Wstęp matematyczny

Leszek Stolarczyk

22 kwietnia 2005

1 Liczby zespolone

Zbiór liczb rzeczywistych oznaczamy przez $\mathbb{R}$, a zbiór liczb zespolonych przez $\mathbb{C}$. Liczba zespolona $z \in \mathbb{C}$ może być zapisana w postaci

$$z = a + bi \ , \quad \text{gdzie} \quad a, b \in \mathbb{R} \ , \tag{1}$$

a i jest tzw. jednostką urojoną, spełniającą warunek

$$i^2 = -1 \ . \tag{2}$$

Potocznie zapisuje się: $i = \sqrt{-1}$.

1. Dla każdej liczby zespolonej z danej w postaci (1) definiuje się liczbę zespoloną sprzężoną względem z:
$$z^* = a - bi \ . \tag{3}$$
Spełnione jest:
$$zz^* = (a + bi)(a - bi) = a^2 + b^2 \ , \tag{4}$$
a więc $zz^* \in \mathbb{R}$ i $zz^* \geq 0$. Oczywiście, $zz^* = 0$ wtedy i tylko wtedy, gdy $z = 0 + 0i = 0$.

2. Moduł liczby zespolonej z definiuje się jako
$$|z| = \sqrt{zz^*} = \sqrt{a^2 + b^2} \ . \tag{5}$$
Z powyższej definicji wynika, że $|z|$ jest liczbą rzeczywistą nieujemną. $|z| = 0$ wtedy i tylko wtedy, gdy $z = 0$.

3. Dla każdej liczby zespolonej z danej w postaci (1) definiuje się
– część rzeczywistą z:
$$\mathrm{Re} z = \frac{1}{2}(z + z^*) = a \ , \tag{6}$$
– część urojoną z:
$$\mathrm{Im} z = \frac{-i}{2}(z - z^*) = b \ . \tag{7}$$
$\mathrm{Re} z$ i $\mathrm{Im} z$ są więc liczbami rzeczywistymi.

4. Działania arytmetyczne w zbiorze liczb zespolonych: dla $z_1 = a + bi$ i $z_2 = c + di$ określa się:
   – sumę:
$$z_1 + z_2 = (a + bi) + (c + di) = (a + c) + (b + d)i \ , \tag{8}$$
   – iloczyn:
$$z_1 z_2 = (a + bi)(c + di) = ac + bci + adi + bdi^2 = (ac - bd) + (bc + ad)i \ . \tag{9}$$
   Oba działania są przemienne. Dla każdej liczby zespolonej $z \neq 0$ można skonstruować odwrotność z^{-1},
$$z^{-1} = \frac{z^*}{zz^*} = \frac{a - bi}{a^2 + b^2} \ , \tag{10}$$
   tak, że zachodzi
$$zz^{-1} = 1 \ . \tag{11}$$
   Działania arytmetyczne w zbiorze liczb zespolonych są więc analogiczne do działań w zbiorze liczb rzeczywistych.

5. Definicja liczby sprzężonej zespolonej (3), zastosowana do sumy i iloczynu liczb zespolonych, daje
$$(z_1 + z_2)^* = z_1^* + z_2^* \ , \tag{12}$$
$$(z_1 z_2)^* = z_1^* z_2^* \ . \tag{13}$$
   Także w zastosowaniu do obliczania odwrotności otrzymujemy
$$(z^{-1})^* = (z^*)^{-1} \ . \tag{14}$$

6. Dowodzi się, że
$$e^{\phi i} = \cos\phi + i \sin\phi \ , \quad \text{gdzie} \quad \phi \in \mathbb{R} \ , \tag{15}$$

7. Liczba zespolona $z \neq 0$ może być zapisana w tzw. postaci wykładniczej,
$$z = r\, e^{\phi i} \ , \tag{16}$$
   oraz w równoważnej postaci trygonometrycznej:
$$z = r \cos\phi + i\, r \sin\phi \ , \tag{17}$$
   gdzie r jest liczbą rzeczywistą nieujemną, a $\phi \in [0, 2\pi)$. Łatwo sprawdzić, że
$$r = |z| = \sqrt{a^2 + b^2} \ . \tag{18}$$
   Wielkość kątowa ϕ nazywana jest argumentem liczby z. Jest ona wyznaczona jako rozwiązanie układu równań
$$\begin{aligned} \cos\phi &= \tfrac{a}{\sqrt{a^2+b^2}} \ , \\ \sin\phi &= \tfrac{b}{\sqrt{a^2+b^2}} \ , \end{aligned} \tag{19}$$

należące do przedziału $[0, 2\pi)$. Dla $z = 0$ mamy $r = 0$, a wartość ϕ jest nieokreślona. Postacie: wykładnicza i trygonometryczna liczby zespolonej z są wygodne przy obliczaniu potęg z o dowolnym wykładniku całkowitym (dodatnim lub ujemnym), oraz ułamkowym (wyciąganie pierwiastków). W szczególności, odwrotność liczby z, patrz równ. (10), przedstawić można w postaciach:

$$z^{-1} = \frac{1}{r}\, e^{-\phi i} = \frac{1}{r}\, \cos\phi - \frac{i}{r}\, \sin\phi\ , \tag{20}$$

które wynikają z równ. (15) i (16).

8. Rozważmy wielomian stopnia n zmiennej zespolonej z, o współczynnikach zespolonych:
$$W^{(n)}(z) = a_n z^n + a_{n-1} z^{n-1} + \ldots + a_0\ . \tag{21}$$
Ważnym problemem jest określenie, czy taki wielomian ma pierwiastki (rozwiązania równania $W^{(n)}(z) = 0$). Tzw. podstawowe twierdzenie algebry glosi, że wielomian (21) ma dokładnie n pierwiastków, $z_1, z_2, \ldots, z_n$, (niekoniecznie różnych), tak, że możliwe jest przedstawienie iloczynowe:
$$W^{(n)}(z) = a_n(z - z_1)(z - z_2)\cdots(z - z_n)\ . \tag{22}$$
Twierdzenie to nie zachodzi dla wielomianów zmiennej rzeczywistej. Np. wielomiany $W^{(2)}(x) = x^2 + 1$ i $W^{(4)}(x) = x^4 + 4x^2 + 3$ nie mają żadnych pierwiastków rzeczywistych.

9. Liczby zespolone spełniające warunek
$$z^* = z\ , \tag{23}$$
(czyli liczby zespolone, których część urojona $\mathrm{Im} z = 0$) mają wszystkie własności arytmetyczne liczb rzeczywistych. Wygodnie będzie więc utożsamiać te liczby z liczbami rzeczywistymi. W tym sensie będziemy dalej uważać zbiór liczb rzeczywistych $\mathbb{R}$ za podzbiór zbioru liczb zespolonych $\mathbb{C}$ (co zapisuje się w postaci $\mathbb{R} \subset \mathbb{C}$), a równanie (23) za warunek definiujący liczbę rzeczywistą.

10. Liczby zespolone spełniające warunek
$$z^* = -z\ , \tag{24}$$
(czyli liczby zespolone, których część rzeczywista $\mathrm{Re} z = 0$) nazywane są liczbami urojonymi. Liczby te można zapisać w postaci $z = bi$, gdzie b jest liczbą rzeczywistą. Suma dwóch liczb urojonych jest liczbą urojoną, ale iloczyn dwóch liczb urojonych jest zawsze liczbą rzeczywistą.

2 Macierze

Macierzą $\boldsymbol{A}$ o wymiarach $M \times N$ nazywa się zbiór pewnych elementów $\{A_{mn}\}_{m=1,n=1}^{m=M,n=N}$, ułożonych w formie dwuwymiarowej tablicy:

$$\boldsymbol{A} = \begin{pmatrix} A_{11} & A_{12} & \cdots & A_{1N} \\ A_{21} & A_{22} & \cdots & A_{2N} \\ \vdots & \vdots & \vdots & \vdots \\ A_{M1} & A_{M2} & \cdots & A_{MN} \end{pmatrix} . \tag{25}$$

Gdy $N \neq M$, macierz taką nazywamy macierzą prostokątną, a gdy $N = M$, macierzą kwadratową. Pierwszy wskaźnik, $m = 1, 2, \ldots, M$, numeruje wiersze macierzy, a drugi wskaźnik, $n = 1, 2, \ldots, N$, numeruje kolumny macierzy. Gdy $M = 1$, macierz nazywa się macierzą (jedno)wierszową, a gdy $N = 1$, macierzą (jedno)kolumnową; takie macierze będą dalej oznaczane małymi literami pogrubionymi, np.

$$\boldsymbol{a} = \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_M \end{pmatrix} . \tag{26}$$

oznacza macierz kolumnową, gdzie dla wygody pominięto drugi wskaźnik (zawsze równy 1), np. $a_{11} = a_1$, $a_{21} = a_2$, itd. Elementy macierzy, A_{mn}, są zwykle liczbami należącymi do pewnego zbioru $\mathbb{K}$, gdzie $\mathbb{K} = \mathbb{R}$ lub $\mathbb{K} = \mathbb{C}$. Zgodnie z uwagą 9 rozdziału 1 przyjmujemy, że $\mathbb{R} \subset \mathbb{C}$, i będziemy traktować przypadek $\mathbb{K} = \mathbb{R}$ jako przypadek szczególny sytuacji ogólnej $\mathbb{K} = \mathbb{C}$.

1. Równość dwu macierzy, $\boldsymbol{A} = \boldsymbol{B}$, zachodzi, gdy mają te same wymiary i ich odpowiednie elementy są równe:
$$A_{mn} = B_{mn} . \tag{27}$$

2. Dodawanie macierzy jest określone dla macierzy o tych samych wymiarach $M \times N$:
$$\boldsymbol{C} = \boldsymbol{A} + \boldsymbol{B} , \tag{28}$$
gdy elementy macierzy $\boldsymbol{C}$ wyznaczone są w postaci:
$$C_{mn} = A_{mn} + B_{mn} . \tag{29}$$
Dodawanie macierzy jest przemienne. Istnieje macierz zerowa $\boldsymbol{O}$ (o elementach równych 0), taka, że dla dowolnej macierzy $\boldsymbol{A}$
$$\boldsymbol{A} + \boldsymbol{O} = \boldsymbol{A} . \tag{30}$$

Dla każdej macierzy $\boldsymbol{A}$ istnieje macierz przeciwna $-\boldsymbol{A}$ (o elementach przeciwnych, równych $-A_{mn}$), tak, że zachodzi

$$\boldsymbol{A} + (-\boldsymbol{A}) = \boldsymbol{A} - \boldsymbol{A} = \boldsymbol{O} \, . \tag{31}$$

Można też zdefiniować mnożenie macierzy przez liczbę:

$$\boldsymbol{C} = c\boldsymbol{A} \, , \tag{32}$$

gdy elementy macierzy $\boldsymbol{C}$ wyznaczone są w postaci:

$$C_{mn} = cA_{mn} \, . \tag{33}$$

3. Niech $\boldsymbol{A}$ jest macierzą o wymiarach $M \times K$, a $\boldsymbol{B}$ macierzą o wymiarach $K \times N$. Iloczynem tych macierzy nazywamy macierz $\boldsymbol{C}$ o wymiarach $M \times N$,

$$\boldsymbol{C} = \boldsymbol{A}\boldsymbol{B} \, , \tag{34}$$

której elementy oblicza się w sposób następujący:

$$C_{mn} = A_{m1}B_{1n} + A_{m2}B_{2n} + \ldots + A_{mK}B_{Kn} = \sum_{k=1}^{K} A_{mk}B_{kn} \, . \tag{35}$$

Można powiedzieć, że element C_{mn} powstaje w wyniku „pomnożenia” m-tego wiersza macierzy $\boldsymbol{A}$ przez n-tą kolumnę macierzy $\boldsymbol{B}$.

