Rozdział 11

Przestrzenie unitarne

Niech X będzie rzeczywistą przestrzenią liniową.

Definicja 11.1 (iloczyn skalarny). *Funkcję $s : X \times X \to \mathbb{R}$ spełniającą warunki:*

(a) $\forall \alpha, \beta \in \mathbb{R}, \forall x, y, z \in X :$
$$s(\alpha x + \beta y, z) = \alpha s(x, z) + \beta s(y, z);$$

(b) $\forall x, y \in X :$
$$s(x, y) = s(y, x);$$

(c) $\forall x \in X :$
$$s(x, x) \geqslant 0 \quad oraz \quad s(x, x) = 0 \Leftrightarrow x = 0$$

nazywamy **iloczynem skalarnym**. *Parę (X, s) nazywamy* **przestrzenią unitarną**.

Iloczyn skalarny wektorów x, y będziemy również oznaczać jako $\langle x, y \rangle$ lub $x \circ y$.

Przykład 11.1. *Odwzorowanie*
$$(x_1, \ldots, x_n) \circ (y_1, \ldots, y_n) = \sum_{k=1}^{n} x_k y_k \tag{11.1}$$
to naturalny iloczyn skalarny w $\mathbb{R}^n$.

Przykład 11.2. *Odwzorowanie*
$$s(f, g) = \int_a^b f(x) g(x) \, dx$$
jest iloczynem skalarnym w przestrzeni $\mathcal{L}_2(a, b)$ tzw. funkcji całkowalnych z kwadratem, tj.
$$\mathcal{L}_2(a, b) = \left\{ f : (a, b) \to \mathbb{R} : \int_a^b f^2(x) \, dx < +\infty \right\}.$$

Przykład 11.3. *Odwzorowanie*
$$s(A, B) = \operatorname{tr}\left(A^T B\right)$$
jest iloczynem skalarnym w przestrzeni $\mathbb{R}^{m \times n}$.

11.1. Norma określona przez iloczyn skalarny

Niech X będzie rzeczywistą przestrzenią liniową.

Definicja 11.2 (norma). *Funkcję $\|\cdot\| : X \to \mathbb{R}$ spełniającą warunki*

(a) $\forall x \in X$:
$$\|x\| \geqslant 0 \quad oraz \quad \|x\| = 0 \Leftrightarrow x = 0;$$

(b) $\forall x \in X, \forall \alpha \in \mathbb{R}$:
$$\|\alpha x\| = |\alpha| \cdot \|x\|\,;$$

(c) $\forall x, y \in X$:
$$\|x + y\| \leqslant \|x\| + \|y\|$$

nazywamy **normą**. *Parę $(X, \|\cdot\|)$ nazywamy* **przestrzenią unormowaną**.

Warunki występujące w powyższej definicji to naturalne wymagania stawiane przed funkcją mierzącą długość wektorów.

Przykład 11.4. *Każda z poniższych par $(X, \|\cdot\|)$ jest przestrzenią unormowaną:*

a) $X = \mathbb{R}^n$ *z normą*
$$\|(x_1, \ldots, x_n)\| = \sqrt{x_1^2 + \cdots + x_n^2};$$

b) $X = \mathcal{C}_{[a,b]}$ *z normą*
$$\|f\| = \max\{|f(x)| : x \in [a, b]\}\,;$$

c) $X = \mathbb{R}^{n\times n}$ *z normą*
$$\|A\| = \sqrt{\max\{|\lambda| : \lambda \in \sigma(A^T A)\}}.$$

Wykażemy teraz, że jeżeli w rzeczywistej przestrzeni liniowej X zdefiniowano iloczyn skalarny s to funkcja $\|\cdot\|_s : X \to \mathbb{R}$ określona wzorem
$$\|x\|_s := \sqrt{s(x, x)} \tag{11.2}$$
jest normą w X.

