Wykład 7 (4 IV 2012)

Bazy ortonormalne i macierze ortogonalne

Przypomnienie:
Dla (dowolnej) podprzestrzeni liniowej $W \subset \mathbb{R}^n$ zostało skonstruowane odwzorowanie $P_W : \mathbb{R}^n \to \mathbb{R}^n$ nazywane *rzutem ortogonalnym na podprzestrzeń* W. Do jego konstrukcji wykorzystana była baza ortonormalna $\{u_1, \ldots, u_k\}$ przestrzeni W, a sam rzut P_W był zadany wzorem (por. wzór (5.6))

$$\mathbb{R}^n \ni x \mapsto P_W(x) = \sum_{j=1}^{k} (x \cdot u_j) u_j \in W.$$

Odwzorowanie P_W realizuje rozkład wektora x na składniki należące do ortogonalnie dopełniających podprzestrzeni W i $W^\perp$ w postaci $x = P_W(x) + x - P_W(x)$ i ma następującą własność minimalności

$$d(x, W) = \inf\{\, d(x, z) \mid z \in W \,\} = |x - P_W(x)|.$$

7.1 Konstrukcja bazy ortonormalnej metodą Grama–Schmidta

Twierdzenie 8 (O ortonormalizacji układu liniowo niezależnych wektorów)
Jeśli $\{v_1, \ldots, v_m\}$ *jest liniowo niezależnym układem wektorów przestrzeni* $\mathbb{R}^n$*, to istnieje układ ortonormalny* $\{u_1, \ldots, u_m\}$*, taki że dla każdego* $k \in \{1, \ldots, m\}$ *wektory* $\{v_1, \ldots, v_k\}$ *i* $\{u_1, \ldots, u_k\}$ *rozpinają tę samą podprzestrzeń w* $\mathbb{R}^n$*.*

ZOB:
http://wazniak.mimuw.edu.pl/images/9/91/kontener.html?nazwa_animacji=/images/2/2d/Ag_10_2a

Przedstawimy tutaj zarys konstrukcji układu ortonormalnego $\{u_1, \ldots, u_m\}$, nazywanej zazwyczaj *metodą ortonormalizacji Grama–Schmidta*. Jest to konstrukcja rekurencyjna, polegająca na dołączaniu krok po kroku kolejnego elementu do już skonstruowanego w poprzednich krokach układu ortonormalnego, przy tym tak, aby w każdym kroku (to jest dla każdego $k = 1, \ldots, m$) był spełniony warunek

$$\mathrm{lin}\{v_1, \ldots, v_k\} = \mathrm{lin}\{u_1, \ldots, u_k\}. \tag{7.1}$$

Zamiast podania formalnego dowodu prześledzimy początkowe etapy tej konstrukcji. Dla zapewnienia pełniejszego obrazu sytuacji założymy, że wyjściowy układ $\{v_1, \ldots, v_k\}$ zawiera nie mniej niż 3 wektory, $m \geqslant 3$. Dla $k = 1$ równość (7.1) oznacza, że $u_1 = \lambda v_1$, a wobec unormowania $|u_1| = 1$ otrzymujemy

$$u_1 = \frac{1}{|v_1|} v_1. \tag{7.2}$$

Dla $k = 2$ przy spełnieniu równości (7.2) warunek (7.1) będzie także spełniony, jeśli przyjąć

$$u_2 = \alpha_1 v_1 + \alpha_2 v_2,$$

Współczynniki α_1, α_2 dobieramy tak, aby wektory u_1 i u_2 tworzyły układ ortonormalny. Wykorzystując równość (7.2) zapisujemy warunek ortogonalności w postaci równania

$$u_2 \cdot u_1 = \alpha_1(v_1 \cdot u_1) + \alpha_2(v_2 \cdot u_1) = \alpha_1|v_1| + \alpha_2(v_2 \cdot u_1) = 0.$$

A zatem $\alpha_1 = -\alpha_2|v_1|^{-1}(v_2 \cdot u_1)$ i jeszcze raz korzystając z (7.2) otrzymujemy

$$u_2 = \alpha_2(v_2 - (v_2 \cdot u_1)u_1). \tag{7.3}$$

Zaobserwujmy, że wyrażenie w nawiasie po prawej stronie jest różnicą v_2 i jego rzutu na kierunek v_1, jest więc różne od zera wobec liniowej niezależności v_1, v_2. Można więc unormować wektor u_2 przyjmując

$$\alpha_2 = \frac{1}{|v_2 - (v_2 \cdot u_1)u_1|}, \tag{7.4}$$

co po podstawieniu obliczonej wartości α_2 daje

$$\begin{aligned} u_2 &= \frac{1}{|v_2 - (v_2 \cdot u_1)u_1|}(v_2 - \frac{v_2 \cdot v_1}{|v_1|^2}v_1), \\ u_2 &= \frac{1}{|v_2 - (v_2 \cdot u_1)u_1|}(v_2 - (v_2 \cdot u_1)u_1). \end{aligned} \tag{7.5}$$

To kończy konstrukcję drugiego wektora poszukiwanego układu ortonormalnego, która zapewnia spełnienie warunku (7.1) dla $k = 2$. Krok trzeci jest rozwinięciem poprzedniego — teraz przyjmujemy

$$u_3 = \alpha_1 v_1 + \alpha_2 v_2 + \alpha_3 v_3, \qquad \text{gdzie} \quad \alpha_3 \neq 0.$$

