Maciej Grzesiak

Iloczyn skalarny

1. Iloczyn skalarny wektorów na płaszczyźnie i w przestrzeni

Iloczyn skalarny wektorów $\vec{a}$ i $\vec{b}$ określamy jako

$$\vec{a} \circ \vec{b} = |\vec{a}||\vec{b}| \cos \varphi.$$

Bezpośrednio z definicji iloczynu skalarnego mamy, że $\vec{i} \circ \vec{i} = \vec{j} \circ \vec{j} = 1$ oraz $\vec{i} \circ \vec{j} = 0$. Stąd otrzymujemy wzór

$$\vec{a} \circ \vec{b} = (a_x\vec{i} + a_y\vec{j}) \circ (b_x\vec{i} + b_y\vec{j}) = a_xb_x + a_yb_y.$$

Zatem

$$|\vec{a}||\vec{b}| \cos \varphi = a_xb_x + a_yb_y,$$

więc

$$\cos \varphi = \frac{a_xb_x + a_yb_y}{|\vec{a}| \cdot |\vec{b}|}$$

Iloczyn skalarny wektorów $\vec{a}$ i $\vec{b}$ w przestrzeni określamy tak jak na płaszczyźnie, tj. $\vec{a} \circ \vec{b} = |\vec{a}| \cdot |\vec{b}| \cos \varphi$. Ponieważ

$$\vec{i} \circ \vec{i} = \vec{j} \circ \vec{j} = \vec{k} \circ \vec{k} = 1$$

oraz

$$\vec{i} \circ \vec{j} = \vec{i} \circ \vec{k} = \vec{j} \circ \vec{k} = 0,$$

więc

$$\begin{aligned} \vec{a} \circ \vec{b} &= (a_x\vec{i} + a_y\vec{j} + a_z\vec{k}) \circ (b_x\vec{i} + b_y\vec{j} + b_z\vec{k}) = \\ &= a_xb_x + a_yb_y + a_zb_z. \end{aligned}$$

Przypomniany ”geometryczny” iloczyn skalarny jest inspiracją do zdefiniowania ogólniejszego pojęcia iloczynu w dowolnej przestrzeni liniowej.
W przeciwieństwie do geometrii, teraz **najpierw określimy iloczyn skalarny, a dopiero potem długość wektora**.

2. Iloczyn skalarny: definicja

Iloczyn skalarny określimy osobno dla przestrzeni rzeczywistej i dla przestrzeni zespolonej.

Definicja 1. W przestrzeni rzeczywistej V określony jest *iloczyn skalarny*, jeśli każdej parze wektorów $\mathbf{v}, \mathbf{w} \in V$ przyporządkowana jest liczba rzeczywista, oznaczona przez $\langle \mathbf{v}, \mathbf{w} \rangle$, przy czym przyporządkowanie to ma następujące własności:
1. $\langle \mathbf{v}, \mathbf{w} \rangle = \langle \mathbf{w}, \mathbf{v} \rangle$ (symetria),
2. $\langle \alpha\mathbf{v}, \mathbf{w} \rangle = \alpha\langle \mathbf{v}, \mathbf{w} \rangle$ dla $\alpha \in \mathbb{R}$ (jednorodność),
3. $\langle \mathbf{v}_1 + \mathbf{v}_2, \mathbf{w} \rangle = \langle \mathbf{v}_1, \mathbf{w} \rangle + \langle \mathbf{v}_2, \mathbf{w} \rangle$ (addytywność),
4. dla dowolnego $\mathbf{v} \in V$ jest $\langle \mathbf{v}, \mathbf{v} \rangle \geqslant 0$, przy czym $\langle \mathbf{v}, \mathbf{v} \rangle = 0$ wtedy i tylko wtedy, gdy $\mathbf{v} = 0$.

Wprawdzie zakładamy tylko addytywność i jednorodność ze względu na pierwszą zmienną, ale aksjomat 1 pozwala wywnioskować to samo dla drugiej zmiennej. Np.

$$\langle \mathbf{v}, \alpha \mathbf{w} \rangle = \langle \alpha \mathbf{w}, \mathbf{v} \rangle = \alpha \langle \mathbf{w}, \mathbf{v} \rangle = \alpha \langle \mathbf{v}, \mathbf{w} \rangle$$