Zauważmy na początek, że na podstawie warunku (c) definicji 10.1, funkcja $\|\cdot\|_s$ jest dobrze określona – wartości $s(x, x)$ są nieujemne; ten sam warunek gwarantuje również, że punkt (a) definicji 10.2 jest spełniony. Ponieważ
$$\begin{aligned}\|\alpha x\|_s &= \sqrt{s(\alpha x, \alpha x)} = \sqrt{\alpha s(x, \alpha x)} = \sqrt{\alpha s(\alpha x, x)} = \sqrt{\alpha^2 s(x, x)}\\ &= |\alpha|\sqrt{s(x, x)} = |\alpha| \cdot \|x\|_s\end{aligned}$$
zatem punkt (b) również zachodzi. Zanim uzasadnimy punkt (c), wykażemy następujące

Twierdzenie 11.1 (nierówność Schwarza). *Dla dowolnych wektorów x, y rzeczywistej przestrzeni liniowej wyposażonej w iloczyn skalarny s zachodzi*
$$|s(x, y)| \leqslant \|x\|_s \cdot \|y\|_s\,. \tag{11.3}$$

Dowód: Dla dowolnych ustalonych wektorów x, y rozważmy funkcję
$$\varphi(t) = s(x + ty, x + ty)$$
zmiennej $t \in \mathbb{R}$. Bez straty ogólności możemy założyć, że $y \neq 0$. Z warunku (c) definicji 10.1 wynika, że
$$\varphi(t) \geqslant 0, \quad \text{dla } t \in \mathbb{R}. \tag{11.4}$$

Ponieważ

$$\varphi(t) = t^2 s(y, y) + 2ts(x, y) + s(x, x)$$

zatem φ jest funkcją kwadratową, która – wobec warunku (11.4) – ma niedodatni wyróżnik, tj.

$$\triangle = 4s^2(x, y) - 4s(x, x)\, s(y, y) \leqslant 0,$$

lub równoważnie

$$s^2(x, y) \leqslant s(x, x)\, s(y, y).$$

Stąd wynika zależność (11.3). ■

Dla dowolnych $x, y \in X$ mamy więc

$$\begin{aligned} \|x+y\|_s^2 &= s(x+y, x+y) = s(x, x) + 2s(x, y) + s(y, y) \leqslant \\ &\leqslant s(x, x) + 2\sqrt{s(x, x)\, s(y, y)} + s(y, y) = \\ &= \left(\sqrt{s(x, x)} + \sqrt{s(y, y)}\right)^2 = \left(\|x\|_s + \|y\|_s\right)^2. \end{aligned}$$

Oznacza to, że wzór (11.2) definiuje normę w dowolnej przestrzeni unitarnej.

11.2. Ortogonalność

Z nierówności Schwarza wynika, że dla niezerowych wektorów x, y rzeczywistej przestrzeni X:

$$-1 \leqslant \frac{s(x, y)}{\|x\|_s \cdot \|y\|_s} \leqslant 1.$$

Wynika stąd, że iloraz $s(x, y) / (\|x\|_s \cdot \|y\|_s)$ jest kosinusem ściśle określonego kąta $\measuredangle(x, y)$:

$$\cos\measuredangle(x, y) = \frac{s(x, y)}{\|x\|_s \cdot \|y\|_s}, \quad \measuredangle(x, y) \in [0, \pi].$$

Na podstawie definicji przyjmujemy, że jest to kąt między wektorami x oraz y. Mamy więc

$$s(x, y) = \|x\|_s \cdot \|y\|_s \cdot \cos\measuredangle(x, y).$$

Definicja 11.3. *Dwa wektory nazywamy* **ortogonalnymi***, jeżeli ich iloczyn skalarny jest równy zero.*

Wektor zerowy jest jedynym wektorem prostopadłym do każdego wektora (również do siebie samego).

Przykład 11.5. *Rozważmy przestrzeń $\mathbb{R}^3$ z naturalnym iloczynem skalarnym (zob. przykład 10.1). Dla wektorów $v_1 = \left(1, \sqrt{3}, 0\right), v_2 = (0, 1, 0)$ mamy*

$$\cos\varphi = \frac{s(v_1, v_2)}{\|v_1\|_s \cdot \|v_2\|_s} = \frac{1 \cdot 0 + \sqrt{3} \cdot 1 + 0 \cdot 0}{\sqrt{1+3} \cdot \sqrt{1}} = \frac{\sqrt{3}}{2}.$$

Tym samym $\measuredangle(v_1, v_2) = \varphi = \frac{\pi}{6}$.