4. Gdy macierze $\boldsymbol{A}$, $\boldsymbol{B}$ są macierzami kwadratowymi o wymiarach $M \times M$, ich iloczyny $\boldsymbol{AB}$ i $\boldsymbol{BA}$ są także macierzami kwadratowymi o tych wymiarach i, naogół, $\boldsymbol{AB} \neq \boldsymbol{BA}$. W zbiorze macierzy kwadratowych, oprócz działań opisanych w p. 2, jest więc określone mnożenie (nieprzemienne). Z definicji dodawania (29) i mnożenia macierzy (35) wynika, że zachodzi rozdzielność dodawania względem mnożenia:

$$\begin{aligned} (\boldsymbol{A} + \boldsymbol{B})\,\boldsymbol{C} &= \boldsymbol{AC} + \boldsymbol{BC} \, , \\ \boldsymbol{C}\,(\boldsymbol{A} + \boldsymbol{B}) &= \boldsymbol{CA} + \boldsymbol{CB} \, . \end{aligned} \tag{36}$$

5. Mając daną macierz zespoloną $\boldsymbol{A}$ o wymiarach $M \times N$, można skonstruować pewne macierze z nią związane:

   (a) Macierz zespolona sprzężona $\boldsymbol{A}^* = \boldsymbol{B}$, o takich samych wymiarach i o elementach sprzężonych zespolonych,

   $$B_{mn} = A^*_{mn} \, . \tag{37}$$

   (b) Macierz transponowaną $\boldsymbol{A}^{\mathrm{T}} = \boldsymbol{B}$, o wymiarach $N \times M$ i elementach

   $$B_{mn} = A_{nm} \, . \tag{38}$$

   Macierz transponowana $\boldsymbol{A}^{\mathrm{T}}$ powstaje z macierzy $\boldsymbol{A}$ przez zamianę wierszy na kolumny i kolumn na wiersze, z pozostawieniem na miejscu tzw. elementów diagonalnych: $A_{11}, A_{22}, \ldots$.

(c) Macierz hermitowsko sprzężoną $\boldsymbol{A}^\dagger = \boldsymbol{B}$, o wymiarach $N \times M$ i elementach

$$B_{mn} = A^*_{nm} \ . \tag{39}$$

Zachodzi oczywiście

$$\boldsymbol{A}^\dagger = (\boldsymbol{A}^*)^{\mathrm{T}} = (\boldsymbol{A}^{\mathrm{T}})^* \ . \tag{40}$$

Przykład: macierz hermitowsko sprzężona względem macierzy kolumnowej $\boldsymbol{a}$, patrz równ. (26), jest macierzą wierszową:

$$\boldsymbol{a}^\dagger = (a_1^*, a_2^*, \ldots, a_M^*) \ . \tag{41}$$

Zauważmy przy okazji, że mnożenie macierzy wierszowej $\boldsymbol{a}^\dagger$ przez macierz kolumnową $\boldsymbol{b}$, analogiczną do macierzy (26), daje macierz $\boldsymbol{a}^\dagger\boldsymbol{b}$, która jest macierzą o wymiarach 1×1, czyli pewną liczbą ze zbioru liczb $\mathbb{K}$: $\boldsymbol{a}^\dagger\boldsymbol{b} \in \mathbb{K}$.

Dwukrotne powtórzenie każdej z operacji (a-c) przywraca stan początkowy:

$$(\boldsymbol{A}^*)^* = (\boldsymbol{A}^{\mathrm{T}})^{\mathrm{T}} = (\boldsymbol{A}^\dagger)^\dagger = \boldsymbol{A} \ . \tag{42}$$

Zastosowanie operacji (a-c) do sumy macierzy daje:

$$\begin{array}{rcl} (\boldsymbol{A}+\boldsymbol{B})^* & = & \boldsymbol{A}^* + \boldsymbol{B}^* \ , \\ (\boldsymbol{A}+\boldsymbol{B})^{\mathrm{T}} & = & \boldsymbol{A}^{\mathrm{T}} + \boldsymbol{B}^{\mathrm{T}} \ , \\ (\boldsymbol{A}+\boldsymbol{B})^\dagger & = & \boldsymbol{A}^\dagger + \boldsymbol{B}^\dagger \ , \end{array} \tag{43}$$

a w przypadku iloczynu macierzy przez liczbę:

$$\begin{array}{rcl} (c\boldsymbol{A})^* & = & c^*\boldsymbol{A}^* \ , \\ (c\boldsymbol{A})^{\mathrm{T}} & = & c\boldsymbol{A}^{\mathrm{T}} \ , \\ (c\boldsymbol{A})^\dagger & = & c^*\boldsymbol{A}^\dagger \ . \end{array} \tag{44}$$

Z kolei, zastosowanie operacji (a-c) do iloczynu macierzy daje:

$$\begin{array}{rcl} (\boldsymbol{AB})^* & = & \boldsymbol{A}^*\boldsymbol{B}^* \ , \\ (\boldsymbol{AB})^{\mathrm{T}} & = & \boldsymbol{B}^{\mathrm{T}}\boldsymbol{A}^{\mathrm{T}} \ , \\ (\boldsymbol{AB})^\dagger & = & \boldsymbol{B}^\dagger\boldsymbol{A}^\dagger \ . \end{array} \tag{45}$$

Odnotujmy zmianę porządku iloczynu w wyniku operacji (b) i (c).

6. Szczególną rolę w algebrze odgrywają macierze kwadratowe ($N = M$). Wśród macierzy kwadratowych wyróżnioną role odgrywają: (kwadratowa) macierz zerowa $\boldsymbol{O}$, o własnościach określonych w równ. (30) i (31), oraz macierz jednostkowa,

$$\mathbf{1} = \begin{pmatrix} 1 & 0 & \ldots & 0 \\ 0 & 1 & \ldots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \ldots & 1 \end{pmatrix} \ , \tag{46}$$

spełniająca rolę „jedynki” w mnożeniu macierzy:

$$\mathbf{1}\boldsymbol{A} = \boldsymbol{A}\mathbf{1} = \boldsymbol{A} \ . \tag{47}$$

7. Każdej macierzy kwadratowej $\boldsymbol{A}$ można przypisać liczbę $\det \boldsymbol{A}$, zwaną wyznacznikiem macierzy $\boldsymbol{A}$, zapisywaną także jako

$$\det \boldsymbol{A} = \begin{vmatrix} A_{11} & A_{12} & \cdots & A_{1M} \\ A_{21} & A_{22} & \cdots & A_{2M} \\ \vdots & \vdots & \vdots & \vdots \\ A_{M1} & A_{M2} & \cdots & A_{MM} \end{vmatrix} , \tag{48}$$

a zdefinowaną formalnie w postaci

$$\det \boldsymbol{A} = \sum_{\sigma} \text{znak}(\sigma) A_{1\sigma(1)} A_{2\sigma(2)} \cdots A_{M\sigma(M)} , \tag{49}$$

gdzie sumowanie rozciąga się na wszystkie $M!$ permutacji zbioru wskaźników $1, 2, \ldots, M$ numerujących kolumny macierzy $\boldsymbol{A}$, a znak(σ) jest znakiem permutacji σ, przyjmującym wartości $+1$ lub -1, w zależności od tego, czy permutacja ta jest parzysta, czy nieparzysta. Prosty przykład: wyznacznik macierzy $\boldsymbol{A}$ o wymiarach 2×2:

$$\det \boldsymbol{A} = \begin{vmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{vmatrix} = A_{11} A_{22} - A_{12} A_{21} . \tag{50}$$

W przypadku macierzy wyższych wymiarów efektywnym sposobem obliczania wyznaczników jest zastosowanie tzw. rozwinięcia Laplace'a (względem kolumn lub wierszy macierzy). Niektóre własności wyznacznika macierzy:

(a) Jeśli z macierzy $\boldsymbol{A}$ utworzyć nową macierz $\boldsymbol{A}'$ przez zamianę dwóch kolumn (lub dwóch wierszy), to $\det \boldsymbol{A}' = -\det \boldsymbol{A}$. Wynika stąd, że gdy dwie kolumny (lub dwa wiersze) macierzy $\boldsymbol{A}$ są identyczne, to $\det \boldsymbol{A} = 0$.

(b) Zachodzi

$$\det \mathbf{1} = 1 . \tag{51}$$

Dla każdej macierzy kwadratowej $\boldsymbol{A}$ mamy

$$\det(\boldsymbol{A}^{\mathrm{T}}) = \det \boldsymbol{A} , \tag{52}$$

$$\det(\boldsymbol{A}^{*}) = \det(\boldsymbol{A}^{\dagger}) = (\det \boldsymbol{A})^{*} , \tag{53}$$

(c) Jeśli macierze $\boldsymbol{A}$ i $\boldsymbol{B}$ są macierzami kwadratowymi o tych samych wymiarach, to

$$\det(\boldsymbol{AB}) = \det(\boldsymbol{BA}) = \det \boldsymbol{A} \det \boldsymbol{B} . \tag{54}$$

8. Jeśli $\boldsymbol{A}$ jest macierzą kwadratową i $\det \boldsymbol{A} \neq 0$, to istnieje macierz odwrotna do macierzy $\boldsymbol{A}$, oznaczana przez $\boldsymbol{A}^{-1}$, taka, że

$$\boldsymbol{A}^{-1} \boldsymbol{A} = \boldsymbol{A} \boldsymbol{A}^{-1} = \mathbf{1} . \tag{55}$$

Z równ. (51), (54) i (55) wynika, że

$$\det(\boldsymbol{A}^{-1}) = (\det \boldsymbol{A})^{-1} \ . \tag{56}$$

Prosty przykład: macierz odwrotna do macierzy $\boldsymbol{A}$ o wymiarach 2×2:

$$\boldsymbol{A}^{-1} = \frac{1}{\det \boldsymbol{A}} \begin{pmatrix} A_{22} & -A_{12} \\ -A_{21} & A_{11} \end{pmatrix} , \tag{57}$$

gdzie $\det \boldsymbol{A}$ jest zdefiniowany w równ. (50).
Jeśli $\boldsymbol{A}$ i $\boldsymbol{B}$ są macierzami kwadratowymi o tych samych wymiarach i są odwracalne (czyli mają odwrotności), to ich iloczyn $\boldsymbol{AB}$ jest także macierzą odwracalną [bo wyznacznik (54) jest różny od zera] i zachodzi

$$(\boldsymbol{AB})^{-1} = \boldsymbol{B}^{-1}\boldsymbol{A}^{-1} \ . \tag{58}$$

Tu, podobnie jak w przypadku operacji (b) i (c) w równ. (45), także zachodzi zmiana porządku iloczynu.

9. Rozważmy macierze zespolone ($\mathbb{K} = \mathbb{C}$). Macierz kwadratową $\boldsymbol{A}$ nazywamy:
– rzeczywistą, gdy
$$\boldsymbol{A}^* = \boldsymbol{A} \quad (\text{czyli } A^*_{mn} = A_{mn}) , \tag{59}$$
– symetryczną, gdy
$$\boldsymbol{A}^T = \boldsymbol{A} \quad (\text{czyli } A_{nm} = A_{mn}) , \tag{60}$$
– hermitowską, gdy
$$\boldsymbol{A}^\dagger = \boldsymbol{A} \quad (\text{czyli } A^*_{nm} = A_{mn}) , \tag{61}$$
– ortogonalną, gdy $\det \boldsymbol{A} \neq 0$ i zachodzi
$$\boldsymbol{A}^T = \boldsymbol{A}^{-1} , \tag{62}$$
– unitarną, gdy $\det \boldsymbol{A} \neq 0$ i zachodzi
$$\boldsymbol{A}^\dagger = \boldsymbol{A}^{-1} . \tag{63}$$

Z równ. (52), (56) i (62) wynika, że w przypadku macierzy ortogonalnej zachodzi

$$(\det \boldsymbol{A})^2 = 1 \ . \tag{64}$$

Podobnie, z równ. (53), (56) i (63) wynika, że w przypadku macierzy unitarnej zachodzi

$$|\det \boldsymbol{A})|^2 = 1 \ . \tag{65}$$

Gdy rozważamy macierze rzeczywiste ($\mathbb{K} = \mathbb{R}$), to równ. (59) jest zawsze spełnione, patrz uwaga 9 w rozdziale 1. W tym przypadku pojęcie macierzy hermitowskiej pokrywa się z pojęciem macierzy symetrycznej, a pojęcie macierzy unitarnej pokrywa się z pojęciem macierzy ortogonalnej.