Na mocy właśnie zakończonego etapu konstrukcji (w istocie wykorzystujemy tu (7.1) dla $k = 2$ jako założenie indukcyjne) możemy zastąpić kombinację wektorów v_1, v_2 przez kombinację wektorów u_1, u_2 otrzymując

$$u_3 = \gamma_1 u_1 + \gamma_2 u_2 + \alpha_3 v_3, \tag{7.6}$$

a nakładając warunek ortogonalności u_3 do u_1 i u_2 otrzymamy równania

$$\begin{aligned} u_3 \cdot u_1 &= \gamma_1 + \alpha_3(v_3 \cdot u_1) = 0, \\ u_3 \cdot u_2 &= \gamma_2 + \alpha_3(v_3 \cdot u_2) = 0, \end{aligned}$$

skąd po wyznaczeniu γ_1 i γ_2 dochodzimy, podobnie jak powyżej, do zależności

$$u_3 = \alpha_3(v_3 - (v_3 \cdot u_1)u_1 - (v_3 \cdot u_2)u_2). \tag{7.7}$$

Na zakończenie wyznaczamy współczynnik α_3, który zapewni normalizację u_3

$$\alpha_3 = \frac{1}{|v_3 - (v_3 \cdot u_1)u_1 - (v_3 \cdot u_2)u_2|}. \tag{7.8}$$

Analiza wykonanej konstrukcji upewnia, że przy zastosowaniu tej procedury do liniowo niezależnego układu $\{v_1, \ldots, v_k\}$ można ją kontynuować tak długo, aż otrzymamy układ ortonormalny $\{u_1, \ldots, u_m\}$ o tej samej liczbie elementów co układ wyjściowy.

Dodajmy jeszcze następującą obserwację. W równościach (7.3) i (7.7) wyrażenie stojące w nawiasie po prawej stronie jest różnicą między wektorem v_2, odpowiednio v_3, wziętym z bazy $\{v_1, \ldots, v_m\}$, a rzutem ortogonalnym tego wektora na podprzestrzeń rozpiętą przez „wcześniejsze” wektory tej bazy. Inaczej mówiąc, ta różnica jest składową danego wektora bazy $\{v_1, \ldots, v_m\}$ w kierunku prostopadłym do podprzestrzeni rozpiętej przez „wcześniejsze” wektory tej bazy;

$$
\begin{aligned}
P_{\mathrm{lin}\{v_1\}}(v_2) &= (v_2 \cdot u_1)u_1, & v_2 - (v_2 \cdot u_1)u_1 &= v_2 - P_{\mathrm{lin}\{v_1\}}(v_2) \perp u_1,\\
P_{\mathrm{lin}\{v_1, v_2\}}(v_3) &= (v_3 \cdot u_1)u_1 + (v_3 \cdot u_2)u_2, & v_3 - (v_3 \cdot u_1)u_1 - (v_3 \cdot u_2)u_2 &= v_3 - P_{\mathrm{lin}\{v_1, v_2\}}(v_3) \perp \{u_1,\ u_2\}.
\end{aligned}
$$

i td.

Przykład 7.1.1 Niech $v_1, v_2 \in \mathbb{R}^2$ będą wektorami

$$v_1 = \begin{bmatrix}1\\1\end{bmatrix}, \qquad v_2 = \begin{bmatrix}3\\-4\end{bmatrix},$$

Normalizujemy v_1 otrzymując

$$u_1 = \frac{1}{\sqrt{2}}\begin{bmatrix}1\\1\end{bmatrix}. \tag{7.9}$$

Zgodnie z drugim krokiem konstrukcji G–S rzutujemy v_2 na kierunek v_1 i odejmujemy rzut od wektora v_2, co daje, po przeniesieniu stałej normalizacyjnej α_2 na lewą stronę równości

$$\frac{1}{\alpha_2}u_2 = v_2 - \frac{(v_2 \cdot v_1)}{|v_1|^2}v_1 = v_2 - (v_2 \cdot u_1)u_1 = \begin{bmatrix}3\\-4\end{bmatrix} - (-\frac{1}{\sqrt{2}})\frac{1}{\sqrt{2}}\begin{bmatrix}1\\1\end{bmatrix} = \frac{7}{2}\begin{bmatrix}1\\-1\end{bmatrix}.$$

Po wyznaczeniu stałej α_2 z warunku normalizacji możemy wyrazić u_2 jako kombinację liniową wektorów wyjściowej bazy;

$$u_2 = \frac{\sqrt{2}}{7}\begin{bmatrix}3\\-4\end{bmatrix} + \frac{1}{7\sqrt{2}}\begin{bmatrix}1\\1\end{bmatrix} = \frac{1}{\sqrt{2}}\begin{bmatrix}1\\-1\end{bmatrix}. \tag{7.10}$$

A zatem w wyniku ortonormalizacji Grama–Schmidta bazy

$$v_1 = \begin{bmatrix}1\\1\end{bmatrix}, \qquad v_2 = \begin{bmatrix}3\\-4\end{bmatrix} \qquad \text{otrzymaliśmy bazę} \qquad u_1 = \frac{1}{\sqrt{2}}\begin{bmatrix}1\\1\end{bmatrix}, \qquad u_2 = \frac{1}{\sqrt{2}}\begin{bmatrix}1\\-1\end{bmatrix}.$$