Przykład 11.6. *W przestrzeni $\pi_n(\mathbb{R})$ definiujemy funkcje:*

$$\begin{aligned} s_1(f, g) &= \int_{-1}^{1} f(x)\, g(x)\, dx, \\ s_2(f, g) &= \sum_{k=0}^{n} f^{(k)}(-1)\, g^{(k)}(-1). \end{aligned}$$

Przestrzenie $(\pi_n(\mathbb{R}), s_1)$ *oraz* $(\pi_n(\mathbb{R}), s_2)$ *są przestrzeniami unitarnymi. Niech* $f(x) = 2x+1$ *oraz* $g(x) = 3x^2 - 3x + 1$. *Wówczas*

$$s_1(f,g) = \int_{-1}^{1} (2x+1)\left(3x^2-3x+1\right)dx = \int_{-1}^{1} 6x^3 - 3x^2 - x + 1dx = 0$$

oraz

$$s_2(f,g) = f(-1)\,g(-1) + f'(-1)\,g'(-1) + f''(-1)\,g''(-1) = -25.$$

Oznacza to, że rozważane wielomiany są ortogonalne w przestrzeni $(\pi_n(\mathbb{R}), s_1)$, *natomiast nie są ortogonalne w przestrzeni* $(\pi_n(\mathbb{R}), s_2)$.

11.3. Ortogonalizacja Grama–Schmidta

Rozważmy ciąg $v_1, \ldots, v_n$ wektorów rzeczywistej przestrzeni liniowej X wyposażonej w iloczyn skalarny s. Jeżeli

$$v_i \perp v_j \quad \text{dla} \quad i \neq j,$$

to mówimy, że ciąg $v_1, \ldots, v_n$ jest **ciągiem wektorów ortogonalnych**. Jeżeli dodatkowo $\|v_i\|_s = 1$ $(i = 1, \ldots, n)$, to ciąg ten nazywamy **ciągiem ortonormalnym**.

Przypuśćmy, że wektory $v_1, \ldots, v_n$ stanowią bazę przestrzeni X. Podamy teraz algorytm modyfikujący tę bazę w taki sposób, że nowo otrzymana baza $\tilde{v}_1, \ldots, \tilde{v}_n$ jest bazą ortonormalną przestrzeni X.

Twierdzenie 11.2 (algorytm Grama–Schmidta). *Niech ciąg* $v_1, \ldots, v_n$ *stanowi bazę rzeczywistej przestrzeni* X *wyposażonej w iloczyn skalarny* s. *Wówczas ciąg wektorów* $\tilde{v}_1, \ldots, \tilde{v}_n$ *określonych wzorami*

$$\tilde{v}_1 = \frac{v_1}{\|v_1\|_s}, \quad \tilde{v}_k = \frac{v_k - \sum_{i=1}^{k-1} s(v_k, \tilde{v}_i)\,\tilde{v}_i}{\left\| v_k - \sum_{i=1}^{k-1} s(v_k, \tilde{v}_i)\,\tilde{v}_i \right\|_s}, \quad k = 2, \ldots, n$$

jest taki, że:

a) *dla każdego* $k \in \{1, \ldots, n\} : \operatorname{span}\{v_1, \ldots, v_k\} = \operatorname{span}\{\tilde{v}_1, \ldots, \tilde{v}_k\}$;
b) *ciąg* $\tilde{v}_1, \ldots, \tilde{v}_n$ *jest bazą ortonormalną przestrzeni* X.

Dowód: Dowód poprowadzimy przez indukcję względem n. Dla $n = 1$ twierdzenie jest prawdziwe, tj. $\|\tilde{v}_1\|_s = 1$ oraz $\operatorname{span}\{v_1\} = \operatorname{span}\{\tilde{v}_1\}$. Przypuśćmy więc, że układ $\tilde{v}_1, \ldots, \tilde{v}_{k-1}$ jest bazą ortonormalną przestrzeni $\operatorname{span}\{v_1, \ldots, v_{k-1}\}$. Niech $v = v_k - \sum_{i=1}^{k-1} s(v_k, \tilde{v}_i)\,\tilde{v}_i$. Wykażemy teraz, że wektor v jest ortogonalny do wektorów $\tilde{v}_1, \ldots, \tilde{v}_{k-1}$. Mamy dla $j = 1, \ldots, k-1$:

$$\begin{aligned} s(v, \tilde{v}_j) &= s\left(v_k - \textstyle\sum_{i=1}^{k-1} s(v_k, \tilde{v}_i)\,\tilde{v}_i, \tilde{v}_j\right) = s(v_k, \tilde{v}_j) - \textstyle\sum_{i=1}^{k-1} s(v_k, \tilde{v}_i)\, s(\tilde{v}_i, \tilde{v}_j) \\ &= s(v_k, \tilde{v}_j) - s(v_k, \tilde{v}_j) = 0. \end{aligned}$$