10. Każdej macierzy kwadratowej $\boldsymbol{A}$ można przypisać liczbę $\mathrm{tr}\,\boldsymbol{A}$, zwaną śladem macierzy $\boldsymbol{A}$, równą sumie elementów diagonalnych macierzy:

$$\mathrm{tr}\,\boldsymbol{A} = \sum_{m=1}^{M} A_{mm} \; . \tag{66}$$

Łatwo wykazać następujące własności śladu macierzy:

$$\mathrm{tr}\,(\boldsymbol{A}+\boldsymbol{B}) = \mathrm{tr}\,\boldsymbol{A} + \mathrm{tr}\,\boldsymbol{B} \; , \tag{67}$$

$$\mathrm{tr}\,(\boldsymbol{A}\boldsymbol{B}) = \mathrm{tr}\,(\boldsymbol{B}\boldsymbol{A}) \; , \tag{68}$$

$$\mathrm{tr}\,(\boldsymbol{A}^{\mathrm{T}}) = \mathrm{tr}\,\boldsymbol{A} \; , \tag{69}$$

$$\mathrm{tr}\,(\boldsymbol{A}^{*}) = \mathrm{tr}\,(\boldsymbol{A}^{\dagger}) = (\mathrm{tr}\,\boldsymbol{A})^{*} \; . \tag{70}$$

11. Transformacja podobieństwa macierzy kwadratowych. Wybierając ustaloną macierz kwadratową $\boldsymbol{X}$, taką, że $\det \boldsymbol{X} \neq 0$, możemy każdej macierzy kwadratowej $\boldsymbol{A}$ przyporządkować macierz kwadratową $\boldsymbol{A'}$ określoną równaniem

$$\boldsymbol{A'} = \boldsymbol{X}\boldsymbol{A}\boldsymbol{X}^{-1} \; . \tag{71}$$

Własności transformacji podobieństwa:

$$(\boldsymbol{A}+\boldsymbol{B})' = \boldsymbol{A'} + \boldsymbol{B'} \; , \tag{72}$$

$$(\boldsymbol{A}\boldsymbol{B})' = \boldsymbol{A'}\boldsymbol{B'} \; . \tag{73}$$

Jeśli istnieje macierz odwrotna $\boldsymbol{A}^{-1}$, to

$$(\boldsymbol{A}^{-1})' = (\boldsymbol{A'})^{-1} \; . \tag{74}$$

Z własności wyznacznika macierzy [równ. (54)] wynika niezmienniczość wyznacznika macierzy względem transformacji podobieństwa:

$$\det \boldsymbol{A'} = \det \boldsymbol{A} \; . \tag{75}$$

Podobnie, z własności śladu macierzy [równ. (68)] wynika niezmienniczość śladu macierzy względem transformacji podobieństwa:

$$\mathrm{tr}\,\boldsymbol{A'} = \mathrm{tr}\,\boldsymbol{A} \; . \tag{76}$$

Na ogół definicje macierzy zespolonej sprzężonej [równ. (37)], macierzy transponowanej [równ. (38)], oraz macierzy hermitowsko sprzężonej [równ. (39)] nie są niezmiennicze względem transformacji podobieństwa. W przypadku szczególnego wyboru macierzy transformacji $\boldsymbol{X}$ można jednak uzyskać niezmienniczość, np. gdy $\boldsymbol{X}$ jest macierzą unitarną [spełnia równ. (63)], to zachodzi

$$(\boldsymbol{A}^{\dagger})' = (\boldsymbol{A'})^{\dagger} \; . \tag{77}$$

3 Układy równań liniowych

Układ M równań liniowych z N niewiadomymi $x_1, x_2, \ldots, x_N$, należącymi do zbioru liczb $\mathbb{K}$ (równego $\mathbb{R}$ lub $\mathbb{C}$), zapisuje się w postaci

$$\begin{array}{cccccc} A_{11}\,x_1 + A_{12}\,x_2 + & \cdots & + A_{1N}\,x_N & = & b_1\,, \\ A_{21}\,x_1 + A_{22}\,x_2 + & \cdots & + A_{2N}\,x_N & = & b_2\,, \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ A_{M1}\,x_1 + A_{M2}\,x_2 + & \cdots & + A_{MN}\,x_N & = & b_M\,, \end{array} \tag{78}$$

gdzie liczby A_{mn}, zwane współczynnikami liniowymi, tworzą macierz prostokątną $\boldsymbol{A}$ o wymiarach $M \times N$, patrz równ. (25), a liczby b_m, zwane wyrazami wolnymi, tworzą macierz kolumnową $\boldsymbol{b}$ o wymiarze M, analogiczną do zdefiniowanej w równ. (26). Zarówno współczynniki liniowe, jak i wyrazy wolne należą do tego samego zbioru liczb $\mathbb{K}$. Wprowadzając macierz kolumnową o wymiarze N, zbudowaną z niewiadomych,

$$\boldsymbol{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_N \end{pmatrix}, \tag{79}$$

można układ równań (78) zapisać w równoważnej postaci macierzowej:

$$\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}\,, \tag{80}$$

gdzie mnożenie macierzy współczynników $\boldsymbol{A}$, o wymiarach $M \times N$, przez macierz niewiadomych $\boldsymbol{x}$, o wymiarach $N \times 1$, daje macierz wyrazów wolnych $\boldsymbol{b}$, o wymiarach $M \times 1$.

1. Układ równań liniowych (78) może: (i) nie mieć rozwiązań (nazywany jest wtedy sprzecznym), (ii) mieć jedno rozwiązanie $\boldsymbol{x}$, (iii) mieć nieskończenie wiele rozwiązań.

2. Układ równań liniowych (78) nazywa się układem jednorodnym, gdy wszystkie wyrazy wolne są równe zeru, $\boldsymbol{b} = \boldsymbol{0}$; w przeciwnym wypadku układ nazywa się układem niejednorodnym. Jednym z rozwiązań jednorodnego układu równań liniowych jest zawsze rozwiązanie zerowe, $\boldsymbol{x} = \boldsymbol{0}$. Uwaga: jeśli $\boldsymbol{x} \neq \boldsymbol{0}$ jest rozwiązaniem jednorodnego układu równań, a c jest dowolną liczbą ze zbioru $\mathbb{K}$, to $\boldsymbol{x}' = c\boldsymbol{x}$ jest także rozwiązaniem tego układu równań.

3. Dowodzi się, że jeśli w układzie równań (78):
   (a) jedno z równań-wierszy pomnożyć stronami przez pewną liczbę $c \neq 0$, to otrzymany w ten sposób nowy układ równań ma ten sam zbiór rozwiązań;
   (b) do jednego z równań-wierszy dodać stronami inne równanie-wiersz, to otrzymany w ten sposób nowy układ równań ma ten sam zbiór rozwiązań.
   Wielokrotne zastosowanie operacji (a) i (b), z odpowiednio dobranymi mnożnikami c, może doprowadzic do eliminacji jednej z niewiadomych, a w konsekwencji do otrzymania nowego układu $M-1$ równań z $N-1$ niewiadomymi; jest to metoda rozwiązywania układu równań liniowych przez kolejną eliminację niewiadomych.

4. Skupimy się teraz na ważnym przypadku szczególnym, gdy liczba równań równa jest liczbie niewiadomych, $M = N$. Macierz $\boldsymbol{A}$ jest wtedy macierzą kwadratową $N \times N$. Rozpatrzymy dwie możliwości:
(1) $\det \boldsymbol{A} \neq 0$. Istnieje wtedy macierz odwrotna $\boldsymbol{A}^{-1}$, i mnożąc lewostronnie obie strony równania macierzowego (80) przez tę macierz otrzymujemy (jedyne) rozwiązanie układu równań (78) w postaci

$$\boldsymbol{x} = \boldsymbol{A}^{-1}\boldsymbol{b}\,. \tag{81}$$

Metoda rozwiązywania układu równań liniowych $N \times N$ przez wyznaczanie macierzy $\boldsymbol{A}^{-1}$ nie jest opłacalna w praktyce, gdy liczba niewiadomych N jest duża; korzystniej jest wtedy stosować odpowiednio zoptymalizowaną metodę kolejnej eliminacji niewiadomych. W szczególnym przypadku jednorodnego układu równań, $\boldsymbol{b} = \boldsymbol{0}$, z równ. (81) otrzymujemy jako (jedyne) rozwiązanie wynik $\boldsymbol{x} = \boldsymbol{0}$.
(2) $\det \boldsymbol{A} = 0$. W tym przypadku można wykazać, że niejednorodny układ równań (78) jest albo sprzeczny, albo ma nieskończenie wiele rozwiązań. Interesujący jest przypadek układu jednorodnego – można wykazać, że taki układ ma w tym przypadu także rozwiązania niezerowe, $\boldsymbol{x} \neq \boldsymbol{0}$. Co więcej, jest wtedy nieskończenie wiele rozwiązań niezerowych proporcjonalnych do siebie, patrz uwaga w punkcie 2. Mogą jednak wystąpić także rozwiązania niezerowe, które nie są proporcjonalne do siebie. Łatwo można wykazać, że jeśli $\boldsymbol{x}$ i $\boldsymbol{x}'$ są takimi dwoma rozwiązaniami, to $\boldsymbol{x}'' = c\boldsymbol{x} + c'\boldsymbol{x}'$, gdzie c i c' są dowolnymi liczbami ze zbioru $\mathbb{K}$, jest także rozwiązaniem niezerowym naszego jednorodnego układu równań liniowych. Podumowując: warunkiem koniecznym i wystarczającym, by jednorodny układ N równań liniowych z N niewiadomymi (78) miał rozwiązania niezerowe, jest $\det \boldsymbol{A} = 0$.

4 Odwzorowania

Mówi się, że zostało określone odwzorowanie f zbioru $\mathbb{X}$ w zbiór $\mathbb{Y}$, jeśli *każdemu* elementowi x zbioru $\mathbb{X}$ został przyporządkowany *jeden* element y ze zbioru $\mathbb{Y}$. Dany element x zbioru $\mathbb{X}$ nazywamy argumentem odworowania f, a przyporządkowany mu element y zbioru $\mathbb{Y}$ nazywamy wartością odwzorowania f odpowiadającą x, co zapisujemy w postaci $y = f(x)$. Zbiór $\mathbb{X}$ nazywamy dziedziną odwzorowania f, a zbiór $\mathbb{Y}$ przeciwdziedziną tego odwzorowania. Odwzorowanie jest określone jednoznacznie przez swoją dziedzinę i przeciwdziedzinę, oraz przez podanie zależności $y = f(x)$. Odwzorowanie zapisuje się symbolicznie w postaci „strzałki":

$$\mathbb{X} \xrightarrow{f} \mathbb{Y} , \tag{82}$$

inna forma to $f\colon\ \mathbb{X} \to \mathbb{Y}$.

Podana wyżej definicją odwzorowania, wykorzystująca bardziej rozbudowaną notację matematyczną, wygląda następująco:

$$\mathbb{X} \ni x \mapsto y = f(x) \in \mathbb{Y} . \tag{83}$$

gdzie zakłada sie, że
(i) każdemu x jest przypisane pewne y,
(ii) jeśli danemu x są przypisane y_1 i y_2, to $y_1 = y_2$.
Dwa odwzorowania f i g są identyczne, co zapisuje się jako $f = g$, wtedy i tylko wtedy gdy:
(i) dziedziny $\mathbb{X}$ obu odwzorowań są identyczne,
(ii) przeciwdziedziny $\mathbb{Y}$ obu odwzorowań są identyczne,
(iii) dla każdego $x \in \mathbb{X}$ zachodzi $f(x) = g(x)$.
Gdy zbiory $\mathbb{X}$ i $\mathbb{Y}$ są zbiorami liczb (podzbiorami $\mathbb{R}$ lub $\mathbb{C}$), to odwzorowanie f nazywa się funkcją. Przykład: poniższe definicje określają trzy *różne* funkcje:
(a) $f_1 = \sin$, o dziedzinie $\mathbb{R}$ i przeciwdziedzinie $\mathbb{R}$,
(b) $f_2 = \sin$, o dziedzinie $\mathbb{R}$ i przeciwdziedzinie $[-1, 1]$,
(c) $f_3 = \sin$, o dziedzinie $[0, 2\pi]$ i przeciwdziedzinie $[-1, 1]$.
UWAGA: z powyższych rozważań wynika, że powinno się rozróżniać symbol f, określający odwzorowanie, od symbolu $f(x)$, określającego wartość odwzorowania f odpowiadającą danemu argumentowi x. Zapisu $f(x)$ na określenie odwzorowania stosuje się czasem, by podkreślić, jak wygląda zbiór argumentów odwzorowania: np. $f(x, y)$ jako symbol odwzorowania wskazuje, że zbiorem argumentów jest zbiór uporządkowanych par elementów (x, y).

1. Jeśli $\mathbb{A}$ jest podzbiorem dziedziny odwzorowania f, $\mathbb{A} \subset \mathbb{X}$, to przez $f(\mathbb{A})$ oznacza się podzbiór przeciwdziedziny $\mathbb{Y}$ złożony ze wszystkich y spełniających warunek: istnieje $x \in \mathbb{A}$ taki, że $y = f(x)$. Zbiór $f(\mathbb{A})$ nazywa się obrazem zbioru $\mathbb{A}$ w odwzorowaniu f. Zbiór $f(\mathbb{X})$, gdzie $\mathbb{X}$ jest dziedziną odwzorowania f, nazywa się po prostu obrazem odwzorowania f. Oczywiście $f(\mathbb{X})$ jest podzbiorem przeciwdziedziny $\mathbb{Y}$, $f(\mathbb{X}) \subset \mathbb{Y}$, ale niekoniecznie równa się $\mathbb{Y}$.

