwóch wektorów mamy zatem jednoznacznie wyznaczony kierunek i długość wektora będącego ich iloczynem wektorowym. Aby mieć jednoznacznie wyznaczony wektor musimy jeszcze określić jego zwrot. W tym celu zastosujemy regułę śruby prawoskrętnej.

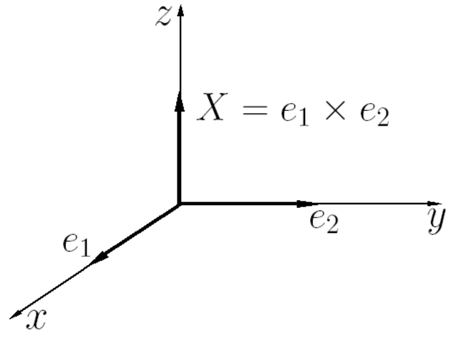
Definicja 3.3.4 (**śruba prawoskrętna**)**.** Śrubą prawoskrętną nazywamy śrubę, która wkręca się, gdy jest obracana w prawo.

Kręcąc śrubą od X do Y wkręca się ona w kierunku Z.



Fakt 3.3.5 (**zwrot wektora** $A \times B$)**.** *Zwrot wektora $A \times B$ jest taki jak kierunek, w jakim wkręca się śruba prawoskrętna, gdy jest obracana od A do B.*

✔ **Przykład.** Spróbujmy teraz, bez odwoływania się do wzoru, korzystając tylko z geometrycznej interpretacji, znaleźć iloczyn wektorowy X wektorów $e_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$, $e_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$. Jak wiemy z Faktu 3.3.1, X będzie wektorem prostopadłym do płaszczyzny wyznaczonej przez e_1 i e_2, czyli do płaszczyzny O_{xy}. Równoległobok wyznaczony przez e_1 i e_2 to kwadrat o boku długości 1, zatem, jak wynika z Faktu 3.3.2, $|A \times B| = 1$. Szukanym wektorem X mogą być zatem wektory $\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$ lub $\begin{pmatrix} 0 \\ 0 \\ -1 \end{pmatrix}$.



Stosując regułę śruby prawoskrętnej, otrzymujemy

$$X = e_1 \times e_2 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}.$$

✍ **Ćwiczenie.** Jako sprawdzenie poprawności tak wyznaczonego iloczynu wektorowego, czytelnik może wyznaczyć ten wektor korzystając ze wzoru (3.1).

Rozdział 4

Wyznacznik macierzy 3×3

W matematyce istnieją pojęcia, które znajdują szerokie zastosowanie w różnych dziedzinach. Jednym z takich pojęć jest wyznacznik macierzy kwadratowej. Odgrywa on ważną rolę między innymi przy rozwiązywaniu układów równań liniowych z wieloma niewiadomymi, oraz w geometrii. Na początek poznamy definicję wyznacznika, następnie omówimy jego najważniejsze własności. Na koniec zajmiemy się niektórymi jego zastosowaniami.

4.1 Podstawowe definicje.

W pierwszej części skryptu wprowadzone zostało pojęcie wyznacznika macierzy rozmiaru 2×2. Podamy teraz definicję odpowiednika tego pojęcia dla macierzy 3×3.

Definicja 4.1.1 (wyznacznik macierzy). Wyznacznikiem macierzy $M = \begin{pmatrix} a & b & c \\ d & e & f \\ g & h & i \end{pmatrix}$ nazywamy wyrażenie algebraiczne postaci

$$\boxed{\det M = aei + bfg + cdh - gec - hfa - idb.} \tag{4.1}$$

✔ **Przykład.** Wyznacznik macierzy $M = \begin{pmatrix} -2 & 3 & 0 \\ 1 & 2 & 5 \\ 7 & 1 & 4 \end{pmatrix}$ wynosi

$$\det M = -16 + 105 + 0 - 0 + 10 - 12 = 87.$$

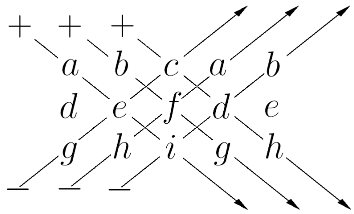
Poniżej podamy łatwy sposób na obliczanie wyznacznika macierzy bez konieczności pamiętania wzoru.

Uwaga 4.1.2 (reguła Sarrusa). *W praktyce w celu obliczenia wyznacznika stosować będziemy regułę Sarrusa.*

Do danej macierzy $\begin{pmatrix} a & b & c \\ d & e & f \\ g & h & i \end{pmatrix}$ *dopisujemy kolumnę pierwszą i drugą. Otrzymujemy układ*

$$\begin{matrix} a & b & c & a & b \\ d & e & f & d & e. \\ g & h & i & g & h \end{matrix}$$

Następnie dodajemy do siebie iloczyny wyrazów znajdujących się na ukośnych prostych wychodzących z trzech pierwszych wyrazów pierwszego wiersza, a od wyniku odejmujemy iloczyny wyrazów znajdujących się na ukośnych prostych wychodzących z trzech pierwszych wyrazów ostatniego wiersza, tak jak to pokazuje poniższy schemat.



Otrzymujemy wówczas podany w Definicji 18.3.1 wzór na wyznacznik danej macierzy.

Można również określić wyznacznik dla trójki wektorów. Niech
$A = \begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix}$, $B = \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix}$, $C = \begin{pmatrix} c_1 \\ c_2 \\ c_3 \end{pmatrix}$.

Definicja 4.1.3 (wyznacznik wektorów). Wyznacznikiem wektorów A, B, C nazywamy liczbę

$$\det(A, B, C) = \det \begin{pmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{pmatrix}.$$

Przykład. Wyznacznik wektorów $A = \begin{pmatrix} 3 \\ 5 \\ 1 \end{pmatrix}$, $B = \begin{pmatrix} -2 \\ 1 \\ 3 \end{pmatrix}$, $C = \begin{pmatrix} 1 \\ 3 \\ 0 \end{pmatrix}$ wynosi
$\det \begin{pmatrix} 3 & -2 & 1 \\ 5 & 1 & 3 \\ 1 & 3 & 0 \end{pmatrix} = 0 - 6 + 15 - 1 - 27 - 0 = -19.$

4.2 Własności wyznacznika.

Poniżej określimy najważniejsze, zarówno arytmetyczne, jak i geometryczne własności wyznacznika.

➢ Wyznacznik jako iloczyn mieszany

Fakt 4.2.1 (wyznacznik a iloczyn wektorów). *Dla $A, B, C \in R^3$ mamy*

$$\det(A, B, C) = A \circ (B \times C) = (A \times B) \circ C.$$

Ze względu na występujące w powyższej tożsamości pojęcia iloczynu skalarnego i wektorowego, wyznacznik trójki wektorów bywa często nazywany iloczynem mieszanym tych wektorów.

Dowód. Korzystając z definicji iloczynu skalarnego i wektorowego mamy

$$A \circ (B \times C) = (A \times B) \circ C = \begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix} \circ \begin{pmatrix} b_2c_3 - c_2b_3 \\ -b_1c_3 + c_1b_3 \\ b_1c_2 - c_1b_2 \end{pmatrix} =$$
$$= a_1(b_2c_3 - c_2b_3) + a_2(-b_1c_3 + c_1b_3) + a_3(b_1c_2 - c_1b_2) =$$
$$= a_1b_2c_3 - a_1c_2b_3 - a_2b_1c_3 + a_2c_1b_3 + a_3b_1c_2 - a_3c_1b_2 = \det(A, B, C)$$

Podobnie pokazujemy, że $\det(A, B, C) = (A \times B) \circ C$. □

Fakt 4.2.2. *Dla $A, B, C \in R^3$ mamy*

$$\det(A, B, C) = -\det(B, A, C) = -\det(A, C, B) = -\det(C, B, A) =$$
$$= \det(B, C, A) = \det(C, A, B).$$

Dowód. Korzystając z poprzedniego faktu oraz z własności iloczynu wektorowego otrzymujemy

$$\det(B, A, C) = (B \times A) \circ C = (-A \times B) \circ C = -(A \times B) \circ C = -\det(A, B, C),$$
$$\det(B, C, A) = (B \times C) \circ C = A \circ (B \times C) = \det(A, B, C).$$

Dla pozostałych równości dowody wyglądają analogicznie. Otrzymaliśmy więc, że zamiana dwóch kolumn w macierzy, której wyznacznik liczymy, powoduje zmianę znaku wyznacznika, natomiast cykliczne przestawienie kolumn macierzy nie zmienia wartości wyznacznika. □

✍ **Ćwiczenie.** Dowody pozostałych równości, pozostawimy do wykonania czytelnikowi.

Uwaga 4.2.3. *Znak „-”pojawia się, gdy dokonujemy zamiany którychkolwiek dwóch wektorów spośród A, B, C. Takie zamiany dwóch wektorów nazywamy **transpozycjami**. Zatem wyznacznik zmienia się na przeciwny przy transpozycji którychkolwiek dwóch wektorów. Wyznacznik nie zmienia się, przy przestawieniu cyklicznym, np. gdy wektory przesuniemy o jedno miejsce do przodu, zaś pierwszy przeniesiemy na koniec.*

➢ Macierz transponowana i jej wyznacznik

Bardzo ważnym pojęciem związanym z macierzami jest macierz transponowana do danej macierzy.

Definicja 4.2.4 (macierz transponowana). Jeśli $m = \begin{pmatrix} a & b & c \\ d & e & f \\ g & h & k \end{pmatrix}$, to macierzą transponowaną do macierzy, oznaczaną symbolem m^T, nazywamy macierz

$$m^T = \begin{pmatrix} a & d & g \\ b & e & h \\ c & f & k \end{pmatrix},$$

otrzymaną z m przez zamianę wyrazów położonych symetrycznie względem przekątnej głównej.

Uwaga 4.2.5. *Macierz transponowaną można również uzyskać poprzez zapisanie kolumn macierzy m jako wierszy w macierzy m^T.*

Jak nietrudno się przekonać wyznaczniki macierzy m i macierzy do niej transponowanej są sobie równe, o czym mówi następujący fakt.

Fakt 4.2.6 (wyznacznik macierzy transponowanej). *Dla macierzy m mamy*

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

✍ **Ćwiczenie.** Dowód tej własności, wynikający wprost z definicji wyznacznika, pozostawimy do wykonania czytelnikowi.

➢ Wyznacznik a liniowa niezależność

Podamy teraz bardzo użyteczną własność wyznacznika, dzięki której w łatwy sposób można sprawdzić, czy trójka wektorów jest liniowo niezależna.

Fakt 4.2.7. ***(wyznacznik a liniowa zależność)*** *Wyznacznik wektorów $A, B, C \in R^3$ jest równy zero wtedy i tylko wtedy gdy wektory te są liniowo zależne.*

Dowód. Niech $\det(A, B, C) = 0$. Gdyby wektory A, B, C były liniowo niezależne, wtedy mielibyśmy, że wektory A, B są niewspółliniowe, czyli wektor $A \times B$ jest niezerowy. Ponadto C nie należałoby do płaszczyzny wyznaczonej przez A i B, więc nie byłby prostopadły do $A \times B$. Zatem $(A \times B) \circ C \neq 0$, ale wiemy, że $(A \times B) \circ C = \det(A, B, C) = 0$. Otrzymaliśmy więc sprzeczność, która oznacza, że jeśli $\det(A, B, C) = 0$, to wektory A, B, C są liniowo zależne.
Udowodnijmy jeszcze implikację w drugą stronę. Załóżmy, że A, B, C są liniowo zależne czyli np. A należy do płaszczyzny wyznaczonej przez B i C. Wynika stąd, że A jest prostopadły do $B \times C$, więc $A \circ (B \times C) = 0$, ale jak wiemy $A \circ (B \times C) = \det(A, B, C)$, stąd $\det(A, B, C) = 0$. Otrzymaliśmy więc, że jeśli wektory są liniowo zależne, to ich wyznacznik musi być równy zero. □

✔ **Przykład.** Sprawdźmy, czy wektory $A = \begin{pmatrix} 2 \\ 0 \\ 1 \end{pmatrix}$, $B = \begin{pmatrix} -1 \\ 1 \\ -1 \end{pmatrix}$, $C = \begin{pmatrix} 3 \\ 1 \\ 2 \end{pmatrix}$ są liniowo niezależne. W tym celu obliczmy wyznacznik tych wektorów.

$$\det \begin{pmatrix} 2 & -1 & 3 \\ 0 & 1 & 1 \\ 1 & -1 & 2 \end{pmatrix} = 4 - 1 + 0 - 3 + 2 - 0 = 2$$

Wyznacznik ten jest niezerowy więc wektory A, B, C są liniowo niezależne.

➢ Wieloliniowość wyznacznika

Określimy teraz własności, które nazywane są wieloliniowością wyznacznika.

Fakt 4.2.8 (wieloliniowość wyznacznika). *Dla wektorów $A, A_1, A_2, B, C \in R^3$ oraz liczby rzeczywistej t mamy*

(a) $\det(tA, B, C) = \det(A, tB, C) = \det(A, B, tC) = t \cdot \det(A, B, C)$,

(b) $\det(A_1 + A_2, B, C) = \det(A_1, B, C) + \det(A_2, B, C)$.

Dowód powyższego faktu łatwo wynika z podstawowych własności iloczynu skalarnego i wektorowego, dlatego go pominiemy.

Dla $A, B, B_1, B_2, C, C_1, C_2 \in R^3$, możemy wyprowadzić wzory podobne do tych z podpunktu (b) w Fakcie 4.2.8.

Wniosek 4.2.9.

$$\det(A, B_1 + B_2, C) = \det(A, B_1, C) + \det(A, B_2, C),$$

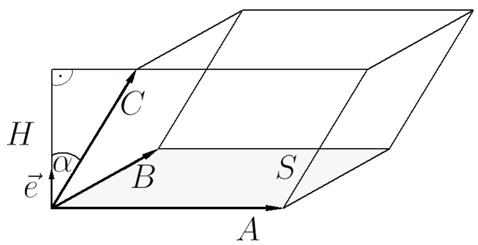
$$\det(A, B, C_1 + C_2) = \det(A, B, C_1) + \det(A, B, C_2).$$

Zajmijmy się geometrycznymi własnościami wyznacznika.

➢ Wyznacznik a objętość równoległościanu i czworościanu

Fakt 4.2.10 (wyznacznik a objętość równoległościanu). *Dla $A, B, C \in R^3$ wartość bezwzględna wyznacznika $|\det(A, B, C)|$ jest równa objętości równoległościanu $R_{A,B,C}$ zbudowanego na wektorach A, B, C.*

Dowód. Niech S będzie polem podstawy równoległościanu $R_{A,B,C}$, czyli polem równoległoboku $R_{A,B}$ zbudowanego na wektorach A, B (patrz rysunek). Niech $\vec{e}$ będzie wektorem jednostkowym prostopadłym do płaszczyzny podstawy, o zwrocie takim jak iloczyn wektorrowy $A \times B$. Jak wiemy z Faktu 3.3.2 $|A \times B| = S$, co daje $(A \times B) = S \cdot \vec{e}$.



Otrzymujemy zatem

$$|\det(A, B, C)| = |(A \times B) \circ C| = |(S \cdot \vec{e}) \circ C| = S \cdot |\vec{e} \circ C| = S \cdot |e| \cdot |C| \cdot \cos\alpha.$$

Jak nietrudno się przekonać $|C| \cdot \cos\alpha = H$ jest wysokością równoległościanu. Wiemy również, że $|\vec{e}| = 1$, a stąd $|\det(A, B, C)| = S \cdot H = V(R_{A,B,C})$. □

✔ **Przykład.** Objętość równoległościanu zbudowanego na wektorach $A = \begin{pmatrix} 1 \\ 1 \\ -1 \end{pmatrix}$, $B = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}$, $C = \begin{pmatrix} 2 \\ 3 \\ 2 \end{pmatrix}$ wynosi $\left|\det \begin{pmatrix} 1 & 1 & 2 \\ 1 & 0 & 3 \\ -1 & 1 & 2 \end{pmatrix}\right| = |0 - 3 + 2 - 0 - 3 - 2| = 6.$

Bezpośrednią konsekwencją powyższego faktu jest następujący wniosek.

Wniosek 4.2.11 (objętość czworościanu). *Objętość czworościanu zbudowanego na wektorach A, B, C wynosi* $\frac{1}{6}|\det(A, B, C)|$.

Dowód. Zauważmy, że pole trójkąta zbudowanego na wektorach A, B jest równe połowie pola równoległoboku zbudowanego na tych wektorach. Oznaczmy ten trójkąt $T_{A,B}$, czworościan zbudowany na A, B, C oznaczmy przez $C_{A,B,C}$. Otrzymujemy więc

$$V(C_{A,B,C}) = \frac{1}{3} \cdot P(T_{A,B}) \cdot H = \frac{1}{3} \cdot \frac{1}{2} \cdot P(R_{A,B}) \cdot H = \frac{1}{6} \cdot P(R_{A,B}) \cdot H = $$
$$= \frac{1}{6} \cdot V(R_{A,B,C}) = \frac{1}{6} \cdot |\det(A, B, C)|.$$

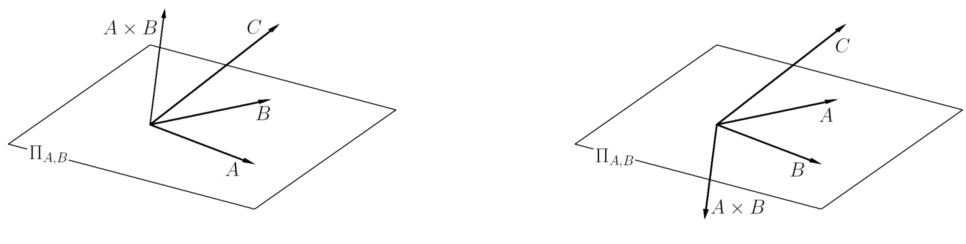
□

➢ Znak wyznacznika i orientacja

Dla trójki liniowo niezależnych wektorów A, B, C wprowadzamy pojęcie ich orientacji w przestrzeni.

Definicja 4.2.12 (orientacja trójki wektorów). Uporządkowana trójka wektorów A, B, C jest prawoskrętna (dodatnio zorientowana), gdy wektor C tworzy z wektorem $A \times B$ kąt $\alpha < 90^0$. Podobnie trójka wektorów jest lewoskrętna (ujemnie zorientowana), gdy kąt α pomiędzy C a $A \times B$ spełnia warunek $\alpha > 90^0$.

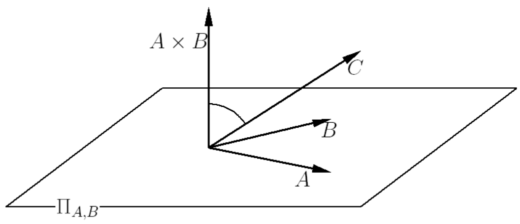
Równoważnie możemy powiedzieć, że układ jest dodatnio zorientowany, gdy C leży po tej samej stronie płaszczyzny wyznaczonej przez A i B, co $A \times B$, natomiast jest ujemnie zorientowany, gdy leży po przeciwnej stronie tej płaszczyzny.



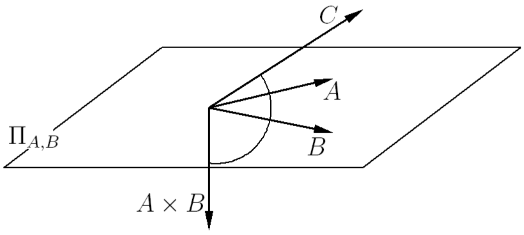
Rysunek po lewej stronie przedstawia trójkę wektorów dodatnio zorientowanych, rysunek po prawej stronie - ujemnie zorientowanych.

Uwaga 4.2.13 (orientacja a wyznacznik). *Trójka wektorów A, B, C jest prawoskrętna wtedy i tylko wtedy gdy* $\det(A, B, C) > 0$, *natomiast lewoskrętna, gdy* $\det(A, B, C) < 0$.

Dowód. Zauważmy, że jeśli wektor $A \times B$ leży po tej samej stronie płaszczyzny co wektor C, to mniejszy z kątów między tymi wektorami jest mniejszy niż 90^0, a w takim wypadku cosinus tego kąta jest dodatni.



Jak nietrudno się przekonać, cosinus kąta między wektorami $A \times B$ a C ma znak taki sam jak iloczyn skalarny tych wektorów, czyli $(A \times B) \circ C > 0$ gdy $(\measuredangle A \times B, C) < 90^0$, ale jak wiemy $(A \times B) \circ C = \det(A, B, C)$. Możemy wnioskować, że trójka wektorów jest dodatnio zorientowana, gdy $\det(A, B, C) > 0$. Podobnie możemy pokazać, że trójka wektorów jest ujemnie zorientowana, gdy mniejszy z kątów pomiędzy wektorami $A \times B$ a C jest większy niż 90^0.



Oznacza to, że jego cosinus jest mniejszy od zera a tym samym $\det(A, B, C) < 0$. $\square$

✔ **Przykład.** Wektory $A = \begin{pmatrix} \frac{1}{2} \\ 1 \\ -2 \end{pmatrix}$, $B = \begin{pmatrix} 3 \\ 0 \\ 1 \end{pmatrix}$, $C = \begin{pmatrix} -2 \\ 2 \\ 1 \end{pmatrix}$ tworzą układ ujemnie zorientowany ponieważ $\det \begin{pmatrix} \frac{1}{2} & 3 & -2 \\ 1 & 0 & 2 \\ -2 & 1 & 1 \end{pmatrix} = 0 - 12 - 2 - 0 - 1 - 3 = -18.$

Uwaga 4.2.14. *Orientacja wektorów zależy od kolejności wektorów. Przykładowo trójka wektorów zmienia orientację na przeciwną po przestawieniu dwóch wektorów.*

✍ **Ćwiczenie.** Ustal dokładnie, w jaki sposób orientacja trzech wektorów zmienia się pod wpływem zmian ich porządku. Kiedy zmienia się na przeciwną a kiedy pozostaje bez zmian. (Porównaj Fakt 4.2.2).

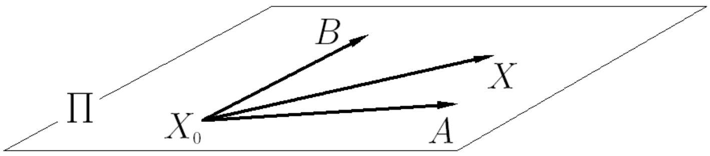
4.3 Zastosowania wyznacznika 3×3.

Zajmijmy się teraz nie wymienionymi wcześniej zastosowaniami wyznacznika w geometrii analitycznej przestrzeni.

Jak wiemy z Faktu 4.2.7 trzy wektory są współpłaszczyznowe jeśli ich wyznacznik jest równy zero. Dzięki temu w łatwy sposób możemy określić warunek na współpłaszczyznowość czterech punktów.

(1) Współpłaszczyznowość czterech punktów w przestrzeni.
Punkty $A, B, C, D \in R^3$ w przestrzeni leżą na jednej płaszczyźnie wtedy i tylko wtedy gdy wektory $\overrightarrow{AB}, \overrightarrow{AC}, \overrightarrow{AD}$ są współpłaszczyznowe , a to zachodzi wtedy i tylko wtedy gdy $\det(\overrightarrow{AB}, \overrightarrow{AC}, \overrightarrow{AD}) = 0$.

(2) Równanie płaszczyzny równoległej do dwóch wektorów.
Z powyższej własności możemy wyprowadzić równanie płaszczyzny równoległej do dwóch danych wektorów $A = \begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix}$, $B = \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix}$ i przechodzącej przez punkt $X_0 = \begin{pmatrix} x_0 \\ y_0 \\ z_0 \end{pmatrix}$. Oznaczmy szukaną płaszczyznę przez Π. Weźmy dowolny punkt $X = \begin{pmatrix} x \\ y \\ z \end{pmatrix}$ należący do tej płaszczyzny, wtedy wektor $\overrightarrow{XX_0}$ również będzie się zawierać w tej płaszczyźnie, zatem wektory $A, B, \overrightarrow{XX_0}$ będą współpłaszczyznowe.



Jak pokazaliśmy w poprzednim podpunkcie dla takich wektorów zachodzi równość $\det(A, B, \overrightarrow{XX_0}) = 0$. Po podstawieniu współrzędnych otrzymujemy więc $\det \begin{pmatrix} a_1 & b_1 & x - x_0 \\ a_2 & b_2 & y - y_0 \\ a_3 & b_3 & z - z_0 \end{pmatrix} = 0$, co rozwijając względem trzeciej kolumny, możemy zapisać jako
(4.2)

$$\boxed{(x - x_0) \det \begin{pmatrix} a_2 & b_2 \\ a_3 & b_3 \end{pmatrix} - (y - y_0) \det \begin{pmatrix} a_1 & b_1 \\ a_3 & b_3 \end{pmatrix} + (z - z_0) \det \begin{pmatrix} a_1 & b_1 \\ a_2 & b_2 \end{pmatrix} = 0}$$

Zaważmy, że równanie to jest spełnione dla każdego wektora z płaszczyzny Π, jak również każdy wektor, który spełnia to równanie jest współpłaszczyznowy z A, B, należy więc do płaszczyzny Π. Możemy więc przyjąć, że równość (4.2) jest równaniem tej płaszczyzny.

(3) Równanie płaszczyzny wyznaczonej przez trzy punkty.
Mając dane trzy punkty A, B, C możemy znaleźć równanie płaszczyzny wyznaczonej przez te punkty. Dla dowolnego X należącego do tej płaszczyzny wektory $\overrightarrow{AB}, \overrightarrow{AC}, \overrightarrow{AX}$ są współpłaszczyznowe , a zatem ich wyznacznik jest równy zero $\det(\overrightarrow{AB}, \overrightarrow{AC}, \overrightarrow{AX}) = 0$, co jednoznacznie wyznacza nam szukaną płaszczyznę.

Poniższy przykład ilustruje jak wygląda tak uzyskiwane równanie płaszczyzny dla punktów A, B, C o konkretnych zadanych współrzędnych.

✔ **Przykład.** Wyznaczmy równanie płaszczyzny zawierającej punkty

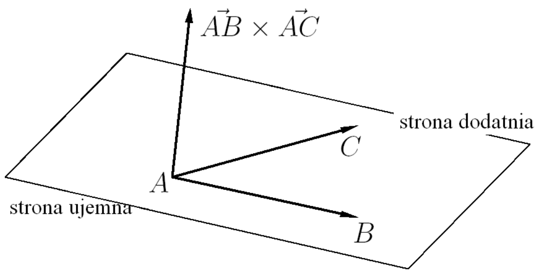
$$A = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}, \ B = \begin{pmatrix} 0 \\ 2 \\ 4 \end{pmatrix}, \ C = \begin{pmatrix} 1 \\ 5 \\ 1 \end{pmatrix}.$$

Dla $X = \begin{pmatrix} x \\ y \\ z \end{pmatrix}$ mamy $\overrightarrow{AB} = \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix}$, $\overrightarrow{AC} = \begin{pmatrix} 0 \\ 3 \\ -2 \end{pmatrix}$, $\overrightarrow{AX} = \begin{pmatrix} x-1 \\ y-2 \\ z-3 \end{pmatrix}$.
Dla takich wektorów otrzymujemy

$$\det \begin{pmatrix} -1 & 0 & x-1 \\ 0 & 3 & y-2 \\ 1 & -2 & z-3 \end{pmatrix} = 0.$$

Mamy więc $-3 \cdot (x-1) - 2 \cdot (y-2) - 3(z-3) = 0$, czyli równanie szukanej płaszczyzny to $-3x - 2y - 3z + 16 = 0$.

(4) Dodatnia i ujemna strona płaszczyzny.
Dla uporządkowanej trójki punktów A, B, C możemy określić strony płaszczyzny przez nie wyznaczonej. Dodatnią stroną płaszczyzny wyznaczonej przez A, B, C nazywamy tę część przestrzeni, w której znajduje się koniec wektora $\overrightarrow{AB} \times \overrightarrow{AC}$. Drugą część tej przestrzeni nazywamy stroną ujemną.



Dla dowolnego punktu $X \in R^3$ możemy łatwo sprawdzić po której stronie płaszczyzny znajduje się ten punkt, mamy bowiem

(a) X leży po dodatniej stronie płaszczyzny wyznaczonej przez A, B, C wtedy i tylko wtedy gdy $\det(\overrightarrow{AB}, \overrightarrow{AC}, \overrightarrow{AX}) > 0$

(b) X leży po ujemnej stronie płaszczyzny wyznaczonej przez A, B, C wtedy i tylko wtedy gdy $\det(\overrightarrow{AB}, \overrightarrow{AC}, \overrightarrow{AX}) < 0$

Dowód. Zauważmy, że X leży po dodatniej stronie wtedy i tylko wtedy gdy trójka wektorów $\overrightarrow{AB}, \overrightarrow{AC}, \overrightarrow{AX}$ jest dodatnio zorientowana. Podobnie X leży po ujemnej stronie wtedy i tylko wtedy gdy ta trójka wektorów jest ujemnie zorientowana. Powyższa własność jest zatem konsekwencją Uwagi 4.2.13. □

4.4 Zastosowanie wyznacznika do rozwiązywania układu 3 równań liniowych z 3 niewiadomymi.

Kolejnym bardzo ważnym zastosowaniem wyznacznika jest użycie go do rozwiązywania układów trzech równań liniowych z trzema niewiadomymi. Układ

$$\begin{cases} a_1x + b_1y + c_1z = d_1 \\ a_2x + b_2y + c_2z = d_2 \\ a_3x + b_3y + c_3z = d_3 \end{cases} \tag{4.3}$$

gdzie a_i, b_i, c_i, d_i są dane, natomiast x, y, z są niewiadomymi, możemy rozwiązać metodą „przez podstawianie", bywa ona jednak często bardzo czasochłonna. Korzystając z pojęcia wyznacznika można wyprowadzić ogólne wzory na rozwiązanie takiego układu. W tym celu posłużymy się wektorową interpretacją danego układu. Możemy go bowiem zapisać jako

$$x \cdot \begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix} + y \cdot \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix} + z \cdot \begin{pmatrix} c_1 \\ c_2 \\ c_3 \end{pmatrix} = \begin{pmatrix} d_1 \\ d_2 \\ d_3 \end{pmatrix},$$

lub krócej

$$x \cdot A + y \cdot B + z \cdot C = D.$$

Rozwiązanie naszego układu sprowadza się zatem do znalezienia współczynników rozkładu wektora D w kombinację liniową wektorów A, B, C. Aby wyznaczyć współczynnik x, pomnóżmy skalarnie obie strony otrzymanego równania przez $B \times C$. Otrzymamy wówczas

$$(x \cdot A + y \cdot B + z \cdot C) \circ (B \times C) = D \circ (B \times C),$$

co, korzystając z własności wyznacznika, możemy zapisać jako

$$\det(x \cdot A + y \cdot B + z \cdot C, B, C) = \det(D, B, C),$$

czyli $x \cdot \det(A, B, C) + y \cdot \det(B, B, C) + z \cdot \det(C, B, C) = \det(D, B, C)$. Zauważmy, że $\det(B, B, C) = (B \times B) \circ C = 0$. Podobnie $\det(C, B, C) = C \circ (B \times C) = 0$, bo $B \times C \perp C$. Otrzymujemy zatem

$$x \cdot \det(A, B, C) = \det(D, B, C).$$

W analogiczny sposób możemy wyprowadzić zależności

$$y \cdot \det(A, B, C) = \det(A, D, C)$$

$$z \cdot \det(A, B, C) = \det(A, B, D).$$

Przy założeniu, że $\det(A, B, C) \neq 0$ rozwiązaniem układu (4.3) jest

$$\begin{cases} x = \frac{\det(D,B,C)}{\det(A,B,C)} \\ y = \frac{\det(A,D,C)}{\det(A,B,C)} \\ z = \frac{\det(A,B,D)}{\det(A,B,C)} \end{cases}.$$

✍ **Ćwiczenie.** Postaraj się wyprowadzić zależności $y \cdot \det(A, B, C) = \det(A, D, C)$, $z \cdot \det(A, B, C) = \det(A, B, D)$ stosując rozumowanie podobne do tego, które pozwoliło uzyskać zależność $x \cdot \det(A, B, C) = \det(D, B, C)$.

Otrzymane rozwiązanie możemy zapisać w ogólniejszej postaci. W tym celu zdefiniujmy pewne użyteczne pojęcia.

Definicja 4.4.1. Macierz $(A, B, C) = \begin{pmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{pmatrix}$ nazywamy **macierzą główną** układu (4.3).
Macierze pomocnicze powstają przez zastąpienie kolumn współczynników przy pewnej niewiadomej, kolumną wyrazów wolnych. Dla układu (4.3) mamy więc trzy macierze pomocnicze: (D, B, C), (A, D, C), (A, B, D).
Wyznacznikiem głównym danego układu nazywamy wyznacznik macierzy głównej, czyli $W = \det(A, B, C)$.
Wyznacznikami pomocniczymi nazywamy wyznaczniki $W_x = \det(D, B, C)$, $W_y = \det(A, D, C)$, $W_z = \det(A, B, D)$.

Zauważmy, że W_x to wyznacznik macierzy powstałej przez zastąpienie w macierzy głównej kolumny odpowiadającej zmiennej x przez kolumnę wyrazów wolnych. Wyznacznik ten jest wykorzystywany do wyznaczenia zmiennej x. Macierz we wzorze na W_y powstaje przez zastąpienie kolumny odpowiadającej zmiennej y przez kolumnę wyrazów wolnych, natomiast sam wyznacznik wykorzystujemy do wyznaczenia zmiennej y. Do znalezienia z używamy wyznacznika W_z, który powstaje w analogiczny sposób.
Możemy teraz sformułować ogólne twierdzenie.

Twierdzenie 4.4.2 (rozwiązanie układu równań). *Jeśli $W \neq 0$ to układ (4.3) ma dokładnie jedno rozwiązanie, które przedstawia się następująco*

$$x = \frac{W_x}{W}, \quad y = \frac{W_y}{W}, \quad z = \frac{W_z}{W}.$$

Uwaga 4.4.3. *Twierdzenie powyższe ma analogiczną postać, jak dla układu dwóch równań liniowych z dwiema niewiadomymi.*

✔ **Przykład.** Rozwiążmy układ

$$\begin{cases} 3x + 4y + z = 1 \\ 2x - 3y + 2z = 5 \\ x - y = 3 \end{cases}$$

Macierzą główną tego układu jest macierz $\begin{pmatrix} 3 & 4 & 1 \\ 2 & -3 & 2 \\ 1 & -1 & 0 \end{pmatrix}$. Wyznacznik tej macierzy wynosi $W = 0 - 2 + 8 + 3 + 6 - 0 = 15$. Wyznaczniki pomocnicze to

$$W_x = \det \begin{pmatrix} 1 & 4 & 1 \\ 5 & -3 & 2 \\ 3 & -1 & 0 \end{pmatrix} = 0 - 5 + 24 + 9 + 2 - 0 = 30,$$

$$W_y = \det \begin{pmatrix} 3 & 1 & 1 \\ 2 & 5 & 2 \\ 1 & 3 & 0 \end{pmatrix} = 0 + 6 + 2 - 5 - 18 - 0 = -15,$$

$$W_z = \det \begin{pmatrix} 3 & 4 & 1 \\ 2 & -3 & 5 \\ 1 & -1 & 3 \end{pmatrix} = -27 - 2 + 20 + 3 + 15 - 24 = -15.$$

Rozwiązaniem naszego układu jest więc trójka liczb

$$x = \frac{30}{15} = 2, \; y = \frac{-15}{15} = -1, \; z = \frac{-15}{15} = -1.$$

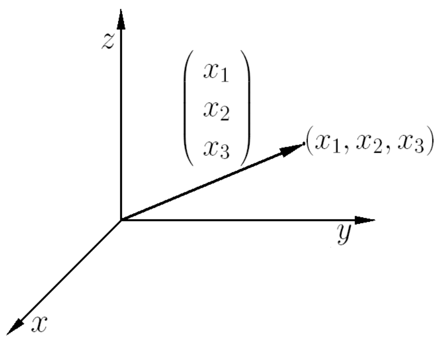
Rozdział 5

Przekształcenia liniowe przestrzeni

Rozdział ten poświęcimy na zapoznanie się z przekształceniami przestrzeni. W szczególności zajmiemy się przekształceniami liniowymi. W pierwszym podrozdziale wprowadzimy podstawowe definicje, w kolejnych poznamy różne, geometryczne przykłady przekształceń liniowych.

5.1 Podstawowe definicje.

Uwaga 5.1.1. *W algebrze liniowej punkt o współrzędnych* (x_1, x_2, x_3) *często utożsamiany jest ze swoim wektorem wodzącym, czyli wektorem o współrzędnych* $\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$.



Definicja 5.1.2 (**przekształcenie przestrzeni**)**.** Przekształcenie T przestrzeni jest to pewna reguła, która każdemu wektorowi X (punktowi) przestrzeni, przyporządkowuje jednoznacznie pewien wektor (punkt), oznaczony $T(X)$ i zwany obrazem wektora X przez przekształcenie T.

Uwaga 5.1.3. *Często będziemy stosować zapis* $T\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} x_1' \\ x_2' \\ x_3' \end{pmatrix}$, *który oznacza, że obrazem wektora* $\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$ *przez przekształcenie* T *jest wektor* $\begin{pmatrix} x_1' \\ x_2' \\ x_3' \end{pmatrix}$.

Najważniejszymi przekształceniami przestrzeni są tzw. przekształcenia liniowe i tylko takimi będziemy zajmować się w dalszej części tego rozdziału.

Definicja 5.1.4 (przekształcenie liniowe). Przekształcenie $T : R^3 \to R^3$ jest przekształceniem liniowym, jeśli można je opisać za pomocą wzoru

$$\boxed{T\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} ax_1 + bx_2 + cx_3 \\ dx_1 + ex_2 + fx_3 \\ gx_1 + hx_2 + ix_3 \end{pmatrix}} \tag{5.1}$$

dla dowolnego wektora $\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \in R^3$, gdzie $a, b, c, d, e, f, g, h, i$ oznaczają stałe liczby.

Każdemu przekształceniu liniowemu możemy jednoznacznie przyporządkować pewną macierz, zwaną macierzą tego przekształcenia. W wielu przypadkach bardzo ułatwi to opis, jak i badanie własności danego przekształcenia.

Uwaga 5.1.5 (macierz przekształcenia liniowego). *Macierz*

$$\boxed{\begin{pmatrix} a & b & c \\ d & e & f \\ g & h & i \end{pmatrix}} \tag{5.2}$$

nazywamy macierzą przekształcenia T, zadanego wzorem (5.1).

✔ **Przykład.** Przekształcenie T, które dowolnemu wektorowi $\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$ przyporządkowuje wektor $\begin{pmatrix} x_1 - 3x_2 + x_3 \\ 2x_1 + x_3 \\ x_2 \end{pmatrix}$ jest przekształceniem liniowym, bo

$$T\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} x_1 - 3x_2 + x_3 \\ 2x_1 + x_3 \\ x_2 \end{pmatrix} = \begin{pmatrix} 1 \cdot x_1 + (-3) \cdot x_2 + 1 \cdot x_3 \\ 2 \cdot x_1 + 0 \cdot x_2 + 1 \cdot x_3 \\ 0 \cdot x_1 + 1 \cdot x_2 + 0 \cdot x_3 \end{pmatrix},$$

czyli da się przedstawić w postaci (5.1) dla

$$a = 1, b = -3, c = 1, d = 2, e = 0, f = 1, g = 0, h = 1, i = 0.$$

Macierzą tego przekształcenia jest

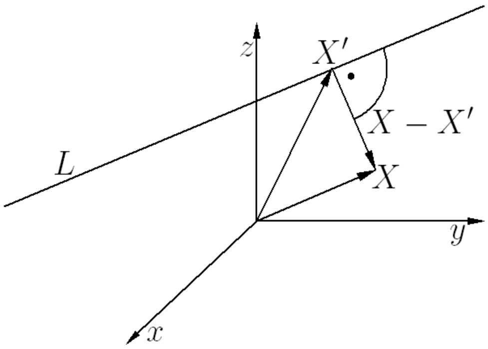
$$\begin{pmatrix} 1 & -3 & 1 \\ 2 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}.$$

W kolejnych podrozdziałach zapoznamy się z rozmaitymi przykładami przekształceń, które okażą się być przekształceniami liniowymi.

5.2 Rzut prostopadły na prostą wzdłuż wektora U.

Na początek zdefiniujmy czym, dla dowolnego punktu X i prostej L jest rzut prostopadły tego punktu na tę prostą.

Definicja 5.2.1 (rzut prostopadły na prostą). Rzutem prostopadłym punktu X na prostą L, nazywamy taki punkt X', o końcu na prostej L, że różnica $X - X'$ jest wektorem prostopadłym do L.



➢ **Rzut prostopadły na prostą wzdłuż wektora** $U = \begin{pmatrix} -3 \\ 5 \\ 1 \end{pmatrix}$

Rozważmy teraz zagadnienie rzutu prostopadłego wektora dla konkretnej prostej L, mianowicie dla prostej wzdłuż wektora U, czyli prostej przechodzącej przez początek układu, o wektorze kierunkowym U. Niech $U = \begin{pmatrix} -3 \\ 5 \\ 1 \end{pmatrix}$. Prostą wzdłuż wektora U można wówczas zapisać jako $t \cdot U = t \cdot \begin{pmatrix} -3 \\ 5 \\ 1 \end{pmatrix} = \begin{pmatrix} -3t \\ 5t \\ t \end{pmatrix}$. Dla takiej prostej L spróbujmy znaleźć rzut prostopadły wektora X na L. Zadanie to sprowadza się do znalezienia odpowiedniej wartości parametru t, tak aby $X' = t \cdot U$, będący rzutem wektora X na prostą wzdłuż U, spełniał warunki określone w Definicji 5.2.1, czyli aby $X - X' = X - tU = \begin{pmatrix} x_1 + 3t \\ x_2 - 5t \\ x_3 - t \end{pmatrix}$ był prostopadły do tej prostej, a tym samym był prostopadły do wektora U. Musi więc zachodzić

$$(X - X') \circ U = 0.$$

Ten warunek możemy zapisać jako

$$\begin{pmatrix} x_1 + 3t \\ x_2 - 5t \\ x_3 - t \end{pmatrix} \circ \begin{pmatrix} -3 \\ 5 \\ 1 \end{pmatrix} = -3x_1 - 9t + 5x_2 - 25t + x_3 - t = 0.$$

Współrzędne x_1, x_2, x_3 traktujemy tutaj jako dane, możemy więc wyznaczyć t

$$(-9-25-1)\cdot t - 3x_1 + 5x_2 + x_3 = 0$$

$$-3x_1 + 5x_2 + x_3 = 35t$$
$$t = \frac{-3x_1 + 5x_2 + x_3}{35}$$

Zatem wektor X' możemy zapisać jako

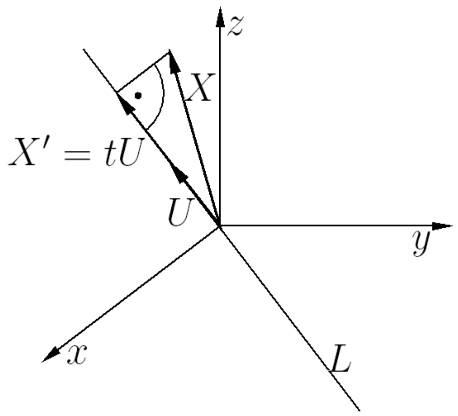
$$X' = tU = \frac{-3x_1 + 5x_2 + x_3}{35}\begin{pmatrix} -3 \\ 5 \\ 1 \end{pmatrix} = \begin{pmatrix} \frac{9}{35}x_1 - \frac{15}{35}x_2 - \frac{3}{35}x_3 \\ -\frac{15}{35} + \frac{25}{35}x_2 + \frac{5}{35}x_3 \\ -\frac{3}{35}x_1 + \frac{5}{35}x_2 + \frac{1}{35}x_3 \end{pmatrix}$$

Jest to wzór postaci zgodnej z definicją przekształcenia liniowego, rzut ten jest więc przekształceniem liniowym, którego macierzą jest

$$\begin{pmatrix} \frac{9}{35} & -\frac{15}{35} & -\frac{3}{35} \\ -\frac{15}{35} & \frac{25}{35} & \frac{5}{35} \\ -\frac{3}{35} & \frac{5}{35} & \frac{1}{35} \end{pmatrix}.$$

➢ Rzut prostopadły na prostą wzdłuż wektora U - przypadek ogólny

Zastanówmy się jednak, jak wygląda wzór takiego przekształcenia, dla prostej L poprowadzonej wzdłuż dowolnego wektora $U \in R^3$ o zadanych współrzędnych $U = \begin{pmatrix} u_1 \\ u_2 \\ u_3 \end{pmatrix}$. Sprawdźmy, czy zawsze będzie to przekształcenie liniowe. Dla przypadku ogólnego, spróbujmy znaleźć takie t, aby $X' = tU$ oraz $(X - X') \circ U = 0$.



Otrzymujemy więc $(X - tU) \circ U = 0$

$$X \circ U - t \cdot U \circ U = 0$$

$$X \circ U = t \cdot U \circ U = t|U|^2$$

a stąd możemy już wyznaczyć t

$$t = \frac{X \circ U}{|U|^2}.$$

Ogólny wzór na rzut to zatem

$$X' = tU = \frac{X \circ U}{|U|^2} \cdot U. \tag{5.3}$$

We współrzędnych, dla $X \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$, $U = \begin{pmatrix} u_1 \\ u_2 \\ u_3 \end{pmatrix}$ wzór ten przyjmuje postać

$$X' = T\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \frac{\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \circ \begin{pmatrix} u_1 \\ u_2 \\ u_3 \end{pmatrix}}{\left|\begin{pmatrix} u_1 \\ u_2 \\ u_3 \end{pmatrix}\right|^2} \cdot \begin{pmatrix} u_1 \\ u_2 \\ u_3 \end{pmatrix} = \frac{x_1u_1 + x_2u_2 + x_3u_3}{u_1^2 + u_2^2 + u_3^2}\begin{pmatrix} u_1 \\ u_2 \\ u_3 \end{pmatrix} =$$

$$\begin{pmatrix} \frac{x_1u_1^2 + x_2u_1u_2 + x_3u_1u_3}{u_1^2+u_2^2+u_3^2} \\ \frac{x_1u_1u_2 + x_2u_2^2 + x_3u_2u_3}{u_1^2+u_2^2+u_3^2} \\ \frac{x_1u_1u_3 + x_2u_2u_3 + x_3u_3^2}{u_1^2+u_2^2+u_3^2} \end{pmatrix} = \begin{pmatrix} \frac{u_1^2}{u_1^2+u_2^2+u_3^2}x_1 + \frac{u_1u_2}{u_1^2+u_2^2+u_3^2}x_2 + \frac{u_1u_3}{u_1^2+u_2^2+u_3^2}x_3 \\ \frac{u_1u_2}{u_1^2+u_2^2+u_3^2}x_1 + \frac{u_2^2}{u_1^2+u_2^2+u_3^2}x_2 + \frac{u_2u_3}{u_1^2+u_2^2+u_3^2}x_3 \\ \frac{u_1u_3}{u_1^2+u_2^2+u_3^2}x_1 + \frac{u_2u_3}{u_1^2+u_2^2+u_3^2}x_2 + \frac{u_3^2}{u_1^2+u_2^2+u_3^2}x_3 \end{pmatrix}$$

Zatem zgodnie z Definicją 5.1 jest to przekształcenie liniowe. Macierzą tego przekształcenia jest

$$\begin{pmatrix} \frac{u_1^2}{u_1^2+u_2^2+u_3^2} & \frac{u_1u_2}{u_1^2+u_2^2+u_3^2} & \frac{u_1u_3}{u_1^2+u_2^2+u_3^2} \\ \frac{u_1u_2}{u_1^2+u_2^2+u_3^2} & \frac{u_2^2}{u_1^2+u_2^2+u_3^2} & \frac{u_2u_3}{u_1^2+u_2^2+u_3^2} \\ \frac{u_1u_3}{u_1^2+u_2^2+u_3^2} & \frac{u_2u_3}{u_1^2+u_2^2+u_3^2} & \frac{u_3^2}{u_1^2+u_2^2+u_3^2} \end{pmatrix}.$$

Macierz tę możemy też zapisać w postaci

$$\frac{1}{u_1^2 + u_2^2 + u_3^2} \cdot \begin{pmatrix} u_1^2 & u_1u_2 & u_1u_3 \\ u_1u_2 & u_2^2 & u_2u_3 \\ u_1u_3 & u_2u_3 & u_3^2 \end{pmatrix}.$$

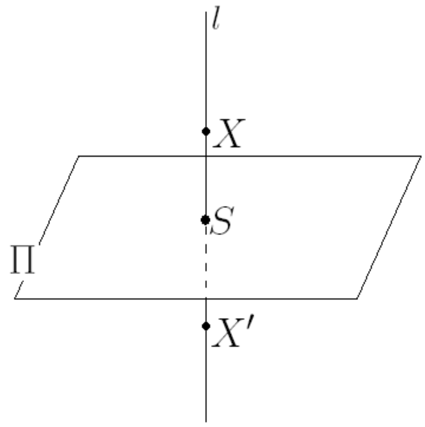
Wniosek 5.2.2. *Rzut prostopadły na prostą wzdłuż dowolnego wektora $U \in R^3$ jest przekształceniem liniowym.*

5.3 Symetria względem płaszczyzny.

Innym przykładem przekształcenia liniowego jest odbicie (symetria) względem płaszczyzny przechodzącej przez początek układu współrzędnych. Rozważmy to przekształcenie na przykładzie odbicia względem płaszczyzny zadanej równaniem $3x - 5y - 2z = 0$. Skorzystamy z następującego opisu warunków, jakie spełnia punkt X' będący obrazem punktu X przez to przekształcenie.

Fakt 5.3.1 (obraz punktu w symetrii względem płaszczyzny). *Obrazem punktu X w odbiciu względem płaszczyzny Π jest punkt X' spełniający następujące dwa warunki*

1. *X' leży na prostej l prostopadłej do Π przechodzącej przez X.*
2. *Środek S odcinka XX' należy do płaszczyzny Π.*



Posługując się powyższymi dwoma warunkami znajdźmy obraz $X' = \begin{pmatrix} x_1' \\ x_2' \\ x_3' \end{pmatrix}$ punktu $X = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$ w symetrii względem płaszczyzny $\Pi : 3x-5y-2z=0$. Wektor normalny płaszczyzny Π, to $N = \begin{pmatrix} 3 \\ -5 \\ -2 \end{pmatrix}$. Prostą l, tzn. prostą prostopadłą do Π przechodzącą przez X, możemy więc zapisać w postaci $l : \ X + t \cdot N$, co po podstawieniu współrzędnych daje

$$l : \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} + t \begin{pmatrix} 3 \\ -5 \\ -2 \end{pmatrix} = \begin{pmatrix} x_1 + 3t \\ x_2 - 5t \\ x_3 - 2t \end{pmatrix}.$$

Punkt X' należy do prostej l, więc dla pewnej wartości t mamy $X' = \begin{pmatrix} x_1 + 3t \\ x_2 - 5t \\ x_3 - 2t \end{pmatrix}$.

Wiemy również, że punkt S, jest środkiem odcinka XX', czyli

$$S = \frac{1}{2}(X + X') = \begin{pmatrix} \frac{1}{2}x_1 + \frac{1}{2}x_1' \\ \frac{1}{2}x_2 + \frac{1}{2}x_2' \\ \frac{1}{2}x_3 + \frac{1}{2}x_3' \end{pmatrix} = \begin{pmatrix} \frac{1}{2}x_1 + \frac{1}{2}x_1 + \frac{3}{2}t \\ \frac{1}{2}x_2 + \frac{1}{2}x_2 - \frac{5}{2}t \\ \frac{1}{2}x_3 + \frac{1}{2}x_3 - t \end{pmatrix} = \begin{pmatrix} x_1 + \frac{3}{2}t \\ x_2 - \frac{5}{2}t \\ x_3 - t \end{pmatrix}.$$

Punkt ten należy do płaszczyzny Π zatem jego współrzędne spełniają warunek

$$3(x_1 + \frac{3}{2}t) - 5(x_2 - \frac{5}{2}t) - 2(x_3 - t) = 0.$$

Przekształćmy tę równość traktując współrzędne x_1, x_2, x_3 jako dane, zaś wartość parametru t jako niewiadomą, którą wyliczamy

$$3x_1 - 5x_2 - 2x_3 + 19t = 0,$$

$$19t = -3x_1 + 5x_2 + 2x_3,$$
$$t = -\frac{3}{19}x_1 + \frac{5}{19}x_2 + \frac{2}{19}x_3.$$

Stąd po wstawieniu wyliczonej wartości dla parametru t otrzymujemy

$$X' = \begin{pmatrix} x_1 + 3t \\ x_2 - 5t \\ x_3 - 2t \end{pmatrix} = \begin{pmatrix} x_1 - \frac{9}{19}x_1 + \frac{15}{19}x_2 + \frac{6}{19}x_3 \\ x_2 + \frac{15}{19}x_1 - \frac{25}{19}x_2 - \frac{10}{19}x_3 \\ x_3 + \frac{6}{19}x_1 - \frac{10}{19}x_2 - \frac{4}{19}x_3 \end{pmatrix} = \begin{pmatrix} -\frac{10}{19}x_1 + \frac{15}{19}x_2 + \frac{6}{19}x_3 \\ \frac{15}{19}x_1 - \frac{6}{19}x_2 - \frac{10}{19}x_3 \\ \frac{6}{19}x_1 - \frac{10}{19}x_2 + \frac{15}{19}x_3 \end{pmatrix}.$$

Symetria względem płaszczyzny $3x - 5y - 2z = 0$ jest więc przekształceniem liniowym. Macierz tego przekształcenia to

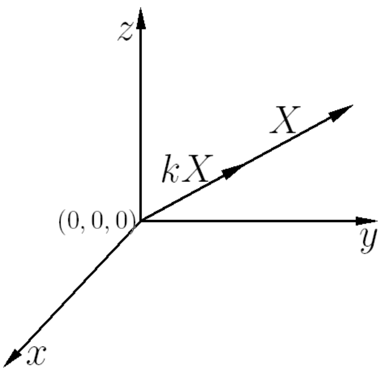
$$\begin{pmatrix} -\frac{10}{19} & \frac{15}{19} & \frac{6}{19} \\ \frac{15}{19} & -\frac{6}{19} & -\frac{10}{19} \\ \frac{6}{19} & -\frac{10}{19} & \frac{15}{19} \end{pmatrix}.$$

5.4 Jednokładność J_k o skali k względem $(0,0,0)$.

Jednokładność J_k o skali k względem $(0,0,0)$, jest również przekształceniem liniowym. Obrazem wektora $X = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$, w takiej jednokładności, jest wektor leżący na prostej wyznaczonej przez wektor X, ale pomnożony przez k, czyli

$$J_k(X) = k \cdot X.$$

Poniżej przedstawiamy przykład takiego przekształcenia dla $k = \frac{1}{2}$.



We współrzędnych jednokładność o skali k możemy przedstawić jako

$$J_k \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = k \cdot \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} kx_1 \\ kx_2 \\ kx_3 \end{pmatrix} = \begin{pmatrix} k \cdot x_1 + 0 \cdot x_2 + 0 \cdot x_3 \\ 0 \cdot x_1 + k \cdot x_2 + 0 \cdot x_3 \\ 0 \cdot x_1 + 0 \cdot x_2 + k \cdot x_3 \end{pmatrix}.$$

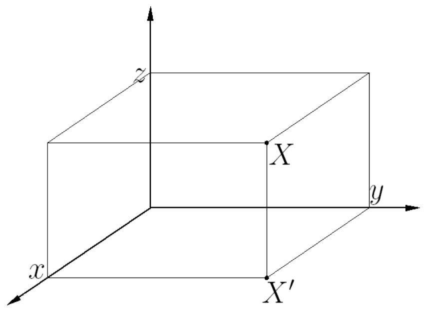
Przekształcenie J_k jest więc rzeczywiście przekształceniem liniowym, którego macierzą jest

$$\begin{pmatrix} k & 0 & 0 \\ 0 & k & 0 \\ 0 & 0 & k \end{pmatrix}.$$

5.5 Rzuty i symetrie związane z osiami i płaszczyznami współrzędnych.

➢ **Rzut na płaszczyznę**

Zgodnie z oznaczeniami, wprowadzonymi w Rozdziale 1, płaszczyzny współrzędnych to O_{xy}, O_{xz}, O_{yz}. Rozważmy rzut prostokątny na płaszczyznę O_{xy}, oznaczmy go P_{xy}.



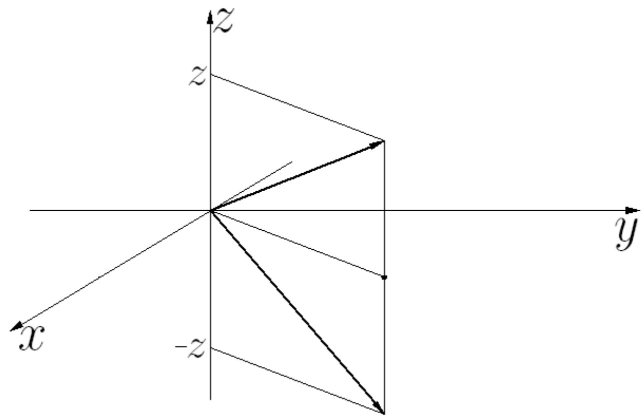
Zauważmy, że obrazem punktu $X = \begin{pmatrix} x \\ y \\ z \end{pmatrix}$ przez takie przekształcenie będzie punkt $X' = \begin{pmatrix} x \\ y \\ 0 \end{pmatrix}$, a więc przekształcenie to możemy opisać wzorem $P_{xy}\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} x \\ y \\ 0 \end{pmatrix}$.
Z postaci tego wzoru możemy wnioskować, że P_{xy} jest przekształceniem liniowym, którego macierzą jest

$$m(P_{xy}) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

W analogiczny sposób otrzymujemy $m(P_{yz}) = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$ oraz $m(P_{xz}) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}$, gdzie P_{yz}, P_{xz} oznaczają rzuty na płaszczyzny O_{yz}, O_{xz}, odpowiednio.

➢ **Symetria względem płaszczyzny**

Rozważmy teraz symetrie względem tych płaszczyzn. Niech S_{xy} oznacza symetrię względem O_{xy}. Zauważmy, że koniec wektora X' będącego obrazem wektora X przez to przekształcenie znajduje się na prostej prostopadłej do płaszczyzny O_{xy} i przechodzącej przez koniec wektora X. Zgodnie z określeniem współrzędnych wektora, podanym w pierwszym rozdziale, oznacza to, że wektory X i X' mają na płaszczyźnie O_{xy} takie same współrzędne, co oznacza, że dwie pierwsze współrzędne tych wektorów są takie same. Ponadto odległość wektora X' od płaszczyzny O_{xy} jest taka sama jak odległość wektora X od tej płaszczyzny i znajdują się one po przeciwnych stronach tej płaszczyzny, tak więc trzecia współrzędna wektora X' będzie miała tę samą wartość ale przeciwny znak co trzecia współrzędna wektora X. Obrazem wektora $X = \begin{pmatrix} x \\ y \\ z \end{pmatrix}$ w takiej symetrii będzie zatem punkt $X' = S_{xy}\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} x \\ y \\ -z \end{pmatrix}$.



Z postaci wzoru wyznaczającego obraz wektora przez to przekształcenie możemy wnioskować, że symetria względem płaszczyzny jest przekształceniem liniowym o macierzy

$$m(S_{xy}) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix}.$$

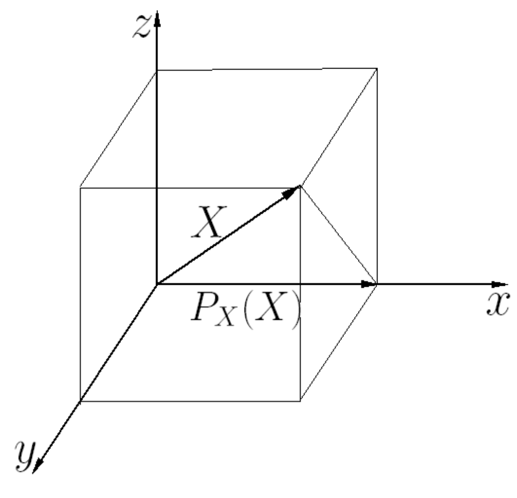
Podobnie otrzymujemy

$$m(S_{yz}) = \begin{pmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad m(S_{xz}) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{pmatrix},$$

gdzie S_{yz}, S_{xz} są symetriami względem płaszczyzn O_{yz}, O_{xz}, odpowiednio.

➢ Rzuty na osie

Zajmijmy się teraz rzutami na osie. Niech $P_x$oznacza rzut na oś O_x.



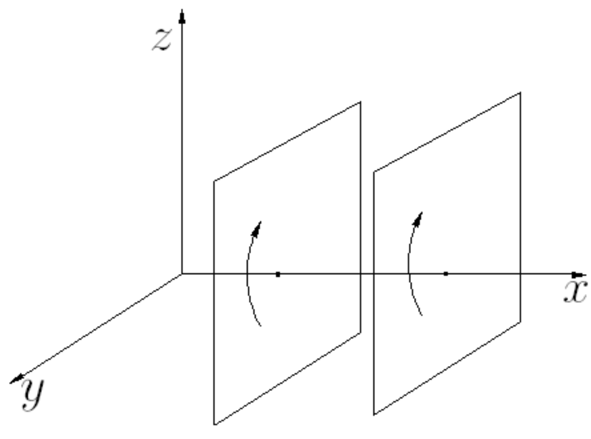
Jak nietrudno się przekonać $P_x\begin{pmatrix} x \\ y \\ z\end{pmatrix} = \begin{pmatrix} x \\ 0 \\ 0\end{pmatrix}$. Rzut na oś jest zatem przekształceniem liniowym, którego macierz to $m(P_x) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0\end{pmatrix}$. Podobnie otrzymujemy

$P_y\begin{pmatrix} x \\ y \\ z\end{pmatrix} = \begin{pmatrix} 0 \\ y \\ 0\end{pmatrix}$, czyli $m(P_y) = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0\end{pmatrix}$,

$P_z\begin{pmatrix} x \\ y \\ z\end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ z\end{pmatrix}$, czyli $m(P_z) = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1\end{pmatrix}$.

5.6 Obroty wokół osi.

Niech $R_{\theta,x}$ oznacza obrót wokół osi O_x o kąt θ. Zauważmy, że dla dowolnego wektora $X = \begin{pmatrix} x \\ y \\ z\end{pmatrix}$ pierwsza współrzędna, określona na osi x, pozostanie pod wpływem obrotu $R_{\theta,x}$ bez zmian. Wynika to stąd, że obrót ten przekształca płaszczyzny prostopadłe do osi O_x na siebie, a płaszczyzny te składają się z punktów o jednakowej współrzędnej x (patrz rysunek).



Na każdej płaszczyźnie prostopadłej do osi O_x ten przestrzenny obrót $R_{\theta,x}$ działa jak obrót o kąt θ tej płaszczyzny. W związku z tym, bazując na macierzach obrotów płaszczyzny, wprowadzonych w pierwszej części skryptu, możemy stwierdzić, że obrazem wektora X jest wektor

$$R_{\theta,X}\begin{pmatrix} x \\ y \\ z\end{pmatrix} = \begin{pmatrix} x \\ \cos\theta\cdot y + \sin\theta\cdot z \\ -\sin\theta\cdot y + \cos\theta\cdot z\end{pmatrix}.$$

Obrót jest więc przekształceniem liniowym o macierzy $\begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & \sin\theta \\ 0 & -\sin\theta & \cos\theta\end{pmatrix}$.

Analogiczną postać przyjmują macierze obrotów wokół osi O_y, O_z. Nie będziemy ich tutaj wprowadzać.

5.7 Przekształcenie tożsamościowe.

Przekształceniem tożsamościowym nazywamy przekształcenie I, które dla każdego $X \in R^3$ spełnia $I(X) = X$, czyli

$$I\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 1 \cdot x_1 + 0 \cdot x_2 + 0 \cdot x_3 \\ 0 \cdot x_1 + 1 \cdot x_2 + 0 \cdot x_3 \\ 0 \cdot x_1 + 0 \cdot x_2 + 1 \cdot x_3 \end{pmatrix}.$$

Macierz tego przekształcenia, nazywana często macierzą identycznościową, to $m(I) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$. Przekształcenie tożsamościowe jest więc przekształceniem liniowym. Czasem na oznaczenie przekształcenia tożsamościowego będziemy używać symbolu Id.

5.8 Przekształcenie zerowe.

Przekształcenie zerowe Z to przekształcenie, które dla każdego $X \in R^3$ spełnia $Z(X) = O$, czyli

$$Z\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \cdot x_1 + 0 \cdot x_2 + 0 \cdot x_3 \\ 0 \cdot x_1 + 0 \cdot x_2 + 0 \cdot x_3 \\ 0 \cdot x_1 + 0 \cdot x_2 + 0 \cdot x_3 \end{pmatrix}.$$

Jest to więc przekształcenie liniowe o macierzy $m(Z) = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$.

Rozdział 6

Własności przekształceń liniowych i ich macierzy

W tym rozdziale własnościami przekształceń liniowych, których przykłady poznaliśmy w poprzednim rozdziale. Nauczymy się składać przekształcenia liniowe. Poznamy określenie iloczynu macierzy i jego związek ze składaniem przekształceń liniowych.

6.1 Charakteryzacja przekształceń liniowych.

Poniższy fakt określa najbardziej charakterystyczne własności przekształceń liniowych, dzięki którym, w łatwy sposób, można sprawdzić czy dane przekształcenie jest przekształceniem liniowym.

Fakt 6.1.1 (charakteryzacja przekształceń liniowych). *Przekształcenie przestrzeni $T : R^3 \to R^3$ jest liniowe wtedy i tylko wtedy, gdy T spełnia następujące warunki:*
(1) ***addytywność:*** $T(X+Y) = T(X) + T(Y)$ *dla dowolnych* X, Y,
(2) ***jednorodność:*** $T(rX) = rT(X)$ *dla dowolnego wektora* X *i dowolnej liczby rzeczywistej* r.

Dowód. Pokażemy, że jeśli T jest przekształceniem liniowym, to rzeczywiście spełnione są warunki (1) i (2). Niech $X = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}, Y = \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix}$ oraz $m(T) = \begin{pmatrix} a & b & c \\ d & e & f \\ g & h & i \end{pmatrix}$.
Wtedy

$$T(X+Y) = T\begin{pmatrix} x_1+y_1 \\ x_2+y_2 \\ x_3+y_3 \end{pmatrix} = \begin{pmatrix} a(x_1+y_1)+b(x_1+y_1)+c(x_1+y_1) \\ d(x_2+y_2)+e(x_2+y_2)+f(x_2+y_2) \\ g(x_3+y_3)+h(x_3+y_3)+i(x_3+y_3) \end{pmatrix} =$$

$$= \begin{pmatrix} ax_1+bx_2+cx_3+ay_1+by_2+cy_3 \\ dx_1+ex_2+fx_3+dy_1+ey_2+fy_3 \\ gx_1+hx_2+ix_3+gy_1+hy_2+iy_3 \end{pmatrix} =$$

$$= \begin{pmatrix} ax_1 + bx_2 + cx_3 \\ dx_1 + ex_2 + fx_3 \\ gx_1 + hx_2 + ix_3 \end{pmatrix} + \begin{pmatrix} ay_1 + by_2 + cy_3 \\ dy_1 + ey_2 + fy_3 \\ gy_1 + hy_2 + iy_3 \end{pmatrix} = T\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} + T\begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} =$$
$$= T(X) + T(Y).$$

Podobnie pokazujemy, że $T(rX) = rT(X)$.

$$T(rX) = T\begin{pmatrix} rx_1 \\ rx_2 \\ rx_3 \end{pmatrix} = \begin{pmatrix} arx_1 + brx_2 + crx_3 \\ drx_1 + erx_2 + frx_3 \\ grx_1 + hrx_2 + irx_3 \end{pmatrix} =$$

$$= \begin{pmatrix} r(ax_1 + bx_2 + cx_3) \\ r(dx_1 + ex_2 + fx_3) \\ r(gx_1 + hx_2 + ix_3) \end{pmatrix} = r\begin{pmatrix} ax_1 + bx_2 + cx_3 \\ dx_1 + ex_2 + fx_3 \\ gx_1 + hx_2 + ix_3 \end{pmatrix} = rT\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = rT(X).$$

Zatem rzeczywiście z faktu liniowości przekształcenia T wynika, że T spełnia warunki (1) i (2). Pokażemy teraz implikację przeciwną. Załóżmy, że przekształcenie T spełnia warunki (1) i (2). Pokażemy, że T jest liniowe. Oznaczmy obrazy wersorów przez przekształcenie T

$$T\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix}, \ T\begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix}, \ T\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} c_1 \\ c_2 \\ c_3 \end{pmatrix}.$$

Mamy wówczas $T\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = T\left(\begin{pmatrix} x_1 \\ 0 \\ 0 \end{pmatrix} + \begin{pmatrix} 0 \\ x_2 \\ 0 \end{pmatrix} + \begin{pmatrix} 0 \\ 0 \\ x_3 \end{pmatrix}\right)$. Korzystając z addytywności przekształcenia T, czyli z warunku (1), otrzymujemy

$$T\left(\begin{pmatrix} x_1 \\ 0 \\ 0 \end{pmatrix} + \begin{pmatrix} 0 \\ x_2 \\ 0 \end{pmatrix} + \begin{pmatrix} 0 \\ 0 \\ x_3 \end{pmatrix}\right) = T\begin{pmatrix} x_1 \\ 0 \\ 0 \end{pmatrix} + T\begin{pmatrix} 0 \\ x_2 \\ 0 \end{pmatrix} + T\begin{pmatrix} 0 \\ 0 \\ x_3 \end{pmatrix} =$$

$$= T\left(x_1\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}\right) + T\left(x_2\begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}\right) + T\left(x_3\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}\right)$$

Z jednorodności przekształcenia T, czyli z warunku (2), otrzymujemy

$$T\left(x_1\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}\right) + T\left(x_2\begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}\right) + T\left(x_3\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}\right) =$$

$$= x_1T\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + x_2T\begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} + x_3T\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} =$$

$$= x_1\begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix} + x_2\begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix} + x_3\begin{pmatrix} c_1 \\ c_2 \\ c_3 \end{pmatrix} = \begin{pmatrix} a_1x_1 + b_1x_2 + c_1x_3 \\ a_2x_1 + b_2x_2 + c_2x_3 \\ a_3x_1 + b_3x_2 + c_3x_3 \end{pmatrix}.$$

Otrzymaliśmy więc, że T jest przekształceniem liniowym o macierzy

$$m(T) = \begin{pmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{pmatrix}.$$

□

Powyższe rachunki jako produkt uboczny dają nam następującą cenną obserwację dotyczącą przekształcenia liniowego.

