# Fizyka dla firm — Matematyka

## 14. Bardzo dużo algebry

P. F. Góra

http://th-www.if.uj.edu.pl/zfs/gora/

26 listopada 2020

Plan:

- Iloczyn skalarny
- Macierze ortogonalne
- Wartości i wektory własne
- Macierze symetryczne, rzeczywiste
- Twierdzenie spektralne i funkcje macierzy
- Potęgi macierzy
- Funkcja wykładnicza macierzy
- Sprzężenie hermitowskie i iloczyn skalarny w $\mathbb{C}^n$
- Macierze unitarne i hermitowskie

## Iloczyn skalarny

Przypomnieliśmy klasyczną definicję iloczynu skalarnego

$$\vec{a} \circ \vec{b} = |\vec{a}| \cdot |\vec{b}| \cdot \cos \measuredangle(\vec{a}, \vec{b}) \,. \tag{1}$$

Pokazaliśmy, że w przestrzeni $\mathbb{R}^n$ można to uogólnić do sumy iloczynów po współrzędnych

$$\mathbf{x} \circ \mathbf{y} = \sum_{i=1}^{n} x_i y_i \,. \tag{2}$$

Teraz poznamy najbardziej ogólną definicję iloczynu skalarnego, której tamte dwie są szczególnymi przypadkami.

## Iloczyn skalarny — ogólna definicja

Niech $V$ będzie przestrzenią wektorową nad $\mathbb{R}$. Iloczynem skalarnym nazywam funkcję $V \times V \to \mathbb{R}$ o następujących własnościach:

1. Symetria: $\forall \mathbf{x}, \mathbf{y} \in V : \mathbf{x} \circ \mathbf{y} = \mathbf{y} \circ \mathbf{x}$.
2. Rozdzielność względem dodawania wektorów:
   $\forall \mathbf{x}, \mathbf{y}, \mathbf{z} \in V : (\mathbf{x} + \mathbf{y}) \circ \mathbf{z} = \mathbf{x} \circ \mathbf{z} + \mathbf{y} \circ \mathbf{z}$.
3. Zgodność z mnożeniem przez skalar:
   $\forall \mathbf{x}, \mathbf{y} \in V,\ \alpha \in \mathbb{R} : (\alpha \cdot \mathbf{x}) \circ \mathbf{y} = \alpha \cdot (\mathbf{x} \circ \mathbf{y})$.
4. Brak dzielników zera: $(\forall \mathbf{x} \in V : \mathbf{x} \circ \mathbf{y} = 0) \Rightarrow \mathbf{y} = \mathbf{0}$.
5. Dodatnia określoność: $\forall \mathbf{x} \in V,\ \mathbf{x} \neq \mathbf{0} : \mathbf{x} \circ \mathbf{x} > 0$.

## Iloczyn skalarny w $\mathbb{R}^n$

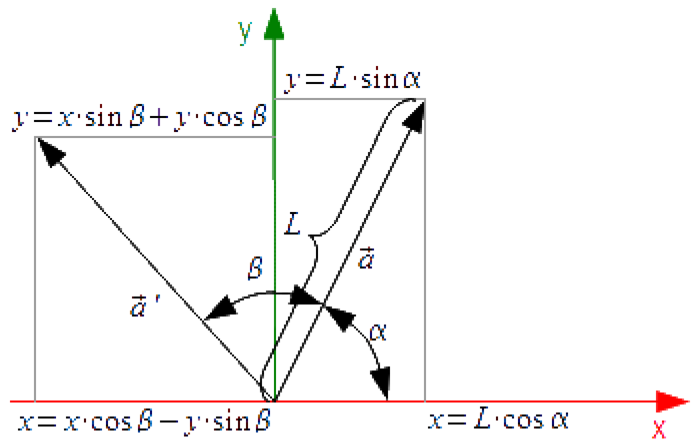
Pamiętając, że $n$-wierszowe macierze jednokolumnowe interpretujemy jak wektory w $\mathbb{R}^n$, możemy powiedzieć, że w tej przestrzeni iloczynem skalarnym jest

$$\mathbf{x} \circ \mathbf{y} = \mathbf{x}^T \mathbf{y} = \sum_{i=1}^{n} x_i y_i \,. \tag{3}$$

Definicja ta spełnia wszystkie warunki podane wyżej. Jest to zgodne ze wzorem (2). Zauważmy, że wektor transponowany $\mathbf{x}^T$ odpowiada jednowierszowej macierzy o $n$ kolumnach.

## Obrót wektora na płaszczyźnie




Dany jest pewien wektor na płaszczyźnie, o współrzędnych $[x, y]$. Chcemy go obrócić o kąt $\beta$. Nowe (po obrocie) współrzędne oznaczymy $[x', y']$. Jak je obliczyć?

Z rysunku widać, że obrócony wektor tworzy z osią $OX$ kąt $\alpha+\beta$, zatem

$$
\begin{aligned}
x' &= L\cos(\alpha+\beta) = L\cos\alpha\cos\beta - L\sin\alpha\sin\beta \\
&= x\cos\beta - y\sin\beta \qquad \text{(4a)} \\
y' &= L\sin(\alpha+\beta) = L\sin\alpha\cos\beta + L\cos\alpha\sin\beta \\
&= y\cos\beta + x\sin\beta \qquad \text{(4b)}
\end{aligned}
$$

gdyż $x = L\cos\beta$, $y = L\sin\beta$. Wyrażenia (4) możemy zapisać macierzowo

$$
\begin{bmatrix} x' \\ y' \end{bmatrix} = \begin{bmatrix} \cos\beta & -\sin\beta \\ \sin\beta & \cos\beta \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} \tag{5}
$$

Macierz występującą we wzorze (5) nazywam macierzą obrotu płaskiego. Oznaczmy ją $\mathbf{O}$.

## Własności macierzy obrotu

$$\det \mathbf{O} = \begin{vmatrix} \cos\beta & -\sin\beta \\ \sin\beta & \cos\beta \end{vmatrix} = \cos^2\beta - (-\sin^2\beta) = 1 \tag{6}$$

Gdybyśmy chcieli obrócić wektor o kąt $-\beta$, to z parzystości funkcji trygonometrycznych od razu widzimy, że macierzą obrotu byłaby macierz

$$\begin{bmatrix} \cos\beta & \sin\beta \\ -\sin\beta & \cos\beta \end{bmatrix} = \mathbf{O}^T \tag{7}$$

Intuicja podpowiada nam, że obrót najpierw o kąt $\beta$, a potem o kąt $-\beta$ łącznie powinny się znosić, czyli dawać transformację tożsamościową. Rzeczywiście,

$$\begin{aligned}
\mathbf{O}\mathbf{O}^T &= \begin{bmatrix} \cos\beta & -\sin\beta \\ \sin\beta & \cos\beta \end{bmatrix} \begin{bmatrix} \cos\beta & \sin\beta \\ -\sin\beta & \cos\beta \end{bmatrix} \\
&= \begin{bmatrix} \cos^2\beta + \sin^2\beta & \cos\beta\sin\beta - \sin\beta\cos\beta \\ \sin\beta\cos\beta - \cos\beta\sin\beta & \sin^2\beta + \cos^2\beta \end{bmatrix} \\
&= \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \mathbb{I}
\end{aligned} \tag{8}$$

i podobnie $\mathbf{O}^T\mathbf{O} = \mathbb{I}$. Widzimy, że macierz obrotu posiada *własność ortogonalności*:

$$\mathbf{O}\mathbf{O}^T = \mathbf{O}^T\mathbf{O} = \mathbb{I} \tag{9}$$

## Obroty w 3D

Obrót w przestrzeni $\mathbb{R}^3$ można przedstawić jako złożenie obrotów w trzech płaszczyznach o określone kąty (kąty Eulera):

Obrót wokół osi $x$ w płaszczyźnie $yz$:

$$\mathbf{O}_x = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & -\sin\theta \\ 0 & \sin\theta & \cos\theta \end{bmatrix} \tag{10a}$$

Obrót wokół osi $y$ w płaszczyźnie $xz$:

$$\mathbf{O}_y = \begin{bmatrix} \cos\varphi & 0 & \sin\varphi \\ 0 & 1 & 0 \\ -\sin\varphi & 0 & \cos\varphi \end{bmatrix} \tag{10b}$$

Obrót wokół osi $z$ w płaszczyźnie $xy$:

$$\mathbf{O}_z = \begin{bmatrix} \cos\psi & -\sin\psi & 0 \\ \sin\psi & \cos\psi & 0 \\ 0 & 0 & 1 \end{bmatrix} \tag{10c}$$

Każda z tych macierzy ma własność ortogonalności: $\mathbf{O}_x^T\mathbf{O}_x = \mathbb{I}$, $\mathbf{O}_y^T\mathbf{O}_y = \mathbb{I}$, $\mathbf{O}_z^T\mathbf{O}_z = \mathbb{I}$. Podobnie — ich złożenia. Na przykład $(\mathbf{O}_x\mathbf{O}_y\mathbf{O}_z)^T\mathbf{O}_x\mathbf{O}_y\mathbf{O}_z = \mathbf{O}_z^T\mathbf{O}_y^T\mathbf{O}_x^T\mathbf{O}_x\mathbf{O}_y\mathbf{O}_z = \mathbf{O}_z^T\mathbf{O}_y^T\mathbf{O}_y\mathbf{O}_z = \mathbf{O}_z^T\mathbf{O}_z = \mathbb{I}$.

## Macierze ortogonalne

Uogólniając obroty w $\mathbb{R}^2$ i $\mathbb{R}^3$ macierzami ortogonalnymi nazywam macierze, dla których

$$\mathbf{O}^T\mathbf{O} = \mathbf{O}\mathbf{O}^T = \mathbb{I}\,. \tag{11}$$

## Nie tylko obroty

Nie tylko macierze obrotu spełniają własność (11). Własność tę posiadają także macierze odbicia, na przykład

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \tag{12}$$

— ta macierz zmiania znak drugiej składowej wektora — oraz macierze permutacji, na przykład

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix} \tag{13}$$

— ta macierz permutuje drugą i trzecią składową wektora.

## Własności macierzy ortogonalnych

$\mathbf{O}$ oznacza macierz ortogonalną.

1. $\mathbf{O}^{-1} = \mathbf{O}^T$.
2. $\det\mathbf{O} = \pm 1$. Istotnie, $1 = \det\mathbb{I} = \det\left(\mathbf{O}^T\mathbf{O}\right) = \left(\det\mathbf{O}^T\right)(\det\mathbf{O}) = (\det\mathbf{O})^2$.
3. Macierze ortogonalne w $\mathbb{R}^n$ tworzą grupę. Działaniem grupowym jest mnożenie macierzy.
4. Kolumny macierzy ortogonalnej są unormowane i wzajemnie prostopadłe. Kolumny macierzy ortogonalnej tworzą bazę w $\mathbb{R}^n$.
5. Wyrażenie $\mathbf{v}' = \mathbf{O}\mathbf{v}$ można interpretować jako przedstawienie wektora $\mathbf{v}$ w bazie kolumn macierzy $\mathbf{O}$.
6. Ortogonalna *transformację podobieństwa* $\mathbf{A}' = \mathbf{O}^T\mathbf{A}\mathbf{O}$ można interpretować jako przedstawienie elementów operatora $\mathbf{A}$ w bazie kolumn macierzy $\mathbf{O}$.

