

WYKŁAD 10

Iloraz skalarny

$$(\cdot, \cdot) : V \times V \rightarrow \mathbb{R}$$

$$(\vec{v}, \vec{v}) = \|\vec{v}\|^2$$

$$\cos \alpha = \frac{(\vec{u}, \vec{v})}{\|\vec{u}\| \cdot \|\vec{v}\|} \leftarrow \text{analog tw. cosinusów}$$

$$\vec{u} \perp \vec{v} \Leftrightarrow (\vec{u}, \vec{v}) = 0$$

- $(\vec{u}, \vec{v}) = (\vec{v}, \vec{u})$
- liniowy po obu mp.
- $(\vec{v}, \vec{v}) \geq 0$
- $(\vec{v}, \vec{v}) = 0 \Leftrightarrow \vec{v} = \vec{0}$

$$V \xrightarrow{[\cdot]_B} B \mathbb{R}^n$$

bazis ortogonalnej

$$v_1, \dots, v_n$$

- baza
- $v_i \perp v_j \quad i \neq j$
- $\|v_i\| = 1$

- $\vec{v} \in B - v_1, \dots, v_n$ ortonorm.

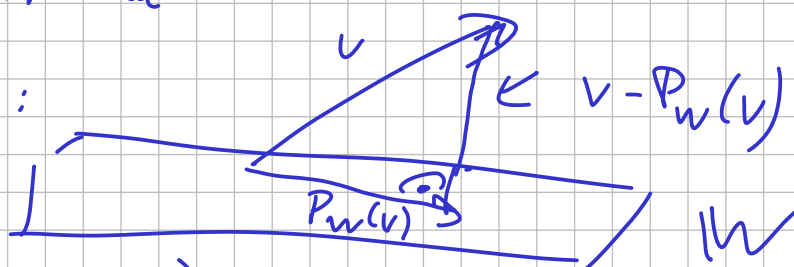
$$\vec{v} = \sum d_i v_i \quad d_i = (\vec{v}, v_i)$$

$$(\vec{v}, \vec{u}) = [\vec{v}]_B \cdot [\vec{u}]_B \quad B\text{-ortonorm.}$$

$\uparrow \mathbb{R} \quad \uparrow \mathbb{R} \quad \text{it. ik. 1k}$

- Istnieje baza ortogonalna

- Rzut prostopadły:



$$P_W(v) \perp v - P_W(v)$$

$\uparrow W$

$$W \subseteq V$$

$$v_1, \dots, v_k, v_{k+1}, \dots, v_n \quad \text{ortonorm.}$$

$$P_W \left(\sum_{i=1}^n \alpha_i v_i \right) = \sum_{i=1}^k \alpha_i \vec{v_i}$$