Wniosek 6.1.2 (kolumny macierzy przekształcenia jako obrazy wersorów). *Kolumny macierzy $m(T)$ przekształcenia liniowego $T : R^3 \to R^3$ są wektorami będącymi obrazami wersorów osiowych przez przekształcenie T, czyli*

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

Możemy więc symbolicznie zapisać

$$m(T) = \left(T\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix},\ T\begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix},\ T\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \right).$$

6.2 Działanie macierzy na wektor.

W tym rozdziale wyjaśnimy, jak znajdować obrazy wektorów przez przekształcenie liniowe T posługując się bezpośrednio macierzą $m(T)$ przekształcenia, bez odwoływania się do wzoru przekształcenia. Służyć temu będzie pojęcie działania macierzy na wektor.

Definicja 6.2.1 (działanie macierzy na wektor). Jeśli $m = \begin{pmatrix} a & b & c \\ d & e & f \\ g & h & i \end{pmatrix}$ jest macierzą rozmiaru 3×3, zaś $T : R^3 \to R^3$ jest przekształceniem liniowym zadanym tą macierzą (tzn. że $m(T) = m$), to określmy działanie macierzy m na wektor $\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$ tak, aby wynikiem tego działania był wektor $T\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$, czyli

$$\begin{pmatrix} a & b & c \\ d & e & f \\ g & h & i \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = T\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} ax_1 + bx_2 + cx_3 \\ dx_1 + ex_2 + fx_3 \\ gx_1 + hx_2 + ix_3 \end{pmatrix}.$$

✔ **Przykład.** Wynikiem działania macierzy $\begin{pmatrix} 2 & -1 & 1 \\ 4 & 2 & 5 \\ 2 & 1 & 0 \end{pmatrix}$ na wektor $\begin{pmatrix} 2 \\ 5 \\ -1 \end{pmatrix}$ jest wektor

$$\begin{pmatrix} 2 & -1 & 1 \\ 4 & 2 & 5 \\ 2 & 1 & 0 \end{pmatrix}\begin{pmatrix} 2 \\ 5 \\ -1 \end{pmatrix} = \begin{pmatrix} 2 \cdot 2 - 1 \cdot 5 - 1 \cdot 1 \\ 4 \cdot 2 + 2 \cdot 5 - 5 \cdot 1 \\ 2 \cdot 2 + 1 \cdot 5 - 0 \cdot 1 \end{pmatrix} = \begin{pmatrix} -2 \\ 13 \\ 9 \end{pmatrix}.$$

W łatwiejszym zapamiętaniu reguły dotyczącej działania macierzy na wektor pomocna może być następująca uwaga.

Uwaga 6.2.2 (działanie macierzy na wektor opisane za pomocą iloczynu skalarnego). *Działanie macierzy na wektor można zinterpretować przy użyciu iloczynu skalarnego. Niech* $m = \begin{pmatrix} W_1 \\ W_2 \\ W_3 \end{pmatrix}$ *będzie daną macierzą rozmiaru* 3×3*, gdzie* W_1, W_2, W_3 *oznaczają wiersze tej macierzy. Wiersze te będziemy traktować jak wektory z* R^3*. Wówczas dla dowolnego wektora* $X \in R^3$ *zachodzi*

$$m \cdot X = \begin{pmatrix} W_1 \\ W_2 \\ W_3 \end{pmatrix} \cdot X = \begin{pmatrix} W_1 \circ X \\ W_2 \circ X \\ W_3 \circ X \end{pmatrix}.$$

6.3 Składanie przekształceń liniowych.

Na wektor w przestrzeni możemy zadziałać kolejno różnymi przekształceniami. Operację taką nazywamy składaniem przekształceń i to właśnie tym pojęciem zajmiemy się w tym podrozdziale. Sprawdzimy, czy składając dwa przekształcenia liniowe również otrzymamy przekształcenie liniowe. Nauczymy się wyznaczać macierz złożenia przekształceń liniowych.

Definicja 6.3.1 (złożenie przekształceń liniowych). Złożeniem przekształcenia T z przekształceniem S nazywamy przekształcenie R, które dla dowolnego wektora X jest określone wzorem $R(X) = S(T(X))$. Przekształcenie takie oznaczać będziemy $R = ST$ lub $R = S \circ T$.

✔ **Przykład.** Rozważmy złożenie rzutu P_{xy} na płaszczyznę O_{xy} z rzutem prostopadłym P_{xz} na płaszczyznę O_{xz}. Przypomnijmy, że dla dowolnego wektora $X = \begin{pmatrix} x \\ y \\ z \end{pmatrix}$ mamy $P_{xy}\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} x \\ y \\ 0 \end{pmatrix}$, $P_{xz}\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} x \\ 0 \\ z \end{pmatrix}$. Wtedy złożenie tych przekształceń to przekształcenie $P_{xy} \circ P_{xz}$, które dowolnemu wektorowi $\begin{pmatrix} x \\ y \\ z \end{pmatrix}$ przyporządkowuje wektor

$$P_{xy}P_{xz}\begin{pmatrix} x \\ y \\ z \end{pmatrix} = P_{xy}\left(P_{xz}\begin{pmatrix} x \\ y \\ z \end{pmatrix}\right) = P_{xy}\begin{pmatrix} x \\ 0 \\ z \end{pmatrix} = \begin{pmatrix} x \\ 0 \\ 0 \end{pmatrix}.$$

Zauważmy, że wzór ten odpowiada określonemu w poprzednim rozdziale rzutowi na oś O_x,

$$P_{xy}P_{xz}\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} x \\ 0 \\ 0 \end{pmatrix} = P_x\begin{pmatrix} x \\ y \\ z \end{pmatrix}.$$

Mamy więc

$$P_{xy}P_{xz} = P_x.$$

Wynikiem takiego złożenia jest więc rzut na oś O_x.

W powyższym przykładzie złożenie przekształceń liniowych, okazało się być przekształceniem liniowym. Zastanówmy się, czy tak jest zawsze.

Fakt 6.3.2 (liniowość złożenia przekształceń liniowych). *Złożenie przekształceń liniowych $S, T : R^3 \to R^3$ jest też przekształceniem liniowym.*

Dowód. Korzystając z Faktu 6.1.1 wystarczy sprawdzić, że przekształcenie ST jest addytywne i jednorodne. Jednorodność tego przekształcenia pokażemy korzystając z jednorodności przekształceń S, T. Mamy bowiem

$$ST(vX) = S(T(vX)) = S(vT(X)) = vS(T(X)) = vST(X).$$

Addytywność również wynika z własności przekształceń S, T, a dokładniej z ich addytywności. Zauważmy, że

$$ST(X+Y) = S(T(X+Y)) = S(T(X)+T(Y)) = S(T(X))+S(T(Y)) = ST(X)+ST(Y).$$

Zatem złożenie przekształceń liniowych jest przekształceniem liniowym. □

➢ Macierz złożenia przekształceń liniowych

Wiemy już, że przekształcenie będące złożeniem przekształceń liniowych jest przekształceniem liniowym. Zastanówmy się jednak, jak wygląda macierz takiego przekształcenia, zakładając oczywiście, że znane są macierze $m(S), m(T)$ przekształceń S, T. Jak wiemy z Wniosku 6.1.2 kolumny tej macierz utworzone zostaną przez obrazy wersorów układu, czyli

$$m(ST) = \left(ST\begin{pmatrix}1\\0\\0\end{pmatrix}, ST\begin{pmatrix}0\\1\\0\end{pmatrix}, ST\begin{pmatrix}0\\0\\1\end{pmatrix} \right).$$

Załóżmy, że

$$m(S) = \begin{pmatrix}a_1 & b_1 & c_1\\ a_2 & b_2 & c_2\\ a_3 & b_3 & c_3\end{pmatrix}, \quad m(T) = \begin{pmatrix}a'_1 & b'_1 & c'_1\\ a'_2 & b'_2 & c'_2\\ a'_3 & b'_3 & c'_3\end{pmatrix}.$$

Znajdźmy obraz wektora $\begin{pmatrix}1\\0\\0\end{pmatrix}$ przez przekształcenie ST. Otrzymujemy

$$ST\begin{pmatrix}1\\0\\0\end{pmatrix} = S\left(T\begin{pmatrix}1\\0\\0\end{pmatrix}\right) = S\begin{pmatrix}a'_1\\a'_2\\a'_3\end{pmatrix} = \begin{pmatrix}a_1a'_1 + b_1a'_2 + c_1a'_3\\ a_2a'_1 + b_2a'_2 + c_2a'_3\\ a_3a'_1 + b_3a'_2 + c_3a'_3\end{pmatrix}.$$

Podobnie możemy wyznaczyć drugą i trzecią kolumnę macierzy $m(ST)$. Dla ułatwienia wprowadźmy oznaczenia

(1) niech w_1, w_2, w_3 będą wierszami macierzy $m(S)$, czyli np. $w_1 = (a_1, b_1, c_1)$,

(2) niech k_1, k_2, k_3 będą kolumnami macierzy $m(T)$, czyli np. $k_1 = \begin{pmatrix} a'_1 \\ a'_2 \\ a'_3 \end{pmatrix}$.

Stosując te oznaczenia możemy pierwszą kolumnę macierzy $m(ST)$ zapisać jako $ST\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} w_1 \circ k_1 \\ w_2 \circ k_1 \\ w_3 \circ k_1 \end{pmatrix}$. Podobnie wyznaczamy pozostałe kolumny tej macierzy

$$ST\begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} = S\left(T\begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}\right) = S(k_2) = \begin{pmatrix} w_1 \circ k_2 \\ w_2 \circ k_2 \\ w_3 \circ k_2 \end{pmatrix},$$

$$ST\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} = S\left(T\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}\right) = S(k_3) = \begin{pmatrix} w_1 \circ k_3 \\ w_2 \circ k_3 \\ w_3 \circ k_3 \end{pmatrix}.$$

Ostatecznie otrzymujemy

$$\boxed{m(ST) = \begin{pmatrix} w_1 \circ k_1 & w_1 \circ k_2 & w_1 \circ k_3 \\ w_2 \circ k_1 & w_2 \circ k_2 & w_2 \circ k_3 \\ w_3 \circ k_1 & w_3 \circ k_2 & w_3 \circ k_3 \end{pmatrix}.} \tag{6.1}$$

Wniosek 6.3.3. *Wyraz macierzy $m(ST)$ w $i-$tym wierszu i $j-$tej kolumnie, jest iloczynem skalarnym $i-$tego wiersza macierzy $m(S)$ i $j-$tej kolumny macierzy $m(T)$.*

6.4 Mnożenie macierzy.

Dotychczas zazwyczaj mówiliśmy o macierzach w kontekście przekształceń przestrzeni. W oparciu o te przekształcenia wprowadzimy w tym podrozdziale abstrakcyjne pojęcie mnożenia macierzy.

Definicja 6.4.1 (iloczyn macierzy). Iloczynem macierzy $m(S)$ i $m(T)$, odpowiadających przekształceniom S i T, nazywamy macierz $m(ST)$, będącą macierzą złożenia przekształceń S i T. Możemy symbolicznie zapisać to w postaci

$$m(S) \cdot m(T) = m(ST).$$

✔ **Przykład.** Wyznaczmy iloczyn macierzy $\begin{pmatrix} 3 & 0 & 8 \\ 7 & 1 & -5 \\ 2 & 0 & 4 \end{pmatrix}$, $\begin{pmatrix} 7 & 1 & 5 \\ 3 & 5 & -1 \\ 0 & -2 & -3 \end{pmatrix}$. Zgodnie ze wzorem (6.1) mamy

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

$$= \begin{pmatrix} 3\cdot 7+0\cdot 3+8\cdot 0 & 3\cdot 1+7\cdot 5+2\cdot(-2) & 3\cdot 5+0\cdot(-1)+8\cdot(-3) \\ 7\cdot 7+1\cdot 3+(-5)\cdot 0 & 7\cdot 1+1\cdot 5+(-5)\cdot(-2) & 7\cdot 5+1\cdot(-1)+(-5)\cdot(-3) \\ 2\cdot 7+0\cdot 3+4\cdot 0 & 2\cdot 1+0\cdot 5+4\cdot(-2) & 2\cdot 5+0\cdot(-1)+4\cdot(-3) \end{pmatrix} =$$

$$= \begin{pmatrix} 21 & 34 & -9 \\ 52 & 22 & 49 \\ 14 & -6 & -2 \end{pmatrix}.$$

Poznamy teraz najważniejsze własności działań na macierzach.

Fakt 6.4.2 **(własności mnożenia macierzy).**

(1) Mnożenie macierzy 3×3 *jest łączne, tzn.* $(m_1 \cdot m_2) \cdot m_3 = m_1 \cdot (m_2 \cdot m_3)$.

(2) Dla dowolnych macierzy 3×3 : m_1, m_2 *oraz wektora* $X \in R^3$ *zachodzi* $(m_1 m_2)X = m_1(m_2 X)$.

(3) Mnożenie macierzy nie jest przemienne, tzn. zazwyczaj $m_1 m_2 \neq m_2 m_1$.

Dowód (1). Mnożenie macierzy jest łaczne, gdyż łaczne jest składanie przekształceń. Dla przekształceń S, T, R mamy

$$(m(S)\cdot m(T))\cdot m(R) = m(ST)\cdot m(R) = m((ST)R) = m(S(TR)) =$$

$$= m(S)\cdot m(TR) = m(S)\cdot(m(T)\cdot m(R)).$$

Sprawdzenie tej własności przez bezpośrednie wyliczenie iloczynów macierzy byłoby bardzo czasochłonne. □

✍ **Ćwiczenie.** Podobny dowód, dla własności drugiej, pozostawimy do wykonania czytelnikowi. Podobnie łatwo sprawdzić, że mnożenie nie jest przemienne np. sprawdzając, że

$$\begin{pmatrix} 3 & 2 & -1 \\ 0 & 1 & 5 \\ 1 & -2 & -3 \end{pmatrix}\begin{pmatrix} 0 & -3 & -1 \\ 2 & -1 & -3 \\ 1 & 0 & 3 \end{pmatrix} \neq \begin{pmatrix} 0 & -3 & -1 \\ 2 & -1 & -3 \\ 1 & 0 & 3 \end{pmatrix}\begin{pmatrix} 3 & 2 & -1 \\ 0 & 1 & 5 \\ 1 & -2 & -3 \end{pmatrix}.$$

Co również pozostawimy do samodzielnego sprawdzenia.

Rozdział 7

Przekształcenia odwrotne

W rozdziale tym zajmiemy się pojęciem przekształcenia odwrotnego. Podobnie jak w przypadku dwuwymiarowym, również w przestrzeni R^3, dla danego przekształcenia możemy spróbować określić przekształcenie do niego odwrotne. W pierwszym podrozdziale zdefiniujmy, czym jest przekształcenie odwrotne i kiedy dane przekształcenie ma przekształcenie odwrotne, czyli kiedy jest odwracalne. Poznamy dwa równoważne określenia odwracalności i przekształcenia odwrotnego. Nauczymy się wyznaczać przekształcenie odwrotne i określać macierz takiego przekształcenia, najpierw dla przypadku ogólnego a następnie dla przekształceń liniowych.

7.1 Podstawowe definicje.

Definicja 7.1.1 (**odwracalność przekształcenia i przekształcenie odwrotne**).

(1) Przekształcenie T nazywamy odwracalnym, jeśli jest wzajemnie jednoznaczne, tzn. różnowartościowe i „na". Przekształceniem odwrotnym do $T : Z \to Z$ nazywamy takie przekształcenie $S : Z \to Z$, że jeśli $T(X) = X'$, to $S(X') = X$.

(2) Przekształcenie $T : Z \to Z$ jest odwracalne, jeśli istnieje przekształcenie $S : Z \to Z$, takie że $S \circ T = Id$ oraz $T \circ S = Id$. Przekształcenie S nazywane jest przekształceniem odwrotnym do T.

Dowód (równoważności definicji (1) i (2)).

Załóżmy, że T jest przekształceniem odwracalnym, a S jest przekształceniem odwrotnym do T w sensie definicji (1). Pokażemy, że wtedy T jest odwracalne w sensie definicji (2), a S jest przekształceniem do niego odwrotnym.
Zgodnie z definicją (1), dla każdego wektora $X \in Z$ mamy

$$S(T(X)) = S(X') = X,$$

oraz dla każdego $X' \in Z$

$$T(S(X')) = T(X) = X'.$$

Zatem przekształcenie T jest odwracalne w sensie definicji (2), a S jest przekształceniem do niego odwrotnym.

Pokażemy teraz implikację przeciwną. Załóżmy, że T jest odwracalne, a S jest przekształceniem odwrotnym do T wg definicji (2). Musimy pokazać, że T jest różnowartościowe i „na". Weźmy dwa różne wektory X, Y z przestrzeni Z. Chcemy pokazać, że wówczas $T(X) \neq T(Y)$. Jak wiemy z definicji (2), $S(T(X)) = X$, $S(T(Y)) = Y$. Gdyby $T(X) = T(Y)$, mielibyśmy

$$X = S(T(X)) = S(T(Y)) = Y,$$

ale to jest sprzeczne z założeniem, więc rzeczywiście przekształcenie T jest różnowartościowe. Pokażemy jeszcze, że dla dowolnego $X \in Z$ istnieje Y, taki że $X = T(Y)$, czyli że przekształcenie T jest „na". Przyjmijmy $Y = S(X)$, mamy wówczas

$$X = T(S(X)) = T(Y).$$

Przekształcenie T jest więc przekształceniem „na". Pozostaje pokazać, że przekształcenie S z definicji (2) ma własność: jeśli $T(X) = X'$, to $S(X') = X$. Przyjmijmy $T(X) = X'$, mamy wówczas $S(X') = S(T(X)) = X$, czyli S jest przekształceniem odwrotnym do T w sensie definicji (1). Podane w Definicji 7.1.1 dwa określenia przekształcenia odwracalnego i odwrotnego są więc równoważne. □

7.2 Wyznaczanie przekształcenia odwrotnego.

Przekształcenie odwrotne, jeśli istnieje, na ogół można wyliczyć. W tym podrozdziale pokażemy przykłady przekształceń odwrotnych, w kolejnych nauczymy się sprawdzać, czy dane przekształcenie ma przekształcenie odwrotne, określimy też macierz tego przekształcenia.

✔ **Przykład.** Niech przekształcenie $T : R^3 \to R^3$ będzie określone wzorem $T(X) = 3X+2W$, gdzie W jest pewnym ustalonym wektorem z R^3. Oznaczmy $T(X)$ jako X', mamy wówczas $X' = 3X + 2W$. Skoro $X \xrightarrow{T} X'$, to szukamy takiego przekształcenia T^{-1}, że $X' \xrightarrow{T^{-1}} X$. Spróbujmy wyrazić X za pomocą X'. Otrzymujemy

$$3X = X' + 2W,$$

$$X = \frac{1}{3}(X' - 2W) = \frac{1}{3}X' - \frac{2}{3}W.$$

Możemy więc stwierdzić, że T jest odwracalne, bo jesteśmy w stanie określić przekształcenie T^{-1} do niego odwrotne. Wyraża się ono wzorem

$$T^{-1}(Y) = \frac{1}{3}Y - \frac{2}{3}W.$$

Sprawdźmy zgodność tak wyznaczonego przekształcenia z Definicją 7.1.1, a zatem sprawdźmy, czy $T^{-1} \circ T = I$. Otrzymujemy
$T^{-1} \circ T(X) = T^{-1}(3X + 2W) = \frac{1}{3}(3X + 2W) - \frac{2}{3}W = X$, zatem rzeczywiście $T^{-1} \circ T = I$.

✍ **Ćwiczenie.** Czytelnik może sam, w analogiczny sposób sprawdzić, że $T \circ T^{-1} = I$.

✔ **Przykład.** Niech $T : R^3 \to R^3$ będzie takim przekształceniem, które wektorowi $\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$ przyporządkowuje wektor $\begin{pmatrix} 2x_1 \\ 3x_1 + x_2 \\ x_1 - x_2 - 2x_3 \end{pmatrix}$, czyli $T\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 2x_1 \\ 3x_1 + x_2 \\ x_1 - x_2 - 2x_3 \end{pmatrix}$. Macierz tego przekształcenia to $\begin{pmatrix} 2 & 0 & 0 \\ 3 & 1 & 0 \\ 1 & -1 & -2 \end{pmatrix}$.

Przyjmijmy, że obrazem wektora $\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$ jest wektor $\begin{pmatrix} x_1' \\ x_2' \\ x_3' \end{pmatrix}$. W takiej konwencji mamy

$$(\star)\begin{cases} x_1' = 2x_1 \\ x_2' = 3x_1 + x_2 \\ x_3' = x_1 - x_2 - 2x_3 \end{cases}.$$

Podobnie jak poprzednio, szukamy takiego przekształcenia T^{-1}, że

$$\begin{pmatrix} x_1' \\ x_2' \\ x_3' \end{pmatrix} \xrightarrow{T^{-1}} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}.$$

Przekształćmy układ $(\star)$. Otrzymujemy

$$\begin{aligned} &x_1 = \tfrac{1}{2}x_1', \\ &x_2 = x_2' - 3x_1 = x_2' - \tfrac{3}{2}x_1', \\ &x_3 = \tfrac{1}{2}(x_1 - x_2 - x_3') = \tfrac{1}{2}(\tfrac{1}{2}x_1' + \tfrac{3}{2}x_1' - x_2' - x_3') = x_1' - \tfrac{1}{2}x_2' - \tfrac{1}{2}x_3'. \end{aligned}$$

Wyznaczyliśmy więc przekształcenie odwrotne do T

$$T^{-1}\begin{pmatrix} x_1' \\ x_2' \\ x_3' \end{pmatrix} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} \frac{1}{2}x_1' \\ -\frac{3}{2}x_1' + x_2' \\ x_1' - \frac{1}{2}x_2' - \frac{1}{2}x_3' \end{pmatrix}.$$

Odpowiadająca temu przekształceniu macierz, to $m(T^{-1}) = \begin{pmatrix} \frac{1}{2} & 0 & 0 \\ -\frac{3}{2} & 1 & 0 \\ 1 & -\frac{1}{2} & -\frac{1}{2} \end{pmatrix}$.

✍ **Ćwiczenie.** Sprawdzenie zgodności tak wyznaczonego przekształcenia odwrotnego z Definicją 7.1.1, pozostawimy jako proste ćwiczenie dla czytelnika.

7.3 Odwracalność przekształceń liniowych.

W poprzednim podrozdziale pokazaliśmy przykłady przekształceń odwrotnych do danych. Niestety nie we wszystkich przypadkach da się określić takie przekształcenie. W tym podrozdziale określimy, kiedy taka operacja jest wykonalna. Podamy teraz bardzo użyteczne twierdzenie, które będzie najczęściej używanym przez nas narzędziem do badania odwracalności przekształceń.

Twierdzenie 7.3.1 (kolumny macierzy a odwracalność przekształcenia). *Niech $T : R^3 \to R^3$ będzie przekształceniem liniowym o macierzy $m(T)$. Mamy wówczas*

(1) jeśli kolumny macierzy $m(T)$ są liniowo zależne, to przekształcenie T nie jest odwracalne,

(2) jeśli kolumny macierzy $m(T)$ są liniowo niezależne, to przekształcenie T jest wzajemnie jednoznaczne, a więc odwracalne.

Dowód.

(1) Macierz przekształcenia T zapiszmy w postaci $m(T) = (K_1, K_2, K_3)$, gdzie $K_i \in R^3$ są jej kolumnami. Korzystając z liniowości przekształcenia T, otrzymujemy

$$T(X) = T\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = T\left[x_1 \cdot \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + x_2 \cdot \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} + x_3 \cdot \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}\right] =$$

$$= x_1 \cdot T\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + x_2 \cdot T\begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} + x_3 \cdot T\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} = x_1K_1 + x_2K_2 + x_3K_3.$$

Przy założeniu, że wektory K_1, K_2, K_3 są liniowo zależne, mamy że jeden z nich jest kombinacją liniową pozostałych. Przyjmijmy, że np. $K_3 = t \cdot K_1 + s \cdot K_2$. Wtedy dla dowolnego wektora $X \in R^3$ mamy

$$T(X) = x_1K_1 + x_2K_2 + x_3(t \cdot K_1 + s \cdot K_2) = (x_1 + x_3t)K_1 + (x_2 + x_3s)K_2.$$

Zatem każdy wektor zostaje przekształcony przez T na wektor będący kombinacją liniową wektorów K_1, K_2. Mogą zajść trzy sytuacje. Wektory K_1, K_2 mogą być

(a) oba zerowe. Wtedy wektory

$$T(X) = (x_1 + x_3t)K_1 + (x_2 + x_3s)K_2 = (x_1 + x_3t) \cdot 0 + (x_2 + x_3s) \cdot 0 = 0$$

również będą wektorami zerowymi.

(b) Mogą być współliniowe. Wtedy wektory $T(X)$ jako ich liniwe kombinacje będą wektorami z prostej wzdłuż wektorów K_1, K_2.

(c) Mogą być niewspółliniowe. Wtedy ich kombinacje liniowe $T(X)$ wyznaczają pewną płaszczyznę.

W każdym z tych przypadków obrazami wektorów X nie mogą być wszystkie wektory z przestrzeni R^3. Tak więc T nie jest przekształceniem „na”, a zatem zgodnie z Definicją 7.1.1 nie jest odwracalne.

(2) Jeśli K_1, K_2, K_3 są wektorami liniowo niezależnymi, to każdy wektor $Y \in R^3$ można jednoznacznie przedstawić w postaci ich kombinacji liniowej $Y = t_1K_1 + t_2K_2 + t_3K_3$. Jednocześnie z pierwszej części dowodu pamiętamy, że $T\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = x_1K_1 + x_2K_2 + x_3K_3$. Wynika stąd, że T jest różnowartościowe, bo jeśli $Y = T(X_1) = T(X_2)$ dla $X_1 \neq X_2$, to Y miałby dwa różne rozkłady w kombinację wektorów K_1, K_2, K_3, a to jak wiemy jest niemożliwe. Przekształcenie T jest również „na”, bo jeśli weźmiemy dowolny $Y = t_1K_1 + t_2K_2 + t_3K_3$, to $Y = T\begin{pmatrix} t_1 \\ t_2 \\ t_3 \end{pmatrix}$, więc każdy Y jest obrazem pewnego wektora przez przekształcenie T. Jak wynika z Definicji 7.1.1, przekształcenie T jest odwracalne.

□

Z powyższego twierdzenia możemy wyprowadzić bardzo pożyteczny wniosek, który bardzo często okaże się być najłatwiejszą metodą sprawdzenia odwracalności przekształcenia.

Wniosek 7.3.2 (wyznacznik a odwracalność przekształcenia). *Przekształcenie T jest odwracalne wtedy i tylko wtedy gdy* $\det(m(T)) \neq 0$.

Dowód.
Jak wiemy z Faktu 4.2.7 kolumny macierzy $m(T)$ są liniowo niezależne wtedy i tylko wtedy gdy $\det(m(T)) \neq 0$, a to, na podstawie Twierdzenia 7.3.1, jest równoważne odwracalności przekształcenia T. □

Uwaga 7.3.3 (macierz nieosobliwa). *Macierze m, dla których* $\det(m) \neq 0$ *nazywane są macierzami nieosobliwymi. Tak więc T jest odwracalne wtedy i tylko wtedy gdy macierz T jest nieosobliwa.*

✔ **Przykład.**

(1) Przekształceniem odwrotnym do jednokładności o skali r i środku w punkcie 0, czyli J^0_r, jest jednokładność o środku w tym samym punkcie i skali $\frac{1}{r}$, czyli $J^0_{\frac{1}{r}}$.

(2) Przekształceniem odwrotnym do obrotu względem osi OX o kąt θ, czyli $R_{OX,\theta}$, jest obrót względem tej samej osi, ale o kąt $-\theta$, czyli $R_{OX,-\theta}$.

W powyższych przykładach przekształcenie odwrotne do przekształcenia liniowego, jest również przekształceniem liniowym. Zastanówmy się, czy tak jest zawsze.

Lemat 7.3.4 (liniowość przekształcenia odwrotnego). *Jeśli T jest odwracalnym przekształceniem liniowym, to T^{-1} jest też przekształceniem liniowym.*

Dowód. Aby pokazać liniowość przekształcenia T^{-1} należy pokazać jego addytywność i jednorodność. Niech X, Y będą dowolnymi wektorami z R^3. Oznaczmy

$W_1 = T^{-1}(X+Y)$, $W_2 = T^{-1}(X) + T^{-1}(Y)$. Chcielibyśmy, aby $W_1 = W_2$, co oznaczałoby, że przekształcenie T^{-1} jest addytywne. Nałóżmy na oba równania przekształcenie T. Otrzymujemy

$$T(W_1) = T(T^{-1}(X+Y)) = X+Y,$$

$$T(W_2) = T(T^{-1}(X) + T^{-1}(Y)) = X+Y.$$

Tak więc $T(W_1) = T(W_2)$. Przekształcenie T jest odwracalne, więc jest różnowartościowe, skąd wnioskujemy, że $W_1 = W_2$. □

Ćwiczenie. Jednorodność przekształcenia T^{-1} pokazuje się podobnie. Zadanie to pozostawimy jako proste ćwiczenie dla czytelnika.

7.4 Macierz przekształcenia odwrotnego.

Dla danego odwracalnego przekształcenia liniowego $T : R^3 \to R^3$ o macierzy $m(T)$ przydatna mogłaby okazać się umiejętność znajdowania macierzy przekształcenia T^{-1}. Spróbujmy zatem zastanowić się, czym jest i jakie ma własności macierz przekształcenia odwrotnego.

Fakt 7.4.1. *Dla dowolnego przekształcenia liniowego T o macierzy $m(T)$ i przekształcenia T^{-1} do niego odwrotnego, mamy*

$$m(T^{-1})m(T) = m(T^{-1}T) = m(I) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix},$$

oraz

$$m(T)m(T^{-1}) = m(TT^{-1}) = m(I) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Definicja 7.4.2 (macierz odwrotna). Macierzą odwrotną do macierzy m rozmiaru 3×3, nazywamy taką macierz n rozmiaru 3×3, dla której

$$m \cdot n = n \cdot m = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Macierz m, mającą macierz odwrotną, nazywamy macierzą odwracalną.

➢ Własności macierzy odwrotnych

Korzystając z tego, że możemy utożsamiać macierze z zadanymi przez nie przekształceniami, a mnożenie macierzy ze składaniem przekształceń, możemy określić następujące własności macierzy odwrotnych.

Fakt 7.4.3.

(1) Macierz odwracalna m, czyli macierz o niezerowym wyznaczniku, ma dokładnie jedną macierz odwrotną, którą oznaczamy m^{-1}.

(2) Macierz m jest odwracalna wtedy i tylko wtedy gdy kolumny tej macierzy są wektorami liniowo niezależnymi, a to zachodzi wtedy i tylko wtedy gdy wyznacznik tej macierzy $\det(m) \neq 0$, *czyli wtedy i tylko wtedy gdy wiersze macierzy m są liniowo niezależne.*

(3) Macierz odwrotna do $m(T)$ jest macierzą przekształcenia T^{-1}, czyli

$$m(T^{-1}) = (m(T))^{-1}.$$

Dowody własności podanych w powyższym fakcie wynikają wprost z udowodnionych wcześniej własności przekształceń odwrotnych. Tutaj je pominiemy.

✔ **Przykład.**

(1) Jeśli $t_1 \neq 0, t_2 \neq 0, t_3 \neq 0$, to $\begin{pmatrix} t_1 & 0 & 0 \\ 0 & t_2 & 0 \\ 0 & 0 & t_3 \end{pmatrix}^{-1} = \begin{pmatrix} \frac{1}{t_1} & 0 & 0 \\ 0 & \frac{1}{t_2} & 0 \\ 0 & 0 & \frac{1}{t_3} \end{pmatrix}$.

(2) $\begin{pmatrix} 1 & a & b \\ 0 & 1 & c \\ 0 & 0 & 1 \end{pmatrix}^{-1} = \begin{pmatrix} 1 & -a & (-b+ac) \\ 0 & 1 & -c \\ 0 & 0 & 1 \end{pmatrix}$.

Wydaje się, że użyteczne mogłoby być wyprowadzenie ogólnego wzoru na macierz odwrotną, a w konsekwencji również na przekształcenie odwrotne do danego.

Fakt 7.4.4. *Jeśli T jest przekształceniem odwracalnym i jeśli przekształcenie S spełnia jeden z warunków:*

(a) $TS = I$,

(b) $ST = I$,

to musi też spełniać drugi z tych warunków, a więc $S = T^{-1}$.

Dowód. Załóżmy, że S spełnia pierwszy warunek, czyli $TS = I$. Wtedy

$$I = T^{-1}T = T^{-1}IT = T^{-1}TST = ST,$$

czyli $ST = I$. Otrzymaliśmy zatem drugi warunek. Podobnie można sprawdzić drugi przypadek. □

✍ **Ćwiczenie.** Sprawdzenie, że jeśli przekształcenie S spełnia drugi warunek, to musi też spełniać pierwszy, pozostawimy do samodzielnego wykonania przez czytelnika.

Uwaga 7.4.5. *Istnieją nieodwracalne przekształcenia T, dla których istnieje przekształcenie S, takie że $TS = I$ ale $ST \neq I$. Przekształceniami takimi nie będziemy się jednak tutaj zajmować.*

Wniosek 7.4.6. *Jeśli m jest macierzą odwracalną, zaś macierz n spełnia jeden z następujących warunków*

$$(a)\quad m\cdot n=\begin{pmatrix}1&0&0\\0&1&0\\0&0&1\end{pmatrix},\quad (b)\quad n\cdot m=\begin{pmatrix}1&0&0\\0&1&0\\0&0&1\end{pmatrix},$$

to

$$n=m^{-1}.$$

➢ Wzór na macierz odwrotną

Podamy teraz wzór an macierz odwrotną do danej. Wprowadźmy jednak najpierw pewne oznaczenia.

$m=\begin{pmatrix}w_1\\w_2\\w_3\end{pmatrix}$, gdzie w_i oznacza i-ty wiersz macierzy m,

$w_i^T=\begin{pmatrix}a_i\\b_i\\c_i\end{pmatrix}$ to kolumna (wektor z R^3) utworzona z wiersza w_i.

Przypomnijmy również, że $\det(m)=det(m^T)$, oraz $v\cdot(k_1\ k_2\ k_3)=(vk_1\ vk_2\ vk_3)$.

Twierdzenie 7.4.7 (wzór na macierz odwrotną). *Dla danej macierzy* $m=\begin{pmatrix}w_1\\w_2\\w_3\end{pmatrix}$ *macierz odwrotna m^{-1} przyjmuje postać*

$$\boxed{m^{-1}=\begin{pmatrix}w_1\\w_2\\w_3\end{pmatrix}^{-1}=\frac{1}{det(m)}\cdot(w_2^T\times w_3^T\quad w_3^T\times w_1^T\quad w_1^T\times w_2^T).}$$

Dowód. Oznaczmy $n=\frac{1}{det(m)}\cdot(w_2^T\times w_3^T\quad w_3^T\times w_1^T\quad w_1^T\times w_2^T)$ i obliczmy $m\cdot n$.

$$m\cdot n=\begin{pmatrix}w_1\\w_2\\w_3\end{pmatrix}\cdot\left(\frac{1}{det(m)}\cdot w_2^T\times w_3^T\quad \frac{1}{det(m)}\cdot w_3^T\times w_1^T\quad \frac{1}{det(m)}\cdot w_1^T\times w_2^T\right)=\begin{pmatrix}x_1&y_1&z_1\\x_2&y_2&z_2\\x_3&y_3&z_3\end{pmatrix}$$

Chcielibyśmy, aby $\begin{pmatrix}x_1&y_1&z_1\\x_2&y_2&z_2\\x_3&y_3&z_3\end{pmatrix}=\begin{pmatrix}1&0&0\\0&1&0\\0&0&1\end{pmatrix}$.

Spróbujmy więc obliczyć kolejno wyrazy tej macierzy. Zacznijmy od pierwszej kolumny.

$$x_1=w_1\circ\left(\frac{1}{det(m)}\cdot w_2^T\times w_3^T\right)=\frac{1}{det(m)}\cdot w_1\circ(w_2\times w_3)=\frac{1}{det(m)}\cdot det(m)=1,$$

$$x_2=w_2\circ\left(\frac{1}{det(m)}\cdot w_2^T\times w_3^T\right)=\frac{1}{det(m)}\cdot w_2\circ(w_2\times w_3)=0,$$

$$x_3=w_3\circ\left(\frac{1}{det(m)}\cdot w_2^T\times w_3^T\right)=\frac{1}{det(m)}\cdot w_3\circ(w_2\times w_3)=0.$$

Otrzymujemy więc $\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$. Podobnie obliczymy, że $\begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$, oraz $\begin{pmatrix} z_1 \\ z_2 \\ z_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$. Tak więc otrzymujemy, że $m \cdot n = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$. Zatem rzeczywiście $n = m^{-1}$. □

✍ **Ćwiczenie.** Czytelnik może sam w podobny sposób sprawdzić, że $\begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$ oraz $\begin{pmatrix} z_1 \\ z_2 \\ z_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$.

✔ **Przykład.** Znajdźmy macierz odwrotną do macierzy $m = \begin{pmatrix} 2 & 3 & 1 \\ 0 & -2 & 1 \\ -1 & 0 & 5 \end{pmatrix}$. Wyznacznik tej macierzy jest równy $\det(m) = -25 \neq 0$. Macierz odwrotna to

$$m^{-1} = -\frac{1}{25}\left(\begin{pmatrix} 0 \\ -2 \\ 1 \end{pmatrix} \times \begin{pmatrix} -1 \\ 0 \\ 5 \end{pmatrix} \quad \begin{pmatrix} -1 \\ 0 \\ 5 \end{pmatrix} \times \begin{pmatrix} 2 \\ 3 \\ 1 \end{pmatrix} \quad \begin{pmatrix} 2 \\ 3 \\ 1 \end{pmatrix} \times \begin{pmatrix} 0 \\ -2 \\ 1 \end{pmatrix}\right) =$$

$$= -\frac{1}{25}\begin{pmatrix} -10 & -15 & 5 \\ -1 & 11 & -2 \\ -2 & -3 & -4 \end{pmatrix}.$$

Rozdział 8

Przekształcenia liniowe a objętość i orientacja

W tym rozdziale zbadamy jaki wpływ na orientację wektorów w przestrzeni oraz objętość równoległościanu ma zadziałanie na daną przestrzeń odwracalnym przekształceniem liniowym.

8.1 Orientacja w przestrzeni.

W Rozdziale 4 wprowadziliśmy pojęcie orientacji wektorów w przestrzeni. Spróbujmy zastanowić się jak zmienia się orientacja wektorów pod wpływem działania przekształcenia liniowego. Określimy tutaj, które przekształcenia zmieniają, a które zachowują orientację wektorów.

Definicja 8.1.1. Przekształcenie $T : R^3 \to R^3$ zachowuje orientację, jeśli dla każdej trójki wektorów prawoskrętny układ wektorów zostaje przekształcony na prawoskrętny, zaś zmienia orientację, gdy prawoskrętny układ wektorów zostaje przekształcony na układ lewoskrętny.

Uwaga 8.1.2. *Jeśli dla jednej trójki wektorów przekształcenie zachowuje (zmienia) ich orientację, to zachodzi to dla wszystkich trójek wektorów. Zatem możemy sprawdzać zachowanie (zmianę) orientacji na jednej wybranej trójce wektorów.*

Nie będziemy tutaj dowodzić powyższej uwagi.

Niech wersory e_1, e_2, e_3 tworzą prawoskrętnym układem wektorów R^3, natomiast u, v, w to obrazy wektorów e_1, e_2, e_3 przez przekształcenie T. Przypomnijmy, że układ wektorów u, v, w jest prawoskrętny wtedy i tylko wtedy gdy $\det(u, v, w) > 0$, a lewoskrętny wtedy i tylko wtedy gdy $\det(u, v, w) < 0$.

Wniosek 8.1.3. *Przekształcenie T zachowuje orientacje, gdy*

$$\det(u, v, w) > 0,$$

zaś zmienia w przeciwnym wypadku.

Przypomnijmy, że $m(T) = (u, v, w)$, gdzie u, v, w są obrazami wersorów przez przekształcenie T. Możemy zatem sformułować następujący:

Wniosek 8.1.4. *Jeśli macierzą przekształcenia T jest $m(T)$, to przekształcenie T zachowuje orientację, gdy $\det(m(T)) > 0$, natomiast zmienia orientację w przeciwnym przypadku.*

Przykład. Przykładem przekształcenia zmieniającego orientację jest symetria względem płaszczyzny O_{xy}. Jak pokazaliśmy w podrozdziale 5.5 macierzą tego przekształcenia jest $m(S_{xy}) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix}$. Wyznacznik tej macierzy wynosi

$$\det m(S_{xy}) = \det \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix} = -1 < 0,$$

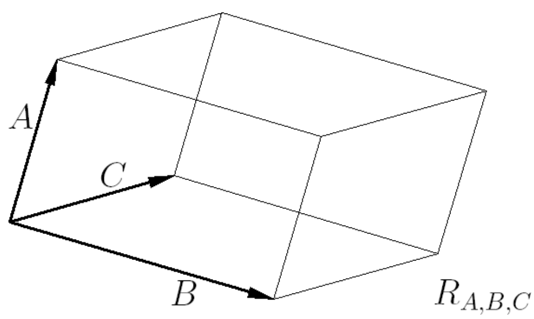
więc zgodnie z Wnioskiem 8.1.4 symetria względem O_{xy} jest przekształceniem zmieniającym orientację.

8.2 Równoległościan i jego obraz przez przekształcenie.

Zajmiemy się teraz równoległościanami w przestrzeni. Sprawdzimy, co jest obrazem równoległościanu przez przekształcenie liniowe i jak zmienia się jego objętość po przejściu przez takie przekształcenie. Zdefiniujmy jednak najpierw, czym jest równoległościan.

Definicja 8.2.1 (równoległościan). Równoległościanem, zbudowanym na liniowo niezależnych wektorach $A, B, C \in R^3$, nazywamy zbiór wektorów będących kombinacją liniową wektorów A, B, C ze współczynnikami należącymi do przedziału $[0, 1]$

$$R_{A,B,C} = \{t \cdot A + s \cdot B + r \cdot C : t \in [0, 1], s \in [0, 1], r \in [0, 1]\}.$$



Twierdzenie 8.2.2 (obraz równoległościanu przez przekształcenie liniowe). *Niech T będzie odwracalnym przekształceniem liniowym R^3, zaś A, B, C liniowo niezależnymi wektorami. Wtedy*

(1) wektory $T(A), T(B), T(C)$ też są liniowo niezależne,

(2) obrazem $T(R_{A,B,C})$ równoległościanu $R_{A,B,C}$ przez przekształcenie T, jest też równoległościan.

Dowód.

(1) Gdyby $T(A), T(B), T(C)$ były liniowo zależne, wtedy jeden z nich byłby kombinacją liniową pozostałych, czyli np. $T(C) = s \cdot T(A) + t \cdot T(B)$. Gdybyśmy teraz nałożyli na obie strony przekształcenie T^{-1} i skorzystali z jego liniowości, otrzymalibyśmy
$$T^{-1}(T(A)) = T^{-1}(sT(A) + tT(B)),$$
ale $T^{-1}(T(A)) = A$ oraz
$$T^{-1}(sT(A) + tT(B)) = sT^{-1}(T(B)) + tT^{-1}(T(C)) = sB + tC.$$
Otrzymujemy więc $A = sB + tC$, co przeczy założeniu o liniowej niezależności wektorów A, B, C. Zatem również wektory $T(A), T(B), T(C)$ muszą być liniowo niezależne.

(2) Z liniowości przekształcenia T otrzymujemy natychmiast
$$T(tA + sB + vC) = tT(A) + sT(B) + vT(C).$$
Zatem
$$T(R_{A,B,C}) = T(\{tA + sB + vC : t, s, v \in [0,1]\}) =$$
$$= \{tT(A) + sT(B) + vT(C) : t, s, v \in [0,1]\} = R_{T(A),T(B),T(C)}.$$

□

➢ Przekształcenie liniowe a objętość brył

Przypomnijmy, że jeśli $m(T) = (K_1,\ K_2,\ K_3)$, gdzie K_i są kolumnami macierzy m, to $T(e_1) = K_1, T(e_2) = K_2,\ T(e_3) = K_3$. Tak więc $T(R_{e_1,e_2,e_3}) = R_{K_1,K_2,K_3}$. Jeśli objętość równoległościanu oznaczymy przez $|R|$, to
$$|R_{K_1,K_2,K_3}| = |det(K_1, K_2, K_3)| = |det(m(T))|.$$
Zauważmy, że R_{e_1,e_2,e_3}, jest sześcianem o krawędzi 1, więc jego objętość jest równa 1. Możemy więc zapisać
$$|T(R_{e_1,e_2,e_3})| = |det(m(T))| \cdot |R_{e_1,e_2,e_3}|.$$
Przekształcenie T zmienia objętość sześcianu R_{e_1,e_2,e_3} w stosunku równym modułowi wyznacznika $|det(m(T))|$. Pokażemy, że to samo jest prawdą dla dowolnego równoległościanu $R_{A,B,C}$.

Twierdzenie 8.2.3 (przekształcenia liniowe a objętość równoległościanu). *Dla dowolnego równoległościanu $R_{A,B,C}$ i dowolnego odwracalnego przekształcenia $T : R^3 \to R^3$ zachodzi*
$$|T(R_{A,B,C})| = |det(m(T))| \cdot |R_{A,B,C}|.$$

Zanim udowodnimy powyższe twierdzenie, wprowadzimy pomocniczy lemat.

Lemat 8.2.4 (wyznacznik iloczynu macierzy). *Dla m_1, m_2, będących dowolnymi macierzami rozmiaru 3×3, mamy*

$$\det(m_1 m_2) = \det(m_1) \cdot \det(m_2).$$

Czysto rachunkowy dowód tego lematu pomijamy. Możemy teraz udowodnić Twierdzenie 8.2.3.

Dowód. Niech m_2 będzie macierzą utworzoną z wektorów A, B, C jako kolumn, czyli $m_2 = (A,\ B,\ C)$. Natommiast m_1 niech będzie macierzą przekształcenia T, czyli $m_1 = m(T)$. Wówczas ich iloczyn możemy zapisać w postaci

$$m_1 \cdot m_2 = m_1 \cdot (A,\ B,\ C) = (m_1 A,\ m_1 B,\ m_1 C) =$$

$$= (m(T)A,\ m(T)B,\ m(T)C) = (T(A),\ T(B),\ T(C)).$$

Otrzymaliśmy więc

$$|T(R_{A,B,C})| = |R_{T(A),T(B),T(C)}| = |\det(T(A), T(B), T(C))| = |\det(m_1 m_2)| =$$

$$|\det(m(T)(A,\ B,\ C))| = |\det(m(T))| \cdot |\det(A,\ B,\ C)| = |\det(m(T))| \cdot |R_{A,B,C}|.$$

□

Poniżej podamy wniosek z powyższego twierdzenia. Nie będziemy przytaczać tutaj jego dowodu, gdyż dowód ten jest analogiczny do dowodu podobnego faktu dla przekształceń i pól równoległoboków, naszkicowanego w pierwszej części skryptu.

Wniosek 8.2.5 (objętość dowolnej bryły przez przekształcenie liniowe). *Dla dowolnej bryły $B \in R^3$ i dowolnego przekształcenia liniowego $T : R^3 \to R^3$ mamy*

$$|T(B)| = |det(m(T))| \cdot |B|.$$

Rozdział 9

Przekształcenia afiniczne

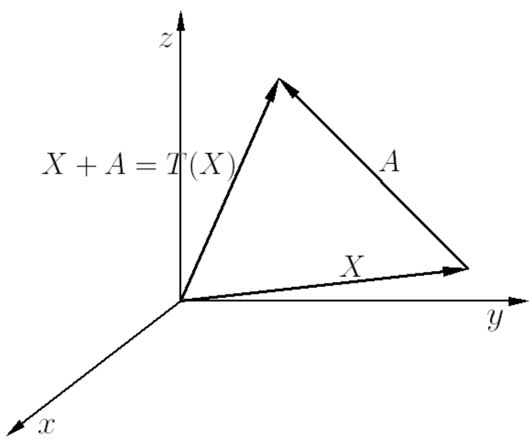
Dotychczas zajmowaliśmy się tylko przekształceniami liniowymi. W tym rozdziale poznamy inny rodzaj przekształceń - przekształcenia afiniczne, zdefiniujemy to pojęcie oraz poznamy przykłady. W kolejnym podrozdziale zajmiemy się własnościami takich przekształceń.

9.1 Podstawowe definicje.

Zanim podamy formalną definicję przekształcenia afinicznego, zapoznajmy się z przykładem takiego przekształcenia.

✔ **Przykład.** Niech przekształcenie T będzie przesunięciem (translacją) o wektor $A = \begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix}$. Dla tak określonego przekształcenia obrazem wektora $X = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$ jest wektor

$$T \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} + \begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix}.$$



Jeśli wektor A nie jest wektorem zerowym, to przekształcenie T, określone powyżej, nie jest przekształceniem liniowym. Podamy teraz dwa uzasadnienia tego stwierdzenia.

(1) Wzór określający przekształcenie T nie ma postaci wymaganej dla przekształceń liniowych.

(2) Przekształcenie T nie jest addytywne. Zauważmy, że

$$T(X+Y) = X+Y+A,$$

$$T(X)+T(Y) = X+A+Y+A = X+Y+2A.$$

Tak więc

$$T(X+Y) \neq T(X)+T(Y).$$

Translacja, podobnie jak wiele innych znanych nam już przekształceń, należy do grupy przekształceń nazywanych przekształceniami afinicznymi.

Definicja 9.1.1 (przekształcenia afiniczne). Przekształcenie $T : R^3 \to R^3$ jest przekształceniem afinicznym jeśli można je opisać wzorem postaci

$$T\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} + \begin{pmatrix} d_1 \\ d_2 \\ d_3 \end{pmatrix} = \begin{pmatrix} a_1x_1 + a_2x_2 + a_3x_3 + d_1 \\ b_1x_1 + b_2x_2 + b_3x_3 + d_3 \\ c_1x_1 + c_2x_2 + c_3x_3 + d_3 \end{pmatrix},$$

dla pewnej macierzy $\begin{pmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{pmatrix}$ i pewnego wektora $\begin{pmatrix} d_1 \\ d_2 \\ d_3 \end{pmatrix}$. Macierz $\begin{pmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{pmatrix}$ nazywamy macierzą części liniowej przekształcenia afinicznego T, zaś wektor $\begin{pmatrix} d_1 \\ d_2 \\ d_3 \end{pmatrix}$ wektorem przesunięcia tego przekształcenia.

Podamy teraz kilka przykładów przekształceń afinicznych.

✔ **Przykład.**

(1) Translacja o wektor $T_A(X) = X + A$.

(2) Każde przekształcenie liniowe jest przekształceniem afinicznym, bo jeżeli $m = m(T)$, to $T(X) = mX = mX + \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$.

(3) Przekształceniami afinicznymi są również rzuty na proste i płaszczyzny nie przechodzące przez $\begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$, symetrie względem takich płaszczyzn i prostych, jednokładność względem punktów różnych od $\begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$, symetrie środkowe względem punktów różnych od $\begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$.

9.2 Własności przekształceń afinicznych.

Poznamy teraz najważniejsze własności przekształceń afinicznych. Na początek, podobnie jak dla przekształceń liniowych określimy własność charakteryzującą złożenie przekształceń afinicznych.

Fakt 9.2.1 (złożenie przekształceń afinicznych). *Złożenie przekształceń afinicznych jest też przekształceniem afinicznym.*

Dowód. Niech T_1, T_2 będą przekształceniami afinicznymi zadanymi wzorami $T_1(X) = m_1X + A_1,\ \ T_2(X) = m_2X + A_2$. Ich złożeniem jest przekształcenie

$$T_2 \circ T_1(X) = T_2(m_1X + A_1) = m_2(m_1X + A_1) + A_2 = (m_2m_1)X + m_2A_1 + A_2.$$

Przekształcenie to również jest afiniczne. Macierzą części liniowej tego przekształcenia jest iloczyn macierzy m_2m_1, natomiast wektorem przesunięcia jest wektor $m_2A_1+A_2$. □

Wniosek 9.2.2 (macierz złożenia przekształceń afinicznych). *Macierz części liniowej złożenia przekształceń afinicznych jest iloczynem części liniowych macierzy tych przekształceń.*

W dalszej części zastanowimy się nad odwracalnością przekształceń afinicznych. Pokażemy kiedy dane przekształcenie jest odwracalne i jaką postać przyjmuje przekształcenie odwrotne do danego.

Fakt 9.2.3 (odwracalność przekształcenia afinicznego). *Przekształcenie afiniczne, dane wzorem $T(X) = mX + A$, jest odwracalne wtedy i tylko wtedy gdy część liniowa tego przekształcenia jest odwracalna, a to jak wiemy zachodzi wtedy i tylko wtedy gdy* $\det(m) \neq 0$.

Dowód. Pokażemy, że jeśli przekształcenie afiniczne T jest odwracalne, to jego część liniowa też jest odwracalna. Gdyby tak nie było, to wtedy część liniowa nie byłaby przekształceniem „na". Moglibyśmy wówczas wybrać taki wektor $Y \in R^3$, który nie byłby obrazem wektora z R^3 przez przekształcenie wyznaczone przez część liniową przekształcenia T. Oznacza to, że dla każdego $X \in R^3$ mielibyśmy $mX \neq Y$. Pokażemy, że wtedy wektor $Y + A$ nie byłby obrazem żadnego wektora z R^3 przez przekształcenie T. Gdyby było inaczej, czyli gdyby istniał taki wektor X, że $T(X) = Y + A$, to wtedy $mX + A = Y + A$, ale wówczas $mX = Y$. Tak nie może się jednak zdarzyć, bo zakładaliśmy, że Y nie jest obrazem żadnego wektora przez przekształcenie wyznaczone przez część liniową przekształcenia T. Tak więc jeśli część liniowa nie jest odwracalna, to przekształcenie T nie jest przekształceniem „na", a więc nie jest przekształceniem odwracalnym. Mamy więc udowodnioną implikację w jedną stronę.

Pokażemy teraz, że jeśli macierz m jest odwracalna, czyli jeśli część liniowa przekształcenia T jest odwracalna, to również przekształcenie T jest odwracalne. Spróbujmy znaleźć wzór na przekształcenie odwrotne. Oznaczmy obraz wektora X przez przekształcenie T jako X'. Mamy wówczas $X' = mX + A$, wtedy $mX = X' - A$. Zatem

$$X = m^{-1}mX - m^{-1}A = m^{-1}(X' - A).$$

Możemy stąd wnioskować, że przekształcenie odwrotne istnieje, bo potrafimy znaleźć jego wzór. Dowiedliśmy więc, że jeśli przekształcenie odpowiadające części liniowej przekształcenia afinicznego T jest odwracalne, to odwracalne musi też być przekształcenie T. Udowodniliśmy zatem drugą implikację, co kończy dowód. □

Na podstawie równości wyprowadzanych w powyższym dowodzie możemy sformułować następujący wniosek.

Wniosek 9.2.4 (przekształcenie odwrotne do przekształcenia afinicznego). *Przekształcenie odwrotne do przekształcenia afinicznego $T(X) = mX + A$, dane jest wzorem*

$$T^{-1}(X) = m^{-1}X - m^{-1}A.$$

Zatem przekształcenie odwrotne do przekształcenia afinicznego, zgodnie z Definicją 9.1.1, samo jest afiniczne.

Określimy teraz jaki wpływ na orientację w przestrzeni ma działanie danego przekształcenia afinicznego.

Fakt 9.2.5 (przekształcenie afiniczne a orientacja w przestrzeni). *Odwracalne przekształcenie afiniczne $T(X) = mX + A$ zachowuje orientację w przestrzeni gdy* $\det(m) > 0$, *zaś zmienia orientację, gdy* $\det(m) < 0$.

Dowód. Odwracalne przekształcenie afiniczne T jest złożeniem przekształcenia liniowego zadanego macierzą m oraz pewnej translacji. Ponieważ translacja zachowuje orientację, więc to czy przekształcenie T zachowuje orientację zależy od tego, czy przekształcenie liniowe zadane macierzą m ją zachowuje. Jak wiemy przekształcenie liniowe zachowuje orientację wtedy, gdy wyznacznik macierzy tego przekształcenia jest dodatni, w tym przypadku $\det(m) > 0$, a zmienia orientację, gdy $\det(m) < 0$. □

Możemy zastanawiać się jaki wpływ na miarowe własności figur, ma działanie na nie przekształceniami afinicznymi. Poniższy fakt określa jak zachowuje się objętość bryły pod wpływem działania takiego przekształcenia.

Fakt 9.2.6 (przekształcenie afiniczne a objętość bryły). *Odwracalne przekształcenie afiniczne $T(X) = mX + A$ zmienia objętość wszystkich brył w proporcji równej* $|\det(m)|$, *tzn. dla dowolnej bryły F mamy*

$$|T(F)| = |\det(m)| \cdot |F|.$$

Dowód. Przekształcenie T jest złożeniem przekształcenia liniowego i translacji. Jak wiemy translacja nie zmienia objętości bryły, zatem cała zmiana objętości jest spowodowana działaniem przekształcenia liniowego, zatem jest dana wzorem określonym we Wniosku 8.2.5, czyli

$$|T(F)| = |\det(m)| \cdot |F|.$$

□

Rozdział 10

Wartości i wektory własne

W rozdziale tym poznamy ważne pojęcia, charakteryzujące przekształcenia, które bardzo często będą się pojawiać w dalszej części skryptu. W oparciu o te pojęcia, uda nam się wyodrębnić pewną grupę przekształceń, o charakterystycznych własnościach, którą zajmiemy się w kolejnym rozdziale.

10.1 Wartości i wektory własne.

Definicja 10.1.1 (wartości i wektory własne). Liczba t jest wartością własną przekształcenia liniowego $T : R^3 \to R^3$, jeśli dla pewnego niezerowego wektora $X \in R^3$ zachodzi $T(X) = t \cdot X$.
Każdy wektor X, dla którego $T(X) = t \cdot X$, jest wektorem własnym dla wartości własnej t.

Zilustrujmy wprowadzone powyżej pojęcia przykładami.

✔ **Przykład.**

(1) Przykład geometryczny.
Niech Π będzie płaszczyzną przechodzącą przez 0. S_Π niech będzie symetrią względem płaszczyzny Π. Zauważmy, że jeśli $X \in \Pi$, to $S_\Pi(X) = X = 1 \cdot X$, więc, zgodnie z Definicją 10.1.1, liczba 1 jest wartością własną przekształcenia S_Π, zaś X jest odpowiadającym tej wartości wektorem własnym. Inną wartością własną tego przekształcenia jest -1. Odpowiadające jej wektory własne będą spełniać warunek $S_\Pi(X) = -X = (-1) \cdot X$. Jak nietrudno się przekonać, będą to wektory z prostej prostopadłej do płaszczyzny Π i przechodzącej przez 0.

(2) Przykład algebraiczny.
Niech $T : R^3 \to R^3$ będzie przekształceniem o macierzy $m(T) = \begin{pmatrix} d_1 & 0 & 0 \\ 0 & d_2 & 0 \\ 0 & 0 & d_3 \end{pmatrix}$.
Zauważmy, że $T(e_1) = \begin{pmatrix} d_1 \\ 0 \\ 0 \end{pmatrix} = d_1 \cdot \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} = d_1 e_1$, czyli d_1 jest wartością własną T, natomiast e_1 jest odpowiadającym jej wektorem własnym. Podobnie

możemy sprawdzić, że $T(e_2) = d_2e_2$ oraz $T(e_3) = d_3e_3$, tak więc d_2, d_3 również są wartościami własnymi T, dla wektorów własnych $e_2,\ e_3$.

10.2 Szukanie wartości własnych przekształcenia liniowego.

Zastanówmy się teraz jak, dla danego przekształcenia liniowego $T : R^3 \to R^3$ o macierzy $m(T) = \begin{pmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{pmatrix}$ znaleźć jego wartości własne. Najkorzystniej byłoby określić sposób pozwalający na wyznaczenie wszystkich wartości własnych danego przekształcenia. Poniżej wyprowadzimy proste kryterium, dzięki któremu będziemy potrafili tego dokonać. Zgodnie z Definicją 10.1.1 szukamy takich liczb t, które spełniają

$$\begin{pmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} tx_1 \\ tx_2 \\ tx_3 \end{pmatrix}$$

dla pewnego niezerowego wektora $\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$. Po przekształceniu powyższej równości otrzymujemy

$$\begin{pmatrix} a_1x_1 + b_1x_2 + c_1x_3 - tx_1 \\ a_2x_1 + b_2x_2 + c_2x_3 - tx_2 \\ a_3x_1 + b_3x_2 + c_3x_3 - tx_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix},$$

co równoznacznie możemy zapisać jako

$$x_1 \cdot \begin{pmatrix} a_1 - t \\ a_2 \\ a_3 \end{pmatrix} + x_2 \cdot \begin{pmatrix} b_1 \\ b_2 - t \\ b_3 \end{pmatrix} + x_3 \cdot \begin{pmatrix} c_1 \\ c_2 \\ c_3 - t \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}.$$

Możemy stąd wnioskować, że t jest wartością własną wtedy i tylko wtedy gdy pewna nietrywialna kombinacja liniowa wektorów $\begin{pmatrix} a_1 - t \\ a_2 \\ a_3 \end{pmatrix}$, $\begin{pmatrix} b_1 \\ b_2 - t \\ b_3 \end{pmatrix}$, $\begin{pmatrix} c_1 \\ c_2 \\ c_3 - t \end{pmatrix}$ daje wektor zerowy. To, jak wiemy, zachodzi wtedy i tylko wtedy gdy wektory te są liniowo zależne, tzn. gdy $\det \begin{pmatrix} a_1 - t & b_1 & c_1 \\ a_2 & b_2 - t & c_2 \\ a_3 & b_3 & c_3 - t \end{pmatrix} = 0$.

Wniosek 10.2.1. *Liczba t jest wartością własną przekształcenia liniowego T wtedy i tylko wtedy gdy* $\det \begin{pmatrix} a_1 - t & b_1 & c_1 \\ a_2 & b_2 - t & c_2 \\ a_3 & b_3 & c_3 - t \end{pmatrix} = 0$.

✔ **Przykład.** Dla przekształcenia liniowego zadanego macierzą $\begin{pmatrix} 1 & 2 & -2 \\ -3 & 2 & 1 \\ 0 & -1 & 2 \end{pmatrix}$, znajdźmy wszystkie jego wartości własne. Jak wiemy, wartość własna tej macierzy, oznaczmy ją t, musi spełniać $\det \begin{pmatrix} 1-t & 2 & -2 \\ -3 & 2-t & 1 \\ 0 & -1 & 2-t \end{pmatrix} = 0$. Korzystając z reguły Sarrusa, obliczmy wyznacznik tej macierzy. Otrzymujemy wówczas

$$(1-t)(2-t)(2-t) - 6 + (1-t) + 6(2-t) = 0,$$

po przekształceniu mamy

$$-t^3 + 5t^2 - 15t + 11 = 0.$$

Równanie to, po rozłożeniu na czynniki, przyjmuje postać

$$(t-1)(-t^2 + 4t - 11) = 0.$$

Jedyną liczbą spełniającą to równanie jest 1, gdyż trójmian kwadratowy $-t^2+4t-11$ ma ujemny wyróżnik więc nie ma rzeczywistych pierwiastków, stąd 1 jest jedyną wartością własną przekształcenia liniowego zadanego macierzą $\begin{pmatrix} 1 & 2 & -2 \\ -3 & 2 & 1 \\ 0 & -1 & 2 \end{pmatrix}$.

10.3 Równanie charakterystyczne.

Ściśle powiązane z rozważanymi wcześniej zagadnieniami, jest pojęcie równania charakterystycznego. Zdefiniujmy je.

Definicja 10.3.1 (równanie charakterystyczne). Równanie

$$\boxed{\det \begin{pmatrix} a_1 - t & b_1 & c_1 \\ a_2 & b_2 - t & c_2 \\ a_3 & b_3 & c_3 - t \end{pmatrix} = 0}$$

nazywane jest równaniem charakterystycznym przekształcenia T (macierzy $m(T)$).

W oparciu o powyższą definicję, Wniosek 10.2.1 możemy więc sformułować w następujący sposób.

Wniosek 10.3.2 (pierwiastki równania charakterystycznego). *Liczba t jest wartością własną przekształcenia T wtedy i tylko wtedy gdy jest ona pierwiastkiem równania charakterystycznego tego przekształcenia.*

Równanie charakterystyczne możemy zawsze przekształcić tak, aby przyjęło postać

$$-t^3 + u_2 t^2 + u_1 t^2 + u_0 = 0,$$

gdzie u_0, u_1, u_2 są stałymi współczynnikami zależnymi od wyrazów macierzy $m(T)$. Wyrażenie $-t^3 + u_2 t^2 + u_1 t^2 + u_0$ jest wielomianem zmiennej t, zwanym **wielomianem charakterystycznym.**

✍ **Ćwiczenie.** Jako ćwiczenie pozostawimy czytelnikowi wyznaczenie współczynników u_0, u_1, u_2 wielomianu charakterystycznego, w zależności od wyrazów macierzy $m(T)$.

