ROZDZIAŁ 2

Oznaczenia i podstawowe fakty z algebry liniowej

2.1 Oznaczenia

W dalszej części używać będziemy następujących oznaczeń i konwencji:

- Element x (punkt) przestrzeni $\mathbb{R}^n$ będziemy nazywać *wektorem* wymiaru n i oznaczać $x = (x_1, ..., x_n) \in \mathbb{R}^n$, gdzie x_j, $j = 1, 2, ..., n$ są współrzędnymi punktu (wektora) x.

- Symbolem
$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}$$
oznaczamy *macierz wymiaru* (*typu*) $m \times n$ o elementach rzeczywistych a_{ij}, $i = 1, 2, ..., m$, $j = 1, 2, ..., n$. Będziemy ją również oznaczać symbolem $[a_{ij}]_{m \times n}$ lub krócej $[a_{ij}]$, jeśli jej wymiar wynika z kontekstu. Jeśli $m = n$, to A nazywamy *macierzą kwadratową stopnia* n.

- Symbolem A^T oznaczamy *transpozycję* macierzy A (*macierz transponowaną* względem A), czyli macierz typu $n \times m$ postaci
$$A^T = \begin{bmatrix} a_{11} & a_{21} & \cdots & a_{m1} \\ a_{12} & a_{22} & \cdots & a_{m2} \\ \vdots & \vdots & \ddots & \vdots \\ a_{1n} & a_{2n} & \cdots & a_{mn} \end{bmatrix}.$$
Piszemy również $A^T = [a_{ji}]_{n \times m}$.

- W zapisie macierzowym wektor $x \in \mathbb{R}^n$ będziemy zazwyczaj identyfikować z macierzą typu $n \times 1$ (*wektor kolumnowy*), tzn. będziemy pisać
$$x = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}.$$
albo $x = [x_1, x_2, ..., x_n]^T$. Wektor x możemy również identyfikować z *wektorem wierszowym*, tzn. $x = [x_1, x_2, ..., x_n]$. Wymiar wektora będzie wynikać z kontekstu. Jeśli nie będzie

prowadzić to do nieporozumień, dla punktów (wektorów) w $\mathbb{R}^2$ lub $\mathbb{R}^3$ będziemy zamiennie używać oznaczeń odpowiednio (x, y) bądź (x, y, z).

- Niech A będzie macierzą rzeczywistą typu $m \times n$. Wektor

$$A^i := [a_{i1}, a_{i2}, ..., a_{in}] \in \mathbb{R}^n$$

nazywamy i-tym *wierszem macierzy* A, $i = 1, 2, ..., m$, zaś wektor

$$A_j := \begin{bmatrix} a_{1j} \\ a_{2j} \\ \vdots \\ a_{mj} \end{bmatrix} \in \mathbb{R}^m$$

– j-tą *kolumną macierzy* A, $j = 1, 2, ..., n$. Jeśli $m = n$, to A nazywamy *macierzą kwadratową stopnia* n. Wymiar macierzy będzie wynikać z kontekstu.

- Dla macierzy kwadratowej $A = [a_{ij}]$ stopnia n wielkość

$$\mathrm{tr}(A) = \sum_{j=1}^{n} a_{ii}$$

nazywamy *śladem* macierzy A.

- Dla macierzy $A = [a_{ij}]$ i liczby α zapis $A + \alpha$ oznacza macierz $B = [b_{ij}]$, gdzie $b_{ij} = a_{ij} + \alpha$ (do każdego elementu macierzy A dodajemy liczbę α).

- Symbolem I_n oznaczamy *macierz jednostkową* stopnia n, czyli

$$I_n := \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{bmatrix}.$$

Macierz jednostkową będziemy również oznaczać symbolem I lub Id, jeśli jej stopień wynika z kontekstu.

- Mówimy, że macierz kwadratowa A jest *odwracalna* jeśli istnieje macierz kwadratowa B taka, że $AB = BA = I$. Wówczas macierz B nazywamy *macierzą odwrotną* do A i zapisujemy $B = A^{-1}$. Oczywiście $(A^{-1})^{-1} = A$.

Ćwiczenie 2.1.1 Pokazać, że dla macierzy odwracalnej A istnieje dokładnie jedna macierz do niej odwrotna.

- Macierz $D = [d_{ij}]$, dla której $d_{ij} = 0$ gdy $i \neq j$ nazywamy *macierzą diagonalną*. Często kwadratową macierz diagonalną będziemy zapisywać w postaci

$$\mathrm{diag}\, d := \begin{bmatrix} d_1 & 0 & \cdots & 0 \\ 0 & d_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & d_n \end{bmatrix},$$

gdzie $d = (d_1, d_2, ...d_n) \in \mathbb{R}^n$.

Ćwiczenie 2.1.2 Pokazać, że iloczyn macierzy diagonalnych jest macierzą diagonalną.

- Macierz kwadratową A nazywamy *macierzą symetryczną* jeśli $A^T = A$.
- Macierz kwadratową $L = [l_{ij}]$, dla której $l_{ij} = 0$ gdy $i < j$ nazywamy *dolna macierzą trójkątną*. Ma ona więc postać

$$L := \begin{bmatrix} l_{11} & 0 & \cdots & 0 \\ l_{21} & l_{22} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ l_{n1} & l_{n2} & \cdots & l_{nn} \end{bmatrix}$$

W podobny sposób definiujemy górną macierz trójkątną.

Ćwiczenie 2.1.3 Pokazać, że iloczyn dolnych (górnych) macierzy trójkątnych jest dolną (górną) macierzą trójkątną.

- Macierze będziemy często zapisywać w *postaci blokowej*, np.

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

pamiętając przy tym, aby bloki występujące w jednym blokowym wierszu miały tyle samo wierszy i bloki występujące w jednej blokowej kolumnie miały tyle samo kolumn. Przykładowo, macierz A typu $m \times n$ możemy przedstawić w postaci blokowej jako blok złożony z jej n kolumn, czyli

$$A = [A_1, A_2, ..., A_n],$$

gdzie $A_j \in \mathbb{R}^m$, $j = 1, 2, ..., n$, albo jako blok złożony z jej m wierszy, czyli

$$\begin{bmatrix} A^1 \\ A^2 \\ \vdots \\ A^m \end{bmatrix},$$

gdzie $A^i \in \mathbb{R}^n$, $i = 1, 2, ..., m$.