Oznaczmy teraz przez G macierz, której kolumnami są wektory v_1 i v_2, tj. $G = [v_1,\ v_2]$, a przez U macierz ortogonalną, której kolumnami są wektory u_1 i u_2, tj. $U = [u_1,\ u_2]$. Związki (7.9) i (7.10) są równoważne macierzowej równości

$$\begin{bmatrix}\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}}\\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}}\end{bmatrix} = \begin{bmatrix}1 & 3\\ 1 & -4\end{bmatrix}\begin{bmatrix}\frac{1}{\sqrt{2}} & \frac{1}{7\sqrt{2}}\\ 0 & \frac{\sqrt{2}}{7}\end{bmatrix}.$$

Oznaczając przez T górnotrójkątną macierz po prawej stronie powyższej równości możemy ją zapisać symbolicznie jako $U = GT$ i po pomnożeniu z prawej strony obu stron równości przez T^{-1} otrzymamy rozkład G na iloczyn macierzy unitarnej i górnotrójkątnej, $G = UT^{-1}$, to jest

$$\begin{bmatrix}1 & 3\\ 1 & -4\end{bmatrix} = \begin{bmatrix}\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}}\\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}}\end{bmatrix}\begin{bmatrix}\sqrt{2} & -\frac{1}{\sqrt{2}}\\ 0 & \frac{7}{\sqrt{2}}\end{bmatrix}.$$

7.2 Macierze ortogonalne

Z pojęciem macierzy ortogonalnej spotkaliśmy się przelotnie w jednym z poprzednich wykładów. Podamy teraz tę definicję w sposób formalny.

Definicja 7.1 (Macierz ortogonalna) Macierz kwadratowa stopnia n, $U \in \mathbf{M}_n(\mathbb{R})$, nazywa się *macierzą ortogonalną*, gdy spełnia warunek

$$U^t U = \mathbf{I}. \tag{7.11}$$

Jak zwykle, $\mathbf{I}$ oznacza macierz jednostkową.

Przypomnijmy (por. Wykład 4), że warunek (7.11) jest zapisem w notacji macierzowej warunku ortonormalności dla układu wektorów w $\mathbb{R}^n$ złożonego z kolumn macierzy U.

Wniosek 9 *Macierz $U \in \mathbf{M}_n(\mathbb{R})$ macierzą ortogonalną wtedy i tylko wtedy, gdy jej kolumny (wiersze) tworzą bazę ortonormalną przestrzeni $\mathbb{R}^n$.*

Stwierdzenie 13
a) *Jeśli $U \in \mathbf{M}_n(\mathbb{R})$ jest ortogonalna, to U i U^t są odwracalne i*

$$U^{-1} = U^t. \qquad \textit{Ponadto} \qquad \left(U^t\right)^{-1} = \left(U^{-1}\right)^t = U.$$

W szczególności wówczas macierze U, $U^t = U^{-1}$ są obie ortogonalne i odwracalne. Ponadto $\det U = \pm 1$.
b) *Jeśli $U, T \in \mathbf{M}_n(\mathbb{R})$ są ortogonalne, to także UT i TU sa ortogonalne oraz*

$$(UT)^t = T^t U^t.$$

c) *Macierz jednostkowa $\mathbf{I}$ stopnia n jest macierzą ortogonalną.*

Zbiór macierzy ortogonalnych stopnia n będziemy w dalszym ciagu oznaczać symbolem $\mathbf{O}(n)$ i nazywać *grupą macierzy ortogonalnych* stopnia n.

Przykład 7.2.1 **a)** Wektory u_1, $u_2 \in \mathbb{R}^2$ dane przez

$$u_1 = \begin{bmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{bmatrix}, \qquad u_2 = \begin{bmatrix} -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{bmatrix},$$

tworzą układ ortonormalny w $\mathbb{R}^2$. Biorąc po uwagę, że $\cos(\pi/4) = \sin(\pi/4) = \frac{1}{\sqrt{2}}$ możemy zapisać macierz odpowiadającą temu układowi w jednej z dwóch postaci:

$$U = \begin{bmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix} = \begin{bmatrix} \cos\frac{\pi}{4} & -\sin\frac{\pi}{4} \\ \sin\frac{\pi}{4} & \cos\frac{\pi}{4} \end{bmatrix}.$$

Po porównaniu tych wzorów z Przykładem 4.2.1 czytelnik zapewne zauważy, że odwzorowanie

$$\mathbb{R}^2 \ni x \mapsto y = \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} \cos\frac{\pi}{4} & -\sin\frac{\pi}{4} \\ \sin\frac{\pi}{4} & \cos\frac{\pi}{4} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \in \mathbb{R}^2$$

przedstawia obrót w płaszczyźnie $\mathbb{R}^2$ o kąt $\pi/4$.