✔ **Przykład.** Wyznaczmy wielomian charakterystyczny przekształcenia liniowego zadanego macierzą $\begin{pmatrix} 1 & 3 & -2 \\ 0 & -5 & 1 \\ 2 & 1 & 1 \end{pmatrix}$. Wielomian ten wyznaczymy przez przekształcenie równania charakterystycznego

$$\det \begin{pmatrix} 1-t & 3 & -2 \\ 0 & -5-t & 1 \\ 2 & 1 & 1-t \end{pmatrix} = 0.$$

Korzystając z metody Sarrusa, możemy to równanie zapisać w postaci

$$(1-t)^2(-5-t) + 6 + 4(-5-t) - (1-t) = 0.$$

Po przekształceniu otrzymujemy $-t^3 - 3t^2 + 4t - 20 = 0$. Tak więc wielomian charakterystyczny tego przekształcenia to

$$W(t) = -t^3 - 3t^2 + 4t - 20.$$

10.4 Liczba rozwiązań równania charakterystycznego.

Zastanówmy się, ile pierwiastków może mieć wielomian charakterystyczny, a tym samym, ile wartości własnych może mieć dane przekształcenie. Spróbujmy ustalić to w ogólnym przypadku:

Lemat 10.4.1. *Równanie charakterystyczne przekształcenia liniowego $T : R^3 \to R^3$ ma zawsze przynajmniej jeden pierwiastek.*

Dowód. Oznaczmy wielomian charakterystyczny przez $f(t)$, czyli

$$f(t) = -t^3 + u_2 t^2 + u_1 t^2 + u_0,$$

dla pewnych u_0, u_1, u_2 zależnych od wyrazów macierzy przekształcenia T. Zauważmy, że wtedy

$$\lim_{t \to -\infty} f(t) = +\infty, \quad \lim_{t \to +\infty} f(t) = -\infty.$$

Tak więc wielomian $f(t)$ przyjmuje wartości zarówno dodatnie, jak i ujemne. Jak wiemy, wielomian jest funkcją ciągłą, tak więc, z własności Darboux wiemy, że musi mieć miejsce zerowe. □

Jeśli t_0 jest miejscem zerowym, czyli pierwiastkiem wielomianu charakterystycznego, to ten wielomian możemy przedstawić w postaci

$$f(t) = (t - t_0)(-t^2 - dt + e),$$

dla pewnych d i e. Jeżeli wielomian f ma inne pierwiastki, to są to pierwiastki trójmianu $-t^2 - dt + e$. Jak wiemy, trójmian może mieć $0, 1$ lub 2 pierwiastki. Możemy zatem podać następujący wniosek.

Wniosek 10.4.2 (liczba wartości własnych przekształcenia). *Przekształcenie $T : R^3 \to R^3$ może mieć jedną, dwie lub trzy wartości własne.*

10.5 Przestrzenie własne dla wartości własnych.

Jak nietrudno się przekonać, danej wartości własnej może odpowiadać więcej niż jeden wektor własny, tak jest np. w geometrycznym przykładzie podanym na początku tego rozdziału. W tym podrozdziale spróbujemy ustalić postać wszystkich wektorów własnych dla danej wartości własnej. W związku z tym wprowadźmy następujące pojęcie:

Definicja 10.5.1 (przestrzeń własna). Przestrzenią własną dla wartości własnej t przekształcenia T nazywamy zbiór $E_t = \{X \in R^3 : T(X) = t \cdot X\}$, utworzony ze wszystkich wektorów własnych tej wartości własnej.

✔ **Przykład.**

(1) Rozważmy symetrię S względem płaszczyzny O_{xy}, czyli płaszczyzny wyznaczonej przez wersory $\begin{pmatrix}1\\0\\0\end{pmatrix}$, $\begin{pmatrix}0\\1\\0\end{pmatrix}$. Jak pokazaliśmy w geometrycznym przykładzie na początku tego rozdziału, wartościami własnymi tego przekształcenia są 1 i -1. Spróbujmy jednak wyznaczyć jeszcze raz te wartości korzystając z Wniosku 10.2.1. W podrozdziale 5.5 pokazaliśmy, że macierz tego przekształcenia to $\begin{pmatrix}1&0&0\\0&1&0\\0&0&-1\end{pmatrix}$. Wartościami własnymi tego przekształcenia będą zatem liczby t spełniające $\det\begin{pmatrix}1-t&0&0\\0&1-t&0\\0&0&-1-t\end{pmatrix} = 0$. Po przekształceniu równanie przyjmuje postać $(1-t)^2(-1-t) = 0$ tak więc $t = 1$ lub $t = -1$. Czyli rzeczywiście wartościami własnymi tego przekształcenia są 1 i -1. Przestrzenie wektorów własnych to odpowiednio:

$$E_1 = \{X \in R^3 : X \in O_{xy}\} = O_{xy},$$

bo tylko te wektory nie zmieniają się przez odbicie względem tej płaszczyzny,

$$E_{-1} = \{X \in R^3 : X \in L\} = L = O_z = \{(0,0,z) : z \in R\},$$

gdzie prosta L jest zbiorem wektorów równoległych do osi O_z, bo tylko dla nich spełniony jest warunek $S(X) = -X$.

(2) Dla jednokładności J_r danej wzorem $J_r(X) = r \cdot X$, wartością własną jest r, natomiast odpowiadającą jej przestrzenią własną jest $E_r = R^3$, bo każdy wektor z przestrzeni, zgodnie z definicją jednokładności, spełnia $J_r(X) = r \cdot X$.

✍ **Ćwiczenie.** Jako ćwiczenie pozostawimy czytelnikowi sprawdzenie, że $-1, 1$ to jedyne wartości własne symetrii względem płaszczyzny O_{XY}.

Poniższe twierdzenie określa, jaką postać może przyjąć przestrzeń wektorów własnych, dla dowolnego przekształcenia.

Twierdzenie 10.5.2 (postać przestrzeni własnych). *Każda przestrzeń własna przekształcenia R^3 jest jednym z następujących zbiorów:*

(1) prostą zawierającą $\begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$,

(2) płaszczyzną zawierającą $\begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$,

(3) całą przestrzenią R^3.

Dowód. Niech E_t będzie przestrzenią własną dla wartości własnej t przekształcenia T. Z definicji wartości własnej wiemy, że istnieje pewien niezerowy wektor X, odpowiadający tej wartości własnej, a więc należący do E_t. Oznaczmy przez L_X prostą wzdłuż tego wektora. Prosta ta składa się z wektorów postaci rX dla $r \in R$. Zauważmy, że $T(rX) = rT(X) = rtX = t(rX)$, zatem rX jest wektorem własnym dla t. Otrzymaliśmy więc $L_X \subset E_t$. Jeśli $E_t = L_X$, to mamy podpunkt (1). Jeśli nie, to E_t zawiera jakiś wektor Y nie należący do L_X. Zauważmy, że ten wektor Y musi być wtedy niewspółliniowy z X. Oznaczmy przez Π_{XY} płaszczyznę wyznaczoną przez wektory X i Y, czyli płaszczyznę składającą się z wektorów postaci $rX + sY$, dla $r, s \in R^3$. Zauważmy, że

$$T(rX + sY) = rT(X) + sT(Y) = rtX + stY = t(rX + sY),$$

tak więc dowolny wektor $rX + sY$ jest wektorem własnym dla wartości własnej t, zatem $\Pi_{XY} \subset E_t$. Jeśli $\Pi_{XY} = E_t$, to otrzymujemy podpunkt (2). Jeśli nie, to E_t zawiera wektor Z nie leżący w płaszczyźnie Π_{XY}. Wtedy X, Y, Z są wektorami liniowo niezależnymi. Każdy wektor $U \in R^3$ możemy więc zapisać jako kombinację liniową X, Y, Z tzn. $U = rX + sY + wZ$, dla pewnych r, s, w. Zauważmy, że

$$\begin{aligned} T(U) = T(rX + sY + wZ) = rT(X) + sT(Y) + wT(Z) = \\ = rtX + stY + wtZ = t(rX + sY + wZ) = tU. \end{aligned}$$

Stąd otrzymujemy, że każdy wektor $U \in R^3$ należy do E_t, zatem $E_t = R^3$, co daje podpunkt (3). $\square$

Rozdział 11

Izometrie liniowe i przekształcenia ortogonalne w R^3

W tym rozdziale zajmiemy się pewnymi rodzajami przekształceń liniowych - izometriami liniowymi i przekształceniami ortogonalnymi. Na początek zdefiniujmy, czym jest izometria liniowa, podamy jedną z jej najbardziej charakterystycznych własności. W kolejnym podrozdziale zdefiniujemy pojęcie przekształcenia ortogonalnego, zbadamy jak zachowuje się kąt między wektorami pod wpływem działania takiego przekształcenia. Poznamy również związek łączący izometrie liniowe z przekształceniami ortogonalnymi. W ostatnim podrozdziale zbadamy jaką postać mogą przyjmować izometrie liniowe.

11.1 Podstawowe definicje.

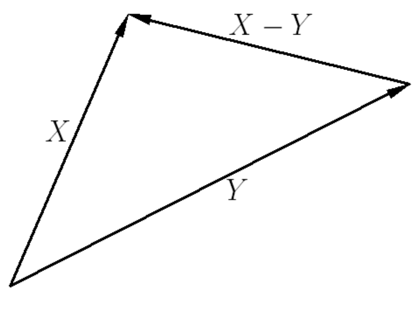
Definicja 11.1.1 (**izometria liniowa**). Izometrią liniową R^3 nazywamy takie przekształcenie liniowe $T : R^3 \to R^3$, które dla każdego wektora $X \in R^3$ spełnia równość $|T(X)| = |X|$, czyli takie, które zachowuje długość wektora.

Prostym wnioskiem z powyższej definicji jest następujący lemat.

Lemat 11.1.2 (**izometrie a odległość między punktami**). *Każda izometria liniowa zachowuje odległości pomiędzy punktami w przestrzeni.*

Dowód. Jak wiemy punkty w przestrzeni możemy reprezentować przez ich wektory wodzące. Zauważmy, że odległość między punktami X, Y to długość wektora $X - Y$, co zapisujemy

$$d(X, Y) = |X - Y|$$



Niech X', Y' będą obrazami wektorów X, Y przez izometrię liniową T. Zauważmy, że wówczas, korzystając z liniowości przekształcenia T mamy

$$d(X', Y') = |X' - Y'| = |T(X) - T(Y)| = |T(X - Y)| = |X - Y| = d(X - Y).$$

Zatem izometria liniowa rzeczywiście nie zmienia odległości między punktami w przestrzeni. □

11.2 Przekształcenia ortogonalne.

Wśród izometrii liniowych wyróżniamy pewien rodzaj przekształceń, zwanych przekształceniami ortogonalnymi, których własnościami zajmiemy się w tym podrozdziale.

Definicja 11.2.1 (przekształcenie ortogonalne). Przekształcenie liniowe T jest przekształceniem ortogonalnym, jeśli dla dowolnych wektorów $X, Y \in R^3$ zachodzi $T(X) \circ T(Y) = X \circ Y$, tzn. gdy T zachowuje iloczyn skalarny.

Zauważmy, że przekształcenia ortogonalne rzeczywiście są izometriami liniowymi. Mówi o tym następujący fakt.

Fakt 11.2.2. *Przekształcenia ortogonalne zachowują długości wektorów, czyli są izometriami.*

Dowód. Dla przekształcenia ortogonalnego T mamy

$$|T(X)| = \sqrt{T(X) \circ T(X)} = \sqrt{X \circ X} = |X|.$$

□

Zajmiemy się teraz innymi własnościami przekształceń ortogonalnych.

Fakt 11.2.3 (przekształcenia ortogonalne a kąt między wektorami). *Przekształcenia ortogonalne zachowują kąty między wektorami, tzn. dla dowolnych wektorów $X, Y \in R^3$ mamy $\measuredangle T(X)T(Y) = \measuredangle XY$.*

Dowód. Dla dowolnych wektorów X, Y z przestrzeni R^3 i przekształcenia ortogonalnego T, korzystając z definicji przekształcenia ortogonalnego i faktu (11.2.2) możemy stwierdzić, że

$$\cos \measuredangle T(X)T(Y) = \frac{T(X) \circ T(Y)}{|T(X)| \cdot |T(Y)|} = \frac{X \circ Y}{|X| \cdot |Y|} = \cos \measuredangle XY.$$

□

Wniosek 11.2.4 (obraz wektorów prostopadłych). *Obrazami wektorów prostopadłych przez przekształcenie ortogonalne, są wektory prostopadłe.*

Dowód. Jeśli wektory X, Y są prostopadłe, czyli $X \circ Y = 0$, to również

$$T(X) \circ T(Y) = X \circ Y = 0,$$

tzn. obrazy tych wektorów przez przekształcenie ortogonalne również są prostopadłe. □

Wiemy, że przekształcenia ortogonalne to pewien rodzaj izometrii liniowych. Okazuje się, że każda izometria liniowa jest przekształceniem ortogonalnym. Mówi o tym poniższy fakt.

Fakt 11.2.5. *Każda izometria liniowa jest przekształceniem ortogonalnym.*

Dowód. Niech X, Y będą dowolnymi wektorami z przestrzeni, a T izometrią liniową. Mamy więc

$$|T(X) - T(Y)| = |T(X - Y)| = |X - Y|,$$

a więc również

$$|T(X) - T(Y)|^2 = |T(X - Y)|^2 = |X - Y|^2.$$

Ale wiemy, że

$$|T(X) - T(Y)|^2 = (T(X) - T(Y)) \circ (T(X) - T(Y)) =$$

$$= |T(X)|^2 + |T(Y)|^2 - 2 \cdot (T(X) \circ T(Y)),$$

oraz

$$|X - Y|^2 = (X - Y) \circ (X - Y) = |X|^2 + |Y|^2 - 2 \cdot (X \circ Y).$$

Z dwóch pierwszych równości mamy

$$|X - Y|^2 = |T(X)|^2 + |T(Y)|^2 - 2 \cdot (T(X) \circ T(Y)).$$

Podstawiając do trzeciej równości mamy

$$|T(X)|^2 + |T(Y)|^2 - 2 \cdot (T(X) \circ T(Y)) = |X|^2 + |Y|^2 - 2 \cdot (X \circ Y).$$

Przekształcając tę równość, przy wykorzystaniu faktu, że $|X| = |T(X)|$, $|Y| = |T(Y)|$, otrzymujemy

$$-2 \cdot (T(X) \circ T(Y)) = -2 \cdot (X \circ Y).$$

Otrzymujemy zatem

$$T(X) \circ T(Y) = X \circ Y,$$

czyli, zgodnie z Definicją 11.2.1, izometria T jest przekształceniem ortogonalnym.

□

Wiemy, że przekształcenia ortogonalne są izometriami liniowymi. Powyżej pokazaliśmy również, że izometrie są przekształceniami ortogonalnymi. Stąd mamy następujący wniosek.

Wniosek 11.2.6. *Zbiory izometrii liniowych i przekształceń ortogonalnych są sobie równe.*

11.3 Opis izometrii liniowych w R^3.

W tym podrozdziale sklasyfikujemy izometrie liniowe. Pomocne okaże się w tym określenie postaci wartości własnych tych przekształceń.

Fakt 11.3.1 (wartości własne izometrii). *Jeśli T jest izometrią liniową, to jej każda wartość własna jest równa* 1 *lub* -1.

Dowód. Niech t będzie wartością własną dla izometrii T. Istnieje więc niezerowy wektor X taki, że $T(X) = t \cdot X$. Dla tego wektora mamy

$$|X| = |T(X)| = |t \cdot X| = |t| \cdot |X|,$$

tak więc $|t| = 1$, czyli $t = 1$ lub $t = -1$. □

Fakt 11.3.2 (wyznacznik macierzy izometrii liniowej). *Jeśli T jest izometrią liniową zadaną macierzą m, to*

$$\det(m) = 1 \quad \textit{lub} \quad \det(m) = -1.$$

Dowód. Wiadomo, że macierz przekształcenia utworzona jest z obrazów wersorów przez przekształcenie T, jako kolumn, tak więc

$$m = (T(e_1) \;\; T(e_2) \;\; T(e_3)).$$

Z interpretacji wyznacznika za pomocą objętości równoległościanu wiemy, że

$$|\det(m)| = V(R_{T(e_1),T(e_2),T(e_3)}).$$

Zauważmy, że $|T(e_i)| = |e_i| = 1$, oraz $T(e_i) \circ T(e_j) = e_i \circ e_j = 0$ dla $i \neq j$. Stąd możemy wnioskować, że wektory $T(e_1), T(e_2), T(e_3)$ mają długość 1 i są parami prostopadłe, a więc równoległościan $R_{T(e_1),T(e_2),T(e_3)}$ jest kostką o krawędzi długości 1. Zatem

$$|\det(m)| = V(R_{T(e_1),T(e_2),T(e_3)}) = 1.$$

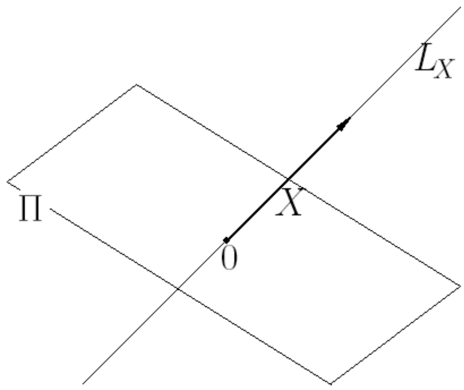
□

Spróbujmy opisać w sposób geometryczny wszystkie izometrie liniowe przestrzeni R^3. Zacznijmy od tych izometrii liniowych, których wyznacznik wynosi 1.

Fakt 11.3.3. *Jeśli izometria liniowa T ma wyznacznik* $\det(m(T)) = 1$, *to T jest obrotem wokół pewnej prostej L, zawierającej* $(0, 0, 0)$.

Dowód. Przekształcenie T ma przynajmniej jedną wartość własną i, jak wiemy z Faktu 11.3.1, jest ona równa 1 lub -1. Rozważmy te przypadki.

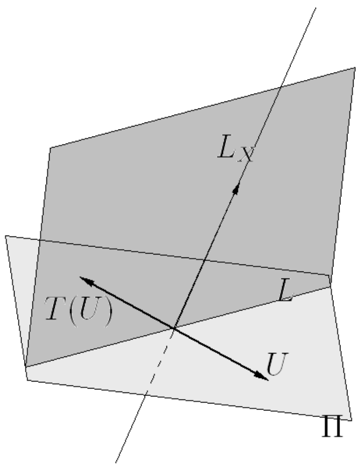
(1) Niech $t = 1$ będzie wartością własną, zaś X odpowiadającym jej wektorem własnym. Oznaczmy przez L_X prostą wzdłuż wektora X, zaś Π niech będzie płaszczyzną prostopadłą do X, przechodzącą przez $(0, 0, 0)$.



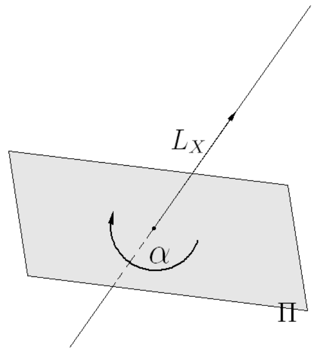
Dla dowolnego wektora Y z płaszczyzny Π, jego obraz przez przekształcenie T również należy do tej płaszczyzny, zauważmy bowiem, że płaszczyzna Π składa się z wektorów prostopadłych do wektora normalnego X tej płaszczyzny. Dla dowolnego wektora Y z płaszczyzny Π określmy przekształcenie

$$T_0 : \Pi \to \Pi : \quad T_0(Y) = T(Y).$$

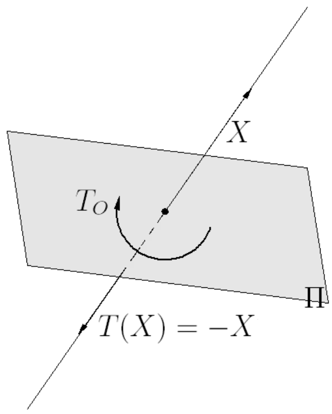
Zauważmy, że tak określone T_0 jest izometrią płaszczyzny Π, ponieważ $|T_0(Y)| = |T(Y)| = |Y|$. Pokażemy teraz, że T_0 jest obrotem płaszczyzny. Gdyby tak nie było, to musiałoby być odbiciem względem pewnej prostej $L \subset \Pi$.



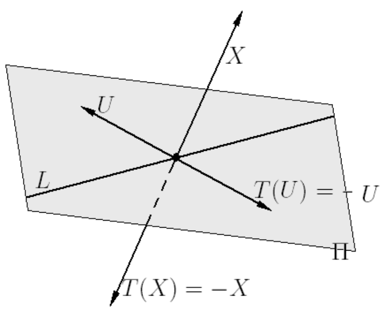
Ale wówczas całe przekształcenie T byłoby, jak widać na rysunku, odbiciem względem płaszczyzny zawierającej proste L i L_X. Jak wiemy symetria zmienia orientację, więc jej wyznacznik jest ujemny, a my założyliśmy, że $\det(m(T)) = 1$. Tak więc T_0 jest obrotem wokół prostej L_X o kąt α, więc T również jest obrotem o taki sam kąt α i wokół tej samej prostej.



(2) Niech $t = -1$ będzie wartością własną, zaś X odpowiadającym jej wektorem własnym. Podobnie jak poprzednio, L_X to prosta wzdłuż wektora X, zaś Π to płaszczyzna prostopadła do X przechodzącą przez $(0,0,0)$. Przekształcenie $T_0 : \Pi \to \Pi$, tak jak poprzednio, jest przekształceniem, które dla dowolnego $Y \in \Pi$ określamy poprzez $T_0(Y) = T(Y)$.
Pokażemy, że T_0 jest symetrią osiową. Gdyby T_0 nie było symetrią, to przekształcenie T musiałoby być złożeniem symetrii względem Π i obrotu wokół L_X.



Symetria zmienia orientację, obrót zachowuje, więc złożenie tych przekształceń zmieniałoby orientację. Wiemy jednak, że $\det(m(T)) = 1$, więc T_0 musi być odbiciem względem prostej L. Przekształcenie T zachowuje prostą L. Zauważmy również, że obrazem wektorów z płaszczyzny Π prostopadłych do prostej L są wektory do nich przeciwne.



Jedyną izometrią spełniającą te warunki, jest obrót wokół prostej L o kąt 180^0.

$\square$

Zanim rozważymy przypadek, gdy $\det(T) = -1$, musimy zdefiniować pewien rodzaj przekształceń.

Definicja 11.3.4 (**odbicie z obrotem**). Odbicie z obrotem jest to izometria przestrzeni, która jest złożeniem obrotu wokół pewnej prostej L z odbiciem względem pewnej płaszczyzny Π prostopadłej do prostej L.

Fakt 11.3.5. *Jeśli izometria liniowa T ma wyznacznik* $\det(T) = -1$, *to T jest albo symetrią względem pewnej płaszczyzny* Π *zawierającej* $(0,0,0)$, *albo odbiciem z obrotem. W tym drugim przypadku T jest złożeniem symetrii względem pewnej płaszczyzny* Π *zawierającej* $(0,0,0)$ *oraz obrotu wokół prostej L przechodzącej przez* $(0,0,0)$ *i prostopadłej do* Π.

Podsumowując przeprowadzony w tym podrozdziale opis poszczególnych rodzajów izometrii liniowych przestrzeni R^3 możemy sformułować następujący:

Wniosek 11.3.6. ***(rodzaje izometrii liniowych)*** *Każda izometria liniowa przestrzeni* R^3 *jest albo obrotem, albo odbiciem, albo odbiciem z obrotem.*

Rozdział 12

Macierz ortogonalna i macierz transponowana

Rozdział ten poświęcimy na zbadanie pewnych klas macierzy. Na początek wprowadzimy pojęcie macierzy ortogonalnej, które jest ściśle związane z poznanymi w poprzednim rozdziale przekształceniami ortogonalnymi (izometriami liniowymi). W drugim podrozdziale dowiemy się, czym jest macierz transponowana do danej, nauczymy się wyznaczać ją na wiele sposobów. Poznamy również związek macierzy transponowanych z macierzami ortogonalnymi. Na koniec podamy kilka własności macierzy transponowanej.

12.1 Macierz ortogonalna.

Definicja 12.1.1 (macierz ortogonalna). Macierz 3×3 jest ortogonalna, jeśli zadane przez tę macierz przekształcenie liniowe $T : R^3 \to R^3$ jest ortogonalne lub, równoważnie, jeśli jest izometrią liniową.

Sprawdzanie ortogonalności macierzy wprost z powyższej definicji, na ogół bywa bardzo kłopotliwe, musimy bowiem odwołać się do pojęcia ortogonalności przekształcenia liniowego. Poniżej podamy bardzo użyteczne kryterium, które pozwoli na łatwe sprawdzenie zarówno ortogonalności macierzy, jak również ortogonalności odpowiadającego mu przekształcenia.

Lemat 12.1.2 (kolumny macierzy ortogonalnej). *Macierz $m = (K_1,\ K_2,\ K_3)$ jest ortogonalna wtedy i tylko wtedy gdy jej kolumny są parami prostopadłymi wektorami jednostkowymi, czyli gdy*

$$|K_1| = |K_2| = |K_3| = 1 \quad oraz \quad K_1 \circ K_2 = K_1 \circ K_3 = K_2 \circ K_3 = 0.$$

Dowód. Załóżmy, że macierz m jest ortogonalna, co jest równoważne temu, że przekształcenie T, zadane tą macierzą, jest ortogonalne. Wówczas $|K_i| = |T(e_i)| = |e_i| = 1$, dla $i = 1, 2, 3$. Mamy również $K_i \circ K_j = T(e_i) \circ T(e_j) = e_i \circ e_j = 0$, gdy $i \neq j$. Pokazaliśmy więc implikację w jedną stronę.
Teraz pokażemy, że jeśli kolumny macierzy są wzajemnie prostopadłymi wektorami

jednostkowymi, to przekształcenie T jest ortogonalne. Niech X, Y będą dowolnymi wektorami z przestrzeni. Zapisujemy je w postaci

$$X = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = x_1e_1 + x_2e_2 + x_3e_3,$$

$$Y = \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} = y_1e_1 + y_2e_2 + y_3e_3.$$

Z jednorodności i addytywności przekształcenia liniowego T mamy

$$T(X) = T(\sum_{i=1}^{3} x_ie_i) = \sum_{i=1}^{3} x_iT(e_i) = \sum_{i=1}^{3} x_iK_i,$$

$$T(Y) = T(\sum_{i=1}^{3} y_ie_i) = \sum_{i=1}^{3} y_iT(e_i) = \sum_{i=1}^{3} y_iK_i.$$

Pokażemy teraz, że przekształcenie to zachowuje iloczyn skalarny

$$T(X) \circ T(Y) = (\sum_{i=1}^{3} x_iK_i) \circ (\sum_{i=1}^{3} y_iK_i) = \sum_{i,j=1}^{3} x_iy_j(K_i \circ K_j),$$

założyliśmy jednak, że $K_i \circ K_j = 0$ dla $i \neq j$, tak więc

$$\sum_{i,j=1}^{3} x_iy_j(K_i \circ K_j) = \sum_{i=1}^{3} x_iy_i(K_i \circ K_j) = \sum_{i=1}^{3} x_iy_i = X \circ Y.$$

Tak więc rzeczywiście, dla przekształcenia T mamy $T(X) \circ T(Y) = X \circ Y$, więc przekształcenie T jest ortogonalne. $\square$

✔ **Przykład.** Macierz $\begin{pmatrix} \frac{4}{5} & 0 & -\frac{3}{5} \\ \frac{3}{5} & 0 & \frac{4}{5} \\ 0 & 1 & 0 \end{pmatrix}$ jest ortogonalna. Sprawdzenie ortogonalności tej macierzy pozostawimy jako ćwiczenie dla czytelnika.

12.2 Macierz transponowana.

Dowolnej macierzy m możemy jednoznacznie przyporządkować inną macierz m^T, zwaną macierzą transponowaną. Pojęcie to pojawiło się już w rozdziale poświęconym wyznacznikom, tutaj trochę szerzej zajmiemy się teorią z nim związaną. Zacznijmy od definicji.

Definicja 12.2.1 (macierz transponowana). Macierzą transponowaną do danej macierzy $m = \begin{pmatrix} a & b & c \\ d & e & f \\ g & h & i \end{pmatrix}$ nazywamy macierz $m^T = \begin{pmatrix} a & d & g \\ b & e & h \\ c & f & i \end{pmatrix}$.

➢ Znajdowanie macierzy transponowanej

Podamy teraz kilka sposobów na utworzenie macierzy transponowanej do danej:

(1) Zamieniamy ze sobą wyrazy macierzy symetryczne względem przekątnej.

(2) Dla każdego $i, j = 1, 2, 3$ wyraz z $i-$tego wiersza i $j-$tej kolumny umieszczamy w $j-$tym wierszu i $i-$tej kolumnie.

(3) Kolumny macierzy m stają się wierszami macierzy transponowanej m^T.

(4) Wiersze macierzy m stają się kolumnami macierzy m^T.

✍ **Ćwiczenie.** Proponujemy czytelnikowi wypróbowanie każdego z tych czterech sposobów tworzenia macierzy transponowanej do danej, i przekonanie się, że w każdym przypadku otrzymujemy tą samą macierz zgodną z Definicją 12.2.1.

✔ **Przykład.** Macierzą transponowaną do macierzy $\begin{pmatrix} 1 & 0 & 6 \\ -8 & 1 & 5 \\ 3 & -3 & 0 \end{pmatrix}$ jest macierz $\begin{pmatrix} 1 & -8 & 3 \\ 0 & 1 & -3 \\ 6 & 5 & 0 \end{pmatrix}$.

Podamy teraz twierdzenie wiążące pojęcia macierzy transponowanej i macierzy ortogonalnej.

Twierdzenie 12.2.2 (macierz transponowana a ortogonalność macierzy). *Macierz m jest ortogonalna wtedy i tylko wtedy gdy $m \cdot m^T = m^T \cdot m = I$, gdzie I jest macierzą jednostkową rozmiaru 3×3.*

Dowód. Ponieważ macierz ortogonalna jest odwracalna, co pozostawimy do uzasadnienia czytelnikowi, wystarczy uzasadnić, że $m^T \cdot m = I$, bo to oznacza, że macierz m^T jest macierzą odwrotną do m. Jeśli $m = (K_1,\ K_2,\ K_3)$, to $m^T = \begin{pmatrix} K_1^T \\ K_2^T \\ K_3^T \end{pmatrix}$, gdzie $K_i = \begin{pmatrix} a_i \\ b_i \\ c_i \end{pmatrix}$, $K_i^T = (a_i,\ b_i,\ c_i)$. Dla tak określonych macierzy mamy

$$m^T \cdot m = \begin{pmatrix} K_1^T \\ K_2^T \\ K_3^T \end{pmatrix} \cdot (K_1 \quad K_2 \quad K_3) = \begin{pmatrix} K_1 \circ K_1 & K_1 \circ K_2 & K_1 \circ K_3 \\ K_2 \circ K_1 & K_2 \circ K_2 & K_2 \circ K_3 \\ K_3 \circ K_1 & K_3 \circ K_2 & K_3 \circ K_3 \end{pmatrix}.$$

Chcemy, aby $m^T \cdot m = I$, tym samym aby

$$\begin{pmatrix} K_1 \circ K_1 & K_1 \circ K_2 & K_1 \circ K_3 \\ K_2 \circ K_1 & K_2 \circ K_2 & K_2 \circ K_3 \\ K_3 \circ K_1 & K_3 \circ K_2 & K_3 \circ K_3 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Zauważmy, że wyrazy znajdujące się na przekątnej głównej macierzy $m^T \cdot m$, czyli wyrazy $K_i \circ K_i$ są równe 1, oraz wyrazy $K_i \circ K_j$ dla $i \neq j$ są równe 0. Zgodnie z Lematem 12.1.2 jest to równoważne temu, że macierz m jest macierzą ortogonalną.

□

Uwaga 12.2.3. *Związek* $m^T \cdot m = m \cdot m^T = I$ *oznacza, że macierz transponowana* m^T *do macierzy ortogonalnej jest jej macierzą odwrotną.*

Z powyższych rozważań możemy wyprowadzić wniosek, który okaże się być bardzo prostym sposobem na sprawdzenie ortogonalności macierzy.

Wniosek 12.2.4 (macierz transponowana a macierz odwrotna). *Macierz* m *jest ortogonalna wtedy i tylko wtedy gdy jest odwracalna oraz* $m^{-1} = m^T$.

12.3 Własności operacji transponowania macierzy.

Podamy teraz kilka najbardziej charakterystycznych własności operacji transponowania macierzy. Zacznijmy od bardzo prostego faktu.

Fakt 12.3.1.

$$(m^T)^T = m$$

✍ **Ćwiczenie.** Dowód tego faktu, wynikający wprost z definicji macierzy transponowanej, pozostawimy jako ćwiczenie dla czytelnika.

Fakt 12.3.2 (transponowanie iloczynu macierzy). *Dla dowolnych macierzy* m, n *rozmiaru* 3×3 *mamy*

$$\boxed{m^T \cdot n^T = (n \cdot m)^T.}$$

Uwaga 12.3.3. *W powyższym wzorze warto zwrócić uwagę na różną kolejność pojawiania się macierzy* m *i* n *po obu stronach równości.*

Dowód. Ze sposobu określenia iloczynu macierzy wiemy, że do wyznaczenia wyrazów macierzy będącej iloczynem macierzy m i n najlepiej będzie zapisać macierz n przy użyciu jej wierszy a macierz m przy użyciu kolumn, tak więc $n = \begin{pmatrix} W_1 \\ W_2 \\ W_3 \end{pmatrix}$, $m = (K_1,\ K_2,\ K_3)$. Iloczyn tych macierzy przedstawia się wówczas wzorem

$$n \cdot m = \begin{pmatrix} W_1 \circ K_1 & W_1 \circ K_2 & W_1 \circ K_3 \\ W_2 \circ K_1 & W_2 \circ K_2 & W_2 \circ K_3 \\ W_3 \circ K_1 & W_3 \circ K_2 & W_3 \circ K_3 \end{pmatrix},$$

tak więc macierz transponowana do tego iloczynu to

$$(n \cdot m)^T = \begin{pmatrix} W_1 \circ K_1 & W_2 \circ K_1 & W_3 \circ K_1 \\ W_1 \circ K_2 & W_2 \circ K_2 & W_3 \circ K_2 \\ W_1 \circ K_3 & W_2 \circ K_3 & W_3 \circ K_3 \end{pmatrix}.$$

Zgodnie z określeniem macierzy transponowanej mamy

$$n^T = (W_1^T,\ W_2^T,\ W_3^T), \quad m^T = \begin{pmatrix} K_1^T \\ K_2^T \\ K_3^T \end{pmatrix},$$

zatem ich iloczyn $m^T \cdot n^T$ przedstawia się wzorem

$$m^T \cdot n^T = \begin{pmatrix} K_1^T \\ K_2^T \\ K_3^T \end{pmatrix} (W_1^T,\ W_2^T,\ W_3^T) = \begin{pmatrix} W_1 \circ K_1 & W_2 \circ K_1 & W_3 \circ K_1 \\ W_1 \circ K_2 & W_2 \circ K_2 & W_3 \circ K_2 \\ W_1 \circ K_3 & W_2 \circ K_3 & W_3 \circ K_3 \end{pmatrix}.$$

Jak widać spełniony jest zatem warunek

$$m^T \cdot n^T = (n \cdot m)^T.$$

□

Poniższy fakt pokaże, jak przy użyciu macierzy odwrotnej do danej macierzy m znaleźć macierz odwrotną do macierzy transponowanej. W związku z tym, że liczenie macierzy odwrotnej bywa bardzo uciążliwe jesteśmy w stanie sobie wyobrazić jak bardzo użyteczna może być taka umiejętność.

Fakt 12.3.4 (macierz odwrotna do macierzy transponowanej). *Jeśli macierz m jest odwracalna, to*

$$\boxed{(m^T)^{-1} = (m^{-1})^T.}$$

Dowód. Sprawdźmy, czy $(m^{-1})^T$ jest rzeczywiście macierzą odwrotną do m^T. W tym celu sprawdźmy, czy

$$m^T \cdot (m^{-1})^T = (m^{-1})^T \cdot m^T = I.$$

Korzystając z Faktu 12.3.2, oraz z tego, że macierzą transponowaną do macierzy identycznościowej jest ona sama, otrzymujemy

$$m^T \cdot (m^{-1})^T = (m^{-1} \cdot m)^T = I^T = I$$

oraz

$$(m^{-1})^T \cdot m^T = (m \cdot m^{-1})^T = I^T = I.$$

□

W poprzednim podrozdziale poznaliśmy pojęcie macierzy ortogonalnej, w tym - macierzy transponowanej do danej. Spróbujmy połączyć te dwa pojęcia. Zastanówmy się jak zachowuje się macierz transponowana do macierzy ortogonalnej.

Wniosek 12.3.5. *Macierz transponowana do macierzy ortogonalnej sama jest ortogonalna.*

Dowód. Zgodnie z Wnioskiem 12.2.4 wystarczy pokazać, że $(m^T)^T = (m^T)^{-1}$. Jak wiemy z Faktu 12.3.4, $(m^T)^{-1} = (m^{-1})^T$, ale macierz m jest ortogonalna, więc $m^T = m^{-1}$, a stąd otrzymujemy, że $(m^T)^{-1} = (m^{-1})^T = (m^T)^T$. Tak więc macierz m^T również jest ortogonalna. □

Rozdział 13

Macierz symetryczna

Ważnym pojęciem, którym zajmiemy się w tym rozdziale, jest pojęcie macierzy symetrycznej. Ten rodzaj macierzy, jak również przekształcenia nimi zadane, mają bardzo charakterystyczne własności z których część tutaj podamy, zajmiemy się głównie wartościami i wektorami własnymi przekształceń zadanych takimi macierzami. W ostatnim podrozdziale zajmiemy się zagadnieniem diagonalizacji macierzy, które okaże się być bardzo użyteczne w dalszej części skryptu.

13.1 Podstawowe definicje.

Definicja 13.1.1 (macierz symetryczna). Macierzą symetryczną rozmiaru 3×3 nazywamy każdą macierz postaci $\begin{pmatrix} a & b & c \\ b & d & e \\ c & e & f \end{pmatrix}$, czyli macierz, w której wyrazy symetryczne względem przekątnej głównej są sobie równe.

Podamy teraz inną, równoważną definicję macierzy symetrycznej.

Definicja 13.1.2. Macierz m jest symetryczna wtedy i tylko wtedy gdy $m = m^T$.

W kolejnym podrozdziale będziemy badać, jak zachowują się wektory i wartości własne przekształceń zadanych macierzami symetrycznymi, będzie w tym pomocny poniższy Lemat, który będzie bardzo użyteczny w dalszych dowodach.

Lemat 13.1.3. *Jeśli $M : R^3 \to R^3$ jest przekształceniem liniowym zadanym macierzą symetryczną m, to dla dowolnych wektorów $X, Y \in R^3$ mamy*

$$M(X) \circ Y = X \circ M(Y).$$

Dowód. Dla dowolnych wektorów $X = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}, Y = \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix}$ oraz macierzy $m = \begin{pmatrix} a & b & c \\ b & d & e \\ c & e & f \end{pmatrix}$ mamy

$$M(X) \circ Y = (m \cdot X) \circ Y = \left[\begin{pmatrix} a & b & c \\ b & d & e \\ c & e & f \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \right] \circ \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} =$$

$$= \begin{pmatrix} ax_1 + bx_2 + cx_3 \\ bx_1 + dx_2 + ex_3 \\ cx_1 + ex_2 + fx_3 \end{pmatrix} \circ \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} =$$

$$= ax_1y_1 + bx_2y_1 + cx_3y_1 + bx_1y_2 + dx_2y_2 + ex_3y_2 + cx_1y_3 + ex_2y_3 + fx_3y_3 =$$

$$= x_1(ay_1 + by_2 + cy_3) + x_2(by_1 + dy_2 + ey_3) + x_3(cy_1 + ey_2 + fy_3) =$$

$$= \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \circ \begin{pmatrix} ay_1 + by_2 + cy_3 \\ by_1 + dy_2 + ey_3 \\ cy_1 + ey_2 + fy_3 \end{pmatrix} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \circ \left(\begin{pmatrix} a & b & c \\ b & d & e \\ c & e & f \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} \right) =$$

$$= X \circ (m \cdot Y) = X \circ M(Y).$$

□

13.2 Wartości i wektory własne przekształceń zadanych macierzami symetrycznymi.

Wartości i wektory własne przekształceń zadanych macierzami symetrycznymi mają wiele charakterystycznych własności, z których najważniejszą dla nas będzie ta podana w poniższym twierdzeniu.

Twierdzenie 13.2.1 (prostopadłość wektorów własnych). *Dla przekształcenia $M : R^3 \to R^3$ zadanego macierzą symetryczną, jeśli X_1, X_2 są wektorami własnymi przekształcenia M dla różnych wartości własnych t_1, t_2, to $X_1 \circ X_2 = 0$, czyli wektory te są prostopadłe.*

Dowód. Zauważmy, że

$$M(X_1) \circ X_2 = (t_1X_1) \circ X_2 = t_1(X_1 \circ X_2),$$

$$X_1 \circ M(X_2) = X_1 \circ (t_2X_2) = t_2(X_1 \circ X_2).$$

Z symetryczności macierzy przekształcenia M, na podstawie równości z Lematu 13.1.3 mamy więc

$$t_1(X_1 \circ X_2) = t_2(X_1 \circ X_2).$$

Ponieważ $t_1 \neq t_2$, więc aby równość była spełniona, musi zachodzić $X_1 \circ X_2 = 0$. □

Okazuje się, że dla przekształcenia zadanego macierzą symetryczną, możemy znaleźć trójkę wektorów własnych, o bardzo charakterystycznych własnościach. Mówi o tym następujące twierdzenie.

Twierdzenie 13.2.2. *Dla dowolnego przekształcenia liniowego $M : R^3 \to R^3$ zadanego macierzą symetryczną istnieją trzy prostopadłe jednostkowe wektory $X_1, X_2, X_3 \in R^3$, będące wektorami własnymi macierzy przekształcenia M.*

Dowód. Przekształcenie M ma przynajmniej jedną wartość własną t_1. Niech v_1 będzie niezerowym wektorem własnym dla tej wartości własnej. Wówczas również $X_1 = \frac{1}{|v_1|} \cdot v_1$ jest wektorem własnym dla t_1 i X_1 ma długość 1. Spróbujmy teraz znaleźć wektory X_2, X_3. Niech Π oznacza płaszczyznę prostopadłą do wektora X_1, zawierającą $(0, 0, 0)$.
Zanim przejdziemy do dalszej części dowodu, podamy następujący fakt.

Fakt 13.2.3. *Jeśli wektor X należy do płaszczyzny Π, to również $M(X)$ należy do płaszczyzny Π.*

Dowód (faktu). Dla każdego wektora X należącego do Π mamy $X \circ X_1 = 0$. Sprawdźmy, czy również $M(X) \circ X_1 = 0$. Otrzymujemy:

$$M(X) \circ X_1 = X \circ M(X_1) = X \circ t_1 X_1 = t_1(X \circ X_1) = t_1 \cdot 0 = 0,$$

więc wektor $M(X)$ jest prostopadły do wektora X_1, co oznacza, że $M(X)$ należy do płaszczyzny Π. $\square$

Wróćmy teraz do dowodu twierdzenia. Określmy przekształcenie $M_0 : \Pi \to \Pi$ jako $M_0(X) = M(X)$ dla $X \in \Pi$. Ponieważ M_0, tak jak przekształcenie M, jest addytywne i jednorodne, więc M_0 jest przekształceniem liniowym płaszczyzny Π. Wprowadźmy na płaszczyźnie Π układ współrzędnych $O_{x_1 x_2}$ wyznaczony przez jednostkowe prostopadłe wektory F_1, F_2 jako wersory. Wektor X ma współrzędne $\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$, czyli $X = x_1 F_1 + x_2 F_2$. Niech $m_0 = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ będzie macierzą przekształcenia M_0 w układzie $O_{x_1 x_2}$. Oznacza to, że

(1) $M_0(F_1) = aF_1 + cF_2$

(2) $M_0(F_2) = bF_1 + dF_2$

Pokażemy teraz, że $b = c$, czyli że macierz m_0 jest symetryczna. Wiemy, że F_1, F_2 są wektorami jednostkowymi wzajemnie prostopadłymi, stąd mamy

$$F_1 \circ F_2 = F_2 \circ F_1 = 0, \ \ F_1 \circ F_1 = 1, \ \ F_2 \circ F_2 = 1.$$

Mnożymy równanie (1) skalarnie przez F_2 i otrzymujemy

$$M_0(F_1) \circ F_2 = (aF_1 + cF_2) \circ F_2 = aF_1 \circ F_2 + cF_2 \circ F_2 = c.$$

Podobnie, mnożąc równanie (2) skalarnie przez F_1 otrzymujemy

$$F_1 \circ M_0(F_2) = F_1 \circ (bF_1 + dF_2) = bF_1 \circ F_1 + dF_1 \circ F_2 = b.$$

Zatem $c = M_0(F_1) \circ F_2 = M(F_1) \circ F_2 = F_1 \circ M(F_2) = F_1 \circ M_0(F_2) = b$.
Przekształcenie M_0 płaszczyzny zadane jest macierzą symetryczną, zatem na podstawie wiadomości o takich macierzach jakie podane zostały w pierwszej części skryptu, wiemy że posiada ono dwa wzajemnie prostopadłe wektory własne U, W. Wówczas $X_2 = \frac{1}{|U|} \cdot U$, $X_3 = \frac{1}{|W|} \cdot W$ są jednostkowymi, wzajemnie prostopadłymi wektorami własnymi dla przekształcenia M_0, czyli również dla M. Wektory X_2, X_3 należą do płaszczyzny Π, tym samym są prostopadłe do wektora X_1, tak więc wektory X_1, X_2, X_3 są parami prostopadłe. $\square$

Uwaga 13.2.4.

(1) *Wartości własne dla wektorów własnych X_1, X_2, X_3, z powyższego twierdzenia, niekoniecznie muszą być różne. Np. jeśli przekształcenie M zadane jest macierzą* $\begin{pmatrix} 5 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 5 \end{pmatrix}$, *wtedy za wektory własne możemy przyjąć*

$$X_1 = e_1,\ X_2 = e_2,\ X_3 = e_3$$

dla wartości własnych $5; -1; 5$. *Natomiast dla macierzy* $\begin{pmatrix} -3 & 0 & 0 \\ 0 & -3 & 0 \\ 0 & 0 & -3 \end{pmatrix}$ *wszystkie wartości własne będą takie same.*

(2) *Wektory X_1, X_2, X_3 na ogół można wybrać na wiele różnych sposobów.*

13.3 Diagonalizacja macierzy.

Poniżej podamy sposób przedstawiania pewnych macierzy (w tym dowolnych macierzy symetrycznych) w postaci iloczynu trzech macierzy, z których jedna jest macierzą diagonalną, a pozostałe są macierzami utworzonymi z wektorów własnych macierzy m. Takie przedstawienie, nazywane diagonalizacją macierzy, ma wiele zastosowań. Podamy jedno przykładowe zastosowanie - do potęgowania macierzy.

Twierdzenie 13.3.1 (diagonalizacja macierzy). *Niech $M : R^3 \to R^3$ będzie przekształceniem liniowym zadanym macierzą m. Załóżmy, że liniowo niezależne wektory X_1, X_2, X_3 są wektorami własnymi przekształcenia M o wartościach własnych t_1, t_2, t_3. Oznaczmy przez $p = (X_1,\ X_2,\ X_3)$ macierz utworzoną z wektorów X_1, X_2, X_3 jako kolumn, zaś przez $d = \begin{pmatrix} t_1 & 0 & 0 \\ 0 & t_2 & 0 \\ 0 & 0 & t_3 \end{pmatrix}$ macierz diagonalną utworzoną z wartości własnych t_1, t_2, t_3. Wówczas macierz m możemy przedstawić w postaci*

$$m = pdp^{-1}.$$

Zanim przejdziemy do dowodu tego twierdzenia, podajmy pomocniczy fakt:

Fakt 13.3.2 (równość przekształceń liniowych). *Niech $T_1, T_2 : R^3 \to R^3$ będą dwoma przekształceniami liniowymi, zaś $A, B, C \in R^3$ liniowo niezależnymi wektorami. Załóżmy ponadto, że $T_1(A) = T_2(A)$, $T_1(B) = T_2(B)$, $T_1(C) = T_2(C)$. Wówczas dla każdego $X \in R^3$ mamy $T_1(X) = T_2(X)$, czyli $T_1 = T_2$.*

Dowód (faktu). Każdy wektor $X \in R^3$ można jednoznacznie przedstawić w postaci $X = t \cdot A + s \cdot B + r \cdot C$ dla pewnych $t, s, r \in R^3$. Wówczas:

$$T_1(X) = T_1(t \cdot A + s \cdot B + r \cdot C) = t \cdot T_1(A) + s \cdot T_1(B) + r \cdot T_1(C) =$$

$$= t \cdot T_2(A) + s \cdot T_2(B) + r \cdot T_2(C) = T_2(t \cdot A + s \cdot B + r \cdot C) = T_2(X).$$

□

Dowód (twierdzenia). Niech P, D będą przekształceniami liniowymi zadanymi macierzami p, d. Mamy wówczas $M(X_i) = t_i X_i$, dla $i = 1, 2, 3$ oraz $P(e_i) = X_i$, dla $i = 1, 2, 3$. Zauważmy, że macierz p jest odwracalna, bo $\det p = \det(X_1, X_2, X_3) \neq 0$, co wynika z liniowej niezależności wektorów X_1, X_2, X_3. Zatem przekształcenie P jest odwracalne i mamy $P^{-1}(X_i) = E_i$. Ponadto zachodzi również $D(e_i) = t_i e_i$, dla $i = 1, 2, 3$. Dla $i = 1, 2, 3$ mamy więc

$$PDP^{-1}(X_i) = PD(e_i) = P(t_i e_i) = t_i P(e_i) = t_i X_i = M(X_i).$$

Ponieważ X_1, X_2, X_3 są liniowo niezależne, z Faktu 13.3.2 wynika $PDP^{-1} = M$, a więc $pdp^{-1} = m$. □

Wniosek 13.3.3. *Niech m będzie macierzą symetryczną rozmiaru 3×3, zaś $M : R^3 \to R^3$ przekształceniem liniowym zadanym tą macierzą. Wówczas*

$$m = pdp^{-1},$$

gdzie p jest pewną macierzą ortogonalną, zaś $d = \begin{pmatrix} t_1 & 0 & 0 \\ 0 & t_2 & 0 \\ 0 & 0 & t_3 \end{pmatrix}$ *jest macierzą diagonalną o wyrazach równych wartościom własnym przekształcenia M.*

Dowód. Zgodnie z Twierdzeniem 13.2.2 przekształcenie M ma trzy parami prostopadłe, jednostkowe wektory własne X_1, X_2, X_3. Macierz $p = (X_1,\ X_2,\ X_3)$, jak wiemy z Lematu 12.1.2, jest macierzą ortogonalną. Na podstawie Twierdzenia 13.3.1 wiemy również, że zachodzi $m = pdp^{-1}$. □

Uwaga 13.3.4 (ortogonalna diagonalizacja macierzy). *Możliwość przedstawienia macierzy m w postaci iloczynu pdp^{-1}, gdzie macierz p jest ortogonalna a macierz d - diagonalna, nazywa się ortogonalną diagonalizacją macierzy m.*

Tak więc Wniosek 13.3.3 oznacza, że każdą macierz symetryczną można zdiagonalizować w sposób ortogonalny.

Rozdział 14

Nowe układy współrzędnych

W celu określenia położenia punktu stosuje się układ współrzędnych. Opisany w pierwszym rozdziale układ współrzędnych to układ kartezjański. Nie jest on jednak jedynym sposobem zapisywania położenia punktów. W wielu sytuacjach wygodniej jest posłużyć się innym układem współrzędnych, tzw. układem ukośnokątnym.

W rozdziale tym bliżej zajmiemy się innymi układami współrzędnych. W pierwszym podrozdziale zdefiniujemy, czym jest układ ukośnokątny na płaszczyźnie, następnie podamy analogiczną definicję w przypadku przestrzeni. Nauczymy się określać współrzędne punktu w tych układach. W kolejnych podrozdziałach poznamy sposoby na zamianę współrzędnych punktów w jednym układzie na współrzędne w innym, pokażemy też, jak znajdować macierz przekształcenia w nowym układzie.

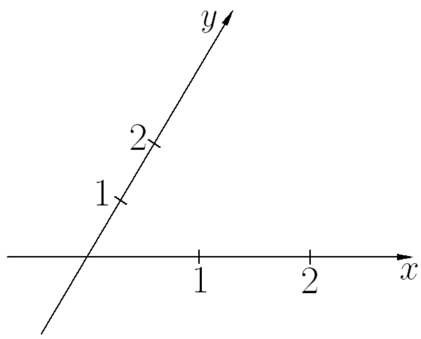
14.1 Układ ukośnokątny na płaszczyźnie.

Definicja 14.1.1 (układ ukośnokątny). Układem ukośnokątnym nazywamy taki układ, w którym:

(1) osie układu nie muszą być prostopadłe,

(2) jednostki na osiach nie muszą być jednakowe.

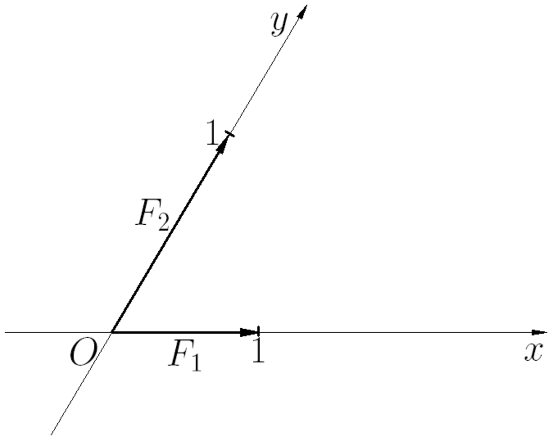
Przykład takiego układu został przedstawiony na rysunku poniżej.



W układzie kartezjańskim wprowadziliśmy pojęcie wersorów. Możemy to zrobić również w układzie ukośnokątnym.

Definicja 14.1.2 (wersory w układzie). Wersorami w układzie nazywamy wektory osiowe o długości równej jednostce na danej osi oraz zwrocie zgodnym z dodatnim kierunkiem na osi.

Na rysunku poniżej wersorami są wektory F_1, F_2.

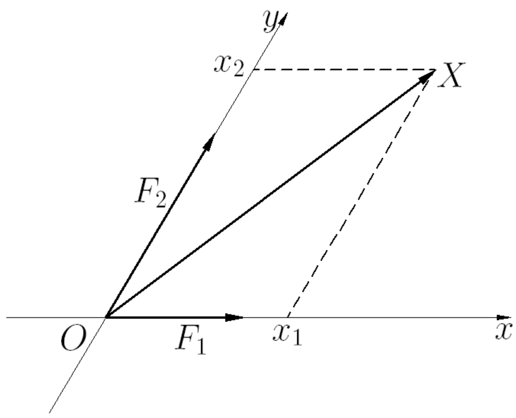


Podobnie jak w układzie kartezjańskim, również w układzie ukośnokątnym możemy przyporządkować każdemu punktowi współrzędne.

Uwaga 14.1.3 (współrzędne punktu). *Współrzędne punktu X wyraża się za pomocą jego rzutów na osie, w kierunku drugiej osi. W układzie ukośnokątnym rzuty te na ogół nie są prostokątne.*

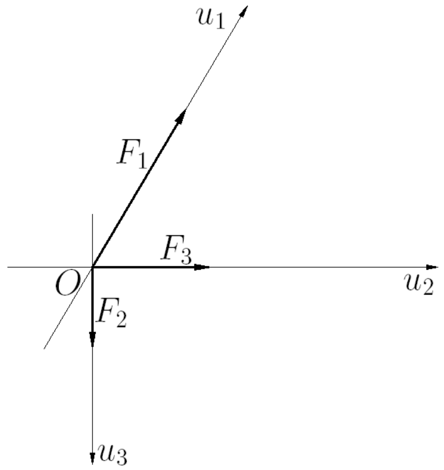
Rozważmy w przyjętym przez nas układzie współrzędnych o wersorach F_1, F_2 punkt X o współrzędnych (x_1, x_2). Możemy te współrzędne zinterpretować w sposób wektorowy i z takiej interpretacji będziemy korzystać w dalszej części. Zauważmy, że punkt X możemy utożsamiać z jego wektorem wodzącym $\overrightarrow{X}$, który wyraża się jako

$$\overrightarrow{X} = x_1 F_1 + x_2 F_2.$$



14.2 Układ ukośnokątny w przestrzeni.

Podobnie jak na płaszczyźnie, tak i w przestrzeni możemy wprowadzić niekartezjański układ współrzędnych. Za wersory takiego układu możemy przyjąć dowolną trójkę liniowo niezależnych wektorów F_1, F_2, F_3. Oznaczmy przez $O_{u_1u_2u_3}$ ukośnokątny układ związany z tymi wektorami (traktowanymi jako wersory).



W przestrzennym układzie ukośnokątnym nie będziemy zajmować się określaniem współrzędnych punktu za pomocą rzutowania. Skorzystamy tutaj z wektorowej charakteryzacji, jaką znamy z płaszczyzny i przez analogię określimy współrzędne w przestrzeni.

Definicja 14.2.1 (współrzędne punktu). Punkt X ma w układzie $O_{u_1u_2u_3}$, o wersorach F_1, F_2, F_3 współrzędne $\begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix}$ wtedy i tylko wtedy gdy $\overrightarrow{X} = v_1F_1 + v_2F_2 + v_3F_3$.

14.3 Macierz przejścia do nowego układu współrzędnych.

Dla danego punktu jesteśmy w stanie określić jego współrzędne w dowolnym układzie współrzędnych. Zastanówmy się jednak, jaki jest związek między współrzędnymi tego samego punktu w różnych układach. Przydatne mogłoby okazać się znalezienie sposobu zamiany jednych współrzędnych na drugie. W naszych rozważaniach zajmować będziemy się tylko takimi parami układów współrzędnych, które mają ten sam początek. Załóżmy, że znamy współrzędne wersorów F_i nowego układu $O_{u_1u_2u_3}$, w układzie kartezjańskim O_{xyz}. Niech $F_i = \begin{pmatrix} f_{1i} \\ f_{2i} \\ f_{3i} \end{pmatrix}$, dla $i = 1, 2, 3$.

Definicja 14.3.1 (macierz przejścia). Macierzą przejścia do układu $O_{u_1u_2u_3}$ nazywamy macierz $f = (F_1\ F_2\ F_3) = \begin{pmatrix} f_{11} & f_{12} & f_{13} \\ f_{21} & f_{22} & f_{23} \\ f_{31} & f_{32} & f_{33} \end{pmatrix}$ utworzoną ze współrzędnych wersorów F_i jako kolumn.

Uwaga 14.3.2. *Zauważmy, że skoro wektory F_1, F_2, F_3 są liniowo niezależne, to macierz przejścia jest macierzą odwracalną.*

14.4 Współrzędne wektorów w nowych układach.

Pokażemy teraz jak mając współrzędne wektora w jednym układzie znaleźć jego współrzędne w drugim. Zacznijmy od łatwiejszego przypadku.

➢ **Zamiana współrzędnych układu $O_{u_1u_2u_3}$ na współrzędne układu O_{xyz}**

Weźmy wektor X, który w nowym układzie $O_{u_1u_2u_3}$ ma współrzędne $\begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix}$. Wiemy, że

$$X = v_1F_1 + v_2F_2 + v_3F_3,$$

co po podstawieniu współrzędnych możemy zapisać jako

$$X = v_1 \begin{pmatrix} f_{11} \\ f_{21} \\ f_{31} \end{pmatrix} + v_2 \begin{pmatrix} f_{12} \\ f_{22} \\ f_{32} \end{pmatrix} + v_3 \begin{pmatrix} f_{13} \\ f_{23} \\ f_{33} \end{pmatrix}.$$

Przekształcając dalej otrzymujemy

$$X = \begin{pmatrix} v_1f_{11} + v_2f_{12} + v_3f_{13} \\ v_1f_{21} + v_2f_{22} + v_3f_{23} \\ v_1f_{31} + v_2f_{32} + v_3f_{33} \end{pmatrix} = \begin{pmatrix} f_{11} & f_{12} & f_{13} \\ f_{21} & f_{22} & f_{23} \\ f_{31} & f_{32} & f_{33} \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix} = f \cdot V.$$

Krótko możemy więc zapisać

$$X = fV.$$

✔ **Przykład.** Rozważmy wektor X, który w układzie O_{u_1,u_2,u_3}, o wersorach

$$F_1 = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \; F_2 = \begin{pmatrix} \frac{1}{2} \\ 1 \\ 0 \end{pmatrix}, \; F_3 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix},$$

ma współrzędne $V = \begin{pmatrix} 1 \\ 5 \\ -2 \end{pmatrix}$. Znajdźmy współrzędne tego wektora w układzie O_{xyz}. Zgodnie z powyższymi rozważaniami współrzędne tego wektora w układzie O_{xyz} wyrażają się wzorem $X = fV$, gdzie f jest macierzą przejścia do układu $O_{u_1u_2u_3}$, w naszym przypadku

$$f = \begin{pmatrix} 1 & \frac{1}{2} & 1 \\ 1 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}.$$

Wektor X ma zatem współrzędne

$$X = fV = \begin{pmatrix} 1 & \frac{1}{2} & 1 \\ 1 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ 5 \\ -2 \end{pmatrix} = \begin{pmatrix} 1\frac{1}{2} \\ 4 \\ -2 \end{pmatrix}.$$

Wektor, który w układzie o wersorach F_1, F_2, F_3 ma współrzędne $V = \begin{pmatrix} 1 \\ 5 \\ -2 \end{pmatrix}$ w układzie O_{xyz} ma współrzędne $X = \begin{pmatrix} 1\frac{1}{2} \\ 4 \\ -2 \end{pmatrix}$.

Zajmijmy się teraz trudniejszym przypadkiem.

➢ **Zamiana współrzędnych układu O_{xyz} na współrzędne układu $O_{u_1u_2u_3}$**

Niech X będzie wektorem, który w układzie O_{xyz} ma współrzędne $X = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$.

Spróbujmy znaleźć współrzędne $\begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix}$ tego wektora w układzie $O_{u_1u_2u_3}$. Oznaczmy w skrócie $V = \begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix}$. Jak wiemy z poprzedniego podpunktu,

$$X = fV.$$

Współrzędne V możemy zatem obliczyć mnożąc tę równość lewostronnie przez macierz f^{-1}. Macierz f^{-1} istnieje, bo macierz f, utworzona z liniowo niezależnych wektorów F_1, F_2, F_3, jest odwracalna. Otrzymujemy zatem

$$f^{-1}X = f^{-1}fV = V,$$

czyli

$$V = f^{-1}X.$$

✔ **Przykład.** Rozważmy ten sam wektor X co w poprzednim przykładzie. Tym razem załóżmy jednak, że mamy dane współrzędne tego wektora w układzie O_{xyz} tzn. $X = \begin{pmatrix} 1\frac{1}{2} \\ 4 \\ -2 \end{pmatrix}$. Chcemy znaleźć współrzędne V tego wektora w układzie $O_{u_1u_2u_3}$ o wersorach $F_1 = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}$, $F_2 = \begin{pmatrix} \frac{1}{2} \\ 1 \\ 0 \end{pmatrix}$, $F_3 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$. Zgodnie z przeprowadzonym powyżej rozumowaniem współrzędne te wyrażają się wzorem

$$V = f^{-1}X.$$

Wyznaczmy zatem macierz odwrotną do macierzy f. Zgodnie z Twierdzeniem 7.4.7 macierz ta zadana jest wzorem

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

Współrzędne wektora X w nowym układzie to zatem

$$V = \begin{pmatrix} 2 & -1 & -1 \\ -2 & 2 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1\frac{1}{2} \\ 4 \\ -2 \end{pmatrix} = \begin{pmatrix} 1 \\ 5 \\ -2 \end{pmatrix}.$$

14.5 Macierze przekształceń w nowych układach współrzędnych.

Potrafimy już określić współrzędne wektora w nowym układzie. Postaramy się teraz, mając dane przekształcenie liniowe M o macierzy m, znaleźć macierz m' tego przekształcenia w nowym układzie. Macierz m' będzie taką macierzą, dla której obrazem wektora $U = \begin{pmatrix} u_1 \\ u_2 \\ u_3 \end{pmatrix}$ (są to współrzędne w nowym układzie) przez przekształcenie M, jest wektor $U' = \begin{pmatrix} u'_1 \\ u'_2 \\ u'_3 \end{pmatrix}$, taki że $U' = m'U$.

Weźmy przekształcenie liniowe $M : R^3 \to R^3$, które w układzie O_{xyz} ma macierz m. Znajdźmy macierz m' tego przekształcenia w układzie $O_{u_1u_2u_3}$. Rozważmy wektor $X = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$. Obraz tego wektora przez przekształcenie M to $M(X) = m \cdot X$. Wektor X ma w układzie $O_{u_1u_2u_3}$ współrzędne $V = f^{-1}X$. Współrzędne obrazu tego wektora w układzie $O_{u_1u_2u_3}$ to $f^{-1}mX$. Zauważmy, że jeśli m' jest macierzą przekształcenia M w układzie $O_{u_1u_2u_3}$, to $M(X)$ ma w tym układzie współrzędne $m'V$. Zatem dla dowolnego wektora $X \in R^3$ mamy

$$f^{-1}mX = m'V,$$

czyli

$$f^{-1}mX = m'f^{-1}X.$$

Równość ta zachodzi dla dowolnego wektora $X \in R^3$.

Zanim przejdziemy dalej podamy pomocniczy lemat.

Lemat 14.5.1. *Dla m_1, m_2 będących dowolnymi macierzami rozmiaru 3×3, jeśli dla każdego wektora $X \in R^3$ zachodzi $m_1X = m_2X$, to $m_1 = m_2$.*

Dowód. Niech $m_1 = \begin{pmatrix} a_1 & b_1 & c_1 \\ d_1 & e_1 & f_1 \\ g_1 & h_1 & k_1 \end{pmatrix}$, $m_2 = \begin{pmatrix} a_2 & b_2 & c_2 \\ d_2 & e_2 & f_2 \\ g_2 & h_2 & k_2 \end{pmatrix}$,

wówczas $m_1 \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} a_1x_1 + b_1x_2 + c_1x_3 \\ d_1x_1 + e_1x_2 + f_1x_3 \\ g_1x_1 + h_1x_2 + k_1x_3 \end{pmatrix}$, $m_2 \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} a_2x_1 + b_2x_2 + c_2x_3 \\ d_2x_1 + e_2x_2 + f_2x_3 \\ g_2x_1 + h_2x_2 + k_2x_3 \end{pmatrix}$.

Skoro $m_1X = m_2X$, to np.

$$a_1x_1 + b_1x_2 + c_1x_3 = a_2x_1 + b_2x_2 + c_2x_3.$$

Wiemy, że równość ta jest spełniona dla dowolnego wektora X, więc w szczególności dla $X = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$. Otrzymujemy wówczas

$$a_1 \cdot 1 + b_1 \cdot 0 + c_1 \cdot 0 = a_2 \cdot 1 + b_2 \cdot 0 + c_2 \cdot 0,$$

czyli

$$a_1 = a_2.$$

Podobnie, dla wektora $X = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$ otrzymujemy $b_1 = b_2$. Postępując analogicznie, możemy wykazać, że macierz m_1 ma wszystkie wyrazy równe odpowiednim wyrazom macierzy m_2, czyli $m_1 = m_2$. □

Wróćmy do wyznaczania macierzy m'. Otrzymaliśmy dla dowolnego $X \in R^3$

$$f^{-1}mX = m'f^{-1}X.$$

Stosując powyższy lemat mamy

$$f^{-1}m = m'f^{-1}.$$

Mnożąc obie strony tej równości z prawej strony przez macierz f dostajemy ostateczny wzór na m':

$$f^{-1}mf = m'.$$

Twierdzenie 14.5.2. *Macierz m' przekształcenia M w nowym układzie współrzędnych wyraża się wzorem*

$$\boxed{f^{-1}mf = m',}$$

gdzie f jest macierzą przejścia do nowego układu współrzędnych, natomiast m to macierz przekształcenia M w wyjściowym układzie.

✔ **Przykład.** Niech $T : R^3 \to R^3$ będzie przekształceniem liniowym mającym trzy liniowo niezależne wektory własne F_1, F_2, F_3 o wartościach własnych t_1, t_2, t_3. Wówczas w układzie $O_{u_1u_2u_3}$ związanym z wektorami F_1, F_2, F_3, przekształcenie T zadane jest macierzą diagonalną $\begin{pmatrix} t_1 & 0 & 0 \\ 0 & t_2 & 0 \\ 0 & 0 & t_3 \end{pmatrix}$.

Podamy teraz dwa uzasadnienia poprawności tak wyznaczonej macierzy.