- Macierze w zapisie blokowym A i B możemy mnożyć blokowo według tych samych reguł, które obowiązują dla zwykłego mnożenia macierzy, przy czym liczba blokowych kolumn macierzy A musi być równa liczbie blokowych wierszy macierzy B oraz mnożone bloki muszą mieć takie wymiary, aby ich iloczyn był dobrze zdefiniowany. Przykładowo, jeśli

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

i

$$B = \begin{bmatrix} B_{11} & B_{12} & B_{13} \\ B_{21} & B_{22} & B_{23} \end{bmatrix}$$

są macierzami blokowymi, to AB jest macierzą blokową i

$$\begin{aligned} AB &= \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} \begin{bmatrix} B_{11} & B_{12} & B_{13} \\ B_{21} & B_{22} & B_{23} \end{bmatrix} \\ &= \begin{bmatrix} A_{11}B_{11} + A_{12}B_{21} & A_{11}B_{12} + A_{12}B_{22} & A_{11}B_{13} + A_{12}B_{23} \\ A_{21}B_{11} + A_{22}B_{21} & A_{21}B_{12} + A_{22}B_{22} & A_{21}B_{13} + A_{22}B_{23} \end{bmatrix}, \end{aligned}$$

przy czym należy założyć, że macierz A_{ij} ma tyle samo kolumn, co macierz B_{jk} ma wierszy, $i, j = 1, 2,\ k = 1, 2, 3$. Oznacza to, że że wszystkie występujące w tym wzorze iloczyny blokowe są dobrze określone. W szczególności, jeśli macierz A typu $m \times p$ zapiszemy w postaci blokowej

$$A = \begin{bmatrix} A^1 \\ A^2 \\ \vdots \\ A_m \end{bmatrix}$$

, gdzie A^i jest i-tym wierszem macierzy A, $i = 1, 2, ..., p$ zaś macierz B typu $p \times n$ zapiszemy w postaci blokowej $B = [B_1, B_2, ...B_n]$, gdzie B_j jest j-tą kolumną macierzy B, to iloczyn AB (macierz typu $m \times n$) możemy zapisać w postaci

$$AB = \begin{bmatrix} A^1B_1 & A^1B_2 & \cdots & A^1B_n \\ A^2B_1 & A^2B_2 & \cdots & A^2B_n \\ \vdots & \vdots & \ddots & \vdots \\ A^mB_1 & A^mB_2 & \cdots & A^mB_n \end{bmatrix};$$

- $a \geq 0$ oznacza, że wszystkie współrzędne wektora $a \in \mathbb{R}^n$ są nieujemne;
- $\alpha_+ = \max\{0, \alpha\}$, $\alpha_- = \max\{0, -\alpha\}$ oznaczają odpowiednio cześć dodatnią i część ujemną liczby $\alpha \in \mathbb{R}$;
- $a_+ = ((a_1)_+, ..., (a_n)_+)$, czyli część dodatnią wektora $a \in \mathbb{R}^n$ otrzymujemy wyznaczając części dodatnie jego współrzędnych a_j, $j = 1, 2, ..., n$;
- $\mathbb{R}^n_+ = \{x \in \mathbb{R}^n : x \geq 0\}$ oznacza *ortant nieujemny*;
- $a > 0$ oznacza, że wszystkie współrzędne wektora $a \in \mathbb{R}^n$ są dodatnie;
- $\mathbb{R}^n_{++} = \{x \in \mathbb{R}^n : x > 0\}$ oznacza *ortant dodatni*;
- Symbolem 0 będziemy oznaczać zarówno liczbę zero jak i macierz zerową (o wszystkich elementach równych zero), w szczególności wektor zerowy.
- E – macierz odpowiedniego wymiaru o wszystkich elementach równych 1;
- $e := (1, 1, ..., 1)$ – wektor odpowiedniego wymiaru o wszystkich współrzędnych równych 1;
- $e_j := (0, ..., 0, 1, 0, ..., 0)$ – wersor jednostkowy odpowiedniego wymiaru z 1 na j-tej współrzędnej;
- Oznaczając j-tą kolumnę macierzy A typu $m \times n$ symbolem A_i, $j = 1, 2, ..., n$, macierz A możemy zapisać blokowo $A = [A_1, A_2, ..., A_n]$. Wówczas dla $x \in \mathbb{R}^n$ mamy $Ax = x_1A_1 + x_2A_2 + ... + x_nA_n$, czyli wektor Ax jest kombinacją liniową kolumn macierzy A, przy czym współrzędne wektora x są współczynnikami tej kombinacji liniowej. W szczególności $Ae_j = A_j$.
- Oznaczając i-ty wiersz macierzy A typu $m \times n$ symbolem A_i, $i = 1, 2, ..., m$, macierz A możemy zapisać blokowo

$$A = \begin{bmatrix} A^1 \\ A^2 \\ \vdots \\ A^m \end{bmatrix}.$$

Wówczas dla wektora wierszowego $y = [y_1, y_2, ..., y_m] \in \mathbb{R}^m$ mamy $yA = y_1A^1 + y_2A^2 + ... + y_mA^m$, czyli wektor wierszowy yA jest kombinacją liniową wierszy macierzy A, przy czym współrzędne wektora y są współczynnikami tej kombinacji liniowej. W szczególności $e_i^T A = A^i$.

- $B(x, r) = \{y \in \mathbb{R}^n : \|y - x\| \leq r\}$ oznacza *kulę domkniętą* o środku w punkcie x i promieniu r;
- $S(x, r) = \{y \in \mathbb{R}^n : \|y - x\| = r\}$ oznacza *sferę* o środku w punkcie x i promieniu r;

2.2 Podstawowe pojęcia i fakty z algebry liniowej

W tym ustępie przypomnimy podstawowe pojęcia i fakty z algebry liniowej, których będziemy używać w dalszej części. Fakty te podajemy najczęściej bez dowodów, gdyż można je znaleźć w większości podręczników do algebry liniowej. Pojęcia i fakty podane niżej zawężamy do przestrzeni euklidesowej, chociaż większość z nich odnosi się do dowolnej przestrzeni liniowej.