$\uparrow$
 $\langle v, v_i \rangle$

$$P_W(\vec{v}) = \sum_{i=1}^k \langle v, v_i \rangle \vec{v_i}$$

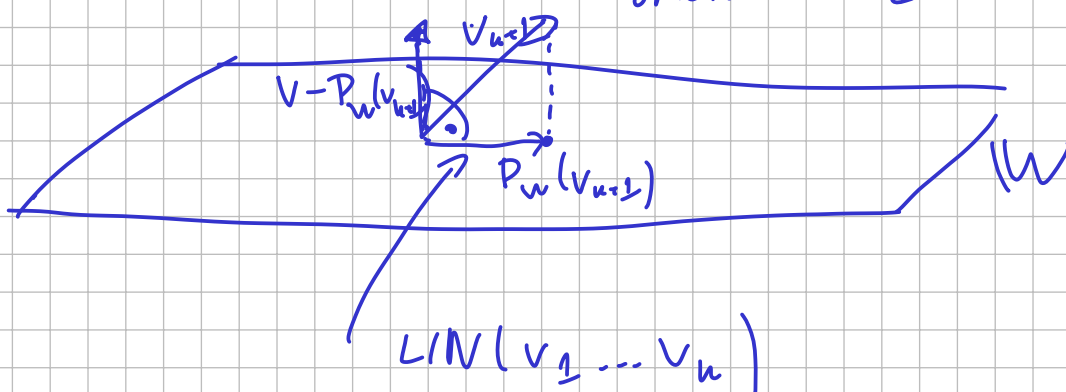
$v_1, \dots, v_k$
- basis orton. W

Konstruktion einer ortonormalen Basis

$$v_1, \dots, v_k, v_{k+1}, \dots, v_n$$

$$\underbrace{v_1, \dots, v_k}_{v_1, \dots, v_k}$$

$$\text{LIN}(v_1, \dots, v_k) = \text{LIN}(v_1', \dots, v_k') \quad \text{ortonormal}$$



$$v_1, \dots, v_n \quad \text{for } k=2 \text{ to } n$$

$$v_k \leftarrow v_k - P_{\text{LIN}(v_1, \dots, v_{k-1})}(v_k)$$

$$v_k \leftarrow v_k / \sqrt{\langle v_k, v_k \rangle} \quad \text{d.h. } \|v_k\| = 1$$

11.4 Algorytm Grama-Schmidta ortonormalizacji bazy

Algorytm 1 Algorytm Gram-Schmidta ortonormalizacji

- 1: $\vec{v}_1 \leftarrow \frac{\vec{v}_1}{\sqrt{\langle \vec{v}_1, \vec{v}_1 \rangle}}$ ▷ Normowanie
norm $\|\vec{v}_1\| = 1$
 - 2: **for** $i \leftarrow 2 \dots n$ **do**
 - 3: $\vec{v}_i \leftarrow \vec{v}_i - \sum_{j=1}^{i-1} \langle \vec{v}_i, \vec{v}_j \rangle \vec{v}_j$ ▷ Odjęcie rzutu na przestrzeń rozpiętą przez $\vec{v}_1, \dots, \vec{v}_{i-1}$
norm
 - 4: **if** $\langle \vec{v}_i, \vec{v}_i \rangle = 0$ **then**
 - 5: **return** Wektory są liniowo zależne.
 - 6: $\vec{v}_i \leftarrow \frac{\vec{v}_i}{\sqrt{\langle \vec{v}_i, \vec{v}_i \rangle}}$ ▷ Normowanie
 $\|\vec{v}_i\| = 1$
-

Uwaga. Ostatni krok, w którym ortonormalizujemy kolejne wektory, nie jest w zasadzie potrzebny (i możemy dostać bazę ortogonalną), jednak w takim przypadku musimy zmienić odpowiednio wyrażenie na rzut prostopadły.

Twierdzenie 11.21. *Jeśli układ na wejściu algorytmu Grama-Schmidta był niezależny, to uzyskane wektory są układem ortonormalnym.*

Jeśli układ $\vec{v}_1, \dots, \vec{v}_i$ był zależny i układ $\vec{v}_1, \dots, \vec{v}_{i-1}$ był niezależny, to w czasie algorytmu przekształcimy $\vec{v}_i$ na $\vec{0}$.

$$\begin{matrix} v_1, \dots, v_n \\ v_1', \dots, v_n' \end{matrix} \quad \text{wejście}$$

$$\begin{aligned} & \bullet v_1', \dots, v_i' \quad \text{układ ortonormalny} \\ & \bullet \text{LIN}(v_1, \dots, v_{i-1}, v_i) = \text{LIN}(v_1', \dots, v_i') \end{aligned}$$

$$\begin{aligned} & \bullet v_i \in \text{LIN}(v_1, \dots, v_{i-1}) \quad P_{W_{i-1}}(v_i) = v_i \\ & \quad v_i - P_{W_{i-1}}(v_i) = \vec{0} \end{aligned}$$

$$\begin{aligned} \text{LIN}(v_1, \dots, v_{i-1}, v_i) &= \text{LIN}(v_1', \dots, v_{i-1}', v_i) \\ v_i &= v_i - \sum_{j=1}^{i-1} \langle v_i, v_j' \rangle v_j' \end{aligned}$$

Przykład 11.22. Dla standardowego iloczynu skalarnego w $\mathbb{R}^4$ zortonormalizujemy układ wektorów

$$\{(4, 4, -2, 0); (1, 4, 1, 0); (5, -4, -7, 1)\}$$

i uzupełnimy go do bazy ortonormalnej.