b) Ogólniej, dla dowolnego $\varphi \in \mathbb{R}$ układ złożony z wektorów

$$u_1 = \begin{bmatrix} \cos\varphi \\ \sin\varphi \\ 0 \end{bmatrix}, \qquad u_2 = \begin{bmatrix} -\sin\varphi \\ \cos\varphi \\ 0 \end{bmatrix}, \qquad u_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

jest układem ortonormalnym w przestrzeni $\mathbb{R}^3$. Macierz U, której kolumnami są podane wektory

$$U = \begin{bmatrix} \cos\varphi & -\sin\varphi & 0 \\ \sin\varphi & \cos\varphi & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

jest macierzą ortogonalną. Łatwo sprawdzić, że przekształcenie

$$\mathbb{R}^3 \ni x \mapsto y = \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} = \begin{bmatrix} \cos\varphi & -\sin\varphi & 0 \\ \sin\varphi & \cos\varphi & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \in \mathbb{R}^3$$

jest obrotem w płaszczyźnie Ox_1x_2 tj. płaszczyżnie prostopadłej do osi Ox_3. Można też sprawdzić, że macierze

$$R_1(\varphi) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\varphi & -\sin\varphi \\ 0 & \sin\varphi & \cos\varphi \end{bmatrix}, \qquad R_2(\varphi) = \begin{bmatrix} \cos\varphi & 0 & \sin\varphi \\ 0 & 1 & 0 \\ -\sin\varphi & 0 & \cos\varphi \end{bmatrix}, \qquad R_3(\varphi) = \begin{bmatrix} \cos\varphi & -\sin\varphi & 0 \\ \sin\varphi & \cos\varphi & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

przedstawiają obroty wokół osi układu (odpowiednio Ox_1, Ox_2, Ox_3) o kąt φ.

Stwierdzenie 14 *Każdą macierz ortogonalną stopnia 2, $U \in \mathbf{O}(2)$, można jednoznacznie zapisać w jednej z dwóch postaci:*

$$\text{gdy } \det U = 1, \text{ to wtedy} \qquad U = \begin{bmatrix} \cos\varphi & -\sin\varphi \\ \sin\varphi & \cos\varphi \end{bmatrix}, \qquad \text{gdzie } \varphi \in [0,\, 2\pi[, \tag{7.12}$$

$$\text{gdy } \det U = -1, \text{ to wtedy} \qquad U = \begin{bmatrix} \cos\varphi & \sin\varphi \\ \sin\varphi & -\cos\varphi \end{bmatrix}, \qquad \text{gdzie } \varphi \in [0,\, 2\pi[. \tag{7.13}$$

Dodajmy, że dla macierzy U przedstawionej w (7.13) odwzorowanie $x \mapsto Ux$ jest odbiciem w prostej przecinającej oś Ox_1 pod kątem $\varphi/2$.

7.2.1 Transponowanie macierzy i skalarne mnożenie wektorów

Stwierdzenie 15 *Jeśli $A \in \mathbf{M}_n(\mathbb{R})$ i $x,\, y \in \mathbb{R}^n$, to mamy*

$$x \cdot (Ay) = (A^t x) \cdot y, \qquad \text{w zapisie macierzowym} \qquad x^t A y = (A^t x)^t y, \tag{7.14}$$

gdzie oczywiście traktujemy $x,\, y \in \mathbb{R}^n$ jako wektory kolumnowe.

Aby odsłonić to, co ukrywa skondensowany, używający minimalnej liczby symboli macierzowy zapis wzoru (7.14), zapiszemy teraz ten wzór za pomocą współrzędnych. Ograniczymy się do przypadku $n = 2$, bo już ten przypadek wystarczy, aby uwidocznić korzyści płynące z użycia macierzowej symboliki. Jeśli $A = \left[\begin{smallmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{smallmatrix}\right]$, to lewą stronę wzoru (7.14) można przekształcić do postaci

$$\begin{bmatrix} x_1 & x_2 \end{bmatrix} \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} x_1 & x_2 \end{bmatrix} \begin{bmatrix} a_{11}y_1 + a_{12}y_2 \\ a_{21}y_1 + a_{22}y_2 \end{bmatrix} = x_1(a_{11}y_1 + a_{12}y_2) + x_2(a_{21}y_1 + a_{22}y_2)$$

Przekształcając w analogiczny sposób prawą stronę tego wzoru dochodzimy do

$$\left(\begin{bmatrix} a_{11} & a_{21} \\ a_{12} & a_{22} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \right)^t \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} a_{11}x_1 + a_{21}x_2 \\ a_{12}x_1 + a_{22}x_2 \end{bmatrix}^t \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = (a_{11}x_1 + a_{21}x_2)y_1 + (a_{12}x_1 + a_{22}x_2)y_2.$$

Otrzymane wyrażenia są równe,

$$x_1(a_{11}y_1 + a_{12}y_2) + x_2(a_{21}y_1 + a_{22}y_2) = (a_{11}x_1 + a_{21}x_2)y_1 + (a_{12}x_1 + a_{22}x_2)y_2,$$

gdyż po otwarciu nawiasów i wykonaniu mnożeń, każde z nich sprowadza się do

$$a_{11}x_1y_1 + a_{12}x_1y_2 + a_{21}x_2y_1 + a_{22}x_2y_2.$$

7.2.2 Dalsze własności macierzy ortogonalnych

Stwierdzenie 16 *Dla macierzy $U \in \mathbf{M}_n(\mathbb{R})$ następujące warunki są równoważne między sobą:*
a) *Macierz U jest macierzą ortogonalną;*
b) *dla każdej pary wektorów $x,\, y \in \mathbb{R}^n$ zachodzi*

$$(Ux) \cdot (Uy) = x \cdot y; \tag{7.15}$$

c) *dla każdego wektora $x \in \mathbb{R}^n$ zachodzi*

$$|Ux| = |x|; \tag{7.16}$$

d) *dla każdej bazy ortonormalnej $\{v_1,\, \ldots,\, v_n\}$ w przestrzeni $\mathbb{R}^n$ układ $\{Uv_1,\, \ldots,\, Uv_n\}$ jest bazą ortonormalną w przestrzeni $\mathbb{R}^n$.*

A zatem macierz $U \in \mathbf{M}_n(\mathbb{R})$, która spełnia jeden (a w konsekwencji wszystkie) z warunków Stwierdzenia 16 jest macierzą ortogonalną. Z punktu **c)** wynika z kolei, że odwzorowania liniowe zadane przez ortogonalną macierz zachowują odległość między punktami przestrzeni.