2.2.1 Podprzestrzeń liniowa

Definicja 2.2.1 Niech $a_1, a_2, ..., a_m \in \mathbb{R}^n$. Wektor

$$b = \alpha_1 a_1 + \alpha_2 a_2 + ... + \alpha_m a_m,$$

gdzie $\alpha_i \in \mathbb{R}$, $i = 1, 2, ..., m$, nazywamy *kombinacją liniową* wektorów $a_1, a_2, ..., a_m$.

Niech $A = [a_1, a_2, ..., a_m]$ będzie macierzą typu $n \times m$ zapisaną w postaci blokowej, gdzie $a_i \in \mathbb{R}^n$, $i = 1, 2, ..., m$. Wówczas kombinację liniową wektorów $a_1, a_2, ..., a_m$ możemy przedstawić w postaci Aa, gdzie współrzędne a_i wektora $a = (\alpha_1, \alpha_2, ..., \alpha_m)$ są współczynnikami w tej kombinacji liniowej. Zatem wektor kolumnowy Aa jest kombinacją liniową kolumn macierzy A. Podobnie wektor wierszowy βA, gdzie $\beta = (\beta_1, \beta_2, ..., \beta_m)$ jest kombinacją liniową wierszy macierzy A.

Definicja 2.2.2 Mówimy, że zbiór $X \subseteq \mathbb{R}^n$ jest *podprzestrzenią liniową* jeśli

$$\forall_{a,b \in X} \forall_{\alpha,\beta \in \mathbb{R}} \quad \alpha a + \beta b \in X.$$

Definicja 2.2.3 Mówimy, że układ wektorów $a_1, a_2, ..., a_m \in X \subseteq \mathbb{R}^n$ *rozpina* podprzestrzeń X jeśli dowolny wektor $a \in X$ można przedstawić w postaci *kombinacji liniowej* wektorów $a_1, a_2, ..., a_m \in X$, czyli istnieją stałe $\alpha_1, \alpha_2, ..., \alpha_m \in \mathbb{R}$ takie, że

$$a = \alpha_1 a_1 + \alpha_2 a_2 + ... + \alpha_m a_m.$$

Mówimy wówczas, że X jest *rozpięta* przez układ wektorów $a_1, a_2, ..., a_m$ i piszemy

$$X = \mathrm{span}\{a_1, a_2, ..., a_m\}.$$

2.2.2 Liniowa niezależność wektorów

Definicja 2.2.4 Układ wektorów $\{a_1, a_2, ..., a_m\} \subseteq \mathbb{R}^n$ nazywamy *liniowo niezależnym*, jeśli

$$\forall_{\alpha_1,\alpha_2,...,\alpha_m \in \mathbb{R}} \quad (\alpha_1 a_1 + \alpha_2 a_2 + ... + \alpha_m a_m = 0 \Longrightarrow \alpha_1 = \alpha_2 = ... = \alpha_m = 0).$$

Układ ten jest liniowo zależny, jeśli nie jest liniowo niezależny, czyli

$$\exists_{\alpha_1,\alpha_2,...,\alpha_m \in \mathbb{R}} \quad (\alpha_1 a_1 + \alpha_2 a_2 + ... + \alpha_m a_m = 0 \text{ i } \alpha_i \neq 0 \text{ dla przynajmniej jednego } i = 1, 2, ..., m).$$

Definicja 2.2.5 Niech $X \subseteq \mathbb{R}^n$ będzie podprzestrzenią liniową i niech układ $\{a_1, a_2, ..., a_m\} \subseteq X$ będzie liniowo niezależny. Jeśli dowolny wektor $x \in X$ można jednoznacznie przedstawić w postaci

$$x = \alpha_1 a_1 + \alpha_2 a_2 + ... + \alpha_m a_m,$$

gdzie $\alpha_1, \alpha_2, ..., \alpha_m \in \mathbb{R}$, to układ ten nazywamy *bazą* tej podprzestrzeni, a liczbę m jej *wymiarem*, co zapisujemy $m = \dim(X)$.

Można pokazać, że wymiar podprzestrzeni nie zależy od przyjętej bazy. W szczególności, jeśli $X = \mathbb{R}^n$, to dowolna jej baza ma n elementów (wymiar przestrzeni $\mathbb{R}^n$ wynosi n).

Definicja 2.2.6 Liczbę liniowo niezależnych wierszy (bądź kolumn) macierzy A nazywamy *rzędem macierzy* i oznaczmy symbolem $\mathrm{rank}(A)$.

W niektórych podręcznikach podaje się równoważną definicję rzędu macierzy A jako maksymalny stopień niezerowego minora macierzy A.

Ćwiczenie 2.2.7 Pokazać, że dla niezerowych wektorów kolumnowych $u, v \in \mathbb{R}^n$ macierz uv^T typu $n \times n$ jest macierzą rzędu pierwszego.

Ćwiczenie 2.2.8 Niech A będzie macierzą typu $m \times n$, zaś B macierzą nieosobliwą typu $m \times m$. Pokazać, że $\mathrm{rank}(AB) = \mathrm{rank}(A)$.

Ćwiczenie 2.2.9 Pokazać, że dla liniowo niezależnych wektorów (kolumnowych) $u, v \in \mathbb{R}^n$ i dla niezerowych stałych $\alpha, \beta \in \mathbb{R}$ macierz $\alpha uu^T + \beta vv^T$ typu $n \times n$ jest macierzą rzędu drugiego.

Definicja 2.2.10 Mówimy, że macierz A typu $m \times n$ ma *pełny rząd kolumnowy* (*pełny rząd wierszowy*), jeśli układ jej kolumn (wierszy) jest liniowo niezależny.

2.2.3 Układ równań liniowych

Poniższa definicja wyznacznika macierzy kwadratowej A stopnia n jest równoważna powszechnie znanym definicjom Laplace'a (rekurencyjna) i Lebniza jako sumy $n!$ składników pomnożonych przez ± 1, z których każdy jest iloczynem n elementów tej macierzy wyznaczonych przez permutację n-elementową (składniki te mnożymy przez $+1$ bądź -1 w zależności od tzw. parzystości permutacji).

Definicja 2.2.11 *Wyznacznik* macierzy kwadratowej jest funkcją określoną na macierzach typu $n \times n$ spełniającą warunki:

(i) Wyznacznik macierzy jednostkowej I wynosi 1;

(ii) Zamiana dwóch wierszy macierzy zmienia znak wyznacznika;

(iii) Pomnożenie wiersza macierzy przez liczbę skutkuje pomnożeniem jej wyznacznika przez tę liczbę;

(iv) Dodanie do wiersza macierzy wielokrotności innego wiersza nie zmienia jej wyznacznika.