Oznaczmy zadane wektory jako $\vec{v}_1, \vec{v}_2, \vec{v}_3$. Dokonamy ortonormalizacji bazy metodą Grama-Schmidta; niech $\vec{v}'_1, \vec{v}'_2, \vec{v}'_3$ to wektory po tym procesie.

$$v_1 = 4, 4, -2, 0$$

$$|\langle v_1, v_1 \rangle| = \sqrt{16 + 16 + 4} = \sqrt{36} = 6$$

$$v'_1 = \frac{1}{6} \langle 4, 4, -2, 0 \rangle = \frac{1}{3} \langle 2, 2, -1, 0 \rangle$$

$$v_2 = (1, 4, 1, 0)$$

$$(1, 4, 1, 0) - \langle v'_1, v_2 \rangle \cdot v'_1$$

$$\frac{1}{3} \langle 2, 2, -1, 0 \rangle \cdot (1, 4, 1, 0)$$

$$= 3$$

$$(1, 4, 1, 0) - 3 \cdot \frac{1}{3} \cdot (2, 2, -1, 0)$$

$$= (-1, 2, 2, 0)$$

$$|\langle -1, 2, 2, 0 \rangle| = \sqrt{1 + 4 + 4} = \sqrt{9} = 3$$

$$v'_2 = \frac{1}{3} (-1, 2, 2, 0)$$

$$v'_1 = \frac{1}{3} (2, 2, -1, 0)$$

$$v_3 = (5, -4, -7, 1)$$

$$v_3 - \langle v'_1, v_3 \rangle v'_1 - \langle v'_2, v_3 \rangle v'_2$$

$$\frac{1}{3} \cdot (5, -4, -7, 1) \cdot (2, 2, -1, 0) = \frac{1}{3} (10 - 8 - 7) = 3$$

$$\frac{1}{3} (5, -4, -7, 1) \cdot (-1, 2, 2, 0) = \frac{1}{3} (-5 - 8 - 14) = -9$$

$$(5, -4, -7, 1) - (2, 2, -1, 0) + 3(-1, 2, 2, 0) = (0, 0, 0, 1)$$

$$v'_3 = (0, 0, 0, 1)$$

Lemat 11.23. Jeśli baza B powstaje z bazy A przez ortonormalizację Grama-Schmidta, to M_{BA} i M_{AB} są macierzami górnotrójkątnymi. Ponadto mają one dodatnie elementy na przekątnej.

Prosty dowód pozostawiamy jako ćwiczenie.

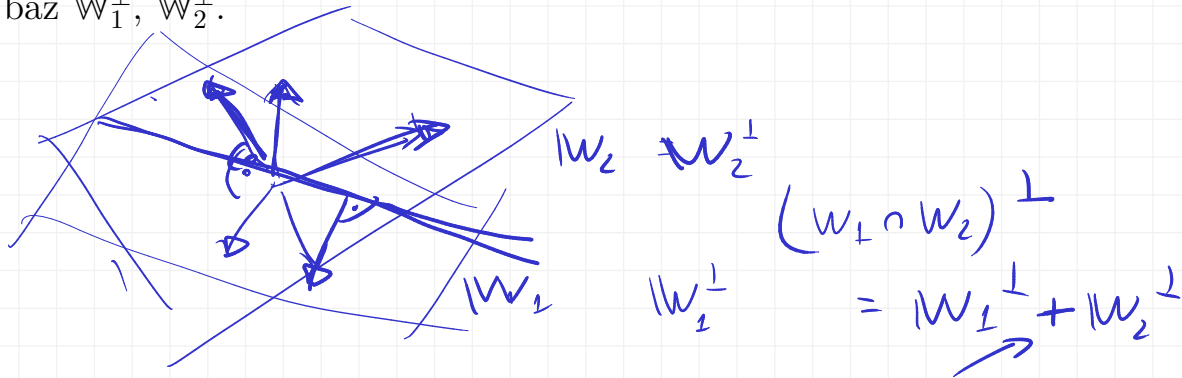
$$M_{AB} = \begin{bmatrix} \times & \times & \times \\ 0 & \times & \times \\ 0 & 0 & \times \end{bmatrix}$$

$$M_{BA} = \begin{bmatrix} \times & 0 & 0 \\ \times & \times & 0 \\ \times & \times & \times \end{bmatrix}$$