7. Macierz ortogonalna zachowuje iloczyn skalarny. Istotnie,
$(\mathbf{Ou}) \circ (\mathbf{Ov}) = (\mathbf{Ou})^T (\mathbf{Ou}) = \mathbf{u}^T \mathbf{O}^T \mathbf{Ov} = \mathbf{u}^T \mathbf{v} = \mathbf{u} \circ \mathbf{v}.$

## Zagadnienie własne

**Definicja**: Niech $\mathbf{A} \in \mathbb{C}^{n\times n}$. Liczbę $\lambda \in \mathbb{C}$ nazywam *wartością własną macierzy* $\mathbf{A}$, jeżeli istnieje niezerowy wektor $\mathbf{x} \in \mathbb{C}^n$ taki, że

$$\mathbf{A}\mathbf{x} = \lambda\mathbf{x}\,. \tag{14}$$

Wektor $\mathbf{x}$ nazywam wówczas *wektorem własnym macierzy* $\mathbf{A}$ *do wartości własnej* $\lambda$.

Definicja jest sformułowana dla macierzy zespolonych, my jednak — poza przypadkiem macierzy hermitowskich — będziemy zajmować się macierzami rzeczywistymi. Należy jednak pamiętać, że także macierze rzeczywiste (niesymetryczne) mogą mieć zespolone wartości własne.

Problem poszukiwania wartości własnych macierzy nazywa się *zagadnieniem własnym*.

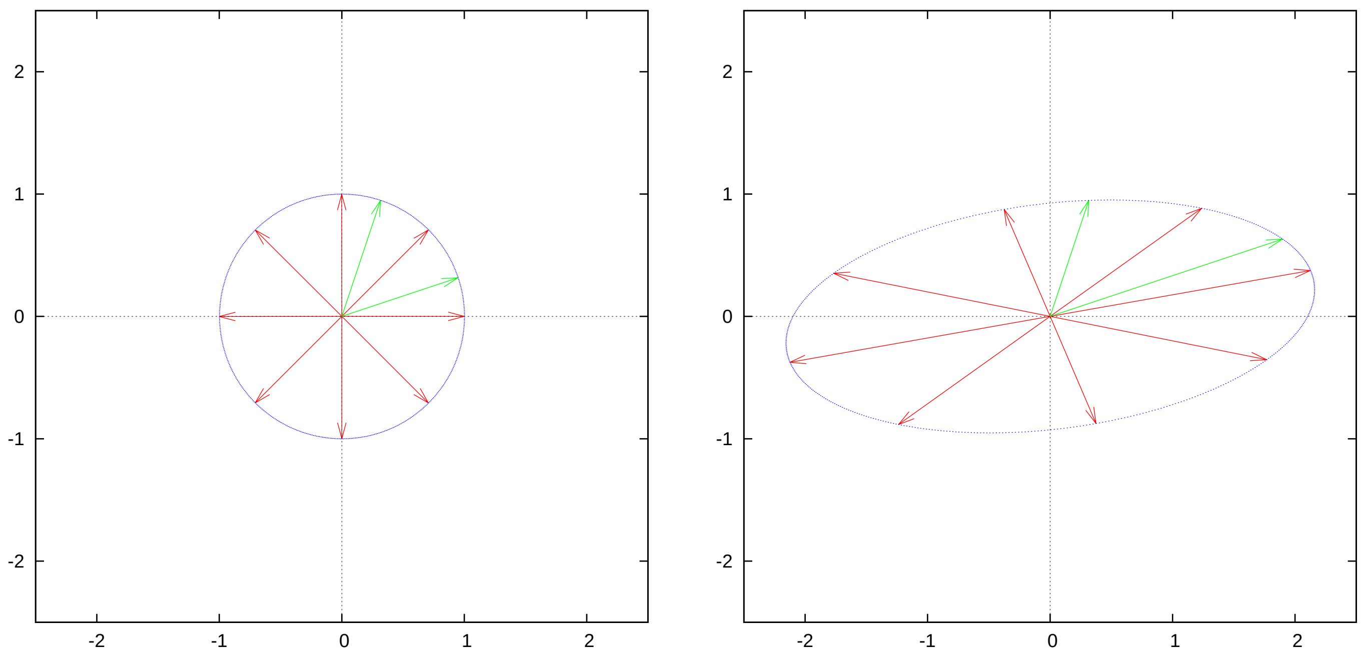
 

## Przykład

Rozważmy macierz

$$\begin{bmatrix} \frac{17}{8} & -\frac{3}{8} \\ \frac{3}{8} & \frac{7}{8} \end{bmatrix}. \tag{15}$$

Jej unormowanymi wektorami własnymi są $\left[\frac{1}{\sqrt{10}}, \frac{3}{\sqrt{10}}\right]^T$, $\left[\frac{3}{\sqrt{10}}, \frac{1}{\sqrt{10}}\right]^T$, a odpowiadającymi im wartościami własnymi liczby $1, 2$. Rysunek na następnej stronie przedstawia rodzinę wektorów jednostkowych (lewy panel) oraz tę samą rodzinę przekształconą przez macierz (15). Wektory własne zaznaczone są na zielono.




Rodzina wektorów jednostkowych przekształconych przez macierz (15). Wektory własne tej macierzy zaznaczone są na zielono. Jeden z nich, odpowiadający wartości własnej 1, nie zmienia się, drugi, odpowiadający wartości własnej 2, zachowuje kierunek, ale jego długość rośnie dwukrotnie.

## Równanie charakterystyczne

Równanie (14) można zapisać w postaci

$$(\mathbf{A} - \lambda \mathbb{I})\, \mathbf{x} = 0\,. \tag{16}$$

Jeżeli $\lambda$ jest wartością własną, równanie to musi mieć niezerowe rozwiązanie ze względu na $\mathbf{x}$. Jest to możliwe tylko w wypadku

$$\det\left(\mathbf{A} - \lambda \mathbb{I}\right) = 0\,. \tag{17}$$

Równanie (17) nazywane jest *równaniem charakterystycznym macierzy* $\mathbf{A}$. Jest to równanie wielomianowe stopnia $n$. Widzimy, że każda wartość własna jest pierwiastkiem równania charakterystycznego. Zbiór wszystkich rozwiązań równania (17) nazywam *widmem macierzy* $\mathbf{A}$.

## Przykład

Znajdźmy wartości własne macierzy obrotu płaskiego. W tym celu znajdujemy równanie charakterystyczne:

$$\begin{vmatrix} \cos\beta - \lambda & -\sin\beta \\ \sin\beta & \cos\beta - \lambda \end{vmatrix} = (\cos\beta - \lambda)^2 + \sin^2\beta \tag{18}$$

Znaleziony wielomian charakterystyczny nie ma pierwiastków rzeczywistych (poza przypadkiem $\sin\beta = 0$), gdyż jest sumą dwu składników nieujemnych. Istnieją dwa pierwiastki zespolone: $\lambda_1 = \cos\beta + i\sin\beta$, $\lambda_2 = \cos\beta - i\sin\beta$. Zauważmy, że $|\lambda_1| = |\lambda_2| = 1$.

Jak widzimy, niesymetryczna macierz rzeczywista może mieć zespolone wartości własne.

## Diagonalizacja

Przypuśćmy, że wektory $\mathbf{x}^{(1)}, \mathbf{x}^{(2)}, \dots, \mathbf{x}^{(n)}$ są wektorami własnymi pewnej macierzy $\mathbf{A}$, a $\lambda_1, \lambda_2, \dots, \lambda_n$ odpowiednimi wartościami własnymi. Utwórzmy macierz $\mathbf{X}$, której **kolumnami są kolejne wektory** $\mathbf{x}^{(i)}$. Każdy z wektorów $\mathbf{x}^{(i)}$ spełnia $\mathbf{A}\mathbf{x}^{(i)} = \lambda_i \mathbf{x}^{(i)}$. Gdybyśmy wszystkie te zależności chcieli zapisać łącznie, to *zgodnie z zasadami mnożenia macierzowego* przybrałoby to postać

$$\mathbf{A}\mathbf{X} = \mathbf{X}\,\mathsf{diag}\{\lambda_1, \lambda_2, \dots, \lambda_n\} \tag{19}$$

Faktycznie, niech $x_k^{(l)}$ oznacza $k$-tą składową $l$-tego wektora, czyli $k$-ty wiersz $l$-tej kolumny macierzy $\mathbf{X}$:

$$
\begin{aligned}
\mathbf{X}\,\mathsf{diag}\{\lambda_1,\lambda_2,\dots,\lambda_n\} &= \begin{bmatrix} x_1^{(1)} & x_1^{(2)} & \dots & x_1^{(n)} \\ x_2^{(1)} & x_2^{(2)} & \dots & x_2^{(n)} \\ \vdots & \vdots & \ddots & \vdots \\ x_n^{(1)} & x_n^{(2)} & \dots & x_n^{(n)} \end{bmatrix} \begin{bmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \lambda_n \end{bmatrix} \\
&= \begin{bmatrix} \lambda_1 x_1^{(1)} & \lambda_2 x_1^{(2)} & \dots & \lambda_n x_1^{(n)} \\ \lambda_1 x_2^{(1)} & \lambda_2 x_2^{(2)} & \dots & \lambda_n x_2^{(n)} \\ \vdots & \vdots & \ddots & \vdots \\ \lambda_1 x_n^{(1)} & \lambda_2 x_n^{(2)} & \dots & \lambda_n x_n^{(n)} \end{bmatrix} \qquad (20)
\end{aligned}
$$

Pierwsza kolumna macierzy $\mathbf{X}$ została wymnożona przez $\lambda_1$, druga przez $\lambda_2$ itd.

Jeżeli wszystkie wektory własne są liniowo niezależne, $\det \mathbf{X} \neq 0$ i istnieje macierz $\mathbf{X}^{-1}$. Wówczas mnożąc (19) od lewej strony przez $\mathbf{X}^{-1}$

$$\mathbf{X}^{-1}\mathbf{A}\mathbf{X} = \text{diag}\{\lambda_1, \lambda_2, \dots, \lambda_n\} \tag{21}$$

Transformację po lewej (21) nazywam transformacją podobieństwa. Macierze, dla których istnieje transformacja podobieństwa sprowadzająca je do postaci diagonalnej, nazywam *normalnymi* lub *diagonalizowalnymi*. Transformację podobieństwa, która sprowadza macierz do postaci diagonalnej, można rozumieć jako znalezienie takiej bazy, w której elementy macierzowe danego operatora przybierają postać macierzy diagonalnej.

