Wykład 14

Macierze symetryczne

Niech $A \in \mathbb{C}^{n\times n}$. Macierz A nazywamy HERMITOWSKĄ iff $A = A^H$ (H oznacza sprzężenie i transpozycję). Dla macierzy hermitowskich iloczyn skalarny można zapisać na dwa sposoby $\langle Ax, y\rangle = \langle x, Ay\rangle$, ogólnie $\langle Ax, y\rangle = \left\langle x, A^H y\right\rangle$. Jeżeli macierz macierz ma elementy rzeczywiste, to zachodzi równość $A = A^T$.

Twierdzenie 1. *Wartości własne macierzy symetrycznej A są rzeczywiste.*

Dowód. Załóżmy, że λ_1 jest wartością własną a x_1 wektorem własnym. Jest wówczas: $\lambda_1 \langle x_1, x_1\rangle = \langle Ax_1, x_1\rangle = \left\langle x_1, A^H x_1\right\rangle = \bar{\lambda_1} \langle x_1, x_1\rangle$. □

Istnienie wartości własnej wynika z zasadniczego twierdzenia algebry.

Twierdzenie 2. *Wektory własne macierzy symetrycznej (hermitowskiej) tworzą bazę ortonormalną przestrzeni $\mathbb{R}^n$ ($\mathbb{C}^n$).*

Dowód. Weźmy wartość własną λ_1 i odpowiadający jej wektor własny x_1. Wówczas istnieją (tw. Steinitza) wektory $y_2, \dots, y_n$, które wraz z x_1 tworzą bazę $\mathbb{R}^n$. Można zatem wybrać (ortogonalizacja Grama-Schmidta) wektory $z_2, \dots, z_n$ – tworzące wraz z x_1 bazę – takie, że $\langle x_1, z_i\rangle = 0$.

Niech $V_{n-1} = \operatorname{span}\{z_2, \dots, z_n\}$ oraz $z \in V_{n-1}$ (czyli $z = \sum_{k=2}^{n} \alpha_k z_k$). Jest teraz

$$0 = \lambda_1 \sum_{k=2}^{n} \alpha_k \langle x_1, z_k\rangle = \lambda_1 \langle x_1, z\rangle = \langle \lambda_1 x_1, z\rangle = \langle Ax_1, z\rangle = \langle x_1, Az\rangle .$$

To oznacza, że $Az \in V_{n-1}$ czyli $AV_{n-1} \subset V_{n-1}$. Do przestrzeni V_{n-1} można teraz powtórzyć wcześniejsze postępowanie (dowód przez indukcję skończoną). □

Zauważmy, że $x_1 \perp V_{n-1}$.

Suma krotności geometrycznych wartości własnych jest równa sumie ich krotności algebraicznych.

Niech $X = [x_1, \dots, x_n]$ będzie macierzą wektorów własnych (ortonormalnych). Wówczas $X \cdot X^{-1} = X^{-1} \cdot X = I_n$ oraz $X^T AX = \Lambda$, gdzie $\Lambda = \operatorname{diag}(\lambda_1, \dots, \lambda_n)$ i można przyjąć, że $\lambda_1 \geqslant \lambda_2 \geqslant \dots \geqslant \lambda_n$.

Przykład:

Znajdźmy wartości i wektory własne macierzy $A = \begin{bmatrix} 3 & 2 & 4 \\ 2 & 0 & 2 \\ 4 & 2 & 3 \end{bmatrix}$. Wielomian charakterystyczny macierzy A to $w(\lambda) = \det(A - \lambda I) = -\lambda^3 + 6\lambda^2 + 15\lambda + 8$. Stąd $\lambda_1 = \lambda_2 = -1, \lambda_3 = 8$.