11.5 Zastosowania: geometria

11.5.1 Reprezentacja przez dopełnienie ortogonalne

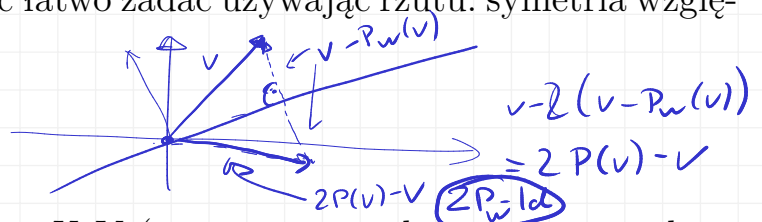
Dopełnienie ortogonalne jest dobrym sposobem reprezentacji płaszczyzn/prostych itp.: dla danej płaszczyzny W reprezentujemy ją jako bazę $W^\perp$. Reprezentacja ta jest o tyle dobra, że można łatwo przecinać tak zadane przestrzenie: dla W_1, W_2 ich przecięcie to $(W_1^\perp + W_2^\perp)^\perp$. Zwartą reprezentację otrzymujemy przez ortonormalizację sumy baz $W_1^\perp, W_2^\perp$.



Płaszczyzny reprezentowane są w ten sposób jako jeden wektor, często w CADach linie trzyma się jako przecięcie płaszczyzn.

11.5.2 Symetrie

Macierz symetrii względem prostej dość łatwo zadać używając rzutu: symetria względem W wyraża się jako $2P_W - \text{Id}$.



11.5.3 Regresja liniowa

Założmy, że mamy dwie zmienne losowe X, Y (np.: cena mieszkania i powierzchnia mieszkania). Dla uproszczenia zakładamy, że zmienne są unormowane:

wartość oczekiwana mają wartość oczekiwaną równą 0:

$$X \rightarrow X - E[X]$$

$$E[X] = E[Y] = 0.$$

wariancja mają wariancję równą 1

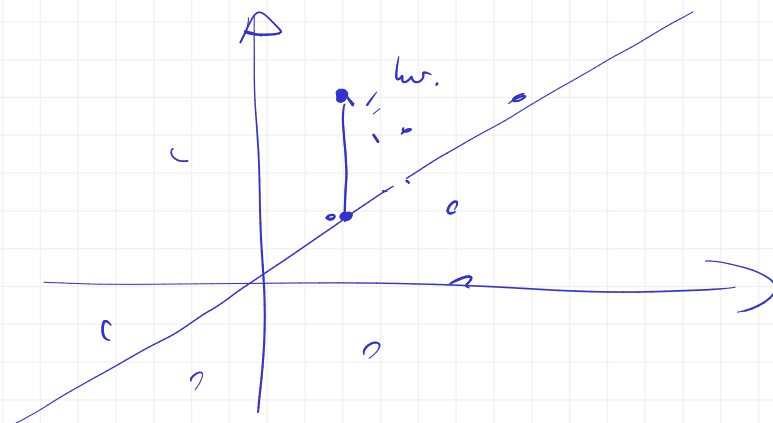
$$\text{Var}[X] = \text{Var}[Y] = 1.$$

$$E[X - E[X]] = E[X^2]$$

Chcemy policzyć „najlepsze” β , takie że $X = \beta Y$. Co to znaczy najlepsze? W tym konkretnym problemie chcemy zminimalizować kwadrat błędu, czyli

$$E[(X - \beta Y)^2] = \text{Var}[X - \beta Y]$$

(bo obie zmienne mają wartość oczekiwaną 0).