Wniosek 10 *Odwzorowanie liniowe $f_U : \mathbb{R}^n \ni x \to f_U(x) = Ux \in \mathbb{R}^n$ zadane przez macierz ortogonalną U ma następującą własność: dla każdej pary $x,\, y \in \mathbb{R}^n$ odległość ich obrazów $d(f_U(x),\, f_U(y))$ jest taka sama, jak odległość $d(x,\, y)$;*

$$d(x,\, y) = d(f_U(x),\, f_U(y)), \qquad \textit{a to jest tym samym, co} \qquad |x - y| = |Ux - Uy|. \tag{7.17}$$

Ta ostatnia własność pozwala wyróżnić niezwykle ważną klasę odwzorowań o bogatych zastosowaniach geometrycznych.

Definicja 7.2 (Izometria przestrzeni $\mathbb{R}^n$) Odwzorowanie $f : \mathbb{R}^n \to \mathbb{R}^n$ nazywa się *izometrią* przestrzeni euklidesowej $\mathbb{R}^n$, gdy dla każdej pary $x,\, y \in \mathbb{R}^n$ zachodzi

$$|x - y| = |f(x) - f(y)|. \tag{7.18}$$

Ponieważ odległość euklidesowa między punktami $x,\, y \in \mathbb{R}^n$ jest dana wzorem $d(x,\, y) = |x - y|$, to równość (7.18) interpretujemy jako wyraz tego faktu, że pod działaniem izometrii f każde dwa punkty przechodzą na parę punktów o tej samej od siebie odległości (mówimy krótko, że odległość nie zmienia się pod działaniem izometrii). Takie odwzorowania stanowią podstawowe narzędzie w badaniu geometrii płaszczyzny Euklidesa (a także trójwymiarowej przestrzeni fizykalnej), bo to one służą do określenia *przystawania figur geometrycznych.* Tu będziemy mogli jedynie dotknąć najbardziej podstawowych ich własności.

Wniosek 10 wyjaśnia, że odwzorowania liniowe z ortogonalną macierzą są izometriami — odnosząc się do tych szczególnych izometrii będziemy je nazywać *liniowymi izometriami* przestrzeni euklidesowej $\mathbb{R}^n$. Podkreślanie liniowości izometrii jest uzasadnione tym, że nie jest ona konieczną własnością izometrii — obszerną klasą izometrii są przesunięcia $x \mapsto \tau_v(x) = x - v$, które nie są odwzorowaniami liniowymi. Okazuje się jednak, że z tych dwóch specjalnych klas izometrii (przesunięć i izometrii liniowych) można otrzymać dowolną izometrię przestrzeni $\mathbb{R}^n$.

Wykład 8 (18 IV 2012)

Izometrie i ortogonalne zmiany współrzędnych

8.1 Klasyfikacja izometrii

8.1.1 Odwzorowania afiniczne

Dalszą analizę struktury izometrii poprzedzimy opisem obszerniejszej klasy odwzorowań.

Definicja 8.1 (Odwzorowania afiniczne przestrzeni $\mathbb{R}^n$) Niech $b \in \mathbb{R}^n$ będzie ustalonym wektorem i niech $A = [a_{ij}]$ będzie macierzą kwadratową stopnia n o wyrazach z $\mathbb{R}$, $A \in \mathbf{M}_n(\mathbb{R})$. Odwzorowanie

$$\mathbb{R}^n \ni x \mapsto f(x) = Ax + b \in \mathbb{R}^n \tag{8.1}$$

nazywamy *odwzorowaniem afinicznym.*

Nie jest trudno wykazać, że we wzorze (8.1) zarówno macierz A jak i wektor b są wyznaczone jednoznacznie. Inaczej mówiąc, jeśli odwzorowanie h jest określone analogicznym wzorem $h(x) = Bx+c$, gdzie $c \in \mathbb{R}^n$ i $B = [b_{ij}] \in \mathbf{M}_n(\mathbb{R})$, to $f(x) = h(x)$ dla wszystkich $x \in \mathbb{R}^n$ wtedy i tylko wtedy, gdy $A = B$ i $b = c$. A zatem każdemu odwzorowaniu afinicznemu odpowiada jednoznacznie wyznaczone odwzorowanie liniowe $\mathbb{R}^n \ni x \mapsto Ax \in \mathbb{R}^n$ o macierzy A, które będziemy nazywali *częścią liniową odwzorowania* f, i wektor przesunięcia $b = f(0)$.

Znaczenie odwzorowań afinicznych dla problemów geometrii wynika z faktu, że przeprowadzają one rozmaitości afiniczne na rozmaitości afiniczne, a w szczególności obrazem prostej w odwzorowaniu afinicznym jest punkt lub prosta. Wykazanie tego faktu pozostawiamy czytelnikowi.