Przestrzeń, w której jest określony iloczyn skalarny nazywamy *przestrzenią euklidesową*.
1. Odwzorowanie $\langle \cdot, \cdot \rangle : \mathbb{R}^3 \times \mathbb{R}^3 \to \mathbb{R}$ określone wzorem:

$$\langle \mathbf{v}, \mathbf{w} \rangle = \alpha_1\beta_1 + \alpha_2\beta_2 + \alpha_3\beta_3, \quad \text{gdzie } \mathbf{v} = (\alpha_1, \alpha_2, \alpha_3), \mathbf{w} = (\beta_1, \beta_2, \beta_3)$$

jest iloczynem skalarnym w $\mathbb{R}^3$. Jest to zwykły, znany z kursu geometrii, iloczyn skalarny. Zamiast $\langle \mathbf{v}, \mathbf{w} \rangle = \langle (\alpha_1, \alpha_2, \alpha_3), (\beta_1, \beta_2, \beta_3) \rangle$ można w tym przypadku pisać $(\alpha_1, \alpha_2, \alpha_3) \cdot (\beta_1, \beta_2, \beta_3)$.
2. Ogólniej, wzór:

$$\langle \mathbf{v}, \mathbf{w} \rangle = \alpha_1\beta_1 + \alpha_2\beta_2 + \cdots \alpha_n\beta_n = \sum_{i=1}^{n} \alpha_i\beta_i$$

dla $\mathbf{v} = (\alpha_1, \alpha_2, \cdots \alpha_n)$, $\mathbf{w} = (\beta_1, \beta_2, \cdots \beta_n)$ określa iloczyn skalarny w $\mathbb{R}^n$.
3. W przestrzeni funkcji ciągłych $C(a, b)$ iloczyn skalarny można wprowadzić wzorem:

$$\langle f, g \rangle = \int_a^b f(x)g(x)\, dx.$$

Np. w przestrzeni $C(0, 2\pi)$:

$$\langle \cos x, \sin x \rangle = \int_0^{2\pi} \cos x \sin x\, dx = 0,$$

$$\langle \cos x, \cos x \rangle = \int_0^{2\pi} \cos^2 x\, dx = \pi.$$

Iloczyn skalarny w przestrzeni zespolonej określamy nastepująco.

Definicja 2. W przestrzeni zespolonej V iloczyn skalarny to funkcja $V \times V \to \mathbb{C}$ której wartość na parze wektorów $(\mathbf{v}, \mathbf{w})$ oznaczymy przez $\langle \mathbf{v}, \mathbf{w} \rangle$, przy czym spełnione są następujące własności:
1. $\langle \mathbf{v}, \mathbf{w} \rangle = \overline{\langle \mathbf{w}, \mathbf{v} \rangle}$ (skośna symetria),
2. $\langle \alpha\mathbf{v}, \mathbf{w} \rangle = \alpha \langle \mathbf{v}, \mathbf{w} \rangle$ dla $\alpha \in \mathbb{C}$ (jednorodność),
3. $\langle \mathbf{v}_1 + \mathbf{v}_2, \mathbf{w} \rangle = \langle \mathbf{v}_1, \mathbf{w} \rangle + \langle \mathbf{v}_2, \mathbf{w} \rangle$ (addytywność),
4. dla dowolnego $\mathbf{v} \in V$ jest $\langle \mathbf{v}, \mathbf{v} \rangle \geqslant 0$, przy czym $\langle \mathbf{v}, \mathbf{v} \rangle = 0$ wtedy i tylko wtedy, gdy $\mathbf{v} = 0$.

Iloczyn skalarny w przestrzeni zespolonej nie jest już jednorodny (a więc nie jest także liniowy) ze względu na drugą zmienną. Mamy bowiem

$$\langle \mathbf{v}, \beta\mathbf{w} \rangle = \overline{\langle \beta\mathbf{w}, \mathbf{v} \rangle} = \overline{\beta \langle \mathbf{w}, \mathbf{v} \rangle} = \bar{\beta} \cdot \overline{\langle \mathbf{w}, \mathbf{v} \rangle} = \bar{\beta} \langle \mathbf{v}, \mathbf{w} \rangle.$$

Przestrzeń zespoloną, w której jest określony iloczyn skalarny, nazywamy *przestrzenią unitarną*.
Wzór:

$$\langle \mathbf{v}, \mathbf{w} \rangle = \alpha_1\overline{\beta_1} + \alpha_2\overline{\beta_2} + \cdots \alpha_n\overline{\beta_n} = \sum_{i=1}^{n} \alpha_i\overline{\beta_i}$$

dla $\mathbf{v} = (\alpha_1, \alpha_2, \cdots \alpha_n)$, $\mathbf{w} = (\beta_1, \beta_2, \cdots \beta_n)$ określa iloczyn skalarny w $\mathbb{C}^n$.