$$E[X] = E[Y] = 0$$

kowariancja: iloczyn składowy

$$E[X \cdot Y]$$

$$\begin{array}{c|c|c} 2 & 3 & 2 \cdot 3 \cdot \frac{1}{2} \\ 7 & 8 & 7 \cdot 8 \cdot \frac{1}{2} \\ \vdots & \vdots & \vdots \end{array}$$

$$X, Y$$

$$X \sim \beta \cdot Y$$

$$X = Y_{\perp} + \alpha \cdot Y \quad \text{ort. part.}$$

$$E[X \cdot Y]$$

$$\sum$$

$$X \rightarrow P_Y(X)$$

$$E[(X - \beta Y)^2] = \langle X - \beta Y, X - \beta Y \rangle$$

$$= \langle Y_{\perp} + (\alpha - \beta)Y, Y_{\perp} + (\alpha - \beta)Y \rangle$$

$$= \langle Y_{\perp}, Y_{\perp} \rangle + \underbrace{\langle (\alpha - \beta)Y, Y_{\perp} \rangle}_0 + \underbrace{\langle Y_{\perp}, (\alpha - \beta)Y \rangle}_0$$

↑
stałe

$$+ \langle (\alpha - \beta)Y, (\alpha - \beta)Y \rangle$$

$$(\alpha - \beta)^2 \langle Y, Y \rangle$$

$$\beta < \alpha$$

Rozdział 12

Izometrie, macierze ortogonalne

12.1 Izometrie

Definicja 12.1 (Izometria). Przekształcenie liniowe $F : \mathbb{V} \rightarrow \mathbb{V}$ na przestrzeni liniowej $\mathbb{V}$ z iloczynem skalarnym $\langle \cdot, \cdot \rangle$, nazywamy *izometrią*, jeśli zachowuje iloczyn skalarny, tj. dla każdych dwóch wektorów $\vec{u}, \vec{v} \in \mathbb{V}$ zachodzi:

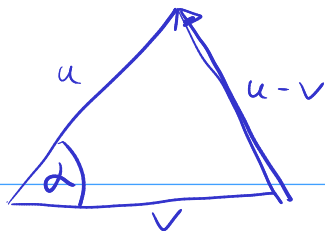
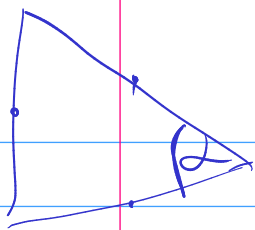
$$\langle F\vec{v}, F\vec{u} \rangle = \langle \vec{v}, \vec{u} \rangle .$$

Przykład 12.2. • obrót o kąt α (na płaszczyźnie)

- symetria względem prostej
- symetria względem płaszczyzny
- symetria względem punktu

Lemat 12.3. *Przekształcenie F jest izometrią wtedy i tylko wtedy gdy zachowuje długość, tj. dla każdego $\vec{v} \in \mathbb{V}$ mamy $\|F(\vec{v})\| = \|\vec{v}\|$.*

Przekształcenie F jest izometrią wtedy i tylko wtedy gdy zachowuje iloczyn skalarny elementów z bazy.



$$\cos \alpha = \frac{\langle u, v \rangle}{\|u\| \cdot \|v\|}$$

$\Rightarrow$ izometria

$$\langle v, v \rangle = \langle Fv, Fv \rangle$$

$$\|v\|^2 = \|Fv\|^2$$

$$\|v\| = \|Fv\|$$

$\Leftarrow$

$$\|u\|^2 = \|Fu\|^2$$

$$\|v\|^2 = \|Fv\|^2$$

$$\|u-v\|^2 = \|Fu-Fv\|^2$$

$\uparrow$

$$\langle u-v, u-v \rangle = \langle Fu-Fv, Fu-Fv \rangle$$

$$\langle u, u \rangle + \langle v, v \rangle - 2\langle u, v \rangle = \langle Fu, Fu \rangle + \langle Fv, Fv \rangle - 2\langle Fu, Fv \rangle$$