2. Jeśli $\mathbb{B}$ jest podzbiorem przeciwdziedziny odwzorowania f, $\mathbb{B} \subset \mathbb{Y}$, to przez $f^{-1}(\mathbb{B})$ oznacza się podzbiór dziedziny $\mathbb{X}$ złożony ze wszystkich x spełniających warunek: istnieje $y \in \mathbb{B}$ taki, że $y = f(x)$. Zbiór $f^{-1}(\mathbb{B})$ nazywa się przeciwobrazem zbioru $\mathbb{B}$ w odwzorowaniu f. Uwaga: zbiór $f^{-1}(\mathbb{B})$ może być zbiorem pustym [gdy nie istnieje żaden x spełniający warunek $y = f(x)$ dla $y \in \mathbb{B}$]. Uwaga: zawsze zachodzi: $f^{-1}(\mathbb{Y}) = \mathbb{X}$.

3. Jeśli f jest odwzorowaniem zbioru $\mathbb{X}$ w zbiór $\mathbb{Y}$, a g jest odwzorowaniem zbioru $\mathbb{Y}$ w zbiór $\mathbb{Z}$, to można zbudować odwzorowanie k zbioru $\mathbb{X}$ w zbiór $\mathbb{Z}$, takie, że $k(x) = g[f(x)]$. Odwzorowanie k nazywamy *złożeniem* odwzorowań g i f (w podanej kolejności), co oznacza się $k = g \circ f$. Dla trzech odwzorowań f, g i h (gdzie to ostatnie jest odwzorowaniem zbioru $\mathbb{Z}$ w jakiś zbiór $\mathbb{W}$) można konstruować odwzorowanie t zbioru $\mathbb{X}$ w zbiór $\mathbb{W}$ określone przez $t(x) = h\{g[f(x)]\}$. Wynika stąd, że składanie odwzorowań spełnia *warunek łączności*:

$$t = h \circ (g \circ f) = (h \circ g) \circ f = h \circ g \circ f \; , \tag{84}$$

gdzie ostatnia równość oznacza, że użycie nawiasów jest zbędne. Jeśli $\mathbb{X} = \mathbb{Y} = \mathbb{Z}$, to oprócz odwzorowania $k = g \circ f$ można też rozważać odwzorowanie $k' = f \circ g$, gdzie $k'(x) = f[g(x)]$. Można się przekonać, że na ogół

$$g \circ f \neq f \circ g \; , \tag{85}$$

a więc operacja składania odwzorowań jest w ogólności *nieprzemienna*.

4. Dla każdego zbioru *tożsamościowe* istnieje odwzorowanie $\mathbb{X} \to \mathbb{X}$, oznaczane $\mathrm{id}_{\mathbb{X}}$, zdefiniowane w następujący sposób:

$$\mathrm{id}_{\mathbb{X}}(x) = x \; , \tag{86}$$

i nazywane odwzorowaniem *tożsamościowym* (w zbiorze $\mathbb{X}$). Dla dowolnego odwzorowania f zbioru $\mathbb{X}$ w zbiór $\mathbb{Y}$ zachodzi:

$$f \circ \mathrm{id}_{\mathbb{X}} = f \; , \tag{87}$$

$$\mathrm{id}_{\mathbb{Y}} \circ f = f \; . \tag{88}$$

5. Odwzorowanie f nazywamy odwzorowaniem różnowartościowym (lub *injekcją*), jeśli z warunku $x_1 \neq x_2$ wynika, że $f(x_1) \neq f(x_2)$.

6. Odwzorowanie f nazywamy odwzorowaniem na (lub *surjekcją*), jeśli dla każdego $y \in \mathbb{Y}$ istnieje $x \in \mathbb{X}$ taki, że $y = f(x)$ [czyli $f(\mathbb{X}) = \mathbb{Y}$].

7. Odwzorowanie f, które jest jednocześnie injekcją i surjekcją nosi nazwę odwzorowania wzajemnie jednoznacznego (lub *bijekcji*). Każde odwzorowanie tożsamościowe $\mathrm{id}_{\mathbb{X}}$ jest bijekcją. Jeśli odwzorowania f i g są bijekcjami, to ich złożenie, $k = g \circ f$, też jest bijekcją. Każda injekcja f zbioru $\mathbb{X}$ w zbiór $\mathbb{Y}$ staje się bijekcją, gdy przeciwdziedzinę $\mathbb{Y}$ ograniczyć do zbioru $f(\mathbb{X})$, czyli obrazu dziedziny $\mathbb{X}$. Ważnym przykładem bijekcji zbiorów skończonych $\mathbb{X} = \mathbb{Y} = \{1, 2, \ldots, M\}$ są permutacje $x \mapsto \sigma(x)$:

$$\begin{array}{ccccc} 1 & 2 & 3 & \cdots & M \\ \sigma(1) & \sigma(2) & \sigma(3) & \cdots & \sigma(M) \end{array}$$

gdzie żadna z wartości $\sigma(1), \sigma(2), \ldots$ się nie powtarza. W przypadku zbioru M-elementowego istnieje $M!$ różnych permutacji tego zbioru. Permutacje wykorzystywane są w definicji wyznacznika macierzy, patrz równ. (49).

8. Niech zbiory $\mathbb{X}$ i $\mathbb{Y}$ maja każdy skończoną liczbę elementów i oznaczmy przez $\mathrm{card}(\mathbb{X})$ liczbę elementów (czyli inaczej *moc*) zbioru $\mathbb{X}$.
(a) Jeśli istnieje odwzorowanie f zbioru $\mathbb{X}$ w zbiór $\mathbb{Y}$ będące injekcją, to musi zachodzić $\mathrm{card}(\mathbb{X}) \leq \mathrm{card}(\mathbb{Y})$,
(b) Jeśli istnieje odwzorowanie f zbioru $\mathbb{X}$ na zbiór $\mathbb{Y}$ będące surjekcją, to musi zachodzić $\mathrm{card}(\mathbb{Y}) \geq \mathrm{card}(\mathbb{Y})$,
Wynika stąd:
(c) Jeśli istnieje odwzorowanie f zbioru $\mathbb{X}$ w zbiór $\mathbb{Y}$ będące bijekcją, to musi zachodzić $\mathrm{card}(\mathbb{X}) = \mathrm{card}(\mathbb{Y})$, czyli zbiory $\mathbb{X}$ i $\mathbb{Y}$ muszą mieć taką samą liczbę elementów (taką samą moc). Ta własność służy do porównywania mocy dowolnych zbiorów (także nieskończonych, jak zbiory liczb całkowitych, rzeczywistych, czy zespolonych): mówi się, że dwa zbiory $\mathbb{X}$ i $\mathbb{Y}$ maja tę samą moc wtedy i tylko wtedy, gdy istnieje bijekcja zbioru $\mathbb{X}$ w zbiór $\mathbb{Y}$.

9. Gdy istnieje bijekcja f zbioru $\mathbb{X}$ w zbiór $\mathbb{Y}$, to można skonstruować tzw. *odwzorowanie odwrotne* do f, oznaczane przez f^{-1}, będące odwzorowaniem zbioru $\mathbb{Y}$ w zbiór $\mathbb{X}$, takie, że zachodzi $f^{-1}[f(x)] = x$ oraz $f[f^{-1}(y)] = y$; warunki te można zapisać jako

$$f^{-1} \circ f = \mathrm{id}_{\mathbb{X}} \ , \tag{89}$$

$$f \circ f^{-1} = \mathrm{id}_{\mathbb{Y}} \ . \tag{90}$$

Można wykazać, że odwzorowanie f^{-1} jest bijekcją.

10. W zbiorze wszystkich bijekcji $\{f\}$ zbioru $\mathbb{X}$ w siebie można wprowadzić działanie, które parze bijekcji f i g przyporządkowuje bijekcję k, będącą ich złożeniem: $k = f \circ g$. Działanie to ma następujące własności:
(a) Jest łączne, patrz równ. (84).
(b) Wśród bijekcji istnieje odwzorowanie tożsamościowe $\mathrm{id}_{\mathbb{X}}$, spełniające warunki:

$$f \circ \mathrm{id}_{\mathbb{X}} = f \ , \qquad \mathrm{id}_{\mathbb{X}} \circ f = f \ ,$$

dla każdej bijekcji f, patrz równ. (87) i (88).
(c) Każda bijekcja f posiada odwzorowanie odwrotne f^{-1}, spełniające warunki:

$$f^{-1} \circ f = \mathrm{id}_{\mathbb{X}} \ , \qquad f \circ f^{-1} = \mathrm{id}_{\mathbb{X}} \ ,$$

patrz równ. (89) i (90).
Wynika z tego, że zbiór wszystkich bijekcji dowolnego zbioru $\mathbb{X}$ ma strukturę *grupy*:

elementami grupy są bijekcje, a działaniem grupowym (nieprzemiennym) jest składanie odwzorowań; w grupie istnieje *element neutralny*, reprezentowany przez $\mathrm{id}_{\mathbb{X}}$, a każdy element grupy f ma *element odwrotny*, równy f^{-1}. Przykładem grupy bijekcji jest grupa permutacji, patrz punkt 7.

11. Inne przykłady grup:
(a) Zbiór liczb rzeczywistych $\mathbb{K} = \mathbb{R}$ lub zespolonych $\mathbb{K} = \mathbb{C}$, z dodawaniem jako (przemiennym) działaniem grupowym; 0 jest tu elementem neutralnym grupy, a $-c$ jest elementem grupy odwrotnym w stosunku do c.
(b) Zbiór liczb rzeczywistych różnych od zera $\mathbb{K}' = \mathbb{R}'$ lub zespolonych różnych od zera $\mathbb{K}' = \mathbb{C}'$, z mnożeniem jako (przemiennym) działaniem grupowym; 1 jest tu elementem neutralnym grupy, a $1/c$ jest elementem grupy odwrotnym w stosunku do c.
(c) Zbiór macierzy $\{\boldsymbol{A}\}$ o wymiarach $M \times N$, o elementach macierzowych ze zbioru liczbowego $\mathbb{K} = (\mathbb{R}$ lub $\mathbb{C})$, z dodawaniem jako (przemiennym) działaniem grupowym; macierz zerowa $\mathbf{O}$ jest tu elementem neutralnym grupy, a macierz $-\boldsymbol{A}$ jest elementem grupy odwrotnym w stosunku do macierzy $\boldsymbol{A}$.
(d) Zbiór kwadratowych macierzy $\{\boldsymbol{A}\}$ o elementach macierzowych ze zbioru liczbowego $\mathbb{K} = (\mathbb{R}$ lub $\mathbb{C})$, spełniających warunek $\det \boldsymbol{A} \neq 0$, z mnożeniem macierzy jako (nieprzemiennym) działaniem grupowym; macierz jednostkowa $\mathbf{1}$ jest tu elementem neutralnym grupy, a macierz odwrotna $\boldsymbol{A}^{-1}$ jest elementem grupy odwrotnym w stosunku do macierzy $\boldsymbol{A}$.
Grupy w przykładach (a)–(c) są grupami *przemiennymi* (czyli inaczej grupami *abelowymi*).

5 Przestrzenie wektorowe

Przestrzenią wektorową nad zbiorem liczb $\mathbb{K}$ (gdzie $\mathbb{K} = \mathbb{R}$ lub $\mathbb{K} = \mathbb{C}$) będziemy nazywać zbiór $\mathbb{V} = \{\boldsymbol{u}, \boldsymbol{v}, \boldsymbol{w}, \ldots\}$ pewnych obiektów zwanych wektorami, spelniający następujące warunki:

1. Określone jest dodawanie wektorów, $\boldsymbol{u} + \boldsymbol{v} = \boldsymbol{w}$, które jest działaniem *łącznym* i *przemiennym*. Istnieje *wektor zerowy* $\mathbf{0}$, taki, że

$$\mathbf{0} + \boldsymbol{v} = \boldsymbol{v} \ . \tag{91}$$

   Dla każdego wektora $\boldsymbol{v}$ istnieje *wektor przeciwny*, oznaczany przez $-\boldsymbol{v}$, taki, że

$$\boldsymbol{v} + (-\boldsymbol{v}) = \boldsymbol{v} - \boldsymbol{v} = \mathbf{0} \ . \tag{92}$$

   gdzie pierwsza równość oznacza, że *odejmowanie* wektora to dodawanie wektora przeciwnego.