Zauważmy ponadto, że $v \neq 0$. W przeciwnym przypadku mielibyśmy

$$v_k \in \operatorname{span}\{\tilde{v}_1, \ldots, \tilde{v}_{k-1}\} = \operatorname{span}\{v_1, \ldots, v_{k-1}\}$$

wbrew liniowej niezależności wektorów $v_1, \ldots, v_n$. Możemy więc przyjąć

$$\tilde{v}_k := \frac{v}{\|v\|_s} = \frac{v_k - \sum_{i=1}^{k-1} s(v_k, \tilde{v}_i)\,\tilde{v}_i}{\left\| v_k - \sum_{i=1}^{k-1} s(v_k, \tilde{v}_i)\,\tilde{v}_i \right\|_s}.$$

Tym samym wektory $\tilde{v}_1, \ldots, \tilde{v}_k$ tworzą układ wektorów ortonormalnych oraz rozpinają tę samą przestrzeń co wektory $v_1, \ldots, v_k$. ■

Przykład 11.7. *Niech* $X = \left\{(x,y,z,w) \in \mathbb{R}^4 : x+2y+3z+w=0, x+y+z=0\right\}$. *Łatwo stwierdzić, że*

$$X = \{(-y-z, y, z, -y-2z) : y, z \in \mathbb{R}\}.$$

Jest to więc podprzestrzeń liniowa przestrzeni $\mathbb{R}^4$, *a ponieważ*

$$(-y-z, y, z, -y-2z) = y\,(-1,1,0,-1) + z\,(-1,0,1,-2)$$

zatem jej bazą są wektory $e_1 = (-1,1,0,-1)\,, e_2 = (-1,0,1,-2)$. *Wyznaczymy bazę ortonormalną przestrzeni* X *w sensie naturalnego iloczynu skalarnego indukowanego z przestrzeni* $\mathbb{R}^4$. *Z twierdzenia 11.2 wynika, że szukana baza* $\tilde{e}_1, \tilde{e}_2$ *może być wyznaczona ze wzorów*

$$\tilde{e}_1 = \frac{e_1}{\|e_1\|}, \quad \tilde{e}_2 = \frac{e_2 - (e_2 \circ \tilde{e}_1)\,\tilde{e}_1}{\|e_2 - (e_2 \circ \tilde{e}_1)\,\tilde{e}_1\|}.$$

Ponieważ $\|e_1\| = \sqrt{e_1 \circ e_1} = \sqrt{3}$, *zatem* $\tilde{e}_1 = \frac{\sqrt{3}}{3}\,(-1,1,0,-1)$. *Podobnie, ponieważ*

$$e_2 - (e_2 \circ \tilde{e}_1)\,\tilde{e}_1 = (-1,0,1,-2) - \left((-1,0,1,-2) \circ \frac{\sqrt{3}}{3}\,(-1,1,0,-1)\right)\frac{\sqrt{3}}{3}\,(-1,1,0,-1) =$$
$$= (-1,0,1,-2) - (-1,1,0,-1) = (0,-1,1,-1)$$

więc

$$\tilde{e}_2 = \frac{(0,-1,1,-1)}{\sqrt{(0,-1,1,-1) \circ (0,-1,1,-1)}} = \frac{\sqrt{3}}{3}\,(0,-1,1,-1).$$

Wektory

$$\tilde{e}_1 = \frac{\sqrt{3}}{3}\,(-1,1,0,-1), \qquad \tilde{e}_2 = \frac{\sqrt{3}}{3}\,(0,-1,1,-1)$$

są bazą ortonormalną przestrzeni X.

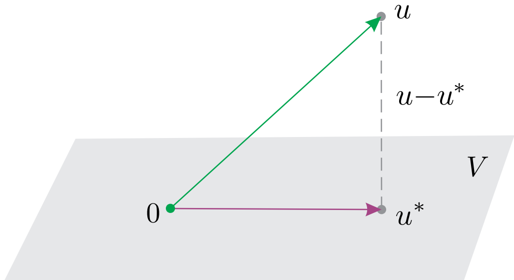
11.4. Rzut prostopadły na podprzestrzeń liniową

Niech V będzie n-wymiarową podprzestrzenią liniową rzeczywistej przestrzeni unitarnej (X, s). Rozważmy dowolny wektor $u \in X \backslash V$.