Powtórzmy, że kolumny macierzy $\mathbf{X}$ tworzą kolejne wektory własne macierzy $\mathbf{A}$.

## Ortogonalne transformacje podobieństwa

Szczególne znaczenie mają takie transformacje podobieństwa (nie wyłącznie transformacje diagonalizujące!), w których macierz $\mathbf{X}$ jest ortogonalna:

$$\mathbf{A}' = \mathbf{O}^T\mathbf{A}\mathbf{O} \tag{22}$$

Własności ortogonalnych transformacji podobieństwa:

1. Wyznacznik macierzy nie zmienia się w wyniku ortogonalnej transformacji podobieństwa.
2. Ślad macierzy nie zmienia się w wyniku ortogonalnej transformacji podobieństwa, gdzie przez "ślad macierzy" rozumiemy sumę jej elementów diagonalnych:

$$\text{Tr}\mathbf{A} = \sum_{i=1}^{n} a_{ii} \tag{23}$$

3. Ortogonalna transformacja podobieństwa nie zmienia widma macierzy. Istotnie,

$$\begin{aligned}\mathrm{det}\left(\mathbf{A}' - \lambda\mathbb{I}\right) &= \mathrm{det}\left(\mathbf{O}^T\mathbf{A}\mathbf{O} - \lambda\mathbf{O}^T\mathbf{O}\right) = \mathrm{det}\left(\mathbf{O}^T\left(\mathbf{A} - \lambda\mathbb{I}\right)\mathbf{O}\right) \\ &= \left(\mathrm{det}\,\mathbf{O}^T\right)\cdot\mathrm{det}(\mathbf{A}-\lambda\mathbb{I})\cdot(\mathrm{det}\,\mathbf{O}) = (\mathrm{det}\,\mathbf{O})^2\,\mathrm{det}(\mathbf{A}-\lambda\mathbb{I}) \\ &= \mathrm{det}(\mathbf{A}-\lambda\mathbb{I}) \end{aligned} \tag{24}$$

a więc ortogonalna transformacja podobieństwa nie zmienia równania charakterystycznego.

## Diagonalizacja dwu macierzy

**Twierdzenie:** Jeżeli dwie macierze komutują, $[\mathbf{A}, \mathbf{B}] = \mathbf{AB} - \mathbf{BA} = 0$, istnieje transformacja podobieństwa diagonalizująca *obie* te macierze. Jeżeli macierze **nie** komutują, nie da się ich jednocześnie zdiagonalizować.

## Przykład

Łatwo sprawdzić, że macierze

$$\begin{bmatrix} \frac{17}{8} & -\frac{3}{8} \\ \frac{3}{8} & \frac{7}{8} \end{bmatrix} \qquad \begin{bmatrix} \frac{35}{8} & -\frac{9}{8} \\ \frac{9}{8} & \frac{5}{8} \end{bmatrix} \tag{25a}$$

komutują (druga z nich jest kwadratem pierwszej). Jednocześnie

$$\begin{bmatrix} -\frac{\sqrt{10}}{8} & \frac{3\sqrt{10}}{8} \\ \frac{3\sqrt{10}}{8} & -\frac{\sqrt{10}}{8} \end{bmatrix} \begin{bmatrix} \frac{17}{8} & -\frac{3}{8} \\ \frac{3}{8} & \frac{7}{8} \end{bmatrix} \begin{bmatrix} \frac{1}{\sqrt{10}} & \frac{3}{\sqrt{10}} \\ \frac{3}{\sqrt{10}} & \frac{1}{\sqrt{10}} \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix} \tag{25b}$$

$$\begin{bmatrix} -\frac{\sqrt{10}}{8} & \frac{3\sqrt{10}}{8} \\ \frac{3\sqrt{10}}{8} & -\frac{\sqrt{10}}{8} \end{bmatrix} \begin{bmatrix} \frac{35}{8} & -\frac{9}{8} \\ \frac{9}{8} & \frac{5}{8} \end{bmatrix} \begin{bmatrix} \frac{1}{\sqrt{10}} & \frac{3}{\sqrt{10}} \\ \frac{3}{\sqrt{10}} & \frac{1}{\sqrt{10}} \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 4 \end{bmatrix} \tag{25c}$$

## Wartości i wektory własne macierzy odwrotnej

Niech $\mathbf{B}$ będzie odwracalna, $\det \mathbf{B} \neq 0$, i niech $\lambda$ będzie jej wartością własną a $\mathbf{x}$ odpowiadającym jej wektorem własnym:

$$\mathbf{B}\mathbf{x} = \lambda\mathbf{x} \tag{26a}$$

Mnożąc powyższe równanie lewostronnie przez $\mathbf{B}^{-1}$ otrzymujemy

$$\begin{aligned} \mathbf{B}^{-1}\mathbf{B}\mathbf{x} &= \lambda\mathbf{B}^{-1}\mathbf{x} && \text{(26b)} \\ \mathbf{x} &= \lambda\mathbf{B}^{-1}\mathbf{x} && \text{(26c)} \\ \mathbf{B}^{-1}\mathbf{x} &= \frac{1}{\lambda}\mathbf{x} && \text{(26d)} \end{aligned}$$

Macierz i macierz odwrotna mają te same wektory własne, a ich wartości własne są swoimi odwrotnościami.

## Macierze symetryczne, rzeczywiste

Macierzą symetryczną, rzeczywistą nazywam macierz $\mathbb{R}^{n\times n}$, która spełnia

$$\mathbf{A}^T = \mathbf{A} \tag{27}$$

Macierze symetryczne, rzeczywiste wyglądają tak samo, niezależnie czy czytać je kolumnami, czy wierszami.

Macierze takie odgrywają **wielką** rolę w fizyce.

## Przykłady

Następujące macierze są symetryczne i rzeczywiste:

$$\begin{bmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \\ 3 & 1 & 2 \end{bmatrix} \qquad \begin{bmatrix} 4 & 1 & 0 & 0 & 0 \\ 1 & 4 & 1 & 0 & 0 \\ 0 & 1 & 4 & 1 & 0 \\ 0 & 0 & 1 & 4 & 1 \\ 0 & 0 & 0 & 1 & 4 \end{bmatrix}$$

$$\begin{bmatrix} 128 & 1 & 1 & 1 & \dots & 1 \\ 1 & 2 & & & & \\ 1 & & 2 & & & \\ 1 & & & 2 & & \\ \vdots & & & & \ddots & \\ 1 & & & & & 2 \end{bmatrix}$$

(niezaznaczone elementy są zerami)

## Wartości własne macierzy symetrycznej, rzeczywistej

**Twierdzenie:** Wartości własne macierzy symetrycznej, rzeczywistej są rzeczywiste.

*Dowód.* Niech $\mathbf{A}^T = \mathbf{A}$ i niech zachodzi

$$\mathbf{A}\mathbf{x} = \lambda\mathbf{x} \tag{28a}$$

Dokonajmy sprzężenia zespolonego (28a). Ponieważ $\mathbf{A}$ jest rzeczywista, $\mathbf{A}^* = \mathbf{A}$. Otrzymujemy

$$\mathbf{A}\mathbf{x}^* = \lambda^*\mathbf{x}^* \tag{28b}$$

Mnożymy (28a) lewostronnie przez $(\mathbf{x}^*)^T$, a (28b) lewostronnie przez $\mathbf{x}^T$. Otrzymujemy

$$(\mathbf{x}^*)^T \mathbf{A}\mathbf{x} = \lambda (\mathbf{x}^*)^T \mathbf{x} \tag{28c}$$

$$\mathbf{x}^T \mathbf{A}\mathbf{x}^* = \lambda^* \mathbf{x}^T \mathbf{x}^* \tag{28d}$$

Zauważmy, że $\mathbf{x}^T\mathbf{x}^*$ jest *liczbą*, a zatem $\mathbf{x}^T\mathbf{x}^* = \left(\mathbf{x}^T\mathbf{x}^*\right)^T = (\mathbf{x}^*)^T\,\mathbf{x}$. Podobnie $\mathbf{x}^T\mathbf{A}\mathbf{x}^*$ jest liczbą, a wobec tego

$$\mathbf{x}^T\mathbf{A}\mathbf{x}^* = \left(\mathbf{x}^T\mathbf{A}\mathbf{x}^*\right)^T = (\mathbf{x}^*)^T\,\mathbf{A}^T\mathbf{x} = (\mathbf{x}^*)^T\,\mathbf{A}\mathbf{x}\,, \tag{28e}$$

gdzie ostatnia równość wynika z symetrii macierzy $\mathbf{A}$: $\mathbf{A}^T = \mathbf{A}$. Równania (28c), (28d) możemy więc zapisać w postaci

$$(\mathbf{x}^*)^T\,\mathbf{A}\mathbf{x} = \lambda\,(\mathbf{x}^*)^T\,\mathbf{x} \tag{28f}$$
$$(\mathbf{x}^*)^T\,\mathbf{A}\mathbf{x} = \lambda^*\,(\mathbf{x}^*)^T\,\mathbf{x} \tag{28g}$$

Odejmując stronami równania (28f), (28g), otrzymujemy ostatecznie

$$0 = (\lambda - \lambda^*)\,(\mathbf{x}^*)^T\,\mathbf{x} \tag{28h}$$

ponieważ $\mathbf{x} \neq 0$, $(\mathbf{x}^*)^T\,\mathbf{x} = \|\mathbf{x}\|^2 > 0$, wnioskujemy, że $\lambda = \lambda^*$, czyli $\lambda \in \mathbb{R}$, co kończy dowód.

$\square$

## Ortogonalne wektory własne

**Twierdzenie:** Wektory własne macierzy symetrycznych, rzeczywistych do *różnych* wartości własnych są ortogonalne.

*Dowód.* Niech $\lambda_1 \neq \lambda_2$ będą wartościami własnymi pewnej macierzy symetrycznej, rzeczywistej:

$$
\begin{aligned}
\mathbf{A}\mathbf{x}_1 &= \lambda_1 \mathbf{x}_1 & \text{(29a)} \\
\mathbf{A}\mathbf{x}_2 &= \lambda_2 \mathbf{x}_2 & \text{(29b)}
\end{aligned}
$$

Mnożymy lewostronnie pierwsze z równań (29) przez $\mathbf{x}_2^T$, drugie zaś przez $\mathbf{x}_1^T$:

$$
\begin{aligned}
\mathbf{x}_2^T \mathbf{A}\mathbf{x}_1 &= \lambda_1 \mathbf{x}_2^T \mathbf{x}_1 & \text{(30a)} \\
\mathbf{x}_1^T \mathbf{A}\mathbf{x}_2 &= \lambda_2 \mathbf{x}_1^T \mathbf{x}_2 & \text{(30b)}
\end{aligned}
$$

Podobnie jak poprzednio, korzystając z symetrii macierzy $\mathbf{A}$, możemy łatwo pokazać, że $\mathbf{x}_2^T\mathbf{A}\mathbf{x}_1 = \mathbf{x}_1^T\mathbf{A}\mathbf{x}_2$. Odejmując równania (30) stronami otrzymujemy

$$0 = (\lambda_1 - \lambda_2)\mathbf{x}_2^T\mathbf{x}_1 \tag{31}$$

a zatem, ponieważ $\lambda_1 \neq \lambda_2$, widzimy, iż $\mathbf{x}_2^T\mathbf{x}_1 = 0$.