Uzasadnienie 1 (bezpośrednie).
Macierz przekształcenia tworzy się z obrazów wersorów $T(F_1), T(F_2), T(F_3)$, rozkłada-

jąc je względem F_1, F_2, F_3:

$$\left.\begin{aligned} T(F_1) &= t_1F_1 = t_1F_1 + 0\cdot F_2 + 0\cdot F_3 = \begin{pmatrix} t_1 \\ 0 \\ 0 \end{pmatrix} \\ T(F_2) &= t_2F_2 = 0\cdot F_1 + t_2F_2 + 0\cdot F_3 = \begin{pmatrix} 0 \\ t_2 \\ 0 \end{pmatrix} \\ T(F_3) &= 0\cdot F_1 + 0\cdot F_2 + t_3F_3 = \begin{pmatrix} 0 \\ 0 \\ t_3 \end{pmatrix} \end{aligned}\right\} \begin{array}{c}\text{współrzędne} \\ \text{w układzie} \\ O_{u_1u_2u_3}\end{array}$$

Tak więc szukana macierz to

$$m' = \begin{pmatrix} t_1 & 0 & 0 \\ 0 & t_2 & 0 \\ 0 & 0 & t_3 \end{pmatrix}.$$

Uzasadnienie 2 (korzystające ze wzoru $m' = f^{-1}mf$).
Ponieważ $f = (F_1\ F_2\ F_3)$, zaś F_i są wektorami własnymi, to korzystając z twierdzenia o diagonalizacji, czyli Twierdzenia 13.3.1, możemy zapisać macierz m przekształcenia T jako $m = fdf^{-1}$, gdzie $d = \begin{pmatrix} t_1 & 0 & 0 \\ 0 & t_2 & 0 \\ 0 & 0 & t_3 \end{pmatrix}$. Wstawiając to do wzoru na m' otrzymujemy

$$m' = f^{-1}mf = f^{-1}(fdf^{-1})f = (f^{-1}f)d(f^{-1}f) = d,$$

czyli

$$m' = d = \begin{pmatrix} t_1 & 0 & 0 \\ 0 & t_2 & 0 \\ 0 & 0 & t_3 \end{pmatrix},$$

tak jak chcieliśmy.

Rozdział 15

Formy kwadratowe i powierzchnie stopnia drugiego

W rozdziale tym wprowadzony zostanie pewien rodzaj funkcji, zwanych formami kwadratowymi. W początkowych podrozdziałach poznamy podstawowe definicje związane z tym pojęciem, poznamy przykłady form kwadratowych, nauczymy się również znajdować ich postać w nowych układach współrzędnych. W dalszej części poznamy sposób na odróżnienie ich od siebie, za pomocą powiązanych z nimi powierzchni charakterystycznych, zwanych kwadrykami. Na koniec szerzej zajmiemy się własnościami form kwadratowych i związanych z nimi macierzy symetrycznych.

15.1 Podstawowe definicje.

Definicja 15.1.1 (forma kwadratowa). Forma kwadratowa w R^3 to funkcja postaci

$$\varphi\begin{pmatrix} x \\ y \\ z \end{pmatrix} = ax^2 + 2bxy + 2cxz + dy^2 + 2eyz + fz^2,$$

gdzie a, b, c, d, e, f są pewnymi ustalonymi liczbami rzeczywistymi, nazywanymi współczynnikami formy φ.

Tak określonemu przekształceniu możemy przyporządkować pewną macierz, zwaną macierzą formy kwadratowej.

Definicja 15.1.2 (macierz formy kwadratowej). Macierzą formy kwadratowej φ postaci $\varphi\begin{pmatrix} x \\ y \\ z \end{pmatrix} = ax^2 + 2bxy + 2cxz + dy^2 + 2eyz + fz^2$ nazywamy macierz symetryczną

$$m = \begin{pmatrix} a & b & c \\ b & d & e \\ c & e & f \end{pmatrix}.$$

✔ **Przykład.**
(1) Dla formy kwadratowej określonej wzorem

$$\varphi\begin{pmatrix} x \\ y \\ z \end{pmatrix} = 4xy + 3xz + 2y^2 - 3z^2$$

odpowiadająca jej macierz to

$$m = \begin{pmatrix} 0 & 2 & \frac{3}{2} \\ 2 & 2 & 0 \\ \frac{3}{2} & 0 & -3 \end{pmatrix}.$$

(2) Macierzy symetrycznej

$$m = \begin{pmatrix} -1 & 3 & 1 \\ 3 & 4 & -2 \\ 1 & -2 & 0 \end{pmatrix}$$

odpowiada forma kwadratowa postaci

$$\varphi\begin{pmatrix} x \\ y \\ z \end{pmatrix} = -x^2 + 6xy + 2xz + 4y^2 - 4yz.$$

✍ Intuicyjnie oczywista wydaje się następująca uwaga, której dowód pozostawimy do wykonania czytelnikowi.

Uwaga 15.1.3. *Każdej formie kwadratowej odpowiada dokładnie jedna macierz symetryczna, jednocześnie każdej macierzy symetrycznej odpowiada jedna forma kwadratowa.*

Podana powyżej definicja macierzy formy kwadratowej nabiera sensu w świetle następującego faktu.

Fakt 15.1.4. *Jeśli m jest macierzą formy kwadratowej φ, zaś $X = \begin{pmatrix} x \\ y \\ z \end{pmatrix}$ jest dowolnym wektorem, to*

$$\varphi\begin{pmatrix} x \\ y \\ z \end{pmatrix} = X \circ (mX).$$

Dowód. W dowodzie skorzystamy z definicji iloczynu skalarnego i podstawowych własności działań na macierzach oraz liczbach rzeczywistych:

$$X \circ (mX) = \begin{pmatrix} x \\ y \\ z \end{pmatrix} \circ \begin{pmatrix} a & b & c \\ b & d & e \\ c & e & f \end{pmatrix}\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} x \\ y \\ z \end{pmatrix} \circ \begin{pmatrix} ax + by + cz \\ bx + dy + ez \\ cx + ey + fz \end{pmatrix} =$$

$$= ax^2 + bxy + cxz + bxy + dy^2 + eyz + cxz + eyz + fz^2 =$$

$$= ax^2 + 2bxy + 2cxz + dy^2 + 2eyz + fz^2 = \varphi\begin{pmatrix} x \\ y \\ z \end{pmatrix}.$$

□

15.2 Formy kwadratowe w nowych układach współrzędnych.

W poprzednim rozdziale nauczyliśmy się znajdować współrzędne wektora oraz postać macierzy przekształcenia w nowym układzie współrzędnych. Teraz pokażemy, jak w nowym układzie znaleźć postać formy kwadratowej.

Lemat 15.2.1 (**o postaci formy kwadratowej w nowym układzie współrzędnych**). *Niech φ będzie formą kwadratową w R^3 zadaną (we współrzędnych O_{xyz}) macierzą m. Niech O_{uvw} będzie nowym układem współrzędnych, o macierzy przejścia f. Wówczas forma φ względem nowych współrzędnych u, v, w ma taką samą ogólną postać $\varphi\begin{pmatrix} u \\ v \\ w \end{pmatrix} = Au^2 + Buv + Cuw + Dv^2 + Evw + Fw^2$ na ogół z zupełnie innymi współczynnikami A, B, C, D, E, F. Oznacza to, że forma φ jest w nowym układzie współrzędnych dalej formą kwadratową. Ponadto macierz m' formy φ w nowym układzie wyraża się wzorem*

$$m' = f^T \cdot m \cdot f.$$

Dowód.

Dowód powyższego lematu przeprowadzimy w trzech etapach. Na początek pokażemy, że forma φ w nowym układzie jest rzeczywiście formą kwadratową, następnie udowodnimy, że macierz tej formy wyraża się podanym wzorem. Na koniec uzasadnimy, że macierz m' jest macierzą symetryczną.

Oznaczmy przez $U = \begin{pmatrix} u \\ v \\ w \end{pmatrix}$ nowe współrzędne wektora $X = \begin{pmatrix} x \\ y \\ z \end{pmatrix}$. Jak wiemy z poprzedniego rozdziału, wówczas $X = fU$, czyli

$$\begin{array}{l} x = f_{11}u + f_{12}v + f_{13}w \\ y = f_{21}u + f_{22}v + f_{23}w \\ z = f_{31}u + f_{32}v + f_{33}w \end{array}.$$

Po podstawieniu tych wyrażeń do wzoru na $\varphi = ax^2 + 2bxy + 2cxz + dy^2 + 2eyz + fz^2$, po uproszczeniu i pogrupowaniu ze względu na $u^2, v^2, w^2, uv, uw, vw$, otrzymujemy

$$\begin{aligned} \varphi = (af_{11}^2 + 2bf_{11}f_{21} + 2cf_{11}f_{31} + df_{21}^2 + 2ef_{21}f_{31} + ff_{31}^2) \cdot u^2 + \\ + (af_{12}^2 + 2bf_{12}f_{22} + 2cf_{12}f_{32} + df_{22}^2 + 2ef_{22}f_{32} + ff_{32}^2) \cdot v^2 + \\ + (af_{13}^2 + 2bf_{13}f_{23} + 2cf_{13}f_{33} + df_{23}^2 + 2ef_{23}f_{33} + ff_{33}^2) \cdot w^2 + \\ +2\cdot(af_{11}f_{12} + b(f_{11}f_{22} + f_{12}f_{21}) + c(f_{11}f_{32} + f_{12}f_{31}) + df_{21}f_{22} + e(f_{21}f_{32} + f_{22}f_{31}) + ff_{32}f_{31}) \cdot uv + \\ +2\cdot(af_{11}f_{13} + b(f_{11}f_{23} + f_{13}f_{21}) + c(f_{11}f_{33} + f_{13}f_{31}) + df_{21}f_{23} + e(f_{21}f_{33} + f_{23}f_{31}) + ff_{33}f_{31}) \cdot uw + \\ +2\cdot(af_{12}f_{13} + b(f_{12}f_{23} + f_{13}f_{22}) + c(f_{12}f_{33} + f_{13}f_{32}) + df_{22}f_{23} + e(f_{22}f_{33} + f_{23}f_{32}) + ff_{32}f_{33}) \cdot vw. \end{aligned}$$

Jak widać, współczynniki przy $u^2, v^2, w^2, uv, uw, vw$ zależą tylko od współczynników a, b, c, d, e, f oraz od wyrazów macierzy przejścia, więc w skrócie możemy macierz φ zapisać w postaci $\varphi = Au^2 + Buv + Cuw + Dv^2 + Evw + Fw^2$.

Przejdźmy do drugiej części dowodu. Wzór $m' = f^T mf$ na macierz formy kwadratowej w nowym układzie wyprowadzimy dla szczególnego przypadku, gdy nowy układ O_{uvw} jest też układem kartezjańskim, tzn. jego wersory F_1, F_2, F_3 są parami prostopadłe i mają długość 1. Zanim przejdziemy dalej, udowodnimy pewien pomocniczy fakt.

Fakt 15.2.2. *Macierz przejścia f do nowego układu kartezjańskiego jest macierzą ortogonalną.*

Dowód (faktu).
Macierz f jest utworzona z wersorów układu traktowanych jako kolumny $f = (F_1\ F_2\ F_3)$. Skoro F_1, F_2, F_3 są parami prostopadłe i jednostkowe, to macierz f, zgodnie z Lematem 12.1.2, jest ortogonalna. □

Wróćmy do dowodu wzoru. Zauważmy, że zgodnie z Faktem 15.1.4, w układzie O_{xyz} mamy

$$\varphi(X) = X \circ (mX),$$

a w układzie O_{uvw}

$$\varphi(X) = U \circ (m'U).$$

Zatem równość $X \circ (mX) = U \circ (m'U)$ zachodzi dla dowolnego wektora $X \in R^3$ i jego współrzędnych U w nowym układzie. Szukamy więc m' takiego, że zachodzi $\varphi = U \circ m'U$. Zauważmy, że skoro macierz f jest ortogonalna, to również f^{-1} jest ortogonalna, tak więc dla dowolnych wektorów A, B mamy $f^{-1}A \circ f^{-1}B = A \circ B$. Wykorzystując tę zależność w naszej równości, otrzymujemy

$$\varphi = X \circ (mX) = fU \circ mfU = f^{-1}(fU) \circ f^{-1}(mfU) = U \circ (f^{-1}mf)U = U \circ (f^{-1}mf)U.$$

Wiemy, że spełniona jest równość $X \circ (mX) = U \circ (m'U)$, tak więc mamy

$$U \circ (f^{-1}mf)U = U \circ (m'U),$$

dla U będących współrzędnymi w nowym układzie dla dowolnego X, czyli dla dowolnego U.
Szukana macierz m' to zatem

$$m' = f^{-1}mf.$$

Ogólny przypadek można uzasadnić bezpośrednim, niestety bardzo obszernym, rachunkiem. Tutaj go pominiemy.

Na koniec pozostaje jeszcze sprawdzić, że macierz m' jest rzeczywiście macierzą symetryczną. W tym celu sprawdźmy, że

$$(m')^T = m'.$$

Pamiętajmy, że macierz m jako macierz formy kwadratowej jest macierzą symetryczną. Mamy wówczas

$$(m')^T = (f^T mf)^T = f^T \cdot m^T \cdot (f^T)^T = f^T \cdot m \cdot f = m',$$

co kończy dowód. □

15.3 Kwadryki form kwadratowych.

Najważniejsze, a zarazem najłatwiejsze do obserwowania, formy kwadratowe to te, które nie zawierają w swoim wzorze wyrazów mieszanych, czyli formy postaci $ax^2 + dy^2 + fz^2$, opisane przez macierze diagonalne. Takie formy można odróżniać od siebie i dzielić na pewne grupy za pomocą tzw. powierzchni charakterystycznych zwanych też kwadrykami form.

Definicja 15.3.1 (**kwadryka formy kwadratowej**). Kwadryką związaną z formą kwadratową φ nazywamy powierzchnię (lub zbiór) w R^3 zadaną wzorem

$$\varphi\begin{pmatrix} x \\ y \\ z \end{pmatrix} = 1.$$

W dalszej części tego rozdziału spróbujemy dokonać przeglądu kwadryk form kwadratowych zadanych macierzami diagonalnymi, czyli kwadryk postaci

$$ax^2 + dy^2 + fz^2 = 1.$$

Podzielimy kwadryki ze względu na znak współczynników a, d, f.

➢ Przypadek (1), gdy $\mathbf{a > 0,\ d > 0,\ f > 0}$

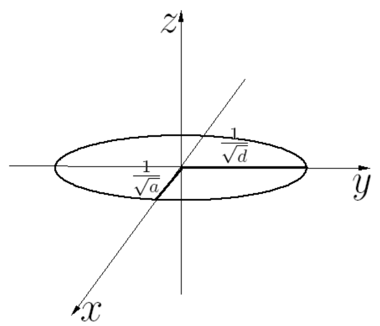
Równanie $ax^2 + dy^2 + fz^2 = 1$ możemy przekształcić do postaci

$$\frac{x^2}{\left(\frac{1}{\sqrt{a}}\right)^2} + \frac{y^2}{\left(\frac{1}{\sqrt{d}}\right)^2} + \frac{z^2}{\left(\frac{1}{\sqrt{f}}\right)^2} = 1.$$

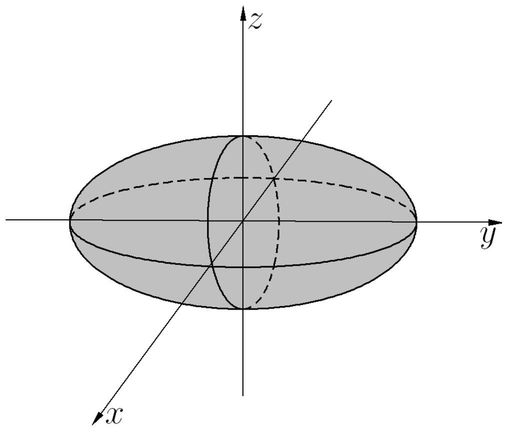
Zauważmy, że równanie zapisane w tej postaci przypomina równanie elipsy. Zobaczmy, czym jest przekrój powierzchni zadanej tym równaniem z płaszczyznami współrzędnych. Przekrój z płaszczyzną $z = 0$ to elipsa o równaniu

$$\frac{x^2}{\left(\frac{1}{\sqrt{a}}\right)^2} + \frac{y^2}{\left(\frac{1}{\sqrt{d}}\right)^2} = 1,$$

czyli elipsa o półosiach długości $\frac{1}{\sqrt{a}}$, $\frac{1}{\sqrt{d}}$.



Przekroje z płaszczyznami $x = 0$, $y = 0$ również będą elipsami. Szukaną przez nas kwadryką jest **elipsoida** o półosiach długości $\frac{1}{\sqrt{a}}$, $\frac{1}{\sqrt{d}}$, $\frac{1}{\sqrt{f}}$.



➢ **Przypadek (2), gdy a $>$ 0, d $>$ 0, f $<$ 0**

Podobnie jak poprzednio, możemy przekształcić równanie opisujące tę kwadrykę. Otrzymujemy

$$\frac{x^2}{\frac{1}{a}} + \frac{y^2}{\frac{1}{d}} + \frac{z^2}{\frac{1}{f}} = 1.$$

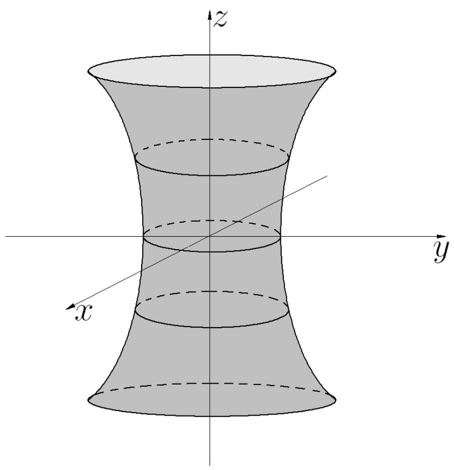
Przekształcając dalej mamy

$$\frac{x^2}{\frac{1}{a}} + \frac{y^2}{\frac{1}{d}} - \frac{z^2}{-\frac{1}{f}} = 1.$$

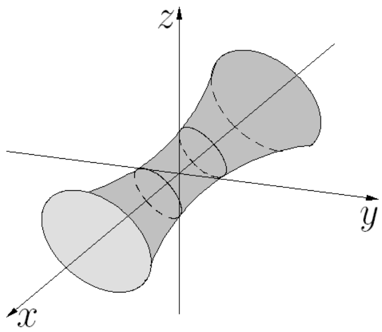
Zauważmy, że $-\frac{1}{f}$ jest dodatni, więc równanie powyższe możemy zapisać jako

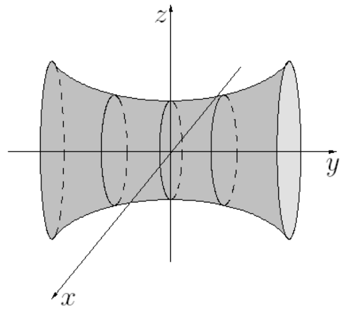
$$\frac{x^2}{\left(\frac{1}{\sqrt{a}}\right)^2} + \frac{y^2}{\left(\frac{1}{\sqrt{d}}\right)^2} - \frac{z^2}{\left(\frac{1}{\sqrt{-f}}\right)^2} = 1.$$

Zbadajmy przekroje tej figury z płaszczyznami współrzędnych. Przekrój z płaszczyzną $z = 0$ to elipsa. Przekroje z płaszczyznami $x = 0$ i $y = 0$ to hiperbole. Szukana kwadryka to **hiperboloida jednopowłokowa eliptyczna**.



Przypadki, gdy $a < 0,\ d > 0,\ f > 0$ oraz $a > 0,\ d < 0,\ f > 0$ można rozważać analogicznie. Otrzymujemy wówczas kwadryki przedstawione na rysunkach poniżej, gdzie pierwszemu warunkowi odpowiada kwadryka przedstawiona na pierwszym rysunku.





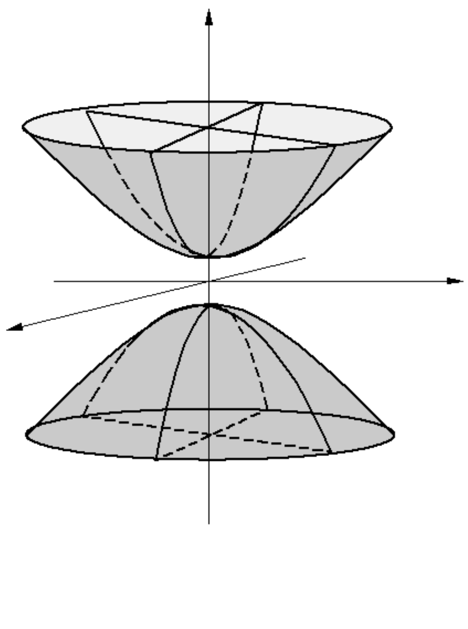
Te trzy rodzaje hiperboloid wyglądają podobnie są tylko obrócone.

➢ **Przypadek (3), gdy $\mathbf{a > 0,\ d < 0,\ f < 0}$**

Tak jak poprzednio, przypadek ten możemy traktować bardziej ogólnie, mianowicie obejmuje on wszystkie kombinacje współczynników, z których jeden jest dodatni, a pozostałe ujemne. Przekształćmy równość opisującą tę kwadrykę. Otrzymujemy

$$\frac{x^2}{\left(\frac{1}{\sqrt{a}}\right)^2} - \frac{y^2}{\left(\frac{1}{\sqrt{-d}}\right)^2} - \frac{z^2}{\left(\frac{1}{\sqrt{-f}}\right)^2} = 1.$$

W tym wypadku nie będziemy przytaczać przekształceń prowadzących do tej równości. Przekroje z płaszczyznami $z = 0, y = 0$ są hiperbolami. Przekrój z płaszczyzną $x = 0$ jest zbiorem pustym. Jeżeli będziemy badać przekroje z płaszczyznami równoległymi do tej płaszczyzny, czyli płaszczyznami $x = c$, dla $c \in R$, to okaże się, że dla $|c| > \frac{1}{\sqrt{a}}$ przekroje te będą elipsami. Kwadryka, którą opisuje to równanie, to **hiperboloida dwupowłokowa eliptyczna**.



➢ Przypadek (4), gdy $\mathbf{a < 0,\ d < 0,\ f < 0}$

W tym wypadku równość opisującą tę kwadrykę możemy przekształcić do postaci

$$-\frac{x^2}{\left(\frac{1}{\sqrt{-a}}\right)^2}-\frac{y^2}{\left(\frac{1}{\sqrt{-d}}\right)^2}-\frac{z^2}{\left(\frac{1}{\sqrt{-f}}\right)^2}=1.$$

Zauważmy, że po lewej stronie mamy wyrażenie, które zawsze przyjmuje wartości ujemne, lub zero, natomiast prawa strona równości to 1, tak więc zbiorem rozwiązań tej równości jest zbiór pusty. W takich przypadkach można rozważać drugą powierzchnię charakterystyczną $\varphi\begin{pmatrix} x \\ y \\ z \end{pmatrix}=-1$. W naszym przypadku otrzymujemy

$$-\frac{x^2}{\left(\frac{1}{\sqrt{-a}}\right)^2}-\frac{y^2}{\left(\frac{1}{\sqrt{-d}}\right)^2}-\frac{z^2}{\left(\frac{1}{\sqrt{-f}}\right)^2}=-1.$$

Po obustronnym przemnożeniu przez -1 daje to

$$\frac{x^2}{\left(\frac{1}{\sqrt{-a}}\right)^2}+\frac{y^2}{\left(\frac{1}{\sqrt{-d}}\right)^2}+\frac{z^2}{\left(\frac{1}{\sqrt{-f}}\right)^2}=1.$$

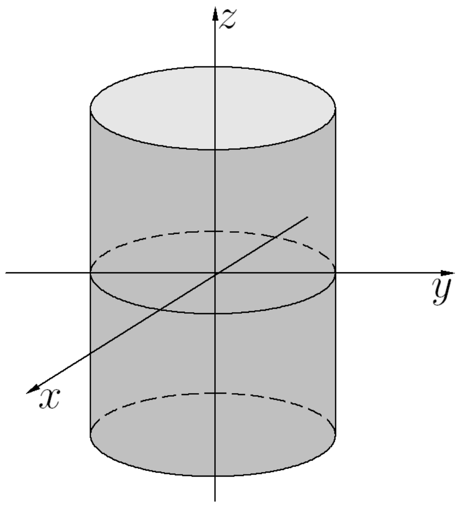
Druga powierzchnia charakterystyczna będzie w tym wypadku elipsoidą.

➢ Przypadek (5), gdy $\mathbf{a > 0,\ d > 0,\ f = 0}$

W tym wypadku równanie kwadryki można przekształcić do postaci

$$\frac{x^2}{\left(\frac{1}{\sqrt{a}}\right)^2}+\frac{y^2}{\left(\frac{1}{\sqrt{d}}\right)^2}=1.$$

Zauważmy, że jest to równanie elipsy na płaszczyźnie. Równanie to nie zależy od zmiennej z. Przekrojem z dowolną płaszczyzną $z=c$ będzie zatem elipsa. Zauważmy, że bez względu na to, jakie c wybierzemy zawsze otrzymamy tę samą elipsę. Zatem dla takich współczynników otrzymujemy figurę nazywaną **walcem eliptycznym**. Warto zwrócić uwagę na fakt, że figura ta rozciąga się nieskończenie w górę i w dół, ale nie ma żadnej podstawy jak znane nam ze szkoły walce.



Możemy rozważać jeszcze dwie analogiczne sytuacje, tzn. gdy $a = 0,\ d > 0,\ f > 0$ oraz $a > 0,\ d = 0,\ f > 0$. Kwadryki otrzymane w ten sposób wyglądają podobnie jak ta na rysunku powyżej, są tylko obrócone.

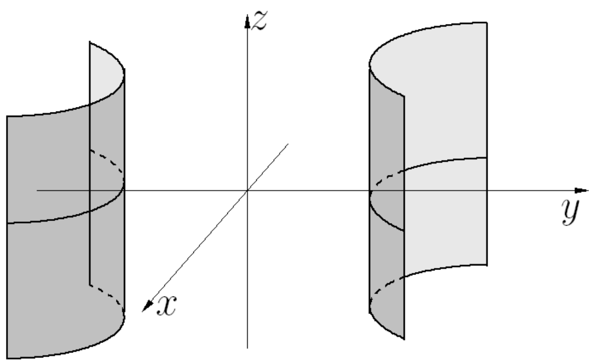
W większości z dalszych przypadków będziemy mogli określić takie dodatkowe kwadryki, które są jedynie obrotami tych rozważanych przez nas. Dlatego pominiemy je w dalszych rozważaniach.

➢ **Przypadek (6), gdy $\mathbf{a > 0,\ d < 0,\ f = 0}$**

Równanie tej kwadryki można przekształcić do postaci

$$\frac{x^2}{\left(\frac{1}{\sqrt{a}}\right)^2} - \frac{y^2}{\left(\frac{1}{\sqrt{-d}}\right)^2} = 1.$$

Zauważmy, że również w tym przypadku równanie nie zależy od zmiennej z. W przekroju z dowolną płaszczyzną równoległą do $z = 0$ otrzymamy tę samą hiperbolę. Opisana tym równaniem kwadryka to **walec hiperboliczny**.



➢ Przypadek (7), gdy $\mathbf{a < 0,\ d < 0,\ f = 0}$

Równanie tej kwadryki możemy przekształcić do postaci

$$-\frac{x^2}{\left(\frac{1}{\sqrt{-a}}\right)^2} - \frac{y^2}{\left(\frac{1}{\sqrt{-d}}\right)^2} = 1.$$

Jak nietrudno się przekonać, podobnie jak w jednym z poprzednich przypadków, zbiór rozwiązań tego równania jest zbiorem pustym. Również tutaj możemy rozważać drugą powierzchnię charakterystyczną $ax^2 + dy^2 = -1$, której równanie możemy sprowadzić do postaci

$$\frac{x^2}{\left(\frac{1}{\sqrt{-a}}\right)^2} + \frac{y^2}{\left(\frac{1}{\sqrt{-d}}\right)^2} = 1.$$

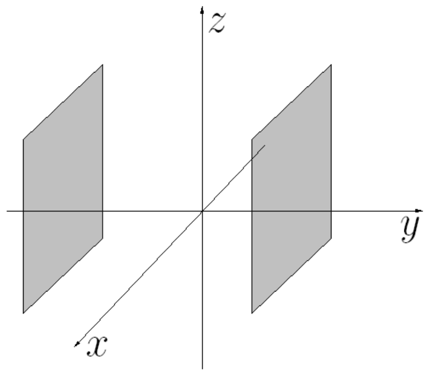
Z postaci tego równania możemy wnioskować, że druga powierzchnia charakterystyczna jest walcem eliptycznym.

➢ Przypadek (8), gdy $\mathbf{a > 0,\ d = 0,\ f = 0}$

Równanie kwadryki przyjmuje postać

$$ax^2 = 1,$$

co możemy zapisać jako $x = \pm\frac{1}{\sqrt{a}}$. Równanie to nie zależy od zmiennej y i zmiennej z, zatem dla $x = \frac{1}{\sqrt{a}}$ przebiega wszystkie wartości y, z jest to więc płaszczyzna równoległa do płaszczyzny $x = 0$ i przecinająca oś O_x w punkcie $(\frac{1}{\sqrt{a}}, 0, 0)$. Podobnie mamy dla $x = -\frac{1}{\sqrt{a}}$. Rozwiązaniem są zatem dwie płaszczyzny równoległe do płaszczyzny $x = 0$.



➢ Przypadek (9), gdy $\mathbf{a < 0,\ d = 0,\ f = 0}$

Podobnie jak w poprzednim przypadku, równanie kwadryki to

$$ax^2 = 1.$$

Zauważmy jednak, że skoro $a < 0$, to również ax^2 jest liczbą ujemną lub zerem, tak więc nigdy nie będzie równe 1, stąd rozwiązaniem jest zbiór pusty. Druga powierzchnia charakterystyczna, której równanie to

$$ax^2 = -1,$$

sprowadza się do poprzedniego przypadku, zatem jest parą płaszczyzn równoległych.

➢ **Przypadek (10), gdy $\mathbf{a} = \mathbf{0}$, $\mathbf{d} = \mathbf{0}$, $\mathbf{f} = \mathbf{0}$**

Dla takich współczynników forma φ to forma zerowa. W takim wypadku zarówno pierwsza, jak i druga powierzchnia charakterystyczna to zbiór pusty.

15.4 Kwadryki w nowych układach współrzędnych.

Dotychczas rozważaliśmy kwadryki form zadanych macierzami diagonalnymi. Poniższe twierdzenie pokazuje, że w ten sposób rozważyliśmy już wszystkie możliwe kwadryki.

Twierdzenie 15.4.1. *Dla każdej formy kwadratowej φ w R^3 można tak dobrać nowy, kartezjański układ wspólrzędnych O_{uvw}, że w tym układzie forma φ ma postać*

$$\varphi\begin{pmatrix} u \\ v \\ w \end{pmatrix} = pu^2 + qv^2 + rw^2,$$

dla pewnych stałych współczynników p, q, r.

Dowód. Macierz m formy φ w układzie O_{xyz} jest symetryczna, a więc z twierdzenia o diagonalizacji, tzn. Twierdzenia 13.3.1, można ją przedstawić w postaci $m = pdp^{-1}$ dla pewnej diagonalnej macierzy $d = \begin{pmatrix} p & 0 & 0 \\ 0 & q & 0 \\ 0 & 0 & r \end{pmatrix}$ oraz pewnej ortogonalnej macierzy p. Rozważmy nowy układ współrzędnych O_{uvw} w R^3, w którym wersorami są wektory $p \cdot e_1, p \cdot e_2, p \cdot e_3$, powstałe poprzez zadziałanie macierzą p na wersory e_1, e_2, e_3. Przypomnijmy, że działanie macierzą ortogonalną na wektory jest przekształceniem ortogonalnym, a więc zachowuje długości wektorów oraz kąty pomiędzy nimi (patrz podrozdział 12.1).Zatem wektory $p{\cdot}e_1, p{\cdot}e_2, p{\cdot}e_3$ są jednostkowe i parami prostopadłe, a więc układ O_{uvw} jest kartezjański. Macierzą przejścia do układu O_{uvw} jest macierz p. Wyznaczmy macierz m' formy φ w nowym układzie współrzędnych. Jak wiemy z Lematu 15.2.1,

$$m' = f^T \cdot m \cdot f,$$

gdzie macierz f to macierz przejścia. W naszym wypadku $f = p$ oraz $m = pdp^{-1}$, stąd

$$m' = p^T pdp^{-1}p.$$

Macierz p jest macierzą ortogonalną dlatego $p^T p = I$. Zatem

$$m' = IdI = d = \begin{pmatrix} p & 0 & 0 \\ 0 & q & 0 \\ 0 & 0 & r \end{pmatrix}.$$

Zatem $\varphi\begin{pmatrix} u \\ v \\ w \end{pmatrix} = pu^2 + qv^2 + rw^2.$ □

Uwaga 15.4.2 (diagonalizacja formy kwadratowej). *Dobranie układu współrzędnych, w którym forma kwadratowa przybiera postać diagonalną nazywamy diagonalizacją formy kwadratowej. Jeśli ponadto nowy układ jest kartezjański, to mówimy o diagonalizacji ortogonalnej.*

Wniosek 15.4.3. *Kwadryka* $\varphi\begin{pmatrix} x \\ y \\ z \end{pmatrix} = 1$ *dla dowolnej formy kwadratowej* φ *w* R^3, *jest jedną z powierzchni wymienionych w poprzednim podrozdziale, być może obróconą w pewien sposób w przestrzeni.*

W dalszej części pokażemy jak praktycznie diagonalizować formy kwadrtowe a także ogólne równania drugiego stopnia. Do tego posłuży nam poniższa uwaga.

Uwaga 15.4.4. *Za układ* O_{uvw} *w Twierdzeniu 15.4.1 można przyjąć dowolny układ, w którym wersorami są jednostkowe parami prostopadłe wektory własne macierzy* m *formy* φ. *Wówczas współczynniki* p, q, r *formy są wartościami własnymi macierzy* m *dla tych wektorów.*

W celu uzasadnienia powyższej uwagi skorzystajmy z dowodu Twierdzenia 15.4.1. Pokazaliśmy tam, że macierz m możemy zdiagonalizować w sposób ortogonalny, przy użyciu pewnej ortogonalnej macierzy p, której postać została określona w Twierdzeniu 13.3.1. Macierz p jest to macierz utworzona z wektorów własnych macierzy m, tak więc rzeczywiście wersorami w nowym układzie, zgodnie z dowodem Twierdzenia 15.4.1, będą wektory własne macierzy m. Diagonalna macierz, która pojawia się przy diagonalizacji macierzy m, jest utworzona z wartości własnych macierzy m, które zgodnie z Twierdzeniem 15.4.1 są współczynnikami formy φ w nowym układzie.

Zobaczmy na przykładzie, jak przebiega diagonalizacja kwadryki.

✔ **Przykład.** Rozważmy kwadrykę zadaną równaniem

$$-x^2 + y^2 - 4yz + 2z^2 = 1.$$

Związana z tym równaniem forma kwadratowa, to

$$\varphi\begin{pmatrix} x \\ y \\ z \end{pmatrix} = -x^2 + y^2 - 4yz + 2z^2,$$

a odpowiadająca jej macierz

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

Dobierzmy teraz nowy układ współrzędnych tak, aby część mieszana w tej formie, czyli $-4yz$ zniknęła. Zgodnie z Uwagą 15.4.4 wersorami tego układu mogą być jednostkowe, parami prostopadłe wektory własne macierzy m. W celu wyznaczenia tych wektorów własnych wyznaczmy równanie charakterystyczne macierzy m. Przedstawia się ono wzorem

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

Przekształcając to równanie otrzymujemy

$$(1-t)(t+2)(t-3) = 0.$$

Tak więc wartościami własnymi tej macierzy są $1, 3, -2$. Znajdźmy teraz odpowiadające tym wartościom przestrzenie własne. Przypomnijmy, że przestrzenią własną dla wartości własnej t nazywamy zbiór $E_t = \{X \in R^3 : mX = t \cdot X\}$, utworzony ze wszystkich wektorów własnych tej wartości własnej. Dla wartości własnej równej 1 jest to zbiór wektorów takich, że

$$\begin{pmatrix} -1 & 0 & -2 \\ 0 & 1 & 0 \\ -2 & 0 & 2 \end{pmatrix}\begin{pmatrix} x \\ y \\ z \end{pmatrix} = 1 \cdot \begin{pmatrix} x \\ y \\ z \end{pmatrix}.$$

Po przekształceniu powyższe równanie przyjmuje postać

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

Jest ono spełnione dla dowolnego wektora postaci $\begin{pmatrix} 0 \\ y \\ 0 \end{pmatrix}$, gdzie y jest dowolną liczbą rzeczywistą. Jednostkowym wektorem własnym dla tej wartości własnej jest np. $X_1 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$.

Podobnie wyznaczamy jednostkowy wektor własny dla wartości własnej 3. Przestrzeń własna dla tej wartości składa się z wektorów spełniających

$$\begin{pmatrix} -1 & 0 & -2 \\ 0 & 1 & 0 \\ -2 & 0 & 2 \end{pmatrix}\begin{pmatrix} x \\ y \\ z \end{pmatrix} = 3 \cdot \begin{pmatrix} x \\ y \\ z \end{pmatrix}.$$

Jak nietrudno się przekonać równość taką spełniają wektory postaci $\begin{pmatrix} x \\ 0 \\ -2x \end{pmatrix}$, dla dowolnej liczby rzeczywistej x. Za jednostkowy wektor własny odpowiadający tej wartości własnej możemy przyjąć $X_2 = \begin{pmatrix} \frac{\sqrt{5}}{5} \\ 0 \\ -\frac{2\sqrt{5}}{5} \end{pmatrix}$. Powtarzając powyższe rozumowanie dla wartości własnej -2, otrzymujemy, że wektory własne to wektory postaci $\begin{pmatrix} 2z \\ 0 \\ z \end{pmatrix}$, tak więc przykładem jednostkowego wektora własnego jest $X_3 = \begin{pmatrix} \frac{2\sqrt{5}}{5} \\ 0 \\ \frac{\sqrt{5}}{5} \end{pmatrix}$.

Macierzą przejścia p do nowego układu współrzędnych O_{uvw}, w którym wersorami są wektory X_1, X_2, X_3 jest

$$p = \begin{pmatrix} 0 & \frac{\sqrt{5}}{5} & \frac{2\sqrt{5}}{5} \\ 1 & 0 & 0 \\ 0 & -\frac{2\sqrt{5}}{5} & \frac{\sqrt{5}}{5} \end{pmatrix}.$$

Macierz diagonalna odpowiadająca naszej formie w nowym układzie to

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

W nowym układzie współrzędnych O_{uvw}, którego wersorami są X_1, X_2, X_3, forma φ przyjmuje postać

$$\varphi \begin{pmatrix} u \\ v \\ w \end{pmatrix} = u^2 + 3v^2 - 2w^2.$$

Znaleźliśmy zatem diagonalną postać kwadryki φ. Kwadryka ta ma dwa współczynniki dodatnie i jeden ujemny, tak więc równanie to odpowiada Przypadkowi (2) spośród omówionych powyżej. Kwadryka ta, w nowym układzie, to zatem hiperboloida jednopowłokowa eliptyczna.

15.5 Inne powierzchnie stopnia drugiego.

Aby mieć listę wszystkich powierzchni stopnia dwa, rozważmy teraz przykłady powierzchni kwadratowych, które nie są kwadrykami.

➢ Przypadek (11), $\mathbf{ax^2 + by^2 + cz - d = 0}$ dla $\mathbf{a > 0, b > 0, c \neq 0}$

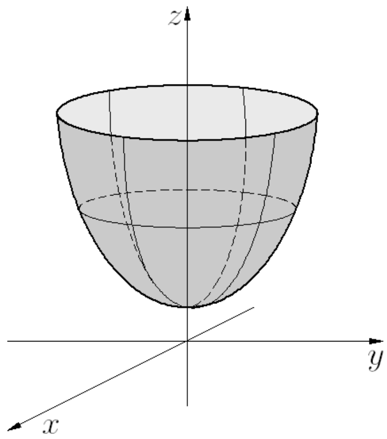
Przekrój tej powierzchni z płaszczyzną poziomą, czyli zadaną równaniem $z = m$, gdzie m jest stałą liczbą, przedstawia się wzorem

$$ax^2 + by^2 = d - m.$$

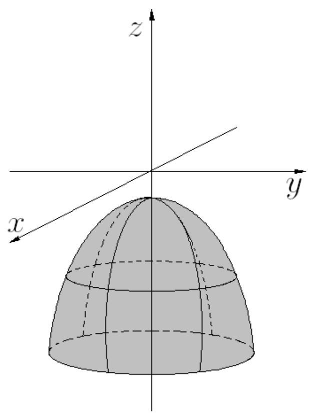
Może to być elipsa lub zbiór pusty. Przekrój z płaszczyzną pionową, tzn. z płaszczyzną zadaną równaniem $x = m$, dla stałej wartości m, przedstawia się wzorem

$$cz = -by^2 + d - am^2,$$

tak więc jest parabolą. Podobnie jest w przypadku przekroju z płaszczyzną $y = m$. Powierzchnia zadana tym równaniem to **paraboloida eliptyczna.**



Podobnie, gdy mamy równanie $\mathbf{ax^2 + by^2 + cz - d = 0}$, a współczynniki spełniają $\mathbf{a < 0, b < 0, c \neq 0}$, to zadana tym równaniem powierzchnia to paraboloida eliptyczna, która jest skierowana w dół.



➢ **Przypadek (12),** $\mathbf{ax^2 + by^2 + cz - d = 0}$ **dla** $\mathbf{a > 0, b < 0, c \neq 0}$

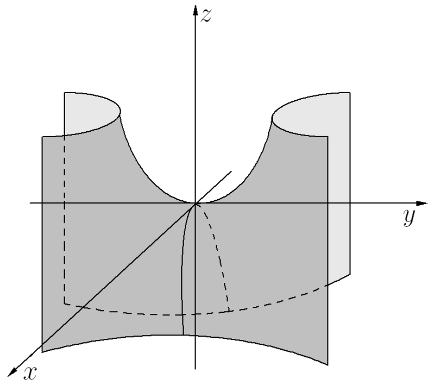
Rozważmy przekrój tej płaszczyzny z płaszczyzną pionową zadaną równaniem $x = m$, gdzie m jest stałą liczbą. Przyjmijmy tutaj, że $c > 0$. Przypadek gdy $c < 0$ można rozważać analogicznie. Przekrój ten przedstawia się wzorem

$$by^2 + cz = d - am^2,$$

po przekształceniu mamy

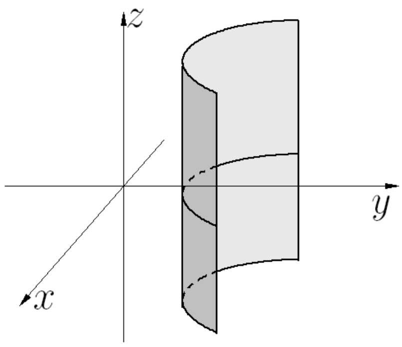
$$cz = -by^2 + d - am^2.$$

W związku z tym, że $b < 0$ a $c > 0$, jest to parabola z ramionami skierowanymi w górę. Podobnie przekrój z płaszczyzną pionową $y = m$ jest parabolą, ale z ramionami skierowanymi w dół.
Równanie powierzchni $\mathbf{ax^2 + by^2 + cz - d = 0}$ dla podanych warunków opisuje **paraboloidę hiperboliczną**, która często bywa nazywana siodłem.



➢ **Przypadek (13),** $\mathbf{ax^2 + by + d = 0}$ **dla** $a \neq 0, b \neq 0$

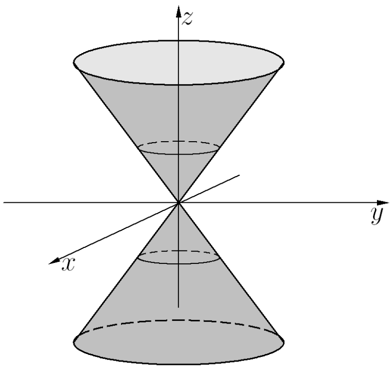
Zadana tym równaniem powierzchnia to **walec paraboliczny**.



Warto zwrócić uwagę na fakt, że postać tej powierzchni nie zależy od znaku współczynników a i b.

➢ **Przypadek (14),** $\mathbf{ax^2 + by^2 + cz^2 = 0}$ **dla** $\mathbf{a > 0, b > 0, c < 0}$

Przekrój tej powierzchni z dowolną płaszczyzną poziomą, tzn. z płaszczyzną o równaniu $z = m$, dla $m \in R$, jest elipsą lub punktem w przypadku $m = 0$. Powstała w ten sposób powierzchnia to **stożek eliptyczny**.



Podobną postać mają formy zadane równaniem $\mathbf{ax^2 + by^2 + cz^2 = 0}$, gdy jedna z liczb a, b, c różni się od znaku dwóch pozostałych. Istotne jest jedynie to, aby współczynniki te były niezerowe.

15.6 Rozpoznawanie krzywych stopnia dwa.

W poprzednich rozdziałach omawialiśmy pewne rodzaje krzywych, okazuje się, że ich znajomość wystarczy do rozpoznania wszystkich powierzchni stopnia dwa. Pokażemy teraz przykłady klasyfikacji krzywych.

✔ **Przykład.** Rozważmy powierzchnię stopnia dwa zadaną równaniem

$$(x-2)^2-(y+1)^2-(2z-3)^2=5.$$

Możemy to równanie przekształcić do postaci

$$\frac{1}{5}(x-2)^2-\frac{1}{5}(y+1)^2-\frac{4}{5}(z-\frac{3}{2})^2=1.$$

Zastosujmy teraz następujące podstawienie

$$u=x-2,\quad v=y+1,\quad w=z-\frac{3}{2}.$$

Równanie powyższe przyjmuje wówczas postać

$$\frac{1}{5}u^2-\frac{1}{5}v^2-\frac{4}{5}w^2=1.$$

Jest to zatem, względem u, v, w, powierzchnia opisana w Przypadku (3), a więc hiperboloida dwupowłokowa eliptyczna. Zauważmy, że układ wyznaczony przez wektory u, v, w powstał z układu x, y, z przez przesunięcie. Więc równanie $(x-2)^2-(y+1)^2-(2z-3)^2=5$ opisuje hiperboloidę dwupowłokową eliptyczną, powstałą przez przesunięcie kwadryki $\frac{1}{5}x^2-\frac{1}{5}y^2-\frac{4}{5}z^2=1$ o wektor $[2,\ -1,\ \frac{3}{2}]$.

Istnieją jednak dużo bardziej skomplikowane powierzchnie, które potrafimy sklasyfikować przy użyciu powierzchni omówionych w Przypadkach (1)-(14). Przykład pokazujący jak rozpoznawać takie powierzchnie pokażemy poniżej.

✔ **Przykład.** Zastanówmy się, jaka powierzchnia w R^3 zadana jest równaniem

$$-x^2+4xy+2y^2-3y+2=0.$$

W tym celu zauważmy, że lewą stronę tej równości możemy podzielić na trzy części:

- część ściśle drugiego stopnia $-x^2+4xy+2y^2$,
- część liniową $-3y$,
- część stałą 2.

Każdą z tych części rozważymy z osobna. Zajmijmy się najpierw częścią kwadratową. Rozważmy związaną z nią pomocniczą formę kwadratową $\varphi\begin{pmatrix}x\\y\\z\end{pmatrix}=-x^2+4xy+2y^2$. Tej formie odpowiada macierz $m=\begin{pmatrix}-1&2&0\\2&2&0\\0&0&0\end{pmatrix}$. Dobierzemy teraz nowy układ współrzędnych tak, aby część mieszana w tej formie, tzn. $4xy$, zniknęła. Zgodnie z Uwagą 15.4.4 wersorami tego układu mogą być jednostkowe, parami prostopadłe wektory własne macierzy m. Znajdźmy zatem te wektory własne. Równanie charakterystyczne macierzy m to

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

Równanie to po przekształceniu możemy zapisać jako

$$t(t+2)(t-3) = 0.$$

Tak więc wartościami własnymi tej macierzy są $-2, 0, 3$. Znajdźmy odpowiadające tym wartościom przestrzenie własne. Przypomnijmy, że przestrzenią własną dla wartości własnej t nazywamy zbiór $E_t = \{X \in R^3 : mX = t \cdot X\}$, utworzony ze wszystkich wektorów własnych tej wartości własnej oraz z wektora zerowego. Dla wartości własnej 0 jest to zatem zbiór takich wektorów, dla których

$$\begin{pmatrix} -1 & 2 & 0 \\ 2 & 2 & 0 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = 0 \cdot \begin{pmatrix} x \\ y \\ z \end{pmatrix}.$$

Po przekształceniu otrzymujemy zatem równanie postaci

$$\begin{pmatrix} -x+2y \\ 2x+2y \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix},$$

które jest spełnione dla wszystkich wektorów postaci $\begin{pmatrix} 0 \\ 0 \\ z \end{pmatrix}$, gdzie z jest dowolną liczbą rzeczywistą. Przestrzenią wektorów własnych dla tej wartości własnej jest zatem prosta

$$E_0 = \{\begin{pmatrix} 0 \\ 0 \\ z \end{pmatrix} : z \in R\}.$$

Jednostkowy wektor własny dla tej wartości własnej to np. $X_1 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$.

Podobnie postępujemy dla wartości własnej równej 3. Przestrzeń własna składa się w tym wypadku z wektorów, które spełniają

$$\begin{pmatrix} -1 & 2 & 0 \\ 2 & 2 & 0 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = 3 \cdot \begin{pmatrix} x \\ y \\ z \end{pmatrix}.$$

Po przekształceniu równość tą możemy zapisać jako:

$$\begin{pmatrix} -x+2y \\ 2x+2y \\ 0 \end{pmatrix} = \begin{pmatrix} 3x \\ 3y \\ 0 \end{pmatrix}.$$

Spełniają ją wszystkie wektory postaci $\begin{pmatrix} x \\ 2x \\ 0 \end{pmatrix}$. Przestrzeń wektorów własnych to

$$E_3 = \{\begin{pmatrix} x \\ 2x \\ 0 \end{pmatrix} : x \in R\}.$$

Jednostkowym wektorem własnym dla tej wartości jest np. $X_2 = \begin{pmatrix} \frac{\sqrt{5}}{5} \\ \frac{2\sqrt{5}}{5} \\ 0 \end{pmatrix}$.

Postępując analogicznie, dla wartości własnej równej -2 otrzymamy, że odpowiadająca jej przestrzeń własna to

$$E_{-2} = \{\begin{pmatrix} -2y \\ y \\ 0 \end{pmatrix} : y \in R\},$$

a przykładowy jednostkowy wektor własny $X_3 = \begin{pmatrix} -\frac{2\sqrt{5}}{5} \\ \frac{\sqrt{5}}{5} \\ 0 \end{pmatrix}$.

Macierzą przejścia do nowego układu O_{uvw}, w którym wersorami są wektory X_1, X_2, X_3 jest

$$p = \begin{pmatrix} 0 & \frac{\sqrt{5}}{5} & -\frac{2\sqrt{5}}{5} \\ 0 & \frac{2\sqrt{5}}{5} & \frac{\sqrt{5}}{5} \\ 1 & 0 & 0 \end{pmatrix}.$$

A macierz diagonalna odpowiadająca naszej formie w nowym układzie to

$$d = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & -2 \end{pmatrix}.$$

W nowym układzie, z wersorami X_1, X_2, X_3, forma kwadratowa φ ma postać

$$\varphi\begin{pmatrix} u \\ v \\ w \end{pmatrix} = 3v^2 - 2w^2.$$

Żeby zorientować się jak przedstawia się część liniowa $-3y$ w nowym układzie O_{uvw}, wystarczy wyrazić x, y, z przez u, v, w i podstawić do tego wyrażenia. Wiemy, że $X = pU$, czyli

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 & \frac{\sqrt{5}}{5} & -\frac{2\sqrt{5}}{5} \\ 0 & \frac{2\sqrt{5}}{5} & \frac{\sqrt{5}}{5} \\ 1 & 0 & 0 \end{pmatrix} \begin{pmatrix} u \\ v \\ w \end{pmatrix}.$$

Otrzymujemy zatem

$$\begin{aligned} x &= \tfrac{\sqrt{5}}{5}v - \tfrac{2\sqrt{5}}{5}w, \\ y &= \tfrac{2\sqrt{5}}{5}v + \tfrac{\sqrt{5}}{5}w, \\ z &= u. \end{aligned}$$

Interesujące nas wyrażenie liniowe to $-3y$, zatem

$$-3y = -3(\frac{2\sqrt{5}}{5}v + \frac{\sqrt{5}}{5}w) = -\frac{6\sqrt{5}}{5}v - \frac{3\sqrt{5}}{5}w.$$

Część stała pozostanie bez zmian. Nasze równanie w nowym układzie współrzędnych ma postać

$$3v^2 - 2w^2 - \frac{6\sqrt{5}}{5}v - \frac{3\sqrt{5}}{5}w + 2 = 0.$$

Po przekształceniu mamy

$$3(v - \frac{3\sqrt{5}}{5})^2 - \frac{27}{5} - 2(w - \frac{3\sqrt{5}}{20})^2 + \frac{9}{40} + 2 = 0,$$

czyli

$$3(v - \frac{3\sqrt{5}}{5})^2 - 2(w - \frac{3\sqrt{5}}{20})^2 = \frac{127}{40}.$$

Po dalszych przekształceniach otrzymujemy

$$\frac{(v - \frac{3\sqrt{5}}{5})^2}{(\sqrt{\frac{127}{120}})^2} - \frac{(w - \frac{3\sqrt{5}}{20})^2}{(\sqrt{\frac{127}{80}})^2} = 1.$$

Spróbujmy teraz ustalić, którą z opisanych w poprzednim podrozdziale form kwadratowych jest to równanie. Zauważmy, że w naszym równaniu jeden ze współczynników jest zerowy, jeden dodatni, a jeden ujemny, jest to zatem walec hiperboliczny.

15.7 Formy kwadratowe oraz macierze dodatnio określone.

Poznamy teraz pojęcia, które wykorzystywane są w analizie wielu zmiennych przy badaniu minimum i maksimum funkcji. Dla funkcji wielu zmiennych zamiast badać znak drugiej pochodnej w punkcie, gdzie pierwsza pochodna się zeruje, bada się formę kwadratową. Dodatnia określoność tej formy, w punkcie gdzie pierwsza pochodna cząstkowa się zeruje oznacza, że punkt ten może być ekstremum funkcji.

W tym podrozdziale podamy definicje formy oraz macierzy dodatnio określonej. W kolejnym poznamy ich najważniejsze własności. Nauczymy się też w łatwy sposób rozpoznawać macierze dodatnio określone.

Definicja 15.7.1 (forma kwadratowa dodatnio określona). Forma kwadratowa φ jest dodatnio określona, jeśli spełnia którykolwiek z następujących równoważnych warunków:

(1) po diagonalizacji forma ta ma postać

$$\varphi\begin{pmatrix} u \\ v \\ w \end{pmatrix} = pu^2 + qv^2 + rw^2,$$

gdzie $p > 0, q > 0, r > 0$,

(2) dla dowolnego wektora $X \neq O$ z R^3 mamy $\varphi(X) > 0$.

Udowodnimy teraz, że te dwa warunki rzeczywiście są równoważne.

Dowód równoważności warunków (1) *i* (2).
Na początek pokażemy, że jeśli forma spełnia warunek pierwszy, to musi też spełniać warunek drugi. Weźmy dowolny, niezerowy wektor X, który w nowym układzie, po diagonalizacji formy φ, ma współrzędne $X = \begin{pmatrix} u \\ v \\ w \end{pmatrix}$. Wówczas forma φ na wektorze X przyjmuje wartość $\varphi(X) = pu^2 + qv^2 + rw^2$. Wiemy, że wszystkie trzy składniki tej sumy są nieujemne oraz, że przynajmniej jeden z nich jest dodatni, bo wektor X jest niezerowy. Otrzymaliśmy zatem, że $\varphi(X) > 0$ dla dowolnego wektora $X \neq 0$.
Pokażemy teraz implikację przeciwną, tzn. że z warunku drugiego wynika warunek pierwszy. Zakładamy zatem, że $\varphi(X) > 0$ dla dowolnego niezerowego wektora X. Forma φ po diagonalizacji ma postać $\varphi = pu^2 + qv^2 + rw^2$. Dla wektora $X = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$ mamy $\varphi\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} = p \cdot 1^2 + q \cdot 0^2 + r \cdot 0^2 = p$. Wiemy, że foma φ przyjmuje wartości dodatnie dla dowolnego wektora X, więc w szczególności $\varphi\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} > 0$ zatem $p > 0$. Przeprowadzając podobne rozumowanie z użyciem wektorów $\begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$, $\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$, możemy pokazać, że $q > 0$ oraz $r > 0$, z czego wynika, że spełniony jest warunek pierwszy naszej definicji.
Pokazaliśmy, że jeżeli zachodzi warunek pierwszy, to musi też zachodzić warunek drugi, jak również, że jeśli zachodzi warunek drugi, to musi zachodzić warunek pierwszy. Udowodniliśmy zatem równoważność tych dwóch warunków. □

Definicja 15.7.2 (macierz dodatnio określona). Macierz symetryczna jest dodatnio określona, jeśli forma kwadratowa φ zadana tą macierzą jest dodatnio określona.

✔ **Przykład.** Macierz $m = \begin{pmatrix} 3 & 0 & 0 \\ 0 & 5 & 0 \\ 0 & 0 & 11 \end{pmatrix}$ jest dodatnio określona ponieważ forma kwadratowa zadana tą macierzą, czyli $\varphi = 3x^2 + 5y^2 + 11z^2$, jak wynika z Definicji 15.7.1, jest dodatnio określona.

15.8 Własności form i macierzy dodatnio określonych.

Własność 15.8.1 (minimum formy dodatnio określonej). *Forma dodatnio określona φ potraktowana jako funkcja przestrzeni $R^3 \to R^3$, ma minimum w punkcie*

$X_0 = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}.$

Dowód. W punkcie $X_0 = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$ forma φ przyjmuje wartość 0. Wiemy również, że dla każdego niezerowego wektora X mamy $\varphi(X) > 0$, tak więc w punkcie X_0 funkcja ma minimum. □

Własność 15.8.2 (wartości własne macierzy dodatnio określonej). *Macierz dodatnio określona ma tylko dodatnie wartości własne.*

Dowód. Macierz m jest symetryczna, więc posiada ortogonalną diagonalizację $m = pdp^{-1}$, gdzie p jest macierzą ortogonalną, natomiast d diagonalną, $d = \begin{pmatrix} t_1 & 0 & 0 \\ 0 & t_2 & 0 \\ 0 & 0 & t_3 \end{pmatrix}$. Wyrazy t_1, t_2, t_3 są przy tym wartościami własnymi macierzy m. Ponieważ dla macierzy ortogonalnej mamy $p^{-1} = p^T$, zatem $m = pdp^T$. Oznacza to, że forma kwadratowa φ zadana macierzą m ma w nowym układzie O_{uvw} z macierzą przejścia p postać diagonalną

$$t_1u^2 + t_2v^2 + t_3w^2.$$

Macierz m, a więc i forma φ są dodatnio określone, więc $t_1 > 0, t_2 > 0, t_3 > 0$, a to są wszystkie wartości własne macierzy m. □

Poniżej podamy lemat, który okaże się bardzo przydatny przy sformułowaniu kryterium pozwalającego na szybkie sprawdzenie dodatniej określoności macierzy bez konieczności zmiany układu współrzędnych.

Lemat 15.8.3. *Jeśli macierz* $m = \begin{pmatrix} a & b & c \\ b & d & e \\ c & e & f \end{pmatrix}$ *jest dodatnio określona, to*

(1) $a > 0$,

(2) $\det \begin{pmatrix} a & b \\ b & d \end{pmatrix} > 0$,

(3) $\det(m) > 0$.

Dowód.

(1) Rozważmy wektor $X = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$. Z dodatniej określoności macierzy m wiemy, że również forma nią zadana $\varphi \begin{pmatrix} x \\ y \\ z \end{pmatrix} = ax^2 + dy^2 + fz^2 + 2bxy + 2cxz + 2eyz$ jest dodatnio określona, a więc $\varphi \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} > 0$. Mamy również $\varphi \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} = a \cdot 1^2 + 0 + 0 + 0 + 0 + 0 =$

a, zatem $a > 0$.

(2) Rozważmy wektory postaci $\begin{pmatrix} x \\ y \\ 0 \end{pmatrix}$. Dla takich wektorów mamy $\varphi\begin{pmatrix} x \\ y \\ 0 \end{pmatrix} = ax^2 + 2bxy + dy^2$. Jest to forma kwadratowa w R^2 , w układzie O_{xy}. Dla form kwadratowych w R^2 zachodzą analogiczne rezultaty jak dla form w R^3. Przykładowo macierz symetryczna $\begin{pmatrix} a & b \\ b & d \end{pmatrix}$ diagonalizuje się w postaci pdp^{-1}, gdzie $d = \begin{pmatrix} t_1 & 0 \\ 0 & t_2 \end{pmatrix}$, zaś p jest ortogonalną macierzą 2×2. Ponieważ φ jest dodatnio określona, dla dowolnego niezerowego wektora $\begin{pmatrix} x \\ y \\ 0 \end{pmatrix}$ mamy $\varphi\begin{pmatrix} x \\ y \\ 0 \end{pmatrix} > 0$. Stąd analogicznie jak dla form w R^3 mamy $t_1 > 0, t_2 > 0$. Wówczas $\det\begin{pmatrix} a & b \\ b & d \end{pmatrix} = \det(pdp^{-1}) = \det(p) \cdot \det\begin{pmatrix} t_1 & 0 \\ 0 & t_2 \end{pmatrix} \cdot \det(p^{-1}) = \det(p) \cdot (t_1t_2) \cdot \frac{1}{\det(p)} = t_1t_2 > 0$.

(3) Z twierdzenia o diagonalizacji oraz z Własności 15.8.2 mamy $m = pdp^{-1}$, gdzie $d = \begin{pmatrix} t_1 & 0 & 0 \\ 0 & t_2 & 0 \\ 0 & 0 & t_3 \end{pmatrix}$ oraz $t_1 > 0, t_2 > 0, t_3 > 0$. Otrzymujemy zatem

$$\det(m) = \det(pdp^{-1}) = \det(p)\det(d)\det(p^{-1}) = \det(p)\det(d)\frac{1}{\det(p)} =$$

$$= \det(d) = \det\begin{pmatrix} t_1 & 0 & 0 \\ 0 & t_2 & 0 \\ 0 & 0 & t_3 \end{pmatrix} = t_1t_2t_3 > 0.$$

□

Lemat 15.8.4 (kryterium dodatniej określoności macierzy). *Macierz* $m = \begin{pmatrix} a & b & c \\ b & d & e \\ c & e & f \end{pmatrix}$, *jest dodatnio określona wtedy i tylko wtedy gdy spełnione są następujące trzy warunki:*

(1) $a > 0$,

(2) $\det\begin{pmatrix} a & b \\ b & d \end{pmatrix} > 0$,

(3) $\det(m) > 0$.

Dowód.
W Lemacie 15.8.3 pokazaliśmy, że dodatnia określoność macierzy m pociąga warunki (1),(2) i (3). Potrzebujemy więc udowodnić implikację w drugą stronę, tzn. że z warunków (1),(2),(3) wynika dodatnia określoność macierzy m. Dzięki diagonalizacji, w pewnym układzie O_{uvw} forma φ zadana macierzą m ma postać $pu^2+qv^2+rw^2$. Z warunku (3) wiemy, że $\det(m) = pqr > 0$. Są więc dwie możliwości:

(i) $p > 0, q > 0, r > 0,$

(ii) dwie spośród tych liczb są ujemne a trzecia dodatnia.

Wykluczmy drugą możliwość. Załóżmy, że $p < 0, q < 0, r > 0$. Wówczas dla niezerowych wektorów postaci $U = \begin{pmatrix} u \\ v \\ 0 \end{pmatrix}$, czyli niezerowych wektorów z płaszczyzny $w = 0$ w układzie O_{uvw}, mamy $\varphi \begin{pmatrix} u \\ v \\ 0 \end{pmatrix} = pu^2 + qv^2 < 0.$

Tak więc mamy całą płaszczyznę na której forma jest ujemna na niezerowych wektorach.

Rozważmy teraz płaszczyznę $z = 0$ w R^3, oraz układ O_{xy} na tej płaszczyźnie. Dla wektorów $\begin{pmatrix} x \\ y \end{pmatrix}$ z tej płaszczyzny, czyli wektorów $\begin{pmatrix} x \\ y \\ 0 \end{pmatrix}$ z R^3, forma φ jest zadana macierzą $\begin{pmatrix} a & b \\ b & d \end{pmatrix}$. Po diagonalizacji tej macierzy dostajemy $\begin{pmatrix} t_1 & 0 \\ 0 & t_2 \end{pmatrix}$. Z warunku (2) wiemy, że $\det \begin{pmatrix} a & b \\ b & d \end{pmatrix} > 0$, a więc

$$0 < \det \begin{pmatrix} a & b \\ b & d \end{pmatrix} = \det \begin{pmatrix} t_1 & 0 \\ 0 & t_2 \end{pmatrix} = t_1 t_2.$$

Tym razem również mamy dwie możliwości, gdy t_1, t_2 są jednocześnie dodatnie lub ujemne. Jeśli zachodzi ten drugi przypadek, to w nowym układzie $O_{x'y'}$ forma ma postać $t_1(x')^2 + t_2(y')^2$ i jest ujemnie określona, tzn. $\varphi(X) < 0$, dla niezerowych wektorów X z tej płaszczyzny. Wiemy jednak, że $a > 0$, a zatem $\varphi \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} > 0$ i otrzymujemy sprzeczność. Ten przypadek nie może zatem zachodzić. Mamy więc, że t_1, t_2 są dodatnie, zatem dla niezerowych wektorów X z tej płaszczyzny $\varphi(X) > 0$. Jak nietrudno zauważyć, płaszczyzny $w = 0$ oraz $z = 0$ przecinają się wzdłuż pewnej prostej, dla wektorów z tej prostej mamy zatem $\varphi(X) < 0$ oraz $\varphi(X) > 0$ co daje sprzeczność. W ten sposób wykluczyliśmy przypadek (2), tak więc mamy $p > 0, q > 0, r > 0$, co jest równoważne dodatniej określoności macierzy m. $\square$

Rozdział 16

Układy równań liniowych z wieloma niewiadomymi

W rozdziale tym poznamy jedną z najszybszych metod rozwiązywania układów równań liniowych - metodę eliminacji Gaussa. Jest ona dużo bardziej efektywna niż podana w Twierdzeniu 4.4.2 metoda zwana metodą Cramera, w której konieczne jest liczenie wyznaczników, co bywa bardzo czasochłonne.

W pierwszym podrozdziale na przykładzie poznamy tę metodę, w kolejnym poznamy sposób kodowania układu równań tak, aby dużo łatwiej można było korzystać z algorytmu Gaussa. W dalszej części pokażemy, w jaki sposób można wykorzystać tę metodę do rozwiązania zagadnienia geometrycznego.

16.1 Metoda eliminacji Gaussa.

Nową metodę rozwiązywania układów równań wprowadzimy na przykładzie konkretnego układu równań. Oznaczmy nasz układ przez $U1$, natomiast poszczególne równania przez I, II, III. Układ $U1$ ma postać:

$$U1: \quad \begin{cases} x + y - 5z = -14 & {\scriptstyle (I)} \\ 2x + 3y + 2z = 7 & {\scriptstyle (II)} \\ -3x + 4y + z = -7 & {\scriptstyle (III)} \end{cases}$$

Przekształćmy ten układ. Otrzymamy nowy układ $U2$ zbudowany z trzech równań I', II', III'. Równanie I pozostawimy bez zmian, co symbolicznie zapiszemy przez $I' = I$. Równanie II' powstanie przez odjęcie od równania II równania I pomnożonego przez 2. Równanie III' powstanie przez dodanie do równania III równania I pomnożonego przez 3. W ten sposób w równaniach II' i III' nie będzie już występować zmienna x. W przyszłości operacje jakie będziemy wykonywać na równaniach, będziemy zapisywać w układzie równań, tak jak to zostało zrobione poniżej.

$$U2: \quad \begin{cases} {\scriptstyle (I'=I)} & x & + & y & - & 5z & = & -14 \\ {\scriptstyle (II'=II-2I)} & & & y & + & 12z & = & 35 \\ {\scriptstyle (III'=III+3I)} & & & 7y & - & 14z & = & -49 \end{cases}$$

Zauważmy, że jeśli jakaś trójka liczb x_0, y_0, z_0 spełnia układ $U1$, to spełnia również układ $U2$. I na odwrót, jeśli x_0, y_0, z_0 spełnia $U2$, to spełnia też $U1$, bo

$II = II' + 2I'$ oraz $III = III' - 3I'$. Zatem układy $U1$ i $U2$ mają te same zbiory rozwiązań. Chcąc rozwiązać układ $U1$ wystarczy rozwiązać układ $U2$. Możemy uogólnić ten wniosek. Dwa układy, które można uzyskać jeden z drugiego i drugi z pierwszego przez operacje podobne, jak w naszym przykładzie, nazywamy układami równoważnymi.