Definicja 11.4 (rzut ortogonalny). *Wektor* $u^* \in V$ *spełniający warunek*

$$\forall v \in V : \quad u - u^* \perp v \tag{11.5}$$

nazywamy **rzutem ortogonalnym wektora u na podprzestrzeń V**.



Wykres 3. Rzut ortogonalny wektora u na podprzestrzeń liniową V.

Przypuśćmy, że wektory $u_1, \ldots, u_n$ są bazą podprzestrzeni V. Warunek (11.5) równoważny jest wówczas warunkowi

$$u - u^* \perp u_i \quad (i = 1, \ldots, n). \tag{11.6}$$

Ponieważ $u^* \in V$ zatem istnieją skalary $\alpha_i \in \mathbb{R}$ $(i = 1, \ldots, n)$ dla których

$$u^* = \alpha_1^* u_1 + \ldots + \alpha_n^* u_n.$$

Aby wyznaczyć rzut ortogonalny u^* wektora u wystarczy więc wyznaczyć jego współrzędne $\alpha_1^*, \ldots, \alpha_n^*$ względem dowolnej bazy przestrzeni V. Zależność (11.6) oznacza, że dla $i = 1, \ldots, n$:

$$0 = s(u - u^*, u_i) = s(u - \textstyle\sum_{k=1}^n \alpha_k^* u_k, u_i) = s(u, u_i) - \sum_{k=1}^n \alpha_k^* s(u_k, u_i),$$

lub równoważnie

$$\textstyle\sum_{k=1}^n s(u_i, u_k)\, \alpha_k^* = s(u, u_i).$$

Szukane wartości $\alpha_1^*, \ldots, \alpha_n^*$ są więc rozwiązaniem układu równań liniowych

$$\begin{bmatrix} s(u_1, u_1) & \cdots & s(u_1, u_n) \\ \vdots & \ddots & \vdots \\ s(u_n, u_1) & \cdots & s(u_n, u_n) \end{bmatrix} \begin{bmatrix} \alpha_1^* \\ \vdots \\ \alpha_n^* \end{bmatrix} = \begin{bmatrix} s(u, u_1) \\ \vdots \\ s(u, u_n) \end{bmatrix}. \tag{11.7}$$

Macierz $G = [s(u_i, u_j)]$ tego układu – tzw. **macierz Grama** – posiada wiele ważnych i interesujących własności. Można na przykład pokazać, że jej wyznacznik jest różny od zera wtedy i tylko wtedy, gdy wektory $u_1, \ldots, u_n$ są liniowo niezależne (zob. zestaw 12, zad. 5). Postać macierzy G zależy od wyboru bazy przestrzeni V. W przypadku, gdy baza $u_1, \ldots, u_n$ jest bazą ortonormalną, macierz G jest macierzą jednostkową, a rozwiązaniem układu (11.7) są skalary

$$\alpha_i^* = s(u, u_i), \quad (i = 1, \ldots, n).$$

Wynika stąd następujące

Twierdzenie 11.3. *Niech $u_1, \ldots, u_n$ będzie bazą ortonormalną podprzestrzeni V przestrzeni liniowej X wyposażonej w iloczyn skalarny s. Dla dowolnego wektora $u \in X$ istnieje dokładnie jeden wektor $u^* \in V$ będący rzutem ortogonalnym wektora u na podprzestrzeń V. Wektor ten określony jest wzorem:*

$$u^* = \sum_{i=1}^n s(u, u_i)\, u_i. \tag{11.8}$$

Przykład 11.8. *Wyznaczymy rzut ortogonalny wektora $u = (1, 1, 1, 1)$ na podprzestrzeń X przestrzeni $\mathbb{R}^4$ rozważaną w przykładzie 11.7 (z naturalnym iloczynem skalarnym). Przypomnijmy, że ortonormalną bazę X stanowią wektory*

$$u_1 = \frac{\sqrt{3}}{3}(-1, 1, 0, -1), \quad u_2 = \frac{\sqrt{3}}{3}(0, -1, 1, -1).$$

Poszukiwany wektor u^* wyznaczymy ze wzoru (11.8). Otrzymujemy

$$u^* = (u \circ u_1)\, u_1 + (u \circ u_2)\, u_2 = -\frac{1}{3}(-1, 1, 0, -1) - \frac{1}{3}(0, -1, 1, -1) = \frac{1}{3}(1, 0, -1, 2).$$