2. Określone jest mnożenie wektorów przez liczby c ze zbioru $\mathbb{K}$, $c\boldsymbol{v} = \boldsymbol{v}c = \boldsymbol{w}$ (porządek czynników w iloczynie nie jest istotny). Mnożenie wektorów przez liczby jest *rozdzielne* ze względu na dodawania wektorów,

$$(\boldsymbol{u} + \boldsymbol{v})c = \boldsymbol{u}c + \boldsymbol{v}c \ . \tag{93}$$

   oraz ze względu na dodawanie liczb,

$$\boldsymbol{v}(c + d) = \boldsymbol{v}c + \boldsymbol{v}d \ . \tag{94}$$

3. Mnożenie wektorów przez liczby spełnia warunek *łączności*:

$$(\boldsymbol{v}c)d = \boldsymbol{v}(cd) = \boldsymbol{v}cd \ , \tag{95}$$

   gdzie ostatnia równość oznacza, że użycie nawiasów jest zbędne.

4. Zachodzi

$$\boldsymbol{v}1 = \boldsymbol{v} \ . \tag{96}$$

Własności opisane w punkcie 1 oznaczają, że przestrzeń wektorowa $\mathbb{V}$ jest grupą przemienną ze względu na dodawanie wektorów, patrz rozdział 4, punkty 10 i 11. Gdy $\mathbb{K} = \mathbb{R}$, $\mathbb{V}$ nazywana jest *rzeczywistą* przestrzenią wektorową, a gdy $\mathbb{K} = \mathbb{C}$ – *zespoloną* przestrzenią wektorową.

Z własności 1 – 4, które musi spełniać każda przestrzeń wektorowa $\mathbb{V}$, wynikają pewne własności pochodne, np.:

$$\boldsymbol{v}0 = \mathbf{0}c = \mathbf{0} \ , \tag{97}$$

$$\boldsymbol{v}(-c) = (-\boldsymbol{v})c = -(\boldsymbol{v}c) \ , \tag{98}$$

dla dowolnych wektorów $\boldsymbol{v}$ i dowolnych liczb c ze zbioru $\mathbb{K}$.

Przykłady przestrzeni wektorowych:

1. Nasze intuicyjne pojmowanie abstrakcyjnych przestrzeni wektorowych związane jest ze znanymi z geometrii przestrzeniami: dwuwymiarową, będącą podstawą planimetrii, i trójwymiarową, będącą podstawą stereometrii. Wektory należące do tych przestrzeni to odcinki skierowane („strzałki”), zaczepione w pewnym wspólnym punkcie; możemy określić dodawanie owych „strzałek,” oraz ich mnożenie przez liczby rzeczywiste. Każdej „strzałce” możemy przyporządkować dwójkę lub trójkę liczb rzeczywistych, będących współrzędnymi końca „strzałki” w pewnym kartezjańskim układzie współrzędnych; stąd te (rzeczywiste) przestrzenie wektorowe oznaczamy, odpowiednio, przez $\mathbb{R}^2$ i $\mathbb{R}^3$.

2. Zbiór macierzy o wymiarach $M \times N$, o elementach ze zbioru liczb $\mathbb{K}$ ma strukturę przestrzeni wektorowej nad $\mathbb{K}$ (patrz punkt 2 na str. 4).

3. Zbiór macierzy kolumnowych o wymiarze M, o elementach ze zbioru liczb $\mathbb{K}$, jest przestrzenią wektorową nad $\mathbb{K}$, oznaczaną symbolem $\mathbb{K}^M$ $(M = 1, 2, \ldots)$. Jest to przypadek szczególny przestrzeni wektorowych macierzy z przykładu 2. Przestrzenie $\mathbb{K}^M$ to bardzo ważna klasa przestrzeni wektorowych. Sam zbiór liczbowy $\mathbb{K}$ może być traktowany jako przestrzeń wektorowa $\mathbb{K}^1$ nad zbiorem liczbowym $\mathbb{K}$. Przestrzenie $\mathbb{R}^2$ i $\mathbb{R}^3$ były wspomniane w przykładzie 1.

4. Rozważmy zbiór wszystkich odwzorowań $\{f\}$ pewnego zbioru $\mathbb{X}$ w zbiór liczbowy $\mathbb{K}$; $\{f\}$ jest więc zbiorem funkcji o wartościach w zbiorze $\mathbb{K}$. Możemy określić sumę dwóch funkcji, $f + g$, jako funkcję, której wartości wyznacza się z wartości funkcji f i g:
$$(f + g)(x) = f(x) + g(x) \ , \tag{99}$$
dla każdego $x \in X$. Łatwo sprawdzić, że dodawanie funkcji jest łączne i przemienne. Istnieje funkcja zerowa (którą oznaczać będziemy po prostu przez 0):
$$0(x) = 0 \ , \tag{100}$$
spełniająca warunki
$$(f + 0)(x) = f(x) + 0(x) = f(x) \ . \tag{101}$$
Ponadto każdej funkcji f odpowiada funkcją przeciwna $-f$:
$$(-f)(x) = -f(x) \ , \tag{102}$$
i zachodzi
$$[f + (-f)](x) = f(x) - f(x) = 0(x) = 0 \ . \tag{103}$$
A więc zbiór funkcji $\{f\}$ jest grupą przemienną ze względu na działanie dodawania funkcji, patrz warunek 1 określający przestrzeń wektorową; łatwo sprawdzić, że spełnione są też pozostałe warunki: 2, 3 i 4. Zbiór $\{f\}$ można więc uważać za przestrzeń wektorową nad zbiorem liczbowym $\mathbb{K}$. Można wykazać, że gdy $\mathbb{X}$ jest zbiorem skończonym o M elementach, wtedy odpowiedni zbiór funkcji ma tę samą strukturę co przestrzeń wektorowa $\mathbb{K}^M$, patrz przykład 3 powyżej.

5.1 Baza przestrzeni wektorowej

W danej przestrzeni wektorowej $\mathbb{V}$ będziemy rozważać skończone (M-elementowe) zbiory wektorów ponumerowanych wskaźnikiem $k = 1, 2, \ldots, M$, czyli *uporządkowane* zbiory wektorów postaci

$$(\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_M)\,. \tag{104}$$

1. *Kombinacją liniową wektorów* ze zbioru (104) nazywamy wektor zapisany w postaci

   $$\boldsymbol{w} = \boldsymbol{v}_1 c_1 + \boldsymbol{v}_2 c_2 + \cdots + \boldsymbol{v}_M c_M = \sum_{k=1}^{M} \boldsymbol{v}_k c_k\,, \tag{105}$$

   gdzie $c_1, c_2, \ldots, c_M$ są *współczynnikami liniowymi*, należącymi do zbioru liczbowego $\mathbb{K}$.

2. Zbiór wektorów (104) nazywamy *zbiorem wektorów liniowo niezależnych*, jeśli kombinacja liniowa wektorów (105) równa jest wektorowi zerowemu, $\boldsymbol{w} = \boldsymbol{0}$, wtedy i tylko wtedy, gdy znikają wszystkie współczynniki liniowe, $c_1 = c_2 = \cdots = c_M = 0$. Ważne: własność liniowej niezależności jest cechą danego zbioru wektorów, a nie pojedynczych wektorów z tego zbioru. Jeśli jeden z wektorów zbioru (104) daje się wyrazić jako kombinacja liniowa pozostałych wektorów, to zbiór ten nie jest zbiorem wektorów liniowo niezależnych. W szczególności, nie jest zbiorem wektorów liniowo niezależnych zbiór w którym jeden z wektorów jest wektorem zerowym, lub jeden z wektorów jest proporcjonalny do innego wektora z tego zbioru.

3. Zbiór wektorów (104) nazywamy *zupełnym zbiorem wektorów* w danej przestrzeni wektorowej $\mathbb{V}$, jeśli dowolny wektor z tej przestrzeni przedstawić można w postaci pewnej kombinacji liniowej (105) wektorów ze zbioru (104). Także własność zupełności jest cechą danego zbioru wektorów, a nie pojedynczych wektorów z tego zbioru.

4. Jeśli zbiór wektorów (104) jest jednocześnie zbiorem wektorów liniowo niezależnych i zbiorem zupełnym, to nazywany jest *bazą* danej przestrzeni wektorowej $\mathbb{V}$. W danej przestrzeni wektorowej $\mathbb{V}$ można wybrać nieskończenie wiele różnych baz, dowodzi się jednak, że każde dwie bazy muszą mieć taką samą liczbę elementów; liczbę tę nazywa się *wymiarem* danej przestrzeni wektorowej i oznacza przez $\dim(\mathbb{V})$. Jeżeli zbiór wektorów (104) jest bazą rozważanej przestrzeni wektorowej $\mathbb{V}$, to

   $$\dim(\mathbb{V}) = M\,. \tag{106}$$

   Rozważamy tu tylko przestrzenie wektorowe skończonego wymiaru, $\dim(\mathbb{V}) < \infty$. Warto zauważyć, że $\dim(\mathbb{V})$ wyznacza maksymalną liczbę wektorów w zbiorze wektorów liniowo niezależnych w danej przestrzeni wektorowej.

5. Jeśli w przestrzeni wektorowej $\mathbb{V}$ dana jest baza (104) i mamy dwa wektory $\boldsymbol{w}$ i $\boldsymbol{w}'$ wyrażone w postaci kombinacji liniowych

   $$\boldsymbol{w} = \sum_{k=1}^{M} \boldsymbol{v}_k c_k\,, \quad \boldsymbol{w}' = \sum_{k=1}^{M} \boldsymbol{v}_k c'_k\,, \tag{107}$$

to z równości $\boldsymbol{w} = \boldsymbol{w}'$ wynika $c_k = c'_k$ dla $k = 1, 2, \ldots, M$. A więc każdy wektor w przestrzeni $\mathbb{V}$ można w sposób jednoznaczny przedstawić w postaci kombinacji liniowej (105) wektorów bazy. Pojęcie bazy przestrzeni wektorowej jest bardzo ważne: zwykle analiza własności danej przestrzeni wektorowej $\mathbb{V}$ poprzedzona jest wyborem pewnej bazy w tej przestrzeni, co pozwala wyrazić wszystkie wektory w postaci kombinacji liniowych wektorów bazy.

6. Dowodzi się, że w przestrzeni wektorowej $\mathbb{V}$ wymiaru N dany zbiór wektorów (104) może być zbiorem wektorów liniowo niezależnych tylko gdy $M \leq N$, a zbiorem zupełnym tylko gdy $M \geq N$. Gdy zbiór (104) jest zbiorem wektorów liniowo niezależnych i $M < N$, to nie każdy wektor z przestrzeni $\mathbb{V}$ da się wyrazić w postaci kombinacji liniowej (105). Z kolei, gdy zbiór (104) jest zbiorem zupełnym wektorów i $M > N$, to choć każdy wektor z przestrzeni $\mathbb{V}$ da się wyrazić w postaci kombinacji liniowej (105), to jednak przedstawienie takie nie spełnia warunku jednoznaczności, patrz punkt 5.

7. W przestrzeni wektorowej $\mathbb{V}$ wymiaru N każdy zbiór N wektorów liniowo niezależnych jest bazą przestrzeni $\mathbb{V}$.

8. Wymiar przestrzeni wektorowej $\mathbb{K}^M$ (patrz przykład 3 na str. 17) wynosi M. Ogólnie, przestrzeń wektorowa macierzy o wymiarach $M \times N$ (patrz przykład 2 na str. 17) ma wymiar MN.

Gdy w przestrzeni wektorowej $\mathbb{V}$ wymiaru $N > M$ wybierze się pewien M-elementowy uporządkowany zbiór wektorów (104) spełniający warunek liniowej niezależności, to zbiór $\mathbb{W}$ wszystkich wektorów, które można otrzymać jako kombinacje liniowe postaci (105), spełniać będzie następujące warunki: (i) suma wektorów należących do $\mathbb{W}$ należy do $\mathbb{W}$, (ii) iloczyn wektora należącego do $\mathbb{W}$ przez liczbę ze zbioru liczbowego$\mathbb{K}$ należy do $\mathbb{W}$. Oznacza to, że zbiór wektorów $\mathbb{W}$ jest domknięty ze względu na działania dodawania wektorów oraz mnożenia wektorów przez liczby. Zbiór $\mathbb{W}$ jest więc przestrzenią wektorową, zaś (104) jest bazą tej przestrzeni [wynika stąd, że $\dim(\mathbb{W}) = M$]. Mówi się, że $\mathbb{W}$ jest pewną *podprzestrzenią* przestrzeni wektorowej $\mathbb{V}$.