Szczególne odwzorowania afiniczne:

Odwzorowanie liniowe

$$\mathbb{R}^n \ni x \mapsto Ax \in \mathbb{R}^n$$

jest odwzorowaniem afinicznym z zerowym wektorem przesunięcia.

Przesunięcie o wektor $b \in \mathbb{R}^n$, które jest dane wzorem

$$\mathbb{R}^n \ni x \mapsto x + b \in \mathbb{R}^n \tag{8.2}$$

jest także odwzorowaniem afinicznym. Częścią liniową tego odwzorowania jest, oczywiście, odwzorowanie identycznościowe $\mathbb{R}^n \ni x \mapsto x = \mathbf{I}x \in \mathbb{R}^n$.

Odwzorowanie afiniczne o zerowym wektorze przesunięcia i diagonalnej macierzy A,

$$A = \begin{bmatrix} k_1 & 0 & \dots & 0 \\ 0 & k_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & k_n \end{bmatrix}$$

nazywa się **skalowaniem**. W takim przypadku liczby k_1, k_2, $\dots$, k_n nazywają się *współczynnikami skali* w kierunkach osi Ox_1, odpowiednio Ox_2 itd. Jeśli wszystkie współczynniki skali są jednakowe, to mamy do czynienia z *podobieństwem*.

Twierdzenie 9 *Niech f_1, f_2 będą dwoma odwzorowaniami afinicznymi, A_1, $A_2 \in \mathbf{M}_n(\mathbb{R})$ odpowiednio ich częściami liniowymi i b_1, b_2 wektorami przesunięcia. Złożenie $f_1 \circ f_2$ jest odwzorowaniem afinicznym danym wzorem*

$$f_1 \circ f_2(x) = (A_1 \cdot A_2)x + A_1 b_2 + b_1. \tag{8.3}$$

Odwzorowanie afiniczne $x \mapsto f(x) = Ax + b$ jest bijekcją wtedy i tylko wtedy, gdy A jest macierzą odwracalną, tj. $\det A \neq 0$. *W takim przypadku odwzorowanie odwrotne jest też afiniczne i jest dane wzorem*

$$f^{-1}(x) = A^{-1}x - A^{-1}b.$$

8.1.2 Analityczna postać izometrii przestrzeni $\mathbb{R}^n$

Twierdzenie 10 *Odwzorowanie $f : \mathbb{R}^n \to \mathbb{R}^n$ jest izometrią wtedy i tylko wtedy, gdy jest odwzorowaniem afinicznym, którego część liniowa jest macierzą ortogonalną. Inaczej mówiąc f jest dane wzorem*

$$f(x) = Ux + b. \tag{8.4}$$

Ponadto każda izometria jest odwzorowaniem bijektywnym i jej odwzorowanie odwrotne jest izometrią postaci

$$f^{-1}(x) = U^t x - U^t b.$$

Przykład 8.1.1 Wyznaczyć odwzorowanie afiniczne postaci $f(x) = Ax + b$, które przeprowadza układ punktów $\{p_0, p_1, p_2, p_3\} \in \mathbb{R}^3$ w podanej kolejności na układ $\{q_0, q_1, q_2, q_3\} \in \mathbb{R}^3$, gdzie

$$p_0 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \quad p_1 = \begin{bmatrix} 2 \\ 1 \\ 1 \end{bmatrix}, \quad p_2 = \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix}, \quad p_3 = \begin{bmatrix} 1 \\ 1 \\ 2 \end{bmatrix} : \qquad q_0 = \begin{bmatrix} 1 \\ 3 \\ 1 \end{bmatrix}, \quad q_1 = \frac{1}{2}\begin{bmatrix} 3 \\ 6 + \sqrt{3} \\ 2 \end{bmatrix}, \quad q_2 = \frac{1}{2}\begin{bmatrix} 2 - \sqrt{3} \\ 7 \\ 2 \end{bmatrix}, \quad q_3 = \begin{bmatrix} 1 \\ 3 \\ 2 \end{bmatrix}.$$

Sprawdzić, że f jest izometrią.

Wsk. Punkty p_i dla $i = 1, 2, 3$ oznaczają końce wersorów osi e_i zaczepionych w punkcie p_0, gdyż mamy $p_i - p_0 = e_i$ dla $i = 1, 2, 3$.

Wykład 9 (25 IV 2012)

Układy współrzędnych; funkcje wielomianowe

9.1 Zamiany współrzędnych w przestrzeni i na płaszczyźnie

Przedstawimy teraz dla przypadku $n = 3$ (lub $n = 2$) nieco inne spojrzenie na wzór (8.1). Załóżmy, że $A \in \mathbf{M}_3(\mathbb{R})$ jest macierzą odwracalną i niech układ liczb rzeczywistych y_1, y_2, y_3 będzie określony wzorem

$$\begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} + \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix}. \tag{9.1}$$

Wzorowi temu nadajemy następującą interpretację. Początkowi układu współrzędnych uznajdującemu się w punkcie $0 = (0, 0, 0) \in \mathbb{R}^3$ odpowiada punkt (b_1, b_2, b_3), którego współrzędnymi są współrzędne wektora przesunięcia b. Punktom wyznaczonym przez wersory osi współrzędnych e_1, e_2, e_3 odpowiadać będą punkty

$$p_1 = \begin{bmatrix} y_1^1 \\ y_2^1 \\ y_3^1 \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} + \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix} = \begin{bmatrix} a_{11} \\ a_{21} \\ a_{31} \end{bmatrix} + \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix}, \tag{9.2}$$

$$p_2 = \begin{bmatrix} y_1^2 \\ y_2^2 \\ y_3^2 \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} + \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix} = \begin{bmatrix} a_{12} \\ a_{22} \\ a_{32} \end{bmatrix} + \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix}, \tag{9.3}$$

$$p_3 = \begin{bmatrix} y_1^3 \\ y_2^3 \\ y_3^3 \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} + \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix} = \begin{bmatrix} a_{13} \\ a_{23} \\ a_{33} \end{bmatrix} + \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix}. \tag{9.4}$$