$$\langle u, v \rangle = \langle Fu, Fv \rangle$$

$$\langle u, v \rangle \stackrel{?}{=} \langle Fu, Fv \rangle$$

$\uparrow$

$$u = \sum \alpha_i \vec{w}_i$$

$$v = \sum \beta_j \vec{w}_j$$

$$\langle u, v \rangle = \langle \sum_i \alpha_i \vec{w}_i, \sum_j \beta_j \vec{w}_j \rangle = \sum_{i,j} \alpha_i \beta_j \langle \vec{w}_i, \vec{w}_j \rangle$$

$$\langle Fu, Fv \rangle = \langle F \sum_i \alpha_i \vec{w}_i, F \sum_j \beta_j \vec{w}_j \rangle = \sum_{i,j} \alpha_i \beta_j \langle F\vec{w}_i, F\vec{w}_j \rangle$$

Macierz iz w bazie ortonormalnej

12.2 Macierze ortogonalne

Definicja 12.4. Macierz kwadratową nazywamy *ortogonalną*, jeśli jej kolumny są parami ortogonalne oraz są długości 1 (w standardowym iloczynie skalarnym).



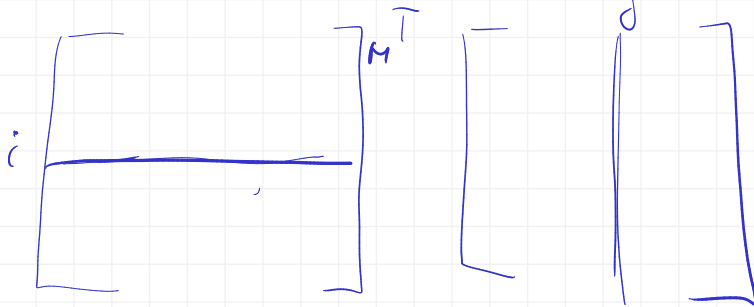
układ ortonormalny

Lemat 12.5. M jest ortogonalna wtedy i tylko wtedy gdy $M^{-1} = M^T$.

$$(M^T \cdot M)_{ij} = M_i \cdot M_j$$

$$V \cdot U = V^T U$$

$$M = [M_1 | M_2 | \dots]$$



$\Rightarrow M$ ortogonalna

$$(M^T \cdot M)_{ij} = \begin{cases} 0 & i \neq j \\ 1 & i = j \end{cases}$$

$$M_i \cdot M_j = \begin{cases} 0 & i \neq j \\ 1 & i = j \end{cases}$$

$$M^T \cdot M = Id \Rightarrow M^T = M^{-1}$$

$$\Leftarrow M^T = M^{-1} \Rightarrow M_i \cdot M_j = \begin{cases} 0 & i \neq j \\ 1 & i = j \end{cases} \Rightarrow \text{układ kol. jest ortonormalny} \\ (M^T \cdot M)_{ij} = \begin{cases} 0 & i \neq j \\ 1 & i = j \end{cases} \Rightarrow M \text{ jest ortogonalna}$$

Lemat 12.6. Macierze ortogonalne są zamknięte na mnożenie, transponowanie i na branie macierzy odwrotnej.

Prosty dowód pozostawiamy jako ćwiczenie.

Lemat 12.7. M jest macierzą ortogonalną wtedy i tylko wtedy gdy indukowane przez nią przekształcenie liniowe $L_M : \vec{V} \mapsto M\vec{V}$ jest izometrią (dla standardowego iloczynu skalarnego).

$\Rightarrow M$ jest ortogonalna

$$\vec{v} \rightarrow M\vec{v} \quad \text{izometria}$$

rozłożyj il. skł. na bazy $E_1, \dots, E_n$

$$E_i \cdot E_j = \begin{cases} 0 & i \neq j \\ 1 & i = j \end{cases} \Rightarrow \text{izometria}$$

$$ME_i \cdot ME_j = M_i \cdot M_j = \begin{cases} 0 & i \neq j \\ 1 & i = j \end{cases}$$

$\Leftarrow \vec{v} \rightarrow M\vec{v}$ jest izometrią

$$E_i \cdot E_j = (ME_i) \cdot (ME_j) = (M_i) \cdot (M_j)$$

$\parallel$

$$\begin{cases} 0 & i \neq j \\ 1 & i = j \end{cases}$$

$\nwarrow \nearrow$
kol.