Wektory własne znajdujemy rozwiązując układy równań $(A - \lambda_k I)x = 0$. Dla $\lambda_{1,2} = -1$ otrzymujemy $\left[\begin{array}{ccc|c} 4 & 2 & 4 & 0 \\ 2 & 1 & 2 & 0 \\ 4 & 2 & 4 & 0 \end{array}\right]$. Rząd macierz jest równy 1, otrzymamy zatem dwa liniowo niezależne rozwiązania. Ortogonalne rozwiązania (wektory własne) to, na przykład $z_1 = [1, -2, 0]^T$, $z_2 = [-4, -2, 5]^T$ (ortogonalizacja Grama-Schmidta wektorów $y_1 = [1, -2, 0]^T$, $y_2 = [0, -2, 1]^T$). Po podzieleniu przez $\sqrt{5}$ i przez $\sqrt{45}$ otrzymamy ortonormalne wektory własne.

Dla $\lambda_3 = 8$ otrzymujemy równania $\left[\begin{array}{rrr|r} -5 & 2 & 4 & 0 \\ 2 & -8 & 2 & 0 \\ 4 & 2 & -5 & 0 \end{array}\right]$. Rozwiązaniem jest wektor $z_3 = [2, 1, 2]^T$. Jest on prostopadły do wcześniej obliczonych wektorów z_1, z_2. Szukanym układem ortonormalnych wektorów własnych jest zatem – dla przykładu – zbiór $\left\{x_1 = \frac{z_1}{\sqrt{5}}, x_2 = \frac{z_2}{\sqrt{45}}, x_3 = \frac{z_3}{3}\right\}$. Dla przykładu, ponieważ można było inaczej wybrać wektory odpowiadające podwójnej wartości własnej.

Niech w końcu $\lambda = \mathrm{diag}(-1, -1, 8)$ natomiast $X = [x_1, x_2, x_3]$. Można sprawdzić, że $X^T AX = X^{-1}AX = \Lambda$.

Macierze nieujemnie (dodatnio) określone

Definicja 1. *Niech $A \in \mathbb{C}^{n\times n}$. Mówimy, że (symetryczna) macierz A jest nieujemnie określona iff $\forall x \in \mathbb{R}^n$ $x^H Ax \geqslant 0$. Jeżeli $x^H Ax = 0$ tylko dla $x = \mathbb{O}$, to A nazywamy macierzą dodatnio określoną.*

Przykład:

Rozpatrzmy macierz $A = \begin{bmatrix} 2 & -1 & 1 \\ -1 & 2 & -1 \\ 1 & -1 & 2 \end{bmatrix}$. Niech $x = [x_1, x_2, x_3]^T$. Bezpośrednim rachunkiem sprawdzamy, że $x^T Ax = 2x_1^2 - 2x_1x_2 + 2x_1x_3 + 2x_2^2 - 2x_2x_3 + 2x_3^2 = = (x_1 - x_2)^2 + (x_1 + x_3)^2 + (x_2 - x_3)^2 \geqslant 0$. Równocześnie, jeżeli $x_1 - x_2 = x_1 + x_3 = x_2 - x_3 = 0$ to $x = \mathbb{O}$. Stąd macierz jest dodatnio określona.

Macierz $B = \begin{bmatrix} 2 & -1 & 1 \\ -1 & 2 & 1 \\ 1 & 1 & 2 \end{bmatrix}$ jest nieujemnie określona. Jest bowiem $x^T Ax = 2x_1^2 - 2x_1x_2 + 2x_1x_3 + 2x_2^2 + 2x_2x_3 + 2x_3^2 = = (x_1 - x_2)^2 + (x_1 + x_3)^2 + (x_2 + x_3)^2 \geqslant 0$. Zarazem dla $x = [a, a, -a]^T$ jest $x^T Ax = 0$.

Zauważmy jeszcze, że wartościami własnymi A są liczby 1, 1 i 4; wartości własne B to 3, 3 oraz 0.

Założymy teraz, że macierz $A \in \mathbb{R}^n$ jest symetryczna. Z twierdzeń 1 i 2 wynika, że wektory własne tej macierzy są rzeczywiste, zaś wektory własne tworzą układ ortonormalny, tzn. $Ax_k = \lambda_k x_k$, $x_k^T x_k = 1$. Z wektorów własnych utwórzmy macierz $X = [x_1, x_2, \ldots, x_n]$. Jest wówczas $X^T X = I_n$ oraz $AX = [\lambda_1 x_1, \lambda_2 x_2, \ldots, \lambda_n] = X\Lambda$. Mnożąc lewostronnie ostatnią równość przez X^T otrzymujemy ostatecznie $X^T AX = \Lambda$, gdzie Λ jest macierzą diagonalną, elementy diagonalne to wartości własne macierzy A.