Układy liczb $\begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix}$ i $\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$ opisują ten sam punkt przestrzeni przez odniesienie do dwóch różnych układów współrzędnych — przy czym środek układu współrzędnych oznaczanych przez y_i przypada w punkcie $b = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix}$. Potocznie mówiąc, ten sam punkt jest tutaj opisany z „dwóch różnych punktów widzenia".

Zamianę współrzędnych opisaną wzorem (9.1) nazywa się *afiniczną zamianą współrzędnych.* Jeśli część liniowa odwzorowania danego tym wzorem jest macierzą ortogonalną — ten przypadek jest dla nas najważniejszy — to mowimy o *ortogonalnej zamianie współrzędnych.*

9.2 Algebra funkcji na przestrzeni $\mathbb{R}^n$

9.2.1 Działania na funkcjach

Przez funkcję będziemy rozumieć odwzorowanie o wartościach w ciele liczbowym, którym może być ciało liczb rzeczywistych $\mathbb{R}$ lub ciało liczb zespolonych $\mathbb{C}$. Interesować nas będą funkcje, których dziedziną jest przestrzeń wektorowa V — najczęściej będziemy przyjmować, że $V = \mathbb{R}^n$ (jeśli funkcje przyjmują wartości w $\mathbb{C}$, to naturalnie będzie brać $V = \mathbb{C}^n$), ale większość konstrukcji można wykonać dla dowolnej przestrzeni skończonego wymiaru.

Zbiór funkcji określonych na przestrzeni V i przyjmujących wartości w wybranym ciele liczbowym $\mathbb{K}$ ($\mathbb{K} = \mathbb{R}$ lub $\mathbb{C}$) oznaczać będziemy symbolem $\mathcal{F}(V, \mathbb{K})$. W zbiorze $\mathcal{F}(V, \mathbb{K})$ wprowadzamy tak zwane działania punktowe — suma dwóch funkcji określonych na zbiorze V i iloczyn takiej funkcji przez liczbę $a \in \mathbb{K}$ są określone „punkt po punkcie” zgodnie z następującą definicją.

Definicja 9.1 (Działania na funkcjach) Dla dowolnych $f,\, g \in \mathcal{F}(V, \mathbb{K})$ oraz $a \in \mathbb{K}$ określimy następującymi wzorami:

$$(f+g)(x) = f(x) + g(x), \qquad \textbf{sumę funkcji} \tag{9.5}$$

$$(af)(x) = af(x), \qquad \textbf{iloczyn funkcji przez liczbę} \tag{9.6}$$

$$(f \cdot g)(x) = f(x) \cdot g(x), \qquad \textbf{iloczyn dwóch funkcji} \tag{9.7}$$

dla dowolnego $\quad x \in V$.

Przy tych określeniach mamy $f,\, g \in \mathcal{F}(V, \mathbb{K}),\ a \in \mathbb{K} \implies f+g,\ af,\ f \cdot g \in \mathcal{F}(V, \mathbb{K})$, a ponadto działania dodawania funkcji i mnożenia przez liczby mają wszystkie własności wymagane w definicji przestrzeni wektorowej. Dzięki temu możemy, rozszerzając zakres definicji przestrzeni liniowej, traktować zbiór $\mathcal{F}(V, \mathbb{K})$ z działaniami określonymi wzorami 9.5–9.6 jako przestrzeń liniową (wektorową) nad ciałem $\mathbb{K}$.

Działenie mnożenia funkcji określone wzorem (9.7) pozwala na wyróżnienie szczególnej klasy funkcji, zwanych funkcjami wielomianowymi, lub prościej, wielomianami. W dalszym ciągu będziemy, jak to się zwykle robi, opuszczać kropkę „$\cdot$” przy oznaczeniu iloczynu funkcji pisząc fg zamiast $f \cdot g$.

Definicja 9.2 (Funkcje wielomianowe na przestrzeni $\mathbb{R}^n$) Niech $i_1, \ldots, i_n$ będzie ciągiem liczb całkowitych nieujemnych. Funkcję określoną wzorem postaci

$$\mathbb{R}^n \ni x = (x_1, \ldots, x_n) \longmapsto x_1^{i_1} x_2^{i_2} \ldots x_n^{i_n} \in \mathbb{R} \tag{9.8}$$

będziemy nazywać *jednomianem*. Stopniem tego jednomianu nazwiemy liczbę $i_1 + i_2 + \ldots + i_n$. Wielomianami na przestrzeni $\mathbb{R}^n$ będziemy nazywać kombinacje liniowe jednomianów, czyli funkcje postaci

$$\mathbb{R}^n \ni x \longmapsto \sum a_{i_1\, i_2 \ldots i_n} x_1^{i_1} x_2^{i_2} \ldots x_n^{i_n}, \qquad \text{gdzie } a_{i_1\, i_2 \ldots i_n} \in \mathbb{R}.$$

Jeśli wszystkie jednomiany wchodzące w skład tej kombinacji liniowej mają jednakowy stopień, powiedzmy d, to mówimy, że jest to *wielomian jednorodny stopnia d*. W ogólności stopniem wielomianu będziemy nazywać największy ze stopni jednomianów wchodzących w skład określającej go kombinacji liniowej.