3. Norma wektora

Wiadomo, że na płaszczyźnie, tj. w $\mathbb{R}^2$ długość wektora $\mathbf{v} = (\alpha, \beta)$ określa się wzorem:

$$\|\mathbf{v}\| = \sqrt{\alpha^2 + \beta^2}$$

Ale $\langle \mathbf{v}, \mathbf{v} \rangle = \alpha^2 + \beta^2$, więc inaczej:

$$\|\mathbf{v}\| = \sqrt{\langle \mathbf{v}, \mathbf{v} \rangle}$$

Definicja 3. Niech V będzie przestrzenią euklidesową lub unitarną. *Normę (długość) wektora* określamy wzorem

$$\| \mathbf{v} \| = \sqrt{\langle \mathbf{v}, \mathbf{v} \rangle}.$$

Wektory $\mathbf{v}$, $\mathbf{w}$ nazywamy *ortogonalnymi (prostopadłymi)* gdy

$$\langle \mathbf{v}, \mathbf{w} \rangle = 0.$$

Piszemy: $\mathbf{v} \perp \mathbf{w}$.

Własności normy

Twierdzenie 1. *Dla dowolnego skalara α i dowolnych wektorów $\mathbf{v}$, $\mathbf{w}$ mamy:*
1. $\|\alpha \mathbf{v}\| = |\alpha| \|\mathbf{v}\|$,
2. $\|\mathbf{v}\| > 0$ *dla* $\mathbf{v} \neq 0$,
3. $|\langle \mathbf{v}, \mathbf{w} \rangle| \leqslant \|\mathbf{v}\| \cdot \|\mathbf{w}\|$ *(nierówność Schwarza*[1]
4. $\|\mathbf{v} + \mathbf{w}\| \leqslant \|\mathbf{v}\| + \|\mathbf{w}\|$ *(nierówność trójkąta).*

Dowód nierówności Schwarza. Jeśli $\mathbf{w} = 0$, to nierówność jest prawdziwa. Załóżmy więc, że $\mathbf{w} \neq 0$. Dla dowolnego $z \in \mathbb{C}$ mamy

$$\langle \mathbf{v} - z\mathbf{w}, \mathbf{v} - z\mathbf{w} \rangle \geqslant 0,$$

czyli

$$\langle \mathbf{v}, \mathbf{v} \rangle - \bar{z} \langle \mathbf{v}, \mathbf{w} \rangle - z \langle \mathbf{w}, \mathbf{v} \rangle + z\bar{z} \langle \mathbf{w}, \mathbf{w} \rangle \geqslant 0.$$

Przyjmijmy $z = \frac{\langle \mathbf{v}, \mathbf{w} \rangle}{\langle \mathbf{w}, \mathbf{w} \rangle}$.

$$\langle \mathbf{v}, \mathbf{v} \rangle - \frac{\overline{\langle \mathbf{v}, \mathbf{w} \rangle}}{\langle \mathbf{w}, \mathbf{w} \rangle} \langle \mathbf{v}, \mathbf{w} \rangle - \frac{\langle \mathbf{v}, \mathbf{w} \rangle}{\langle \mathbf{w}, \mathbf{w} \rangle} \langle \mathbf{w}, \mathbf{v} \rangle + \frac{\langle \mathbf{v}, \mathbf{w} \rangle}{\langle \mathbf{w}, \mathbf{w} \rangle} \frac{\overline{\langle \mathbf{v}, \mathbf{w} \rangle}}{\langle \mathbf{w}, \mathbf{w} \rangle} \langle \mathbf{w}, \mathbf{w} \rangle \geqslant 0,$$

czyli

$$\|\mathbf{v}\|^2 - \frac{|\langle \mathbf{v}, \mathbf{w} \rangle|^2}{\|\mathbf{w}\|^2} \geqslant 0,$$

więc $|\langle \mathbf{v}, \mathbf{w} \rangle| \leqslant \|\mathbf{v}\| \cdot \|\mathbf{w}\|$.
Dla zwykłego iloczynu skalarnego w $\mathbb{R}^n$ nierówność Schwarza przyjmuje postać:

$$\Big| \sum_{k=1}^{n} x_k y_k \Big| \leqslant \sqrt{\sum_{k=1}^{n} x_k^2} \sqrt{\sum_{k=1}^{n} y_k^2}$$