Wniosek 2.1. *Symetryczna (hermitowska) macierz $A \in \mathbb{R}^n$ ($A \in \mathbb{C}^n$) jest diagonalizowalna.*

Załóżmy teraz, że symetryczna macierz A jest nieujemnie (dodatnio) określona. Weźmy pod uwagę równanie $AX = [\lambda_1 x_1, \lambda_2 x_2, \ldots, \lambda_n] = X\Lambda$.

Jest $(Ax)_k = Ax_k = \lambda_k x_k$. Jest też $x_j^T Ax_k = \delta_{jk}$, $j = 1, \ldots, n$. Tę równość – dla ustalonego j – można zapisać macierzowo w postaci

$$x_j^T AX = [x_j^T Ax_1, \ldots, x_j^T Ax_n] = [0, \ldots, 0, \lambda_j, 0, \ldots, 0],$$

gdzie jedyna niezerowa pozycja występuje w miejscu j-tym. Łącząc ostatnie równanie dla $(j = 1, \dots, n)$ otrzymujemy równanie

$$X^T \cdot A \cdot X = \begin{bmatrix} x_1^T \\ \dots \\ x_n^T \end{bmatrix} X\Lambda = \Lambda.$$

Odwołamy się teraz do zadania 5. Pamiętajmy też, że macierz A jest nieujemnie określona. Dla każdego z wektorów x_k jest $0 \leqslant x_k^T A x_k = \lambda_k$. Niech $v \in \mathbb{R}^n$. Wówczas $v = \sum_{k=1}^{n} \beta_k x_k$ (x_k tworzą bazę). Stąd $v^T AV = \langle v, Av \rangle = \left\langle \sum_{k=1}^{n} \beta_k x_k, A \cdot \sum_{s=1}^{n} \beta_s x_s \right\rangle$. Kolejne przekształcenia dają $v^T AV = \sum_{k=1,s=1}^{n} \lambda_s \beta_k \beta_s \langle x_k, x_s \rangle = \sum_{k=1}^{n} \lambda_k \beta_k^2 \geqslant 0$.

Wniosek 2.2. *Symetryczna (hermitowska) macierz $A \in \mathbb{R}^n$ ($A \in \mathbb{C}^n$) jest nieujemnie (dodatnio) określona iff jej wartości własne są nieujemne (dodatnie).*

Rozkład SVD

Niech $A \in \mathbb{R}^{m \times n}(\mathbb{C}^{m \times n})$. Rozważmy dwie macierze: $B_1 = A^T A \in \mathbb{R}^{n \times n}$ oraz $B_2 = AA^T \in \mathbb{R}^{m \times m}$. Zarówno B_1 jak i B_2 są symetryczne. Z poprzednich rozważań wynika, że $B_1(B_2)$ ma $n(m)$ wartości własnych i $n(m)$ ortonormalnych wektorów własnych.

Niech $v \in \mathbb{R}^n$. Z własności iloczynu skalarnego wynika, że

$$0 \leqslant \langle Av, Av \rangle = \left\langle v, A^T Av \right\rangle = \langle v, B_1 v \rangle = v^T B_1 v.$$

Podobnie, dla $w \in \mathbb{R}^m$ jest $0 \leqslant \left\langle A^T w, A^T w \right\rangle = \left\langle w, AA^T w \right\rangle = w^T AA^T w$. Macierze B_1, B_2 są zatem nieujemnie określone.