Wyznacznik macierzy kwadratowej A oznaczamy symbolem $\det(A)$.

Można pokazać, że istnieje dokładnie jedna funkcja spełniająca warunki (i)-(iv) powyższej definicji. Wyznacznik macierzy kwadratowej można zdefiniować równoważnie zastępując w powyższej definicji wiersze macierz jej kolumnami.

Definicja 2.2.12 Macierz kwadratową o niezerowym wyznaczniku nazywamy macierzą *nieosobliwą*.

Twierdzenie 2.2.13 *Niech A będzie macierzą kwadratową. Jeśli macierz odwrotna do macierzy A istnieje, to jest ona określona jednoznacznie. Ponadto następujące warunki są równoważne:*

1. *Macierz A jest odwracalna;*
2. *Macierz A jest pełnego rzędu kolumnowego;*
3. *Macierz A jest pełnego rzędu wierszowego;*
4. *Wyznacznik macierzy A jest niezerowy (macierz A jest nieosobliwa);*
5. *Dla dowolnego $b \in \mathbb{R}^n$ układ równań $Ax = b$ posiada dokładnie jedno rozwiązanie.*

Twierdzenie 2.2.14 *Niech macierz A będzie typu $m \times n$. Wówczas A jest pełnego rzędu kolumnowego wtedy i tylko wtedy, gdy jedynym rozwiązaniem układu równań $Ax = 0$ jest rozwiązanie zerowe.*

Ćwiczenie 2.2.15 Udowodnić twierdzenie 2.2.14.

Ćwiczenie 2.2.16 Pokazać, że macierz trójkątna jest niosobliwa wtedy i tylko wtedy, gdy wszystkie elementy na jej głównej przekątnej są niezerowe.

Ćwiczenie 2.2.17 Niech $A := [a_{ij}]_{n\times n}$ będzie nieosobliwą górną (dolną) macierzą trójkątną. Pokazać, że macierz $B = [b_{ij}]_{n\times n}$ odwrotna do A jest górną (dolną) macierzą trójkątną. Ponadto jeśli $a_{ii} > 0$ (< 0), to $b_{ii} > 0$ (< 0), $i = 1, 2, ..., n$.

Definicja 2.2.18 Podzbiór $X \subseteq \mathbb{R}^n$ nazywamy *podprzestrzenią afiniczną* przestrzeni $\mathbb{R}^n$, jeśli

$$x, y \in X \text{ oraz } \alpha \in \mathbb{R} \Longrightarrow (1-\alpha)x + \alpha y \in X.$$

Ćwiczenie 2.2.19 Niech A będzie macierzą typu $m \times n$ i $b \in \mathbb{R}^m$. Pokazać, że zbiór rozwiązań układu równań liniowych $Ax = b$ jest podprzestrzenią afiniczną.

Twierdzenie 2.2.20 *Niech $X \subseteq \mathbb{R}^n$ będzie niepusty i niech $a \in X$. Podzbiór X jest podprzestrzenią afiniczną wtedy i tylko wtedy $X - a$ jest podprzestrzenią liniową.*

Ćwiczenie 2.2.21 Udowodnić twierdzenie 2.2.20.

Ćwiczenie 2.2.22 Pokazać, że dla macierzy A typu $m\times n$ zbiór rozwiązań układu jednorodnego $Ax = 0$ jest podprzestrzenią liniową.

Twierdzenie 2.2.23 (Kronecker–Capelli) *Niech A będzie macierzą typu $m \times n$ i $b \in \mathbb{R}^m$. Układ równań $Ax = b$ posiada rozwiązanie $x \in \mathbb{R}^n$ wtedy i tylko wtedy, gdy* $\operatorname{rank}([A, b]) = \operatorname{rank}(A)$, *przy czym:*

(a) *jeśli* $\operatorname{rank}([A, b]) = \operatorname{rank}(A) = n$, *to układ posiada dokładnie jedno rozwiązanie;*

(b) *jeśli* $\operatorname{rank}([A, b]) = \operatorname{rank}(A) = r < n$, *to układ posiada nieskończeniewiele rozwiązań zależnych od $n - r$ parametrów.*

Ćwiczenie 2.2.24 Niech A będzie macierzą typu $m \times n$ i $b \in \mathbb{R}^m$. Pokazać, że jeśli A jest macierzą pełnego rzędu wierszowego, to zbiór rozwiązań układu równań $Ax = b$ jest jest niepusty i jest podprzestrzenią afiniczną. Jaka jest postać dowolnego rozwiązania tego układu i jaki jest wymiar tej podprzestrzeni.

2.2.4 Iloczyn skalarny i norma wektora

Definicja 2.2.25 Funkcję $\langle \cdot, \cdot \rangle : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ nazywamy *iloczynem skalarnym* jeśli spełnia ona następujące warunki (aksjomaty):

(i) $\langle x, y \rangle = \langle y, x \rangle$ dla dowolnych $x, y \in \mathbb{R}^n$;

(ii) $\langle x + y, z \rangle = \langle x, z \rangle + \langle y, z \rangle$ dla dowolnych $x, y, z \in \mathbb{R}^n$;

(iii) $\langle \alpha x, y \rangle = \alpha \langle x, y \rangle$ dla dowolnych $x, y \in \mathbb{R}^n$ i $\alpha \in \mathbb{R}$;

(iv) $\langle x, x \rangle \geq 0$ dla dowolnego $x \in \mathbb{R}^n$, przy czym równość zachodzi wtedy i tylko wtedy, gdy $x = 0$.

Definicja 2.2.26 Funkcja $\| \cdot \| : \mathbb{R}^n \to \mathbb{R}$ nazywa się *normą*, jeśli spełnia ona nastepujące warunki (aksjomaty):

(i) $\|x + y\| \leq \|x\| + \|y\|$ dla dowolnych $x, y, z \in \mathbb{R}^n$ (nierówność trójkąta);

(ii) $\|\alpha x\| = |\alpha| \cdot \|x\|$ dla dowolnych $x \in \mathbb{R}^n$ i $\alpha \in \mathbb{R}$ (dodatnia jednorodność);

(iii) $\|x\| \geq 0$ dla dowolnego $x \in \mathbb{R}^n$, przy czym równość zachodzi wtedy i tylko wtedy, gdy $x = 0$.