Natomiast dla iloczynu skalarnego w przestrzeni funkcji ciągłych $C(a, b)$ określonego wzorem $\langle f, g \rangle = \int_a^b f(x)g(x)\, dx$ wygląda ona tak

$$\Big| \int_a^b f(x)g(x)\, dx \Big| \leqslant \sqrt{\int_a^b f^2(x)\, dx} \sqrt{\int_a^b g^2(x)\, dx}.$$

[1] Herman Schwarz (1843-1921)

Dowód nierówności trójkąta.

$$\|\mathbf{v}+\mathbf{w}\|^2 = \langle \mathbf{v}+\mathbf{w}, \mathbf{v}+\mathbf{w}\rangle = \|\mathbf{v}\|^2 + 2\mathrm{re}\,\langle \mathbf{v}, \mathbf{w}\rangle + \|\mathbf{w}\|^2,$$

ale

$$2\mathrm{re}\,\langle \mathbf{v}, \mathbf{w}\rangle \leqslant 2|\langle \mathbf{v}, \mathbf{w}\rangle| \leqslant 2\|\mathbf{v}\| \cdot \|\mathbf{w}\|.$$

Stąd $\|\mathbf{v}+\mathbf{w}\|^2 \leqslant (\|\mathbf{v}\| + \|\mathbf{w}\|)^2$. Obliczając pierwiastki otrzumujemy nierówność trójkąta.
Nierówność trójkąta zawdzięcza swą nazwę oczywistej interpretacji geometrycznej.
W geometrii mamy też *regułę równoległoboku*: suma kwadratów długości czterech boków równoległoboku równa jest sumie kwadratów długości dwóch przekątnych. Jej wersja dla przestrzeni euklidesowej bądź unitarnej nazywa się *tożsamością równoległoboku* i ma postać:

$$\|\mathbf{v}+\mathbf{w}\|^2 + \|\mathbf{v}-\mathbf{w}\|^2 = 2(\|\mathbf{v}\|^2 + \|\mathbf{w}\|^2)$$

Dowód tożsamości równoległoboku.

$$\begin{aligned} \|\mathbf{v}+\mathbf{w}\|^2 + \|\mathbf{v}-\mathbf{w}\|^2 &= \langle \mathbf{v}+\mathbf{w}, \mathbf{v}+\mathbf{w}\rangle + \langle \mathbf{v}-\mathbf{w}, \mathbf{v}-\mathbf{w}\rangle = \\ &= \|\mathbf{v}\|^2 + \|\mathbf{w}\|^2 + \langle \mathbf{v}, \mathbf{w}\rangle + \langle \mathbf{w}, \mathbf{v}\rangle + \|\mathbf{v}\|^2 + \|\mathbf{w}\|^2 - \langle \mathbf{v}, \mathbf{w}\rangle - \langle \mathbf{w}, \mathbf{v}\rangle = \\ &= 2(\|\mathbf{v}\|^2 + \|\mathbf{w}\|^2) \end{aligned}$$

4. Baza ortogonalna.

Definicja 4. Dwa wektory $\mathbf{v}$ i $\mathbf{w}$ nazywają się *ortogonalnymi*, gdy $\langle \mathbf{v}, \mathbf{w}\rangle = 0$. Zbiór $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ nazywa się *ortogonalnym zbiorem wektorów*, gdy:

1. wszystkie wektory $\mathbf{v}_i$, $i = 1, 2, \ldots, n$ są niezerowe,
2. $\langle \mathbf{v}_i, \mathbf{v}_j\rangle = 0$ dla $i \neq j$.

Ortogonalny zbiór wektorów, w którym wszystkie wektory mają długość jeden, nazywa się zbiorem *ortonormalnym*.
Zadanie. Wykazać, że w $\mathbb{R}^4$ zbiór wektorów $(2, -1, 4, 5)$, $(0, -1, 1, -1)$, $(0, 3, 2, -1)$ jest ortogonalny.
Jak zwykle, gdy iloczyn nie jest wyraźnie określony, przyjmujemy, że chodzi o standardowy iloczyn skalarny. Obliczamy iloczyny:

$$(2, -1, 4, 5) \cdot (0, -1, 1, -1) = 2 \cdot 0 + (-1) \cdot (-1) + 4 \cdot 1 + 5 \cdot (-1) = 0,$$

$$(2, -1, 4, 5) \cdot (0, 3, 2, -1) = 2 \cdot 0 + (-1) \cdot 3 + 4 \cdot 2 + 5 \cdot (-1) = 0,$$

$$(0, -1, 1, -1) \cdot (0, 3, 2, -1) = 0 \cdot 0 + (-1) \cdot 3 + 1 \cdot 2 + (-1) \cdot (-1) = 0.$$