W ramach przykładu 4 na str. 17 rozważana byą przestrzeń wektorowa utworzona przez wszystkie odwzorowania pewnego zbioru $\mathbb{X}$ w zbiór liczbowy $\mathbb{K}$. Gdy zbiór $\mathbb{X}$ ma nieskończoną liczbę elementów, przestrzeń taka jest wymiaru nieskończonego; jest tak, np., w przypadku funkcji zmiennej rzeczywistej ($\mathbb{X} = \mathbb{R}$) o wartościach rzeczywistych ($\mathbb{K} = \mathbb{R}$) lub zespolonych ($\mathbb{K} = \mathbb{C}$). Jeśli wybierze się pewien skończony M-elementowy uporządkowany zbiór funkcji

$$(\phi_1, \phi_2, \ldots, \phi_M)\,, \tag{108}$$

to utworzyć można podprzestrzeń powyższej przestrzeni wektorowej, do której należą wszystkie funkcje mające postać kombinacji liniowych funkcji ze zbioru (108):

$$f = \sum_{k=1}^{M} \phi_k c_k\,, \tag{109}$$

gdzie współczynniki c_k należą do zbioru liczbowego $\mathbb{K}$. Zbiór (108) jest zbiorem funkcji liniowo niezależnych, wtedy i tylko wtedy, gdy z r'owności $f = 0$ [patrz równ. (100)] wynika, że wszystkie współczynniki c_k równe są zeru. W takim przypadku zbiór (108) jest bazą przestrzeni wektorowej kombinacji liniowych (109), a wymiar tej przestrzeni równy jest M. Przestrzeń kombinacji liniowych (109) jest więc pewną $M-$wymiarową podprzestrzenią (nieskończeniewymiarowej) przestrzeni wektorowej wszystkich odwzorowań $\mathbb{X} \to \mathbb{K}$.

Przykłady przestrzeni wektorowych i ich baz:

1. Przestrzeń $\mathbb{K}^1 = \mathbb{K}$, gdzie $\mathbb{K} = \mathbb{R}$ lub $\mathbb{C}$, jest jednowymiarowa. Bazę w tej przestrzeni stanowi każda różna od zera liczba $x \in \mathbb{K}$.

2. Wymiar przestrzeni $\mathbb{K}^2$ wynosi 2. Dwa wektory (macierze kolumnowe)

$$\boldsymbol{r}_1 = \begin{pmatrix} x_1 \\ y_1 \end{pmatrix}, \quad \boldsymbol{r}_2 = \begin{pmatrix} x_2 \\ y_2 \end{pmatrix},$$

   tworzą bazę w przestrzeni $\mathbb{K}^2$, gdy $x_1 y_2 - x_2 y_1 \neq 0$.

3. Ogólnie: w przestrzeni $\mathbb{K}^M$ zbiór M wektorów

$$\boldsymbol{b}_1 = \begin{pmatrix} b_{11} \\ b_{21} \\ \vdots \\ b_{M1} \end{pmatrix}, \quad \boldsymbol{b}_2 = \begin{pmatrix} b_{12} \\ b_{22} \\ \vdots \\ b_{M2} \end{pmatrix}, \quad \cdots, \quad \boldsymbol{b}_M = \begin{pmatrix} b_{1M} \\ b_{2M} \\ \vdots \\ b_{MM} \end{pmatrix}, \tag{110}$$

   jest bazą wtedy i tylko wtedy, gdy odpowiedni wyznacznik jest różny od zera:

$$\det \begin{pmatrix} b_{11} & b_{12} & \cdots & b_{1M} \\ b_{21} & b_{22} & \cdots & b_{2M} \\ \vdots & \vdots & \vdots & \vdots \\ b_{M1} & b_{M2} & \cdots & b_{MM} \end{pmatrix} \neq 0\,. \tag{111}$$

4. W przestrzeni wektorowej funkcji zmiennej rzeczywistej o wartościach rzeczywistych ($\mathbb{R} \to \mathbb{R}$) wybieramy zbiór funkcji

$$(\phi_1(x) = 1,\ \phi_2(x) = x,\ \phi_3(x) = x^2,\ \ldots,\ \phi_M(x) = x^{M-1})\,. \tag{112}$$

   Jest to M-elementowy zbiór liniowo niezależnych funkcji zmiennej rzeczywistej x. Odpowiednie kombinacje liniowe (109), o współczynnikach rzeczywistych c_k, są wielomianami stopnia $\leq M-1$ zmiennej x. Zbiór (112) stanowi więc bazę M-wymiarowej (rzeczywistej) przestrzeni wektorowej wielomianów zmiennej rzeczywistej x. Analogicznie konstruuje się (zespoloną) przestrzeń wektorową wielomianów zmiennej zespolonej.

5.2 Iloczyn skalarny wektorów

W przestrzeni wektorowej $\mathbb{V}$ nad zbiorem liczbowym $\mathbb{K}$ ($= \mathbb{R}$ lub $\mathbb{C}$) można wprowadzić dodatkową strukturę algebraiczną przez zdefiniowanie tzw. iloczynu skalarnego wektorów. Iloczyn ten określony jest przez przyporządkowanie każdej parze wektorów $\boldsymbol{v}$ i $\boldsymbol{w}$ liczby ze zbioru $\mathbb{K}$, oznaczanej przez $\langle \boldsymbol{v}|\boldsymbol{w}\rangle$. Na iloczyn skalarny nałożone są następujące warunki:

$$\langle \boldsymbol{v}|\boldsymbol{w}_1 + \boldsymbol{w}_2\rangle = \langle \boldsymbol{v}|\boldsymbol{w}_1\rangle + \langle \boldsymbol{v}|\boldsymbol{w}_2\rangle\,, \tag{113}$$

$$\langle \boldsymbol{v}|\boldsymbol{w}c\rangle = \langle \boldsymbol{v}|\boldsymbol{w}\rangle\, c\,. \tag{114}$$

Zakłada się, że w zespolonej przestrzeni wektorowej $\mathbb{V}$ zachodzi

$$\langle \boldsymbol{w}|\boldsymbol{v}\rangle = \langle \boldsymbol{v}|\boldsymbol{w}\rangle^*\,, \tag{115}$$

gdzie $\langle \boldsymbol{v}|\boldsymbol{w}\rangle^*$ oznacza wartość sprzężoną liczby zespolonej $\langle \boldsymbol{v}|\boldsymbol{w}\rangle$. Natomiast w rzeczywistej przestrzeni wektorowej $\mathbb{V}$, gdzie iloczyn skalarny przyjmuje wartości rzeczywiste, zachodzi

$$\langle \boldsymbol{w}|\boldsymbol{v}\rangle = \langle \boldsymbol{v}|\boldsymbol{w}\rangle\,; \tag{116}$$

w tym wypadku iloczyn skalarny $\langle \boldsymbol{v}|\boldsymbol{w}\rangle$ jest *symetryczną* funkcją wektorów $\boldsymbol{v}$ i $\boldsymbol{w}$.

Z równ. (115) i uwagi 9 na str. 3 wynika, że i w zespolonej przestrzeni wektorowej iloczyn skalarny wektora $\boldsymbol{v}$ przez siebie, $\langle \boldsymbol{v}|\boldsymbol{v}\rangle$, jest liczbą rzeczywistą, $\langle \boldsymbol{v}|\boldsymbol{v}\rangle \in \mathbb{R}$. Dalsze warunki nałożone na iloczyn skalarny (tak w rzeczywistej, jak i zespolonej przestrzeni wektorowej) mają postać:

$$\langle \boldsymbol{v}|\boldsymbol{v}\rangle \geq 0\,, \tag{117}$$

$$\langle \boldsymbol{v}|\boldsymbol{v}\rangle = 0 \quad \text{wtedy i tylko wtedy, gdy} \quad \boldsymbol{v} = \boldsymbol{0}\,. \tag{118}$$

Korzystając z własności (113-115) otrzymujemy, że w zespolonej przestrzeni wektorowej zachodzi

$$\langle \boldsymbol{v}_1 + \boldsymbol{v}_2|\boldsymbol{w}\rangle = \langle \boldsymbol{v}_1|\boldsymbol{w}\rangle + \langle \boldsymbol{v}_2|\boldsymbol{w}\rangle\,, \tag{119}$$

$$\langle \boldsymbol{v}c|\boldsymbol{w}\rangle = c^*\,\langle \boldsymbol{v}|\boldsymbol{w}\rangle\,, \tag{120}$$

zaś w rzeczywistej przestrzeni wektorowej powyższe równanie przyjmuje postać

$$\langle \boldsymbol{v}c|\boldsymbol{w}\rangle = c\,\langle \boldsymbol{v}|\boldsymbol{w}\rangle\,, \tag{121}$$

co wynika z równ. (114) i (116). Z własności (113-115) wynika też, że

$$\langle \boldsymbol{v}|\boldsymbol{0}\rangle = \langle \boldsymbol{0}|\boldsymbol{v}\rangle = 0\,, \tag{122}$$

dla każdego wektora $\boldsymbol{v}$.

1. W przestrzeni wektorowej, w której został określony iloczyn skalarny, zdefiniować można *normę* czyli *długość* wektora $\boldsymbol{v}$:

$$\|\boldsymbol{v}\| = \langle \boldsymbol{v}|\boldsymbol{v}\rangle^{1/2} = \sqrt{\langle \boldsymbol{v}|\boldsymbol{v}\rangle}\,. \tag{123}$$

Norma (długość) wektora przyjmuje zawsze wartości rzeczywiste. Z warunku (118) wynika, że każdy niezerowy wektor $\boldsymbol{v}$ ma niezerową długość. W zespolonej przestrzeni wektorowej zachodzi

$$\|\boldsymbol{v}c\| = \|\boldsymbol{v}\|\,|c|\,, \tag{124}$$

gdzie $|c| = (c^*c)^{1/2}$ jest wartością bezwzględną (modułem) liczby zespolonej c, patrz równ. (5). Można stąd wyciągnąć wniosek, że pomnożenie wektora $\boldsymbol{v}$ przez czynnik fazowy $e^{\phi i}$, patrz równ. (15), nie zmienia długości wektora, $\|\boldsymbol{v}e^{\phi i}\| = \|\boldsymbol{v}\|$. W rzeczywistej przestrzeni wektorowej odpowiednikiem tej własności jest równość $\|-\boldsymbol{v}\| = \|\boldsymbol{v}\|$. Wektory $\boldsymbol{v}$ o długości jednostkowej, $\|\boldsymbol{v}\| = 1$, nazywa się *wektorami unormowanymi* (lub, niekiedy, *wersorami*). *Normalizacja* niezerowego wektora $\boldsymbol{v}$ polega na przekształceniu go w wektor unormowany przez pomnożenie przez *czynnik normalizujący* równy $\langle\boldsymbol{v}|\boldsymbol{v}\rangle^{-1/2}$:

$$\boldsymbol{e} = \boldsymbol{v}\,\langle\boldsymbol{v}|\boldsymbol{v}\rangle^{-1/2} = \frac{\boldsymbol{v}}{\sqrt{\langle\boldsymbol{v}|\boldsymbol{v}\rangle}}\,, \tag{125}$$

a otrzymany wektor spełnia warunek $\|\boldsymbol{e}\| = 1$.

2. Można wykazać spełnienie następujących nierówności związanych z iloczynem skalarnym:
   (A) Nierówność Schwartza:

   $$\langle\boldsymbol{v}|\boldsymbol{w}\rangle\langle\boldsymbol{w}|\boldsymbol{v}\rangle \le \langle\boldsymbol{v}|\boldsymbol{v}\rangle\langle\boldsymbol{w}|\boldsymbol{w}\rangle\,, \tag{126}$$

   którą można zapisać także w postaci

   $$|\langle\boldsymbol{v}|\boldsymbol{w}\rangle| \le \|\boldsymbol{v}\|\,\|\boldsymbol{w}\|\,. \tag{127}$$

   (B) Nierówność Minkowskiego (zwana też nierównością trójkąta):

   $$\|\boldsymbol{v} - \boldsymbol{w}\| \le \|\boldsymbol{v}\| + \|\boldsymbol{w}\|\,. \tag{128}$$

3. Dwa wektory $\boldsymbol{v}$ i $\boldsymbol{w}$ nazywamy wzajemnie *ortogonalnymi*, gdy

   $$\langle\boldsymbol{v}|\boldsymbol{w}\rangle = 0\,. \tag{129}$$

   Wektor $\mathbf{0}$ jest ortogonalny do każdego wektora $\boldsymbol{v}$, wliczając siebie samego, patrz równ. (118) i (122).

