

Wykład 8

Rzutem ortogonalnym wektora u na wektor v nazywamy wektor $u_v \in \text{Lin}(v)$ taki, że $u - u_v \perp v$.

Twierdzenie 1 *W każdej przestrzeni euklidesowej dla dowolnych wektorów u, v istnieje dokładnie jeden rzut wektora u na wektor v .*

Dowód Jeśli $u = 0$ lub $v = 0$ to twierdzenie jest oczywiste. Niech u i v będą wektorami niezerowymi. Wektor u_v ma postać kv . Ponieważ $(u - u_v|v) = 0$ to $0 = (u - kv|v) = (u|v) - k(v|v)$ i mamy $k = \frac{(u|v)}{(v|v)}$. To dowodzi istnienia i jednoznaczności. \square

Twierdzenie 2 *Niech v_1, v_2, \dots, v_n będzie bazą ortonormalną przestrzeni euklidesowej V . Wtedy każdy wektor $v \in V$ może być przedstawiony w postaci $v = (v|v_1)v_1 + (v|v_2)v_2 + \dots + (v|v_n)v_n$.*

Dowód Ponieważ v_1, v_2, \dots, v_n jest bazą przestrzeni V to dla dowolnego wektora v istnieją skalary k_1, k_2, \dots, k_n , takie że:

$$v = k_1v_1 + k_2v_2 + \dots + k_nv_n$$

Obliczmy $(v|v_1) = (k_1v_1 + k_2v_2 + \dots + k_nv_n|v_1) = k_1(v_1|v_1) + k_2(v_2|v_1) + \dots + k_n(v_n|v_1)$, ponieważ baza jest ortonormalna to ostatnie wyrażenie jest równe k_1 , a więc:

$$(v|v_1) = k_1$$

Podobnie udowadnia się, że $k_i = (v|v_i)$. \square

Rzutem ortogonalnym wektora v na podprzestrzeń $U < V$ nazywamy wektor v_U , taki że $v_U \in U$ i $v - v_U \in U^\perp$.

Twierdzenie 3 *Niech U będzie podprzestrzenią skończenie wymiarowej, unormowanej przestrzeni euklidesowej V . Wtedy dla dowolnego wektora $v \in V$ istnieje wyznaczony jednoznacznie rzut wektora v na podprzestrzeń U .*

Dowód Niech v_1, \dots, v_n będzie bazą ortonormalną przestrzeni U . Wtedy poszukiwany wektor v_U da się zapisać w postaci $v_U = k_1v_1 + k_2v_2 + \dots + k_nv_n$. Na podstawie poprzedniego twierdzenia mamy: $k_i = (v_U|v_i)$. Ponieważ $v - v_U \in U^\perp$ to mamy $(v - v_U|v_i) = 0$ dla $i \in \{1, \dots, n\}$, zatem $(v_U|v_i) = (v|v_i)$. Zatem poszukiwanym rzutem wektora v na podprzestrzeń U jest wektor:

$$k_1v_1 + k_2v_2 + \dots + k_nv_n$$

taki, że $k_i = (v|v_i)$. \square

Dowody powyższych Twierdzeń dają również algorytmy szukania rzutów.

Nierówność Bessela

Twierdzenie 4 Niech v_1, v_2, \dots, v_n będzie bazą ortonormalną podprzestrzeni U przestrzeni euklidesowej V . Wtedy dla każdego wektora $v \in V$ zachodzi nierówność Bessela:

$$(v|v_1)^2 + (v|v_2)^2 + \dots + (v|v_n)^2 \leq \|v\|^2$$

przy czym równość zachodzi wtedy i tylko wtedy gdy $v \in U$.

Dowód Wiemy, że $(v - v_U|v - v_U) \geq 0$. Z poprzedniego twierdzenia mamy:

$$(v - v_U|v - v_U) = (v - \sum_{i=1}^n (v|v_i)v_i|v - \sum_{i=1}^n (v|v_i)v_i) = (v|v) - \sum_{i=1}^n (v|v_i)^2 \geq 0$$

a stąd mamy rządaną nierówność. \square

Nierówność Bessela mówi, że norma rzutu ortogonalnego jest mniejsza bądź równa od normy wektora.