Przykład 9.2.1 **a)** Wielomianami pierwszego stopnia są funkcje

$$\mathbb{R}^n \ni x \longmapsto \sum_{i=1}^{n} a_i x_i + b$$

b) Wielomianami jednorodnymi drugiego stopnia są funkcje postaci

$$\mathbb{R}^n \ni x \longmapsto \sum_{j=1}^{n} \sum_{i=1}^{n} a_{ij} x_i x_j .$$

Wielomiany jednorodne stopnia 2 nazywamy też formami kwadratowymi.
c) Dla przestrzeni $\mathbb{R}^2$ formę kwadratową zapisujemy w postaci

$$q(x_1, x_2) = ax_1^2 + bx_1x_2 + cx_2^2, \qquad \text{gdzie} \quad a, b, c \in \mathbb{R},$$

a dla przestrzeni $\mathbb{R}^3$ w postaci

$$q(x_1, x_2, x_3) = a_{11}x_1^2 + a_{12}x_1x_2 + a_{13}x_1x_3 + a_{22}x_2^2 + a_{23}x_2x_3 + a_{33}x_3^2, \qquad \text{gdzie} \quad a_{ij} \in \mathbb{R}.$$

Stwierdzenie 17 *Każda formę kwadratową na przestrzeni $\mathbb{R}^n$ można zapisać w postaci*

$$q(x) = x^t A x = \sum_{j,m=1}^{n} a_{jm} x_j x_m$$

przy czym macierz $A \in \mathbf{M}_n(\mathbb{R})$ jest wyznaczona jednoznacznie, pod warunkiem symetrii.

9.2.2 Wyznacznikowe kryterium dodatniej określoności macierzy

Niech $A \in \mathbf{M}_n(\mathbb{R})$ będzie macierzą kwadratową stopnia n. Dla $k = 1, \ldots, n$ przez $D_k(A)$ oznaczmy *minor główny* stopnia k macierzy A, tj. minor złożony z elementów macierzy A występujących w pierwszych k kolumnach i pierwszych k wierszach. Ogólnie $D_k(A) = \det[a_{j\,l}]_{j\leqslant k, l\leqslant k}$, a w szczególności dla $k = 1, 2, 3$ mamy

$$\begin{aligned}
D_1(A) &= a_{11}, \\
D_2(A) &= \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = a_{11}a_{22} - a_{12}a_{21}, \\
D_3(A) &= \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = \\
&\quad + a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} - a_{13}a_{22}a_{31} - a_{12}a_{21}a_{33} - a_{11}a_{32}a_{23}.
\end{aligned}$$

Stwierdzenie 18 *Niech $A \in \mathbf{M}_n(\mathbb{R})$ będzie macierzą symetryczną, tzn. $A^t = A$. Wówczas*

$$x^t A x = \sum_{j,m=1}^{n} a_{jm} x_j x_m > 0, \quad \text{dla wszystkich } x \neq 0, \quad \Longleftrightarrow \quad D_k(A) > 0 \quad \text{dla } k = 1, \ldots, n; \tag{9.9}$$

$$x^t A x = \sum_{j,m=1}^{n} a_{jm} x_j x_m < 0, \quad \text{dla wszystkich } x \neq 0, \quad \Longleftrightarrow \quad (-1)^k D_k(A) > 0 \quad \text{dla } k = 1, \ldots, n. \tag{9.10}$$

Definicja 9.3 (Macierze dodatnio (ujemnie) określone) Symetryczną macierz kwadratową $A \in \mathbf{M}_n(\mathbb{R})$ nazywamy macierzą:
dodatnio określoną gdy $x^t A x = \sum_{j,m=1}^{n} a_{jm} x_j x_m > 0$ dla każdego $x \neq 0$;
ujemnie określoną gdy $x^t A x = \sum_{j,m=1}^{n} a_{jm} x_j x_m < 0$ dla każdego $x \neq 0$.

Przykład 9.2.2 Rozważmy iloczyn $x^t Ax$ dla macierzy symetrycznej

$$A = \begin{bmatrix} 1 & 2 \\ 2 & 5 \end{bmatrix} \qquad \text{i dowolnego wektora} \quad x = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \in \mathbb{R}^2.$$

Prosty rachunek daje

$$x^t Ax = \begin{bmatrix} x_1 \, x_2 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 2 & 5 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} x_1 + 2x_2 \; 2x_1 + 5x_2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = x_1^2 + 4x_1x_2 + 5x_2^2 = (x_1 + 2x_2)^2 + x_2^2$$

To ostatnie wyrażenie jest dodatnie dla każdego różnego od zera wektora $x \in \mathbb{R}^2$, co dowodzi, że macierz A jest dodatnio określona. Z drugiej strony minory główne macierzy A spełniają

$$D_1(A) = 1 > 0, \qquad D_2(A) = \begin{vmatrix} 1 & 2 \\ 2 & 5 \end{vmatrix} = 1 > 0,$$

zgodnie z treścią Stwierdzenia 19.