Twierdzenie 2. *Niech $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ będzie ortogonalnym zbiorem wektorów. Wtedy*

1. *zbiór $\{\lambda_1\mathbf{v}_1, \lambda_2\mathbf{v}_2, \ldots, \lambda_n\mathbf{v}_n\}$ jest także ortogonalny dla dowolnych skalarów $\lambda_i \neq 0$,*
2. *zbiór $\{\frac{1}{\|\mathbf{v}_1\|}\mathbf{v}_1, \frac{1}{\|\mathbf{v}_2\|}\mathbf{v}_2, \ldots, \frac{1}{\|\mathbf{v}_n\|}\mathbf{v}_n\}$ jest ortonormalny.*

Ważne jest następujące twierdzenie.

Twierdzenie 3. *Każdy ortogonalny zbiór wektorów jest liniowo niezależny.*

D o w ó d. Niech $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ będzie ortogonalny i przypuśćmy, że

$$\mathbf{v} = \lambda_1\mathbf{v}_1 + \lambda_2\mathbf{v}_2 + \cdots + \lambda_n\mathbf{v}_n = \mathbf{0}.$$

Obliczamy iloczyn skalarny wektorów $\mathbf{v}$, $\mathbf{v}_1$:

$$\begin{aligned} 0 &= \langle 0, \mathbf{v}_1\rangle = \langle \lambda_1\mathbf{v}_1 + \lambda_2\mathbf{v}_2 + \cdots + \lambda_n\mathbf{v}_n, \mathbf{v}_1\rangle = \\ &= \lambda_1\langle \mathbf{v}_1, \mathbf{v}_1\rangle + \lambda_2\langle \mathbf{v}_2, \mathbf{v}_1\rangle + \cdots + \lambda_n\langle \mathbf{v}_n, \mathbf{v}_1\rangle = \\ &= \lambda_1\|\mathbf{v}_1\|^2 + 0 + \cdots + 0 = \lambda_1\|\mathbf{v}_1\|^2. \end{aligned}$$

Stąd $\lambda_1 = 0$ i podobnie $\lambda_2 = \lambda_3 = \ldots = \lambda_n = 0$.

Definicja 5. Bazę przestrzeni V składającą się z wektorów ortogonalnych nazywamy *bazą ortogonalną.*

Najważniejszą cechą tej bazy jest, że współrzędne wektora w bazie ortogonalnej są łatwe do wyznaczenia. Istnieją proste wzory, które podamy w następującym twierdzeniu.

Twierdzenie 4. (o rozwinięciu) *Niech $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ będzie bazą ortogonalną przestrzeni V z iloczynem skalarnym $\langle \cdot, \cdot \rangle$. Jeśli $\mathbf{v}$ jest dowolnym wektorem przestrzeni V, to:*

$$\mathbf{v} = \frac{\langle \mathbf{v}, \mathbf{v}_1 \rangle}{\|\mathbf{v}_1\|^2}\mathbf{v}_1 + \frac{\langle \mathbf{v}, \mathbf{v}_2 \rangle}{\|\mathbf{v}_2\|^2}\mathbf{v}_2 + \cdots + \frac{\langle \mathbf{v}, \mathbf{v}_n \rangle}{\|\mathbf{v}_n\|^2}\mathbf{v}_n$$

jest przedstawieniem $\mathbf{v}$ jako kombinacji liniowej wektorów bazy.

D o w ó d. Wektory $\mathbf{v}_i$ stanowią bazę, więc

$$\mathbf{v} = \lambda_1 \mathbf{v}_1 + \lambda_2 \mathbf{v}_2 + \cdots + \lambda_n \mathbf{v}_n$$

dla pewnych skalarów λ_i.
Mnożąc tę równość skalarnie przez $\mathbf{v}_i$ otrzymujemy

$$\langle \mathbf{v}, \mathbf{v}_i \rangle = \lambda_i \|\mathbf{v}_i\|^2,$$

więc

$$\lambda_i = \frac{\langle \mathbf{v}, \mathbf{v}_i \rangle}{\|\mathbf{v}_i\|^2}.$$