Metoda najmniejszych kwadratów

Rozważmy układ m równań liniowych z n niewiadomymi o współczynnikach rzeczywistych:

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + \dots + a_{2n}x_n = b_2 \\ \dots\dots\dots \\ a_{m1}x_1 + a_{m2}x_2 + \dots + a_{mn}x_n = b_m \end{cases}$$

Założymy, że $m > n$, a rząd macierzy współczynników jest równy n . Może się zdarzyć, że w wyniku pewnych pomiarów fizycznych pojawi się taki układ równań. Z powodu niedokładności pomiarów taki układ może nie mieć rozwiązań. Trzeba więc znaleźć rozwiązanie przybliżone. Jedną z metod rozwiązania przybliżonego jest tak zwana metoda najmniejszych kwadratów to znaczy znalezienie takich elementów y_1, y_2, \dots, y_n , żeby suma kwadratów:

$$\sum_{i=1}^m \left(\sum_{j=1}^n a_{ij}y_j - b_i \right)^2$$

była najmniejsza. Niech $f = (b_1, b_2, \dots, b_m)$, $h_i = (a_{1i}, a_{2i}, \dots, a_{mi})$, $i = 1, 2, \dots, n$ będą wektorami z przestrzeni \mathbb{R}^m . Wtedy powyższa suma przyjmuje najmniejszą wartość jeśli (y_1, y_2, \dots, y_n) jest rzutem wektora f na podprzestrzeń rozpiętą na wektorach h_1, h_2, \dots, h_n . Zatem elementy y_1, y_2, \dots, y_n

spełniają układ:

$$\begin{cases} (h_1|h_1)y_1 + (h_1|h_2)y_2 + \dots + (h_1|h_n)y_n = (h_1|f) \\ (h_2|h_1)y_1 + (h_2|h_2)y_2 + \dots + (h_2|h_n)y_n = (h_2|f) \\ \dots\dots\dots \\ (h_n|h_1)y_1 + (h_n|h_2)y_2 + \dots + (h_n|h_n)y_n = (h_n|f) \end{cases}$$

gdzie $(|)$ oznacza standardowy iloczyn skalarny w przestrzeni \mathbb{R}^n (jest to również sposób wyznaczania rzutu wektora na podprzestrzeń). Wyznacznik macierzy współczynników tego układu nazywamy wyznacznikiem Grama.

Zadanie Rozwiązać metodą najmniejszych kwadratów układ:

$$\begin{cases} x + y = 3 \\ 2x + y = 5 \\ x - y = 2 \end{cases}$$

Rozwiązanie Układ ten jest sprzeczny. Poszukamy więc takiego rozwiązania przybliżonego x_1, y_1 aby wyrażenie:

$$(x_1 + y_1 - 3)^2 + (2x_1 + y_1 - 5)^2 + (x_1 - y_1 - 2)^2$$

przyjmowało najmniejszą wartość. Do tego wykorzystamy metodę najmniejszych kwadratów. Przyjmijmy $f = (3, 5, 2)$, $h_1 = (1, 2, 1)$, $h_2 = (1, 1, -1)$. Obliczmy:

$$\begin{aligned} (h_1|h_1) &= 6, & (h_1|h_2) &= 2, & (h_1|f) &= 15, \\ (h_2|h_1) &= 2, & (h_2|h_2) &= 3, & (h_2|f) &= 6 \end{aligned}$$

Zatem przybliżone rozwiązanie (w sensie najmniejszych kwadratów) spełnia układ:

$$\begin{cases} 6x_1 + 2y_1 = 15 \\ 2x_1 + 3y_1 = 6 \end{cases}$$

Formy dwuliniowe

Niech V będzie przestrzenią liniową nad ciałem K ($=\mathbb{R}$ lub \mathbb{C}). Funkcję:

$$f : V \times V \rightarrow K$$

nazywamy formą dwuliniową jeśli $\forall u, v, w \in V, \forall k \in K$ mamy:

- (i) $f(u + v, w) = f(u, w) + f(v, w)$,
- (ii) $f(ku, v) = kf(u, v)$,
- (iii) $f(u, v + w) = f(u, v) + f(u, w)$,
- (iv) $f(u, kv) = \bar{k}f(u, v)$.

Przykładem formy dwuliniowej może być iloczyn skalarny w przestrzeni euklidesowej.