Uwaga 16.1.1. *Równoważne układy równań mają te same zbiory rozwiązań.*

Aby rozwiązać układ równań, wystarczy przekształcić go w prostszy, ale równoważny układ równań i rozwiązać ten prostszy. Wróćmy do naszego przykładu. Przekształcamy $U2$ tak, aby wyeliminować niewiadomą y z trzeciego równania. Otrzymujemy nowy układ $U3$.

$$U3: \left\{ \begin{array}{lccccccc} {\scriptstyle (I''=I')} & x & + & y & - & 5z & = & -14 \\ {\scriptstyle (II''=II')} & & & y & + & 12z & = & 35 \\ {\scriptstyle (III'=III'-7II')} & & & & - & 98z & = & -294 \end{array} \right.$$

Układ $U3$ możemy już bardzo łatwo rozwiązać. Z trzeciego równania mamy bowiem

$$z = \frac{-294}{-98} = 3.$$

Możemy teraz wyznaczyć pozostałe niewiadome. Z drugiego równania obliczamy

$$y = 35 - 12z = 35 - 36 = -1.$$

Z trzeciego równania mamy

$$x = 14 - y + 5z = -14 + 1 + 15 = 2.$$

Układ $U3$ ma dokładnie jedno rozwiązanie

$$x = 2, y = -1, z = 3.$$

16.2 Metoda modyfikacji macierzy reprezentujących układ równań.

W praktyce metodę zaprezentowaną w poprzednim podrozdziale upraszcza się. Zamiast na równaniach możemy bowiem operować na macierzach. Sposób ten znów poznamy poprzez rozwiązanie przykładowego układu równań. Weźmy następujący układ $U1$:

$$U1: \left\{ \begin{array}{c} x + 2y + z + 3t = 5 \\ 3x + 4y - 2z + 5t = -11 \\ -2x + 2y + 5z + t = 28 \\ 2x - 6y + 3z - t = 10 \end{array} \right.$$

Układowi temu przyporządkowujemy tablicę, której wiersze stanowią współczynniki w poszczególnych równaniach, a kolumny są utworzone przez współczynniki przy poszczególnych zmiennych w kolejnych równaniach. Dla układu $U1$ jest to zatem tablica $T1$

$$T1: \quad \begin{array}{lrrrrr} \scriptstyle(I) & 1 & 2 & 1 & 3 & 5 \\ \scriptstyle(II) & 3 & 4 & -2 & 5 & -11 \\ \scriptstyle(III) & -2 & 2 & 5 & 1 & 28 \\ \scriptstyle(IV) & 2 & -6 & 3 & -1 & 10 \end{array}$$

Możemy teraz wykonywać działania na tej tablicy, podobnie jak robiliśmy to z układem równań.

$$T2: \quad \begin{array}{lrrrrr} \scriptstyle(I'=I) & 1 & 2 & 1 & 3 & 5 \\ \scriptstyle(II'=II-3I) & 0 & -2 & -5 & -4 & -26 \\ \scriptstyle(III'=III+2I) & 0 & 6 & 7 & 7 & 38 \\ \scriptstyle(IV'=IV-2I) & 0 & -10 & 1 & -7 & 0 \end{array}$$

$$T3: \quad \begin{array}{lrrrrr} \scriptstyle(I''=I') & 1 & 2 & 1 & 3 & 5 \\ \scriptstyle(II''=II') & 0 & -2 & -5 & -4 & -26 \\ \scriptstyle(III''=III'+3II') & 0 & 0 & -8 & -5 & -40 \\ \scriptstyle(IV''=IV'-5II') & 0 & 0 & 26 & 13 & 130 \end{array}$$

$$T4: \quad \begin{array}{lrrrrr} \scriptstyle(I'''=I'') & 1 & 2 & 1 & 3 & 5 \\ \scriptstyle(II'''=II'') & 0 & -2 & -5 & -4 & -26 \\ \scriptstyle(III'''=III'') & 0 & 0 & -8 & -5 & -40 \\ \scriptstyle(IV'''=IV'':13) & 0 & 0 & 2 & 1 & 10 \end{array}$$

$$T5: \quad \begin{array}{lrrrrr} \scriptstyle(I'^{\nu}=I''') & 1 & 2 & 1 & 3 & 5 \\ \scriptstyle(II'^{\nu}=II''') & 0 & -2 & -5 & -4 & -26 \\ \scriptstyle(III'^{\nu}=IV''') & 0 & 0 & 2 & 1 & 10 \\ \scriptstyle(IV'^{\nu}=III''') & 0 & 0 & -8 & -5 & -40 \end{array}$$

$$T6: \quad \begin{array}{lrrrrr} \scriptstyle(I^{\nu}=I^{\nu}) & 1 & 2 & 1 & 3 & 5 \\ \scriptstyle(II^{\nu}=II^{\nu}) & 0 & -2 & -5 & -4 & -26 \\ \scriptstyle(III^{\nu}=III'^{\nu}) & 0 & 0 & 2 & 1 & 10 \\ \scriptstyle(IV^{\nu}=IV'^{\nu}+4III'^{\nu}) & 0 & 0 & 0 & -1 & 0 \end{array}$$

Przekształcenia tablicy odpowiadają przekształceniom układu równań. Możemy teraz odtworzyć końcowy układ równań po przekształceniach.

$$\left\{ \begin{array}{rcrcrcrcl} x & + & 2y & + & z & + & 3t & = & 5 \\ & - & 2y & - & 5z & - & 4t & = & -26 \\ & & & & 2z & + & t & = & 10 \\ & & & & & - & t & = & 0 \end{array} \right.$$

Rozwiązanie takiego układu możemy już bardzo łatwo znaleźć:
$t = 0$
$2z = 10 - t = 10$ stąd $z = 5$
$-2y = -26 + 5z + 4t = -1$ stąd $y = -\frac{1}{2}$
$x = 5 - 2y - z - 3t = -1$

Rozwiązaniem naszego układu są

$$x = -1,\ y = -\frac{1}{2},\ z = 5,\ t = 0.$$

Przekształcając układ równań lub odpowiadającą jemu tablicę, staramy się wyeliminować jak największą liczbę niewiadomych w równaniach, co w języku operacji na tablicach oznacza wyzerowanie jak największej liczby współczynników. Taki zabieg, jak można było zauważyć powyżej, bardzo ułatwia znalezienie rozwiązania.

Początkowo staramy się wyeliminować pierwszą niewiadomą z jak największej liczby równań. W tym celu w tablicy zerujemy współczynniki w pierwszej kolumnie, która jak wiemy odpowiada pierwszej niewiadomej, dokonujemy tego poprzez odjęcie od kolejnych wierszy tablicy, zaczynając od drugiego wiersza, odpowiednich wielokrotności pierwszego. Następnie zerujemy współczynniki w drugiej kolumnie, odejmując od kolejnych wierszy, zaczynając tym razem od trzeciego wiersza, odpowiednie wielokrotności drugiego. Analogicznie postępujemy z kolejnymi kolumnami. Ostatecznie otrzymamy tablicę, w której znaczna część współczynników jest równa zero. Po odtworzeniu układu w ostatnim równaniu będziemy mieć jedną niewiadomą, o ile układ ma jednoznaczne rozwiązanie. W przedostatnim równaniu będą dwie niewiadome itd. Taki układ możemy już bardzo łatwo rozwiązać, wyznaczając kolejne niewiadome, zaczynając od ostatniego równania.

16.3 Wieloparametrowe rodziny rozwiązań.

W tym podrozdziale, na przykładzie, poznamy typowe zjawisko, z jakim możemy spotkać się przy rozwiązywaniu układów równań, które mają nieskończenie wiele rozwiązań. Rozważmy układ równań $U1$.

$$U1: \begin{cases} 2x + 3y + z + 5t - u = 2 \\ -2x - 3y + 2z - 6t + 5u = 0 \\ 4x + 6y + 8z + 9t + 6u = 1 \end{cases}$$

Zauważmy, że w układzie tym występuje więcej niewiadomych niż jest równań. W takim wypadku rozwiązań będzie na ogół nieskończenie wiele i będą one zależne od parametrów.
Układowi temu odpowiada tablica współczynników $T1$.

$$T1: \begin{array}{crrrrrr} {\scriptstyle (I)} & 2 & 3 & 1 & 5 & -1 & 2 \\ {\scriptstyle (II)} & -2 & -3 & 2 & -6 & 5 & 0 \\ {\scriptstyle (III)} & 4 & 6 & 8 & 9 & 6 & 1 \end{array}$$

Przekształćmy tę tablicę zgodnie z algorytmem poznanym w poprzednim podrozdziale.

$$T2: \quad \begin{array}{lrrrrrr} {\scriptstyle (I'=I)} & 2 & 3 & 1 & 5 & -1 & 2 \\ {\scriptstyle (II'=II+I)} & 0 & 0 & 3 & -1 & 4 & 2 \\ {\scriptstyle (III'=III-2I)} & 0 & 0 & 6 & -1 & 8 & -3 \end{array}$$

$$T3: \quad \begin{array}{lrrrrrr} {\scriptstyle (I''=I')} & 2 & 3 & 1 & 5 & -1 & 2 \\ {\scriptstyle (II''=II')} & 0 & 0 & 3 & -1 & 4 & 2 \\ {\scriptstyle (III''=III'-2II')} & 0 & 0 & 0 & 1 & 0 & -7 \end{array}$$

Zauważmy, że dalsze przekształcenia nie mają już sensu, gdyż nie uda nam się już w żadnym wierszu wyeliminować niewiadomych. Możemy teraz odtworzyć układ równań.

$$U3: \quad \begin{cases} 2x + 3y + z + 5t - u = 2 \\ 3z - t + 4u = 2 \\ t = -7 \end{cases}$$

Zauważmy, że w pierwszym wierszu tablicy $T3$ pierwszy niezerowy współczynnik odpowiada zmiennej x, w drugim wierszu pierwszy niezerowy współczynnik odpowiada zmiennej z, w trzecim wierszu zmiennej t. Wyznaczmy zatem x, z, t z układu $U3$, traktując zmienne y, u jak parametry.
$t = -7$
$3z = 2 + t - 4u = 2 - 7 - 4u = -5 - 4u$ zatem $z = -\frac{5}{3} - \frac{4}{3}u$
$2x = 2 - 3y - z - 5t + u = 2 - 3y + \frac{5}{3} + \frac{4}{3}u + 35 + u = 38\frac{2}{3} - 3y + 2\frac{1}{3}u$ więc $x = 19\frac{1}{3} - 1\frac{1}{2}y + 1\frac{1}{6}u$
Zbiór rozwiązań naszego układu to zatem

$$\begin{cases} x = 19\frac{1}{3} - 1\frac{1}{2}y + 1\frac{1}{6}u \\ y = y \\ z = -\frac{5}{3} - \frac{4}{3}u \\ t = -7 \\ u = u \end{cases}$$

Zauważmy, że w równościach

$$y = y, \ u = u$$

y, u pojawiające się po lewej stronie równości oznaczają zmienne, natomiast po prawej stronie występują w charakterze parametrów. Dla przejrzystości zapisu lepiej będzie jeśli wprowadzimy dodatkowe parametry $\alpha, \ \mu \in R^3$. Nasz układ przyjmuje wówczas postać

$$U: \quad \begin{cases} x = 19\frac{1}{3} - \frac{3}{2}\alpha + 1\frac{1}{6}\mu \\ y = \alpha \\ z = -\frac{5}{3} - \frac{4}{3}\mu \\ t = -7 \\ u = \mu \end{cases}$$

Układ ten ma nieskończenie wiele rozwiązań zależnych od parametrów α, μ, parametry te przyjmują wszystkie wartości rzeczywiste. Dla każdej pary α, μ układ liczb U jest rozwiązaniem układu $U1$. Są to wszystkie rozwiązania układu $U1$.

16.4 Układ sprzeczny.

Układ równań sprzeczny to taki układ, który nie ma rozwiązań. Przykład takiego układu przedstawiony został poniżej.

$$U1: \begin{cases} x + 2y + z = 3 \\ 5x - 3y + 4z = 1 \\ 3x - 7y + 2z = -1 \end{cases}$$

Przekonajmy się, że układ ten rzeczywiście jest sprzeczny. Spróbujmy go rozwiązać metodą eliminacji. Odpowiadająca mu tablica to

$$T1: \begin{array}{lrrrr} {\scriptstyle (I)} & 1 & 2 & 1 & 3 \\ {\scriptstyle (II)} & 5 & -3 & 4 & 1 \\ {\scriptstyle (III)} & 3 & -7 & 2 & -1 \end{array}$$

Przekształćmy tę tablicę stosując poznaną wcześniej metodę. Otrzymujemy wówczas

$$T2: \begin{array}{lrrrr} {\scriptstyle (I'=I)} & 1 & 2 & 1 & 3 \\ {\scriptstyle (II'=II-5I)} & 0 & -13 & -1 & -14 \\ {\scriptstyle (III'=III-3I)} & 0 & -13 & -1 & -10 \end{array}$$

$$T3: \begin{array}{lrrrr} {\scriptstyle (I''=I')} & 1 & 2 & 1 & 3 \\ {\scriptstyle (II''=II')} & 0 & -13 & -1 & -14 \\ {\scriptstyle (III''=III'-II')} & 0 & 0 & 0 & 4 \end{array}$$

Możemy teraz odtworzyć układ równań.

$$U3: \begin{cases} x + 2y + z = 3 \\ \quad -13y - z = -14 \\ \quad\quad 0 = 4 \end{cases}$$

Otrzymaliśmy sprzeczność w trzecim równaniu. Te trzy równania nie mogą być zatem jednocześnie spełnione dla żadnej trójki liczb x, y, z.

16.5 Układ zależny.

Rozważmy teraz następujący układ równań $U1$:

$$U1: \begin{cases} x + 2y + z = -5 \\ 5x - 3y + 4z = 15 \\ 2x + 5y - 9z = -2 \\ -x - 4y + 5z = 5 \end{cases}$$

Odpowiadająca temu układowi tablica $T1$ to

$$T1: \begin{array}{lrrrr} {\scriptstyle (I)} & 1 & 2 & 1 & -5 \\ {\scriptstyle (II)} & 5 & -3 & 4 & 15 \\ {\scriptstyle (III)} & 2 & 5 & -9 & -2 \\ {\scriptstyle (IV)} & -1 & -4 & 5 & 5 \end{array}$$

Możemy tę tablicę przekształcić.

$$T2: \quad \begin{array}{lrrrr} {\scriptstyle (I'=I)} & 1 & 2 & 1 & -5 \\ {\scriptstyle (II'=II-5I)} & 0 & -13 & -1 & 40 \\ {\scriptstyle (III'=III-2I)} & 0 & 1 & -11 & 8 \\ {\scriptstyle (IV'=IV+I)} & 0 & -2 & 6 & 0 \end{array}$$

$$T3: \quad \begin{array}{lrrrr} {\scriptstyle (I''=I)} & 1 & 2 & 1 & -5 \\ {\scriptstyle (II''=III')} & 0 & 1 & -11 & 8 \\ {\scriptstyle (III''=II')} & 0 & -13 & -1 & 40 \\ {\scriptstyle (IV''=IV')} & 0 & -2 & 6 & 0 \end{array}$$

$$T4: \quad \begin{array}{lrrrr} {\scriptstyle (I'''=I'')} & 1 & 2 & 1 & -5 \\ {\scriptstyle (II'''=II'')} & 0 & 1 & -11 & 8 \\ {\scriptstyle (III'''=III''+13II'')} & 0 & 0 & -144 & 144 \\ {\scriptstyle (IV'''=IV''+2II'')} & 0 & 0 & -16 & 16 \end{array}$$

$$T5: \quad \begin{array}{lrrrr} {\scriptstyle (I^{IV}=I''')} & 1 & 2 & 1 & -5 \\ {\scriptstyle (II^{IV}=II''')} & 0 & 1 & -11 & 8 \\ {\scriptstyle (III^{IV}=III^{III}:144)} & 0 & 0 & -1 & 1 \\ {\scriptstyle (IV^{IV}=IV''':16)} & 0 & 0 & -1 & 1 \end{array}$$

$$T6: \quad \begin{array}{lrrrr} {\scriptstyle (I^{V}=I^{IV})} & 1 & 2 & 1 & -5 \\ {\scriptstyle (II^{V}=II^{IV})} & 0 & 1 & -11 & 8 \\ {\scriptstyle (III^{V}=III^{IV})} & 0 & 0 & -1 & 1 \\ {\scriptstyle (IV^{V}=IV^{IV}-III^{IV})} & 0 & 0 & 0 & 0 \end{array}$$

Możemy teraz odtworzyć układ równań

$$U6: \quad \left\{ \begin{array}{rcrcrcl} x & + & 2y & + & z & = & -5 \\ & & y & - & 11z & = & 8 \\ & & & & -z & = & 1 \\ & & & & 0 & = & 0 \end{array} \right.$$

Ostatnie równanie występujące w tym układzie to tożsamość. Możemy je opuścić, gdyż jako zawsze spełnione, nie wnosi istotnych zmian do naszego układu. Układ, w którym jedno z równań można opuścić bez zmiany zbioru rozwiązań (lub każdy układ równoważny z takim), nazywamy układem zależnym. Układ taki możemy rozwiązać przechodząc do nowego układu z mniejszą liczbą równań. W przypadku układu $U6$, po odrzuceniu ostatniego równania otrzymujemy równoważny układ

$$U7: \quad \left\{ \begin{array}{rcrcrcl} x & + & 2y & + & z & = & -5 \\ & & y & - & 11z & = & 8 \\ & & & & -z & = & 1 \end{array} \right.$$

Możemy teraz rozwiązać ten układ.
$-z = 1$ więc $z = -1$
$y = 8 + 11z = 8 - 11 = -3$
$x = -5 - 2y - z = -5 + 6 + 1 = 2$
Rozwiązaniem naszego układu jest zatem trójka liczb $x = 2, y = -3, z = -1$.

16.6 Operacje przekształcające dany układ na równoważny.

W poprzednich podrozdziałach wykonywaliśmy operacje na układach równań i odpowiadających im tablicach, przekształcając je w taki sposób, aby otrzymać układy równoważne. Wymieńmy teraz, jakie operacje są dozwolone przy takich przekształceniach. Ograniczymy się do wymienienia operacji, jakie możemy wykonać na tablicy reprezentującej układ. W nawiasach podane zostaną odpowiadające tym operacją działania jakie możemy wykonywać na układach równań.

(1) Zamiana kolejności wierszy w tablicy (co odpowiada zmianie kolejności równań w układzie).

(2) Pomnożenie wiersza tablicy przez niezerową stałą (co odpowiada pomnożeniu obu stron równania przez tą stałą).

(3) Dodanie do jednego wiersza wielokrotności innego (co odpowiada dodaniu do jednego równania wielokrotności innego).

(4) Pominięcie wiersza złożonego z samych zer (co odpowiada pominięciu równania postaci $0 = 0$).

(5) Pominięcie wiersza, jeśli po przemnożeniu przez stałą jest on równy innemu wierszowi (co odpowiada pominięciu jednego równania, jeśli po przemnożeniu przez liczbę jest ono równe innemu równaniu).

(6) Zmiana kolejności kolumn w tablicy. Nie możemy jednak przestawiać ostatniej kolumny macierzy (co odpowiada zmianie kolejności pisania niewiadomych we wszystkich równaniach).

Uwaga 16.6.1. *Przy odtwarzaniu układu równań po modyfikacji numer sześć musimy pamiętać o zmianie kolejności niewiadomych.*

16.7 Zastosowanie metody eliminacji w geometrii.

W tym podrozdziale pokażemy, jak można wykorzystywać układy równań przy rozwiązywaniu zagadnień geometrycznych. Te same problemy można rozwiązać w sposób geometryczny, przedstawiona tutaj metoda jest jednak metodą czysto algebraiczną, a tym samym najbardziej bezpośrednią.

✔ **Przykład (1).**
Znajdźmy równanie parametryczne prostej L przecięcia płaszczyzn

$$\Pi_1 : \ x + y - z - 5 = 0, \quad \Pi_2 : \ 2x - y - 1 = 0.$$

Punkty prostej L będą jednocześnie należeć do obu płaszczyzn, więc muszą spełniać układ

$$\begin{cases} x + y - z = 5 \\ 2x - y = 1 \end{cases}$$

Rozwiążmy ten układ

$$\begin{cases} x + y - z = 5 \\ \quad - 3y + 2z = -9 \end{cases}$$

Wyznaczmy zmienne x, y traktując z jak parametr. Otrzymujemy
$z = z$,
$3y = 9 + 2z$ więc $y = 3 + \frac{2}{3}z$,
$x = 5 - y + z = 5 - 3 - \frac{2}{3}z + z = 2 + \frac{1}{3}z$.
Podobnie jak w jednym z poprzednich przykładów lepiej będzie wprowadzić tutaj nowy parametr $\lambda \in R$. Prostą przecięcia L będziemy mogli wtedy zapisać równaniem parametrycznym postaci

$$\begin{cases} x = 2 + \frac{1}{3}\lambda \\ y = 3 + \frac{2}{3}\lambda \\ z = \lambda \end{cases}$$

Wektorem kierunkowym tej prostej jest wektor $[\frac{1}{3}, \frac{2}{3}, 1]$, zaś jako punkt początkowy możemy przyjąć punkt $(2, 3, 0)$.

✔ **Przykład (2).**
Znajdźmy równanie prostej L w R^3 będącej przekrojem dwóch płaszczyzn zadanych równaniami parametrycznymi.

$$\Pi_1 : \begin{cases} x = t_1 + 2t_2 \\ y = -2t_1 + 5 \\ z = 2t_1 - t_2 - 1 \end{cases} \qquad \Pi_2 : \begin{cases} x = -s_1 - s_2 + 2 \\ y = s_1 - s_2 + 3 \\ z = 3s_1 + s_2 \end{cases}$$

Bardzo ważne jest, aby przy rozważaniu tego typu zadań we wzorach różnych płaszczyzn pojawiały się różne parametry. Istotne jest bowiem to, że parametry te muszą być niezależne.

Zajmijmy się rozwiązaniem naszego problemu. Szukane punkty, czyli punkty prostej L, będą jednocześnie należeć do obu układów, będą się zatem wyrażać na dwa sposoby. Wartości parametrów będą zatem spełniać związki

$$\begin{cases} t_1 + 2t_2 = -s_1 - s_2 + 2 \\ -2t_1 + 5 = s_1 - s_2 + 3 \\ 2t_1 - t_2 - 1 = 3s_1 + s_2 \end{cases}$$

Uporządkujmy ten układ.

$$\begin{cases} t_1 + 2t_2 + s_1 + s_2 = 2 \\ -2t_1 - s_1 + s_2 = -2 \\ 2t_1 - t_2 - 3s_1 - s_2 = 1 \end{cases}$$

Naszym celem jest znalezienie odpowiednich wartości parametrów. Zauważmy jednak, że wystarczy znaleźć dowolną parę: t_1, t_2 lub s_1, s_2. Tablica odpowiadająca temu układowi to

$$T1 : \begin{array}{lrrrrr} {}^{(I)} & 1 & 2 & 1 & 1 & 2 \\ {}^{(II)} & -2 & 0 & -1 & 1 & -2 \\ {}^{(III)} & 2 & -1 & -3 & -1 & 1 \end{array}$$

Przekształćmy ją.

$$T2:\quad \begin{array}{lccccc} {\scriptstyle (I'=I)} & 1 & 2 & 1 & 1 & 2 \\ {\scriptstyle (II'=II+2I)} & 0 & 4 & 1 & 3 & 2 \\ {\scriptstyle (III'=III-2I)} & 0 & -5 & -5 & -3 & -3 \end{array}$$

$$T3:\quad \begin{array}{lccccc} {\scriptstyle (I''=I')} & 1 & 2 & 1 & 1 & 2 \\ {\scriptstyle (II''=II')} & 0 & 4 & 1 & 3 & 2 \\ {\scriptstyle (III''=III'+\frac{5}{4}II')} & 0 & 0 & -3\frac{3}{4} & \frac{3}{4} & -\frac{1}{2} \end{array}$$

Odtwórzmy układ równań.

$$\begin{cases} t_1 + 2t_2 + s_1 + s_2 = 2 \\ \quad 4t_2 + s_1 + 3s_2 = 2 \\ \quad -3\frac{3}{4}s_1 + \frac{3}{4}s_2 = -\frac{1}{2} \end{cases}$$

Jak widać musimy użyć parametru, niech będzie nim s_2. Mamy wówczas $-\frac{15}{4}s_1 = -\frac{1}{2} - \frac{3}{4}s_2$, stąd $s_1 = \frac{2}{15} + \frac{1}{5}s_2$.
Wiemy, że

$$\begin{cases} x = -s_1 - s_2 + 2 \\ y = s_1 - s_2 + 3 \\ z = 3s_1 + s_2 \end{cases}$$

Wstawmy odpowiednie wartości parametrów. Otrzymujemy wówczas

$$\begin{aligned} x &= -\frac{2}{15} - \frac{1}{5}s_2 - s_2 + 2 = -1\frac{1}{5}s_2 + 1\frac{13}{15} \\ y &= \frac{2}{15} + \frac{1}{5}s_2 - s_2 + 3 = 3\frac{2}{15} - \frac{4}{5}s_2 \\ z &= \frac{2}{5} + \frac{3}{5}s_2 + s_2 = \frac{2}{5} + 1\frac{3}{5}s_2 \end{aligned}$$

Te wartości x, y, z to współrzędne punktów prostej L. Zależą one od parametru s_2.
Układ

$$\begin{cases} x = -1\frac{1}{5}s_2 + 1\frac{13}{15} \\ y = 3\frac{2}{15} - \frac{4}{5}s_2 \\ z = \frac{2}{5} + 1\frac{3}{5}s_2 \end{cases}$$

możemy przyjąć za parametryczne równanie prostej L.

Rozdział 17

Przestrzenie R^n

Pierwszą część tego rozdziału poświęcimy na zdefiniowanie podstawowych pojęć, którymi posługiwać będziemy się w dalszej części skryptu, przy badaniu przestrzeni R^n. Będą one uogólnieniem pojęć podanych w przypadku R^2 i R^3. W drugiej części rozdziału zajmiemy się iloczynem skalarnym. Podamy jego definicję oraz podstawowe własności, następnie pokażemy, jak jest on użyteczny. Nauczymy się np. przy użyciu iloczynu skalarnego, znajdować długość wektora z R^n, czy kąt między wektorami z R^n, dla dowolnego n.

17.1 Podstawowe definicje.

Definicja 17.1.1 (przestrzeń R^n). Dla dowolnego $n \in N$, przestrzenią R^n nazywamy zbiór wszystkich obiektów postaci $[x_1, x_2, \dots x_n]$, gdzie x_i są liczbami rzeczywistymi.

Definicja 17.1.2 (wektory w R^n). Obiekty postaci $[x_1, x_2, \dots, x_n]$, gdzie $x_i \in R$ nazywamy wektorami w R^n.

✔ **Przykład.** Uporządkowana piątka liczb rzeczywistych $[4, -2, 0, 2, 4]$ jest wektorem w przestrzeni R^5.

Uwaga 17.1.3 (kolumnowy zapis wektorów). *Wektor $[x_1, x_2, \dots, x_n]$ możemy zapisać w postaci $\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$. Takie przedstawienie nazywamy kolumnowym zapisem wektora. Najczęściej właśnie w taki sposób będziemy przedstawiać wektory.*

Definicja 17.1.4 (współrzędne wektora). Liczbę x_i nazywamy $i-$ tą współrzędną wektora $X = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \in R^n$.

Definicja 17.1.5 (wersory w R^n). Wersorami w przestrzeni R^n nazywamy wektory postaci $E_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$, $E_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \ldots, E_n = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix}$.

✔ **Przykład.** Wersorami w przestrzeni R^4 są

$$E_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \ E_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \ E_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix}, E_4 = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}.$$

17.2 Działania na wektorach.

Poniżej, w sposób algebraiczny, zdefiniujemy działania na wektorach w R^n. Będą one uogólnieniem działań na wektorach w R^2 i R^3.

➢ Dodawanie wektorów i mnożenie wektora przez skalar

Dla wektorów $X = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$, $Y = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} \in R^n$ oraz liczby rzeczywistej t określmy

$$X + Y = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} + \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} = \begin{pmatrix} x_1 + y_1 \\ x_2 + y_2 \\ \vdots \\ x_n + y_n \end{pmatrix}, \quad t \cdot X = t \cdot \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} tx_1 \\ tx_2 \\ \vdots \\ tx_n \end{pmatrix}.$$

✔ **Przykład.** Sumą wektorów $A = \begin{pmatrix} 7 \\ 2 \\ -2 \\ 4 \end{pmatrix}$, $B = \begin{pmatrix} -3 \\ 0 \\ 5 \\ -2 \end{pmatrix}$ nazywamy wektor $A + B = \begin{pmatrix} 4 \\ 2 \\ 3 \\ 2 \end{pmatrix}$. Natomiast iloczyn wektora A i liczby rzeczywistej $t = -5$, to wektor

$$tA = (-5) \cdot \begin{pmatrix} 7 \\ 2 \\ -2 \\ 4 \end{pmatrix} = \begin{pmatrix} -35 \\ -10 \\ 10 \\ -20 \end{pmatrix}.$$

➢ Własności działań na wektorach

Okazuje się, że działania na wektorach z R^n spełniają te same własności, co działania na wektorach z R^2 i R^3. Tak więc, dla dowolnych wektorów X, U, V z R^n oraz dowolnych liczb rzeczywistych t, s zachodzą zależności

(1) $X + U = U + X$

(2) $(X + U) + V = X + (U + V)$

(3) $t(X + U) = tX + tU$

(4) $(t + s)X = tX + sX$

(5) $t(sX) = (ts)X$

(6) $1 \cdot X = X$

✍ **Ćwiczenie.** Udowodnienie tych własności wymaga skorzystania z algebraicznych definicji działań na wektorach. Pozostawimy je do wykonania czytelnikowi.

➢ **Wektor zerowy**

Wektor $\overrightarrow{O} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$ nazywamy wektorem zerowym z R^n.

Fakt 17.2.1. *Dla każdego wektora $X \in R^n$ mamy*

$$X + \overrightarrow{O} = X.$$

➢ **Wektor przeciwny**

Dla każdego wektora $X \in R^n$ możemy dobrać taki wektor $-X$, że

$$X + (-X) = \overrightarrow{O}.$$

Wektor $-X$ o takiej własności nazywamy wektorem przeciwnym do wektora X. Jeśli $X = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$, to wektor przeciwny do X ma postać $-X = \begin{pmatrix} -x_1 \\ -x_2 \\ \vdots \\ -x_n \end{pmatrix}$.

✔ **Przykład.** Wektorem przeciwnym do $A = \begin{pmatrix} 3 \\ 7 \\ 0 \\ -5 \\ -2 \end{pmatrix}$ jest wektor $-A = \begin{pmatrix} -3 \\ -7 \\ 0 \\ 5 \\ 2 \end{pmatrix}$.

➢ Przedstawienie wektora jako kombinacji liniowej wersorów

Obserwacja 17.2.2. *Zauważmy, że każdy wektor $X \in R^n$ przedstawia się jednoznacznie jako kombinacja liniowa wersorów z tej przestrzeni, tzn.*

$$X = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = x_1E_1 + x_2E_2 + \ldots x_nE_n.$$

✔ **Przykład.** Przedstawmy wektor $X = \begin{pmatrix} 3 \\ -7 \\ 6 \\ 2 \\ -3 \end{pmatrix}$ w postaci kombinacji liniowej wersorów. Wektor X jest wektorem z przestrzeni R^5. Wersorami w tej przestrzeni są $E_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}$, $E_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}$, $E_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}$, $E_4 = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \\ 0 \end{pmatrix}$, $E_5 = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}$. Wektor X możemy zapisać jako

$$X = \begin{pmatrix} 3 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} + \begin{pmatrix} 0 \\ -7 \\ 0 \\ 0 \\ 0 \end{pmatrix} + \begin{pmatrix} 0 \\ 0 \\ 6 \\ 0 \\ 0 \end{pmatrix} + \begin{pmatrix} 0 \\ 0 \\ 0 \\ 2 \\ 0 \end{pmatrix} + \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ -3 \end{pmatrix} =$$

$$= 3 \cdot \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} + (-7) \cdot \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} + 6 \cdot \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \\ 0 \end{pmatrix} + 2 \cdot \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \\ 0 \end{pmatrix} + (-3) \cdot \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 1 \end{pmatrix} =$$

$$= 3E_1 + (-7)E_2 + 6E_3 + 2E_4 + (-3)E_5$$

17.3 Liniowa niezależność wektorów.

Podobnie jak w przypadku wektorów z przestrzeni R^2 i R^3, również w przestrzeni R^n możemy mówić o liniowej niezależności wektorów. Przypomnijmy, że dla wektorów z R^2 i R^3, w przypadku dwóch wektorów, liniowa niezależność oznaczała ich niewspółliniowość. Podobnie jest w przypadku wektorów z R^n.

➢ Współliniowość wektorów

Wektory X i Y z R^n są współliniowe, jeśli $X = tY$ dla pewnego $t \in R$. W przeciwnym wypadku wektory są niewspółliniowe.

✔ **Przykład (1).**

Wektory $X = \begin{pmatrix} 1 \\ 5 \\ -7 \\ 2 \\ -4 \end{pmatrix}$, $Y = \begin{pmatrix} -1,5 \\ -7,5 \\ 10,5 \\ -3 \\ 6 \end{pmatrix} \in R^5$ są współliniowe, bo $Y = -1,5 \cdot X$, a stąd $X = -\frac{2}{3}Y$.

✔ **Przykład (2).**

Wektory $X = \begin{pmatrix} -5 \\ 3 \\ 4 \\ 2 \end{pmatrix}$, $Y = \begin{pmatrix} 7 \\ 0 \\ -1 \\ 3 \end{pmatrix} \in R^4$ nie są współliniowe. Musiałaby istnieć taka liczba t, aby pierwsza współrzędna wektora Y była iloczynem tej liczby i pierwszej współrzędnej wektora X, co oznacza, że $7 = t \cdot (-5)$, tak więc $t = -\frac{7}{5}$. Drugie współrzędne powinny spełniać tę samą zależność, zatem w tym wypadku $0 = t \cdot 3$, więc $t = 0$. Zatem $t = 0$ i $t = -\frac{7}{5}$, tak więc nie istnieje odpowiednia wartość t, co daje wniosek, że wektory X i Y są niewspółliniowe.

➢ Liniowa niezależność wektorów

Uogólnieniem pojęcia niewspółliniowości na większą liczbę wektorów jest liniowa niezależność. Podamy teraz dwie równoważne definicje tego pojęcia.

Definicja 17.3.1 (liniowa niezależność wektorów). Układ $V_1, V_2, \ldots V_k$ wektorów z R^n jest liniowo niezależny, jeśli żaden wektor V_i nie jest równy kombinacji liniowej pozostałych.

Uwaga 17.3.2. *Jeśli $n > 3$, to w R^n mogą istnieć liniowo niezależne układy złożone z więcej niż trzech wektorów.*

✔ **Przykład.** Wersory $E_1, E_2, \ldots, E_n$ tworzą układ liniowo niezależny złożony z n wektorów. Aby to uzasadnić, pokażemy najpierw, że E_1 nie jest kombinacją liniową pozostałych wersorów. Załóżmy nie wprost, że $E_1 = t_2E_2 + t_3E_3 + \ldots + t_nE_n$, zatem

$$\begin{pmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix} = t_2 \begin{pmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} + \ldots + t_n \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix}.$$

Mamy zatem

$$\begin{pmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ t_2 \\ t_3 \\ \vdots \\ t_n \end{pmatrix}.$$

Otrzymujemy zatem $1 = 0$ co jest sprzecznością. Podobnie możemy pokazać, że inne wersory nie są kombinacją liniową pozostałych, więc zgodnie z powyższą definicją, układ ten rzeczywiście jest liniowo niezależny.

Definicja 17.3.3 (**liniowa niezależność a kombinacja liniowa**)**.** Układ złożony z wektorów $V_1, V_2, \ldots, V_k$ jest liniowo niezależny, jeśli jedyna kombinacja liniowa $t_1V_1 + t_2V_2 + \ldots + t_nV_n$ dająca wektor zerowy, to kombinacja trywialna, tzn. ze wszystkimi współczynnikami zerowymi.

Dowód równoważności powyższych definicji jest podobny do dowodu Lematu 1.9.2. Tutaj go pominiemy.

Okazuje się, że druga z tych definicji jest dużo bardziej użyteczna niż pierwsza i częściej będziemy z niej korzystać.

✔ **Przykład** (**1**)**.** Rozstrzygnijmy, czy układ złożony z wektorów

$$\begin{pmatrix} -1 \\ 3 \\ 2 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 5 \\ -1 \\ 7 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 2 \\ 0 \\ -1 \\ 0 \end{pmatrix},$$

jest liniowo niezależny. Rozważmy kombinację liniową tych wektorów i przyrównajmy ją do zera.

$$t_1 \begin{pmatrix} -1 \\ 3 \\ 2 \\ 0 \end{pmatrix} + t_2 \begin{pmatrix} 0 \\ 5 \\ -1 \\ 7 \end{pmatrix} + t_3 \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix} + t_4 \begin{pmatrix} 2 \\ 0 \\ -1 \\ 0 \end{pmatrix} = 0.$$

Rozstrzygnięcie liniowej niezależności podanych wektorów, sprowadza się zatem do rozwiązania układu równań. Jeżeli okaże się, że jedynymi rozwiązaniami tego układu są $t_1 = t_2 = t_3 = t_4 = 0$, to układ ten jest liniowo niezależny. Jeśli natomiast znajdziemy niezerowe rozwiązanie, oznaczać to będzie, że układ jest liniowo zależny.

W celu rozwiązania tego układu, skorzystamy z poznanej w poprzednim rozdziale metody Gaussa. Odpowiadająca temu układowi tablica, to

$$\begin{array}{rrrrr} -1 & 0 & 1 & 2 & 0 \\ 3 & 5 & 1 & 0 & 0 \\ 2 & -1 & 1 & -1 & 0 \\ 0 & 7 & 1 & 0 & 0 \end{array}$$

Po kilku przekształceniach otrzymujemy

$$\begin{array}{rrrrr} -1 & 0 & 1 & 2 & 0 \\ 0 & -1 & 3 & 3 & 0 \\ 0 & 0 & 14 & 15 & 0 \\ 0 & 0 & 0 & -2\frac{4}{7} & 0 \end{array}$$

Możemy teraz odtworzyć przekształcony układ równań

$$\begin{cases} -t_1 \qquad\quad + t_3 + 2t_4 = 0 \\ \qquad - t_2 + 3t_3 + 3t_4 = 0 \\ \qquad\qquad 14t_3 + 15t_4 = 0 \\ \qquad\qquad\qquad - 2\frac{4}{7}t_4 = 0 \end{cases}$$

Ostatecznie otrzymujemy

$$\begin{cases} t_1 = 0 \\ t_2 = 0 \\ t_3 = 0 \\ t_4 = 0 \end{cases}$$

Nasz układ wektorów jest zatem rzeczywiście liniowo niezależny.

✔ **Przykład (2).** Weźmy teraz inny układ wektorów

$$\begin{pmatrix} 1 \\ 3 \\ 0 \\ 4 \end{pmatrix}, \begin{pmatrix} -1 \\ 2 \\ -2 \\ 2 \end{pmatrix}, \begin{pmatrix} 0 \\ 7 \\ -3 \\ 2 \end{pmatrix}, \begin{pmatrix} 7 \\ 3 \\ 7 \\ 0 \end{pmatrix}.$$

Przekonamy się, że jest on liniowo zależny. Rozważmy kombinację liniową tych wektorów i przyrównajmy ją do zera.

$$t_1 \begin{pmatrix} 1 \\ 3 \\ 0 \\ 4 \end{pmatrix} + t_2 \begin{pmatrix} -1 \\ 2 \\ -2 \\ 2 \end{pmatrix} + t_3 \begin{pmatrix} 0 \\ 7 \\ -3 \\ 2 \end{pmatrix} + t_4 \begin{pmatrix} 7 \\ 3 \\ 7 \\ 0 \end{pmatrix} = 0.$$

Podobnie jak w poprzednim przypadku rozwiązanie tego zagadnienia sprowadza się do rozwiązania odpowiedniego układu równań. Odpowiadająca temu układowi tablica, to

$$\begin{array}{rrrrr} 1 & -1 & 0 & 7 & 0 \\ 3 & 2 & 7 & 3 & 0 \\ 0 & -2 & -3 & 7 & 0 \\ 4 & 2 & 2 & 0 & 0 \end{array}.$$

Po przekształceniach tablica ta może być zapisana w postaci

$$\begin{array}{rrrrr} 1 & -1 & 0 & 7 & 0 \\ 0 & 5 & 7 & -18 & 0 \\ 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{array}.$$

Jak wiemy układ taki będzie miał całą rodzinę rozwiązań, które przy użyciu parametru $s \in R$ możemy zapisać jako

$$\begin{cases} t_1 = -2s \\ t_2 = 5s \\ t_3 = -s \\ t_4 = s \end{cases}$$

Zatem istnieje nieskończenie wiele nietrywialnych kombinacji danych wektorów dających zero. Przykładem takiej kombinacji jest $t_1 = -2, t_2 = 5, t_3 = -1, t_4 = 1$, mamy wówczas

$$-2 \cdot \begin{pmatrix} 1 \\ 3 \\ 0 \\ 4 \end{pmatrix} + 5 \cdot \begin{pmatrix} -1 \\ 2 \\ -2 \\ 2 \end{pmatrix} - \begin{pmatrix} 0 \\ 7 \\ -3 \\ 2 \end{pmatrix} + \begin{pmatrix} 7 \\ 3 \\ 7 \\ 0 \end{pmatrix} = 0.$$

Dane wektory rzeczywiście są liniowo zależne.

17.4 Definicja i podstawowe własności iloczynu skalarnego w R^n.

W tym i kolejnych podrozdziałach zajmiemy się iloczynem skalarnym. Na początek podamy definicję tego pojęcia, następnie wprowadzimy wiele charakterystycznych własności i sposobów wykorzystania iloczynu skalarnego.

Definicja 17.4.1 (iloczyn skalarny). Iloczynem skalarnym wektorów $X = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$, $Y = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}$ z przestrzeni R^n, nazywamy liczbę

$$\boxed{X \circ Y = x_1 y_1 + x_2 y_2 + \dots x_n y_n.}$$

✔ **Przykład.** Iloczynem skalarnym wektorów $A = [3, -5, -7, 4]$, $B = [2, 1, 0, -3]$ jest $A \circ B = 3 \cdot 2 + (-5) \cdot 1 + (-7) \cdot 0 + 4 \cdot (-3) = -11$.

➢ Własności iloczynu skalarnego

Dla dowolnych wektorów $X, Y \in R^n$ oraz dowolnej liczby rzeczywistej t zachodzą równości

(1) $X \circ Y = Y \circ X$

(2) $X \circ (Y + Z) = X \circ Y + X \circ Z$

(3) $(tX) \circ Y = t(X \circ Y)$

Są to własności analogiczne jak dla zwykłego iloczynu skalarnego w R^2 i R^3.

✍ **Ćwiczenie.** Udowodnienie tych własności za pomocą Definicji 17.4.1 pozostawimy do wykonania czytelnikowi.

➢ Długość wektora z R^n

Dla dowolnego wektora $X = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \in R^n$ jego długość określa się wzorem

$$\boxed{|X| = \sqrt{\textstyle\sum_{i=1}^{n} x_i^2}.} \tag{17.1}$$

Zauważmy, że wzór ten jest uogólnieniem wzorów na długość wektorów z przestrzeni R^2 i R^3.

✔ **Przykład.** Długość wektora $A = [1, -5, 3, 0, -2]$ to

$$|A| = \sqrt{1^2 + (-5)^2 + 3^2 + 0^2 + (-2)^2} = \sqrt{1 + 25 + 9 + 4} = \sqrt{39}.$$

Własność 17.4.2. *Dla dowolnego wektora z przestrzeni R^n mamy*

(17.2) $$\boxed{|X|^2 = X \circ X.}$$

✍ **Ćwiczenie.** Prosty dowód tej własności pozostawimy do wykonania czytelnikowi.

➢ Prostopadłość wektorów w R^n

Fakt 17.4.3. *Mówimy, że dwa wektory X, Y z przestrzeni R^n są prostopadłe, jeśli $X \circ Y = 0$.*

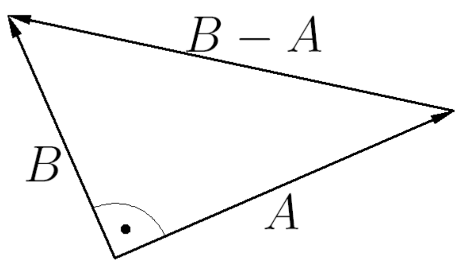
Uwaga 17.4.4. *W przestrzeniach R^n częściej niż określenia prostopadłość będziemy używać zwrotu ortogonalność.*

✔ **Przykład.** Wektory $A = [5, 2, -7, 1]$, $B = [-3, 4, 0, 7]$ są ortogonalne, gdyż

$$A \circ B = 5 \cdot (-3) + 2 \cdot 4 + (-7) \cdot 0 + 1 \cdot 7 = -15 + 8 + 0 + 7 = 0.$$

➢ Twierdzenie Pitagorasa w R^n

W dowolnej przestrzeni R^n możemy podać twierdzenie, będące odpowiednikiem twierdzenia Pitagorasa na płaszczyźnie. Na początek zauważmy, że rolę trójkąta prostokątnego mogą w przestrzeni spełniać dwa prostopadłe wektory A i B. Przeciwprostokątną będzie w tym wypadku wektor $B - A$, co łatwo można zauważyć ze schematycznego rysunku przedstawionego poniżej.



Twierdzenie 17.4.5 (Pitagorasa). *Jeśli wektory $A, B \in R^n$ są prostopadłe, to*

$$|B - A|^2 = |A|^2 + |B|^2.$$

Dowód. Z własności iloczynu skalarnego wiemy, że

$$|B - A|^2 = (B - A) \circ (B - A).$$

Korzystając z rozdzielności iloczynu skalarnego względem odejmowania mamy

$$(B - A) \circ (B - A) = B \circ B - B \circ A - A \circ B + A \circ A.$$

Wiemy również, że wektory A i B są ortogonalne, zatem

$$A \circ B = B \circ A = 0.$$

Ostatecznie otrzymujemy

$$|B - A|^2 = B \circ B + A \circ A = |B|^2 + |A|^2.$$

□

Uwaga 17.4.6. *Czasami nazwę twierdzenie Pitagorasa nosi następująca zależność. Jeśli wektory $A, B \in R^n$ są prostopadłe, to*

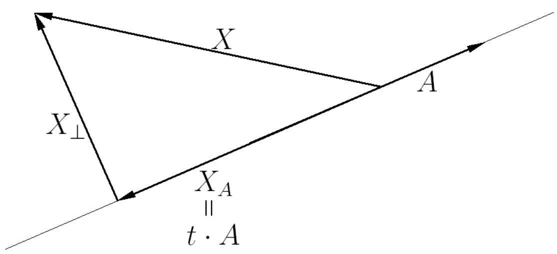
$$|A + B|^2 = |A|^2 + |B|^2.$$

✍ **Ćwiczenie.** Dowód powyższej uwagi jest analogiczny do dowodu Twierdzenia 17.4.5. Pozostawimy go do wykonania czytelnikowi.

17.5 Rozkład wektora na składowe.

Lemat 17.5.1. *Niech A będzie ustalonym, niezerowym wektorem z R^n. Wówczas dowolny wektor $X \in R^n$ można przedstawić jako sumę wektorów $X = X_A + X_\perp$, gdzie X_A to wektor współliniowy z A (składowa równoległa do A), zaś $X_\perp$ jest prostopadły do wektora A (składowa prostopadła do A).*

Uwaga 17.5.2. *Takie przedstawienie wektora X nazywamy jego rozkładem na składowe równoległą i prostopadłą do A. Rozkład taki jest jednoznaczny.*



Dowód (Lematu). Składowa równoległa X_A, jako współliniowa z A, ma postać $X_A = t \cdot A$. Obliczmy t. Skorzystamy z tego, że składowa $X_\perp$ jest prostopadła do A, tak więc $X_\perp \circ A = 0$. Ponieważ $X = X_A + X_\perp$, to $X_\perp = X - X_A = X - t \cdot A$. Otrzymujemy więc

$$(X - tA) \circ A = 0$$

a zatem

$$X \circ A - tA \circ A = 0.$$

Możemy z tej równości wyznaczyć t.

$$t = \frac{X \circ A}{A \circ A} = \frac{X \circ A}{|A|^2}.$$

Stąd składowa równoległa wyraża się wzorem

$$X_A = t \cdot A = \frac{X \circ A}{|A|^2} \cdot A,$$

natomiast składowa prostopadła, to

$$X_\perp = X - \frac{X \circ A}{|A|^2} \cdot A.$$

Wzory te pokazują istnienie i jednoznaczność rozkładu. □

17.6 Kąt między wektorami.

Dla wektorów w przestrzeni R^n, podobnie jak w przypadku R^2 i R^3, zdefiniujemy kąt między wektorami, poprzez podanie jego cosinusa.

Definicja 17.6.1 (cosinus kąta między wektorami). Cosinusem kąta pomiędzy niezerowymi wektorami $U, W \in R^n$ nazywamy liczbę

$$\boxed{\cos(\measuredangle(U,W)) = \frac{U \circ W}{|U| \cdot |W|} = \frac{\sum_{i=1}^n u_i w_i}{\sqrt{\sum_{i=1}^n u_i^2}\sqrt{\sum_{i=i}^n w_i^2}}.} \tag{17.3}$$

✔ **Przykład.** Znajdźmy kąt między wektorami $A = [1, 5, 7, -3], B = [3, 0, -2, -1]$. Cosinus tego kąta to

$$\cos(\measuredangle(A,B)) = \frac{1 \cdot 3 + 5 \cdot 0 + 7 \cdot (-2) + (-3) \cdot (-1)}{\sqrt{1^2 + 5^2 + 7^2 + (-3)^2} \cdot \sqrt{3^2 + 0^2 + (-2)^2 + (-1)^2}} =$$

$$= \frac{-8}{\sqrt{84}\sqrt{14}} \approx -0,2333.$$

Z tablic możemy teraz odczytać przybliżoną wartość kąta między tymi wektorami

$$\measuredangle(A,B) \approx 103^0.$$

Uwaga 17.6.2. *Możemy zastanawiać się, czy cosinus kąta między wektorami, otrzymany ze wzoru (17.3), spełnia warunek*

$$-1 \leqslant \cos(\measuredangle(U,W)) \leqslant 1,$$

co możemy zapisać jako

$$-1 \leqslant \frac{U \circ W}{|U| \cdot |W|} \leqslant 1.$$

Chcielibyśmy aby tak było, gdyż funkcja cosinus przyjmuje wartości z przedziału $[-1, 1]$. *Przekształcając powyższą nierówność, otrzymamy nierówności równoważne.*

(1) $-1 \leqslant \frac{U \circ W}{|U| \cdot |W|} \leqslant 1$

(2) $-|U| \cdot |W| \leqslant U \circ W \leqslant |U| \cdot |W|$

(3) $|U \circ W| \leqslant |U| \cdot |W|$

(4) $|U \circ W|^2 \leqslant |U|^2 \cdot |W|^2$

Trzecia z tych nierówności nazywana jest nierównością Schwartza.

Skoro podane nierówności są równoważne, wystarczy, że udowodnimy jedną z nich. Udowodnijmy zatem czwartą nierówność.

Dowód (4). Rozłóżmy wektor W na składową równoległą $W_U = tU$ oraz prostopadłą $W_\perp$ do wektora U. Wiemy, że wówczas $W = W_U + W_\perp$. Przekształćmy obie strony nierówności wstawiając za W powyższy rozkład. Niech L oznacza lewą stronę nierówności, a R prawą stronę.

$$L = |U \circ W|^2 = |U \circ (W_U + W_\perp)|^2 = |U \circ (tU + W_\perp)|^2 = |tU \circ U + U \circ W_\perp|^2$$

Wiemy, że iloczyn skalarny wektorów prostopadłych jest zerowy. Lewa strona nierówności przyjmuje zatem postać

$$L = |t \cdot |U|^2|^2 = |t|^2 \cdot |U|^4.$$

Natomiast prawa strona to

$$P = |U|^2 \cdot |W|^2 = |U|^2 \cdot |tU + W_\perp|^2.$$

Korzystając z udowodnionego wcześniej twierdzenia Pitagorasa, możemy prawą stronę zapisać jako

$$P = |U|^2 \cdot (|tU|^2 + |W_\perp|^2) = |t|^2|U|^4 + |U|^2|W_\perp|^2 = L + |U|^2|W_\perp|^2.$$

Drugi składnik tej sumy jest nieujemny, dlatego

$$L \leqslant L + |U|^2|W_\perp|^2 = P.$$

□

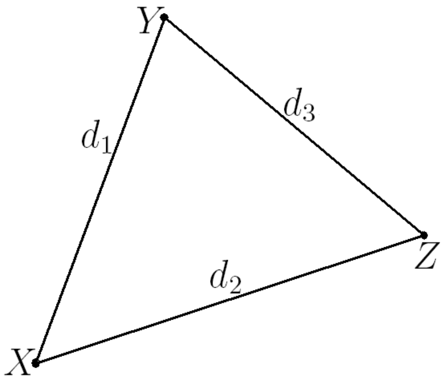
Uwaga 17.6.3. *W nierówności Schwartza zachodzi równość dokładnie wtedy gdy*

$$|U|^2|W_\perp|^2 = 0.$$

Skoro wektory U, W są niezerowe, to równość ta zachodzi, gdy $|W_\perp|^2 = 0$, czyli $|W_\perp| = 0$, a stąd $W_\perp = O$, a to zachodzi gdy $W = tU$. Równość w nierówności Schwartza zachodzi dokładnie wtedy, gdy wektory są współliniowe.

17.7 Nierówność trójkąta.

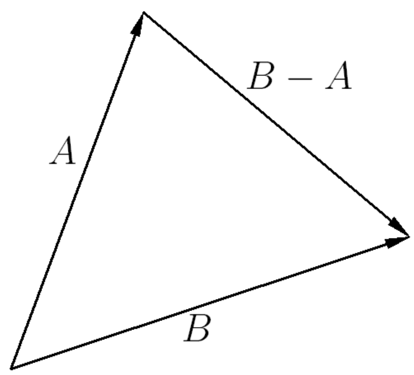
Jeżeli mamy dane trzy punkty X, Y, Z,



to nierówność trójkąta możemy zapisać jako

$$d_3 \leqslant d_1 + d_2.$$

Przy badaniu przestrzeni R^n częściej zajmujemy się wektorami, a nie punktami. Spróbujmy zatem wyrazić tę nierówność w języku wektorów.



Dla wektorów $A, B \in R^n$ mamy zatem

$$\boxed{|B - A| \leqslant |A| + |B|}. \tag{17.4}$$

Dowód (nierówności trójkąta). Zauważmy, że

$$|B - A| \leqslant |A| + |B|$$

wtedy i tylko wtedy gdy

$$|B - A|^2 \leqslant (|A| + |B|)^2$$

ale

$$|B - A|^2 = (B - A) \circ (B - A) = |B|^2 + |A|^2 - 2A \circ B$$

zatem

$$|B|^2 + |A|^2 - 2A \circ B \leqslant |A|^2 + 2|A||B| + |B|^2$$

stąd

$$-2A \circ B \leqslant 2|A||B|$$

$$A \circ B \leqslant |A||B|,$$

a ta nierówność wynika z nierówności Schwartza. Wiemy bowiem, że

$$|A \circ B| \leqslant |A||B|$$

czyli

$$-|A||B| \leqslant A \circ B \leqslant |A||B| \quad / \cdot (-1)$$

$$|A||B| \geqslant -A \circ B \geqslant -|A||B|.$$

Zatem spełniona jest nierówność trójkąta. □

Uwaga 17.7.1. *Równość w nierówności trójkąta zachodzi tylko wtedy, gdy wektory A i B są współliniowe i mają przeciwne zwroty.*

Rozdział 18

Macierze

Dotychczas z pojęciem macierzy mieliśmy do czynienia głównie przy okazji omawiania przekształceń przestrzeni. W tym rozdziale zajmiemy się abstrakcyjnym pojęciem macierzy. Na początek zdefiniujemy ogólnie macierze, poznamy podstawową terminologię z nimi związaną. Bliżej zajmiemy się pewnym rodzajem macierzy - macierzami kwadratowymi. Dla takich macierzy określimy wyznacznik i poznamy jego własności. W drugiej części tego rozdziału poznamy niektóre z bardzo wielu zastosowań wyznacznika.

18.1 Podstawowe definicje i oznaczenia.

Definicja 18.1.1 (macierz $m \times n$). Macierzą rozmiaru $m \times n$ nazywamy prostokątną tablicę, złożoną z liczb lub innych wyrażeń algebraicznych, zwanych elementami macierzy, składającą się z m wierszy i n kolumn.

✔ **Przykład.** Macierzą rozmiaru 2×3 jest np.

$$\begin{bmatrix} 3 & 0 & x \\ -1 & \frac{y}{2} & 2 \end{bmatrix}.$$

Oznaczenia.

- Macierze będziemy zazwyczaj oznaczać dużymi literami alfabetu np. A, B, C.
- Wyrazy macierzy oznaczamy a_{ij}, b_{ij}, c_{ij}, gdzie pierwszy z indeksów oznacza numer wiersza a drugi numer kolumny. W takiej konwencji wyrazy macierzy A rozmiaru $m \times n$ możemy oznaczyć

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

- Macierz A będziemy też często oznaczać przez $[a_{ij}]$.

- Macierz możemy też zapisywać przy użyciu jej kolumn np. $A = [A_1, A_2, \ldots, A_n]$, gdzie $A_i = \begin{bmatrix} a_{1i} \\ a_{2i} \\ \vdots \\ a_{mi} \end{bmatrix} \in R^m$ to i-ta kolumna macierzy A. Przedstawienie takie nazywamy kolumnowym zapisem macierzy A.

- Podobnie możemy macierz opisać przy użyciu jej wierszy, wówczas $A = \begin{bmatrix} W_1 \\ W_2 \\ \vdots \\ W_m \end{bmatrix}$, gdzie $W_j = [a_{j1}, a_{j2}, \ldots, a_{jn}] \in R^n$ jest j-tym wierszem macierzy A.

➢ **Zastosowania macierzy**

Mieliśmy już wcześniej okazję poznać zastosowania macierzy rozmiaru 2×2 i 3×3. Okazuje się, że uogólniają się one na macierze większych rozmiarów. Przypomnijmy niektóre z tych zastosowań. Macierze występowały jako

- tablice współczynników układów równań,
- macierze przekształceń liniowych,
- macierze form kwadratowych,
- macierze przejścia do nowych układów współrzędnych.

W kolejnych podrozdziałach i rozdziałach zapoznamy się z wszystkimi tymi zastosowaniami dla przypadku macierzy dowolnych rozmiarów.

18.2 Macierze kwadratowe.

Szczególnie ważną klasą macierzy są macierze kwadratowe, czyli macierze rozmiaru $n \times n$, a więc takie, które mają tyle samo wierszy co kolumn.

✔ **Przykład.** Macierzą kwadratową rozmiaru 4×4 jest np.

$$\begin{bmatrix} 1 & -5 & 2 & 7 \\ 4 & \frac{3}{4} & -2 & 0 \\ 7 & 4 & -\frac{5}{6} & 1 \\ 0 & -2 & 1 & 1 \end{bmatrix}.$$

Dla macierzy kwadratowej wprowadzimy dodatkowo bardzo użyteczne pojęcie.

Definicja 18.2.1 (przekątna główna). Przekątna główna macierzy składa się z elementów a_{ii} dla $i = 1, 2, \ldots, n$, czyli takich, które mają ten sam numer kolumny co wiersza.

Uwaga 18.2.2. *Wyrazy znajdujące się pod przekątną główną to wyrazy* a_{ij}*, takie że* $i > j$ *natomiast wyrazy nad główną przekątną to te* a_{ij}*, dla których* $i < j$*.*

✔ **Przykład.** W macierzy $A = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}$ przekątna główna składa się z wyrazów a_{11}, a_{22}, a_{33}. Element a_{32} jest przykładem wyrazu znajdującego się pod główną przekątną i rzeczywiście $3 > 2$ co oznacza, że numer wiersza jest większy niż numer kolumny. Z kolei dla a_{12}, który jak widać znajduje się nad główną przekątną, numer wiersza jest mniejszy od numeru kolumny.

➢ Macierz transponowana

W Rozdziale 12 poznaliśmy pojęcie macierzy transponowanej do macierzy rozmiaru 3×3. Macierz transponowaną do danej możemy również określić dla macierzy większego rozmiaru.

Definicja 18.2.3 (macierz transponowana). Macierz transponowana A^T macierzy kwadratowej $A = [a_{ij}]$ jest to taka macierz $B = [b_{ij}]$ tego samego rozmiaru $n \times n$ dla której $b_{ij} = a_{ji}$.

Tę definicję transponowania macierzy uogólnia się, na macierz dowolnego rozmiaru, niekoniecznie kwadratową.

✔ **Przykład.** Macierzą transponowaną macierzy

$$A = \begin{bmatrix} 5 & -7 & 8 & 4 \\ -2 & 0 & 1 & 3 \end{bmatrix}$$

jest macierz

$$A^T = \begin{bmatrix} 5 & -2 \\ -7 & 0 \\ 8 & 1 \\ 4 & 3 \end{bmatrix}.$$

Uwaga 18.2.4. *Jeśli macierz* A *jest rozmiaru* $m \times n$*, to macierz transponowana* A^T *jest rozmiaru* $n \times m$*.*

18.3 Wyznacznik macierzy kwadratowej.

W podrozdziale tym zdefiniujemy wyznacznik macierzy kwadratowej dowolnego rozmiaru. Zastanówmy się jednak jaką funkcję może on pełnić. Przypomnijmy w jakich sytuacjach używany był wyznacznik macierzy 2×2.

- Użycie wyznacznika jest bardzo dobrą metodą rozstrzygania liniowej niezależności, mianowicie wektory A_1, A_2 są liniowo niezależne wtedy i tylko wtedy gdy wyznacznik macierzy $A = [A_1 \ A_2]$ rozmiaru 2×2, w której A_1, A_2 są kolumnami, jest niezerowy, czyli $\det(A) \neq 0$.

- Układ równań
$$\begin{cases} a_{11}x + a_{12}y = b_1 \\ a_{21}x + a_{22}y = b_2 \end{cases},$$
który wektorowo możemy zapisać jako
$$x \cdot A_1 + y \cdot A_2 = B$$
ma rozwiązanie
$$x = \frac{\det[B, A_2]}{\det[A_1, A_2]}, \; y = \frac{\det[A_1, B]}{\det[A_1, A_2]}$$
jeśli wyznacznik główny tego układu, czyli $\det[A_1, A_2]$ jest niezerowy.

- Macierz odwrotna do macierzy $A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$ wyraża się wzorem
$$A^{-1} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}^{-1} = \frac{1}{\det A} \begin{bmatrix} a_{22} & -a_{12} \\ -a_{21} & a_{11} \end{bmatrix}.$$

Wyznacznik macierzy wykorzystywany był jeszcze w wielu innych sytuacjach. Podobne zastosowania miał wyznacznik macierzy 3×3. Zdefiniujemy teraz wyznacznik $\det A$ macierzy kwadratowej A dowolnego rozmiaru $n \times n$ jako pewną liczbę, lub wyrażenie, zależne od wyrazów macierzy A tak aby znalazł on zastosowanie do celów analogicznych jak te wymienione powyżej dla macierzy rozmiaru 2×2.

➢ Indukcyjna definicja wyznacznika

Okazuje się, że definicja wyznacznika, spełniającego wymienione powyżej własności, dla macierzy większych rozmiarów jest dużo bardziej skomplikowana niż w przypadku macierzy 2×2 i 3×3. Wyznacznik ten zdefiniujemy w sposób indukcyjny. Wcześniej wprowadźmy jednak pomocne oznaczenia.

Oznaczenie. Dla macierzy $A = [a_{ij}]$ rozmiaru $n \times n$ niech A_{ij} oznacza macierz rozmiaru $(n-1) \times (n-1)$ powstałą z macierzy A przez wykreślenie i-tego wiersza i j-tej kolumny.

✔ **Przykład.** Dla macierzy $A = \begin{bmatrix} 5 & -7 & 2 \\ 1 & -3 & 4 \\ 0 & 1 & -2 \end{bmatrix}$ macierz $A_{23} = \begin{bmatrix} 5 & -7 \\ 0 & 1 \end{bmatrix}$ jest macierzą powstałą z macierzy A przez wykreślenie drugiego wiersza i trzeciej kolumny.

Definicja 18.3.1 (indukcyjna definicja wyznacznika).

1^0 Dla macierzy $A = [a_{11}]$ rozmiaru 1×1 wyznacznikiem nazywamy liczbę
$$\det[a_{11}] = a_{11}.$$

2^0 Jeśli wiemy już czym jest wyznacznik macierzy rozmiaru $(n-1)\times(n-1)$, to określamy wyznacznik $\det A$ macierzy rozmiaru $n \times n$ wzorem

$$\det A = a_{11}\det A_{11} - a_{12}\det A_{12} + a_{13}\det A_{13} - \ldots + (-1)^{1+n}a_{1n}\det A_{1n} =$$

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

✔ **Przykład.** Sprawdźmy, czy to określenie wyznacznika macierzy zgadza się z podanymi wcześniej wzorami na wyznacznik macierzy 2×2 i 3×3. W przypadku macierzy $A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$ wyznacznik, zgodnie z powyższą definicją, wyraża się wzorem

$$\det A = a_{11}\det A_{11} - a_{12}\det A_{12} = a_{11}\det[a_{22}] - a_{12}\det[a_{21}] = a_{11}a_{22} - a_{12}a_{21}$$

co zgadza się ze znanym już nam wzorem na wyznacznik macierzy rozmiaru 2×2.

Podobnie w przypadku macierzy $A = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}$ możemy obliczyć jej wyznacznik korzystając z Definicji 18.3.1. Otrzymujemy

$$\det A = a_{11}\det A_{11} - a_{12}\det A_{12} + a_{13}\det A_{13} =$$

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

$$= a_{11}[a_{22}a_{33} - a_{23}a_{32}] - a_{12}[a_{21}a_{33} - a_{31}a_{23}] + a_{13}[a_{21}a_{32} - a_{31}a_{22}] =$$

$$= a_{11}a_{22}a_{33} - a_{11}a_{23}a_{32} - a_{12}a_{21}a_{33} + a_{12}a_{31}a_{23} + a_{13}a_{21}a_{32} - a_{13}a_{31}a_{32}$$

a to zgadza się z wyznaczonym przy użyciu metody Sarrusa wyznacznikiem tej macierzy.

W kolejnym przykładzie zastosujemy Definicję 18.3.1 do obliczenia wyznacznika macierzy rozmiaru 4×4.

✔ **Przykład.** Niech $A = \begin{bmatrix} 1 & 5 & 0 & 3 \\ 4 & -2 & 1 & -3 \\ 0 & 7 & -3 & -1 \\ 2 & -2 & 1 & 3 \end{bmatrix}$. Zgodnie z definicją wyznacznika mamy

$$\det A = a_{11}\det A_{11} - a_{12}\det A_{12} + a_{13}\det A_{13} - a_{14}\det A_{14} =$$

$$= 1\cdot\det\begin{bmatrix} -2 & 1 & -3 \\ 7 & -3 & -1 \\ -2 & 1 & 3 \end{bmatrix} - 5\cdot\det\begin{bmatrix} 4 & 1 & -3 \\ 0 & -3 & -1 \\ 2 & 1 & 3 \end{bmatrix} + 0\cdot\det\begin{bmatrix} 4 & -2 & -3 \\ 0 & 7 & -1 \\ 2 & -2 & 3 \end{bmatrix} - 3\cdot\det\begin{bmatrix} 4 & -2 & 1 \\ 0 & 7 & -3 \\ 2 & -2 & 1 \end{bmatrix} =$$

$$= 1\cdot(18+2-21+18-2-21) - 5\cdot(-36-2+0-18+4-0) +$$

$$+0\cdot(84+4+0+42-8-0) - 3\cdot(28+12+0-14-24-0) =$$

$$= 1\cdot(-6) - 5\cdot(-52) + 0\cdot(122) - 3\cdot(2) = -6+260+0-6 = 248.$$

Uwaga 18.3.2. *Zauważmy, że w celu obliczenia wyznacznika macierzy rozmiaru* 4×4 *musimy policzyć cztery wyznaczniki macierzy* 3×3. *Do wyznaczenia wyznacznika macierzy rozmiaru* 5×5 *będziemy potrzebowali pięciu wyznaczników macierzy rozmiaru* 4×4 *a tym samym 20 wyznaczników macierzy rozmiaru* 3×3.