Przykład. Wykazać, że $B = \{(1,-1,3), (-2,1,1), (4,7,1)\}$ jest bazą ortogonalną przestrzeni $\mathbb{R}^3$ i przedstawić wektor $\mathbf{x} = (\xi_1, \xi_2, \xi_3)$ w tej bazie.
Obliczamy:

$$(1,-1,3) \cdot (-2,1,1) = 0, \ (1,-1,3) \cdot (4,7,1) = 0, \ (-2,1,1) \cdot (4,7,1) = 0.$$

Zatem wektory są parami ortogonalne, więc tworzą bazę. Obliczamy iloczyny skalarne:

$$(\xi_1, \xi_2, \xi_3) \cdot (1,-1,3) = \xi_1 - \xi_2 + 3\xi_3,$$

$$(\xi_1, \xi_2, \xi_3) \cdot (-2,1,1) = -2\xi_1 + \xi_2 + \xi_3,$$

$$(\xi_1, \xi_2, \xi_3) \cdot (4,7,1) = 4\xi_1 + 7\xi_2 + \xi_3,$$

a następnie normy wektorów bazy. Są to kolejno $\sqrt{11}$, $\sqrt{6}$, $\sqrt{66}$. Zatem

$$\mathbf{x} = \left(\frac{\xi_1 - \xi_2 + 3\xi_3}{11}, \frac{-2\xi_1 + \xi_2 + \xi_3}{6}, \frac{4\xi_1 + 7\xi_2 + \xi_3}{66} \right).$$

Jeżeli dana baza nie jest ortogonalna, to można ją zastąpić bazą ortogonalną stosując algorytm nazywany *procedurą ortogonalizacji Grama – Schmidta*[2]. Nie będziemy jej jednak opisywać.

5. Macierz ortogonalna

W tym rozdziale zakładamy, że ciałem skalarów jest $\mathbb{R}$.

Twierdzenie 5. *Dla dowolnej macierzy $\mathbf{A}$ stopnia n następujące warunki są równoważne:*
1) $\mathbf{A}$ jest odwracalna i $\mathbf{A}^{-1} = \mathbf{A}^T$,
2) wiersze macierzy $\mathbf{A}$ są ortonormalne,
3) kolumny macierzy $\mathbf{A}$ są ortonormalne.

[2] Jorgen Gram (1850-1916), Erhard Schmidt (1876-1959)

D o w ó d. Pierwszy warunek jest równoważny równości

$$\mathbf{A}\mathbf{A}^T = \mathbf{A}^T\mathbf{A} = \mathbf{I}. \tag{1}$$

Niech $\mathbf{v}_1$, $\mathbf{v}_2$, ..., $\mathbf{v}_n$ oznaczają wiersze macierzy $\mathbf{A}$. Wtedy $\mathbf{v}_j^T$ jest j-tą kolumną macierzy $\mathbf{A}^T$, więc elementem (i, j) macierzy $\mathbf{A}\mathbf{A}^T$ jest $\langle \mathbf{v}_i, \mathbf{v}_j \rangle$.
Zatem warunek $\mathbf{A}\mathbf{A}^T = \mathbf{I}$ znaczy, że

$$\langle \mathbf{v}_i, \mathbf{v}_j \rangle = \begin{cases} 0 & \text{gdy } i \neq j, \\ 1 & \text{gdy } i = j, \end{cases}$$

więc (1) $\Leftrightarrow$ (2).
Podobnie dowodzi się, że (1) $\Leftrightarrow$ (3).
Macierz stopnia n nazywamy *ortogonalną*, jeśli spełnia jeden (a więc i wszystkie) z powyższych warunków. Np. macierz

$$\mathbf{A} = \begin{bmatrix} \frac{1}{2} & \frac{\sqrt{3}}{2} \\ -\frac{\sqrt{3}}{2} & \frac{1}{2} \end{bmatrix}$$

jest ortogonalna, bo $\|\mathbf{v}_1\| = \|\mathbf{v}_2\| = 1$ i $\langle \mathbf{v}_1, \mathbf{v}_2 \rangle = 0$. Zatem

$$\mathbf{A}^{-1} = \mathbf{A}^T = \begin{bmatrix} \frac{1}{2} & -\frac{\sqrt{3}}{2} \\ \frac{\sqrt{3}}{2} & \frac{1}{2} \end{bmatrix}.$$