Przyjmijmy teraz $V = \mathbb{R}^n$, $u = (x_1, x_2, \dots, x_n)$, $v = (y_1, y_2, \dots, y_n)$ wtedy jeśli f jest formą dwuliniową to istnieją współczynniki g_{ij} , $i, j \in \{1, \dots, n\}$, że:

$$f(u, v) = \sum_{i=1}^n \sum_{j=1}^n g_{ij} x_i y_j$$

dla każdych u, v . Wtedy dla macierzy $G = [g_{ij}]$ mamy:

$$f(u, v) = (x_1, x_2, \dots, x_n) G \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}$$

Można zauważyć, że jeśli e_1, e_2, \dots, e_n oznacza bazę kanoniczną to $g_{ij} = f(e_i, e_j)$.

Formę f nazywamy **formą symetryczną** jeśli $\forall u, v \in V$ mamy $f(u, v) = f(v, u)$.

Stwierdzenie 1 *Forma f jest symetryczna wtedy i tylko wtedy gdy macierz G tej formy jest symetryczna (to znaczy $G = G^T$).*

Niech $A = [a_{ij}]$ będzie macierzą nad ciałem liczb zespolonych. Wtedy $A^* = \overline{A^T} = [\overline{a_{ij}}]^T$. Macierz A nazywamy **macierzą hermitowską** jeśli $A^* = A$.

Formę f nazywamy formą hermitowską jeśli jej macierz współczynników G jest hermitowska. Zatem forma hermitowska spełnia tożsamość: $\forall u, v \in V : f(u, v) = \overline{f(v, u)}$.

Macierz A nazywamy macierzą ortogonalną (unitarną) jeśli $A^T A = I$ ($AA^* = I$).

Macierz A nazywamy macierzą normalną jeśli $AA^T = A^T A$ ($AA^* = A^* A$).

Forma kwadratowa

Funkcję:

$$g : V \rightarrow \mathbb{R}$$

nazywamy **formą kwadratową** jeśli istnieje forma dwuliniowa $f : V \times V \rightarrow \mathbb{R}$, taka że $\forall v \in V$, $g(v) = f(v, v)$.

Twierdzenie 5 *Jeśli g jest formą kwadratową w przestrzeni $V = \mathbb{R}^n$ to istnieje dokładnie jedna symetryczna forma dwuliniowa f , że dla każdego $v \in V$ mamy $g(v) = f(v, v)$.*

Symetryczną formę dwuliniową f nazywamy **formą biegunową** formy kwadratowej g .

Dowód Wystarczy przyjąć:

$$f(u, v) = \frac{1}{2}(g(u + v) - g(u) - g(v)) \square$$

Zadanie Niech

$$g(x_1, x_2, x_3) = (x_1, x_2, x_3) \begin{bmatrix} 1 & 2 & 3 \\ 1 & -1 & 2 \\ 2 & 1 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$$

Znaleźć biegunową formę dwuliniową tej formy kwadratowej.

Można zauważyć, że jeśli g jest formą kwadratową w przestrzeni \mathbb{R}^n to:

$$g(x_1, x_2, \dots, x_n) = \sum_{i=1}^n g_{ii}x_i^2 + 2 \sum_{i=1, j=1, i < j}^n g_{ij}x_ix_j = (x_1, x_2, \dots, x_n)G \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$$

gdzie G jest macierzą symetryczną, o współczynnikach g_{ij} . Wtedy biegunowa forma dwuliniowa ma postać:

$$f(u, v) = (x_1, x_2, \dots, x_n)G \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}$$

Każdą formę kwadratową możemy skojarzyć z pewną macierzą symetryczną G .

Oznaczmy przez G_i macierz wymiaru $i \times i$, która powstaje z macierzy G przez wykreślenie wierszy i kolumn o numerach $i + 1, i + 2, \dots, n$. Wtedy $\det G_i$ nazywamy minorem głównym macierzy G .

Formę kwadratową $g : \mathbb{R}^n \rightarrow \mathbb{R}$ nazywamy **dodatnio określona** jeśli $\forall v \in \mathbb{R}^n \setminus \{(0, 0, \dots, 0)\}$ mamy $g(v) > 0$.

Twierdzenie 6 *Forma g wyznaczona przez macierz symetryczną G jest dodatnio określona wtedy i tylko wtedy gdy wszystkie minory główne $\det G_i$ są dodatnie.*