Przykład:

Niech $A = \begin{bmatrix} 3 & 2 & 2 \\ 2 & 3 & -2 \end{bmatrix}$. Obliczymy wartości własne i wektory własne macierzy $AA^T = \begin{bmatrix} 17 & 8 \\ 8 & 17 \end{bmatrix}$. Wielomian charakterystyczny to $w(\lambda) = \lambda^2 - 34\lambda + 225$. Stąd ($\Delta = 256$) otrzymujemy $\lambda_1 = 9$ oraz $\lambda_2 = 25$. Wyznaczenie wektorów własnych to rozwiązanie układów równań $(AA^T - \lambda_k I_2)u_k = \mathbb{O}$. Pozostaje zatem znaleźć i unormować rozwiązania układów równań

$$\left[\begin{array}{cc|c} 8 & 8 & 0 \\ 8 & 8 & 0 \end{array}\right], \left[\begin{array}{cc|c} -8 & 8 & 0 \\ 8 & -8 & 0 \end{array}\right].$$

Stąd $u_1 = [1, -1]^T/\sqrt{2}$, $u_2 = [1, 1]^T/\sqrt{2}$. Zauważamy, że wektory u_1, u_2 są prostopadłe.

Przechodzimy do obliczeń związanych z macierzą $A^T A = \begin{bmatrix} 13 & 12 & 2 \\ 12 & 13 & -2 \\ 2 & -2 & 8 \end{bmatrix}$. Wielomian charakterystyczny to $w(\lambda) = -\lambda^3 + 34\lambda^2 - 225\lambda$. Pierwiastki tego wielomianu to 25, 9 oraz 0. Znajdujemy też wektory własne.

$$\left[\begin{array}{rrr|r} -12 & 12 & 2 & 0 \\ 12 & -12 & -2 & 0 \\ 2 & -2 & -17 & 0 \end{array}\right], \quad \left[\begin{array}{rrr|r} 4 & 12 & 2 & 0 \\ 12 & 4 & -2 & 0 \\ 2 & -2 & -1 & 0 \end{array}\right], \quad \left[\begin{array}{rrr|r} 13 & 12 & 2 & 0 \\ 12 & 13 & -2 & 0 \\ 2 & -2 & 8 & 0 \end{array}\right].$$

Rozwiązaniami są wektory: $v_1 = [1,1,0]^T/\sqrt{2}$, $v_2 = [1,-1,4]^T/\sqrt{18}$ oraz wektor $v_3 = [2,-2,-1]^T/3$.

Utwórzmy macierze U, V łącząc wektory u_k, v_l. Otrzymamy:

$$U = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}, \; V = \begin{bmatrix} 1 & 1 & 2 \\ 1 & -1 & -2 \\ 0 & 4 & -1 \end{bmatrix} \cdot \begin{bmatrix} 1/\sqrt{2} & 0 & 0 \\ 0 & 1/\sqrt{18} & 0 \\ 0 & 0 & 1/3 \end{bmatrix}.$$

Co ważne, zachodzi równość:

$$A = U \cdot \begin{bmatrix} 5 & 0 & 0 \\ 0 & 3 & 0 \end{bmatrix} \cdot V^T = U\,\Sigma\,V^T$$

Przechodzimy teraz do głównego twierdzenia.

Rozważmy macierz $A \in \mathbb{R}^{m\times m}$. Niech $p = \min\{m,n\}$. Macierz $B_2 = AA^T$ jest symetryczną macierzą rozmiaru $m \times n$. Z poprzednich rozważań wynika też, że jest ona nieujemnie określona. Jej wartości własne są zatem nieujemne, a wektory własne tworzą ortonormalną bazę przestrzeni $\mathbb{R}^m$. Załóżmy, że $\lambda_1 \geqslant \ldots \geqslant \lambda_r > 0$ oraz $\lambda_{r+1} = \ldots = \lambda_m = 0$. Oczywiście $r \leqslant p$, szczegóły zależą od rzędu macierzy A.

Macierz U to macierz ortonormalnych wektorów własnych, czyli $U \in \mathbb{R}^{m\times m}$ oraz $U^TU = I_m$. Podzielmy U na dwie podmacierze, to znaczy $U = [U_1, U_2]$. Macierz U_1 zawiera wektory własne odpowiadające niezerowym wartościom własnym macierzy AA^T, macierz U_2 wektory własne odpowiadające $\lambda_{r+1}, \ldots, \lambda_m$.