Uwaga 2.2.27 Dla ustalonego iloczynu skalarnego $\langle \cdot, \cdot \rangle$ funkcja $\| \cdot \|$ określona równością $\|x\| = \sqrt{\langle x, x \rangle}$ spełnia aksjomaty normy. Mówimy, że jest to *norma indukowana* przez iloczyn skalarny $\langle \cdot, \cdot \rangle$. Pokdreślmy jednak, że nie każda norma jest indukowana przez iloczyn skalarny.

Iloczyn skalarny można określić na wiele sposobów, na przykład

$$\langle x, y \rangle = x^T y = \sum_{i=1}^{n} x_i y_i, \tag{2.1}$$

gdzie $x, y \in \mathbb{R}^n$. Tak zdefiniowaną funkcję nazywamy *standardowym iloczynem skalarnym*. Norma przez niego idukowana ma postać

$$\|x\| = \sqrt{x^T x} = \sqrt{\sum_{i=1}^{n} x_i^2} \tag{2.2}$$

i nazywa się *normą euklidesową*. Dla oznaczenia normy euklidesowej będziemy używać również symbolu $\| \cdot \|_2$ szczególnie w sytuacji, gdy wraz z normą euklidesową będziemy rozważać inne normy. Innym przykładem iloczynu skalarnego jest

$$\langle x, y \rangle_A = x^T A y, \tag{2.3}$$

gdzie A jest macierzą określoną dodatnio, czyli macierzą symetryczną spełniającą $x^T A x > 0$ dla dowolnego $x \neq 0$. Norma indukowana przez ten iloczyn skalarny ma więc postać

$$\|x\|_A = \sqrt{x^T A x}. \tag{2.4}$$

W szczególności, gdy $A = I$ otrzymujemy standardowy iloczyn skalarny i normę euklidesową.

Ćwiczenie 2.2.28 Pokazać, że funkcja $\langle \cdot, \cdot \rangle_A$ określona równościa (2.3) spełnia aksjomaty iloczynu skalarnego.

W przestrzeni $\mathbb{C}^n$ używa się ogólniejszej definicji iloczynu skalarnego, w której warunek (i) w definicji 2.2.25 przyjmuje postać $\langle x, y\rangle = \overline{\langle y, x\rangle}$, gdzie $\bar{\alpha}$ oznacza sprzężenie liczby $\alpha \in \mathbb{C}$. W konsekwencji, $\langle x, \alpha y\rangle = \bar{\alpha}\langle x, y\rangle$, $\alpha \in \mathbb{C}$. Przykładem iloczynu skalarnego określonego w $\mathbb{C}^n$ jest $\langle x, y\rangle = x^T\bar{y}$, $x, y \in \mathbb{C}^n$.

Powtórzmy, że iloczyn skalarny można zdefiniować na wiele sposobów. Podobnie normę można zdefiniować na wiele sposobów i nie musi być ona indukowana przez iloczyn skalarny. Jeśli wyraźnie nie zaznaczymy inaczej, to w dalszych rozważaniach będziemy ograniczać się do standardowego iloczynu skalarnego w $\mathbb{R}^n$ i do indukowanej przezeń normy euklidesowej. Zazwyczą normę tę oznacza się symbolem $\|\cdot\|_2$ lub $\|\cdot\|$, o ile nie prowadzi to do nieporozumień. Inne ważne normy w $\mathbb{R}^n$, to norma $\|\cdot\|_p$ zwana również normą l_p, gdzie $p \in [1, +\infty]$ zdefiniowana równościa

$$\|x\|_p := (\sum_{j=1}^{n} |x_j|^p)^{\frac{1}{p}}$$

dla $p \in [1, +\infty)$ i

$$\|x\|_\infty := \max_j |x_j| \,.$$

W szczególności norma $\|\cdot\|_1$ zdefiniowana jest równością

$$\|x\|_1 := \sum_{j=1}^{n} |x_j| \,.$$

Norma $\|\cdot\|_p$ jest indukowana przez iloczyn skalarny wtedy i tylko wtedy gdy $p = 2$. W dalszej części, o ile nie będzie to prowadzić do nieporozumień, symbolem $\|\cdot\|$ będziemy oznaczać normę euklidesową.

W uczeniu maszynowym używana jest również norma $\|x\|_0$ równa liczbie niezerowych współrzędnych wektora $x \in \mathbb{R}^n$, zwana również normą l_0. Należy jednak podkreślić, że nie jest ona normą w sensie definicji 2.2.26. Używana jest ona w szczególności w sytuacji, gdy mamy wyznaczyć wektor x mający możliwie mało niezerowych współrzędnych, spełniający 'w przybliżeniu' układ równań liniowych $Ax = b$.

Ćwiczenie 2.2.29 Narysować w $\mathbb{R}^2$ kule jednostkowe $B(0, 1)$ dla norm l_p, gdzie $p = 1, \frac{3}{2}, 2, 3, \infty$.

Jeśli nie będzie powiedziane inaczej, będziemy zakładać, że wszystkie rozpatrywane macierze mają elementy rzeczywiste.

Definicja 2.2.30 Mówimy, że układ wektorów $\{a_1, a_2, ..., a_m\} \subseteq \mathbb{R}^n$, w której określony jest iloczyn skalarny $\langle\cdot,\cdot\rangle$, jest *ortogonalny*, jeśli

$$\langle a_i, a_j\rangle = \begin{cases} 1, & \text{jeśli } i = j, \\ 0, & \text{jeśli } i \neq j. \end{cases}$$

Jeśli ponadto a_i, $i = 1, 2, ..., m$, są unormowane, czyli $\|a_i\| = 1$, $i = 1, 2, ..., m$, to układ ten nazywamy *ortonormalnym*.

Najprostszym przykładem układu ortogonalnego w przestrzeni $\mathbb{R}^n$, w której zdefiniowano standardowy iloczyn skalarny, jest układ wersorów $\{e_j : j = 1, 2, ..., n\}$.