$\square$

## Inne własności macierzy symetrycznych, rzeczywistych

1. Odwrotność macierzy symetrycznej jest symetryczna. Jeżeli $\mathbf{A}^T = \mathbf{A}$ oraz $\det \mathbf{A} \neq 0$, wówczas $\left(\mathbf{A}^{-1}\right)^T = \mathbf{A}^{-1}$.

2. Ortogonalna transformacja podobieństwa zachowuje symetrię. Niech $\mathbf{A}^T = \mathbf{A}$ oraz $\mathbf{O}^T\mathbf{O} = \mathbb{I}$. Wówczas

$$\mathbf{A}' = \mathbf{O}^T\mathbf{A}\mathbf{O} \tag{32a}$$

$$\left(\mathbf{A}'\right)^T = \left(\mathbf{O}^T\mathbf{A}\mathbf{O}\right)^T = \mathbf{O}^T\mathbf{A}^T\mathbf{O} = \mathbf{O}^T\mathbf{A}\mathbf{O} = \mathbf{A}' \tag{32b}$$

## Prosty przykład

Chcemy znaleźć wartości własne i odpowiadające im unormowane wektory własne macierzy

$$\mathbf{A} = \begin{bmatrix} 7 & -4 \\ -4 & 13 \end{bmatrix} \tag{33a}$$

Wielomian charakterystyczny ma postać

$$P(\lambda) = (7-\lambda)(13-\lambda) - (-4)(-4) = \lambda^2 - 20\lambda + 75\,, \tag{33b}$$

a jego pierwiastkami są $\lambda = 5$, $\lambda = 15$. Są to poszukiwane wartości własne.

Oznaczając przez $a, b$ składowe wektora własnego i podstawiając $\lambda = 5$ do równania $\mathbf{Ax} = \lambda\mathbf{x}$ otrzymuję

$$\left\{ \begin{array}{rrcl} 2a & -4b & = & 0 \\ -4a & +8b & = & 0 \end{array} \right. \tag{33c}$$

Rozwiązaniem jest $a = 2b$, więc po unormowaniu okazuje się, że wektorem własnym macierzy (33a) do wartości własnej $\lambda = 5$ jest wektor

$$\mathbf{e}_1 = \begin{bmatrix} \frac{2}{\sqrt{5}} \\ \frac{1}{\sqrt{5}} \end{bmatrix} \tag{33d}$$

Podobnie, podstawiając $\lambda = 15$ do równania $\mathbf{A}\mathbf{x} = \lambda\mathbf{x}$, otrzymujemy

$$\left\{ \begin{array}{rrcl} -8a & -4b & = & 0 \\ -4a & -2b & = & 0 \end{array} \right. \tag{33e}$$

Tu z kolei rozwiązaniem jest $b = -2a$, więc po unormowaniu jako wektor własny do wartości własnej $\lambda = 15$ otrzymujemy

$$\mathbf{e}_2 = \begin{bmatrix} -\frac{1}{\sqrt{5}} \\ \frac{2}{\sqrt{5}} \end{bmatrix} \tag{33f}$$

## Przykład: Wartości i wektory własne macierzy $4 \times 4$

Dana jest macierz

$$\mathbf{A} = \begin{bmatrix} 4 & 1 & 1 & 1 \\ 1 & -1 & 3 & 0 \\ 1 & 3 & -1 & 0 \\ 1 & 0 & 0 & 2 \end{bmatrix} \tag{34a}$$

Jest to macierz symetryczna i rzeczywista, wiemy zatem, że musi mieć rzeczywiste wartości własne. Aby je znaleźć, musimy najpierw znaleźć wielomian charakterystyczny:

$$P(\lambda) = \begin{vmatrix} 4-\lambda & 1 & 1 & 1 \\ 1 & -1-\lambda & 3 & 0 \\ 1 & 3 & -1-\lambda & 0 \\ 1 & 0 & 0 & 2-\lambda \end{vmatrix} \tag{34b}$$

Jest to wyznacznik 4×4, będziemy go obliczać korzystając z rozwinięcia Laplace'a, dla wygody, według czwartej kolumny.

$$
\begin{aligned}
P(\lambda) &= (-1)^{1+4}\cdot 1\cdot \begin{vmatrix} 1 & -1-\lambda & 3 \\ 1 & 3 & -1-\lambda \\ 1 & 0 & 0 \end{vmatrix} \\
&+ (-1)^{4+4}\cdot(2-\lambda)\cdot \begin{vmatrix} 4-\lambda & 1 & 1 \\ 1 & -1-\lambda & 3 \\ 1 & 3 & -1-\lambda \end{vmatrix} \\
&= -(\lambda^2+2\lambda-8)+(2-\lambda)\left(-\lambda^3+2\lambda^2+18\lambda-24\right) \\
&= \lambda^4-4\lambda^3-15\lambda^2+58\lambda-40 \qquad \text{(34c)}
\end{aligned}
$$

Wielomian (34c) ma współczynniki całkowite, warto więc spróbować poszukać jego miejsc zerowych badając podzielniki wyrazu wolnego. Istotnie, $P(1)=0$. Dzielimy teraz $P(\lambda)$ przez dwumian $(\lambda-1)$. W tym celu rozwiązujemy układ równań

$$
\begin{aligned}
b_3 &= 1 & \text{(34d)}\\
-b_3 + b_2 &= -4 & \text{(34e)}\\
-b_2 + b_1 &= -15 & \text{(34f)}\\
-b_1 + b_0 &= 58 & \text{(34g)}
\end{aligned}
$$

skąd wnioskujemy, że

$$P(\lambda) = (\lambda - 1)\left(\lambda^3 - 3\lambda^2 - 18\lambda + 40\right) \tag{34h}$$

Występujący tu wielomian trzeciego stopnia także ma współczynniki całkowite, a jednym z jego miejsc zerowych (i zarazem podzielnikiem wyrazu wolnego) jest $\lambda = 2$. Po raz kolejny obniżamy stopień wielomianu, rozwiązując

$$
\begin{aligned}
c_2 &= 1 & \text{(34i)}\\
-2c_2 + c_1 &= -3 & \text{(34j)}\\
-2c_1 + c_0 &= -18 & \text{(34k)}
\end{aligned}
$$

czyli

$$P(\lambda) = (\lambda - 1)(\lambda - 2)\left(\lambda^2 - \lambda - 20\right) \tag{34l}$$

Teraz w znany sposób znajdujemy pierwiastki wielomianu stopnia drugiego: $\lambda = 5$ oraz $\lambda = -4$. Ostatecznie

$$P(\lambda) = (\lambda - 5)(\lambda - 2)(\lambda - 1)(\lambda + 4) \tag{34m}$$

a więc wartościami własnymi macierzy (34a) są liczby $5, 2, 1, -4$ (zauważmy, że wszystkie one są dzielnikami wyrazu wolnego wielomianu (34c)).

Przystępujemy do obliczania wektorów własnych. Oznaczmy symbolicznie poszukiwany wektor własny przez $\mathbf{x} = [a, b, c, d]$. Podstawiamy $\lambda = 5$ do wyrażenia $\mathbf{Ax} = \lambda\mathbf{x}$ i otrzymujemy

$$\left\{ \begin{array}{rrrrcl} -a & + \ b & + \ c & + \ d & = & 0 \\ a & -6b & +3c & & = & 0 \\ a & +3b & -6c & & = & 0 \\ a & & & -3d & = & 0 \end{array} \right. \tag{34n}$$

Rozwiązania zależą od jednego parametru: $a = 3d$, $b = c = d$. Po unormowaniu jako wektor własny macierzy (34a) do wartości własnej $\lambda = 5$ dostajemy

$$\mathbf{e}_1 = \begin{bmatrix} \frac{3}{2\sqrt{3}} \\ \frac{1}{2\sqrt{3}} \\ \frac{1}{2\sqrt{3}} \\ \frac{1}{2\sqrt{3}} \end{bmatrix} \tag{34o}$$

Aby znaleźć wektor własny do wartości własnej $\lambda = 2$, podstawiamy tę wartość do $\mathbf{A}\mathbf{x} = \lambda\mathbf{x}$ i otrzymujemy

$$\left\{ \begin{array}{rrrrcl} 2a & + \; b & + \; c & +d & = & 0 \\ a & -3b & +3c & & = & 0 \\ a & +3b & -3c & & = & 0 \\ a & & & & = & 0 \end{array} \right. \tag{34p}$$

skąd $a = 0$, $b = c$, $d = -2c$, co po unormowaniu daje jako wektor własny do wartości własnej $\lambda = 2$

$$\mathbf{e}_2 = \begin{bmatrix} 0 \\ \frac{1}{\sqrt{6}} \\ \frac{1}{\sqrt{6}} \\ -\frac{2}{\sqrt{6}} \end{bmatrix} \tag{34q}$$

Podobnie postępujemy dla wartości własnej $\lambda = 1$, otrzymując

$$\left\{ \begin{array}{rrrrcl} 3a & + \; b & + \; c & +d & = & 0 \\ a & -2b & +3c & & = & 0 \\ a & +3b & -2c & & = & 0 \\ a & & & +d & = & 0 \end{array} \right. \tag{34r}$$

skąd wynika $a = -d$, $b = c = d$, a po unormowaniu

$$\mathbf{e}_3 = \begin{bmatrix} -\frac{1}{2} \\ \frac{1}{2} \\ \frac{1}{2} \\ \frac{1}{2} \end{bmatrix} \tag{34s}$$

Wreszcie dla $\lambda = -4$ dostajemy

$$\left\{\begin{array}{rrrrcl} 8a & + \; b & + \; c & + \; d & = & 0 \\ a & +3b & +3c & & = & 0 \\ a & +3b & +3c & & = & 0 \\ a & & & +6d & = & 0 \end{array}\right. \tag{34t}$$

więc $a = d = 0$, $b = -c$, a po unormowaniu

$$\mathbf{e}_3 = \begin{bmatrix} 0 \\ \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} \\ 0 \end{bmatrix} \tag{34u}$$

## Zdegenerowane wartości własne

Twierdzenie o ortogonalności wektorów własnych głosi, że wektory własne macierzy symetrycznej, rzeczywistej do *różnych* wartości własnych są ortogonalne. Co jednak, gdy jakaś wartość własna się powtarza? Ma to miejsce, gdy równanie charakterystyczne ma pierwiastek wielokrotny. Taką wartość własną nazywam *zdegenerowaną*.