Poniżej pokażemy bardziej ogólny sposób na obliczanie wyznacznika, który w wielu przypadkach pozwoli na znaczne uproszczenie obliczeń. W tym celu wprowadźmy dodatkowe pojęcie.

Definicja 18.3.3 (dopełnienie algebraiczne). Dopełnienie algebraiczne wyrazu a_{ij} macierzy $A=[a_{ij}]$ to wyrażenie

$$D_{ij}=(-1)^{i+j}\det A_{ij}.$$

Uwaga 18.3.4. *Zauważmy, że w świetle tej definicji wyznacznik macierzy* A *wyraża się wzorem*

$$\det A=\sum_{i=1}^{n}a_{1i}D_{1i}.$$

Twierdzenie 18.3.5 (rozwinięcie Laplace'a). *Wyznacznik możemy policzyć indukcyjnie stosując rozwinięcie względem dowolnego wiersza lub dowolnej kolumny.*

- *rozwinięcie względem* p*-tego wiersza*

$$\det A=\sum_{i=1}^{n}a_{pi}D_{pi},$$

- *rozwinięcie względem* q*-tej kolumny*

$$\det A=\sum_{i=1}^{n}a_{iq}D_{iq}.$$

Poznany przez nas w Definicji 18.3.1 sposób na obliczenie wyznacznika jest rozwinięciem względem pierwszego wiersza.

W poniższym przykładzie będziemy mieć okazję zaobserwować dlaczego może mieć znaczenie wybór kolumny lub wiersza względem której rozwijamy wyznacznik.

✔ **Przykład.** Obliczmy wyznacznik macierzy $A=\begin{bmatrix}-7&1&5&1\\4&-2&0&3\\-2&1&0&-5\\3&2&0&1\end{bmatrix}$. Skorzystanie z Definicji 18.3.1, czyli z rozwinięcia względem pierwszego wiersza wymagałoby policzenia czterech wyznaczników macierzy rozmiaru 3×3. Zauważmy, że jeśli zastosujemy rozwinięcie względem trzeciej kolumny, to wyznacznik macierzy A przedstawia się wzorem

$$\det A=(-1)^{1+3}\cdot 5\cdot\det A_{13}+(-1)^{2+3}\cdot 0\cdot\det A_{23}+(-1)^{3+3}\cdot 0\cdot\det A_{33}+(-1)^{3+4}\cdot 0\cdot\det A_{43}=$$

$$=(-1)^{4}\cdot 5\cdot\det A_{13}=5\cdot\det\begin{bmatrix}4&-2&3\\-2&1&-5\\3&2&1\end{bmatrix}=5\cdot(4+30-12-9+40-4)=5\cdot 49=245.$$

18.4 Wyznacznik macierzy górnotrójkątnej i dolnotrójkątnej.

Obliczenie wyznacznika w wielu przypadkach okazuje się być bardzo kłopotliwe. Istnieją jednak macierze, których szczególne własności pozwalają łatwo go wyznaczyć. Takimi macierzami są macierze górnotrójkątne i dolnotrójkątne.

Fakt 18.4.1. *Jeśli $A = [a_{ij}]$ jest macierzą rozmiaru $n \times n$ górnotrójkątną, co oznacza, że wszystkie wyrazy poniżej głównej przekątnej są zerami, lub dolnotrójkątną, tzn. taką że wszystkie wyrazy nad główną przekątną są zerami, to wyznacznik tej macierzy wyraża się wzorem*

$$\boxed{\det A = a_{11} \cdot a_{22} \cdot \ldots \cdot a_{nn},}$$

jest zatem iloczynem wyrazów na głównej przekątnej.

Dowód. Przeprowadzimy teraz indukcyjny dowód tego faktu. Dla macierzy rozmiaru 1×1, 2×2, 3×3 łatwo sprawdzić, że ten fakt zachodzi. Pozostawimy to do wykonania czytelnikowi.

Przejdźmy do sprawdzenia głównego kroku indukcyjnego. Załóżmy, że Fakt jest prawdziwy dla macierzy górnotrójkątnych rozmiaru $(n-1) \times (n-1)$. Weźmy górnotrójkątną macierz A rozmiaru $n \times n$, gdzie $A = \begin{bmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ 0 & a_{22} & \ldots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & \ldots & 0 & a_{nn} \end{bmatrix}$. Rozwijając tę macierz względem pierwszej kolumny otrzymujemy

$$\det A = \sum_{i=1}^{n} a_{i1} D_{i1} = a_{11} D_{11} + a_{21} D_{21} + \ldots + a_{n1} D_{n1}.$$

Ponieważ $a_{21} = a_{31} = \ldots = a_{n1} = 0$ więc

$$\det A = a_{11} D_{11} = (-1)^{1+1} \cdot a_{11} \det A_{11} = a_{11} \det A_{11}.$$

Ponieważ A_{11} jest macierzą górnotrójkątną rozmiaru $(n-1) \times (n-1)$, która na przekątnej głównej ma wyrazy $a_{22}, a_{33}, \ldots a_{nn}$ więc z założenia indukcyjnego wiemy, że $\det A_{11} = a_{22} \cdot a_{33} \cdot \ldots \cdot a_{nn}$. Po podstawieniu do wzoru na wyznacznik macierzy A otrzymujemy

$$\det A = a_{11} \cdot (a_{22} \cdot a_{33} \cdot \ldots \cdot a_{nn}) = \prod_{i=1}^{n} a_{ii}.$$

Możemy zatem stwierdzić, że podany wzór na wyznacznik macierzy górnotrójkątnej jest prawdziwy dla dowolnego n. □

✍ **Ćwiczenie.** Analogiczny dowód dla macierzy dolnotrójkątnej, w którym wykorzystuje się rozwinięcie względem pierwszego wiersza, pozostawimy do wykonania czytelnikowi.

18.5 Własności wyznacznika.

Poniżej podamy kilka własności wyznacznika, które w znaczący sposób mogą ułatwić jego obliczanie, co na przykładzie pokażemy w kolejnym podrozdziale. W poniższych opisach $A_i = \begin{pmatrix} a_{1i} \\ \vdots \\ a_{mi} \end{pmatrix}$ oznaczać będą kolumny danej macierzy A. Dla poprawy czytelności w wielu miejscach pod wektorem oznaczającym kolumnę zapisany zostanie również numer tej kolumny. Uzasadnienie tych własności podamy w dalszej części podrozdziału.

Własność 18.5.1. *Wyznacznik macierzy A, w której pewna kolumna A_i składa się z samych zer wynosi* 0*, co symbolicznie możemy zapisać*

$$\det[\underset{1}{A_1}, \underset{...}{\ldots}, \underset{i}{\overrightarrow{O}}, \underset{...}{\ldots}, \underset{m}{A_m}] = 0.$$

Własność 18.5.2. *Wyznacznik macierzy A, w której dwie kolumny są jednakowe wynosi* 0.

$$\det[\underset{1}{A_1}, \underset{...}{\ldots}, \underset{i}{A_i}, \underset{...}{\ldots}, \underset{j}{A_i}, \underset{...}{\ldots}, \underset{m}{A_m}] = 0.$$

Własność 18.5.3. *Wyznacznik macierzy, w której przestawiono ze sobą dwie kolumny jest przeciwny do wyznacznika macierzy wyjściowej.*

$$\det[\underset{1}{A_1}, \underset{...}{\ldots}, \underset{i}{A_j}, \underset{...}{\ldots}, \underset{j}{A_i}, \underset{...}{\ldots}, \underset{m}{A_m}] = -\det[\underset{1}{A_1}, \underset{...}{\ldots}, \underset{i}{A_i}, \underset{...}{\ldots}, \underset{j}{A_j}, \underset{...}{\ldots} \underset{m}{A_m}].$$

Własność 18.5.4. *Wyznacznik macierzy, w której jedna z kolumn została przemnożona przez stałą jest iloczynem tej stałej i wyznacznika wyjściowej macierzy.*

$$\det[\underset{1}{A_1}, \underset{...}{\ldots}, \underset{i}{t \cdot A_i}, \underset{...}{\ldots}, \underset{m}{A_m}] = t \cdot \det[\underset{1}{A_1}, \underset{...}{\ldots}, \underset{i}{A_i}, \underset{...}{\ldots} \underset{m}{A_m}].$$

Własność 18.5.5. *Wyznacznik macierzy, w której i-ta kolumna jest sumą wektorów A_i i A_i' jest równy sumie wyznaczników dwóch macierzy, tak jak to opisuje poniższy wzór*

$$\det[\underset{1}{A_1}, \underset{...}{\ldots}, \underset{i-1}{A_{i-1}}, \underset{i}{A_i + A_i'}, \underset{i+1}{A_{i+1}}, \underset{...}{\ldots}, \underset{m}{A_m}] =$$

$$= \det[\underset{1}{A_1}, \underset{...}{\ldots}, \underset{i-1}{A_{i-1}}, \underset{i}{A_i}, \underset{i+1}{A_{i+1}}, \underset{...}{\ldots} \underset{m}{A_m}] + \det[\underset{1}{A_1}, \underset{...}{\ldots}, \underset{i-1}{A_{i-1}}, \underset{i}{A_i'}, \underset{i+1}{A_{i+1}}, \underset{...}{\ldots} \underset{m}{A_m}].$$

Własność 18.5.6. *Wyznacznik macierzy nie zmieni się, jeśli do jednej z kolumn dodamy pewną wielokrotność innej kolumny.*

$$\det[\underset{1}{A_1}, \underset{...}{\ldots}, \underset{i}{A_i + t \cdot A_j}, \underset{...}{\ldots}, \underset{m}{A_m}] = \det[\underset{1}{A_1}, \underset{...}{\ldots}, \underset{i}{A_i}, \underset{...}{\ldots} \underset{m}{A_m}].$$

Poniżej przedstawimy dowody niektórych spośród tych własności. W dowodach tych wykorzystamy rozwinięcia Laplace'a.

Dowód Własności 18.5.1. Obliczając wyznacznik macierzy

$$\det A = \det[\underset{1}{A_1}, \quad \underset{...}{\ldots} \quad, \underset{i}{\overrightarrow{O}}, \quad \underset{...}{\ldots} \quad, \underset{m}{A_m}]$$

zastosujemy rozwinięcie względem i-tej kolumny. Otrzymujemy wówczas

$$\det[A_1, \ldots, \overrightarrow{O}, \ldots, A_m] = a_{1i}D_{1i} + a_{2i}D_{2i} + \ldots + a_{mi}D_{mi} =$$

$$= 0 \cdot D_{1i} + 0 \cdot D_{2i} + \ldots + 0 \cdot D_{mi} = 0.$$

□

Dowód Własności 18.5.5. Obliczając wyznacznik macierzy

$$A = [\underset{1}{A_1}, \quad \underset{...}{\ldots} \quad, \underset{i}{A_i + A_i'}, \quad \underset{...}{\ldots} \quad, \underset{m}{A_m}]$$

również zastosujemy rozwinięcie względem i-tej kolumny, tzn. względem kolumny $A_i + A_i'$. Mamy wówczas

$$\det[A_1, \ldots, A_i + A_i', \ldots, A_m] = (a_{1i} + a_{1i}')D_{1i} + (a_{2i} + a_{2i}')D_{2i} + \ldots + (a_{mi} + a_{mi}')D_{mi} =$$

$$= a_{1i}D_{1i} + a_{1i}'D_{1i} + a_{2i}D_{2i} + a_{2i}'D_{2i} + \ldots + a_{mi}D_{mi} + a_{mi}'D_{mi} =$$
$$\det[A_1, \ldots, A_i, \ldots, A_m] + \det[A_1, \ldots, A_i', \ldots, A_m].$$

□

Dowód Własności 18.5.4. Rozwijając wyznacznik

$$\det A = \det[\underset{1}{A_1}, \quad \underset{...}{\ldots} \quad, \underset{i}{t \cdot A_i}, \quad \underset{...}{\ldots} \quad, \underset{m}{A_m}]$$

względem i-tej kolumny mamy

$$\det[A_1, \ldots, t \cdot A_i, \ldots, A_m] = ta_{1i}D_{1i} + \ldots + ta_{mi}D_{mi} = t \cdot (a_{1i}D_{1i} + \ldots + a_{mi}D_{mi}) =$$

$$= t \cdot \det[A_1, \ldots, A_i, \ldots, A_m].$$

□

Ćwiczenie. Analogiczne dowody pozostałych własności pozostawimy do wykonania czytelnikowi.

Własność 18.5.7. *Wyznacznik macierzy transponowanej do danej macierzy A jest równy wyznacznikowi macierzy A.*

$$\det A = \det A^T$$

Uwaga 18.5.8. *Wyznacznik ma podobne własności ze względu na wiersze jak te przedstawione powyżej dla kolumn.*

18.6 Praktyczne sposoby obliczania wyznacznika.

Wcześniej pokazaliśmy, że odpowiedni wybór wiersza lub kolumny względem której rozwijamy wyznacznik może bardzo ułatwić obliczenia. Niestety nie we wszystkich przypadkach metoda ta okazuje się być skuteczna. Podamy teraz dwa sposoby na znaczne uproszczenie obliczeń wyznacznika, które są skuteczne we wszystkich przypadkach. Będą one polegały na stosowaniu przekształceń na wierszach i kolumnach, tak aby uzyskać macierz takiej postaci, dla której jesteśmy już w stanie łatwo obliczyć wyznacznik. Przy przekształcaniu tych macierzy możemy wykonywać mniej operacji niż przy rozwiązywaniu układów równań, nie możemy bowiem mnożyć wierszy czy kolumn przez liczbę. Tutaj możemy za to swobodnie wykonywać operacje zarówno na wierszach jak i kolumnach.

Dla łatwiejszego zapisu wprowadźmy dodatkowe oznaczenie.

Oznaczenie. Wyznacznik macierzy $A = \begin{bmatrix} a_{11} & \dots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \dots & a_{nn} \end{bmatrix}$ będziemy zapisywać w postaci

$$\det A = |A| = \begin{vmatrix} a_{11} & \dots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \dots & a_{nn} \end{vmatrix}.$$

➢ **Sposób I**

Będziemy starali się wyzerować, stosując operacje nie zmieniające wyznacznika, wszystkie wyrazy pewnej kolumny lub pewnego wiersza, z wyjątkiem być może jednego, a następnie zastosujemy rozwinięcie Laplace'a względem tej kolumny (wiersza).
Zobaczmy na przykładzie jak działa ta metoda.

✔ **Przykład.** Zastosujmy powyższy sposób do obliczenia wyznacznika macierzy

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

Zauważmy, że zgodnie z Własnością 18.5.6 możemy do wierszy dodawać wielokrotności innych wierszy, a wyznacznik nie zmieni się. Niech w_1, w_2, w_3, w_4 oznaczają kolejne wiersze macierzy A. W trakcie przekształcania macierzy będziemy obok symbolicznie notować operacje jakie na niej wykonujemy.

$$\det A = \begin{vmatrix} 1 & 7 & 3 & -2 \\ 5 & -2 & 4 & 3 \\ 6 & -1 & -3 & -2 \\ 3 & 1 & 4 & 1 \end{vmatrix} = \begin{vmatrix} 1 & 7 & 3 & -2 \\ 0 & -37 & -11 & 13 \\ 0 & -43 & -18 & 8 \\ 0 & -20 & -5 & 7 \end{vmatrix} \begin{matrix} \\ w_2 - 5w_1 \\ w_3 - 6w_1 \\ w_4 - 3w_1 \end{matrix}$$

Możemy teraz zastosować rozwinięcie względem pierwszej kolumny.

$$\det A = 1 \cdot (-1)^{1+1} \cdot \begin{vmatrix} -37 & -11 & 13 \\ -43 & -21 & 10 \\ -20 & -5 & 7 \end{vmatrix}$$

Wyznacznik macierzy rozmiaru 3×3 możemy obliczyć stosując regułę Sarrusa.

$$\det A = 1 \cdot (5439 + 2200 + 2795 - 5460 - 1850 - 3311) = -187$$

Wyznacznik tej macierzy wynosi $\det A = -187$.

Uwaga 18.6.1. *Jeśli macierz ma rozmiar większy niż 4×4, to powyższy sposób należy zastosować kilkakrotnie.*

➢ Sposób II

Sposób ten wykorzystuje własności macierzy górnotrójkątnej i dolnotrójkątnej. Polega ona na tym, że zerujemy wszystkie wyrazy powyżej lub poniżej przekątnej głównej a następnie bierzemy iloczyn wyrazów na tej przekątnej. Pokażemy to na przykładzie.

✔ **Przykład.** Obliczmy tą metodą wyznacznik macierzy A z poprzedniego przykładu.

$$\det A = \begin{vmatrix} 1 & 7 & 3 & -2 \\ 5 & -2 & 4 & 3 \\ 6 & -1 & -3 & -2 \\ 3 & 1 & 4 & 1 \end{vmatrix} = \begin{vmatrix} 1 & 7 & 3 & -2 \\ 0 & -37 & -11 & 13 \\ 0 & -43 & -18 & 8 \\ 0 & -20 & -5 & 7 \end{vmatrix} \begin{matrix} \\ w_2 - 5w_1 \\ w_3 - 6w_1 \\ w_4 - 3w_1 \end{matrix} =$$

$$= \begin{vmatrix} 1 & 7 & 3 & -2 \\ 0 & -37 & -11 & 13 \\ 0 & 0 & -\frac{304}{37} & -\frac{189}{37} \\ 0 & 0 & \frac{35}{37} & -\frac{1}{37} \end{vmatrix} \begin{matrix} \\ \\ w_3 - \frac{43}{37}w_2 \\ w_4 - \frac{20}{37}w_2 \end{matrix} = \begin{vmatrix} 1 & 7 & 3 & -2 \\ 0 & -37 & -11 & 13 \\ 0 & 0 & -\frac{304}{37} & -\frac{189}{37} \\ 0 & 0 & 0 & -\frac{187}{304} \end{vmatrix} \begin{matrix} \\ \\ \\ w_4 + \frac{35}{304}w_3 \end{matrix}$$

Taki wyznacznik możemy już łatwo obliczyć.

$$\det A = 1 \cdot (-37) \cdot (-\frac{304}{37}) \cdot (-\frac{187}{304}) = -187$$

Wynik ten zgadza się z wartością wyznacznika obliczoną przy użyciu pierwszego sposobu.

Uwaga 18.6.2. *Jeśli modyfikując macierz zamieniamy ze sobą miejscami dwie kolumny lub dwa wiersze, to wyznacznik zmienia się na przeciwny. Jeśli natomiast mnożymy wiersz lub kolumnę przez liczbę, to wyznacznik musimy podzielić przez tę liczbę.*

✔ **Przykład.** Wykorzystanie pierwszej części tej uwagi pomaga przy obliczeniu następującego wyznacznika, w którym zamienimy miejscami drugą i czwartą kolumnę

$$\begin{vmatrix} 2 & 5 & 3 & 1 \\ 0 & 2 & -3 & -1 \\ 0 & 4 & 7 & 0 \\ 0 & 3 & 0 & 0 \end{vmatrix} = -\begin{vmatrix} 2 & 1 & 3 & 5 \\ 0 & -1 & -3 & 2 \\ 0 & 0 & 7 & 4 \\ 0 & 0 & 0 & 3 \end{vmatrix} = -[2 \cdot (-1) \cdot 7 \cdot 3] = -(-42) = 42.$$

18.7 Cramerowski układ równań liniowych.

W tym podrozdziale zajmiemy się jedną z własności jaką chcieliśmy aby spełniał wyznacznik. Rozważmy układ

$$(\bigstar) \quad x_1 \cdot A_1 + x_2 \cdot A_2 + \ldots + x_n \cdot A_n = B,$$

gdzie $A_i \in R^n, B \in R^n$, który jest układem n równań z n niewiadomymi $x_1, x_2, \ldots, x_n$, natomiast A_i, B są kolumnowymi wektorami utworzonymi ze współczynników tego układu. Macierzą główną dla układu $(\bigstar)$ jest macierz $A = [A_1, A_2, \ldots, A_n]$, zaś wyznacznik główny to $W = \det A$.

Definicja 18.7.1 (układ cramerowski). Układ $(\bigstar)$ nazywamy cramerowskim jeśli wyznacznik główny W tego układu jest niezerowy:

$$W = \det A \neq 0.$$

Podobnie jak w przypadku układu trzech równań z trzema niewiadomymi, dla układu $(\bigstar)$ możemy oprócz macierzy głównej i jej wyznacznika określić również macierze pomocnicze i ich wyznaczniki. Ze zmienną x_1 związana jest macierz P_{x_1}, która powstaje przez zastąpienie pierwszej kolumny macierzy głównej przez kolumnę wyrazów wolnych B, tak więc $P_{x_1} = [B, A_2, \ldots, A_n]$. Wyznacznik tej macierzy $W_{x_1} = \det P_{x_1}$ wykorzystywany będzie do obliczenia zmiennej x_1. Podobnie powstają i wykorzystywane są inne macierze pomocnicze. Mamy więc
$P_{x_1} = [B, A_2, A_3, \ldots, A_n]$
$P_{x_2} = [A_1, B, A_3, \ldots, A_n]$
$P_{x_3} = [A_1, A_2, B, \ldots, A_n]$
$\vdots$
$P_{x_n} = [A_1, A_2, \ldots, A_{n-1}, B]$
a odpowiednie wyznaczniki to

$$W_{x_i} = \det P_{x_i}.$$

Możemy już teraz sformułować twierdzenie, które pozwoli nam rozwiązywać układy równań liniowych.

Twierdzenie 18.7.2 (Cramera). *Układ $(\bigstar)$, gdy $W \neq 0$ ma dokładnie jedno rozwiązanie dane wzorami*

$$x_i = \frac{W_{x_i}}{W}$$

dla $i = 1, 2, \ldots, n$, gdzie $W = \det A$ jest wyznacznikiem głównym tego układu, natomiast W_{x_i} są odpowiednimi wyznacznikami pomocniczymi.

Dowód tego twierdzenia jest analogiczny do dowodu dla układu cramerowskiego rozmiaru 3×3 przedstawionego w Rozdziale 4, tutaj go pominiemy.

✔ **Przykład.** Zastosujmy twierdzenie Cramera do rozwiązania następującego układu równań

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

Wyznacznik główny tego układu to

$$W = \begin{vmatrix} 2 & 3 & 4 & 5 \\ -3 & 4 & -1 & -2 \\ -1 & -1 & 3 & -1 \\ 5 & -4 & 2 & 3 \end{vmatrix} = -140.$$

Wyznacznik główny jest niezerowy a więc układ jest układem Cramerowskim i ma jedno rozwiązanie. Odpowiednie wyznaczniki pomocnicze to

$$W_{x_1} = \begin{vmatrix} 3 & 3 & 4 & 5 \\ -1 & 4 & -1 & -2 \\ -6 & -1 & 3 & -1 \\ 4 & -4 & 2 & 3 \end{vmatrix} = -280, \quad W_{x_2} = \begin{vmatrix} 2 & 3 & 4 & 5 \\ -3 & -1 & -1 & -2 \\ -1 & -6 & 3 & -1 \\ 5 & -4 & 2 & 3 \end{vmatrix} = -140,$$

$$W_{x_3} = \begin{vmatrix} 2 & 3 & 3 & 5 \\ -3 & 4 & -1 & -2 \\ -1 & -1 & -6 & -1 \\ 5 & -4 & 4 & 3 \end{vmatrix} = 140, \qquad W_{x_4} = \begin{vmatrix} 2 & 3 & 4 & 3 \\ -3 & 4 & -1 & -1 \\ -1 & -1 & 3 & -6 \\ 5 & -4 & 2 & 3 \end{vmatrix} = 0.$$

rozwiązaniem naszego układu jest zatem czwórka liczb

$$x_1 = \frac{W_{x_1}}{W} = 2, \quad x_2 = \frac{W_{x_2}}{W} = 1, \quad x_3 = \frac{W_{x_3}}{W} = -1, \quad x_4 = \frac{W_{x_4}}{W} = 0.$$

18.8 Wyznacznik a liniowa niezależność.

Związek wyznacznika z liniową niezależnością wektorów jest konsekwencją faktów, które sformułujemy i udowodnimy poniżej.

Fakt 18.8.1. *Jeśli kolumny $A_1, A_2, \dots A_n$ macierzy A rozmiaru $n \times n$ są liniowo zależne, to* $\det A = 0$.

Dowód. Przypomnijmy, że wektory są liniowo zależne, jeśli jeden z nich jest liniową kombinacją pozostałych. Skoro wektory $A_1, A_2, \dots A_n$ mają być liniowo zależne, to jeden z nich, przyjmijmy A_n, jest liniową kombinacją pozostałych. Jeśli inna kolumna jest kombinacją pozostałych, rozumowanie będzie podobne. Tak więc mamy $A_n = t_1A_1 + t_2A_2 + \ldots + t_{n-1}A_{n-1}$. Wówczas wyznacznik tej macierzy przyjmuje postać

$$\det A = \det[A_1, A_2, \dots A_n] = \det[A_1, A_2, \dots, A_{n-1}, t_1A_1 + t_2A_2 + \ldots + t_{n-1}A_{n-1}]$$

Z Własności 18.5.6 wiemy, że odjęcie od kolumny wielokrotności innej kolumny nie zmienia wyznacznika tak więc odejmujemy od n-tej kolumny iloczyn t_1A_1, następnie t_2A_2 i tak kolejno aż do $t_{n-1}A_{n-1}$, otrzymujemy wówczas

$$\det A = \det[A_1, A_2, \dots, A_{n-1}, t_2A_2 + \ldots + t_{n-1}A_{n-1}] =$$

$$= \det[A_1, A_2, \dots, A_{n-1}, t_3A_3 + \ldots + t_{n-1}A_{n-1}] = \ldots = \det[A_1, A_2, \dots, A_{n-1}, O] = 0.$$

□

Fakt 18.8.2. *Jeśli kolumny $A_1, A_2, \dots, A_n$ macierzy A są liniowo niezależne w R^n, to $det A \neq 0$.*

Dowód tego faktu pomijamy.

Z Faktów 18.8.1 i 18.8.2 możemy wyprowadzić następujący wniosek.

Wniosek 18.8.3. *Wektory* $A_1, A_2, \ldots, A_n$ *z* R^n *są liniowo niezależne wtedy i tylko wtedy gdy* $\det[A_1, A_2, \ldots, A_n] \neq 0$.

18.9 Działania na macierzach.

W podrozdziale tym określimy podstawowe działania jakie możemy wykonywać na macierzach. Podamy również ich podstawowe własności.

➢ **Iloczyn liczby i macierzy**

Iloczynem $t \cdot [a_{ij}]$ liczby rzeczywistej t i macierzy $[a_{ij}]$ nazywamy macierz $[b_{ij}]$ tego samego rozmiaru co macierz $[a_{ij}]$, dla której $b_{ij} = t \cdot a_{ij}$ dla każdego i, j. Stosowany bywa również skrótowy zapis $t \cdot [a_{ij}] = [t \cdot a_{ij}]$.

✔ **Przykład.** Iloczynem liczby $t = -2$ i macierzy $A = \begin{bmatrix} 2 & -4 & -1 & 3 \\ 5 & 0 & -7 & 1 \end{bmatrix}$ jest macierz B postaci $B = (-2) \cdot \begin{bmatrix} 2 & -4 & -1 & 3 \\ 5 & 0 & -7 & 1 \end{bmatrix} = \begin{bmatrix} -4 & 8 & 2 & -6 \\ -10 & 0 & 14 & -2 \end{bmatrix}$.

➢ **Suma macierzy**

Jeżeli macierze $[a_{ij}]$ i $[b_{ij}]$ mają ten sam rozmiar, to ich sumą $[a_{ij}] + [b_{ij}]$ nazywamy macierz $[c_{ij}]$ tego samego rozmiaru, dla której $c_{ij} = a_{ij} + b_{ij}$, dla wszystkich i, j. Piszemy też $[a_{ij}] + [b_{ij}] = [a_{ij} + b_{ij}]$.

✔ **Przykład.** Sumą macierzy $A = \begin{bmatrix} 2 & 4 \\ -3 & 0 \\ 1 & -2 \end{bmatrix}$, $B = \begin{bmatrix} -3 & -2 \\ 1 & -6 \\ 0 & 1 \end{bmatrix}$ jest macierz

$C = A + B = \begin{bmatrix} 2 & 4 \\ -3 & 0 \\ 1 & -2 \end{bmatrix} + \begin{bmatrix} -3 & -2 \\ 1 & -6 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} -1 & 2 \\ -2 & -6 \\ 1 & -1 \end{bmatrix}$.

➢ **Iloczyn macierzy**

Iloczynem macierzy $A = [a_{ij}]$ rozmiaru $m \times n$ i macierzy $B = [b_{ij}]$ rozmiaru $n \times k$ nazywamy macierz $C = [c_{ij}]$ rozmiaru $m \times k$, dla której

$$c_{ij} = a_{i1}b_{1j} + a_{i2}b_{2j} + \ldots + a_{in}b_{nj}.$$

Uwaga 18.9.1. *Mnożenie macierzy* A *i* B *jest wykonalne tylko wtedy gdy macierz* A *ma tyle samo kolumn co macierz* B *wierszy. Zauważmy, że wyraz* c_{ij} *macierzy* C *jest iloczynem skalarnym* i*-tego wiersza macierzy* A *i* j*-tej kolumny macierzy* B.

✔ **Przykład.** Iloczynem macierzy $A = \begin{bmatrix} 2 & 1 & -3 & 4 \\ -7 & 1 & 2 & 0 \end{bmatrix}$ i macierzy $B = \begin{bmatrix} 3 & -7 \\ 4 & 2 \\ -8 & 1 \\ 1 & -3 \end{bmatrix}$

jest macierz

$$C = \begin{bmatrix} 2 & 1 & -3 & 4 \\ -7 & 1 & 2 & 0 \end{bmatrix} \cdot \begin{bmatrix} 3 & -7 \\ 4 & 2 \\ -8 & 1 \\ 1 & -3 \end{bmatrix} =$$

$$= \begin{bmatrix} 2\cdot 3 + 1\cdot 4 + (-3)\cdot(-8) + 4\cdot 1 & 2\cdot(-7) + 1\cdot 2 + (-3)\cdot 1 + 4\cdot(-3) \\ (-7)\cdot 3 + 1\cdot 4 + 2\cdot(-8) + 0\cdot 1 & (-7)\cdot(-7) + 1\cdot 2 + 2\cdot 1 + 0\cdot(-3) \end{bmatrix} = \begin{bmatrix} 38 & -27 \\ -33 & 53 \end{bmatrix}.$$

➢ **Własności działań na macierzach**

Podamy teraz podstawowe własności działań na macierzach. A, B, C oznaczać będą macierze, natomiast t to dowolna liczba rzeczywista.

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

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

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

(4) $(t \cdot A) \cdot B = t \cdot (A \cdot B) = A \cdot (t \cdot B)$

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

✍ Dowody czterech pierwszych własności wynikają z własności działań na liczbach rzeczywistych i z określenia działań na macierzach. Pozostawimy je do wykonania czytelnikowi. Dowód piątej własności jest dużo bardziej skomplikowany. Tutaj go pominiemy.

18.10 Macierz odwrotna.

W tym podrozdziale wprowadzimy pojęcie macierzy odwrotnej, zostanie ono zdefiniowane w sposób analogiczny, jak to zostało zrobione w przypadku macierzy rozmiaru 3×3. Macierz odwrotna będzie przez nas wielokrotnie wykorzystywana w kolejnych rozdziałach, np. przy badaniu przekształceń przestrzeni wektorowych.

Definicja 18.10.1 (macierz jednostkowa). Macierzą jednostkową rozmiaru $n \times n$ nazywamy macierz I postaci

$$I = I_{n\times n} = \begin{bmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \dots & 0 & 1 \end{bmatrix}.$$

Definicja 18.10.2 (macierz odwrotna). Macierzą odwrotną do macierzy A rozmiaru $n \times n$ nazywamy taką macierz B rozmiaru $n \times n$, że $A \cdot B = B \cdot A = I_{n \times n}$. Macierz B oznaczamy symbolem A^{-1}.

Szukanie macierzy odwrotnej, przy użyciu definicji wiązałoby się z rozwiązaniem n układów złożonych z n równań z n niewiadomymi (łącznie n^2 niewiadomych). Poniższe twierdzenie podaje łatwiejszy sposób na wyznaczenie macierzy odwrotnej.

Twierdzenie 18.10.3. *Dla macierzy A rozmiaru $n \times n$ mamy*

(a) Jeśli $\det A = 0$*, to A nie ma macierzy odwrotnej.*

(b) Jeśli $\det A \neq 0$*, to macierz odwrotna do macierzy A istnieje, jest jedyna i wyraża się wzorem $A^{-1} = [b_{ij}]$, gdzie $b_{ij} = \frac{1}{\det A} D_{ji}$, czyli*

$$\boxed{A^{-1} = \frac{1}{\det A} \begin{bmatrix} D_{11} & D_{21} & \dots & D_{n1} \\ D_{12} & D_{22} & \dots & D_{n2} \\ \vdots & \vdots & \ddots & \vdots \\ D_{1n} & D_{2n} & \dots & D_{nn} \end{bmatrix}.}$$

Uwaga 18.10.4. *Zauważmy, że do wyznaczenia wyrazu b_{ij} używamy wyrażenia D_{ji}. Następuje tu zamiana kolejności występowania indeksów.*

Dowód.

(a) Gdyby istniała macierz odwrotna A^{-1}, to z Własności 18.5.7 mielibyśmy

$$\det A \cdot \det(A^{-1}) = \det(A \cdot A^{-1}) = \det(I_{n \times n}) = 1.$$

Zatem otrzymalibyśmy

$$0 \cdot \det(A^{-1}) = 1.$$

Ponieważ takie równanie nie ma rozwiązań, więc w tym wypadku nie istnieje macierz A^{-1}.

(b) Niech $B = [b_{ij}]$ będzie macierzą określoną wzorem z części (b) Twierdzenia 18.10.3. Obliczmy $A \cdot B = [c_{ij}]$.

$$c_{ij} = \sum_{k=1}^{n} a_{ik} b_{kj} = \sum_{k=1}^{n} a_{ik} \cdot \frac{1}{\det A} \cdot D_{jk} = \frac{1}{\det A} \sum_{k=1}^{n} a_{ik} \cdot (-1)^{j+k} \cdot \det A_{jk}$$

Rozważmy dwa przypadki.
1. Gdy $i = j$ wówczas

$$c_{ii} = \frac{1}{\det A} \sum_{k=1}^{n} a_{ik} \cdot (-1)^{i+k} \cdot \det A_{ik}$$

Zauważmy, że w powyższym wyrażeniu występuje rozwinięcie Laplace'a wyznacznika macierzy A względem i-tego wiersza. Otrzymujemy zatem

$$c_{ii} = \frac{1}{\det A} \cdot \det A = 1.$$

2. Gdy $i \neq j$ wówczas

$$c_{ij} = \frac{1}{\det A} \sum_{k=1}^{n} a_{ik} \cdot (-1)^{j+k} \cdot \det A_{jk}.$$

Wyrażenie to zawiera rozwinięcie względem j-tego wiersza wyznacznika macierzy $\begin{array}{cl} \left[\begin{array}{c} w_1 \\ \vdots \\ w_i \\ \vdots \\ w_i \\ \vdots \\ w_n \end{array}\right] & \begin{array}{l} \\ \\ i\text{–}ty\ wiersz \\ \\ j\text{–}ty\ wiersz \\ \\ \\ \end{array} \end{array}$. Jak wiemy wyznacznik macierzy, która ma dwa takie same wiersze jest zerowy, więc

$$c_{ij} = \frac{1}{\det A} \cdot 0 = 0.$$

Zatem

$$A \cdot B = [c_{ij}] = I_{n\times n}.$$

Dowód tego, że $B \cdot A = I_{n\times n}$ wygląda analogicznie. Tutaj go pominiemy. Zatem B jest macierzą odwrotną do A, czyli $B = A^{-1}$.

□

✔ **Przykład.** Sprawdźmy, czy wzór macierzy odwrotnej do macierzy rozmiaru 2×2, wyznaczony tym sposobem, zgadza się ze znanym nam już wzorem. Dla macierzy $A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$ macierz odwrotna to

$$A^{-1} = \frac{1}{\det A} \begin{bmatrix} D_{11} & D_{21} \\ D_{12} & D_{22} \end{bmatrix} = \frac{1}{\det A} \begin{bmatrix} a_{22} & -a_{12} \\ -a_{21} & a_{11} \end{bmatrix}.$$

Zatem rzeczywiście zgadza się ze znanym wzorem.

✍ **Ćwiczenie.** Podobne sprawdzenie dla macierzy rozmiaru 3×3 pozostawimy do wykonania czytelnikowi.

Rozdział 19

Abstrakcyjne przestrzenie wektorowe

Do tej pory zajmowaliśmy się przestrzeniami R^n, a w szczególności przestrzeniami R^2 i R^3. Przestrzenie R^n są tylko przykładami dużo szerszej klasy - przestrzeni wektorowych. W rozdziale tym zdefiniujemy pojęcie przestrzeni wektorowych, poznamy przykłady. Dowiemy się, czym są podprzestrzenie wektorowe. Zajmiemy się również abstrakcyjnym pojęciem liniowej niezależności.

19.1 Określenie przestrzeni wektorowych.

Przestrzenie wektorowe będziemy najczęściej oznaczać przez U, V, W, natomiast elementy tych przestrzeni, czyli abstrakcyjne wektory, będziemy oznaczać v, w, u, $\overrightarrow{a}$, $\overrightarrow{e}$, $\overrightarrow{f}$.

Definicja 19.1.1 (**przestrzeń wektorowa**). Zbiór V pewnego rodzaju obiektów v nazywamy przestrzenią wektorową, jeśli obiekty te możemy dodawać i mnożyć przez liczby. Muszą być przy tym spełnione następujące własności:

(a) przemienność dodawania, czyli że dla dowolnych wektorów $v_1, v_2 \in V$ mamy

$$v_1 + v_2 = v_2 + v_1,$$

(b) łączność dodawania, czyli że dla dowolnych wektorów $v_1, v_2, v_3 \in V$ mamy

$$(v_1 + v_2) + v_3 = v_1 + (v_2 + v_3),$$

(c) rozdzielność mnożenia względem dodawania, czyli że dla dowolnych wektorów $v_1, v_2 \in V$ oraz dowolnej liczby rzeczywistej t mamy

$$t(v_1 + v_2) = tv_1 + tv_2,$$

(d) rozdzielność mnożenia względem dodawania, czyli że dla dowolnego wektora $v \in V$ i dowolnych liczb rzeczywistych t, s mamy

$$(t + s)v = tv + sv,$$

(e) łączność mnożenia, czyli że dla dowolnego wektora $v \in V$ i dowolnych liczb rzeczywistych t, s mamy

$$(t \cdot s) \cdot v = t \cdot (s \cdot v),$$

(f) istnieje tzw. wektor zerowy $\overrightarrow{O} \in V$ taki, że dla każdego wektora $v \in V$ zachodzi

$$\overrightarrow{O} + v = v,$$

(g) dla dowolnego wektora $v \in V$ mamy

$$1 \cdot v = v \quad \text{oraz} \quad 0 \cdot v = \overrightarrow{O}.$$

Z powyższych własności możemy wyprowadzić dalsze.

Własność 19.1.2. *Każdy wektor ma wektor przeciwny.*

Dowód. Dla wektora $v \in V$ wektorem przeciwnym jest taki wektor $-v \in V$, że

$$v + (-v) = \overrightarrow{O}.$$

Określmy wektor $-v$ jako

$$-v = (-1) \cdot v.$$

Mamy wówczas

$$v + (-v) = v + (-1) \cdot v,$$

możemy teraz skorzystać z podanych własności przestrzeni V, a w szczególności z własności (g) i własności (d). Otrzymujemy

$$v + (-1) \cdot v = 1 \cdot v + (-1) \cdot v = (1 + (-1)) \cdot v = 0 \cdot v = \overrightarrow{O}.$$

Wektor $-v$ jest rzeczywiście wektorem przeciwnym do wektora v, zatem dla dowolnego wektora da się taki wektor znaleźć. □

Własność 19.1.3. *Wektory w przestrzeni wektorowej możemy odejmować.*

Dowód. Odejmowanie wektorów możemy określić jako dodawanie wektora przeciwnego. Dla wektorów $v_1, v_2 \in V$ ich różnica to

$$v_1 - v_2 = v_1 + (-v_2).$$

Z faktu istnienia wektora przeciwnego i określoności sumy wektorów wynika wykonalność tego działania. □

Podamy teraz przykłady przestrzeni wektorowych, wraz z określeniem, w jaki sposób dodawać i mnożyć przez liczbę wektory z tej przestrzeni.

✔ **Przykład.**

(1) Przestrzeń R^n z działaniami określonymi w poprzednich rozdziałach.

(2) Przestrzeń R^∞ złożona z nieskończonych ciągów $(a_n) = (a_1, a_2, \ldots, a_n, \ldots)$ o wyrazach rzeczywistych. Pojedynczym elementem tej przestrzeni jest cały ciąg (a_n). Dodawanie dwóch elementów tej przestrzeni, czyli dwóch ciągów odbywa się poprzez dodanie do siebie kolejnych wyrazów obu ciągów, dla ciągów $(a_n), (b_n)$ mamy

$$(a_1, a_2, a_3 \ldots) + (b_1, b_2, b_3, \ldots) = (a_1 + b_1, a_2 + b_2, a_3 + b_3, \ldots),$$

co krócej możemy zapisać jako

$$(a_n) + (b_n) = (a_n + b_n).$$

Podobnie wykonuje się mnożenie przez liczbę. Dla liczby t i wektora (a_n) mamy

$$t \cdot (a_1, a_2, a_3, \ldots) = (t \cdot a_1, t \cdot a_2, t \cdot a_3, \ldots),$$

a krócej

$$t(a_n) = (ta_n).$$

Wektorem zerowym w tej przestrzeni jest ciąg stale równy zero

$$\overrightarrow{O} = (0, 0, 0, \ldots).$$

(3) Przestrzeń macierzy rozmiaru $n \times m$ o wyrazach rzeczywistych, symbolicznie oznaczmy ją $\mathcal{M}_{n\times m}$. Dodawanie i mnożenie przez liczbę w tej przestrzeni odbywa się wyraz po wyrazie, natomiast wektorem zerowym jest macierz zerowa $\begin{pmatrix} 0 & \ldots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \ldots & 0 \end{pmatrix}$.

(4) Przestrzeń $R_n[X]$ złożona z wielomianów stopnia mniejszego bądź równego n, o współczynnikach rzeczywistych. Dodawanie wektorów z tej przestrzeni odbywa się poprzez dodanie odpowiednich współczynników. Sumą wektorów $a_nx^n + a_{n-1}x^{n-1} \ldots + a_1x + a_0$ oraz $b_nx^n + b_{n-1}x^{n-1} \ldots + b_1x + b_0$ jest wektor

$$(a_nx^n + a_{n-1}x^{n-1} \ldots + a_1x + a_0) + (b_nx^n + b_{n-1}x^{n-1} \ldots + b_1x + b_0) =$$
$$= (a_n + b_n)x^n + (a_{n-1} + b_{n-1})x^{n-1} + \ldots + (a_1 + b_1)x + (a_0 + b_0).$$

Wektorem zerowym w tej przestrzeni jest wielomian zerowy tzn. funkcja tożsamościowo równa zero.

(5) Przestrzeń $C[0, 1]$ złożona z funkcji ciągłych określonych na odcinku $[0, 1]$ o wartościach rzeczywistych. Dla dowolnych elementów f, g z tej przestrzeni ich sumą nazywamy funkcję daną wzorem

$$(f + g)(x) = f(x) + g(x),$$

natomiast iloczynem funkcji f i liczby t jest funkcja

$$(tf)(x) = t \cdot f(x).$$

Wektorem zerowym w tej przestrzeni jest funkcja stała równa 0.

(6) Przestrzeń trywialna V składająca się tylko z jednego wektora $V = \{\overrightarrow{O}\}$.

✍ **Ćwiczenie.** Sprawdzenie, że tak określone przestrzenie wektorowe spełniają wszystkie własności podane w Definicji 19.1.1 pozostawimy do wykonania czytelnikowi.

19.2 Podprzestrzenie wektorowe.

W danej przestrzeni wektorowej możemy wyróżniać pewne podzbiory. W tym podrozdziale zajmiemy się szczególnym rodzajem takich podzbiorów - podprzestrzeniami wektorowymi. Poznamy definicję tego pojęcia, zbadamy jego własności. W kolejnym podrozdziale będziemy również badać podprzestrzenie, jakie możemy tworzyć z danych wektorów przestrzeni.

Definicja 19.2.1 (podprzestrzeń wektorowa). Niepusty podzbiór W przestrzeni wektorowej V jest podprzestrzenią wektorową, jeśli suma wektorów z W zawsze należy do W oraz iloczyn każdego wektora z W przez dowolną liczbę też należy do W.

Uwaga 19.2.2. *W podprzestrzeni W własności (a)-(g) z Definicji 19.1.1 są automatycznie spełnione, tak więc podprzestrzeń W sama jest przestrzenią wektorową.*

Oznaczenie. Fakt, że przestrzeń W jest podprzestrzenią przestrzeni V będziemy oznaczać

$$W < V.$$

Ze sposobu określenia podprzestrzeni wektorowej możemy wywnioskować następujące własności.

Własność 19.2.3. *Wektor zerowy należy do każdej podprzestrzeni $W < V$.*

Dowód. Weźmy wektor w z podprzestrzeni W. Zgodnie z definicją podprzestrzeni wówczas również wektor $0 \cdot w$ należy do tej podprzestrzeni, ale $0 \cdot w = \overrightarrow{O}$ tak więc $\overrightarrow{O} \in W$. □

Własność 19.2.4. *Jeśli $w_1, w_2, \ldots, w_k$ są wektorami z podprzestrzeni W, to dowolna kombinacja liniowa $t_1w_1 + t_2w_2 + \ldots + t_kw_k$ również należy do tej podprzestrzeni.*

Dowód. Jeśli w_1 należy do podprzestrzeni W, to również t_1w_1 należy do tej podprzestrzeni, podobnie $t_2w_2, t_3w_3, \ldots, t_nw_n$ należą do podprzestrzeni W. Z Definicji 19.2.1 wiemy, że wówczas również suma tych wektorów należy do tej podprzestrzeni, tak więc

$$t_1w_1 + t_2w_2 + \ldots + t_nw_n \in W.$$

□

✔ **Przykład.**

(1) Niech $W = \{[x, y, 0, 0] \in R^4 : x \in R, y \in R\}$. Sprawdźmy, że W jest podprzestrzenią przestrzeni R^4. Rzeczywiście W jest podzbiorem R^4. Ponadto, dla dowolnych wektorów z W mamy również

$$[x_1, y_1, 0, 0] + [x_2, y_2, 0, 0] = [x_1 + x_2, y_1 + y_2, 0, 0] \in W,$$

$$t \cdot [x, y, 0, 0] = [tx, ty, 0, 0].$$

(2) Rozważmy podzbiór W przestrzeni R^∞ składający się dokładnie z tych ciągów, których granica jest równa 0. Pokażemy, że W jest podprzestrzenią R^∞. Jak wiemy dla dowolnych ciągów $(a_n),(b_n)$ i dowolnej liczby t mamy

$$\lim_{n\to\infty}(a_n+b_n)=\lim_{n\to\infty}a_n+\lim_{n\to\infty}b_n,$$

$$\lim_{n\to\infty}ta_n=t\cdot\lim_{n\to\infty}a_n.$$

Weźmy dowolne ciągi $(a_n),(b_n)$ należące do W, czyli takie których granica jest równa 0. Wówczas granica sumy tych ciągów to

$$\lim_{n\to\infty}(a_n+b_n)=\lim_{n\to\infty}a_n+\lim_{n\to\infty}b_n=0,$$

a granica iloczynu ciągu (a_n) i dowolnej liczby t wynosi

$$\lim_{n\to\infty}ta_n=t\lim_{n\to\infty}a_n=t\cdot 0=0.$$

Tak więc zarówno (a_n+b_n) jak i $t(a_n)$ należą do W, stąd

$$W<R^\infty.$$

(3) Podprzestrzenią przestrzeni $\mathcal{M}_{n\times m}$ jest podzbiór W, złożony z macierzy rozmiaru $n\times m$, w których pierwsza kolumna jest zerowa.

(4) Podprzestrzenią przestrzeni $C[0,1]$ jest podzbiór W składający się z funkcji ciągłych na odcinku $[0,1]$, których wartość w punkcie 0 jest równa 0.

✍ **Ćwiczenie.** Sprawdzenie, że występujące w przykładach 3 i 4 podzbiory są rzeczywiście podprzestrzeniami podanych przestrzeni, pozostawimy do wykonania czytelnikowi.

19.3 Podprzestrzenie generowane przez układ wektorów.

Niech V będzie przestrzenią wektorową , a $v_1,v_2,\dots,v_k$ będą wektorami z tej przestrzeni. Rozważmy zbiór składający się z wszystkich wektorów będących kombinacją liniową wektorów $v_1,v_2,\dots,v_k$. Zbiór taki będziemy oznaczać $Lin\{v_1,v_2,\dots,v_k\}$, tak więc

$$Lin\{v_1,v_2,\dots,v_k\}=\{t_1v_1+t_2v_2+\ldots+t_kv_k\ :\ t_1,t_2,\dots,t_k\in R\}.$$

Fakt 19.3.1. *Zbiór $Lin\{v_1,v_2,\dots,v_k\}$ jest podprzestrzenią wektorową przestrzeni V.*

Dowód. Sprawdźmy czy suma dwóch wektorów z tego zbioru również do niego należy. Weźmy $t_1v_1+t_2v_2+\ldots+t_kv_k,\ s_1v_1+s_2v_2+\ldots+s_kv_k\in Lin\{v_1,v_2,\dots,v_k\}$. Wówczas ich suma

$$(t_1v_1+t_2v_2+\ldots+t_kv_k)+(s_1v_1+s_2v_2+\ldots+s_kv_k)=(t_1+s_1)v_1+(t_2+s_2)v_2\ldots+(t_k+s_k)v_k$$

również jest kombinacją liniową wektorów $v_1, v_2, \ldots, v_k$ a więc jest wektorem z $Lin\{v_1, v_2, \ldots, v_k\}$. Sprawdźmy jeszcze, czy wynik iloczynu również należy do tego zbioru.

$$t \cdot (t_1 v_1 + t_2 v_2 + \ldots + t_k v_k) = (tt_1)v_1 + (tt_2)v_2 \ldots + (tt_k)v_k$$

a więc również jest wektorem z $Lin\{v_1, v_2, \ldots, v_k\}$. Zgodnie z Definicją 19.2.1 zbiór $Lin\{v_1, v_2, \ldots, v_k\}$ jest podprzestrzenią przestrzeni V. □

Uwaga 19.3.2 (podprzestrzeń generowana). *Przestrzeń $Lin\{v_1, v_2, \ldots, v_k\}$ jest najmniejszą podprzestrzenią V, do której należą $v_1, v_2, \ldots, v_k$. Nazywamy ją podprzestrzenią generowaną przez wektory $v_1, v_2, \ldots, v_k$. Czasem mówimy też, że przestrzeń $Lin\{v_1, v_2, \ldots, v_k\}$ jest rozpinana przez wektory $v_1, v_2, \ldots, v_k$.*

✔ **Przykład.**

(1) Weźmy przestrzeń $V = R^3$ i wektory $v_1 = [1, 0, 0]$, $v_2 = [0, 1, 0]$ z tej przestrzeni. Przestrzeń $Lin\{v_1, v_2\}$ generowana przez te wektory to

$$Lin\{v_1, v_2\} = \{t_1 v_1 + t_2 v_2 \ : \ t_1, t_2 \in R\} = \{[t_1, t_2, 0] \ : \ t_1, t_2 \in R\},$$

jest to zatem płaszczyzna O_{xy}.

(2) Niech V będzie przestrzenią funkcji ciągłych określonych na zbiorze liczb rzeczywistych, czyli $V = C(R)$. Weźmy wektory z tej przestrzeni $f_1(x) = 1$, $f_2(x) = x$, $f_3(x) = x^2$ wówczas podprzestrzeń generowana przez te wektory to

$$Lin\{f_1, f_2, f_3\} = \{a_1 + a_2 x + a_3 x^2 \ : \ a_1, a_2, a_3 \in R\},$$

jest to zatem przestrzeń złożona z wszystkich wielomianów stopnia co najwyżej drugiego.

(3) Niech V będzie dowolną przestrzenią R^n, natomiast v niezerowym wektorem z tej przestrzeni. Wówczas podprzestrzeń generowana przez ten wektor składa się z wektorów będących wielokrotnościami v, jest to zatem prosta wzdłuż tego wektora

$$Lin\{v\} = \{tv \ : \ t \in R\}.$$

Uwaga 19.3.3. *Zauważmy, że podprzestrzenią generowaną przez zbiór pusty jest*

$$Lin(\emptyset) = \{0\}.$$

19.4 Abstrakcyjna liniowa niezależność.

W przestrzeniach wektorowych, podobnie jak w przestrzeniach R^n możemy określić liniową niezależność wektorów.

Definicja 19.4.1 (liniowa niezależność wektorów). Układ $v_1, v_2, \ldots, v_k$ wektorów z przestrzeni V jest liniowo niezależny, jeśli

$$\begin{aligned} &v_1 \neq \overrightarrow{O} \\ &v_2 \notin Lin\{v_1\} \\ &v_3 \notin Lin\{v_1, v_2\} \\ &\vdots \\ &v_k \notin Lin\{v_1, v_2, \ldots, v_{k-1}\} \end{aligned}$$

W przeciwnym wypadku mówimy, że układ $v_1, v_2, \ldots, v_k$ jest liniowo zależny.

Przykład. Układ złożony z funkcji $\mathbb{1}$, x, x^2, x^3 z przestrzeni $R_3[x]$ wielomianów zmiennej x stopnia co najwyżej trzeciego, jest liniowo niezależny. Zauważmy bowiem, że

$$\mathbb{1} \neq 0.$$

Przestrzeń $Lin\{\mathbb{1}\} = \{t \ : t \in R\}$ jest to zbiór złożony z funkcji stałych. Mamy zatem

$$x \notin Lin\{\mathbb{1}\}.$$

Przestrzeń $Lin\{\mathbb{1}, x\} = \{t_1x + t_0 \ : t_1, t_0 \in R\}$ jest to zbiór funkcji liniowych zmiennej x (t_1, t_2 to współczynniki równania funkcji), tak więc

$$x^2 \notin Lin\{\mathbb{1}, x\}.$$

Przestrzeń $Lin\{\mathbb{1}, x, x^2\} = \{t_2x^2 + t_1x + t_0 \ : t_0, t_1, t_2 \in R\}$ jest zbiorem złożonym z wielomianów stopnia co najwyżej drugiego, więc

$$x^3 \notin Lin\{\mathbb{1}, x, x^2\}.$$

Zgodnie z Definicją 19.4.1 układ $\mathbb{1}, x, x^2, x^3$ jest zatem liniowo niezależny.

Taki sposób sprawdzania liniowej niezależności wektorów jest bardzo uciążliwy. Poniżej przedstawimy lemat, który okaże się być łatwiejszą metodą sprawdzania liniowej niezależności. Lemat ten przyjmuje analogiczną postać jak kryteria sprawdzania liniowej niezależności w przypadku przestrzeni R^n.

Lemat 19.4.2. *Układ $v_1, v_2, \ldots, v_k$ jest liniowo niezależny wtedy i tylko wtedy gdy jedyna kombinacja liniowa wektorów $v_1, v_2, \ldots, v_k$ dająca wektor $\overrightarrow{O}$, to kombinacja ze wszystkimi współczynnikami zerowymi, co symbolicznie możemy zapisać*

$$t_1v_1 + t_2v_2 + \ldots + t_kv_k = 0 \Leftrightarrow t_1 = t_2 = \cdots = t_k = 0.$$

Dowód. Załóżmy, że wektory $v_1, v_2, \ldots, v_k$ tworzą układ liniowo niezależny. Pokażemy, że wówczas jedyna kombinacja liniowa tych wektorów dająca wektor zerowy, to kombinacja ze wszystkimi współczynnikami zerowymi. Załóżmy nie wprost, że $t_1v_1 + t_2v_2 + \ldots + t_kv_k = 0$ i jeden ze współczynników, np. t_1 jest niezerowy. Wówczas mamy

$$t_1v_1 = -t_2v_2 - t_3v_3 - \ldots - t_kv_k,$$

co po obustronnym podzieleniu przez t_1 daje

$$v_1 = -\frac{t_2}{t_1}v_2 - \frac{t_3}{t_1}v_3 - \ldots - \frac{t_k}{t_1}v_k.$$

Z powyższej równości wynika, że v_1 jest kombinacją liniową wektorów $v_2, v_3 \ldots, v_k$, należy więc do zbioru $Lin\{v_2, v_3, \ldots, v_k\}$, co przeczy założeniu o liniowej niezależności wektorów $v_1, v_2, \ldots, v_k$.
Dowiedziemy teraz przeciwnej implikacji. Załóżmy, że jedyną kombinacją liniową wektorów $v_1, v_2, \ldots, v_k$ dającą wektor zerowy jest kombinacja trywialna. Pokażemy, że wówczas wektory $v_1, v_2, \ldots, v_k$ są liniowo niezależne. Gdyby wektory $v_1, v_2, \ldots, v_k$ były liniowo zależne, czyli np. $v_1 \in Lin\{v_2, v_3, \ldots, v_k\}$, a tym samym

$$v_k = t_1v_1 + t_2v_2 + \ldots + t_{k-1}v_{k-1},$$

wówczas mielibyśmy

$$-t_1v_1 - t_2v_2 - \ldots - t_{k-1}v_{k-1} + 1 \cdot v_k = \overrightarrow{O}$$

co przeczy założeniu o trywialności kombinacji dającej wektor zerowy. □

Przekonajmy się, że taki sposób sprawdzania liniowej niezależności rzeczywiście jest dużo szybszy.

Przykład. Rozważmy ten sam układ, co w poprzednim przykładzie, a zatem układ złożony z funkcji $\mathbb{1},\ x,\ x^2,\ x^3$. Weźmy kombinację tych funkcji dającą zero

$$t_1 \cdot \mathbb{1} + t_2 \cdot x + t_3 \cdot x^2 + t_4 \cdot x^3 = 0.$$

Zauważmy, że 0 możemy zapisać jako kombinację liniową wektorów $\mathbb{1},\ x,\ x^2,\ x^3$, otrzymujemy

$$0 \cdot \mathbb{1} + 0 \cdot x + 0 \cdot x^2 + 0 \cdot x^3.$$

Szukamy zatem takich t_1, t_2, t_3, t_4, aby

$$t_1 \cdot \mathbb{1} + t_2 \cdot x + t_3 \cdot x^2 + t_4 \cdot x^3 = 0 \cdot \mathbb{1} + 0 \cdot x + 0 \cdot x^2 + 0 \cdot x^3.$$

Wiemy, że wielomiany są równe, gdy mają równe współczynniki przy tych samych potęgach, więc

$$t_1 = t_2 = t_3 = t_4 = 0,$$

co zgodnie z powyższym lematem daje nam liniową niezależność wektorów funkcji $\mathbb{1},\ x,\ x^2,\ x^3$.

W Definicji 19.4.1 sprawdzaliśmy liniową niezależność wektorów przez sprawdzanie, czy kolejne wektory należą do generowanych przez poprzednie podprzestrzeni, można więc było sądzić, że przy badaniu liniowej niezależności istotna jest kolejność w jakiej ustawiamy wektory. Na mocy powyższego Lematu możemy stwierdzić, że kolejność ta nie ma wpływu na liniową zależność lub niezależność wektorów.

Wniosek 19.4.3. *Liniowa niezależność nie zależy od kolejności ustawienia wektorów w układzie.*

Lemat 19.4.4 (o jednoznaczności rozkładu). *Jeśli* $v_1, v_2, \ldots, v_k$ *jest układem liniowo niezależnym, to każdy wektor* v *z przestrzeni generowanej przez te wektory, czyli z przestrzeni* $Lin\{v_1, v_2, \ldots, v_k\}$*, przedstawia się jako kombinacja liniowa wektorów* $v_1, v_2, \ldots, v_k$ *w sposób jednoznaczny.*

Dowód. Skoro $v \in Lin\{v_1, v_2, \ldots, v_k\}$ to $v = t_1v_1 + t_2v_2 + \ldots + t_kv_k$, dla pewnych współczynników $t_1, t_2, \ldots t_k$. Pokażemy, że przedstawienie takie jest jedyne. Załóżmy, że istnieje drugie przedstawienie $v = s_1v_1 + s_2v_2 + \ldots + s_kv_k$. A zatem

$$t_1v_1 + t_2v_2 + \ldots + t_kv_k = s_1v_1 + s_2v_2 + \ldots + s_kv_k.$$

Po przekształceniu tej równości otrzymujemy

$$(t_1 - s_1)v_1 + (t_2 - s_2)v_2 + \ldots + (t_k - s_k)v_k = 0.$$

Wiemy, że wektory $v_1, v_2, \ldots, v_k$ są liniowo niezależne tak więc, zgodnie z Lematem 19.4.2

$$t_1 - s_1 = t_2 - s_2 = \cdots = t_k - s_k = 0,$$

a stąd mamy

$$t_1 = s_1, t_2 = s_2, \ldots, t_k = s_k.$$

Rozkłady te są więc takie same, co kończy dowód. □

Rozdział 20

Baza i wymiar przestrzeni wektorowej

W przestrzeniach R^2 i R^3 określiliśmy, czym jest układ współrzędnych, poznaliśmy również pojęcie wersorów w układzie. W tym rozdziale podamy definicję obiektu, który będzie dla przestrzeni wektorowej pełnił rolę podobną do roli zbioru wersorów układu współrzędnych. Obiektem tym będzie baza przestrzeni wektorowej. Umożliwi nam ona wprowadzenie współrzędnych dla wektorów z danej przestrzeni. W dalszej części poznamy pojęcie wymiaru przestrzeni.

20.1 Baza przestrzeni wektorowej.

Definicja 20.1.1 (baza przestrzeni wektorowej). Bazą (skończoną) przestrzeni wektorowej nazywamy każdy układ wektorów $v_1, v_2, \ldots, v_k$ z tej przestrzeni taki, że

(1) układ ten generuje przestrzeń V tzn. $V = Lin\{v_1, v_2, \ldots, v_k\}$,

(2) układ ten jest liniowo niezależny.

✔ **Przykład.**

(1) Dla przestrzeni $V = R^3$ bazą jest zbiór złożony z wektorów $v_1 = [1, 0, 0], v_2 = [0, 1, 0], v_3 = [0, 0, 1]$. Zauważmy bowiem, że

$$Lin\{v_1, v_2, v_3\} = \{t_1v_1 + t_2v_2 + t_3v_3 \ : \ t_1, t_2, t_3 \in R\} =$$

$$= \{t_1 \cdot [1, 0, 0] + t_2 \cdot [0, 1, 0] + t_3 \cdot [0, 0, 1]\} = \{[t_1, t_2, t_3] \ : \ t_1, t_2, t_3 \in R\} = R^3.$$

Jak nietrudno sprawdzić układ ten jest liniowo niezależny. Zatem v_1, v_2, v_3 tworzą bazę przestrzeni R^3.

(2) Niech $V = R_3[x]$, czyli V jest przestrzenią wielomianów stopnia mniejszego lub równego trzy o współczynnikach rzeczywistych. Weźmy teraz zbiór złożony z wielomianów

$$v_0 = \mathbb{1}, v_1 = x, v_2 = x^2, v_3 = x^3.$$

Układ ten generuje całą przestrzeń V gdyż

$$Lin\{v_0, v_1, v_2, v_3\} = Lin\{\mathbb{1}, x, x^2, x^3\} =$$

$$= \{a_0 + a_1x + a_2x^2 + a_3x^3 \ : \ a_0, a_1, a_2, a_3 \in R^3\} = R_3[x] = V.$$

Wiemy również, że układ ten jest liniowo niezależny - sprawdziliśmy to w poprzednim rozdziale. Tak więc układ złożony z wektorów $\mathbb{1}, x, x^2, x^3$ jest bazą przestrzeni V.

20.2 Współrzędne wektorów w ustalonej bazie.

Dzięki wprowadzonemu w poprzednim podrozdziale pojęciu bazy przestrzeni wektorowej, możemy każdy wektor tej przestrzeni jednoznacznie przedstawić przy użyciu jego współrzędnych.

Rozważmy przestrzeń V, której bazę stanowią wektory $v_1, v_2, \dots, v_k$. Weźmy dowolny wektor a z tej przestrzeni. Ponieważ $V = Lin\{v_1, v_2, \dots, v_k\}$, to wektor a możemy przedstawić w postaci kombinacji liniowej wektorów $v_1, v_2, \dots, v_k$ tzn. $a = a_1v_1 + a_2v_2 + \ldots + a_kv_k$, dla pewnych $a_1, a_2, \dots, a_k \in R$. Z definicji bazy wynika również, że wektory $v_1, v_2, \dots, v_k$ są liniowo niezależne a więc, na mocy Lematu 19.4.4 przedstawienie wektora a w postaci kombinacji liniowej wektorów $v_1, v_2, \dots, v_k$ jest jednoznaczne.

Definicja 20.2.1 (współrzędne wektora). Współczynniki $a_1, a_2, \dots, a_k$ kombinacji liniowej przedstawionej powyżej, nazywamy współrzędnymi wektora a w bazie $v_1, v_2, \dots, v_k$.

Uwaga 20.2.2. *W ustalonej bazie współrzędne każdego wektora są jednoznacznie określone.*

Oznaczenie. Współrzędne wektora a będziemy najczęściej zapisywać wierszowo tzn. $a = [a_1, a_2, \dots, a_k]$.

Określimy teraz jak wyrażają się działania na wektorach przy użyciu współrzędnych w ustalonej bazie $v_1, v_2, \dots, v_k$.

Własność 20.2.3. *Rozważmy wektory $\overrightarrow{a} = [a_1, a_2, \dots, a_k]$, $\overrightarrow{b} = [b_1, b_2, \dots, b_k]$. Suma tych wektorów, określona w tej samej bazie, ma współrzędne*

$$\overrightarrow{a} + \overrightarrow{b} = (a_1v_1 + a_2v_2 + \ldots + a_kv_k) + (b_1v_1 + b_2v_2 + \ldots + b_kv_k) =$$

$$= (a_1 + b_1)v_1 + (a_2 + b_2)v_2 + \ldots + (a_k + b_k)v_k = [a_1 + b_1, a_2 + b_2, \dots, a_k + b_k].$$

Własność 20.2.4. *Dla wektora $\overrightarrow{a} = [a_1, a_2, \dots, a_k]$ i liczby rzeczywistej t, współrzędne wektora będącego ich iloczynem wyrażają się, w tej samej bazie, jako*

$$t\cdot \overrightarrow{a} = t(a_1v_1+a_2v_2+\ldots+a_kv_k) = (ta_1)v_1+(ta_2)v_2+\ldots+(ta_k)v_k = [ta_1, ta_2, \dots, ta_k].$$

Możemy zatem stwierdzić, że działania na wektorach wyrażają się we współrzędnych w ustalonej bazie tak samo, jak dla wektorów z R^n.

20.3 Znajdowanie bazy przestrzeni wektorowej.

W tym rozdziale zajmiemy się problemem istnienia bazy dla danej przestrzeni wektorowej. Sprawdzimy również, czy baza taka jest wyznaczona jednoznacznie. Pokażemy sposób na wyznaczanie bazy dla danej przestrzeni wektorowej.

Definicja 20.3.1 (**przestrzeń skończenie wymiarowa**). Przestrzeń wektorowa V jest skończenie wymiarowa, jeśli jest generowana przez skończony układ wektorów, czyli $V = Lin\{v_1, v_2, \ldots, v_k\}$, dla pewnego skończonego układu wektorów $v_1, v_2, \ldots, v_k$.

Twierdzenie 20.3.2. *Każda skończenie wymiarowa przestrzeń wektorowa ma bazę.*

Rozumowanie przedstawione w dowodzie tego twierdzenia będzie bardzo często pojawiać się przy dalszym badaniu przestrzeni wektorowych. Zawiera ono procedurę, dzięki której będziemy w stanie wyznaczyć bazę danej przestrzeni.

Dowód. Niech $v_1, v_2, \ldots, v_k$ będzie układem wektorów generujących przestrzeń V. Układ ten może być liniowo niezależny i wtedy stanowi bazę tej przestrzeni. Na ogół jednak będzie liniowo zależny. W takiej sytuacji bazę otrzymamy przez usunięcie z tego układu zbędnych wektorów.