Twierdzenie 2.2.31 *Ortogonalny układ wektorów jest liniowo niezależny.*

Dowód. Niech $\{a_1, a_2, ..., a_m\} \subseteq \mathbb{R}^n$ będzie układem ortogonalnym. Niech α_i, $i = 1, 2, ..., m$, będą stałymi takimi, że
$$\alpha_1 a_1 + \alpha_2 a_2 + ... + \alpha_m a_m = 0.$$
Z aksjomatów iloczynu skalarnego dla dowolnego $i = 1, 2, ..., m$ otrzymujemy
$$0 = \langle \alpha_1 a_1 + \alpha_2 a_2 + ... + \alpha_m a_m, a_i\rangle = \alpha_1\langle a_1, a_i\rangle + \alpha_2\langle a_2, a_i\rangle + ... + \alpha_m\langle a_m, a_i\rangle = \alpha_i,$$
co oznacza, że układ $\{a_1, a_2, ..., a_m\}$ jest liniowo niezależny. ∎

Twierdzenie 2.2.32 (nierówność Cauchy'ego–Schwarza) *Niech w $\mathbb{R}^n$ określony będzie iloczyn skalarny $\langle\cdot,\cdot\rangle$ i indukowana przezeń norma $\|\cdot\|$. Wówczas dla dowolnych $x, y \in \mathbb{R}^n$ zachodzi nierówność*
$$-\|x\| \cdot \|y\| \le \langle x, y\rangle \le \|x\| \cdot \|y\|, \tag{2.5}$$
przy czym:

(i) *w pierwszej nierówności zachodzi równość wtedy i tylko wtedy, gdy x i y są ujemnie liniowo zależne ($x = \alpha y$ lub $y = \alpha x$ dla pewnego $\alpha < 0$),*

(ii) *w drugiej nierówności zachodzi równość wtedy i tylko wtedy, gdy x i y są dodatnio liniowo zależne ($x = \alpha y$ lub $y = \alpha x$ dla pewnego $\alpha > 0$).*

Dowód. Jeśli $x = 0$ lub $y = 0$, twierdzenie jest oczywiste. Niech więc $x, y \neq 0$. Dla dowolnego $\lambda \in \mathbb{R}$ mamy
$$0 \le \|x - \lambda y\|^2 = \langle x - \lambda y, x - \lambda y\rangle = \|x\|^2 - 2\lambda\langle x, y\rangle + \lambda^2\|y\|^2. \tag{2.6}$$
Ponieważ ostatnie wyrażenie w (2.6) jest funkcją kwadratową zmiennej λ,
$$f(\lambda) = \|y\|^2\lambda^2 - 2\langle x, y\rangle\lambda + \|x\|^2,$$
więc
$$\Delta := 4\langle x, y\rangle^2 - 4\|x\|^2 \cdot \|y\|^2 \le 0,$$
z czego wynika pierwsza część twierdzenia. Funkcja ta osiąga minimum
$$f_{\min} = \frac{-4\langle x, y\rangle^2 + 4\|x\|^2 \cdot \|y\|^2}{4\|y\|^2} = \frac{\|x\|^2 \cdot \|y\|^2 - \langle x, y\rangle^2}{\|y\|^2} = 0$$
dokładnie wtedy, gdy $\lambda = \frac{\langle x,y\rangle}{\|y\|^2}$. Z drugiej strony, $f(\lambda) = \|x - \lambda y\|^2 = 0$ wtedy i tylko wtedy, gdy $x = \lambda y$, czyli x i y są liniowo zależne. W przypadku dodatniej liniowej zależności, $\lambda > 0$ i druga nierówność w (2.5) jest równością, zaś w przypadku ujemnej liniowej zależności, $\lambda < 0$ i druga nierówność w (2.5) jest równością. ∎

2.2.5 Określoność macierzy

Definicja 2.2.33 Niech A będzie macierzą symetryczną typu $n \times n$. Mówimy, że A jest *określona nieujemnie*, jeśli
$$\forall_{s\in\mathbb{R}^n} \quad s^T A s \ge 0.$$
Mówimy, że A jest *określona dodatnio*, jeśli
$$\forall_{s\in\mathbb{R}^n, s\neq 0} \quad s^T A s > 0.$$

Ćwiczenie 2.2.34 Pokazać, że dla wektora d o współrzędnych nieujemnych (dodatnich) macierz $\operatorname{diag}(d)$ jest określona nieujemnie (dodatnio).

Definicja 2.2.35 Niech $\{a_1, a_2, ..., a_n\}$ będzie układem wektorów w przestrzeni $\mathbb{R}^n$ z określonym iloczynem skalarnym $\langle \cdot, \cdot \rangle$. Macierz G postaci

$$G := \begin{bmatrix} \langle a_1, a_1 \rangle & \langle a_1, a_2 \rangle & \cdots & \langle a_1, a_n \rangle \\ \langle a_2, a_1 \rangle & \langle a_2, a_2 \rangle & \cdots & \langle a_2, a_n \rangle \\ \vdots & \vdots & \ddots & \vdots \\ \langle a_n, a_1 \rangle & \langle a_n, a_2 \rangle & \cdots & \langle a_n, a_n \rangle \end{bmatrix}$$

nazywa się *macierzą Grama* układu wektorów $\{a_1, a_2, ..., a_n\}$.

Uwaga 2.2.36 Jeśli w przestrzeni $\mathbb{R}^n$ określony jest standardowy iloczyn skalarny, to macierz Grama układu kolumn macierzy A ma postać $G = A^T A$, zaś macierz Grama układu wierszy macierzy A ma postać $G = AA^T$. W tym celu wystarczy macierz A zapisać w postaci blokowej $A = [A_1, A_2, ..., A_n]$ bądź $A = \begin{bmatrix} A^1 \\ A^2 \\ \vdots \\ A^m \end{bmatrix}$ i zastosować reguły mnożenia blokowego macierzy.

Twierdzenie 2.2.37 *Niech w przestrzeni $\mathbb{R}^n$ określony będzie standardowy iloczyn skalarny i niech A będzie macierzą typu $m \times n$. Wówczas macierz Grama układu kolumn macierzy A:*

(i) *jest określona nieujemnie;*

(ii) *jest określona dodatnio wtedy i tylko wtedy, gdy A ma pełny rząd kolumnowy.*

Dowód. Dla dowolnego $s \in \mathbb{R}^n$ mamy

$$s^T A^T A s = (As)^T (As) = \|As\|^2 \geq 0. \tag{2.7}$$

(i) Z równości (2.7) wynika, że macierz Grama kolumn macierzy A jest określona nieujemnie.
(ii) Równoważność wynika z równości (2.7) i z twierdzenia 2.2.14. ■

Następujące twierdzenie podajemy bez dowodu.