Jeśli któraś wartość własna jest zdegenerowana, wektory własne do tej wartości własnej rozpinają pewną podprzestrzeń (podprzestrzeń własną) o wymiarowości większej niż 1, ortogonalną do pozostałych wektorów własnych. W takiej podprzestrzeni własnej zawsze można wybrać bazę ortonormalną.

## Przykład: macierz ze zdegenerowanymi wartościami własnymi

Znajdźmy wartości własne macierzy

$$\mathbf{A} = \begin{bmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{bmatrix} \tag{35}$$

Jest to macierz symetryczna, rzeczywista, więc jej wartości własne są rzeczywiste, a wektory własne do *różnych* wartości własnych są ortogonalne.

Równaniem charakterystycznym jest

$$\begin{vmatrix} -\lambda & 1 & 1 \\ 1 & -\lambda & 1 \\ 1 & 1 & -\lambda \end{vmatrix} = -\lambda^3 + 3\lambda + 2 = 0 \tag{36}$$

Korzystając z odpowiedniego twierdzenia, łatwo sprawdzić, że $\lambda = 2$ jest całkowitym pierwiastkiem wielomianu (36), gdyż $-2^3 + 3 \cdot 2 + 2 = 0$.

Algorytm dzielenia wielomianu (36) przez dwumian $(\lambda-2)$ prowadzi do układu równań

$$\begin{aligned} b_2 &= -1 && \text{(37a)}\\ -2b_2+b_1 &= 0 && \text{(37b)}\\ -2b_1+b_0 &= 3 && \text{(37c)} \end{aligned}$$

skąd $b_2=-1$, $b_1=-2$, $b_0=-1$, a zatem równanie charakterystyczne przybiera postać

$$-\lambda^3+3\lambda+2=(\lambda-2)(-\lambda^2-2\lambda-1)=-(\lambda-2)(\lambda+1)^2=0\,. \quad (38)$$

Ostatecznie widzimy, że wartościami własnymi macierzy (35) są $\lambda = 2$ oraz $\lambda=-1$, przy czym ta druga jest dwukrotnie zdegenerowana.

Przystępujemy do szukania wektorów własnych. Oznaczmy symbolicznie poszukiwany wektor przez $\mathbf{x} = [a, b, c]$. Bierzemy $\lambda = 2$ i wstawiamy do równania $\mathbf{A}\mathbf{x} = \lambda\mathbf{x}$:

$$
\begin{aligned}
-2a + b + c &= 0 \quad &(39a)\\
a - 2b + c &= 0 \quad &(39b)\\
a + b - 2c &= 0 \quad &(39c)
\end{aligned}
$$

Wyznacznik główny tego układu — jak wiemy — wynosi 0, a rozwiązania zależą od jednego parametru: $a = b = c$. Po unormowaniu dostajemy

$$
\mathbf{e}_1 = \begin{bmatrix} \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{3}} \end{bmatrix} \tag{40}
$$

jako unormowany wektor własny macierzy (35) do wartości własnej $\lambda = 2$.

Teraz bierzemy $\lambda = -1$. Wszystkie trzy równania definiujące składowe wektora własnego redukują się do jednego równania:

$$a + b + c = 0 \tag{41}$$

i możemy wziąć dowolne liczby spełniające ten warunek; zauważmy, że jest to warunek ortogonalności do wektora (40), co nie powinno dziwić. Weźmy $b = -a$, $c = 0$, co po unormowaniu daje

$$\mathbf{e}_2 = \begin{bmatrix} \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} \\ 0 \end{bmatrix} \tag{42}$$

Drugiego wektora własnego do wartości własnej $\lambda = -1$ musimy szukać jako wektora ortogonalnego do (40) i liniowo niezależnego od (42). *Wygodnie* jest przyjąć, że ten trzeci wektor własny jest także ortogonalny do

(42), czyli jego składowe muszą spełniać

$$a + b + c = 0 \tag{43a}$$

$$a - b = 0 \tag{43b}$$

Unormowanym rozwiązaniem jest

$$\mathrm{e}_3 = \begin{bmatrix} \frac{1}{\sqrt{6}} \\ \frac{1}{\sqrt{6}} \\ -\frac{2}{\sqrt{6}} \end{bmatrix} \tag{44}$$

Układ wektorów $\{\mathrm{e}_1, \mathrm{e}_2, \mathrm{e}_3\}$ stanowi układ wzajemnie ortogonalnych wektorów własnych macierzy (35). Jak widzimy, wektory własne do różnych wartości własnych koniecznie muszą być ortogonalne, natomiast wektory własne do tej samej, zdegenerowanej wartości własnej możemy dobrać tak, aby były do siebie ortogonalne.

## Zadanie ☺

Znajdź wartości własne i unormowane, wzajemnie ortogonalne wektory własne macierzy

$$\begin{bmatrix} 4 & 1 & 1 & 1 \\ 1 & 2 & 0 & 0 \\ 1 & 0 & 2 & 0 \\ 1 & 0 & 0 & 2 \end{bmatrix} \tag{45}$$

Wskazówki:

1. Szukając wielomianu charakterystycznego, proszę zastosować rozwinięcie Laplace'a względem czwartej kolumny.
2. Wielomian charakterystyczny ma współczynniki całkowite i, jak się okazuje, *wszystkie* jego pierwiastki są całkowite. Proszę sprawdzić podzielniki wyrazu wolnego.

## Podsumowanie

Podsumowując rozważania dotyczące wartości i wektorów własnych macierzy symetrycznych, rzeczywistych, możemy powiedzieć, że

1. Wartości własne macierzy symetrycznych, rzeczywistych są rzeczywiste i jest ich tyle (licząc z krotnościami), ile wynosi wymiarowość przestrzeni;
2. Wektory własne macierzy symetrycznych, rzeczywistych do różnych wartości własnych są ortogonalne;

3. W przypadku zdegenerowanych wartości własnych, zawsze możemy dobrać wektory własne tak, aby były ortogonalne,

z czego wynika, że

Unormowane wektory własne macierzy symetrycznej, rzeczywistej stanowią bazę ortonormalną w $\mathbb{R}^n$.

## Diagonalizacja macierzy symetrycznej, rzeczywistej

Każdą macierz symetryczną, rzeczywistą daje się zdiagonalizować. Diadonalizacja ma formę ortogonalnej transformacji podobieństwa

$$\mathrm{diag}\left\{\lambda_1, \lambda_2, \dots, \lambda_n\right\} = \mathbf{S}^T \mathbf{A} \mathbf{S} \tag{46}$$

gdzie kolejne unormowane wektory własne macierzy $\mathbf{A}$ tworzą kolejne kolumny macierzy ortogonalnej $\mathbf{S}$.

## Uogólniony iloczyn skalarny

Iloczyn skalarny w $\mathbb{R}^n$ definiujemy jako $\mathbf{x}^T\mathbf{y} = \mathbf{x}^T\mathbb{I}\mathbf{y}$. Macierz $\mathbb{I}$ jest oczywiście symetryczna, rzeczywista. Co by się stało, gdyby zamiast macierzy jednostkowej wstawić tam jakąś *inną* macierz?

$$\mathbf{x} \circ \mathbf{y} = \mathbf{x}^T\mathbf{A}\mathbf{y} \tag{47}$$

Macierz $\mathbf{A}$ koniecznie musi być symetryczna, gdyż definicja iloczynu skalarnego wymaga symetrii:

$$\mathbf{y} \circ \mathbf{x} = \mathbf{y}^T\mathbf{A}\mathbf{x} = \left(\mathbf{y}^T\mathbf{A}\mathbf{x}\right)^T = \mathbf{x}^T\mathbf{A}^T\mathbf{y} = \mathbf{x}^T\mathbf{A}\mathbf{y} = \mathbf{x} \circ \mathbf{y}\,, \tag{48}$$

co jest możliwe wtedy i tylko wtedy, gdy $\mathbf{A}^T = \mathbf{A}$.

Pojawia się inny problem: Iloczyn skalarny musi być dodatnio określony:

$$\forall \mathbf{x} \neq \mathbf{0} : \mathbf{x}^T \mathbf{A} \mathbf{x} > 0 \tag{49}$$

Macierze symetryczne spełniające warunek (49) nazywam *dodatnio określonymi*. Jeżeli macierz jest symetryczna i dodatnio określona, wyrażenie (47) definiuje iloczyn skalarny. Jest to *inny* iloczyn skalarny, niż "euklidesowy" iloczyn $\mathbf{x}^T \mathbf{y}$, naturalny dla przestrzeni $\mathbb{R}^n$, ale formalnie jest on poprawnie zdefiniowany.

Sens tego iloczynu skalarnego (47) najlepiej zobaczyć przechodząc do bazy, w której macierz $\mathbf{A}$ jest diagonalna. Z wyrażenia (46) otrzymujemy

$$\mathbf{A} = \mathbf{S} \, \mathrm{diag} \left\{ \lambda_1, \dots, \lambda_n \right\} \mathbf{S}^T \tag{50}$$

co po wstawieniu do (47) daje

$$\mathbf{x}^T \mathbf{S} \, \mathrm{diag} \left\{ \lambda_1, \dots, \lambda_n \right\} \mathbf{S}^T \mathbf{y} = \left( \mathbf{S}^T \mathbf{x} \right)^T \mathrm{diag} \left\{ \lambda_1, \dots, \lambda_n \right\} \left( \mathbf{S}^T \mathbf{y} \right) \tag{51}$$

$\mathbf{S}^T\mathbf{x}$, $\mathbf{S}^T\mathbf{y}$ to wektory $\mathbf{x}, \mathbf{y}$ wyrażone w innej bazie. W tej bazie uogólniony iloczyn skalarny działa podobnie, jak zwykły, z tym, że przyczynki od poszczególnych, wzajemnie prostopadłych, kierunków, odpowiadających kierunkom wektorów własnych macierzy $\mathbf{A}$, są domnażane przez dodatnie wagi, równe wartościom własnym $\mathbf{A}$.

**Obserwacja:** Macierz symetryczna jest dodatnio określona wtedy i tylko wtedy, gdy jej wszystkie wartości własne są większe od zera.

Wektory ortogonalne w sensie uogólnionego iloczynu skalarnego (47) nazywa się *wektorami sprzężonymi względem macierzy* $\mathbf{A}$.