Ogólniej, macierz obrotu

$$\begin{bmatrix} \cos\varphi & -\sin\varphi \\ \sin\varphi & \cos\varphi \end{bmatrix}$$

jest ortogonalna dla dowolnego φ.
Natomiast macierz

$$\mathbf{B} = \begin{bmatrix} \frac{1}{2} & \frac{\sqrt{3}}{2} \\ \frac{\sqrt{3}}{2} & \frac{1}{2} \end{bmatrix}$$

nie jest ortogonalna, bo $\langle \mathbf{v}_1, \mathbf{v}_2 \rangle = \frac{\sqrt{3}}{2} \neq 0$.
Przykład. Niech $\pi = (\pi_1, \pi_2, \ldots \pi_n)$ będzie permutacją liczb $\{1, 2, \ldots, n\}$. Określamy macierz $\mathbf{A} = [a_{ij}]$ wzorem

$$a_{ij} = \begin{cases} 1, & \text{gdy } j = \pi_i \\ 0, & \text{gdy } j \neq \pi_i \end{cases}$$

Macierz $\mathbf{A}$ nazywamy *macierzą permutacyjną*.
Np. gdy $\pi = (2, 3, 1)$, to

$$\mathbf{A} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix}$$

Dla dowolnej macierzą permutacyjnej $\mathbf{A} = [a_{ij}]$ mamy $\mathbf{A}\mathbf{A}^T = \mathbf{I}$, ponieważ w każdym wierszu jest dokładnie jeden element różny od 0. Zatem $\mathbf{A}$ jest macierzą ortogonalną. W szczególności $\mathbf{A}^{-1} = \mathbf{A}^T$.
Z warunku (1) można wywnioskować dwie ważne własności macierzy ortogonalnych.

Twierdzenie 6. *1.) Jeżeli macierz* $\mathbf{A}$ *jest ortogonalna, to* $\det \mathbf{A} = 1$ *lub* $\det \mathbf{A} = -1$.
2.)Jeżeli $\mathbf{A}$ *i* $\mathbf{B}$ *są macierzami ortogonalnymi tego samego stopnia, to* $\mathbf{AB}$ *też jest macierzą ortogonalną.*

D o w ó d. 1. Ponieważ zawsze $\det \mathbf{A} = \det \mathbf{A^T}$ a dla macierzy ortogonalnej $\det \mathbf{A}^{-1} = \det \mathbf{A^T}$, więc

$$1 = \det \mathbf{I} = \det \mathbf{A}\mathbf{A^{-1}} = \det \mathbf{A}\mathbf{A^T} = \det \mathbf{A} \cdot \det \mathbf{A^T} = (\det \mathbf{A})^2.$$

Stąd $\det \mathbf{A} = \pm 1$.

2. Mamy

$$(\mathbf{AB})(\mathbf{AB})^T = (\mathbf{AB})\mathbf{B}^T\mathbf{A}^T = \mathbf{A}(\mathbf{BB}^T)\mathbf{A}^T = \mathbf{AA^T} = 1$$

a więc $(\mathbf{AB})^{-1} = (\mathbf{AB})^T$.
Uwaga. Podana definicja macierzy ortogonalnej jest wygodna w tym sensie, że łatwo jest sprawdzić, czy dana macierz jest ortogonalna. Geometryczny sens jest taki, że przekształcenie określone macierzą ortogonalną zachowuje długość wektora, tzn. $\|\mathbf{Av}\| = \|\mathbf{v}\|$. Takie przekształcenie nazywamy *izometrią*.
W szczególności łatwo jest interpretować macierze ortogonalne stopnia 2. Jeżeli $\det \mathbf{A} = 1$, to $\mathbf{A}$ jest macierzą obrotu płaszczyzny, a jeżeli $\mathbf{A} = -1$, to $\mathbf{A}$ określa symetrię płaszczyzny względem pewnej prostej.

6. Ortogonalna diagonalizacja macierzy symetrycznych

Diagonalizacją macierzy $\mathbf{A}$ nazywamy znalezienie macierzy nieosobliwej $\mathbf{P}$ takiej, że macierz $\mathbf{P}^{-1}\mathbf{AP}$ jest diagonalna. Pokażemy, że jeśli macierz $\mathbf{A}$ jest symetryczna, to zawsze można znaleźć ortogonalną macierz diagonalizującą $\mathbf{P}$.
Macierz symetryczną określa warunek $\mathbf{A} = \mathbf{A}^T$.
Dla macierzy 1-kolumnowych $\mathbf{p}$, $\mathbf{q}$ iloczyn skalarny określamy wzorem

$$\langle \mathbf{p}, \mathbf{q}\rangle = \mathbf{p}^T \cdot \mathbf{q}$$

Twierdzenie 7. *Niech* $\mathbf{A}$ *będzie rzeczywistą macierzą symetryczną stopnia* n*,* $\mathbf{u}, \mathbf{v} \in \mathbb{R}^n$*. Wtedy*

$$\langle \mathbf{Au}^T, \mathbf{v}^T\rangle = \langle \mathbf{u}^T, \mathbf{Av}^T\rangle.$$