Twierdzenie 2.2.38 (Schur) *Jeśli $A = [a_{ij}]_{n \times n}$ i $B = [b_{ij}]_{n \times n}$ są macierzami określonymi dodatnio (nieujemnie), to macierz $C := [a_{ij} b_{ij}]_{n \times n}$ (zwana produktem Hadamarda macierzy A i B) jest określona dodatnio (nieujemnie).*

2.2.6 Norma macierzy

Definicja 2.2.39 Niech A będzie macierzą typu $m \times n$ i niech w $\mathbb{R}^n$ i $\mathbb{R}^m$ będą określone normy $\| \cdot \|$. Wielkość

$$\|A\| := \sup_{x \neq 0} \frac{\|Ax\|}{\|x\|} \tag{2.8}$$

albo równoważnie

$$\|A\| := \sup_{\|x\|=1} \|Ax\| \tag{2.9}$$

nazywamy *normą macierzy* A. W szczególności, jeśli obie normy po prawej stronie równości (2.8) są normami euklidesowymi, to tak określoną normę nazywamy *normą spektralną* macierzy A i oznaczamy symbolem $\|A\|_2$.

W dalszej części normę euklidesową wektora i indukowaną przez nią normę spektralną macierzy będziemy dla uproszczenia zapisywać bez indeksu 2. Można również zdefiniować inne normy macierzowe, np.

$$\|A\|_p := \sup_{x \neq 0} \frac{\|Ax\|_p}{\|x\|_p}, \tag{2.10}$$

gdzie $p \in [1, +\infty]$. W szczególności norma macierzowa $\|\cdot\|_2$ jest normą spektralną. Można pokazać, że dla macierzy $A = [a_{ij}]$ typu $m \times n$ zachodzi $\|A\|_1 = \max_j \sum_{i=1}^m |a_{ij}|$ oraz $\|A\|_\infty = \max_i \sum_{j=1}^n |a_{ij}|$.

Definicja 2.2.40 Niech A będzie macierzą typu $m \times n$. Wielkość

$$\|A\|_F := \sqrt{\sum_{i,j} |a_{ij}|^2} \tag{2.11}$$

nazywamy *normą Frobeniusa* macierzy A.

Nietrudno zauważyć, że $\|A\|_F^2 = \sum_{j=1}^n \|A_j\|^2 = \sum_{i=1}^m \|A^i\|^2$, czyli kwadrat normy Frobeniusa macierzy A jest sumą kwadratów norm euklidesowych jej kolumn lub wierszy.

Uwaga 2.2.41 Dowolna norma macierzowa zdefiniowana równością (2.10) spełnia wszystkie aksjomaty normy, czyli

1. $\|A + B\| \leq \|A\| + \|B\|$
2. $\|\alpha A\| = |\alpha| \cdot \|A\|$
3. $\|A\| \geq 0$, przy czym równość zachodzi wtedy i tylko wtedy, gdy $A = 0$ (A jest macierzą zerową).

Ćwiczenie 2.2.42 Pokazać, że norma macierzy spełnia aksjomaty normy.

Przykład 2.2.43 Niech $u, v \in \mathbb{R}^n$ będą wektorami unormowanymi. Z nierówności Cauchy'ego–Schwarza wynika, że

$$\begin{aligned} \|uv^T\|^2 &= \sup_{x \neq 0} \frac{\|uv^T x\|^2}{\|x\|^2} = \sup_{x \neq 0} \frac{x^T v u^T u v^T x}{\|x\|^2} = \sup_{x \neq 0} \frac{x^T v v^T x}{\|x\|^2} \\ &= \sup_{x \neq 0} \frac{\|v^T x\|^2}{\|x\|^2} = \sup_{x \neq 0} \frac{\|v\|^2 \cdot \|x\|^2}{\|x\|^2} = \|v\|^2 = 1. \end{aligned}$$

Oznacza, to że norma spektralna macierzy uv^T wynosi 1.

2.2.7 Macierze ortogonalne

Definicja 2.2.44 Niech w $\mathbb{R}^n$ określony będzie standardowy iloczyn skalarny, czyli $\langle x, y \rangle = x^T y$, $x, y \in \mathbb{R}^n$. Macierz kwadratowa U o kolumnach U_i, $i = 1, 2, ..., n$, nazywa się macierzą ortogonalną, jeśli kolumny U_i, $i = 1, 2, ..., n$, są wzajemnie ortogonalne i unormowane, czyli

$$\langle U_i, U_j \rangle = U_i^T U_j = \begin{cases} 1, & \text{gdy } i = j, \\ 0, & \text{gdy } i \neq j, \end{cases}$$

albo równoważnie $U^T U = I$ (macierz Grama kolumn macierzy U jest macierzą jednostkową).

Z twierdzenia 2.2.31 wynika następujący wniosek.

Wniosek 2.2.45 *Kolumny macierzy ortogonalnej U są liniowo niezależne, czyli macierz ortogonalna jest odwracalna.*

Twierdzenie 2.2.46 *Wiersze macierzy ortogonalnej U są wzajemnie ortogonalne, czyli $UU^T = I$.*

Dowód. Ponieważ kolumny U_i, $i = 1, 2, ..., n$, macierzy ortogonalnej są liniowo niezależne, to stanowią one bazę przestrzeni $\mathbb{R}^n$. Zatem dowolny wektor $x \in \mathbb{R}^n$ możemy przedstawić w postaci

$$x = \beta_1 U_1 + ... + \beta_n U_n$$

albo w zapisie macierzowym $x = Ub$, gdzie $b = [\beta_1, ..., \beta_n]^T$. Mnożąc obie strony tej równości skalarnie przez U_j otrzymamy $\beta_j = \langle U_j, x\rangle = U_j^T x$, $j = 1, 2, ..., n$, albo w zapisie macierzowym $b = U^T x$. Zatem

$$x = Ub = UU^T x.$$

Ponieważ x jest dowolnym wektorem, otrzymujemy $UU^T = I$, co oznacza, że wiersze macierzy U są wzajemnie ortogonalne. ■

Ćwiczenie 2.2.47 Pokazać, że dla macierzy ortogonalnej U zachodzi $\|Ux\| = \|x\|$ dla dowolnego $x \in \mathbb{R}^n$. Oznacza to, że przekształcenie ortogonalne nie zmienia długości wektorów.