## Rozkład jedności

Niech $\{\mathbf{e}_i\}_{i=1}^n$ stanowi bazę ortonormalną w $\mathbb{R}^n$. Wówczas

$$\sum_{i=1}^{n} \mathbf{e}_i \mathbf{e}_i^T = \mathbb{I}\,. \tag{52}$$

(Zauważmy, że obiekty $\mathbf{e}_i\mathbf{e}_i^T$ to operatory (macierze). Nazywa się je operatorami rzutowymi, gdyż rzutują dowolny wektor na wektor $\mathbf{e}_i$.)

*Dowód.* Skoro wektory $\{\mathbf{e}_i\}_{i=1}^n$ stanowi bazę, każdy wektor można przedstawić jako kombinację liniową tych wektorów bazowych:

$$\mathbf{v} = \sum_{j=1}^{n} \alpha_j \mathbf{e}_j\,. \tag{53a}$$

Liczymy

$$
\begin{aligned}
\left(\sum_{i=1}^{n}\mathbf{e}_i\mathbf{e}_i^T\right)\mathbf{v} &= \left(\sum_{i=1}^{n}\mathbf{e}_i\mathbf{e}_i^T\right)\sum_{j=1}^{n}\alpha_j\mathbf{e}_j = \sum_{i,j=1}^{n}\alpha_j\mathbf{e}_i\underbrace{\mathbf{e}_i^T\mathbf{e}_j}_{\delta_{ij}} \\
&= \sum_{i,j=1}^{n}\alpha_j\delta_{ij}\mathbf{e}_i = \sum_{j=1}^{n}\alpha_j\mathbf{e}_j = \mathbf{v}
\end{aligned}
\tag{53b}
$$

a zatem działanie operatora (52) nie zmienia wektora $\mathbf{v}$. Ponieważ zachodzi to dla dowolnego wektora, operator (52) musi być operatorem tożsamościowym, $\mathbb{I}$.

□

## Twierdzenie spektralne

Niech $\mathbf{A} \in \mathbb{R}^{n\times n}$ będzie macierzą symetryczną, rzeczywistą. $\lambda_1, \dots, \lambda_n$ są jej wartościami własnymi, a $\mathbf{e}_1, \dots, \mathbf{e}_n$ odpowiadającymi im unormowanymi wektorami własnymi. Wówczas

$$\mathbf{A} = \sum_{i=1}^{n} \lambda_i \, \mathbf{e}_i \mathbf{e}_i^T \, . \tag{54}$$

Operator po prawej stronie (54) nazywam *rozkładem spektralnym* operatora $\mathbf{A}$.

*Dowód.* Korzystam z reprezentacji (53a) i obliczam

$$\mathbf{A}\mathbf{v} = \mathbf{A} \sum_{j=1}^{n} \alpha_j \mathbf{e}_j = \sum_{j=1}^{n} \alpha_j \mathbf{A}\mathbf{e}_j = \sum_{j=1}^{n} \alpha_j \lambda_j \mathbf{e}_j \tag{55a}$$

Z drugiej strony

$$\left(\sum_{i=1}^{n} \lambda_i\, \mathbf{e}_i \mathbf{e}_i^T\right) \mathbf{v} = \sum_{i,j=1}^{n} \lambda_i \alpha_j \mathbf{e}_i \mathbf{e}_i^T \mathbf{e}_j = \sum_{i,j=1}^{n} \lambda_i \alpha_j \delta_{ij} \mathbf{e}_i = \sum_{j=1}^{n} \alpha_j \lambda_j \mathbf{e}_j \tag{55b}$$

Ponieważ prawe strony (55a), (55b) są sobie równe i zachodzi to dla *dowolnego* wektora $\mathbf{v}$, równość (54) musi być prawdziwa tożsamościowo.

$\square$

**Uwaga:** Równość (54) jest najprostszą formą twierdzenia spektralnego. W przypadkach operatorów niehermitowskich oraz w przestrzeniach nieskończenie wymiarowych twierdzenie spektralne przybiera bardziej skomplikowaną formę.

## Przykład

Weźmy dwa wektory z $\mathbb{R}^2$:

$$\mathbf{e}_1 = \begin{bmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{bmatrix}, \qquad \mathbf{e}_1 = \begin{bmatrix} \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} \end{bmatrix} \tag{56a}$$

Jak łatwo sprawdzić, są one wzajemnie ortogonalne i unormowane, stanowią więc bazę w $\mathbb{R}^2$. Mamy

$$\begin{aligned} \mathbf{e}_1\mathbf{e}_1^T + \mathbf{e}_2\mathbf{e}_2^T &= \begin{bmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{bmatrix} \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix} + \begin{bmatrix} \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} \end{bmatrix} \begin{bmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{bmatrix} \\ &= \begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{bmatrix} + \begin{bmatrix} \frac{1}{2} & -\frac{1}{2} \\ -\frac{1}{2} & \frac{1}{2} \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \end{aligned} \tag{56b}$$

Podobnie

$$
\begin{aligned}
2\,\mathbf{e}_1\mathbf{e}_1^T + 4\,\mathbf{e}_2\mathbf{e}_2^T &= 2\cdot\begin{bmatrix}\frac{1}{\sqrt{2}}\\ \frac{1}{\sqrt{2}}\end{bmatrix}\begin{bmatrix}\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}}\end{bmatrix} + 4\cdot\begin{bmatrix}\frac{1}{\sqrt{2}}\\ -\frac{1}{\sqrt{2}}\end{bmatrix}\begin{bmatrix}\frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}}\end{bmatrix} \\
&= 2\cdot\begin{bmatrix}\frac{1}{2} & \frac{1}{2}\\ \frac{1}{2} & \frac{1}{2}\end{bmatrix} + 4\cdot\begin{bmatrix}\frac{1}{2} & -\frac{1}{2}\\ -\frac{1}{2} & \frac{1}{2}\end{bmatrix} \\
&= \begin{bmatrix}1 & 1\\ 1 & 1\end{bmatrix} + \begin{bmatrix}2 & -2\\ -2 & 2\end{bmatrix} = \begin{bmatrix}3 & -1\\ -1 & 3\end{bmatrix} \qquad \text{(56c)}
\end{aligned}
$$

Macierz po prawej stronie (56c) jest symetryczna, rzeczywista, a jej równanie charakterystyczne ma postać $(3-\lambda)^2-1=0$. Wartościami własnymi są $\lambda_1 = 2$, $\lambda_2 = 4$, a wektory (56a) są odpowiadającymi im wektorami własnymi. Wyrażenie po lewej stronie (56c) jest rozkładem spektralnym tej macierzy.

## Funkcje macierzy

Niech $\mathbf{A}$ będzie macierzą symetryczną, rzeczywistą, posiadającą rozkład spektralny (54). Niech $f : \mathbb{R} \to \mathbb{R}$ będzie jakąś funkcją. **Definiuję** funkcję macierzy następująco:

$$f(\mathbf{A}) = \sum_{i=1}^{n} f(\lambda_i)\, \mathbf{e}_i \mathbf{e}_i^T \tag{57}$$

Definicja (57) ma sens tylko wtedy, gdy dla każdej wartości własnej wyrażenie $f(\lambda_i)$ jest dobrze określone. Na przykład tylko macierze symetryczne i dodatnio określone można w powyższym sensie pierwiastkować lub logarytmować.

## Potęgowanie macierzy kwadratowych

Naturalne potęgi macierzy definiujemy jako mnożenie macierzy przez siebie:

$$\mathbf{A}^n = \underbrace{\mathbf{A} \cdot \mathbf{A} \cdots \mathbf{A}}_{n \text{ razy}} \tag{58}$$

Zauważmy, że dla macierzy symetrycznych, rzeczywistych, a więc posiadających rozkład spektralny, definicja ta zgadza się z wyrażeniem (57). Na przykład dla $\mathbf{A}^2$

$$\begin{aligned} \mathbf{A}^2 &= \left(\sum_{i=1}^{n} \lambda_i \, \mathbf{e}_i \mathbf{e}_i^T\right)^2 = \sum_{i,j=1}^{n} \lambda_i \lambda_j \, \mathbf{e}_i \mathbf{e}_i^T \mathbf{e}_j \mathbf{e}_j^T = \sum_{i,j=1}^{n} \lambda_i \lambda_j \delta_{ij} \mathbf{e}_i \mathbf{e}_j^T \\ &= \sum_{i=1}^{n} \lambda_i^2 \, \mathbf{e}_i \mathbf{e}_i^T \end{aligned} \tag{59}$$

Dla macierzy odwracalnych można też definiować potęgi ujemne, jako potęgi macierzy odwrotnej. Ponadto dla $\mathbf{A} \neq \mathbf{0}$ przyjmujemy $\mathbf{A}^0 = \mathbb{I}$. W ten sposób mamy zdefiniowane wszystkie całkowitoliczbowe potęgi macierzy (potęgi ujemne nie zawsze istnieją!) i zachodzą zwykłe zasady składania potęg.

Niecałkowite potęgi macierzy można definiować tylko w sensie rozkładu spektralnego.

## Macierze nilpotentne

Wśród macierzy, które **nie** są symetryczne, zdarzają się macierze *nilpotentne*, to znaczy takie, których pewna potęga (a zatem i wszystkie wyższe) są macierzami zerowymi.

## Przykłady

$$\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}^2 = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} \tag{60a}$$

$$
\begin{aligned}
\begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}^3 &= \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix} \\
&= \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix} \\
&= \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}
\end{aligned} \tag{60b}
$$

## Macierze idempotentnte

Macierz jest *idempotentna*, jeśli $\mathbf{A}^2 = \mathbf{A}$. Macierz jednostkowa jest idempotentnta. Macierz idempotentna, która nie jest macierzą jednostkową, jest osobliwa.