D o w ó d.

$$\begin{aligned}\langle \mathbf{Au}^T, \mathbf{v}^T\rangle &= \big(\mathbf{Au}^T\big)^T \cdot \mathbf{v}^T = \big((\mathbf{u}^T)^T\mathbf{A}^T\big) \cdot \mathbf{v}^T = \\ &= (\mathbf{u}^T)^T\big(\mathbf{A} \cdot \mathbf{v}^T\big) = \langle \mathbf{u}^T, \mathbf{Av}^T\rangle.\end{aligned}$$

Twierdzenie 8. *Dla dowolnej macierzy symetrycznej* $\mathbf{A}$ *wektory własne odpowiadające różnym wartościom własnym są ortogonalne.*

D o w ó d. Niech $\mathbf{Au}^T = \lambda\mathbf{u}^T$, $\mathbf{Av}^T = \mu\mathbf{v}^T$. Wtedy:

$$\begin{aligned}\lambda\langle \mathbf{u}^T, \mathbf{v}^T\rangle &= \langle \lambda\mathbf{u}^T, \mathbf{v}^T\rangle = \langle \mathbf{Au}^T, \mathbf{v}^T\rangle = \langle \mathbf{u}^T, \mathbf{Av}^T\rangle = \\ &= \langle \mathbf{u}^T, \mu\mathbf{v}^T\rangle = \mu\langle \mathbf{u}^T, \mathbf{v}^T\rangle,\end{aligned}$$

więc

$$(\lambda - \mu)\langle \mathbf{u}^T, \mathbf{v}^T\rangle = 0,$$

a stąd

$$\langle \mathbf{u}^T, \mathbf{v}^T\rangle = 0$$

Macierz $\mathbf{A}$ nazywamy *ortogonalnie diagonalizowalną* gdy można znaleźć macierz ortogonalną $\mathbf{P}$ taką, że $\mathbf{P}^{-1}\mathbf{AP}$ jest diagonalna.

Twierdzenie 9. (spektralne) *Niech* $\mathbf{A}$ *będzie macierzą kwadratową stopnia* n*. Następujące warunki są równoważne:*
a) $\mathbf{A}$ *ma ortonormalny zbiór wektorów własnych,*
b) $\mathbf{A}$ *jest ortogonalnie diagonalizowalna,*
c) $\mathbf{A}$ *jest symetryczna.*

Przykład. Dla macierzy

$$\mathbf{A} = \begin{bmatrix} 1 & -2 & 0 \\ -2 & 0 & 2 \\ 0 & 2 & -1 \end{bmatrix}$$

znaleźć macierz ortogonalną $\mathbf{P}$ taką, że $\mathbf{P}^{-1}\mathbf{AP}$ jest diagonalna.
Mamy

$$c_{\mathbf{A}}(\lambda) = \begin{vmatrix} 1-\lambda & -2 & 0 \\ -2 & -\lambda & 2 \\ 0 & 2 & -1-\lambda \end{vmatrix} = \lambda^3 - 9\lambda = \lambda(\lambda-3)(\lambda+3).$$

Wartościami własnymi są $\lambda = 0, 3, -3$. Odpowiednie wektory własne:

$$\mathbf{v}_1 = (2,1,2),\ \mathbf{v}_2 = (-2,2,1),\ \mathbf{v}_3 = (1,2,-2)$$

są ortogonalne.
Ponieważ $\|\mathbf{v}_1\| = \|\mathbf{v}_2\| = \|\mathbf{v}_3\| = 3$, więc $\frac{1}{3}\mathbf{v}_1,\ \frac{1}{3}\mathbf{v}_2,\ \frac{1}{3}\mathbf{v}_3$ są wektorami ortonormalnymi. Stąd macierz

$$\mathbf{P} = \frac{1}{3}\begin{bmatrix} 2 & -2 & 1 \\ 1 & 2 & 2 \\ 2 & 1 & -2 \end{bmatrix}$$

jest ortogonalna (czyli $\mathbf{P}^{-1} = \mathbf{P}^T$) oraz

$$\mathbf{P}^T\mathbf{AP} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & -3 \end{bmatrix}$$

na podstawie algorytmu diagonalizacji.