Korzystając z ortogonalności macierzy U oraz z zadania 1 otrzymujemy

$$AA^T = U\Lambda U^T = \begin{bmatrix} U_1 & U_2 \end{bmatrix} \begin{bmatrix} \bar{\Lambda} & \mathbb{O} \\ \mathbb{O} & \mathbb{O} \end{bmatrix} \begin{bmatrix} U_1^T \\ U_2^T \end{bmatrix} = U_1\,\bar{\Lambda}\,U_1^T. \tag{1}$$

Macierz $\bar{\Lambda} \in \mathbb{R}^{r\times r}$, pozostałe macierze są takie, że $\Lambda \in \mathbb{R}^{m\times m}$. Pomnożenie równania (1) lewostronnie przez U^T i prawostronnie przez U daje równość

$$\begin{bmatrix} U_1^T \\ U_2^T \end{bmatrix} \cdot AA^T \cdot \begin{bmatrix} U_1 & U_2 \end{bmatrix} = \begin{bmatrix} \bar{\Lambda} & \mathbb{O} \\ \mathbb{O} & \mathbb{O} \end{bmatrix}, \tag{2}$$

skąd wynika, że

$$\begin{bmatrix} U_1^TAA^TU_1 & U_1^TAA^TU_2 \\ U_2^TAA^TU_1 & U_2^TAA^TU_2 \end{bmatrix} = \begin{bmatrix} \bar{\Lambda} & \mathbb{O} \\ \mathbb{O} & \mathbb{O} \end{bmatrix},$$

porównując bloki w prawym dolnym rogu otrzymujemy $\mathbb{O} = U_2^TAA^TU_2 = U_2^TA\left(U_2^TA\right)^T$. Jest to możliwe wtedy, gdy $U_2^TA = \mathbb{O}$.

Zdefiniujmy dwie macierze: $\bar{\Sigma} = \bar{\Lambda}^{1/2} = \mathrm{diag}\left(\sqrt{\lambda_1}, \ldots, \sqrt{\lambda_r}\right)$ oraz $V_1 = A^TU_1\bar{\Sigma}^{-1}$, gdzie $V_1 \in \mathbb{R}^{n\times r}$. Sprawdzamy, że $V_1^TV_1 = I_r$ (zadanie 8). Macierz V_1 można uzupełnić wektorami $v_{r+1}, \dots, v_n$ tak aby macierz $V = [V_1\ V_2]$ była ortogonalna (tw. Steinitza oraz ortogonalizacja Grama-Schmidta).

Macierze V_1, V_2 spełniają równania:

$$U_1^T A V_1 = \bar{\Sigma}, \;\; U_1^T A V_2 = \mathbb{O}. \tag{3}$$

Dla dowodu pierwszej równości zauważmy, że

$$U_1^T A V_1 = U_1^T \left(A A^T U_1 \right) \bar{\Sigma}^{-1} = U_1^T U_1 \, \bar{\Sigma}^2 \, \bar{\Sigma}^{-1} = \bar{\Sigma}.$$

Druga równość wynika z faktu, że można przyjąć $V_2 = A^T U_2 \bar{\Sigma}^{-1}$ (zadanie 9).

Na zakończenie rozważmy iloczyn $U^T\, A\, V$:

$$U^T\, A\, V = \begin{bmatrix} U_1^T \\ U_2^T \end{bmatrix} A \begin{bmatrix} V_1\ V_2 \end{bmatrix} = \begin{bmatrix} U_1^T A V_1 & U_1^T A V_2 \\ U_2^T A V_1 & U_2^T A V_2 \end{bmatrix} = \begin{bmatrix} \bar{\Sigma} & \mathbb{O} \\ \mathbb{O} & \mathbb{O} \end{bmatrix}.$$

Ponieważ macierze U, V są ortogonalne więc ostatnią równość można przepisać jako

$$A = U\, \Sigma\, V^T. \tag{4}$$

Tym samym udowodniliśmy twierdzenie

Twierdzenie 3. *Niech $A \in \mathbb{R}^{m \times n}$. Wówczas istnieją macierze $U \in \mathbb{R}^{m \times m}$, $\Sigma \in \mathbb{R}^{m \times n}$ oraz $V \in \mathbb{R}^{n \times n}$ takie, że*

$$A = U\, \Sigma\, V^T,$$

gzie macierze U, V są ortogonalne, macierz $\Sigma = diag(\sigma_1, \dots, \sigma_p)$, $\sigma_1 \geqslant \dots, \sigma_p \geqslant 0$ oraz $p = min\{m, n\}$.