Ćwiczenie 2.2.48 Pokazać, że dla macierzy ortogonalnej U zachodzi $U^{-1} = U^T$. Oznacza to, że aby odwrócić macierz ortogonalną wystarczy ją transponować.

Przykład 2.2.49 Następujące macierze są ortogonalne:

1. **Macierz obrotu** o kąt α w $\mathbb{R}^2$

$$R(\alpha) := \begin{bmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{bmatrix} \tag{2.12}$$

2. **Macierz obrotu Givensa** w $\mathbb{R}^n$

$$G(i,j,\alpha) := \begin{bmatrix} 1 & \cdots & 0 & \cdots & 0 & \cdots & 0 \\ \vdots & \ddots & \vdots & & \vdots & & \vdots \\ 0 & \cdots & \cos\alpha & \cdots & -\sin\alpha & \cdots & 0 \\ \vdots & & \vdots & \ddots & \vdots & & \vdots \\ 0 & \cdots & \sin\alpha & \cdots & \cos\alpha & \cdots & 0 \\ \vdots & & \vdots & & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & \cdots & 0 & \cdots & 1 \end{bmatrix} \tag{2.13}$$

Jest to macierz obrotu o kąt α w płaszczyźnie określonej przez osie x_i i x_j, $i, j = 1, 2, ..., n$, $i \neq j$. W szczególności dla $n = 3$ macierz

$$G(1,3,\alpha) := \begin{bmatrix} \cos\alpha & 0 & -\sin\alpha \\ 0 & 1 & 0 \\ \sin\alpha & 0 & \cos\alpha \end{bmatrix}$$

jest macierzą obrotu w $\mathbb{R}^3$ w płaszczyźnie $0xz$ albo inaczej wokół osi y.

3. **Macierz permutacji**. Jest to macierz powstała przez permutację wierszy, bądź kolumn macierzy jednostkowej. Niech $p : \{1, 2, ..., n\} \to \{1, 2, ..., n\}$ będzie permutacją n-elementową. Wówczas macierz powstała przez permutację kolumn macierzy I ma postać

$$P_p := \left[e_{p(1)}, e_{p(2)}, ..., e_{p(n)}\right] \tag{2.14}$$

natomiast macierz powstała przez permutację wierszy macierzy I ma postać

$$P_p^T := \begin{bmatrix} e_{p(1)}^T \\ e_{p(2)}^T \\ \vdots \\ e_{p(n)}^T \end{bmatrix} \tag{2.15}$$

4. **Macierz odbicia**, zwana również *macierzą Householdera*

$$R_v = I - \frac{2}{\|v\|^2} vv^T \tag{2.16}$$

gdzie $v \in \mathbb{R}^n$, $v \neq 0$. Jest to macierz odbicia względem hiperpłaszczyzny prostopadłej do wektora v. Istotnie. Zauważmy najpierw, że dla dowolnego wektora $a \in \mathbb{R}^n$ ortogonalnego do v (czyli takiego, że $a^T v = 0$) zachodzi $R_v a = a$ oraz dla dowolnego wektora $b \in \mathbb{R}^n$ współliniowego z v (czyli w postaci $b = \alpha v$, dla pewnego $\alpha \in \mathbb{R}$) mamy $R_v b = -b$. Zapiszmy teraz dowolny wektor $x \in \mathbb{R}^n$ w postaci

$$x = \underbrace{x - \alpha v}_{a} + \underbrace{\alpha v}_{b},$$

gdzie $\alpha = \frac{\langle v, x\rangle}{\|v\|^2}$. Wówczas a jest ortogonalny do v i b jest wpółliniowy z v. Mamy więc $R_v a = a$ oraz $R_v b = -b$. Zatem

$$y := R_v x = R_v(a + b) = R_v a + R_v b = a - b.$$

Widzimy więc, że wektor $y = a - b$ jest odbiciem wektora $x = a + b$, gdyż wektor $x - y = 2b$ jest współliniowy z v oraz $\|x - a\| = \|y - a\|$.

Ćwiczenie 2.2.50 Pokazać, że macierze zdefiniowane w przykładzie 2.2.49 są ortogonalne.

Ćwiczenie 2.2.51 Niech $p = (p(1), p(2), ..., p(n))$ będzie permutacją n-elementową. Pokazać, że dla macierzy permutacji P_p typu $n \times n$ i dla dowolnej macierzy A typu $m \times n$ macierz AP_p jest macierzą powstałą z macierzy A przez permutację jej kolumn, czyli

$$AP_p = [A_{p(1)}, A_{p(2)}, ..., A_{p(n)}].$$

Podobnie, dla dowolnej macierzy B typu $n \times m$ macierz $P_p^T B$ jest macierzą powstałą z macierzy B przez permutację jej wierszy, czyli

$$P_p^T B = \begin{bmatrix} B^{p(1)} \\ B^{p(2)} \\ \vdots \\ B^{p(n)} \end{bmatrix}.$$

Ćwiczenie 2.2.52 Niech $x, y \in \mathbb{R}^n$ będą wektorami unormowanymi i niech

$$v = \begin{cases} x - y, & \text{jeśli } x \neq y, \\ \text{dowolny niezerowy wektor ortogonalny do } x, & \text{jeśli } x = y. \end{cases}$$

Pokazać, że $R_v x = y$ i $R_v y = x$, gdzie macierz R_v jest zdefiniowana równością (2.16).

Wskazówka: Przyjąć $v := x - y$, $a = \frac{1}{2}(x + y)$, $b = x - a$. Wówczas a jest ortogonalny do v (gdyż x i y są unormowane) i b jest współliniowy z v oraz $y = 2a - x = 2a - (a + b) = a - b$. Dalej można skorzystać z argumentacji umieszczonej w przykładzie 2.2.49, p.4. Alternatywnie, wystarczy zauważyć, że $v = 2b$ i skorzystać ze wzoru (2.16).

Ćwiczenie 2.2.53 Pokazać, że iloczyn macierzy ortogonalnych jest macierzą ortogonalną.