4. W rzeczywistej przestrzeni wektorowej można zdefiniować kąt $\alpha \in [0, \pi]$, jaki tworzą dwa niezerowe wektory $\boldsymbol{v}$ i $\boldsymbol{w}$. Cosinus tego kąta wynosi

   $$\cos\alpha = \frac{\langle\boldsymbol{v}|\boldsymbol{w}\rangle}{\|\boldsymbol{v}\|\,\|\boldsymbol{w}\|}\,. \tag{130}$$

   Gdy wektory $\boldsymbol{v}$ i $\boldsymbol{w}$ są ortogonalne, to $\alpha = \pi/2$, a gdy jeden wektor jest proporcjonalny do drugiego, np. $\boldsymbol{w} = \boldsymbol{v}c$, $c \ne 0$, to $\alpha = 0$ $(c > 0)$ lub $\alpha = \pi$ $(c < 0)$.

5. W zespolonej przestrzeni wektorowej $\mathbb{C}^M$ (patrz przykład 3 na str. 17), wektorami są macierze o wymiarach $M \times 1$ (macierze kolumnowe), patrz równ. (26), o elementach będących liczbami zespolonymi. W przestrzeni tej *standardowy iloczyn skalarny* definiuje się jako

$$\langle \boldsymbol{a} | \boldsymbol{b} \rangle = \boldsymbol{a}^\dagger \boldsymbol{b} = a_1^* b_1 + a_2^* b_2 + \cdots + a_M^* b_M = \sum_{k=1}^{M} a_k^* b_k \,, \tag{131}$$

gdzie macierz wierszowa $\boldsymbol{a}^\dagger$ jest macierzą hermitowsko sprzężoną względem macierzy kolumnowej $\boldsymbol{a}$, patrz równ. (41) oraz uwaga poniżej tego równania. W rzeczywistej przestrzeni wektorowej $\mathbb{R}^M$ definicja standardowego iloczynu skalarnego przyjmuje postać:

$$\langle \boldsymbol{a} | \boldsymbol{b} \rangle = \boldsymbol{a}^{\mathrm{T}} \boldsymbol{b} = a_1 b_1 + a_2 b_2 + \cdots + a_M b_M = \sum_{k=1}^{M} a_k b_k \,, \tag{132}$$

gdzie macierz wierszowa $\boldsymbol{a}^{\mathrm{T}}$ jest macierzą transponowaną względem macierzy kolumnowej $\boldsymbol{a}$.

6. W przykładzie 4 na str. 17 rozważany był przykład przestrzeni wektorowej odwzorowań $\mathbb{X} \to \mathbb{K}$. W przypadku $\mathbb{X} = \mathbb{R}$ i $\mathbb{K} = \mathbb{C}$ mamy do czynienia z przestrzenią wektorową funkcji f o argumencie rzeczywistym i wartościach zespolonych. Zbiór tych funkcji, które dodatkowo spełniają warunek

$$\int_{-\infty}^{\infty} dx\, f^*(x)\, f(x) = \int_{-\infty}^{\infty} dx\, |f(x)|^2 < \infty \,, \tag{133}$$

tworzy pewną podprzestrzeń we wspomnianej przestrzeni wektorowej. Ta podprzestrzeń pełnej przestrzeni nosi nazwę *przestrzeni wektorowej funkcji całkowalnych w kwadracie*, albo *przestrzeni Hilberta*. Jest to zespolona przestrzeń wektorowa o wymiarze nieskończonym, w której określony jest iloczyn skalarny w postaci całki:

$$\langle f | g \rangle = \int_{-\infty}^{\infty} dx\, f^*(x)\, g(x) \,. \tag{134}$$

Definicję tę można uogólnić na przestrzenie Hilberta odwzorowań $\mathbb{R}^n \to \mathbb{C}$, zastępując w równ. (133) i (134) całki jednokrotne całkami n-krotnymi.
Uwaga: funkcje wielomianowe w przykładzie 4 na str. 20 nie spełniają warunku (133), a więc nie da się w zbiorze tych funkcji określić iloczynu skalarnego opartego na definicji (134).

7. W zespolonej przestrzeni wektorowej $\mathbb{V}$ wymiaru M, w której określony jest iloczyn skalarny, wybieramy pewną baze (104). Dla dowolnej pary wektorów $\boldsymbol{w}$ i $\boldsymbol{w}'$ w tej przestrzeni, przedstawionych w formie kombinacji liniowych (107) wektorów bazy,

można skorzystać z własności iloczynu skalarnego (113), (114), (119) oraz (120), i obliczyć $\langle \boldsymbol{w}|\boldsymbol{w}'\rangle$ jako

$$\langle \boldsymbol{w}|\boldsymbol{w}'\rangle = \langle \sum_{k=1}^{M} \boldsymbol{v}_k c_k | \sum_{l=1}^{M} \boldsymbol{v}_l c'_l \rangle = \sum_{k=1}^{M}\sum_{l=1}^{M} c_k^* S_{kl} c'_l \,, \tag{135}$$

gdzie

$$S_{kl} = \langle \boldsymbol{v}_k|\boldsymbol{v}_l\rangle \,, \tag{136}$$

są elementami tzw. *macierzy nakrywania* $\boldsymbol{S}$ i spełniać muszą warunek $S_{kl} = S_{lk}^*$, który wynika z równ. (115). Macierz $\boldsymbol{S}$ jest więc hermitowską macierzą kwadratową o wymiarach $M \times M$, patrz definicja (61). Dowodzi się, że gdy (104) jest zbiorem wektorów liniowo niezależnych, to zachodzi $\det \boldsymbol{S} > 0$. Co więcej, w takim przypadku każda macierz $\boldsymbol{S}_{(M-m)}$ wymiaru $(M-m) \times (M-m)$, powstała z macierzy $\boldsymbol{S}$ przez usunięcie dowolnych m wierszy i m kolumn ($m = 1, 2, \ldots, M-1$), spełnia warunek $\det \boldsymbol{S}_{(M-m)} > 0$. W związku z tym mówi się, że macierz nakrywania $\boldsymbol{S}$ jest macierzą *dodatnio określoną*. Z warunku $\det \boldsymbol{S} > 0$ wynika, że istnieje macierz odwrotna $\boldsymbol{S}^{-1}$. W rzeczywistej przestrzeni wektorowej, gdzie iloczyn skalarny spełnia warunek (116), macierz $\boldsymbol{S}$ jest rzeczywistą macierzą symetryczną, patrz definicja (60), i jest dodatnio określona.

8. Wyznaczenie macierzy nakrywania $\boldsymbol{S}$ dla danej bazy, patrz równ. (136), wymaga obliczenia iloczynów skalarnych $M(M+1)/2$ par wektorów bazy [korzystamy z tego, że $\boldsymbol{S}$ jest macierzą hermitowską (lub symetryczną)]. Włożony nakład pracy pozwala potem w sposób łatwy, przy wykorzystaniu równ. (135), obliczać iloczyny skalarne dowolnych par wektorów. Taki sposób postępowania jest korzystny, gdy (i) dana przestrzeń wektorowa generowana jest jako przestrzeń kombinacji liniowych pewnej ustalonej bazy wektorów i (ii) bezpośrednie obliczanie iloczynu skalarnego jest dosyć pracochłonne. Typowym przykładem może być przestrzeń wektorowa $\mathbb{V}$ kombinacji liniowych M-elementowego uporządkowanego zbioru funkcji (108); zakładamy, że rozważane funkcje są odwzorowaniami $\mathbb{R} \to \mathbb{C}$, zbiór (108) zawiera funkcje liniowo niezależne, a wektory w przestrzeni $\mathbb{V}$ są funkcjami postaci (109), gdzie $c_k \in \mathbb{C}$ ($\mathbb{V}$ jest więc zespoloną przestrzenią wektorową wymiaru M). Jeśli funkcje bazy (108) spełniają warunek (133), a więc należą do pewnej przestrzeni Hilberta, to można obliczyć elementy macierzy nakrywania dla tej bazy:

$$S_{kl} = \langle \phi_k|\phi_l\rangle = \int_{-\infty}^{\infty} dx\, \phi_k^*(x)\, \phi_l(x) \,, \tag{137}$$

a następnie obliczać iloczyny skalarne $\langle f|g\rangle$ dla funkcji postaci (109) korzystając z przepisu (135).

9. Niech $\boldsymbol{S}$ będzie pewną hermitowską, dodatnio określoną macierzą o wymiarach $M \times M$. W zespolonej przestrzeni wektorowej $\mathbb{C}^M$ (patrz przykład 5 na str. 23), zdefinio-

wać można „niestandardowy” iloczyn skalarny

$$\langle \boldsymbol{a}|\boldsymbol{b}\rangle = \boldsymbol{a}^\dagger \boldsymbol{S}\boldsymbol{b} = \sum_{k=1}^{M}\sum_{l=1}^{M} a_k^* S_{kl} b_l \,. \tag{138}$$

Standardowy iloczyn skalarny (131) w przestrzeni $\mathbb{C}^M$ jest szczególnym przypadkiem powyższej definicji, odpowiadającym $\boldsymbol{S} = \mathbf{1}$. Ogólnie: różne macierze hermitowskie $\boldsymbol{S}$ (dodatnio określone) definiują różne iloczyny skalarne w przestrzeni $\mathbb{C}^M$. W przestrzeni $\mathbb{R}^M$ definicja „niestandardowych” iloczynów skalarnych jest analogiczna.

5.3 Baza ortonormalna przestrzeni wektorowej

W przestrzeni wektorowej $\mathbb{V}$ wymiaru M nad zbiorem liczbowym $\mathbb{K}$ (= $\mathbb{R}$ lub $\mathbb{C}$), w której określony jest iloczyn skalarny, można wybrać bazę, której wektory spełniają dwa warunki:
(i) są parami ortogonalne,
(ii) są unormowane.
Jeśli zachodzi tylko (i), to odpowiednia baza nazywana jest *bazą ortogonalną*. Gdy dana baza spełnia (i) i (ii), to jest nazywana *bazą ortonormalną*.

Gdy uporządkowany zbiór wektorów

$$(\boldsymbol{e}_1, \boldsymbol{e}_2, \ldots, \boldsymbol{e}_M)\,, \tag{139}$$

jest bazą ortonormalną w przestrzeni wektorowej $\mathbb{V}$ wymiaru M, to iloczyny skalarne wektorów tej bazy mają, z definicji, następującą postać:

$$\langle \boldsymbol{e}_k|\boldsymbol{e}_l\rangle = \delta_{kl}\,, \tag{140}$$

gdzie po prawej stronie równania występuje tzw. *symbol Kroneckera*:

$$\delta_{kl} = \begin{cases} 1 & \text{gdy} \quad k = l\,, \\ 0 & \text{gdy} \quad k \neq l\,. \end{cases} \tag{141}$$

Wielkości δ_{kl} można uważać za elementy macierzy jednostkowej $\mathbf{1}$ o wymiarach $M \times M$. Baza ortonormalna jest więc taką szczególną bazą w przestrzeni wektorowej, że odpowiadająca jej macierz całek nakrywania, patrz równ. (136), równa jest macierzy jednostkowej, $\boldsymbol{S} = \mathbf{1}$. Wykażemy dalej, że każdą bazę, dla której $\boldsymbol{S} \neq \mathbf{1}$ (bazę nieortonormalną), przekształcić można w pewną bazę ortonormalną. Pokażemy też, że w danej przestrzeni wektorowej istnieje wiele różnych baz ortonormalnych.