Zadania

1. Załóżmy, że wartościami i wektorami własnymi macierzy A są liczby $\lambda_1, \dots, \lambda_n$ oraz wektory własne $x_1, \dots, x_n$. Niech $X = [x_1, \dots, x_n]$. Sprawdzić, że $AX = X\Lambda$, gdzie Λ to macierz diagonalna z liczbami $\lambda_1, \dots, \lambda_n$ na przekątnej.
2. Znaleźć wartości własne oraz wektory własne macierzy $A = \begin{bmatrix} 3 & 1 & 0 & 1 \\ 1 & 3 & 1 & 0 \\ 0 & 1 & 3 & 1 \\ 1 & 0 & 1 & 3 \end{bmatrix}$. Sprawdzić, czy wektory własne są ortogonalne.
3. Sprawdzić określoność poniższych macierzy:

$$Z_1 = \begin{bmatrix} 3 & 1 & 0 \\ 1 & 4 & 2 \\ 0 & 2 & 5 \end{bmatrix}, \quad Z_2 = \begin{bmatrix} \sqrt{2}+2 & 1 & 0 & 1 \\ 1 & \sqrt{2}+2 & 1 & 0 \\ 0 & 1 & \sqrt{2}+2 & 1 \\ 1 & 0 & 3 & \sqrt{2} \end{bmatrix}.$$

4. Podać wartości i wektory własne macierzy A podanej poniżej. Jakie są krotności geometryczne i algebraiczne wartości własnych?

$$A = \begin{bmatrix} a & 1 & 0 & 0 & \dots & 0 \\ 0 & a & 1 & 0 & \dots & 0 \\ 0 & 0 & a & 1 & \dots & 0 \\ \vdots & \vdots & & \ddots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & a & 1 \\ 0 & 0 & 0 & \dots & 0 & a \end{bmatrix}.$$

5. Niech A będzie symetryczną macierzą $n \times n$ o elementach rzeczywistych. Czy macierz wektorów własnych X jest nieosobliwa?
6. Weźmy macierz diagonalną, dla przykładu $\Lambda = \operatorname{diag}(16, 8, 4, 1)$ oraz macierz Q:

$$Q = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 & 0 & 0 \\ -1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & -1 & 1 \end{bmatrix}.$$

   (a) Sprawdzić, że $Q^T Q = QQ^T = I_4$;
   (b) Obliczyć macierz A: $A = Q\Lambda Q^T$;
   (c) Jakie są wartości i wektory własne macierzy A?
7. Niech $A = \begin{bmatrix} 1 & 1 & 0 & 1 \\ 0 & 0 & 0 & 1 \\ 1 & 1 & 0 & 0 \end{bmatrix}$. Znaleźć wartości własne macierzy AA^T i $A^T A$.
8. Niech $V_1 = A^T U_1 \bar{\Sigma}^{-1}$, gdzie $A \in \mathbb{R}^{m \times n}$, $U_1 \in \mathbb{R}^{m \times r}$, $\bar{\Sigma} = \operatorname{diag}\left(\sqrt{\lambda_1}, \dots, \sqrt{\lambda_r}\right)$. Kolumny macierzy U_1 to ortonormalne wektory własne, odpowiadające niezerowym wartościom własnym macierzy AA^T. Sprawdzić, że $V_1^T V_1 = I_r$.
9. Niech $V_2 = A^T U_2 \bar{\Sigma}^{-1}$. Wykazać, że $V_2 T V_2 = I$ oraz $V_1^T V_2 = \mathbb{O}$.