## Przykłady

$$\begin{bmatrix} 2 & -1 \\ 2 & -1 \end{bmatrix}^2 = \begin{bmatrix} 2 & -1 \\ 2 & -1 \end{bmatrix} \begin{bmatrix} 2 & -1 \\ 2 & -1 \end{bmatrix} = \begin{bmatrix} 4-2 & -2+1 \\ 4-2 & -2+1 \end{bmatrix} = \begin{bmatrix} 2 & -1 \\ 2 & -1 \end{bmatrix} \tag{61a}$$

Operator rzutowy $\mathbf{e}\mathbf{e}^T$, gdzie $\|\mathbf{e}\| = 1$, jest idempotentny. Istotnie,

$$\left(\mathbf{e}\mathbf{e}^T\right)^2 = \mathbf{e}\,\underbrace{\mathbf{e}^T\mathbf{e}}_{=1}\,\mathbf{e}^T = \mathbf{e}\mathbf{e}^T \tag{61b}$$

## Macierzowa funkcja wykładnicza

Funkcję wykładniczą *definiujemy* jako sumę nieskończonego szeregu

$$e^x \equiv \exp(x) = \sum_{k=0}^{\infty} \frac{x^k}{k!} \tag{62}$$

Szereg (62) zawiera wyłącznie naturalne potęgi $x^n$, nie ma więc przeszkód, by tę definicję uogólnić na dowolne macierze kwadratowe:

$$\exp(\mathbf{A}) = \sum_{k=0}^{\infty} \frac{1}{k!}\mathbf{A}^k \tag{63}$$

## Własności macierzowej funkcji wykładniczej

$$\exp(\mathbf{A}^T) = (\exp(\mathbf{A}))^T \tag{64a}$$

$$\exp(\mathbf{A}^*) = (\exp(\mathbf{A}))^* \tag{64b}$$

$$[\mathbf{A}, \mathbf{B}] = 0 \Rightarrow \exp(\mathbf{A}) \exp(\mathbf{B}) = \exp(\mathbf{A} + \mathbf{B}) \tag{64c}$$

Ostatnia zależność jest bardzo ważna. Mówi ona, że jeśli dwie macierze komutują, eksponenta ich sumy jest równa iloczynowi eksponent, tak, jak dla liczb. Jeśli macierze nie komutują, eksponenta sumy **nie** jest równa iloczynowi eksponent. Za to zawsze zachodzi

$$\exp(\mathbf{A}) \exp(-\mathbf{A}) = \exp(\mathbf{0}) = \mathbb{I} \tag{64d}$$

## Przykład

Obliczmy

$$\exp\left(\boldsymbol{\sigma}_z\beta\right) = \exp\left(\begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}\beta\right) \tag{65a}$$

Zauważmy, że

$$\boldsymbol{\sigma}_z^2 = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}\begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix} = -\mathbb{I} \tag{65b}$$

a wobec tego

$$\boldsymbol{\sigma}_z^3 = -\boldsymbol{\sigma}_z\,, \quad \boldsymbol{\sigma}_z^4 = \mathbb{I}\,, \quad \boldsymbol{\sigma}_z^5 = \boldsymbol{\sigma}_z\,, \quad \text{itd} \tag{65c}$$

Podstawiając do definicji macierzowej funkcji wykładniczej otrzymujemy

$$
\begin{aligned}
\exp(\boldsymbol{\sigma}_z\beta) &= \sum_{k=0}^{\infty}\frac{\beta^k}{k!}\boldsymbol{\sigma}_z^k \\
&= \sum_{k=0}^{\infty}\frac{\beta^{2k}}{(2k)!}\boldsymbol{\sigma}_z^{2k} + \sum_{k=0}^{\infty}\frac{\beta^{2k+1}}{(2k+1)!}\boldsymbol{\sigma}_z^{2k+1} \\
&= \sum_{k=0}^{\infty}(-1)^k\frac{\beta^{2k}}{(2k)!}\mathbb{I} + \sum_{k=0}^{\infty}(-1)^k\frac{\beta^{2k+1}}{(2k+1)!}\boldsymbol{\sigma}_z \\
&= \cos\beta\,\mathbb{I} + \sin\beta\,\boldsymbol{\sigma}_z \\
&= \begin{bmatrix} \cos\beta & -\sin\beta \\ \sin\beta & \cos\beta \end{bmatrix} \qquad (65d)
\end{aligned}
$$

gdzie skorzystaliśmy z rozwinięcia Taylora funkcji sinus i kosinus. Jak widzimy, otrzymalismy macierz obrotu płaskiego, co *nie* jest przypadkową koincydencją ☺.

## Funkcja wykładnicza macierzy symetrycznych, rzeczywistych

Jeśli macierz jest symetryczna, rzeczywista, zachodzi (50). Wobec tego

$$
\begin{aligned}
\mathbf{A}^k &= \left(\mathbf{S}\,\mathsf{diag}\,\{\lambda_1,\dots,\lambda_n\}\,\mathbf{S}^T\right)^k \\
&= \mathbf{S}\mathsf{diag}\{\dots\}\mathbf{S}^T\mathbf{S}\mathsf{diag}\{\dots\}\mathbf{S}^T\cdots\mathbf{S}\mathsf{diag}\{\dots\}\mathbf{S}^T
\end{aligned} \tag{66a}
$$

Wewnątrz tego wyrażenia, pomiędzy każdymi kolejnymi wystąpieniami macierzy diagonalnej $\mathsf{diag}\{\dots\}$, pojawiają się iloczyny $\mathbf{S}^T\mathbf{S}$, które są równe $\mathbb{I}$. Zatem

$$
\begin{aligned}
\mathbf{A}^n &= \mathbf{S}\,\underbrace{\mathsf{diag}\{\dots\}\cdots\mathsf{diag}\{\dots\}}_{k \text{ razy}}\,\mathbf{S}^T \\
&= \mathbf{S}\,(\mathsf{diag}\,\{\lambda_1,\dots,\lambda_n\})^n\,\mathbf{S}^T \\
&= \mathbf{S}\mathsf{diag}\,\{\lambda_1^k,\dots,\lambda_n^k\}\mathbf{S}^T
\end{aligned} \tag{66b}
$$

Wstawiając to wyrażenie do definicji funkcji wykładniczej dostajemy

$$
\begin{aligned}
\exp(\mathbf{A}) &= \sum_{k=0}^{\infty} \frac{1}{k!}\mathbf{A}^k = \sum_{k=0}^{\infty} \frac{1}{k!}\mathbf{S}\,\mathrm{diag}\,\{\lambda_1^k, \dots, \lambda_n^k\}\mathbf{S}^T \\
&= \mathbf{S}\left(\sum_{k=0}^{\infty} \frac{1}{k!}\mathrm{diag}\,\{\lambda_1^k, \dots, \lambda_n^k\}\right)\mathbf{S}^T \\
&= \mathbf{S}\,\mathrm{diag}\left\{\sum_{k=0}^{\infty} \frac{\lambda_1^k}{k!}, \dots, \sum_{k=0}^{\infty} \frac{\lambda_1^k}{k!}\right\}\mathbf{S}^T \\
&= \mathbf{S}\begin{bmatrix} e^{\lambda_1} & & & \\ & e^{\lambda_2} & & \\ & & \ddots & \\ & & & e^{\lambda_n} \end{bmatrix}\mathbf{S}^T
\end{aligned}
\tag{67}
$$

## Przykład

Chcemy obliczyć

$$\exp\left(\begin{bmatrix} 3 & -4 \\ -4 & 3 \end{bmatrix} t\right), \tag{68a}$$

gdzie $t$ jest liczbą rzeczywistą. Liczba ta domnaża macierz, a więc domnaża jej wartości własne, nie wpływając na wektory własne. Równanie charakterystystyczne macierzy, której eksponentę chcemy obliczyć, ma postać

$$\begin{vmatrix} 3-\lambda & -4 \\ -4 & 3-\lambda \end{vmatrix} = (3-\lambda)^2 - 16 = 0 \tag{68b}$$

a więc wartościami własnymi są $\lambda = 7$ i $\lambda = -1$. Jak łatwo sprawdzić, wektorami własnymi są

$$\mathbf{e}_1 = \begin{bmatrix} \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} \end{bmatrix}, \qquad \mathbf{e}_2 = \begin{bmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{bmatrix} \tag{68c}$$

a wobec tego macierzą diagonalizującą jest

$$\mathbf{S} = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix} \tag{68d}$$

Ostatecznie

$$\begin{aligned} \exp\left(\begin{bmatrix} 3 & -4 \\ -4 & 3 \end{bmatrix} t\right) &= \frac{1}{2} \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} e^{7t} & 0 \\ 0 & e^{-t} \end{bmatrix} \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix} \\ &= \frac{1}{2} \begin{bmatrix} e^{-t} + e^{7t} & e^{-t} - e^{7t} \\ e^{-t} - e^{7t} & e^{-t} + e^{7t} \end{bmatrix} \end{aligned} \tag{68e}$$

## Sprzężenie hermitowskie

Sprzężenie hermitowskie jest złożeniem sprzężenia zespolonego oraz transpozycji. Dla macierzy rzeczywistych sprzężenie hermitowskie jest zwykłą transpozycją.

## Przykłady

$$\begin{bmatrix} 1 & 2 \\ -3i & 4i \end{bmatrix}^\dagger = \begin{bmatrix} 1 & 3i \\ 2 & -4i \end{bmatrix} \qquad \begin{bmatrix} 1 \\ 2i \\ -3i \\ 4 \end{bmatrix}^\dagger = \begin{bmatrix} 1 & -2i & 3i & 4 \end{bmatrix}$$

$$\begin{bmatrix} 1 & i & 2 \\ -i & -3 & 4i \end{bmatrix}^\dagger = \begin{bmatrix} 1 & i \\ -i & -3 \\ 2 & -4i \end{bmatrix}$$

## Własności sprzężenia hermitowskiego

Sprzężenie hermitowskie odwraca kolejność iloczynu:

$$(\mathbf{AB})^\dagger = \mathbf{B}^\dagger \mathbf{A}^\dagger \tag{69}$$

Z tego wynika, że

$$\left(\mathbf{A}^\dagger\right)^k = \left(\mathbf{A}^k\right)^\dagger \tag{70}$$

a zatem

$$\exp\left(\mathbf{A}^\dagger\right) = (\exp(\mathbf{A}))^\dagger \tag{71}$$

## Iloczyn skalarny w $\mathbb{C}^n$

Dopuszczając wektory i skalary zespolone, musimy nieco zmodyfikować definicję iloczynu skalarnego, tak by uwzględniała ona to, że część urojona zmienia znak przy sprzężeniu zespolonym. W szczególności

- $\forall \mathbf{x}, \mathbf{y} \in \mathbb{C} : \mathbf{x} \circ \mathbf{y} = (\mathbf{y} \circ \mathbf{x})^*$.
- $\forall \mathbf{x}, \mathbf{y} \in \mathbb{C}\,,\ \alpha \in \mathbb{C} : (\alpha \mathbf{x}) \circ \mathbf{y} = \mathbf{x} \circ (\alpha^* \mathbf{y})$

W przypadku rzeczywistych wektorów i skalarów zasady te sprowadzają się do zasad, które już znamy. **Nie zmienia** się wymóg **dodatniej określoności**:

- $\forall \mathbf{x} \in \mathbb{C}^n\,,\ \mathbf{x} \neq \mathbf{0} : \mathbf{x} \circ \mathbf{x} > 0$.