PROCEDURA USUWANIA ZBĘDNYCH WEKTORÓW:

Krok (1). Jeśli pierwszy wektor v_1 jest zerowy, to usuwamy go. Pozostałe wektory dalej generują V, bo pierwszy wektor nic nie wnosił do kombinacji liniowej. Powtarzamy Krok (1) tak długo, aż pierwszy wektor będzie niezerowy. Po każdym usunięciu wektora, w którymkolwiek kroku procedury, od nowa numerujemy pozostałe wektory jako $(v_1, v_2, \ldots)$.

Krok (2). Jeśli pierwszy wektor v_1 jest już niezerowy, to patrzymy na drugi wektor v_2. Jeśli $v_2 = t \cdot v_1$, to usuwamy v_2, pamiętając o przenumerowaniu wektorów. Zauważmy, że pozostałe wektory nadal generują przestrzeń V, bo jeśli $\overrightarrow{a} = a_1v_1 + a_2v_2 + a_3v_3 + \ldots + a_kv_k$, to

$$\overrightarrow{a} = a_1v_1 + a_2 \cdot tv_1 + a_3v_3 + \ldots + a_kv_k = (a_1 + ta_2)v_1 + a_3v_3 + \ldots + a_kv_k.$$

Powtarzamy Krok (2) tak długo aż $v_2 \notin Lin\{v_1\}$.

Krok ogólny. Jeśli po wykonaniu poprzednich kroków mamy układ, w którym

$$\begin{array}{l} v_1 \neq \overrightarrow{O} \\ v_2 \notin Lin\{v_1\} \\ \vdots \\ v_m \notin Lin\{v_1, v_2, \ldots, v_{m-1}\} \end{array}$$

to patrzymy na wektor v_{m+1}. Jeśli $v_{m+1} \notin Lin\{v_1, v_2, \ldots, v_m\}$, to zostawiamy go. Jeśli $v_{m+1} \in Lin\{v_1, v_2, \ldots, v_m\}$, to usuwamy go. W tym drugim przypadku możemy zauważyć, że pozostałe wektory dalej generują całą przestrzeń, bo każda kombinacja $a_1v_1 + \ldots + a_mv_m + a_{m+1}v_{m+1} \ldots + a_kv_k$ może też być zapisana jako

$$a_1v_1 + \ldots + a_mv_m + a_{m+1}(t_1v_1 + \ldots + t_mv_m) + a_{m+2}v_{m+2} + \ldots + a_kv_k =$$

$$(a_1 + a_{m+1}t_1)v_1 + \ldots + (a_m + a_{m+1}t_m)v_m + a_{m+2}v_{m+2} + \ldots + a_kv_k.$$

Jak widać jest to kombinacja wektorów $v_1, \ldots, v_m, v_{m+2}, \ldots, v_k$. Ogólny krok wykonujemy tak długo, aż uzyskamy układ $v_1, v_2, \ldots, v_k$ w dalszym ciągu generujący V i taki, że

$$\begin{array}{l} v_1 \neq \overrightarrow{O} \\ v_2 \notin Lin\{v_1\} \\ \vdots \\ v_k \notin Lin\{v_1, \ldots, v_{k-1}\} \end{array}$$

czyli liniowo niezależny. Zauważmy, że indeks k występujący przy ostatnim z tych wektorów nie jest tym samym indeksem k, który mieliśmy na początku rozważań. Wektorów tych jest mniej niż na początku, jednak dla zachowania czytelności zapisu indeksy pozostały niezmienione.

Tak uzyskany układ wektorów generuje przestrzeń V i jest liniowo niezależny, więc jest bazą przestrzeni V. $\square$

✔ **Przykład.** Sprawdźmy działanie tej procedury na przykładzie. Niech naszą przestrzenią V będzie przestrzeń R^3 a układem generującym tę przestrzeń jest

$$\left\{ v_1 = \begin{pmatrix} -1 \\ 2 \\ 1 \end{pmatrix}, v_2 = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}, v_3 = \begin{pmatrix} -2 \\ 4 \\ 2 \end{pmatrix}, v_4 = \begin{pmatrix} 3 \\ 1 \\ -2 \end{pmatrix}, v_5 = \begin{pmatrix} 1 \\ 5 \\ 0 \end{pmatrix}, v_6 = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} \right\}.$$

Sprawdzenie, że układ ten rzeczywiście generuje przestrzeń R^3 pominiemy. W pierwszym kroku sprawdzamy czy $v_1 \neq \overrightarrow{O}$. W tym wypadku tak jest, więc przechodzimy do Kroku (2). Sprawdzamy czy $v_2 = tv_1$. Zauważmy, że $v_2 = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} = 0 \cdot \begin{pmatrix} -1 \\ 2 \\ 1 \end{pmatrix} = 0 \cdot v_1$. Możemy zatem usunąć wektor v_2. Nasz układ wektorów to teraz

$$\left\{ v_1 = \begin{pmatrix} -1 \\ 2 \\ 1 \end{pmatrix}, v_2 = \begin{pmatrix} -2 \\ 4 \\ 2 \end{pmatrix}, v_3 = \begin{pmatrix} 3 \\ 1 \\ -2 \end{pmatrix}, v_4 = \begin{pmatrix} 1 \\ 5 \\ 0 \end{pmatrix}, v_5 = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} \right\}.$$

Powtarzamy procedurę z Kroku (2). Zauważmy, że

$$v_2 = \begin{pmatrix} -2 \\ 4 \\ 2 \end{pmatrix} = 2 \cdot \begin{pmatrix} -1 \\ 2 \\ 1 \end{pmatrix} = 2 \cdot v_1$$

usuwamy zatem v_2. Otrzymujemy zestaw wektorów

$$\left\{ v_1 = \begin{pmatrix} -1 \\ 2 \\ 1 \end{pmatrix}, v_2 = \begin{pmatrix} 3 \\ 1 \\ -2 \end{pmatrix}, v_3 = \begin{pmatrix} 1 \\ 5 \\ 0 \end{pmatrix}, v_4 = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} \right\}.$$

Jeszcze raz sprawdzamy Krok (2). Tym razem $v_2 \neq tv_1$. Tak więc zostawiamy v_2. Sprawdzamy czy $v_3 \in Lin\{v_1, v_2\}$. Zauważmy, że

$$v_3 = \begin{pmatrix} 1 \\ 5 \\ 0 \end{pmatrix} = 2 \cdot \begin{pmatrix} -1 \\ 2 \\ 1 \end{pmatrix} + \begin{pmatrix} 3 \\ 1 \\ -2 \end{pmatrix} = 2v_1 + v_2,$$

tak więc $v_3 \in Lin\{v_1, v_2\}$, zatem go usuwamy. Nasz układ to teraz

$$\left\{ v_1 = \begin{pmatrix} -1 \\ 2 \\ 1 \end{pmatrix}, v_2 = \begin{pmatrix} 3 \\ 1 \\ -2 \end{pmatrix}, v_3 = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} \right\}.$$

Sprawdźmy, czy $v_3 \in Lin\{v_1, v_2\}$. Jak nietrudno się przekonać $v_3 \notin Lin\{v_1, v_2\}$. Możemy to zrobić np. poprzez policzenie wyznacznika

$$\begin{vmatrix} -1 & 3 & 1 \\ 2 & 1 & 0 \\ 1 & -2 & 1 \end{vmatrix} = -12.$$

Wyznacznik tej macierzy jest niezerowy, więc wektory v_1, v_2, v_3 są liniowo niezależne. Mamy więc

$$\begin{array}{l} v_1 \neq \overrightarrow{O} \\ v_2 \notin Lin\{v_1\} \\ v_3 \notin Lin\{v_1, v_2\} \end{array}.$$

Układ złożony z wektorów v_1, v_2, v_3 jest zatem bazą przestrzeni R^3.

Procedura opisana w dowodzie Twierdzenia 20.3.2 pozwoli nam uzasadnić rozmaite dalsze własności. Jedną z nich jest następujący wniosek.

Wniosek 20.3.3. *Każdy liniowo niezależny układ $e_1, \ldots, e_p$ wektorów skończenie wymiarowej przestrzeni V można uzupełnić wektorami $e_{p+1}, \ldots, e_k$ do bazy przestrzeni V.*

Dowód. Przestrzeń V jest skończenie wymiarowa, więc zgodnie z Twierdzeniem 20.3.2 posiada bazę. Niech bazą tą będzie układ $v_1, v_2, \ldots, v_k$. Mamy wówczas $V = Lin\{v_1, v_2, \ldots, v_k\}$ zatem tym bardziej $V = Lin\{e_1, e_2, \ldots, e_p, v_1, \ldots, v_k\}$. Możemy teraz zastosować procedurę usuwania wektorów do układu $e_1, e_2, \ldots, e_p, v_1, \ldots, v_k$. Zauważmy, że skoro wektory $e_1, \ldots, e_p$ są liniowo niezależne, co oznacza, że $e_1 \neq \overrightarrow{O}$, $e_2 \notin Lin\{e_1\}$, $\ldots, e_p \notin Lin\{e_1, \ldots, e_{p-1}\}$, to żaden z tych wektorów nie zostanie usunięty z układu. Usunięte będą dopiero pewne spośród wektorów $v_1, v_2, \ldots, v_k$. Te które pozostaną, razem z wektorami $e_1, \ldots, e_p$, utworzą bazę. □

Powyższy wniosek oznacza między innymi, że w danej przestrzeni wektorowej jest dużo baz. Zauważmy, że na pierwszy wektor bazy możemy wybrać dowolny niezerowy wektor. Na wybór drugiego wektora również mamy wiele możliwości, wystarczy aby nie był on współliniowy z pierwszym. Podobnie możemy wybierać kolejne wektory. Na podstawie powyższego wniosku stwierdzamy, że jest to wykonalne. Tak więc pokazaliśmy, że rzeczywiście bazę danej przestrzeni możemy wybrać na wiele różnych sposobów.

20.4 Wymiar przestrzeni wektorowej.

W poprzednim podrozdziale pokazaliśmy, że przestrzeń wektorowa może mieć wiele baz, w tym pokażemy, że wszystkie te bazy składają się z takiej samej liczby wektorów. Tę liczbę nazywamy wymiarem przestrzeni i właśnie nim zajmiemy się w tym podrozdziale.

Twierdzenie 20.4.1. *Jeśli $e_1, \ldots, e_n$ jest bazą przestrzeni V zaś $v_1, v_2, \ldots, v_k$ jest dowolnym liniowo niezależnym układem, to $k \leqslant n$.*

Dowód tego twierdzenia pomijamy.

Na mocy tego twierdzenia możemy wyprowadzić bardzo użyteczny wniosek.

Wniosek 20.4.2. *Dowolne dwie bazy skończenie wymiarowej przestrzeni V są równoliczne.*

Dowód. Niech $e_1, e_2, \ldots, e_n$ i $f_1, f_2, \ldots, f_m$ będą bazami przestrzeni V. Ponieważ układ $f_1, f_2, \ldots, f_m$ jest liniowo niezależny, więc z Twierdzenia 20.4.1 mamy $m \leqslant n$. Podobnie wiemy, że układ $e_1, e_2, \ldots, e_n$ jest liniowo niezależny, stąd $n \leqslant m$. Ostatecznie otrzymujemy więc $n = m$. □

Definicja 20.4.3 (wymiar przestrzeni liniowej). Wymiarem (skończenie wymiarowej) przestrzeni V nazywamy liczbę n będącą liczbą wektorów w dowolnej bazie tej przestrzeni.

Oznaczenie. Wymiar przestrzeni V oznaczać będziemy symbolem $dimV$.

✔ **Przykład.**

(1) $dimR^3 = 3$ (patrz Przykład (1) w podrozdziale 20.1)

(2) $dimR^n = n$ (bo, jak wiemy, układ n wektorów jest bazą)

(3) $dimR_3[X] = 4$ (patrz Przykład (2) w podrozdziale 20.1)

Wniosek 20.4.4. *Jeśli W jest podprzestrzenią skończenie wymiarowej przestrzeni V, to W też jest skończenie wymiarowa, i ponadto*

$$dimW \leqslant dimV.$$

Dowód. Niech wymiar przestrzeni V będzie równy n ($dimV = n$). Z podprzestrzeni W wybieramy kolejno wektory $w_1, w_2, \ldots$ w taki sposób, że

$$\begin{array}{l} w_1 \neq \overrightarrow{O} \\ w_2 \notin Lin\{w_1\} \\ \vdots \\ w_k \notin Lin\{w_1, w_2, \ldots, w_{k-1}\} \end{array}$$

Zauważmy, że to dobieranie musi się skończyć po k krokach, dla pewnego $k \leqslant n$, bo $w_1, w_2, \ldots, w_k$ są liniowo niezależne zatem nie może ich być więcej niż wynosi wymiar przestrzeni V. Tak wybrany układ $w_1, w_2, \ldots, w_k$ jest bazą podprzestrzeni W, gdyż jest liniowo niezależny i generuje W. Gdyby ten układ nie generował W, to zgodnie z Wnioskiem 20.3.3 dałoby się dobrać inne wektory liniowo niezależne z pozostałymi. Otrzymujemy zatem

$$dim W = k \leqslant n = dim V.$$

□

Wniosek 20.4.5. *Jeśli V jest skończenie wymiarową przestrzenią wektorową, zaś W jej podprzestrzenią, to każdą bazę podprzestrzeni W można uzupełnić do bazy całej przestrzeni V.*

Dowód. Niech $w_1, w_2, \ldots, w_k$ będzie bazą podprzestrzeni W. Jako baza W układ ten jest liniowo niezależny więc, zgodnie z Wnioskiem 20.3.3 można go uzupełnić do bazy przestrzeni V. □

Rozdział 21

Teoria minorów

Rozdział ten poświęcimy na wprowadzenie bardzo użytecznej teorii, zwanej teorią minorów. Na początek, zanim zdefiniujemy pojęcie minora, nauczymy się sprawdzać przy jego użyciu liniową niezależność, oraz badać wymiar podprzestrzeni generowanej przez układ wektorów. W kolejnym podrozdziale zdefiniujemy, czym są minory, wprowadzimy też bardzo ważne pojęcie, które często będziemy wykorzystywać w dalszej części tego rozdziału - pojęcie rzędu macierzy. W ostatnim podrozdziale pokażemy najważniejsze zastosowanie teorii minorów, a mianowicie wykorzystywanie ich do rozwiązywania układów równań.

21.1 Wyznacznik a liniowa niezależność.

Jak już mieliśmy okazję się przekonać, rozpoznawanie liniowej niezależności zbioru wektorów wprost z definicji, bywa bardzo pracochłonne. W przypadku przestrzeni R^3 do badania liniowej niezależności używaliśmy wyznacznika macierzy. Metoda ta może być równie dobrze wykorzystywana w przypadku dowolnej przestrzeni R^n.

Niech $v_1, v_2, \ldots, v_k$ będą wektorami z przestrzeni R^n, przy czym k niekoniecznie jest równe n. Możemy z tych wektorów utworzyć macierz $(v_1, v_2, \ldots, v_k)$ rozmiaru $n \times k$, poprzez wpisanie współrzędnych tych wektorów w kolejne kolumny macierzy. Rozważmy trzy przypadki.

(1) Gdy $k = n$ wektory $v_1, v_2, \ldots, v_k$ są liniowo niezależne wtedy i tylko wtedy gdy wyznacznik $\det(v_1, v_2, \ldots, v_k)$ jest niezerowy, co wiemy z Lematu 19.4.2.

(2) Gdy $k > n$ układ jest liniowo zależny, gdyż wymiar przestrzeni $dimR^n = n$, a jak wiemy układy liniowo niezależne nie mogą składać się z większej liczby wektorów, niż wynosi wymiar przestrzeni.

(3) Gdy $k < n$, tzn. wektorów jest mniej niż wynosi wymiar przestrzeni, to macierz utworzona z tych wektorów $(v_1, v_2, \ldots, v_k)$, ma rozmiar $n \times k$, więc nie jest macierzą kwadratową, w związku z czym nie posiada wyznacznika. Metoda rozpoznawania liniowej niezależności takiego układu wektorów podana została w poniższym twierdzeniu.

Twierdzenie 21.1.1. *Układ złożony z wektorów $v_1, v_2, \ldots, v_k$ jest liniowo niezależny wtedy i tylko wtedy gdy przynajmniej jedna z macierzy rozmiaru $k \times k$ utworzona z macierzy $(v_1, v_2, \ldots, v_k)$ przez skreślenie $n-k$ wierszy, ma niezerowy wyznacznik.*

Zanim udowodnimy to twierdzenie zobaczmy jego działanie na przykładzie.

✔ **Przykład.** Rozważmy wektory $v_1 = \begin{pmatrix} 3 \\ -2 \\ 4 \\ 0 \end{pmatrix}, v_2 = \begin{pmatrix} 1 \\ -1 \\ 7 \\ 5 \end{pmatrix}, v_3 = \begin{pmatrix} 8 \\ -1 \\ -2 \\ -3 \end{pmatrix}$ z przestrzeni R^4. Macierz utworzona z tych wektorów to

$$(v_1, v_2, v_3) = \begin{pmatrix} 3 & 1 & 8 \\ -2 & -1 & -1 \\ 4 & 7 & -2 \\ 0 & 5 & -3 \end{pmatrix}.$$

Policzmy wyznacznik macierzy powstałej z powyższej, przez wykreślenie pierwszego wiersza.

$$\det \begin{pmatrix} -2 & -1 & -1 \\ 4 & 7 & -2 \\ 0 & 5 & -3 \end{pmatrix} = -10 \neq 0.$$

Wyznacznik ten jest niezerowy, więc zgodnie z powyższym twierdzeniem, wektory v_1, v_2, v_3 są liniowo niezależne.

Dowód Twierdzenie 21.1.1. Pokażemy, że jeśli któraś z macierzy m rozmiaru $k \times k$, powstała z macierzy $(v_1, v_2, \ldots, v_k)$ przez wykreślenie odpowiedniej liczby wierszy, ma niezerowy wyznacznik, to wektory $v_1, v_2, \ldots, v_k$ są liniowo niezależne. Oznaczmy przez $v_1', v_2', \ldots, v_k'$ wektory z przestrzeni R^k będące kolumnami macierzy m jak wyżej o niezerowym wyznaczniku. Każdy z wektorów v_i' powstaje z v_i przez usunięcie $n-k$ odpowiednich współrzędnych. Mamy uzasadnić, że układ $v_1, v_2, \ldots, v_k$ jest liniowo niezależny, czyli $t_1 v_1 + t_2 v_2 + \ldots + t_k v_k = \overrightarrow{O}$ tylko wtedy gdy $t_1 = t_2 = \ldots = t_k = 0$.

Skoro wyznacznik $\det(v_1', v_2', \ldots, v_k') \neq 0$, więc $v_1', v_2', \ldots, v_k'$ są liniowo niezależne, co oznacza, że jedyna kombinacja liniowa tych wektorów dająca wektor zerowy

$$t_1 v_1' + t_2 v_2' + \ldots + t_k v_k' = \overrightarrow{O},$$

to kombinacja trywialna, czyli taka w której

$$t_1 = t_2 = \ldots = t_k = 0.$$

Jak nietrudno się przekonać wówczas również

$$t_1 v_1 + t_2 v_2 + \ldots + t_k v_k = \overrightarrow{O}.$$

Zauważmy, że są to jedyne wartości współczynników, dla których $t_1v_1 + t_2v_2 + \ldots + t_kv_k = \overrightarrow{O}$, bo jeśli ta kombinacja jest zerowa to również $t_1v_1' + t_2v_2' + \ldots + t_kv_k' = \overrightarrow{O}$, a jak wiemy ten drugi warunek zachodzi tylko wtedy gdy $t_1 = t_2 = \ldots = t_k = 0$. Wektory $v_1, v_2, \ldots, v_k$ są zatem liniowo niezależne. Dowód implikacji przeciwnej jest bardziej złożony i w tym skrypcie go pomijamy. □

21.2 Wymiar podprzestrzeni generowanej przez układ wektorów.

Mając dany układ wektorów z R^n możemy zastanawiać się, jaką podprzestrzeń generuje ten układ. W podrozdziale tym określimy jedną z cech, która charakteryzuje taką podprzestrzeń - jej wymiar. Rozważmy układ, niekoniecznie liniowo niezależnych wektorów $v_1, v_2, \ldots, v_m$ z przestrzeni R^n.

Fakt 21.2.1. *Wymiar przestrzeni generowanej przez wektory $v_1, v_2, \ldots, v_m \in R^n$, czyli $dim(Lin\{v_1, v_2, \ldots, v_m\})$, jest równy maksymalnej liczbie liniowo niezależnych wektorów spośród $v_1, v_2, \ldots, v_m$.*

Dowód. Rozważmy bazę podprzestrzeni $Lin\{v_1, v_2, \ldots, v_k\}$ uzyskaną przez usunięcie części wektorów z $v_1, v_2, \ldots, v_m$ generujących tę podprzestrzeń. Baza ta jest liniowo niezależna, a jak wiemy każdy inny zbiór wektorów liniowo niezależnych nie może być liczniejszy niż baza. Tak więc baza ta realizuje maksimum liczby liniowo niezależnych wektorów spośród $v_1, v_2, \ldots, v_m$, a liczba wektorów w tej bazie to $dim(Lin\{v_1, v_2, \ldots, v_m\})$. □

Fakt 21.2.2. *Wymiar podprzestrzeni $Lin\{v_1, v_2, \ldots, v_m\}$ to maksymalny rozmiar k macierzy $k \times k$ o niezerowym wyznaczniku, utworzonej z macierzy $(v_1, v_2, \ldots, v_m)$ przez skreślenie odpowiedniej liczby kolumn oraz wierszy.*

Dowód. Jeżeli wymiar podprzestrzeni $Lin\{v_1, v_2, \ldots, v_m\}$ wynosi k, to jak wynika z Faktu 21.2.1 znajdziemy k liniowo niezależnych wektorów spośród $v_1, v_2, \ldots, v_m$. Wówczas jak wiemy z Twierdzenia 21.1.1, istnieje macierz rozmiaru $k \times k$ utworzona z macierzy $(v_1, v_2, \ldots, v_m)$ przez skreślenie odpowiedniej liczby kolumn oraz wierszy, której wyznacznik jest niezerowy. Gdyby istniała macierz rozmiaru $(k+1) \times (k+1)$ utworzona z macierzy $(v_1, v_2, \ldots, v_m)$, której wyznacznik byłby niezerowy, wówczas z Twierdzenia 21.1.1 mielibyśmy, że układ złożony z wektorów tworzących kolumny tej macierzy jest liniowo niezależny. Wtedy na mocy Faktu 21.2.1 mielibyśmy, że wymiar podprzestrzni generowanej przez $v_1, v_2, \ldots, v_m$ wynosi co najmniej $(k+1)$. Co daje sprzeczność. Tak więc wymiar podprzestrzeni generowanej przez $v_1, v_2, \ldots, v_m$ jest równy maksymalnemu rozmiarowi macierzy o niezerowym wyznaczniku powstałej z macierzy $(v_1, v_2, \ldots, v_m)$. □

21.3 Rząd macierzy - minory.

W tym podrozdziale zdefiniujemy, przy użyciu pojęcia minora, czym jest rząd danej macierzy. Pojęcie to okaże się bardzo użyteczne w kolejnym podrozdziale przy określaniu, czy dany układ równań liniowych ma rozwiązanie. Podajmy najpierw definicję minora macierzy.

Definicja 21.3.1 (minor macierzy). Niech A będzie macierzą dowolnego rozmiaru. Minorem macierzy A nazywamy wyznacznik dowolnej kwadratowej macierzy B, utworzonej z A przez wykreślenie pewnej ilości wierszy i kolumn. Macierz B nazywamy macierzą minora. Jeśli macierz minora ma rozmiar $k \times k$, to liczbę k nazywamy stopniem minora.

✔ **Przykład.** Rozważmy macierz

$$A = \begin{bmatrix} 1 & 7 & 1 & 0 & 1 \\ 3 & 6 & -5 & 2 & -1 \\ -2 & -4 & 0 & -3 & 1 \end{bmatrix}.$$

Jeżeli z tej macierzy wykreślimy pierwszą, trzecią i piątą kolumnę, oraz drugi wiersz, to otrzymamy macierz $A' = \begin{bmatrix} 7 & 0 \\ -4 & -3 \end{bmatrix}$. Wyznacznik tej macierzy to

$$\det A' = \det \begin{bmatrix} 7 & 0 \\ -4 & -3 \end{bmatrix} = -21,$$

tak więc jeden z minorów stopnia dwa macierzy A jest równy -21.

Uwaga 21.3.2. *Jeżeli pewien minor macierzy A jest niezerowy, to kolumny (a także wiersze) tej macierzy, występujące w macierzy tego minora są liniowo niezależne. I na odwrót - jeśli pewien układ kolumn (wierszy) macierzy A jest liniowo niezależny, to z tych kolumn (wierszy) można utworzyć macierz niezerowego minora.*

✔ **Przykład.** Rozważmy macierz A z poprzedniego przykładu. Znaleźliśmy niezerowy minor tej macierzy, zawierał on pierwszy i trzeci wiersz macierzy A, tak więc wiersze pierwszy i trzeci są liniowo niezależne.

Możemy przy użyciu pojęcia minora macierzy sformułować Fakt 21.2.2. Czytelnikowi pozostawimy przetłumaczenie sformułowania jednego faktu na drugi.

Fakt 21.3.3. *Dla macierzy A rozmiaru $n \times m$ wymiar podprzestrzeni w R^n generowanej przez kolumny macierzy A jest równy maksymalnemu stopniowi niezerowego minora macierzy A.*

Definicja 21.3.4 (rząd macierzy). Rzędem macierzy $C = (C_1, \ldots, C_m)$, gdzie $C_i \in R^n$, rozmiaru $n \times m$ nazywamy wymiar $dim(Lin\{C_1, \ldots, C_m\})$ podprzestrzeni w R^n generowanej przez kolumny tej macierzy. Rząd macierzy C oznaczamy przez $r(C)$.

Uwaga 21.3.5.

(1) Rząd $r(C)$ macierzy C jest równy maksymalnej liczbie liniowo niezależnych kolumn w macierzy C (wynika to z Faktu 21.2.1).

(2) Rząd $r(C)$ macierzy C jest równy maksymalnemu stopniowi niezerowego minora macierzy C (wynika to z Faktu 21.3.3).

(3) *Jeżeli macierz C ma rozmiar $m \times n$, to $r(C) \leqslant \min\{m,n\}$ (jest to konsekwencją powyższej własności (2), bo stopień k dowolnego minora spełnia $k \leqslant min(m,n)$).*

(4) *Rząd macierzy C jest równy rzędowi macierzy do niej transponowanej,*

$$r(C) = r(C^T),$$

bo jeśli M jest macierzą niezerowego minora w C, to M^T jest macierzą analogicznego minora w C^T (bo $\det M^T = \det M$).

(5) *Każda macierz ma tyle samo liniowo niezależnych kolumn co wierszy i ilość ta odpowiada rzędowi tej macierzy.*

Poniżej podamy uwagi, które okażą się być bardzo pomocne, przy obliczaniu rzędu macierzy. Przykład obliczania rzędu zobaczymy w kolejnym podrozdziale, przy okazji omawiania sposobu rozwiązywania układu równań.

Uwaga 21.3.6.

(1) *Rząd macierzy nie zmienia się pod wpływem operacji na wierszach lub kolumnach tej macierzy, takich jak mnożenie wiersza lub kolumny przez niezerową liczbę, dodawania do wiersza (kolumny) wielokrotności innego wiersza (kolumny) czy zamiany kolejności wierszy (kolumn).*

(2) *Każda macierz niezerowego minora zawiera się w macierzy pewnego niezerowego minora maksymalnego stopnia. Niezerowych minorów maksymalnego stopnia należy więc szukać wśród macierzy okalających macierze poprzednio znalezionych niezerowych minorów.*

21.4 Rozwiązywanie układu równań metodą minorów.

Do tej pory poznaliśmy dwie metody rozwiązywania układów równań. Pierwsza z nich, metoda eliminacji Gaussa, wprowadzona została w Rozdziale 16, druga, metoda Cramera, w podrozdziale 18.7. Tę drugą metodę mogliśmy jednak stosować tylko do układów, w których liczba równań była taka sama jak liczba równań, oraz których wyznacznik główny był niezerowy. W tym podrozdziale poznamy inną metodę rozwiązywania układów równań, która częściowo będzie bazować na metodzie Cramera, jednak będzie można ją stosować do dowolnych układów równań liniowych. Na początek nauczymy się przy użyciu pojęcia rzędu macierzy sprawdzać, czy dany układ ma rozwiązanie. W dalszej części podrozdziału na przykładzie poznamy algorytm rozwiązywania układu równań.

Rozważmy układ n równań z m niewiadomymi

$$(\bigstar) \begin{cases} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1m}x_m = b_1 \\ a_{21}x_1 + a_{22}x_2 + \ldots + a_{2m}x_m = b_2 \\ \vdots \\ a_{n1}x_1 + a_{n2}x_2 + \ldots + a_{nm}x_m = b_n \end{cases}$$

w którym niewiadomymi są $x_1, x_2, \ldots, x_m$, natomiast a_{ij}, b_j to współczynniki tego układu. Układ ten możemy zapisać wektorowo

$$x_1 \cdot A_1 + x_2 \cdot A_2 + \ldots + x_m \cdot A_m = B,$$

gdzie $A_i = \begin{bmatrix} a_{1i} \\ a_{2i} \\ \vdots \\ a_{ni} \end{bmatrix}$, $B = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix}$.

Uwaga 21.4.1. *Układ (★) ma rozwiązanie wtedy i tylko wtedy gdy wektor B jest kombinacją liniową wektorów $A_1, A_2, \ldots, A_m$, czyli $B \in Lin\{A_1, A_2, \ldots, A_m\}$, a to zachodzi wtedy i tylko wtedy gdy*

$$dim(Lin\{A_1, A_2, \ldots, A_m\}) = dim(Lin\{A_1, A_2, \ldots, A_m, B\}).$$

Dowód. Pierwsza równoważność wynika wprost z wektorowego zapisu układu (★). Udowodnijmy zatem tylko drugą równoważność.
Zauważmy, że jeśli $B \in Lin\{A_1, A_2, \ldots, A_m\}$, to

$$Lin\{A_1, A_2, \ldots, A_m\} = Lin\{A_1, A_2, \ldots, A_m, B\},$$

zatem

$$dim(Lin\{A_1, A_2, \ldots, A_m\}) = dim(Lin\{A_1, A_2, \ldots, A_m, B\}),$$

tym samym mamy udowodnioną jedną z implikacji.
Pokażmy teraz, że jeśli $dim(Lin\{A_1, A_2, \ldots, A_m\}) = dim(Lin\{A_1, A_2, \ldots, A_m, B\})$, to $B \in Lin\{A_1, A_2, \ldots, A_m\}$.
Niech $A_{i_1}, A_{i_2}, \ldots, A_{i_p}$ będzie bazą $Lin\{A_1, A_2, \ldots, A_m\}$. Wówczas wymiar tej przestrzeni $dim(Lin\{A_1, A_2, \ldots, A_m\}) = p$, a zatem również

$$dim(Lin\{A_1, A_2, \ldots, A_m, B\}) = p.$$

Układ złożony z wektorów $A_{i_1}, A_{i_2}, \ldots, A_{i_p}$ jest liniowo niezależny również w przestrzeni $Lin\{A_1, A_2, \ldots, A_m, B\}$. Ponieważ ma tyle wektorów ile wynosi wymiar tej przestrzeni, więc jest w niej bazą. Oznacza to, że $B \in Lin\{A_{i_1}, A_{i_2}, \ldots, A_{i_p}\}$, więc tym bardziej $B \in Lin\{A_1, A_2, \ldots, A_m\}$. Co kończy dowód. □

Posługując się pojęciem rzędu macierzy możemy przeformułować powyższą uwagę. Oznaczmy $A = [A_1, A_2, \ldots, A_m]$, $U = [A_1, A_2, \ldots, A_m, B]$. Macierz A nazywa się **macierzą główną**, zaś macierz U **macierzą rozszerzoną** układu (★).

Twierdzenie 21.4.2 (Kroneckera - Capelli'ego). *Układ (★) ma rozwiązanie wtedy i tylko wtedy gdy*

$$r(A) = r(U),$$

czyli gdy rząd macierzy rozszerzonej tego układu jest równy rzędowi macierzy głównej.

✔ **Przykład.** Przy użyciu powyższego twierdzenia sprawdźmy, czy układ

$$\begin{cases} 2x - 3y = 12 \\ x + 4y = -5 \\ x - y = 3 \end{cases}$$

ma rozwiązanie. Układ ten możemy wektorowo zapisać w postaci

$$x \cdot \begin{bmatrix} 2 \\ 1 \\ 1 \end{bmatrix} + y \cdot \begin{bmatrix} -3 \\ 4 \\ -1 \end{bmatrix} = \begin{bmatrix} 12 \\ -5 \\ 3 \end{bmatrix}.$$

Zgodnie z wprowadzonymi wcześniej oznaczeniami mamy

$$A = \begin{bmatrix} 2 & -3 \\ 1 & 4 \\ 1 & -1 \end{bmatrix}, \quad U = \begin{bmatrix} 2 & -3 & 12 \\ 1 & 4 & -5 \\ 1 & -1 & 3 \end{bmatrix}.$$

Policzmy rzędy macierzy A i U. Zauważmy, że kolumny macierzy A są niewspółliniowe, więc są liniowo niezależne, stąd

$$r(A) = r \begin{bmatrix} 2 & -3 \\ 1 & 4 \\ 1 & -1 \end{bmatrix} = 2.$$

Macierz U jest macierzą kwadratową, możemy zatem obliczyć jej wyznacznik

$$\det U = \det \begin{bmatrix} 2 & -3 & 12 \\ 1 & 4 & -5 \\ 1 & -1 & 3 \end{bmatrix} = -22,$$

jak widać wyznacznik ten jest niezerowy, więc kolumny macierzy U są liniowo niezależne, a tym samym

$$r(U) = r \begin{bmatrix} 2 & -3 & 12 \\ 1 & 4 & -5 \\ 1 & -1 & 3 \end{bmatrix} = 3.$$

Otrzymujemy więc

$$r(A) \neq r(U),$$

a więc dany układ równań, zgodnie z Twierdzeniem Kroneckera - Capelli'ego nie ma rozwiązania (jest sprzeczny).

Zanim przejdziemy do rozwiązywania danego układu musimy najpierw sprawdzić, czy układ ten ma rozwiązanie, tym samym musimy policzyć rzędy odpowiednich macierzy. Podamy teraz kilka praktycznych uwag, które mogą bardzo ułatwić te obliczenia w przypadku macierzy większych rozmiarów.

Uwaga 21.4.3. *Rząd macierzy nie zmienia się pod wpływem operacji na wierszach i kolumnach macierzy, możemy więc mnożyć wiersz lub kolumnę przez niezerową liczbę, możemy dodawać do wiersza (kolumny) wielokrotność innego wiersza (kolumny), dopuszczalne są również operacje zamiany kolejności wierszy lub kolumn.*

Uwaga 21.4.4. *Jeżeli macierz A jest postaci*

$$A = \begin{pmatrix} x_1 & \star & \dots & \star & \dots & \star \\ & x_2 & \dots & \star & \dots & \star \\ & & \ddots & \vdots & & \vdots \\ & 0 & & x_m & \dots & \star \\ & & & 0 & \dots & 0 \end{pmatrix}$$

gdzie pod zaznaczoną linią wszystkie wyrazy są zerowe, wyrazy x_i są niezerowe, natomiast inne wyrazy znajdujące się nad linią są dowolne, wówczas rząd tej macierzy wynosi

$$r(A) = m.$$

Powyższa uwaga wynika z tego, że w macierzy A istnieje niezerowy minor stopnia m. Minorem tym jest lewa górna podmacierz A' rozmiaru $m \times m$, która jest górnotrójkątna, i której wyznacznik $\det A' = x_1 \cdot x_2 \cdot \ldots \cdot x_m$ jest niezerowy, bo $x_i \neq 0$. Zauważmy również, że każdy minor stopnia większego od m jest zerowy, gdyż jego macierz zawiera przynajmniej jeden wiersz złożony z samych zer. Tak więc m jest maksymalnym stopniem niezerowego minora, a więc jest to rząd macierzy.

Wniosek 21.4.5. *Uwagi 21.4.3 i 21.4.4 pozwalają na bardzo sprawne wyliczenie rzędu danej macierzy.*

Uwaga 21.4.6. *Niezerowych minorów maksymalnego stopnia wystarczy szukać wśród macierzy okalających macierze poprzednio znalezionych niezerowych minorów.*

Przykład. Obliczmy rząd macierzy

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

Możemy to zrobić na dwa sposoby.
Sposób I. Szukamy niezerowego minora maksymalnego stopnia. Znajdźmy najpierw minor stopnia jeden. Zauważmy, że

$$\det(1) = 1 \neq 0.$$

Następnie szukamy minora stopnia dwa. Na początek sprawdźmy macierz okalającą macierz poprzednio znalezionego minora. Mamy

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

Znaleźliśmy zatem minor stopnia dwa. Szukamy teraz minora stopnia trzy. Szukamy maksymalnego niezerowego minora wśród macierzy 3×3 okalających macierz $\begin{pmatrix} 1 & 2 \\ 1 & 1 \end{pmatrix}$. Macierze takie powstają poprzez wykreślenie z macierzy A jednej z kolumn $3, 4$ i jednego z wierszy $3, 4$. Są cztery takie macierze. Na początek sprawdźmy wyznacznik macierzy powstałej przez wykreślenie z macierzy A czwartej kolumny i czwartego wiersza. Mamy

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

Sprawdźmy pozostałe macierze. Otrzymujemy

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

Na podstawie Uwagi 21.4.6 możemy stwierdzić, że macierz $\begin{pmatrix}1 & 2\\ 1 & 1\end{pmatrix}$ daje maksymalny niezerowy minor, czyli

$$r(A) = 2.$$

Sposób II. Możemy wykonywać operacje na wierszach i kolumnach macierzy tak długo, aż dostaniemy postać taką jak macierz w Uwadze 21.4.4. W przypadku macierzy $A = \begin{pmatrix}1 & 2 & 1 & -1\\ 1 & 1 & 0 & 1\\ 2 & 3 & 1 & 0\\ 0 & 1 & 1 & -2\end{pmatrix}$ otrzymujemy kolejno

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

W pierwszym przekształceniu powyżej od wiersza drugiego odjeliśmy wiersz pierwszy, zaś od wiersza trzeciego podwojony wiersz pierwszy. W drugim przekształceniu od wiersza trzeciego odejmujemy wiersz drugi natomiast do czwartego wiersza dodajemy drugi. Z ostatecznej postaci możemy wnioskować, że rząd tej macierzy $r(A) = 2$.

Ogólny sposób rozwiązywanie układów równań metodą minorów poznamy na przykładzie rozwiązania układu

$$(\star)\left\{\begin{array}{rcrcrcrcl} x_1 & + & 2x_2 & + & x_3 & - & x_4 & = & 1\\ x_1 & + & x_2 & & & + & x_4 & = & 0\\ 2x_1 & + & 3x_2 & + & x_3 & & & = & 1\\ & & x_2 & + & x_3 & - & 2x_4 & = & 1\end{array}\right.$$

Na początek sprawdźmy, czy układ ten ma rozwiązanie. W tym celu obliczmy rzędy macierzy

$$A = \begin{pmatrix}1 & 2 & 1 & -1\\ 1 & 1 & 0 & 1\\ 2 & 3 & 1 & 0\\ 0 & 1 & 1 & -2\end{pmatrix},\ U = \begin{pmatrix}1 & 2 & 1 & -1 & 1\\ 1 & 1 & 0 & 1 & 0\\ 2 & 3 & 1 & 0 & 1\\ 0 & 1 & 1 & -2 & 1\end{pmatrix}.$$

W poprzednim przykładzie obliczyliśmy rząd macierzy A otrzymaliśmy

$$r(A) = 2.$$

✍ Jako ćwiczenie pozostawimy czytelnikowi do sprawdzenia, że $r(U) = 2$.
Zgodnie z Twierdzeniem 21.4.2 możemy stwierdzić, że podany powyżej układ ma rozwiązanie, bo $r(A) = r(U)$.

Kiedy już wiemy, że dany układ ma rozwiązanie, możemy przystąpić do wyznaczania go. Rozwiązanie układu przeprowadzimy w czterech etapach.

ETAPY ROZWIĄZYWANIA UKŁADU $(\star)$.
Krok 1. Na początek znajdujemy macierz niezerowego minora maksymalnego stopnia macierzy głównej A. Jak już pokazaliśmy w poprzednim przykładzie szukaną macierzą jest lewa górna podmacierz

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

Krok 2. Następnie z wyjściowego układu usuwamy równania reprezentowane przez te wiersze macierzy głównej, które nie należą do macierzy wybranego w Kroku 1 minora. Odrzucone równania są zależne od równań, które pozostawiamy, dlatego nie są potrzebne do znalezienia rozwiązań układu. W naszym układzie usuwamy zatem równanie trzecie i czwarte. Otrzymujemy układ

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

Krok 3. Niewiadome z tych kolumn, które nie należą do wybranego minora, traktujemy jako parametry i przenosimy je na drugą stronę równań. Otrzymujemy

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

gdzie x_1, x_2 traktujemy jako niewiadome, natomiast x_3, x_4 jako parametry. Otrzymaliśmy zatem układ dwóch równań z dwoma niewiadomymi. Macierzą główną tego układu jest macierz wybranego w kroku pierwszym minora. Macierz ta ma niezerowy wyznacznik, tak więc układ ten jest układem Cramerowskim z parametrami.
Krok 4. Tak zmodyfikowany układ równań rozwiązujemy metodą Cramera. Wyznacznik główny tego układu to

$$W = \det \begin{pmatrix} 1 & 2 \\ 1 & 1 \end{pmatrix} = -1.$$

Pozostałe wyznaczniki to

$$W_{x_1} = \det \begin{pmatrix} 1 - x_3 + x_4 & 2 \\ -x_4 & 1 \end{pmatrix} = 1 - x_3 + x_4 + 2x_4 = 1 - x_3 + 3x_4,$$

$$W_{x_2} = \det \begin{pmatrix} 1 & 1 - x_3 + x_4 \\ 1 & -x_4 \end{pmatrix} = -x_4 - 1 + x_3 - x_4 = -1 + x_3 - 2x_4.$$

Rozwiązaniem naszego układu jest zatem

$$\begin{aligned} x_1 &= \tfrac{W_{x_1}}{W} = -1 + x_3 - 3x_4 \\ x_2 &= \tfrac{W_{x_2}}{W} = 1 - x_3 + 2x_4 \\ x_3 &= x_3 \\ x_4 &= x_4 \end{aligned}$$

Zauważmy, że w dwóch ostatnich równościach po lewej stronie x_3, x_4 występują jako zmienne, a po prawej jako parametry, dlatego lepiej będzie wprowadzić w tym miejscu dodatkowe zmienne. Niech $x_3 = t, x_4 = s$, wówczas rozwiązanie naszego układu możemy zapisać w postaci

$$\begin{aligned} x_1 &= \tfrac{W_{x_1}}{W} = -1 + t - 3s \\ x_2 &= \tfrac{W_{x_2}}{W} = 1 - t + 2s \\ x_3 &= t \\ x_4 &= s \end{aligned}$$

Zbiór rozwiązań zadanego układu równań zależy więc od dwóch parametrów.

Przedstawioną metodę można stosować do dowolnych układów równań liniowych.

Rozdział 22

Przekształcenia liniowe dowolnych przestrzeni wektorowych

Omawiając przestrzeń R^3 dość dużo uwagi poświęciliśmy przekształceniom, w szczególności przekształceniom liniowym. W wielu przypadkach przekształcenia te definiowaliśmy w sposób geometryczny. W dowolnych przestrzeniach wektorowych również będziemy zajmować się przekształceniami przestrzeni, jednak w tym wypadku będzie to podejście czysto algebraiczne. W pierwszym podrozdziale zdefiniujemy pojęcie przekształcenia liniowego dowolnej przestrzeni wektorowej, nauczymy się również opisywać takie przekształcenie przy użyciu macierzy. Drugi podrozdział poświęcimy na wprowadzenie podstawowej terminologii związanej z przekształceniami liniowymi.

22.1 Definicje i oznaczenia.

Definicja 22.1.1 (przekształcenie liniowe). Niech V, W będą przestrzeniami wektorowymi. Odwzorowanie $T : V \to W$ nazywamy przekształceniem liniowym, jeśli spełnia następujące warunki:

(1) przekształcenie to jest addytywne, czyli dla dowolnych wektorów v_1, v_2 z przestrzeni V zachodzi warunek

$$T(v_1 + v_2) = T(v_1) + T(v_2),$$

(2) przekształcenie to jest jednorodne, czyli dla dowolnego wektora v z przestrzeni V oraz dowolnej liczby rzeczywistej t mamy

$$T(t \cdot v) = t \cdot T(v).$$

✔ **Przykład.** Niech przestrzenią V będzie zbiór wielomianów o współczynnikach rzeczywistych stopnia co najwyżej trzeciego, tzn. $R_3[x]$, natomiast przestrzeń W to zbiór wielomianów o współczynnikach rzeczywistych stopnia co najwyżej drugiego, czyli $R_2[x]$. Określmy przekształcenie $T : V \to W$ wzorem

$$T(f) = f',$$

co oznacza, że każdemu wielomianowi z przestrzeni V przyporządkowujemy jego pochodną. Sprawdźmy, czy tak określone przekształcenie jest liniowe. Z własności pochodnej, znanych z analizy wiemy, że

$$T(f+g) = (f+g)' = f' + g' = T(f) + T(g),$$

zatem przekształcenie T jest addytywne. Podobnie sprawdzamy jednorodność tego przekształcenia

$$T(c \cdot f) = (c \cdot f)' = c \cdot f' = c \cdot T(f).$$

Przekształcenie T, czyli operacja brania pochodnej, zgodnie z Definicją 22.1.1, jest przekształceniem liniowym.

➢ Macierz przekształcenia liniowego w bazach.

Przekształcenie liniowe przestrzeni R^3 bardzo często opisywaliśmy przy użyciu macierzy. Możemy wyobrazić sobie, że takie przedstawienie mogłoby być użyteczne również w przypadku przekształceń dowolnych przestrzeni wektorowych. Przypomnijmy, że macierz przekształcenia w przestrzeni R^3 tworzyliśmy przy użyciu obrazów wersorów. W przypadku przestrzeni wektorowych możemy postąpić podobnie, rolę wersorów spełniać tu będą wektory bazowe.

Definicja 22.1.2 (macierz przekształcenia). Niech $B = (v_1, v_2, \ldots, v_k)$ będzie bazą przestrzeni V, natomiast $C = (w_1, w_2, \ldots, w_l)$ bazą przestrzeni W. Macierz przekształcenia T względem baz B, C będziemy tworzyć w dwóch etapach:

(1) obrazy $T(v_i)$ wektorów z bazy B przestrzeni V wyrażamy za pomocą współrzędnych w bazie C przestrzeni W, tzn.

$$T(v_i) = a_{1i}\overrightarrow{w_1} + a_{2i}\overrightarrow{w_2} + \ldots + a_{li}\overrightarrow{w_l} = \begin{pmatrix} a_{1i} \\ a_{2i} \\ \vdots \\ a_{li} \end{pmatrix}_C ,$$

(2) następnie współrzędne te umieszczamy kolejno w kolumnach macierzy

$$A = \begin{bmatrix} a_{11} & a_{12} & \ldots & a_{1k} \\ a_{21} & a_{22} & \ldots & a_{2k} \\ \vdots & \vdots & \ddots & \vdots \\ a_{l1} & a_{l2} & \ldots & a_{lk} \end{bmatrix}.$$

Macierz tę, czyli macierz przekształcenia T w bazach B w V i C w W, będziemy symbolicznie oznaczać przez $[T]_{BC}$.

✔ **Przykład.** Znajdźmy macierz przekształcenia T określonego w poprzednim przykładzie, czyli macierz przekształcenia, które każdemu wielomianowi o stopniu co najwyżej trzy, przyporządkowuje jego pochodną

$$T : R_3[x] \to R_2[x],\ T(f) = f'.$$

Niech $B = \{x^3, x^2, x, \mathbb{1}\}$ będzie bazą przestrzeni $R_3[x]$, natomiast $C = \{x^2, x, \mathbb{1}\}$ to baza przestrzeni $R_2[x]$. Znajdźmy obraz wektora x^3 z bazy B przez przekształcenie T

$$T(x^3) = (x^3)' = 3x^2,$$

następnie znajdźmy jego współrzędne w bazie C

$$T(x^3) = 3x^2 = 3 \cdot (x^2) + 0 \cdot (x) + 0 \cdot (\mathbb{1}) = \begin{pmatrix} 3 \\ 0 \\ 0 \end{pmatrix}_C .$$

Podobnie postępujemy z pozostałymi wektorami bazy B

$$T(x^2) = (x^2)' = 2x = 0 \cdot (x^2) + 2 \cdot (x) + 0 \cdot (\mathbb{1}) = \begin{pmatrix} 0 \\ 2 \\ 0 \end{pmatrix}_C ,$$

$$T(x) = (x)' = \mathbb{1} = 0 \cdot (x^2) + 0 \cdot (x) + 1 \cdot (\mathbb{1}) = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}_C ,$$

$$T(\mathbb{1}) = (\mathbb{1})' = 0 = 0 \cdot (x^2) + 0 \cdot (x) + 0 \cdot (\mathbb{1}) = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}_C .$$

Tworzymy teraz macierz tego przekształcenia

$$[T]_{BC} = \begin{bmatrix} 3 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}.$$

Uwaga 22.1.3. *Zauważmy, że gdy przekształcenie T przekształca przestrzeń V na samą siebie $T : V \to V$, to możemy wybrać tylko jedną bazę B, czyli uznać, że $C = B$. Obrazy wektorów v_i wyrażamy wówczas w bazie B*

$$T(v_i) = \begin{pmatrix} a_{1i} \\ a_{2i} \\ \vdots \\ a_{ki} \end{pmatrix}_B ,$$

zatem

$$[T]_{BB} = [T]_B = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1k} \\ a_{21} & a_{22} & \dots & a_{2k} \\ \vdots & \vdots & \ddots & \vdots \\ a_{k1} & a_{k2} & \dots & a_{kk} \end{bmatrix}.$$

22.2 Własności przekształceń liniowych.

W rozdziale tym poznamy kilka pojęć charakteryzujących liniowe przekształcenia przestrzeni. Na początek zdefiniujemy pewien rodzaj przekształceń przestrzeni o bardzo charakterystycznych własnościach, które wprowadzimy w końcowej części tego podrozdziału. Następnie zdefiniujemy bardzo istotne pojęcia - obraz i jądro przekształcenia, w oparciu o nie wyprowadzimy kilka charakterystycznych własności przestrzeni.

➢ **Przekształcenie odwracalne**

Definicja 22.2.1 (przekształcenie odwracalne). Przekształcenie liniowe $T : V \to W$ jest odwracalne, jeśli jest wzajemnie jednoznaczne tzn. różnowartościowe i „na".

Uwaga 22.2.2. *Inną nazwą stosowaną w odniesieniu do przekształceń odwracalnych jest izomorfizm przestrzeni wektorowych.*

Rozważmy izomorfizm T przestrzeni V, W. Skoro jest to przekształcenie odwracalne, to istnieje przekształcenie do niego odwrotne. Poniższy fakt przedstawia najbardziej charakterystyczną własność przekształcenia odwrotnego do przekształcenia liniowego. Dowód tego faktu jest analogiczny do dowodu przedstawionego wcześniej dla przypadków R^2 i R^3. Tutaj go pominiemy.

Fakt 22.2.3. *Przekształcenie odwrotne $T^{-1} : W \to V$ do odwracalnego przekształcenia liniowego $T : V \to W$ samo też jest przekształceniem liniowym.*

➢ **Obraz przekształcenia liniowego**

W poprzednim podrozdziale używaliśmy sformułowania obraz wektora przez przekształcenie. Możemy jednak zastanawiać się, jakim zbiorem jest zbiór obrazów wszystkich wektorów przestrzeni V przez dane przekształcenie.

Definicja 22.2.4 (obraz przekształcenia liniowego). Obrazem przekształcenia liniowego $T : V \to W$ oznaczanym $T(V)$ lub imT jest zbiór

$$\{T(v) \in W : v \in V\}.$$

Zauważmy, że obraz przekształcenia składa się z wektorów należących do przestrzeni W jest więc pewnym podzbiorem tej przestrzeni. Jak pokaże poniższy fakt ma on jeszcze inne własności.

Fakt 22.2.5. *Obraz przekształcenia liniowego $T : V \to W$ jest podprzestrzenią w przestrzeni wektorowej W.*

Dowód. Zgodnie z definicją podprzestrzeni wektorowej, czyli Definicją 19.2.1, wystarczy sprawdzić, że dla dowolnych wektorów należących do obrazu przekształcenia ich suma również należy do tego obrazu, oraz dla dowolnej liczby rzeczywistej iloczyn jej i wektora z obrazu przekształcenia jest wektorem z obrazu przekształcenia. Weźmy dowolne wektory w_1, w_2 należące do imT. Skoro wektory te należą do obrazu przekształcenia T, to istnieją takie wektory $v_1, v_2 \in V$, że $w_1 = T(v_1), w_2 = T(v_2)$. Wówczas

$$w_1 + w_2 = T(v_1) + T(v_2) = T(v_1 + v_2).$$

Zauważmy, że $v_1 + v_2$ jest wektorem z przestrzeni V, więc $T(v_1 + v_2)$ należy do imT, zatem $w_1 + w_2 = T(v_1 + v_2)$ należy do obrazu przekształcenia T. Zauważmy, że również

$$t \cdot w_1 = t \cdot T(v_1) = T(t \cdot v_1)$$

dla dowolnego t jest obrazem pewnego wektora z przestrzeni v, tak więc należy do obrazu imT przekształcenia T. Otrzymujemy więc, że obraz imT jest podprzestrzenią w przestrzeni W. □

Poniżej przedstawimy własność obrazu przekształcenia, która pozwala na bardzo szybkie wyznaczenie go.

Lemat 22.2.6. *Niech $T : V \to W$ będzie przekształceniem liniowym i niech wektory $v_1, v_2, \ldots, v_n$ będą układem generującym przestrzeń V. Wówczas*

$$imT = Lin\{T(v_1), T(v_2), \ldots, T(v_n)\}.$$

Dowód. Pokażemy, że dowolny wektor ze zbioru imT rzeczywiście da się przedstawić w postaci kombinacji liniowej wektorów $T(v_1), T(v_2), \ldots, T(v_n)$. Niech $w \in imT$, co oznacza, że $w = T(v)$ dla pewnego wektora $v \in V$. Ponieważ $v_1, v_2, \ldots, v_n$ generują przestrzeń V, więc $v = t_1v_1 + t_2v_2 + \ldots + t_nv_n$, dla pewnych $t_1, t_2, \ldots, t_n \in R$. Mamy wtedy

$$w = T(v) = T(t_1v_1 + t_2v_2 + \ldots + t_nv_n) = t_1T(v_1) + t_2T(v_2) + \ldots + t_nT(v_n),$$

zatem $w \in Lin\{T(v_1), T(v_2), \ldots, T(v_n)\}$. Ponieważ w był dowolnie wybranym wektorem ze zbioru imT, więc

$$imT = Lin\{T(v_1), T(v_2), \ldots, T(v_n)\}.$$

□

➢ Rząd przekształcenia liniowego

Dla dowolnej przestrzeni wektorowej zdefiniowaliśmy, czym jest jej wymiar. Zauważmy, że obraz przekształcenia jako podprzestrzeń przestrzeni wektorowej sam jest przestrzenią wektorową, więc możemy dla niego określić wymiar, który w tym przypadku ma specjalną nazwę.

Definicja 22.2.7 (rząd przekształcenia). Rzędem $r(T)$ przekształcenia liniowego $T : V \to W$ nazywamy wymiar obrazu tego przekształcenia, tzn.

$$r(T) = dim(imT).$$

Uwaga 22.2.8. *Rząd przekształcenia $T : R^n \to R^n$ jest równy rzędowi macierzy przekształcenia T, zauważmy bowiem, że*

$$r(T) = dim(imT) = dim(Lin\{T(E_1), T(E_2), \ldots T(E_n)\}) = r(m(T)).$$

Powyższą uwagę możemy sformułować dla bardziej ogólnego przypadku.

Uwaga 22.2.9. *Dla dowolnego $T : V \to W$ i dowolnych baz B w V i C w W mamy*

$$r(T) = r([T]_{B,C}).$$

Uzasadnienie powyższej uwagi jest analogiczne do uzasadnienia Uwagi 22.2.8 dlatego je pominiemy.

➢ Jądro przekształcenia liniowego

Definicja 22.2.10 (jądro przekształcenia). Jądrem przekształcenia liniowego $T : V \to W$, oznaczanym $KerT$, jest zbiór

$$\{v \in V : T(v) = 0\}.$$

Fakt 22.2.11. *Dla dowolnego przekształcenia liniowego T mamy $0 \in KerT$.*

✍ **Ćwiczenie.** Łatwy dowód powyższego faktu pozostawimy do wykonania czytelnikowi.

Fakt 22.2.12. *Jądro przekształcenia liniowego T jest podprzestrzenią w przestrzeni wektorowej V.*

Dowód. Podobnie jak w dowodzie Faktu 22.2.5 musimy pokazać, że zbiór $KerT$ jest zamknięty na dodawanie i mnożenie przez liczbę. Weźmy dowolne wektory $v_1, v_2 \in KerT$. Wtedy, zgodnie z definicją $T(v_1) = 0, T(v_2) = 0$. Zauważmy, że

$$T(v_1 + v_2) = T(v_1) + T(v_2) = 0 + 0 = 0,$$

a więc $v_1 + v_2 \in KerT$. Podobnie otrzymujemy

$$T(t \cdot v_1) = t \cdot T(v_1) = t \cdot 0 = 0,$$

czyli $t \cdot v_1 \in KerT$.
Zbiór ten rzeczywiście jest więc podprzestrzenią przestrzeni V. □

Lemat 22.2.13. *Niech $T : V \to W$ będzie przekształceniem liniowym, $e_1, e_2, \ldots, e_k$ bazą jądra $KerT$ tego przekształcenia, natomiast $e_{k+1}, \ldots, e_n$ będą wektorami uzupełniającymi układ $e_1, e_2, \ldots, e_n$ do bazy $e_1, e_2, \ldots, e_k, e_{k+1}, \ldots, e_k$ całej przestrzeni V. Wówczas*

(1) wektory $T(e_{k+1}), \ldots, T(e_n)$ są liniowo niezależne w przestrzeni W,

(2) wektory $T(e_{k+1}), \ldots, T(e_n)$ generują obraz imT tego przekształcenia.

Dowód. Zacznijmy od dowodu drugiej części powyższego lematu. Jak wiemy z Lematu 22.2.6,

$$imT = Lin\{T(e_1), T(e_2), \ldots, T(e_k), T(e_{k+1}), \ldots, T(e_n)\}.$$

Wiemy również, że $T(e_1) = T(e_2) = \ldots = T(e_k) = 0$, bo wektory $e_1, e_2, \ldots, e_k$ należą do jądra przekształcenia T. Otrzymujemy zatem

$$imT = Lin\{T(e_1), T(e_2), \ldots, T(e_k), T(e_{k+1}), \ldots, T(e_n)\} = Lin\{T(e_{k+1}), \ldots, T(e_n)\}.$$

Pokażemy teraz, że zbiór złożony z wektorów $T(e_{k+1}), \ldots, T(e_n)$ jest liniowo niezależny. Gdyby wektory te były liniowo zależne, to dla pewnych nietrywialnych współczynników $t_{k+1}, \ldots, t_n$ mielibyśmy

$$t_{k+1}e_{k+1} + \ldots + t_n e_n = 0.$$

Wtedy przy użyciu własności wynikających z liniowości przekształcenia T mielibyśmy

$$T(t_{k+1}e_{k+1} + \ldots + t_n e_n) = 0,$$

czyli $t_{k+1}e_{k+1} + \ldots + t_n e_n \in KerT$. Z drugiej strony wiemy, że wektory $e_1, e_2, \ldots, e_k$ tworzą bazę $KerT$, stąd dla pewnych współczynników $s_1, s_2, \ldots s_k$ mamy

$$t_{k+1}e_{k+1} + \ldots + t_n e_n = s_1 e_1 + s_2 e_2 + \ldots + s_k e_k,$$

co po przekształceniu daje

$$s_1 e_1 + s_2 e_2 + \ldots + s_k e_k + (-t_{k+1})e_{k+1} + \ldots + (-t_n)e_n = 0.$$

Znaleźliśmy zatem nietrywialną kombinację wektorów $e_1, e_2, \ldots, e_k, e_{k+1}, \ldots, e_n$ dającą wektor zerowy więc wektory te tworzą układ liniowo zależny. Wiemy jednak, że wektory te są bazą przestrzeni V, a co za tym idzie są liniowo zależne. Otrzymaliśmy zatem sprzeczność, czyli wektory $T(e_{k+1}), \ldots, T(e_n)$ są liniowo niezależne. $\square$

W oparciu o oznaczenia wprowadzone w powyższym Lemacie podamy teraz kilka wniosków dotyczących poznanych w tym podrozdziale pojęć. Pierwszy wniosek poniżej jest bezpośrednią konsewencją udowodnionych już własności (1) i (2) z Lematu 22.2.13.

Wniosek 22.2.14. *Przy założeniach Lematu 22.2.13, układ wektorów $T(e_{k+1}), \ldots, T(e_n)$ jest bazą podprzestrzeni imT.*

Wniosek 22.2.15 (zasada zachowania wymiaru). *Dla przekształcenia $T : V \to W$ mamy*

$$dim(imT) + dim(KerT) = dimV.$$

Dowód. Przy oznaczeniach takich jak w Lemacie 22.2.13 mamy

$$dimV = n,\ dim(KerT) = k,\ dim(imT) = n - k,$$

więc skoro $k + (n - k) = n$, to mamy tezę wniosku. Zauważmy, że rozważenie takiej sytuacji wystarczy, gdyż w Lemacie rozważaliśmy sytuację ogólną. $\square$

Wniosek 22.2.16. *Jeżeli przestrzenie V, W są izomorficzne, to wówczas ich wymiary muszą być sobie równe*

$$dimV = dimW.$$

Dowód. Niech $T : V \to W$ będzie izomorfizmem przestrzeni V, W. Z określenia takiego przekształcenia wiemy, że T jest różnowartościowe, więc $KerT = \{0\}$. Wtedy

$dim(KerT) = 0$. Wiemy również, że T jest „na”, stąd $imT = W$. Możemy teraz zastosować poprzedni wniosek i otrzymujemy

$$dim(KerT) + dim(imT) = dimV,$$

$$0 + dimW = dimV.$$

Tak więc

$$dimW = dimV.$$

□

Wniosek 22.2.17. *Jeśli $T : V \to W$ jest izomorfizm przestrzeni wektorowych, zaś wektory $v_1, v_2, \ldots, v_n$ są bazą przestrzeni V, to wektory $T(v_1), T(v_2), \ldots, T(v_n)$ są bazą przestrzeni W.*

Dowód. Jak wiemy jądrem izomorfizmu, które jest przekształceniem różnowartościowym, jest $KerT = \{0\}$. Jak wiemy z Uwagi 19.3.3, zbiór $\{0\}$ jest generowany przez zbiór pusty

$$Lin(\emptyset) = \{0\}.$$

Tak więc

$$KerT = Lin\{\emptyset\},$$

stąd zbiór wektorów bazowych tej podprzestrzeni jest pusty, a tym samym jej wymiar jest równy 0. Możemy pustą bazę $KerT$ uzupełnić do bazy przestrzeni V wektorami $e_1 = v_1, e_2 = v_2, \ldots, e_n = v_n$. Wówczas, jak wiemy z Lematu 22.2.6, układ $T(e_1) = T(v_1), T(e_2) = T(v_2), \ldots, T(e_n) = T(v_n)$ generuje przestrzeń imT, dodatkowo z drugiej części Lematu 22.2.13 wiemy, że układ ten jest liniowo niezależny, tak więc jest on bazą przestrzeni $imT = W$. □

Fakt 22.2.18. *Niech V, W będą przestrzeniami wektorowymi, i niech $v_1, v_2, \ldots, v_n$ będzie bazą przestrzeni V. Przekształcenie liniowe $T : V \to W$ jest jednoznacznie określone przez podanie obrazów $T(v_1), T(v_2), \ldots, T(v_n) \in W$ wektorów bazowych.*

Dowód. Sprawdźmy, jak wygląda obraz dowolnego wektora z przestrzeni V przez przekształcenie T. Ponieważ wektory $v_1, v_2, \ldots, v_k$ tworzą bazę przestrzeni V, to wektor v można w jednoznaczny sposób przedstawić w postaci $v = t_1v_1 + t_2v_2 + \ldots + t_nv_n$. Wówczas

$$T(v) = T(t_1v_1 + t_2v_2 + \ldots + t_nv_n) = t_1T(v_1) + t_2T(v_2) + \ldots + t_nT(v_n) =$$

$$= t_1w_1 + t_2w_2 + \ldots + t_nw_n,$$

więc T jest jednoznacznie określone dla dowolnego wektora $v \in V$. Sprawdzenie, że przekształcenie określone wzorem

$$T(t_1v_1 + t_2v_2 + \ldots + t_nv_n) = t_1w_1 + t_2w_2 + \ldots + t_nw_n$$

✍ jest liniowe pozostawimy do wykonania czytelnikowi. □

Twierdzenie 22.2.19. *Przestrzenie wektorowe V, W są izomorficzne wtedy i tylko wtedy gdy*

$$dimV = dimW.$$

Dowód. Jedna z implikacji składających się na to twierdzenie, to Wniosek 22.2.16. Wystarczy zatem udowodnić drugą implikację. Musimy pokazać, że jeśli wymiary przestrzeni V, W są równe, to przestrzenie te są izomorficzne. Niech $v_1, v_2, \ldots, v_n$ będzie bazą przestrzeni V. Ponieważ $dimV = dimW$, to możemy wybrać bazę $w_1, w_2, \ldots, w_n$ dla przestrzeni W złożoną z tej samej liczby wektorów. Określmy przekształcenie liniowe $T : V \to W$ warunkiem $T(v_i) = w_i$ dla $i = 1, 2, \ldots, n$. Wówczas

$$T(t_1v_1 + t_2v_2 + t_nv_n) = t_1T(v_1) + t_2T(v_2) + \ldots + t_nT(v_n) = t_1w_1 + t_2w_2 + \ldots + t_nw_n.$$

Stąd już łatwo widać, że przekształcenie T jest różnowartościowe i „na”, czyli jest izomorfizmem. □

Lemat 22.2.20. *Przekształcenie liniowe jest różnowartościowe wtedy i tylko wtedy gdy* $KerT = \{0\}$.

Dowód. Jak wiadomo $T(0) = 0$ więc z różnowartościowości przekształcenia T wynika, że dla $v \neq 0$ mamy $T(v) \neq 0$, stąd $KerT = \{0\}$. Pokażmy teraz przeciwną implikację. Załóżmy, że $KerT = \{0\}$ i załóżmy nie wprost, że przekształcenie T nie jest różnowartościowe. Wówczas dla pewnych różnych v, v' mamy $T(v) = T(v')$, zatem $0 = T(v) - T(v') = T(v - v')$ a więc niezerowy wektor $v - v'$ byłby wektorem z jądra przekształcenia T. Otrzymujemy sprzeczność, bo założyliśmy, że jedynym wektorem należącym do $KerT$ jest wektor zerowy. □

Wniosek 22.2.21. *Jeśli przekształcenie* $T : V \to W$ *jest liniowe,* $KerT = \{0\}$ *i* $dimV = dimW$*, wówczas* T *jest izomorfizmem.*

Dowód. Aby udowodnić, że przekształcenie T jest izomorfizmem wystarczy pokazać, że jest różnowartościowe i „na”. Jądro $KerT = \{0\}$, więc jak wiemy z poprzedniego lematu przekształcenie T jest różnowartościowe.
Z Wniosku 22.2.15 wiemy, że

$$dim(imT) + dim(KerT) = dimV,$$

ale skoro $KerT = \{0\}$ to jego wymiar jest równy 0, tak więc

$$dimV = dim(imT).$$

Z założenia mamy zatem

$$dimW = dimV = dim(imT).$$

Wiemy również, że imT jest podprzestrzenią W, zatem w związku z tym, że $dim(imT) = dimW$, to $imT = W$. Czyli T jest przekształceniem „na”. W związku z tym, że przekształcenie T jest różnowartościowe i „na” możemy wnioskować, że jest ono izomorfizmem. □

Rozdział 23

Iloczyn skalarny i przestrzenie euklidesowe

W tym rozdziale zajmiemy się pojęciem iloczynu skalarnego. Na jego podstawie określimy, czym są przestrzenie euklidesowe. Dzięki iloczynowi skalarnemu będziemy mogli w tych przestrzeniach wprowadzić elementy geometrii podobnej do tej, jaką znamy z przestrzeni R^2 i R^3. W dalszej części rozdziału wprowadzimy bardzo ważne pojęcia układów ortogonalnych i ortonormalnych, wraz z ich zastosowaniami. Określimy np. czym jest rzut ortogonalny na podprzestrzeń i przekształcenia ortogonalne.