1. W przestrzeni wektorowej $\mathbb{V}$ z iloczynem skalarnym bazy ortonormalne stanowią wyróżnioną klasą baz, a jak pokażemy poniżej, ich zastosowanie bardzo upraszcza opis wektorów w takiej przestrzeni. W tym celu przedstawimy dowolny wektor $\boldsymbol{w} \in \mathbb{V}$ w postaci kombinacji liniowej wektorów bazy ortonormalnej (139):

$$\boldsymbol{w} = \sum_{k=1}^{M} \boldsymbol{e}_k c_k\,. \tag{142}$$

Obliczymy iloczyn skalarny

$$\langle \boldsymbol{e}_l | \boldsymbol{w} \rangle = \langle \boldsymbol{e}_l | \sum_{k=1}^{M} \boldsymbol{e}_k c_k \rangle = \sum_{k=1}^{M} \langle \boldsymbol{e}_l | \boldsymbol{e}_k \rangle c_k = \sum_{k=1}^{M} \delta_{lk} c_k = c_l \, . \tag{143}$$

Jak widać, współczynniki rozwinięcia dowolnego wektora $\boldsymbol{w}$ w bazie ortonormalnej wyznaczyć można obliczając odpowiednie iloczyny skalarne,

$$c_k = \langle \boldsymbol{e}_k | \boldsymbol{w} \rangle \, . \tag{144}$$

2. Gdy baza (139) jest bazą ortonormalną, iloczyn skalarny wektorów $\boldsymbol{w}$ i $\boldsymbol{w}'$, przedstawionych w postaci (142), obliczyć można korzystając z ogólnego równ. (135) i uwzględniając, że $\boldsymbol{S} = \boldsymbol{1}$:

$$\langle \boldsymbol{w} | \boldsymbol{w}' \rangle = \langle \sum_{k=1}^{M} \boldsymbol{e}_k c_k | \sum_{l=1}^{M} \boldsymbol{e}_l c_l' \rangle = \sum_{k=1}^{M} \sum_{l=1}^{M} c_k^* \delta_{kl} c_l' = \sum_{k=1}^{M} c_k^* c_k' \, . \tag{145}$$

3. Zdefiniujemy teraz pewne bardzo ważne odwzorowanie $\mathcal{P}$ naszej M-wymiarowej zespolonej przestrzeni wektorowej $\mathbb{V}$ (z iloczynem skalarnym), w przestrzeń $\mathbb{C}^M$ ze standardowym iloczynem skalarnym (patrz przykład 5 na str. 23):

$$\mathbb{V} \ni \boldsymbol{w} \to \mathcal{P}(\boldsymbol{w}) = \begin{pmatrix} c_1 \\ c_2 \\ \vdots \\ c_M \end{pmatrix} = \boldsymbol{c} \in \mathbb{C}^M \, , \tag{146}$$

gdzie liczby zespolone c_k tworzące macierz kolumnową $\boldsymbol{c}$ są współczynnikami liniowymi w kombinacji liniowej (142), obliczonymi z równ. (144).
Odwzorowanie $\mathcal{P}$ ma następujące własności:
(a) jest wzajemnie jednoznaczne (jest bijekcją),
(b) zachowuje strukturę przestrzeni wektorowych $\mathbb{V}$ i $\mathbb{C}^M$, gdyż zachodzi

$$\mathcal{P}(\boldsymbol{w} + \boldsymbol{w}') = \mathcal{P}(\boldsymbol{w}) + \mathcal{P}(\boldsymbol{w}') = \boldsymbol{c} + \boldsymbol{c}' \, , \tag{147}$$

$$\mathcal{P}(\boldsymbol{w} a) = \mathcal{P}(\boldsymbol{w}) a = \boldsymbol{c} a \, , \tag{148}$$

gdzie wektory $\boldsymbol{w}$ i $\boldsymbol{w}'$ są dowolnymi wektorami z przestrzeni $\mathbb{V}$, zapisanymi w postaci kombinacji liniowych (142), $\boldsymbol{c}$ i $\boldsymbol{c}'$ s'a odpowiadającymi tym wektorom macierzami kolumnowymi z przestrzeni $\mathbb{C}^M$, a a jest dowolną liczbą zespoloną.
(c) konfrontacja równań (145) i (131) wskazuje, że iloczyn skalarny w przestrzeni $\mathbb{V}$ równy jest standardowemu iloczynowi skalarnemu w przestrzeni $\mathbb{C}^M$, czyli zachodzi

$$\langle \boldsymbol{w} | \boldsymbol{w}' \rangle = \langle \mathcal{P}(\boldsymbol{w}) | \mathcal{P}(\boldsymbol{w}') \rangle = \langle \boldsymbol{c} | \boldsymbol{c}' \rangle \, . \tag{149}$$

Uwaga: w powyższym równaniu $\langle \boldsymbol{w} | \boldsymbol{w}' \rangle$ jest iloczynem skalarnym w przestrzeni $\mathbb{V}$, zaś $\langle \mathcal{P}(\boldsymbol{w}) | \mathcal{P}(\boldsymbol{w}') \rangle = \langle \boldsymbol{c} | \boldsymbol{c}' \rangle$ oznacza standardowy iloczyn skalarny w przestrzeni $\mathbb{C}^M$. Odwzorowanie $\mathcal{P}$ pozwala więc reprezentować wektory z abstrakcyjnej M-wymiarowej zespolonej przestrzeni wektorowej $\mathbb{V}$ przez macierze kolumnowe z przestrzeni $\mathbb{C}^M$. Analogicznie konstruuje się reprezentację M-wymiarowej rzeczywistej przestrzeni wektorowej w przestrzeni $\mathbb{R}^M$. Należy jednak odnotować, że odwzorowanie $\mathcal{P}$ zależy od wyboru bazy ortonormalnej (104) w przestrzeni $\mathbb{V}$, czyli że różnym bazom ortonormalnym w $\mathbb{V}$ odpowiadają różne odwzorowania $\mathcal{P}$.

4. W przestrzeni wektorowej $\mathbb{K}^M$ (gdzie $\mathbb{K} = \mathbb{R}$ lub $\mathbb{C}$), w której określono standardowy iloczyn skalarny, patrz równ. (132) lub (131), istnieje pewna wyróżniona baza ortonormalna, zwana *bazą kanoniczną* przestrzeni wektorowej $\mathbb{K}^M$, złożona z wektorów:

$$\underline{\boldsymbol{e}}_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \\ 0 \end{pmatrix}, \quad \underline{\boldsymbol{e}}_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \\ 0 \end{pmatrix}, \quad \cdots, \quad \underline{\boldsymbol{e}}_M = \begin{pmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix}. \tag{150}$$

Jest to szczególna postać bazy (110) przestrzeni wektorowej $\mathbb{K}^M$. Dowolną macierz kolumnowa $\boldsymbol{c} \in \mathbb{K}^M$ o elementach c_k można przedstawić w postaci następującej kombinacji liniowej wektorów bazy kanonicznej:

$$\boldsymbol{c} = \sum_{k=1}^{M} \underline{\boldsymbol{e}}_k c_k \, . \tag{151}$$

Liczby c_k nazywane są *współrzędnymi* wektora $\boldsymbol{c}$ w bazie kanonicznej przestrzeni wektorowej $\mathbb{K}^M$.
Korzystając z własności odwzorowania $\mathcal{P}$ (patrz poprzedni punkt), obliczymy

$$\mathcal{P}(\boldsymbol{w}) = \sum_{k=1}^{M} \mathcal{P}(\boldsymbol{e}_k) c_k = \boldsymbol{c} = \sum_{k=1}^{M} \underline{\boldsymbol{e}}_k c_k \, . \tag{152}$$

Stąd, biorąc szczególny przypadek $\boldsymbol{w} = \boldsymbol{e}_l$ (czyli przyjmując $c_k = \delta_{kl}$, gdzie l jest ustalonym indeksem), dostajemy

$$\mathcal{P}(\boldsymbol{e}_l) = \underline{\boldsymbol{e}}_l \, , \tag{153}$$

co oznacza, że odwzorowanie $\mathcal{P}$ przyporządkowuje wektorom bazy ortonormalnej (139) przestrzeni wektorowej $\mathbb{V}$ wektory kanonicznej bazy ortonormalnej przestrzeni $\mathbb{C}^M$.

5. W M-wymiarowej przestrzeni wektorowej z iloczynem skalarnym wybieramy pewną bazę (104). Wykażemy poniżej, że bazę tę można przekształcić w pewną bazę ortonormalną w drodze konstrukcji zwanej *ortonormalizacją Schmidta.* Pierwszy krok w tej konstrukcji polega na wyznaczeniu wektora $\boldsymbol{e}_1$ poprzez normalizację pierwszego wektora bazy (104):
$$\boldsymbol{e}_1 = \boldsymbol{v}_1 \langle \boldsymbol{v}_1 | \boldsymbol{v}_1 \rangle^{-1/2} , \tag{154}$$
patrz równ. (125). W drugim kroku konstruujemy wektor $\boldsymbol{v}_2'$, ortogonalny do wektora $\boldsymbol{e}_1$, a następnie wyznaczamy wektor $\boldsymbol{e}_2$ poprzez normalizację wektora $\boldsymbol{v}_2'$:
$$\boldsymbol{v}_2' = \boldsymbol{v}_2 - \boldsymbol{e}_1 \langle \boldsymbol{e}_1 | \boldsymbol{v}_2 \rangle , \quad \boldsymbol{e}_2 = \boldsymbol{v}_2' \langle \boldsymbol{v}_2' | \boldsymbol{v}_2' \rangle^{-1/2} . \tag{155}$$
Ogólnie, krok $k+1$ polega na konstrukcji wektora $\boldsymbol{v}_{k+1}'$, ortogonalnego do otrzymanych wcześniej wektorów $\boldsymbol{e}_1, \boldsymbol{e}_2, \ldots, \boldsymbol{e}_k$, a następnie na wyznaczeniu wektora $\boldsymbol{e}_{k+1}$ poprzez normalizację wektora $\boldsymbol{v}_{k+1}'$:
$$\boldsymbol{v}_{k+1}' = \boldsymbol{v}_{k+1} - \sum_{i=1}^{k} \boldsymbol{e}_i \langle \boldsymbol{e}_i | \boldsymbol{v}_{k+1} \rangle , \quad \boldsymbol{e}_{k+1} = \boldsymbol{v}_{k+1}' \langle \boldsymbol{v}_{k+1}' | \boldsymbol{v}_{k+1}' \rangle^{-1/2} . \tag{156}$$
Krok M, w którym wyznaczamy wektor $\boldsymbol{e}_M$, kończy konstrukcję bazy ortonormalnej (139). A więc w każdej przestrzeni wektorowej $\mathbb{V}$ wymiaru M nad zbiorem liczbowym $\mathbb{K}$ ($= \mathbb{R}$ lub $\mathbb{C}$), w której określony jest iloczyn skalarny, istnieje pewna baza ortonormalna (139).
Warto zauważyć, że otrzymana baza ortonormalna zależy nie tylko od zbioru wektorów w bazie wyjściowej (104), ale także od ich uporządkowania. Istnieją także inne niż ortonormalizacja Schmidta metody przekształcania danej bazy w bazę ortonormalną, m.in. tzw. *ortonormalizacja symetryczna* (zwana też *ortonormalizacją Löwdina*).

6. Mając w przestrzeni wektorowej $\mathbb{V}$ wymiaru M bazę ortonormalną (139), utworzymy M nowych wektorów w postaci kombinacji liniowych
$$\boldsymbol{e}_k' = \sum_{l=1}^{M} \boldsymbol{e}_l U_{lk} , \tag{157}$$
gdzie macierz $\boldsymbol{U}$ (o elementach U_{lk}) jest macierzą unitarną, patrz równ. (63), w przypadku zespolonej przestrzeni wektorowej ($\mathbb{K} = \mathbb{C}$), lub rzeczywistą macierzą ortogonalną, patrz równ. (62), w przypadku rzeczywistej przestrzeni wektorowej ($\mathbb{K} = \mathbb{R}$). W przypadku zespolonej przestrzeni wektorowej obliczenie iloczynów skalarnych wektorów (157) daje:
$$\langle \boldsymbol{e}_k' | \boldsymbol{e}_l' \rangle = \langle \sum_{m=1}^{M} \boldsymbol{e}_m U_{mk} | \sum_{n=1}^{M} \boldsymbol{e}_n U_{nl} \rangle = \sum_{m=1}^{M} \sum_{n=1}^{M} U_{mk}^* \delta_{mn} U_{nl} = \sum_{m=1}^{M} U_{mk}^* U_{nl} = \delta_{mn} , \tag{158}$$

gdzie w ostatnia równość wynika z własności macierzy unitarnej, patrz równ. (63). Wynika stąd, że transformacja (157) zastosowana do bazy ortonormalnej (139) daje nową bazę ortonormalną reprezentowaną przez uporządkowany zbiór wektorów

$$(\boldsymbol{e}'_1, \boldsymbol{e}'_2, \ldots, \boldsymbol{e}'_M)\,. \tag{159}$$

Analogiczny wynik otrzymujemy w przypadku rzeczywistej przestrzeni wektorowej. W danej przestrzeni wektorowej każde dwie bazy ortonormalne można powiązać przy pomocy transformacji (157), jest więc tyle różnych baz ortonormalnych, ile jest różnych macierzy $\boldsymbol{U}$ (unitarnych lub rzeczywistych ortogonalnych, zależnie od przypadku).