Jako definicję iloczynu skalarnego w $\mathbb{C}^n$ przyjmujemy uogólnienie definicji "euklidesowej":

$$\mathbf{x} \circ \mathbf{y} = \mathbf{x}^\dagger \mathbf{y} \tag{72}$$

W tej sytuacji długością wektora jest

$$\|\mathbf{x}\| = \sqrt{\mathbf{x}^\dagger \mathbf{x}} = \sqrt{\sum_{i=1}^{n} |x_i|^2} \tag{73}$$

## Przykład

Znajdźmy wartości własne i wektory własne macierzy obrotu płaskiego. W tym celu znajdujemy równanie charakterystyczne:

$$\begin{vmatrix} \cos\beta-\lambda & -\sin\beta \\ \sin\beta & \cos\beta-\lambda \end{vmatrix} = (\cos\beta-\lambda)^2+\sin^2\beta \tag{74}$$

Znaleziony wielomian charakterystyczny nie ma pierwiastków rzeczywistych (poza przypadkiem $\sin\beta = 0$), gdyż jest sumą dwu składników nieujemnych. Istnieją dwa pierwiastki zespolone: $\lambda_1 = \cos\beta + i\sin\beta$, $\lambda_2 = \cos\beta - i\sin\beta$. Zauważmy, że $|\lambda_1| = |\lambda_2| = 1$. W celu znalezienia wektorów własnych podstawiamy jedną z wartości własnych do równania definiującego wektory własne:

$$a\,(\cos\beta-(\cos\beta+i\sin\beta))+b\,(-\sin\beta)=0 \tag{75a}$$

$$-ia-b=0 \tag{75b}$$

Równanie na drugi wektor własny będzie się różnić jedynie znakiem przy $i$. Wektorami własnymi są

$$\mathbf{e}_1 = \begin{bmatrix} \frac{1}{\sqrt{2}} \\ -\frac{i}{\sqrt{2}} \end{bmatrix} \qquad \mathbf{e}_2 = \begin{bmatrix} \frac{1}{\sqrt{2}} \\ \frac{i}{\sqrt{2}} \end{bmatrix} \tag{76}$$

Zauważmy, że wektory (76) są unormowane w sensie (73) i ortogonalne w sensie (72).

Sprawdzamy:

$$
\begin{aligned}
\begin{bmatrix} \cos\beta & -\sin\beta \\ \sin\beta & \cos\beta \end{bmatrix} \mathbf{e}_1 &= \frac{1}{\sqrt{2}} \begin{bmatrix} \cos\beta & -\sin\beta \\ \sin\beta & \cos\beta \end{bmatrix} \begin{bmatrix} 1 \\ -i \end{bmatrix} \\
&= \frac{1}{\sqrt{2}} \begin{bmatrix} \cos\beta + i\sin\beta \\ \sin\beta - i\cos\beta \end{bmatrix} \\
&= \frac{1}{\sqrt{2}} \begin{bmatrix} \cos\beta + i\sin\beta \\ -i(i\sin\beta + \cos\beta) \end{bmatrix} \\
&= \frac{1}{\sqrt{2}} (\cos\beta + i\sin\beta) \begin{bmatrix} 1 \\ -i \end{bmatrix} \\
&= (\cos\beta + i\sin\beta)\mathbf{e}_1
\end{aligned}
\tag{77}
$$

i analogicznie dla drugiej pary wartość własna-wektor własny.

## Macierze unitarne

Macierz $\mathbf{U} \in \mathbb{C}^{n\times n}$ nazywam macierzą unitarną, jeżeli

$$\mathbf{U}^\dagger\mathbf{U} = \mathbf{U}\mathbf{U}^\dagger = \mathbb{I} \tag{78}$$

Macierze unitarne są uogólnieniem pojęcia macierzy ortogonalnych na przypadek zespolony. Własności macierzy ortogonalnych:

1. $\mathbf{U}^{-1} = \mathbf{U}^\dagger$.
2. $|\det \mathbf{U}| = 1$. Istotnie, $1 = \det \mathbb{I} = \det\left(\mathbf{U}^\dagger\mathbf{U}\right) = \left(\det \mathbf{U}^\dagger\right)(\det \mathbf{U})$ $= (\det \mathbf{U})^* (\det \mathbf{U}) = |\det \mathbf{U}|^2$.
3. Macierze unitarne w $\mathbb{C}^{n\times n}$ tworzą grupę. Działaniem grupowym jest mnożenie macierzy.

4. Kolumny macierzy unitarnej są unormowane w sensie (73) i wzajemnie prostopadłe w sensie (72). Kolumny macierzy unitarnej tworzą bazę w $\mathbb{C}^n$.
5. Macierz unitarna zachowuje iloczyn skalarny. Istotnie, $(\mathbf{U}\mathbf{u}) \circ (\mathbf{U}\mathbf{v}) = (\mathbf{U}\mathbf{u})^\dagger (\mathbf{U}\mathbf{u}) = \mathbf{u}^\dagger \mathbf{U}^\dagger \mathbf{U}\mathbf{v} = \mathbf{u}^\dagger \mathbf{v} = \mathbf{u} \circ \mathbf{v}$.
6. Każda *rzeczywista* macierz ortogonalna jest unitarna.

## Macierze hermitowskie

Mówię, że macierz $\mathbf{H} \in \mathbb{C}^{n\times n}$ jest hermitowska, jeżeli

$$\mathbf{H}^{\dagger} = \mathbf{H} \tag{79}$$

Macierze hermitowskie są uogólnieniem pojęcia macierzy symetrycznych, rzeczywistych na przypadek zespolony i dzielą z nimi szereg własności.

- Każda macierz symetryczna, rzeczywista jest hermitowska.
- Wartości własne macierzy hermitowskich są rzeczywiste, choć odpowiadające im wektory własne mogą być zespolone.

- Wektory własne macierzy hermitowskich do różnych wartości własnych są ortogonalne w sensie (72).
- Unormowane wektory własne macierzy hermitowskiej tworzą bazę w $\mathbb{C}^n$.
- Macierz hermitowską można zdiagonalizować za pomocą *unitarnej transformacji podobieństwa*

$$\mathrm{diag}\,\{\lambda_1, \lambda_2, \dots, \lambda_n\} = \mathbf{U}^\dagger \mathbf{H} \mathbf{U} \tag{80}$$

  gdzie kolejne unormowane wektory własne $\mathbf{H}$ stanowią kolejne kolumny unitarnej macierzy $\mathbf{U}$.
- Odwrotność macierzy hermitowskiej, jeśli istnieje, jest hermitowska.
- Każda unitarna transformacja podobieństwa (nie tylko transformacja diagonalizująca) zachowuje hermitowskość.
- Jeżeli macierz $\mathbf{H}$ jest hermitowska, macierz $\exp(it\mathbf{H})$ jest unitarna.

Istotnie,

$$(\exp(it\mathbf{H}))^\dagger = \exp\left((it\mathbf{H})^\dagger\right) = \exp\left(-it\mathbf{H}^\dagger\right) = \exp(-it\mathbf{H}) \tag{81a}$$

wobec czego

$$\exp(it\mathbf{H})\,(\exp(it\mathbf{H}))^\dagger = \exp(it\mathbf{H})\,\exp(-it\mathbf{H}) = \mathbb{I} \tag{81b}$$

## Przykład

Znajdźmy wartości i wektory własne macierzy

$$\mathbf{H} = \begin{bmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & i & 0 \\ 0 & -i & 0 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix} \tag{82}$$

Jest to macierz hermitowska. W celu znalezienia wartości własnych obliczamy

$$P(\lambda) = \left| \begin{array}{cccc} -\lambda & 0 & 0 & 1 \\ 0 & -\lambda & i & 0 \\ 0 & -i & -\lambda & 0 \\ 1 & 0 & 0 & -\lambda \end{array} \right] \tag{83}$$

Jest to wyznacznik $4\times4$. Obliczymy go korzystając z rozwinięcia Laplace'a względem pierwszej kolumny

$$
\begin{aligned}
P(\lambda) &= (-\lambda)(-1)^{1+1}\begin{vmatrix} -\lambda & i & 0 \\ -i & -\lambda & 0 \\ 0 & 0 & -\lambda \end{vmatrix} + 1\cdot(-1)^{4+1}\begin{vmatrix} 0 & 0 & 1 \\ -\lambda & i & 0 \\ -i & -\lambda & 0 \end{vmatrix} \\
&= -\lambda(-\lambda^3+\lambda) - (\lambda^2-1) = \lambda^4 - 2\lambda^2 + 1 \\
&= (\lambda^2-1)^2 = (\lambda-1)^2(\lambda+1)^2
\end{aligned}
\tag{84}
$$

Wartościami własnymi są $\lambda = 1$, $\lambda = -1$, obie dwukrotnie zdegenerowane.

Aby znaleźć wektory własne odpowiadające $\lambda = 1$ podstawiamy tę wartość i wektor o składowych oznaczonych $a, b, c, d$ do równania $\mathbf{A}\mathbf{x} = \lambda\mathbf{x}$

i otrzymujemy

$$
\begin{aligned}
-a + d &= \; = 0 &\quad (85a)\\
-b + ic &= 0 &\quad (85b)\\
-ib - c &= 0 &\quad (85c)\\
a - d &= 0 &\quad (85d)
\end{aligned}
$$

z których tylko dwa są niezależne. Wektory własne do wartości własnej $\lambda = 1$ zależą od dwóch parametrów i muszą mieć postać

$$
\begin{bmatrix} a \\ ic \\ c \\ a \end{bmatrix} \tag{86}
$$

i jak długo mają postać (86), można je wybrać dowolnie. *Wygodnie* je dobrać tak, aby były ortogonalne, na przykład (po unormowaniu)

$$\mathbf{e}_1 = \begin{bmatrix} \frac{1}{2} \\ \frac{i}{2} \\ \frac{1}{2} \\ \frac{1}{2} \end{bmatrix}, \quad \mathbf{e}_2 = \begin{bmatrix} \frac{1}{2} \\ -\frac{i}{2} \\ -\frac{1}{2} \\ \frac{1}{2} \end{bmatrix} \tag{87}$$

Powtarzamy tę samą procedurę dla $\lambda = -1$:

$$\begin{aligned} a + d &= 0 && \text{(88a)} \\ b + ic &= 0 && \text{(88b)} \\ -ib + c &= 0 && \text{(88c)} \\ a + d &= 0 && \text{(88d)} \end{aligned}$$

wobec czego wektory własne do wartości własnej $\lambda = -1$ muszą mieć postać

$$\begin{bmatrix} a \\ -ic \\ c \\ -a \end{bmatrix} \tag{89}$$

Zauważmy, że wszystkie wektory postaci (89) są ortogonalne do już znalezionych wektorów $\mathbf{e}_1$, $\mathbf{e}_2$. Jako wzajemnie ortogonalną parę wektorów postaci (89) wybieramy

$$\mathbf{e}_3 = \begin{bmatrix} \frac{1}{2} \\ -\frac{i}{2} \\ \frac{1}{2} \\ -\frac{1}{2} \end{bmatrix}, \quad \mathbf{e}_4 = \begin{bmatrix} \frac{1}{2} \\ \frac{i}{2} \\ -\frac{1}{2} \\ -\frac{1}{2} \end{bmatrix} \tag{90}$$