23.1 Podstawowe definicje.

Podrozdział ten poświęcimy zdefiniowaniu podstawowych pojęć, którymi będziemy posługiwać się w dalszej części tego rozdziału. W przestrzeniach R^2, R^3 a następnie w dowolnej przestrzeni R^n wprowadziliśmy iloczyn skalarny. Okazuje się, że analogiczne pojęcie możemy zdefiniować dla dowolnej przestrzeni wektorowej.

Definicja 23.1.1 (**iloczyn skalarny**). Iloczynem skalarnym na przestrzeni wektorowej V nazywamy każde odwzorowanie $\Big\langle\ ,\ \Big\rangle : V \times V \to R$, czyli odwzorowanie, które dowolnej parze wektorów v, w, z przestrzeni V, przyporządkowuje pewną liczbę rzeczywistą, oznaczoną symbolem $\Big\langle v, w\Big\rangle$, spełniające następujące warunki:

(1) dla dowolnych wektorów v, w z przestrzeni V mamy

$$\Big\langle v, w\Big\rangle = \Big\langle w, v\Big\rangle,$$

(2) dla dowolnych wektorów v, w z przestrzeni V oraz dowolnej liczby rzeczywistej λ mamy

$$\Big\langle \lambda v, w\Big\rangle = \lambda\Big\langle v, w\Big\rangle,$$

(3) dla dowolnych wektorów v_1, v_2, w z przestrzeni V mamy

$$\Big\langle v_1 + v_2, w\Big\rangle = \Big\langle v_1, w\Big\rangle + \Big\langle v_2, w\Big\rangle$$

(4) dla dowolnego, niezerowego wektora v z przestrzeni V mamy

$$\left\langle v, v \right\rangle > 0.$$

Warunek pierwszy nazywamy symetrycznością iloczynu skalarnego. Warunek drugi i trzeci to liniowość względem pierwszej współrzędnej, przy czym warunek drugi nazywamy jednorodnością a warunek trzeci addytywnością.

Zauważmy, że iloczyn skalarny określony dla przestrzeni R^3 spełnia wszystkie warunki z powyższej definicji.

Obserwacja 23.1.2. *Zauważmy, że z własności* $(1), (2), (3)$ *wynika liniowość względem drugiej zmiennej, tzn:*

(a) $\left\langle v, \lambda v \right\rangle = \lambda \left\langle w, v \right\rangle$, *bo*
$\left\langle v, \lambda w \right\rangle = \left\langle \lambda w, v \right\rangle = \lambda \left\langle w, v \right\rangle = \lambda \left\langle v, w \right\rangle$,

(b) $\left\langle v, w_1 + w_2 \right\rangle = \left\langle v, w_1 \right\rangle + \left\langle v, w_2 \right\rangle$, *bo*
$\left\langle v, w_1 + w_2 \right\rangle = \left\langle w_1 + w_2, v \right\rangle = \left\langle w_1, v \right\rangle + \left\langle w_2, v \right\rangle = \left\langle v, w_1 \right\rangle + \left\langle v, w_2 \right\rangle$.

Definicja 23.1.3 (**przestrzeń euklidesowa**)**.** Przestrzeń wektorowa V, w której określony jest iloczyn skalarny nazywamy przestrzenią euklidesową.

Poznamy teraz kilka przykładów przestrzeni euklidesowych wraz z określonym w nich iloczynem skalarnym.

✔ **Przykład.**

(A) Jeśli $V = R^n$, to iloczyn skalarny, dany wzorem

$$\left\langle v, w \right\rangle = \sum_{i=1}^{n} v_i w_i,$$

dla $v = [v_1, v_2, \ldots, v_n], w = [w_1, w_2, \ldots, w_n]$, spełnia warunki $(1) - (4)$. Przestrzeń R^n z tym iloczynem skalarnym oznaczać będziemy symbolem E^n.

(B) W przestrzeni $V = C[a, b]$, czyli w przestrzeni funkcji ciągłych na odcinku $[a, b]$, iloczyn skalarny można określić za pomocą całkowania:

$$\left\langle f, g \right\rangle = \int_a^b f(x) \cdot g(x) dx.$$

Sprawdźmy, czy iloczyn ten spełnia warunki podane w Definicji 23.1.1.

(1) $\left\langle g, f \right\rangle = \int_a^b g(x) \cdot f(x) dx = \int_a^b f(x) \cdot g(x) dx = \left\langle f, g \right\rangle$.

(2) Dla dowolnej liczby rzeczywistej t mamy

$$\left\langle t \cdot f, g \right\rangle = \int_a^b t \cdot f(x) \cdot g(x) dx = t \cdot \int_a^b f(x) \cdot g(x) dx = t \cdot \left\langle f, g \right\rangle.$$

(3) $\left\langle f_1 + f_2, g\right\rangle = \int_a^b (f_1 + f_2)(x) \cdot g(x)dx = \int_a^b (f_1(x) + f_2(x)) \cdot g(x)dx = \int_a^b (f_1(x) \cdot g(x) + f_2(x) \cdot g(x))dx = \int_a^b f_1(x) \cdot g(x)dx + \int_a^b f_2(x) \cdot g(x)dx = \left\langle f_1, g\right\rangle + \left\langle f_2, g\right\rangle$.

(4) Niech f będzie funkcją, która nie jest tożsamościowo równa zero, $f \not\equiv 0$. Mamy wówczas

$$\left\langle f, f\right\rangle = \int_a^b f(x) \cdot f(x)dx = \int_a^b |f(x)|^2 dx.$$

Z analizy wiemy, że jeśli funkcja ciągła $h : [a, b] \to R$ jest nieujemna, a w pewnym miejscu ściśle dodatnia to $\int_a^b h(x)dx > 0$. Stosując to do funkcji $h(x) = |f(x)|^2$ otrzymujemy $\left\langle f, f\right\rangle = \int_a^b |f(x)|^2 dx > 0$.

(C) W przestrzeni R^n możemy wprowadzić inny iloczyn skalarny. Dla wektorów $v = [v_1, v_2, \ldots, v_n], w = [w_1, w_2, \ldots, w_n]$ jest nim

$$\left\langle v, w\right\rangle = \sum_{i=1}^{n} a_i \cdot v_i \cdot w_i,$$

✍ gdzie a_i są stałymi dodatnimi. Sprawdzenie, że tak określony iloczyn skalarny spełnia warunki $(1)-(4)$ podane w Definicji 23.1.1 pozostawimy do wykonania czytelnikowi.

Przy użyciu takiego iloczynu skalarnego można do przestrzeni wektorowej wprowadzić geometrię. Zacznijmy od wprowadzenia normy, czyli odpowiednika znanej nam z przestrzeni R^2 i R^3 długości wektora.

➢ Norma w przestrzeni euklidesowej

Definicja 23.1.4 (norma wektora). Normą (długością) wektora v przestrzeni euklidesowej E nazywamy liczbę

$$\|v\| = \sqrt{\left\langle v, v\right\rangle}.$$

Uwaga 23.1.5. *Z własności* (4) *mamy, że jeśli* $v \neq 0$*, to* $\left\langle v, v\right\rangle > 0$*, czyli* $\|v\| > 0$*.*

Fakt 23.1.6. *Dla tak określonej normy, mamy*

$$\|\lambda \cdot v\| = |\lambda| \cdot \|v\|,$$

w szczególności $\| - v\| = \|v\|$*.*

Dowód. Powyższy fakt udowodnimy korzystając z drugiej własności iloczynu skalarnego określonej w Definicji 23.1.1. Z określenia normy mamy

$$\|\lambda \cdot v\| = \sqrt{\left\langle \lambda \cdot v, \lambda \cdot v\right\rangle}.$$

Z wspomnianej już własności iloczynu skalarnego mamy zatem

$$\sqrt{\left\langle \lambda \cdot v, \lambda \cdot v\right\rangle} = \sqrt{\lambda\left\langle v, \lambda \cdot v\right\rangle} = \sqrt{\lambda^2\left\langle v, v\right\rangle} = |\lambda| \cdot \|v\|.$$

Co kończy dowód. □

Możemy wprowadzać dalsze geometryczne pojęcia.

➢ **Kąt między wektorami**

Definicja 23.1.7 (kąt między wektorami). Cosinusem kąta między wektorami v_1, v_2 nazywamy liczbę

$$cos(\measuredangle v_1, v_2) = \frac{\left\langle v_1, v_2 \right\rangle}{\|v_1\|\|v_2\|}.$$

✔ **Przykład.** W przestrzeni $V = C[0,1]$ iloczyn skalarny to $\left\langle f, g \right\rangle = \int_0^1 f(x) \cdot g(x)dx$. Rozważmy funkcje $f_1(x) = x, f_2(x) = 1$, które w tej przestrzeni są wektorami. Do obliczenia kąta między tymi wektorami musimy najpierw obliczyć normy tych wektorów.

$$\|f_1\| = \sqrt{\int_0^1 f_1(x) \cdot f_1(x)dx} = \sqrt{\int_0^1 x^2 dx} = \sqrt{\frac{1}{3}} = \frac{\sqrt{3}}{3}$$

$$\|f_2\| = \sqrt{\int_0^1 f_2(x) \cdot f_2(x)dx} = \sqrt{\int_0^1 1^2 dx} = 1.$$

Wówczas cosinus kąta między tymi wektorami to

$$cos(\measuredangle f_1, f_2) = \frac{\left\langle f_1, f_2 \right\rangle}{\|f_1\|\|f_2\|} = \frac{\int_0^1 x \cdot 1dx}{\frac{\sqrt{3}}{3} \cdot 1} = \frac{\frac{1}{2}}{\frac{\sqrt{3}}{3}} = \frac{\sqrt{3}}{2}.$$

Tak więc szukany kąt to

$$\|\measuredangle f_1, f_2\| = \frac{\pi}{6}.$$

Dla dowolnej przestrzeni wektorowej wprowadzimy teraz, znaną nam dla przestrzeni R^n (Uwaga 17.6.2), nierówność Schwartza.

Twierdzenie 23.1.8 (nierówność Schwartza). *W przestrzeni euklidesowej E, dla dowolnych wektorów $v_1, v_2 \in E$ zachodzi*

$$\left\langle v_1, v_2 \right\rangle^2 \leqslant \|v_1\|^2 \|v_2\|^2.$$

Dowód tego twierdzenia jest analogiczny do dowodu Uwagi 17.6.2. Tutaj go pominiemy.

Uwaga 23.1.9. *Dla dowolnych wektorów v_1, v_2 mamy*

$$-1 \leqslant cos(\measuredangle v_1, v_2) \leqslant 1,$$

a więc miara kąta jest zawsze dobrze określona.

Dowód. Z nierówności Szwartza mamy

$$\left\langle v_1, v_2 \right\rangle^2 \leqslant \|v_1\|^2 \|v_2\|^2,$$

zatem

$$0 \leqslant \frac{\left\langle v_1, v_2 \right\rangle^2}{\|v_1\|^2 \|v_2\|^2} \leqslant 1,$$

stąd
$$\left| \frac{\langle v_1, v_2 \rangle}{\|v_1\| \|v_2\|} \right| \leqslant 1.$$
Możemy zatem stwierdzić, że kąt między wektorami v_1, v_2 jest zawsze dobrze określony.

□

➢ Wektory ortogonalne

Iloczyn skalarny, w przestrzeniach R^n, często wykorzystywaliśmy do stwierdzenia prostopadłości, czyli inaczej ortogonalności, wektorów. Podobnie jest w przypadku dowolnej przestrzeni wektorowej.

Definicja 23.1.10 (wektory ortogonalne). Wektory $v_1, v_2 \in E$ są ortogonalne (prostopadłe) jeśli kąt między tymi wektorami jest kątem prostym, $|\measuredangle v_1, v_2| = \frac{\pi}{2}$, a to zachodzi wtedy i tylko wtedy gdy $\cos(\measuredangle v_1, v_2) = 0$, a co za tym idzie wtedy i tylko wtedy gdy $\langle v_1, v_2 \rangle = 0$.

✔ **Przykład.** W przestrzeni euklidesowej $V = C[-1, 1]$ pokażemy, że funkcje $f_1(x) = x, f_2(x) = x^2$ są ortogonalne. Przypomnijmy, że iloczyn skalarny w tej przestrzeni określony jest wzorem
$$\langle f_1, f_2 \rangle = \int_{-1}^{1} f_1(x) \cdot f_2(x) dx.$$
Sprawdźmy, że $\langle f_1, f_2 \rangle = 0$.
$$\langle f_1, f_2 \rangle = \int_{-1}^{1} x \cdot x^2 dx = \int_{-1}^{1} x^3 dx = \frac{1}{4} x^4 \Big|_{-1}^{1} = \frac{1}{4} - \frac{1}{4} = 0,$$
a więc funkcje x i x^2 są w tej przestrzeni ortogonalne.

➢ Odległość między wektorami

Definicja 23.1.11 (odległość pomiędzy wektorami). Odległością pomiędzy wektorami $v, w \in E$ nazywamy liczbę
$$d(v, w) = \|v - w\|.$$

Uwaga 23.1.12. *Łatwo sprawdzić, że tak zdefiniowana odległość spełnia warunki*

(i) $d(v, w) = d(w, v)$,

(ii) jeśli $v \neq w$, *to* $d(v, w) > 0$, *oraz* $d(v, v) = 0$,

(iii) $d(u, v) + d(v, w) \geqslant d(u, w)$.

Trzecia z podanych powyżej własności nosi nazwę nierówności trójkąta i daje się wyprowadzić z nierówności Schwartza w taki sam sposób jak nierówność trójkąta w R^n (podrozdział 17.7).

✍ **Ćwiczenie.** Jako ćwiczenie pozostawimy czytelnikowi udowodnienie tych własności.

✔ **Przykład.** W przestrzeni $V = C[0,1]$ znajdźmy odległość między wektorami $f_1(x) = x, f_2(x) = x^2$.

$$d(f_1, f_2) = \sqrt{\left\langle f_1 - f_2, f_1 - f_2 \right\rangle} = \sqrt{\int_0^1 (f_1(x) - f_2(x))^2 dx} =$$

$$= \sqrt{\int_0^1 (x - x^2)^2 dx} = \sqrt{\int_0^1 x^2 - 2x^3 + x^4 dx} = \sqrt{\frac{1}{3}x^3 - \frac{1}{2}x^4 + \frac{1}{5}x^5 \Big|_0^1} =$$

$$= \sqrt{\frac{1}{3} - \frac{1}{2} + \frac{1}{5}} = \sqrt{\frac{1}{30}}.$$

Odległość między wektorami f_1, f_2 wynosi $\sqrt{\frac{1}{30}}$.

➢ **Składowa równoległa i prostopadła do wektora**

Lemat 23.1.13. *Niech $v_0 \neq 0$ będzie ustalonym wektorem przestrzeni euklidesowej E. Każdy wektor $v \in E$ rozkłada się, i to jednoznacznie, w sumę*

$$v = \lambda \cdot v_0 + v_\perp,$$

gdzie $\left\langle v_\perp, v_0 \right\rangle = 0$.

Uwaga 23.1.14. *Wektor λv_0 nazywamy składową wektora v równoległą do v_0, zaś $v_\perp$ składową prostopadłą do v_0.*

Dowód (Lematu). Znajdźmy takie λ, aby $v = \lambda \cdot v_0 + v_\perp$. Zauważmy, że mamy wówczas

$$v_\perp = v - \lambda v_0.$$

Wektor $v_\perp$ ma być prostopadły do wektora v_0, więc musi zachodzić

$$\left\langle v_\perp, v_0 \right\rangle = 0.$$

Mamy zatem

$$\left\langle v - \lambda v_0, v_0 \right\rangle = 0,$$

co po przekształceniu, przy wykorzystaniu liniowości iloczynu skalarnego, daje

$$\left\langle v, v_0 \right\rangle = \lambda \left\langle v_0, v_0 \right\rangle.$$

Szukany współczynnik λ możemy więc jednoznacznie wyznaczyć ze wzoru

$$\lambda = \frac{\left\langle v, v_0 \right\rangle}{\left\langle v_0, v_0 \right\rangle}.$$

Składowa równoległa to zatem

$$\frac{\left\langle v, v_0 \right\rangle}{\left\langle v_0, v_0 \right\rangle} \cdot v_0,$$

zaś składowa prostopadła to

$$v - \frac{\left\langle v, v_0 \right\rangle}{\left\langle v_0, v_0 \right\rangle} \cdot v_0.$$

□

23.2 Układy ortogonalne i ortonormalne.

W poprzednim podrozdziale wprowadziliśmy pojęcie ortogonalności wektorów, a w tym szerzej się nim zajmiemy. Zdefiniujemy czym są układy wektorów ortogonalnych, poznamy ich własności, a następnie nauczymy się znajdować takie układy.

Definicja 23.2.1 (**układ ortogonalny i ortonormalny**). Układ niezerowych wektorów $v_1, v_2, \ldots, v_k$ przestrzeni euklidesowej E nazywamy ortogonalnym jeśli iloczyn skalarny $\left\langle v_i, v_j \right\rangle = 0$ dla $i \neq j$, czyli wtedy gdy wektory są parami ortogonalne. Układ taki jest ortonormalny jeśli dodatkowo normy $\|v_i\| = 1$, tzn. wektory v_i są jednostkowe.

Podamy teraz ciekawą zależność między ortogonalnością a liniową niezależnością wektorów.

Lemat 23.2.2. *Każdy ortogonalny układ wektorów jest liniowo niezależny.*

Dowód. Niech kombinacja liniowa $t_1v_1 + t_2v_2 + \ldots + t_kv_k$ wektorów z ortogonalnego układu daje wektor $\overrightarrow{O}$. Musimy pokazać, że wówczas $t_1 = t_2 = \ldots = t_k = 0$. Zauważmy, że z liniowości iloczynu skalarnego mamy

$$0 = \left\langle \overrightarrow{O}, v_i \right\rangle = \left\langle t_1v_1 + t_2v_2 + \ldots + t_kv_k, v_i \right\rangle =$$

$$= t_1\left\langle v_1, v_i \right\rangle + t_2\left\langle v_2, v_i \right\rangle + \ldots + t_k\left\langle v_k, v_i \right\rangle = t_i\left\langle v_i, v_i \right\rangle.$$

Wiemy jednak, że $\left\langle v_i, v_i \right\rangle > 0$, więc musi zachodzić $t_i = 0$. Tak jest dla wszystkich współczynników i stąd wynika liniowa niezależność układu $v_1, v_2, \ldots, v_k$. □

Jak wiemy w danej przestrzeni wektorowej liniowo niezależny układ nie może składać się z większej liczby wektorów niż wynosi wymiar przestrzeni. Możemy zatem sformułować następujący wniosek.

Wniosek 23.2.3. *W n-wymiarowej przestrzeni euklidesowej ortogonalny układ wektorów nie może mieć więcej niż n wektorów.*

Twierdzenie 23.2.4. *W każdej n-wymiarowej przestrzeni euklidesowej E istnieje baza ortogonalna.*

Dowód powyższego twierdzenia zawierać będzie bardzo użyteczny algorytm znajdowania bazy ortogonalnej. Algorytm ten nosi nazwę metody ortogonalizacji Grama - Schmidta.

Dowód. Niech $g_1, g_2, \ldots, g_n$ będzie dowolną bazą przestrzeni euklidesowej E. Zmodyfikujemy tę bazę tak, aby stała się bazą ortogonalną $e_1, e_2, \ldots, e_n$.
Krok (1). Pierwszy wektor pozostawimy bez zmian

$$e_1 = g_1.$$

Krok (2). Jeżeli wektor g_2 jest ortogonalny do $g_1 = e_1$, to przyjmujemy $e_2 = g_2$. Jeśli g_2 nie jest ortogonalny do $g_1 = e_1$, to zastępujemy g_2 przez jego składową ortogonalną do e_1, czyli bierzemy

$$e_2 = g_2 - \frac{\left\langle g_2, e_1 \right\rangle}{\|e_1\|^2} \cdot e_1.$$

Tak określony wektor e_2 jest niezerowy, bo wektory g_2 i e_1 nie były współliniowe.
Krok ogólny. Załóżmy, że wektory $e_1, e_2, \ldots, e_k$ są już wybrane i to tak, że

(i) są one parami ortogonalne,

(ii) $Lin\{e_1, e_2, \ldots, e_k\} = Lin\{g_1, g_2, \ldots, g_k\}$.

Zauważmy, że warunki te są spełnione dla $k = 1$ i $k = 2$.
Wektor e_{k+1} tworzymy z wektora g_{k+1} przez odjęcie od g_{k+1} składowych równoległych do $e_1, e_2, \ldots, e_k$, tak więc

$$e_{k+1} = g_{k+1} - \frac{\left\langle g_{k+1}, e_1 \right\rangle}{\|e_1\|^2} \cdot e_1 - \frac{\left\langle g_{k+1}, e_2 \right\rangle}{\|e_2\|^2} \cdot e_2 - \ldots - \frac{\left\langle g_{k+1}, e_k \right\rangle}{\|e_k\|^2} \cdot e_k.$$

Pokażemy, że tak utworzony wektor jest odpowiedni, tzn. że spełnia następujące warunki:

(1) $e_{k+1} \neq \overrightarrow{O}$,

(2) $\left\langle e_{k+1}, e_j \right\rangle = 0$, dla $j = 1, 2, \ldots, k$,

(3) $Lin\{e_1, e_2, \ldots, e_{k+1}\} = Lin\{g_1, g_2, \ldots, g_{k+1}\}$.

Pierwszy warunek wynika z faktu, że $g_{k+1} \notin Lin\{e_1, e_2, \ldots, e_k\}$.
Drugą własność wyprowadzimy korzystając z liniowości iloczynu skalarnego.

$$\left\langle e_{k+1}, e_j \right\rangle = \left\langle g_{k+1} - \sum_{i=1}^{k} \frac{\left\langle g_{k+1}, e_i \right\rangle}{< e_i, e_i >} \cdot e_i, e_j \right\rangle =$$

$$= \left\langle g_{k+1}, e_j \right\rangle - \sum_{i=1}^{k} \frac{\left\langle g_{k+1}, e_i \right\rangle}{\|e_i\|^2} \cdot \left\langle e_i, e_j \right\rangle =$$

$$= \left\langle g_{k+1}, e_j \right\rangle - \frac{\left\langle g_{k+1}, e_j \right\rangle}{\|e_j\|^2} \cdot \left\langle e_j, e_j \right\rangle =$$

$$= \left\langle g_{k+1}, e_j \right\rangle - \frac{\left\langle g_{k+1}, e_j \right\rangle}{\|e_j\|^2} \cdot \|e_j\| = 0.$$

Trzecia własność wynika z założenia indukcyjnego, tzn.

$$Lin\{g_1, g_2, \ldots, g_k, g_{k+1}\} = Lin\{e_1, e_2, \ldots, e_k, g_{k+1}\}.$$

Z określenia wektora e_{k+1} możemy wnioskować, że przestrzeń generowana przez wektory $\{e_1, e_2, \ldots, e_k, g_{k+1}\}$ to ta sama przestrzeń co ta definiowana przez wektory $\{e_1, e_2, \ldots, e_k, e_{k+1}\}$, tak więc

$$Lin\{e_1, e_2, \ldots, e_k, g_{k+1}\} = Lin\{e_1, e_2, \ldots, e_k, e_{k+1}\}.$$

Z własności trzeciej wynika, że procedurę tę możemy kontynuować, aż g_n zamienimy na e_n. Wtedy baza $\{e_1, e_2, \ldots, e_n\}$ będzie ortogonalna. □

Uwaga 23.2.5.

(1) *Każdy niezerowy wektor v można unormować, tzn. pomnożyć przez taką liczbę $\lambda > 0$, by wektor λv był jednostkowy, $\|\lambda v\| = 1$. Mamy bowiem $1 = \|\lambda v\| = \lambda\|v\|$, czyli $\lambda = \frac{1}{\|v\|}$. Tak więc po unormowaniu wektora v otrzymujemy jednostkowy wektor $\frac{1}{\|v\|} \cdot v$ lub krótko $\frac{v}{\|v\|}$.*

(2) *Bazę ortogonalną $e_1, e_2, \ldots, e_n$ można przekształcić w bazę ortonormalną, poprzez unormowanie układu wektorów $e_1, e_2, \ldots, e_n$. Otrzymujemy wówczas układ*

$$\frac{e_1}{\|e_1\|}, \frac{e_2}{\|e_2\|}, \ldots, \frac{e_n}{\|e_n\|}.$$

Wniosek 23.2.6. *W każdej n-wymiarowej przestrzeni euklidesowej istnieje baza ortonormalna.*

Wniosek 23.2.7. *Każdy ortogonalny układ wektorów z E można rozszerzyć do bazy ortogonalnej w E. Dodatkowo skoro każdą bazę ortogonalną można przekształcić do bazy ortonormalnej (Uwaga 23.2.5), więc każdy ortogonalny układ wektorów z E można rozszerzyć do bazy ortonormalnej.*

Dowód. Niech E będzie przestrzenią euklidesową, natomiast $e_1, e_2, \ldots, e_k$ będzie układem niezerowych parami ortogonalnych wektorów w E. Jak wiemy z Lematu 23.2.2 układ ten jest liniowo niezależny, zatem można go w dowolny sposób uzupełnić do bazy $e_1, e_2, \ldots, e_k, g_{k+1}, \ldots, g_n$, nie starając się aby była to baza ortogonalna. Możemy następnie zastosować do tego układu metodę modyfikacji Grama - Schmidta, przy czym pierwszych k wektorów pozostawimy bez zmian. Dostaniemy więc układ ortogonalny $e_1, e_2, \ldots, e_k, g'_{k+1}, \ldots, g'_n$ i to jest żądane uzupełnienie. Podobnie można postąpić w przypadku ortonormalnym, z tym że wektory $g'_{k+1}, \ldots, g'_n$ trzeba będzie wtedy dodatkowo znormalizować. □

23.3 Współrzędne wektora oraz iloczyn skalarny w bazie ortonormalnej.

Fakt 23.3.1 (współrzędne wektora w bazie ortonormalnej). *Jeśli E jest przestrzenią euklidesową, zaś $e_1, e_2, \ldots, e_n$ jej bazą ortonormalną, wówczas dla każdego $v \in E$ mamy*

$$v = \langle v, e_1\rangle e_1 + \langle v, e_2\rangle e_2 + \ldots + \langle v, e_n\rangle e_n.$$

Dowód. Niech $v = t_1e_1 + t_2e_2 + \ldots + t_ne_n$. Obliczmy t_i.

$$\Big\langle v, e_i \Big\rangle = \Big\langle t_1e_1 + t_2e_2 + \ldots + t_ne_n, e_i \Big\rangle = t_1\Big\langle e_1, e_i\Big\rangle + t_2\Big\langle e_2, e_i\Big\rangle + \ldots + t_n\Big\langle e_n, e_i\Big\rangle =$$

$$= t_i\Big\langle e_i, e_i\Big\rangle = t_i\|e_i\|^2 = t_i.$$

Ostatecznie $t_i = \Big\langle v, e_i\Big\rangle$. □

Pokażemy teraz ciekawą własność iloczynu skalarnego. Sprawdzimy jak zachowuje się on we współrzędnych w bazie ortonormalnej.

Niech E będzie przestrzenią euklidesową z iloczynem skalarnym. Natomiast $B = \{e_1, e_2, \ldots, e_n\}$ niech będzie ortonormalną bazą tej przestrzeni. Dla wektorów v, w niech $v = [v_1, v_2, \ldots, v_n], w = [w_1, w_2, \ldots, w_n]$ będą ich współrzędnymi w bazie B.

Fakt 23.3.2. *Iloczyn skalarny wektorów v, w wyraża się wzorem*

$$\Big\langle v, w\Big\rangle = v_1w_1 + v_2w_2 + \ldots + v_nw_n.$$

Dowód. Skoro $v = (v_1, v_2, \ldots, v_n), w = (w_1, w_2, \ldots, w_n)$, to możemy te wektory zapisać w postaci $v = v_1e_1 + v_2e_2 + \ldots + v_ne_n, w = w_1e_1 + w_2e_2 + \ldots + m_ne_n$. Iloczyn skalarny tych wektorów przedstawia się zatem wzorem

$$\Big\langle v, w\Big\rangle = \Big\langle \sum_{i=1}^n v_ie_i, \sum_{j=1}^n w_ie_i \Big\rangle = \sum_{i=1}^n\sum_{j=1}^n v_iw_j\Big\langle e_i, e_j\Big\rangle.$$

Zauważmy, że dla $i \neq j$ iloczyn skalarny $\Big\langle e_i, e_j\Big\rangle$ jest równy 0 oraz, że $\|e_i\| = 1$ dla każdego i. Tak więc

$$\Big\langle v, w\Big\rangle = \sum_{i=1}^n v_iw_i\Big\langle e_i, e_i\Big\rangle = \sum_{i=1}^n v_iw_i\|e_i\|^2 = \sum_{i=1}^n v_iw_i.$$

□

Uwaga 23.3.3. *We współrzędnych względem ortonormalnej bazy, iloczyn skalarny wyraża się takim samym wzorem jak standardowy iloczyn skalarny w R^n.*

23.4 Rzut ortogonalny na podprzestrzeń.

W tym podrozdziale pokażemy do czego może przydać się baza ortogonalna. Wprowadzimy tutaj uogólnienie rzutu prostopadłego, znanego z przestrzeni R^3, na przypadek dowolnej przestrzeni euklidesowej.

Definicja 23.4.1 (wektor ortogonalny do podprzestrzeni). Niech E będzie przestrzenią euklidesową, zaś U jej podprzestrzenią wektorową. Wektor $v \in E$ nazywamy ortogonalnym do podprzestrzeni U, jeśli jest ortogonalny do każdego wektora z tej podprzestrzeni, tzn. dla każdego $u \in U$ zachodzi $\Big\langle v, u\Big\rangle = 0$. Oznaczymy to symbolem $v \perp U$.

Lemat 23.4.2 (charakteryzacja ortogonalności do podprzestrzeni). *Niech układ $u_1, u_2, \ldots, u_k$ będzie dowolną bazą podprzestrzeni U. Wówczas v jest ortogonalny do U wtedy i tylko wtedy gdy v jest ortogonalny do każdego z wektorów u_i.*

Dowód. Zauważmy, że dowód jednej z implikacji jest oczywisty, bo jeśli v jest ortogonalny do każdego wektora z U, to w szczególności jest ortogonalny do wszystkich u_i. Sprawdźmy zatem przeciwną implikację, czyli że wektor v jest ortogonalny do każdego wektora z podprzestrzeni U, o ile jest ortogonalny do wektorów u_i. Musimy zatem sprawdzić, że dla każdego $u \in U$ zachodzi

$$\Big\langle v, u \Big\rangle = 0.$$

Zapiszmy $u = t_1u_1 + t_2u_2 + \ldots + t_ku_k$. Obliczamy, że

$$\Big\langle v, u \Big\rangle = \Big\langle v, t_1u_1 + t_2u_2 + \ldots + t_ku_k \Big\rangle = t_1\Big\langle v, u_1 \Big\rangle + t_2\Big\langle v, u_2 \Big\rangle + \ldots + t_k\Big\langle v, u_k \Big\rangle =$$

$$= t_1 \cdot 0 + t_2 \cdot 0 + \ldots + t_k \cdot 0 = 0,$$

czyli $v \perp u$. Zatem $v \perp U$. □

Lemat 23.4.3. *Każdy wektor v można, i to jednoznacznie, rozłożyć w sumę*

$$v = v_U + v_\perp,$$

gdzie $v_U \in U$ jest składową z podprzestrzeni U, natomiast $(v_\perp) \perp U$ to składowa ortogonalna do U.

Dowód.
Istnienie rozkładu.
Zauważmy, że podprzestrzeń U sama jest przestrzenią euklidesową, z tym samym iloczynem skalarnym, co przestrzeń E. Niech $e_1, e_2, \ldots, e_k$ będzie bazą ortogonalną w U, natomiast $e_{k+1}, \ldots, e_n$ będzie uzupełnieniem tej bazy do bazy ortogonalnej całej przestrzeni E. Rozkładamy v względem tej bazy.

$$v = t_1e_1 + t_2e_2 + \ldots + t_ke_k + t_{k+1}e_{k+1} + \ldots + t_ne_n.$$

Przyjmijmy, że $v_U = t_1e_1 + t_2e_2 + \ldots + t_ke_k$ jest składową z podprzestrzeni U, natomiast $v_\perp = t_{k+1}e_{k+1} + \ldots + t_ne_n$ będzie składową ortogonalną do podprzestrzeni U. Mamy $v = v_U + v_\perp$, $v_U \in U$. Pozostaje sprawdzić, że $v_\perp$ jest rzeczywiście ortogonalny do U. W tym celu wystarczy sprawdzić ortogonalność tego wektora do wektorów $e_1, e_2, \ldots, e_k$. Niech $m \in \{1, 2, \ldots, k\}$, wiemy, że baza $e_1, e_2, \ldots, e_n$ jest bazą ortogonalną dlatego mamy

$$\Big\langle v_\perp, e_m \Big\rangle = \Big\langle t_{k+1}e_{k+1} + \ldots + t_ne_n, e_m \Big\rangle = t_{k+1}\Big\langle e_{k+1}, e_m \Big\rangle + \ldots + t_n\Big\langle e_n, e_m \Big\rangle =$$

$$= t_{k+1} \cdot 0 + \ldots + t_n \cdot 0 = 0.$$

Zatem $v_\perp$ jest rzeczywiście ortogonalny do U, co oznacza, że rozkład $v = v_\perp + v_U$ jest właściwy.

Jednoznaczność rozkładu.
Niech $v = v_\perp + v_U$ i $v = v'_\perp + v'_U$, gdzie $v_U \in U, v'_U \in U$ to składowe ortogonalne do U, tzn. $v_\perp \perp U, v'_\perp \perp U$, będą dwoma rozkładami wektora v na składowe. Z równości $v_\perp + v_U = v'_\perp + v'_U$ wynika, że

$$v_\perp - v'_\perp = v'_U - v_U.$$

Zdefinujmy wektor z taki, że $z = v_\perp - v'_\perp = v'_U - v_U$. Zauważmy, że skoro $z = v'_U - v_U$, to wektor z należy do podprzestrzeni U, natomiast z tego, że $z = v_\perp - v'_\perp$ wynika, że z jest ortogonalny do U (łatwo sprawdzić, że różnica dwóch wektorów ortogonalnych do U jest wektorem ortogonalnym do U). Te dwa warunki są spełnione tylko wtedy gdy $z = 0$, a z tego mamy już $v_U = v'_U$ oraz $v_\perp = v'_\perp$. Tak więc nie istnieją dwa różne rozkłady wektora v. □

Składowe $v_U, v_\perp$ będziemy odpowiednio nazywać składową równoległą i składową prostopadłą do U.

Definicja 23.4.4 (rzut ortogonalny). Odwzorowanie, które dowolnemu wektorowi $v \in E$ przyporządkowuje jego składową v_U równoległą do U, nazywamy rzutem ortogonalnym na podprzestrzeń U. Rzut na podprzestrzeń U oznaczamy symbolem $P_U : E \to U$.

W poprzednim podrozdziale określiliśmy współrzędne wektora w bazie ortonormalnej. Na podstawie tego określenia możemy wyznaczyć wzór na rzut ortogonalny na podprzestrzeń.

Wniosek 23.4.5. *Niech U będzie podprzestrzenią przestrzeni euklidesowej E i niech $e_1, e_2, \dots, e_n$ będzie bazą ortonormalną w U. Wtedy rzut ortogonalny na podprzestrzeń U wyraża się wzorem*

$$P_U(v) = \sum_{i=1}^{k} \Big\langle v, e_i \Big\rangle e_i.$$

Dowód. Dla bazy $e_1, e_2, \dots, e_k$ w U niech jej uzupełnieniem do bazy ortonormalnej w E będą wektory $e_{k+1}, \dots, e_n$. Wówczas na podstawie Faktu 23.3.1, mamy rozkład

$$v = \Big\langle v, e_1 \Big\rangle e_1 + \Big\langle v, e_2 \Big\rangle e_2 + \ldots + \Big\langle v, e_k \Big\rangle e_k + \Big\langle v, e_{k+1} \Big\rangle e_{k+1} + \ldots + \Big\langle v, e_n \Big\rangle e_n.$$

Składowa równoległa do podprzestrzeni U to

$$v_U = \Big\langle v, e_1 \Big\rangle + \Big\langle v, e_2 \Big\rangle + \ldots + \Big\langle v, e_k \Big\rangle e_k.$$

Ostatecznie więc

$$P_U(v) = \sum_{i=1}^{k} \Big\langle v, e_i \Big\rangle e_i.$$

□

Ćwiczenie. Z powyższego wzoru, poprzez bezpośrednie sprawdzenie addytywności i jednorodności, można przekonać się, że rzut P_U jest przekształceniem liniowym. Pozostawimy to do wykonania czytelnikowi.

23.5 Dopełnienie ortogonalne podprzestrzeni.

Dla danej podprzestrzeni U przestrzeni E, możemy określić pewien zbiór wektorów, o bardzo charakterystycznych własnościach. Zbiór ten to dopełnienie ortogonalne podprzestrzeni. Kilka jego własności podamy i udowodnimy poniżej.

Definicja 23.5.1 (**przekształcenie ortogonalne**)**.** Dopełnieniem ortogonalnym, oznaczanym $U^{\perp}$, podprzestrzeni U w przestrzeni euklidesowej U nazywamy zbiór

$$U^{\perp} = \{v \in E : v \perp U\}.$$

Własność 23.5.2. *Zbiór $U^{\perp}$ jest podprzestrzenią wektorową w E.*

Dowód. Musimy pokazać, że zbiór wektorów $U^{\perp}$ jest zamknięty ze względu na dodawanie i mnożenie przez skalar. Rozważmy dowolne dwa wektory v_1, v_2 ze zbioru $U^{\perp}$. Zauważmy, że skoro są one prostopadłe do U to znaczy, że dla każdego $u \in U$ mamy

$$\Big\langle v_1, u \Big\rangle = \Big\langle v_2, u \Big\rangle = 0.$$

Z własności iloczynu skalarnego mamy $\Big\langle v_1 + v_2, u \Big\rangle = \Big\langle v_1, u \Big\rangle + \Big\langle v_2, u \Big\rangle = 0 + 0 = 0$, zatem suma dowolnych dwóch wektorów z $U^{\perp}$ jest ortogonalna do U, więc należy do $U^{\perp}$.
Podobnie otrzymamy, że dla dowolnego wektora $v \in U^{\perp}$ i dowolnego wektora U mamy

$$\Big\langle tv, u \Big\rangle = t\Big\langle v, u \Big\rangle = t \cdot 0 = 0,$$

zatem $tv \in U^{\perp}$. Zbiór $U^{\perp}$ jest zatem podprzestrzenią wektorową w przestrzeni E. □

Lemat 23.5.3. *Niech $e_1, e_2, \ldots, e_k$ będzie bazą ortogonalną podprzestrzeni U, natomiast $e_{k+1}, \ldots, e_n$ jej uzupełnieniem do bazy ortogonalnej w przestrzeni E. Wówczas*

$$U^{\perp} = Lin\{e_{k+1}, \ldots, e_n\}$$

i co więcej układ $e_{k+1}, \ldots, e_n$ jest bazą $U^{\perp}$.

Dowód. Na początek pokażemy, że $Lin\{e_{k+1}, \ldots, e_n\}$ jest podprzestrzenią $U^{\perp}$. Weźmy dowolny wektor v z przestrzeni $Lin\{e_{k+1}, \ldots, e_n\}$ czyli $v = s_{k+1}e_{k+1} + \ldots + s_n e_n$. Chcemy pokazać, że wektor ten należy do przestrzeni $U^{\perp}$. Obliczmy $\Big\langle v, e_i \Big\rangle$, dla $i = 1, 2, \ldots, k$. Wiemy, że baza $e_1, e_2, \ldots, e_n$ jest bazą ortogonalną E, zatem dla dowolnych e_i, e_j takich, że $i \neq j$ mamy $\Big\langle e_i, e_j \Big\rangle = 0$ zatem

$$\Big\langle v, e_i \Big\rangle = \Big\langle s_{k+1}e_{k+1} + \ldots + s_n e_n, e_i \Big\rangle =$$

$$= s_{k+1}\Big\langle e_{k+1}, e_i \Big\rangle + \ldots + s_n\Big\langle e_n, e_i \Big\rangle = s_{k+1} \cdot 0 + \ldots + s_n \cdot 0 = 0.$$

Tak więc wektor v jest ortogonalny do dowolnego wektora bazowego przestrzeni U, jest więc ortogonalny do dowolnego wektora będącego kombinacją liniową tych wektorów. Wektor v jest prostopadły do całej podprzestrzeni U, należy zatem do $U^{\perp}$. Pokazaliśmy zatem, że

$$Lin\{e_{k+1}, \ldots, e_n\} \subset U^{\perp}.$$

Chcemy teraz pokazać przeciwne zawieranie. Weźmy dowolny wektor $v \in U^{\perp}$ i niech

$$v = t_1e_1 + t_2e_2 + \ldots + t_ke_k + t_{k+1}e_{k+1} + \ldots + t_ne_n.$$

Wiemy, że dla $i = 1, 2, \ldots k$ zachodzi $\Big\langle v, e_i \Big\rangle = 0$ bo wektor v jest ortogonalny do podprzestrzeni U, więc w szczególności do każdego z jej wektorów bazowych. Mamy więc

$$\Big\langle t_1e_1 + t_2e_2 + \ldots + t_ne_n, e_i \Big\rangle = 0,$$

co po przekształceniu możemy zapisać jako

$$t_1\Big\langle e_1, e_i\Big\rangle + t_2\Big\langle e_2, e_i\Big\rangle + \ldots + t_i\Big\langle e_i, e_i\Big\rangle + \ldots + t_n\Big\langle e_n, e_i\Big\rangle = 0.$$

Dowolne dwa wektory bazy $e_1, e_2, \ldots, e_n$ są ortogonalne, więc $\Big\langle e_i, e_j\Big\rangle = 0$ dla $i \neq j$ i $\Big\langle e_i, e_i\Big\rangle \neq 0$. Otrzymujemy więc

$$t_1\Big\langle e_1, e_i\Big\rangle + t_2\Big\langle e_2, e_i\Big\rangle + \ldots + t_i\Big\langle e_i, e_i\Big\rangle + \ldots + t_n\Big\langle e_n, e_i\Big\rangle = t_i\Big\langle e_i, e_i\Big\rangle = 0,$$

a skoro $\Big\langle e_i, e_i\Big\rangle \neq 0$, to $t_i = 0$.
Tak więc

$$v = t_1e_1 + t_2e_2 + \ldots + t_ne_n = 0 \cdot e_1 + 0 \cdot e_2 + \ldots + 0 \cdot e_k + t_{k+1}e_{k+1} + \ldots + t_ne_n =$$

$$= t_{k+1}e_{k+1} + \ldots + t_ne_n.$$

Wektor v jest więc kombinacją liniową wektorów $e_{k+1}, \ldots, e_n$ a zatem należy do podprzestrzeni $Lin\{e_{k+1}, \ldots, e_n\}$. Ostatecznie otrzymujemy, że

$$U^{\perp} \subset \{e_{k+1}, \ldots, e_n\}.$$

Ostatecznie, mając oba zawierania możemy stwierdzić, że

$$U^{\perp} = \{e_{k+1}, \ldots, e_n\}.$$

Zatem skoro układ $e_{k+1}, \ldots, e_n$ jest liniowo niezależny i generuje $U^{\perp}$ jest więc bazą $U^{\perp}$. □

Wniosek 23.5.4. *Dla podprzestrzeni U i jej dopełnienia ortogonalnego $U^{\perp}$ w przestrzeni E mamy*

$$dimU + dimU^{\perp} = dimE.$$

Lemat 23.4.3 możemy przy użyciu pojęcia dopełnienia ortogonalnego podprzestrzeni wyrazić w postaci następującego wniosku.

Wniosek 23.5.5. *Każdy wektor z przestrzeni E jednoznacznie przedstawia się jako suma*

$$v = v_U + v_{\perp},$$

gdzie $v_U \in U$, $v_{\perp} \in U^{\perp}$.

23.6 Izomorfizm przestrzeni euklidesowych.

Zauważmy, że dla jednej przestrzeni wektorowej możemy wprowadzić kilka iloczynów skalarnych, otrzymując w ten sposób różne przestrzenie euklidesowe. W tym podrozdziale pokażemy, że wszystkie takie przestrzenie są izomorficzne, co więcej pokażemy, że dowolne przestrzenie euklidesowe tego samego wymiaru są izomorficzne.

Definicja 23.6.1 (przestrzenie izomorficzne). Dwie przestrzenie E_1, E_2 ze swoimi iloczynami skalarnymi $\langle\ ,\ \rangle_{E_1}$ i $\langle\ ,\ \rangle_{E_2}$ nazywamy izomorficznymi, jeśli istnieje liniowy izomorfizm $\varphi : E_1 \to E_2$, który zachowuje iloczyn skalarny, tzn. dla dowolnych wektorów $v, w \in E_1$ mamy

$$\Big\langle \varphi(v), \varphi(w) \Big\rangle_{E_2} = \Big\langle v, w \Big\rangle_{E_1}.$$

Własność 23.6.2. *Każda n-wymiarowa przestrzeń euklidesowa E jest izomorficzna z przestrzenią E^n, czyli z R^n ze standardowym iloczynem skalarnym.*

Dowód. Niech $e_1, e_2, \ldots, e_n$ będzie bazą ortonormalną w E. Sprawdźmy, że odwzorowanie $\varphi : E \to E^n$ określone wzorem

$$\varphi(x_1e_1 + x_2e_2 + \ldots + x_ne_n) = (x_1, x_2, \ldots, x_n),$$

jest izomorfizmem przestrzeni euklidesowych.

✍ Sprawdzenie, że jest to izomorfizm przestrzeni liniowych pozostawimy do wykonania czytelnikowi. Sprawdźmy tylko, że spełniony jest warunek zachowania iloczynu skalarnego podany w Definicji 23.6.1. Niech $v = v_1e_1 + v_2e_2 + \ldots + v_ne_n, w = w_1e_1 + w_2e_2 + \ldots + w_ne_n$. Wtedy, korzystając z Faktu 23.3.2 wyliczającego iloczyn skalarny w bazie ortonormalnej mamy

$$\Big\langle v, w \Big\rangle_E = \sum_{i=1}^{n} v_iw_i = \Big\langle (v_1, v_2, \ldots, v_n), (w_1, w_2, \ldots, w_n) \Big\rangle_{E_n}.$$

Zatem φ jest dobrze określonym izomorfizmem przestrzeni euklidesowych. □

23.7 Przekształcenia ortogonalne w przestrzeniach E^n.

W tym podrozdziale zajmiemy się pewnym rodzajem przekształceń przestrzeni E^n - przekształceniami ortogonalnymi. Podamy definicję takich przekształceń oraz ich najważniejsze własności. Określimy również macierz przekształceń ortogonalnych.

Definicja 23.7.1 (przekształcenia ortogonalne). Przekształcenie liniowe $\varphi : E^n \to E^n$ nazywamy ortogonalnym, jeśli dla każdej pary wektorów v, w z przestrzeni E^n mamy

$$\Big\langle v, w \Big\rangle = \Big\langle \varphi(v), \varphi(w) \Big\rangle.$$

Obserwacja 23.7.2. *Przekształcenie liniowe $\varphi : E^n \to E^n$ jest ortogonalne wtedy i tylko wtedy gdy obraz $\varphi(e_1), \varphi(e_2), \ldots, \varphi(e_n)$ dowolnej bazy ortonormalnej $e_1, e_2, \ldots, e_n$ jest bazą ortonormalną.*

Dowód. Pokażemy na początek, że dla przekształcenia ortogonalnego φ obraz $\varphi(e_1), \varphi(e_2), \ldots, \varphi(e_n)$ dowolnej bazy ortonormalnej $e_1, e_2, \ldots, e_n$ jest też bazą. Musimy zatem pokazać, że każdy z wektorów $\varphi(e_i)$ jest jednostkowy i że wektory te są do siebie ortogonalne. Weźmy dowolny wektor e_i spośród $e_1, e_2, \ldots, e_n$. Chcemy aby jego obraz przez przekształcenie φ był wektorem jednostkowym. Sprawdźmy jego długość

$$\|\varphi(e_i)\| = \sqrt{\Big\langle \varphi(e_i), \varphi(e_i) \Big\rangle}.$$

Przekształcenie φ jest ortogonalne a więc zachowuje iloczyn skalarny, stąd

$$\|\varphi(e_i)\| = \sqrt{\Big\langle \varphi(e_i), \varphi(e_i) \Big\rangle} = \sqrt{\Big\langle e_i, e_i \Big\rangle} = \|e_i\|.$$

Skoro baza $e_1, e_2, \ldots, e_n$ jest ortonormalna, to $\|e_i\| = 1$. Zatem również

$$\|\varphi(e_i)\| = \|e_i\| = 1.$$

Pokażmy jeszcze ortogonalność wektorów $\varphi(i), \varphi(j)$ dla $i \neq j$. Jak wiemy wektory są ortogonalne, gdy ich iloczyn skalarny jest równy zero. Pamiętając o tym, że przekształcenie φ jest przekształceniem ortogonalnym, dla $i \neq j$ mamy

$$\Big\langle \varphi(e_i), \varphi(e_j) \Big\rangle = \Big\langle e_i, e_j \Big\rangle = 0.$$

Układ wektorów $\varphi(e_1), \varphi(e_2), \ldots, \varphi(e_n)$ jest zatem ortonormalny.
Pokażemy teraz implikację przeciwną. Załóżmy, że $\varphi(e_1), \varphi(e_2), \ldots, \varphi(e_n)$ jest bazą ortonormalną. Chcemy pokazać, że wówczas φ jest przekształceniem ortogonalnym, czyli

$$\Big\langle v, w \Big\rangle = \Big\langle \varphi(v), \varphi(w) \Big\rangle.$$

Lewą stronę tej równości możemy zapisać w postaci

$$L = \Big\langle v, w \Big\rangle = \Big\langle \sum_{i=1}^{n} v_i e_i, \sum_{j=1}^{n} w_j e_j \Big\rangle.$$

Z własności iloczynu skalarnego możemy lewą stronę przekształcić do postaci

$$L = \sum_{i=1}^{n} \sum_{j=1}^{n} v_i w_i \Big\langle e_i, e_j \Big\rangle,$$

natomiast z ortogonormalności bazy $e_1, e_2, \ldots, e_n$ mamy

$$L = \sum_{i=1}^{n} v_i w_i.$$

Podobnie przekształcamy prawą stronę

$$P = \Big\langle \varphi(v), \varphi(w) \Big\rangle = \Big\langle \varphi(\sum_{i=1}^{n} v_i e_i), \varphi(\sum_{j=1}^{n} w_j e_j) \Big\rangle.$$

Przekształcenie φ jest przekształceniem liniowym, więc jest jednorodne i addytywne, stąd

$$P = \Big\langle \sum_{i=1}^{n} v_i\varphi(e_i), \sum_{j=1}^{n} w_i\varphi(e_j)\Big\rangle.$$

Podobnie jak poprzednio otrzymujemy więc

$$P = \sum_{i=1}^{n}\sum_{j=1}^{n} v_i w_i \Big\langle \varphi(e_i), \varphi(e_j)\Big\rangle.$$

Z ortogonormalności bazy $\varphi(e_1), \varphi(e_2), \ldots, \varphi(e_n)$ mamy ostatecznie

$$P = \sum_{i=1}^{n} v_i w_i.$$

Tak więc po prawej i lewej stronie równości otrzymujemy po przekształceniach to samo wyrażenie, co kończy dowód. □

➢ Macierz przekształcenia ortogonalnego w standardowej bazie

Niech $A = [a_{ij}]$ będzie macierzą przekształcenia ortogonalnego $\varphi : E^n \to E^n$ w standardowej bazie. Ponieważ obrazy wektorów bazowych tworzą bazę ortonormalną, to kolumny $A_1, A_2, \ldots, A_n$ macierzy A są wektorami jednostkowymi i wzajemnie prostopadłymi. Mamy wtedy

$$(\star)\begin{cases} \Big\langle A_j, A_j\Big\rangle = \sum_{i=1}^{n} a_{ij}a_{ij} = \sum_{i=1}^{n} a_{ij}^2 = 1, \\ \Big\langle A_j, A_k\Big\rangle = \sum_{i=1}^{n} a_{ij}a_{ik} = 0, \quad dla \quad j \neq k. \end{cases}$$

Zależność $(\star)$ można krótko zapisać jedną równością macierzową:

$$A^T \cdot A = I.$$

✍ **Ćwiczenie.** Udowodnienie powyższej równości pozostawimy jako ćwiczenie dla czytelnika.

Wniosek 23.7.3. *Macierz A jest macierzą przekształcenia ortogonalnego wtedy i tylko wtedy gdy $A^T \cdot A = I$. Macierz taką nazywamy* ***macierzą ortogonalną***.

Rozdział 24

Lista uogólnień

W wielu miejscach tego skryptu, przy omawianiu jakiegoś pojęcia w przestrzeniach R^n wskazywaliśmy na analogię do sytuacji, jaką omawialiśmy w przestrzeniach R^2 i R^3. Podobnie w dowolnych przestrzeniach wektorowych analogie te odnosiły się do przestrzeni R^n. Przeniesienie rozmaitych faktów dotyczących przestrzeni R^2 i R^3 najpierw na przestrzenie R^n, a następnie na abstrakcyjne przestrzenie wektorowe było możliwe dzięki użyciu pojęć, takich jak współrzędne wektorów w bazie i macierze przekształceń w bazie. Rozdział ten poświęcimy na podanie listy dalszych takich uogólnień, które omówimy tu bez dowodów.

24.1 Macierz przejścia do nowej bazy.

Definicja 14.3.1 określa, jak wygląda macierz przejścia do nowego układu współrzędnych w przestrzeni R^3. Tworzą ją współrzędne wersorów nowego układu wyrażone w starym układzie. Podobnie powstaje macierz przejścia do nowej bazy w dowolnej przestrzeni wektorowej. Niech $B = \{v_1, v_2, \ldots, v_k\}, C = \{w_1, w_2, \ldots, w_k\}$ będą bazami przestrzeni wektorowej V. Macierz przejścia $F = F_{BC}$ od bazy B do bazy C to macierz, której kolumny są utworzone ze współrzędnych wektorów $w_1, w_2, \ldots, w_k$ w wyjściowej bazie $v_1, v_2, \ldots, v_k$. Dokładniej, jeśli

$$w_i = \begin{pmatrix} f_{1i} \\ f_{2i} \\ \vdots \\ f_{ki} \end{pmatrix}_B ,$$

to macierz przejścia F ma postać

$$\boxed{F = F_{BC} = \begin{pmatrix} f_{11} & f_{12} & \ldots & f_{1k} \\ f_{21} & f_{22} & \ldots & f_{2k} \\ \vdots & \vdots & \ddots & \vdots \\ f_{k1} & f_{k2} & \ldots & f_{kk} \end{pmatrix} = ([w_1]_B, [w_2]_B, \ldots, [w_k]_B) .}$$

Fakt 24.1.1. *Macierz przejścia F jest macierzą kwadratową. Ponadto macierz jest nieosobliwa, tzn. jej wyznacznik jest niezerowy* $\det F \neq 0$.

24.2 Współrzędne wektora w nowej bazie.

W podrozdziale 14.4 pokazaliśmy jak znaleźć współrzędne danego wektora w nowym układzie w przestrzeni R^3. Rozważmy teraz przestrzeń wektorową V, w której bazami są B i C. Dla dowolnego wektora $u \in V$ niech $[u]_B$, $[u]_C$ będą współrzędnymi (kolumnowymi) wektora u w bazach B i C. Macierzą przejścia od bazy B do bazy C jest $F = F_{BC}$. Wówczas

$$\boxed{[u]_B = F \cdot [u]_C, \quad [u]_C = F^{-1} \cdot [u]_B.}$$

24.3 Macierz przekształcenia w nowej bazie.

Twierdzenie 14.5.2 określa jak wygląda macierz przekształcenia liniowego przestrzeni R^3 w nowym układzie współrzędnych. Podobną zależność spełnia macierz przekształcenia dowolnej przestrzeni wektorowej. Dla przekształcenia $T : V \to V$ oraz baz B, C tej przestrzeni, macierz przejścia to $F = F_{BC}$. Jeżeli $[T]_B$ jest macierzą przekształcenia T w bazie B, to macierz tego przekształcenia w bazie C wyraża się wzorem

$$\boxed{[T]_C = F^{-1} \cdot [T]_B \cdot F.}$$

24.4 Odwracalność przekształceń $T : V \to V$.

Przy okazji omawiania przekształceń liniowych przestrzeni R^3 sformułowaliśmy kryterium, które pozwala łatwo sprawdzić, czy dane przekształcenie jest odwracalne czy nie za pomocą wyznacznika jego macierzy. Mówi o tym Wniosek 7.3.2. Określimy teraz analogiczne kryterium dla przekształceń liniowych $T : V \to V$ dowolnych przestrzeni wektorowych.

Fakt 24.4.1. *Jeżeli B jest dowolną bazą przestrzeni V, wówczas przekształcenie T jest odwracalne wtedy i tylko wtedy gdy wyznacznik* $\det([T]_B)$ *jest niezerowy.*

Możemy uogólnić powyższy fakt na przypadek przekształcenia dowolnych przestrzeni. Dla przestrzeni V z bazą B i przestrzeni W z bazą C, oraz dowolnego przekształcenia $T : V \to W$, przekształcenie to jest odwracalne wtedy i tylko wtedy gdy wyznacznik $\det([T]_{BC})$ jest niezerowy.

Oprócz kryterium odwracalności możemy również określić macierz przekształcenia odwrotnego.

Fakt 24.4.2. *Dla przekształcenia odwracalnego $T : V \to V$ zachodzi*

$$[T^{-1}]_B = ([T]_B)^{-1}.$$

24.5 Macierz złożenia przekształceń liniowych.

Niech $T_1, T_2 : V \to V$ będą przekształceniami liniowymi, natomiast B niech będzie dowolną bazą przestrzeni V. Macierz złożenia tych przekształceń $T_1 \circ T_2$ wyraża się w bazie B wzorem

$$\boxed{[T_1 \circ T_2]_B = [T_1]_B \cdot [T_2]_B.}$$

24.6 Wartości własne i wielomian charakterystyczny.

Dla przekształceń $T : V \to V$ definicja wektorów i wartości własnych pozostaje taka sama, jak ta wprowadzona dla przekształceń przestrzeni R^3 (Definicja 10.1.1). Również określenie wielomianu charakterystycznego odpowiada temu wprowadzonemu dla przypadku R^3. Możemy zatem sformułować fakt, który pozwala na szybkie znajdowanie wartości własnych danego przekształcenia. Fakt ten dla przekształceń przestrzeni wektorowej przyjmuje postać analogiczną do Wniosku 10.3.2.

Fakt 24.6.1. *Niech V będzie przestrzenią wektorową wymiaru n i niech B będzie bazą V. Liczba r jest wartością własną przekształcenia liniowego $T : V \to V$ wtedy i tylko wtedy gdy r jest pierwiastkiem wielomianu charakterystycznego, przy czym wielomian ten przyjmuje postać*

$$\det([T]_B - t \cdot I_{n\times n}) = \det \begin{pmatrix} a_{11}-t & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22}-t & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn}-t \end{pmatrix}.$$

24.7 Izometrie i macierze ortogonalne w R^n.

Definicje izometrii i przekształceń ortogonalnych, jakie wprowadziliśmy w przypadku przestrzeni R^3, możemy bezpośrednio przenieść na przypadek dowolnej przestrzeni R^n.

Definicja 24.7.1 (izometria i przekształcenie ortogonalne). Przekształcenie liniowe $T : R^n \to R^n$ jest ortogonalne, jeśli zachowuje iloczyn skalarny w R^n. Przekształcenie $T : R^n \to R^n$ jest izometrią, jeśli zachowuje odległość w R^n.

Powyższa definicja przyjmuje postać analogiczną jak Definicje 11.1.1 i 11.2.1, wprowadzone dla przekształceń przestrzeni R^3. Scharakteryzowaliżmy też macierze przekształceń ortogonalnych. Podobnie jak dla przestrzeni R^3 (Lemat 12.1.2), możemy określić macierz ortogonalną.

Definicja 24.7.2 (macierz ortogonalna). Macierz A rozmiaru $n \times n$ jest ortogonalna jeśli $A \cdot A^T = I$, lub równoważnie, jeśli kolumny macierzy A tworzą układ wzajemnie prostopadłych jednostkowych wektorów z R^n.

Dla przestrzeni R^3 pokazaliśmy, że zbiory izometrii liniowych i przekształceń ortogonalnych są takie same. Podobny fakt zachodzi również dla przekształceń przestrzeni R^n.

Fakt 24.7.3. *Przekształcenie liniowe $T : R^n \to R^n$ jest izometrią wtedy i tylko wtedy gdy jest ortogonalne. T jest ortogonalne wtedy i tylko wtedy gdy macierz $m(T)$ przekształcenia T jest macierzą ortogonalną.*

24.8 Diagonalizacja macierzy symetrycznych.

Podrozdział 13.3 poświęciliśmy diagonalizacji macierzy symetrycznych rozmiaru 3×3. Podobnie można diagonalizować macierze symetryczne rozmiaru $n \times n$.

Fakt 24.8.1. *Jeśli A jest macierzą symetryczną rozmiaru $n \times n$, to można ją przedstawić w postaci $A = P \cdot D \cdot P^{-1}$, gdzie D jest macierzą diagonalną rozmiaru $n \times n$, zaś P jest macierzą ortogonalną. Ponadto wyrazy na przekątnej macierzy D są wartościami własnymi macierzy A, zaś macierz P jest utworzona z jednostkowych wektorów własnych dla tych wartości własnych.*

24.9 Formy kwadratowe.

Rozdział 15 poświęciliśmy na zbadanie własności pewnego rodzaju funkcji - form kwadratowych. Podobne pojęcie możemy wprowadzić w przestrzeniach R^n.

Definicja 24.9.1 (forma kwadratowa). Funkcja $\varphi : R^n \to R$ jest formą kwadratową jeśli ma postać

$$\varphi \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij} x_i x_j ,$$

przy czym zakładamy, że $a_{ji} = a_{ij}$ dla $i \neq j$.
Macierz $A = [a_{ij}]_{i,j}$ nazywamy macierzą formy kwadratowej φ.

Uwaga 24.9.2. *Macierz $A = [a_{ij}]_{i,j}$ jest macierzą symetryczną.*

Uwaga 24.9.3. *Wzór określający formę kwadratową zawiera składniki $a_{ij}x_ix_j$ oraz $a_{ji}x_jx_i$, które są równe. Zatem ogólną postać formy kwadratowej można też zapisać jako*

$$\varphi \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \sum_{k=1}^{n} a_{kk} x_k^2 + \sum_{1 \leqslant i < j \leqslant n} 2 a_{ij} x_i x_j .$$

Formę kwadratową możemy również zdefiniować w dowolnej przestrzeni wektorowej.

Definicja 24.9.4. Funkcja $\varphi : V \to R$ jest formą kwadratową jeśli dla pewnej bazy B przestrzeni V ma ona w tej bazie postać

$$\varphi \left(\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}_B \right) = \sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij} x_i x_j .$$

Fakt 24.9.5. *W każdej innej bazie B' w przestrzeni V funkcja φ będzie miała podobną postać, ale z innymi współczynnikami a_{ij}.*

Własność mówiąca o tym, że dana funkcja jest formą kwadratową jest całkowicie niezależna od wyboru bazy. Jest to chyba jedna z najmocniejszych cech tej funkcji.

24.10 Macierz formy kwadratowej w nowej bazie i diagonalizacja formy kwadratowej.

W rozdziale 15 pokazaliśmy, że każdą macierz formy kwadratowej można w pewnym nowym układzie współrzędnych wyrazić w postaci diagonalnej. Podobnie możemy postąpić w przypadku formy kwadratowej w dowolnej przestrzeni wektorowej.

Fakt 24.10.1. *Niech V będzie przestrzenią wektorową, natomiast B, C niech będą bazami tej przestrzeni. Macierzą przejścia z bazy B do bazy C jest $F = F_{BC}$. Niech $\varphi : V \to R$ będzie formą kwadratową o macierzy A względem bazy B, oraz o macierzy A' względem bazy C. Wówczas*

$$A' = F^T \cdot A \cdot F.$$

Wniosek 24.10.2. *Dla każdej formy kwadratowej φ istnieje baza ortonormalna, w której forma ta ma postać diagonalną, tzn. macierz formy φ w tej bazie jest macierzą diagonalną.*

Dowód. Weźmy dowolną bazę ortonormalną B i formę φ, która w tej bazie ma macierz A. Macierz A jako macierz formy kwadratowej jest macierzą symetryczną, więc posiada ortogonalną diagonalizację

$$A = PDP^{-1},$$

gdzie P jest macierzą ortogonalną. Potraktujmy macierz P jako macierz przejścia z bazy B do jakiejś bazy C. Zauważmy, że wówczas macierz C też musi być macierzą ortonormalną. Skoro zatem $A+PDP^{-1}$, to $D = P^{-1}AP$. Macierz P jest ortogonalna więc $P^{-1} = P^T$ a stąd

$$D = P^T AP.$$

Z Faktu 24.10.1 wynika, że macierz diagonalna D jest macierzą formy φ w bazie C. □

Uwaga 24.10.3. *W takiej bazie $e_1, e_2, \ldots, e_n$ forma ma postać*

$$\varphi(x_1e_1 + x_2e_2 + \ldots + x_ne_n) = a_{11}x_1^2 + a_{22}x_2^2 + \ldots + a_{nn}x_n^2.$$

24.11 Dodatnia określoność form kwadratowych i macierzy.

W rozdziale 15 wprowadziliśmy pojęcia formy i macierzy dodatnio określonych. Wspominaliśmy również o możliwościach wykorzystania tych pojęć w analizie wielu zmiennych. Możemy w ten sposób wykorzystać również formy i macierze dodatnio określone w przypadku przestrzeni R^n. Pojęcia te będą miały definicję analogiczną do tej wprowadzonej w przestrzeni R^3.

Definicja 24.11.1. Forma kwadratowa $\varphi : R^n \to R$ jest dodatnio określona, jeśli dla każdego niezerowego wektora X z przestrzeni R^n zachodzi $\varphi(X) > 0$.
Macierz symetryczna rozmiaru $n \times n$ jest dodatnio określona, jeśli jest macierzą dodatnio określonej formy kwadratowej w R^n.

Podstawowe własności oraz sposób rozpoznawania macierzy dodatnio określonych przy użyciu wyznaczników przenosi się bez zmian z przypadku 3×3 na przypadek $n \times n$.

Fakt 24.11.2. *Macierz symetryczna A rozmiaru $n \times n$ jest dodatnio określona wtedy i tylko wtedy gdy wszystkie wartości własne macierzy A są dodatnie a to zachodzi wtedy i tylko wtedy gdy*

$$\det A_1 > 0, \det A_2 > 0, \ldots, \det A_{n-1} > 0, \det A > 0,$$

gdzie A_i jest macierzą rozmiaru $i \times i$ będącą lewą górną częścią macierzy A.

Nie będziemy podawać więcej uogólnień, choć niewątpliwie dałoby się jeszcze wiele takich znaleźć. Jak widać wiele zagadnień z przestrzeni R^3 możemy bardzo łatwo przenieść na przypadek R^n a następnie na dowolne przestrzenie wektorowe, dzięki temu możemy badać własności abstrakcyjnych przestrzeni przy pomocy analogii do przypadku bliższej nam przestrzeni R^3.

Bibliografia

[1] F.Leja, *Geometria analityczna*, PWN, **1961**

[2] T.Huskowski, H.Korczowski, H.Matuszczyk, *Algebra liniowa*, Wydawnictwo Politechniki Wrocławskiej, **1980**

[3] A.Białynicki-Birula, *Algebra liniowa z geometrią*, PWN, **1976**

[4] W.Pogorzelski, *Geometria analityczna w przestrzeni*, Akademicka Spółdzielnia Wydawnicza, **1948**

[5] R.W.Fuller, F.W.Byron, *Matematyka w fizyce klasycznej i kwantowej*, PWN, **1975**