Lemat 12.8. Niech $B = \{\vec{b}_1, \dots, \vec{b}_n\}$ będzie bazą ortonormalną.

F jest izometrią $\iff M_{BB}(F)$ jest macierzą ortogonalną. W szczególności, $M_{BB}(F)^{-1}$ to $M_{BB}(F)^T$.

$$\vec{b}_1, \dots, \vec{b}_n$$

$\Rightarrow F$ jest izometrią

$$M_{BB}(F) \stackrel{\text{def.}}{=} [(F\vec{b}_1)_B, (F\vec{b}_2)_B, \dots, (F\vec{b}_n)_B]$$

$$(F\vec{b}_i)_B \cdot (F\vec{b}_j)_B = \langle F\vec{b}_i, F\vec{b}_j \rangle \stackrel{i.z.}{=} \langle \vec{b}_i, \vec{b}_j \rangle = \begin{cases} 0 & i \neq j \\ 1 & i = j \end{cases}$$

$\Rightarrow M_{BB}(F)$ jest ortogonalna

$\Leftarrow M_{BB}(F)$ jest ortogonalna

$$(F\vec{b}_i)_B \cdot (F\vec{b}_j)_B = \begin{cases} 0 & i \neq j \\ 1 & i = j \end{cases}$$

$\parallel$

$$\langle F\vec{b}_i, F\vec{b}_j \rangle = \langle \vec{b}_i, \vec{b}_j \rangle$$

il. in. jest zach.
ma barię

$\Rightarrow F$ jest izometrią

M ortogonalna $\iff V \rightarrow MV$ to izometria

F jest iz. $\iff M_{BB}(F)$ jest ortogonalna
bazą ortonormalną

Rozdział 13

Macierze dodatnio określone: zadawanie iloczynu skalarnego przez macierz

$$u = \sum \alpha_i v_i$$
$$v = \sum \beta_j v_j$$

$$\langle u, v \rangle = \left\langle \sum_i \alpha_i v_i, \sum_j \beta_j v_j \right\rangle =$$
$$= \sum_{i,j} \alpha_i \beta_j \langle v_i, v_j \rangle$$

Definicja 13.1 (Macierz iloczynu skalarnego, macierz Grama). Dla bazy $B = \vec{v}_1, \dots, \vec{v}_n$ oraz iloczynu skalarnego $\langle \cdot, \cdot \rangle$ określamy macierz tego iloczynu w bazie B jako

$$M^B = (\langle \vec{v}_i, \vec{v}_j \rangle)_{i,j=1,\dots,n}$$

Lemat 13.2. Niech B : baza przestrzeni z iloczynem skalarnym $\langle \cdot, \cdot \rangle$. Wtedy

$$\langle \vec{u}, \vec{v} \rangle = (\vec{u})_B^T M^B (\vec{v})_B$$

$$(u \cdot v = u^T v)$$

Definicja 13.3 (Funkcjonał dwuliniowy, forma dwuliniowa). Niech $\mathbb{V}$ będzie przestrzenią liniową nad ciałem $\mathbb{F}$.

Funkcja $F : \mathbb{V} \times \mathbb{V} \rightarrow \mathbb{F}$ jest funkcjonałem dwuliniowym (formą dwuliniową), jeśli jest liniowa po każdej współrzędnej, tj.:

- $F(\alpha \vec{u}, \vec{v}) = \alpha F(\vec{u}, \vec{v})$;
- $F(\vec{u} + \vec{u}', \vec{v}) = F(\vec{u}, \vec{v}) + F(\vec{u}', \vec{v})$;
- $F(\vec{u}, \alpha \vec{v}) = \alpha F(\vec{u}, \vec{v})$;
- $F(\vec{u}, \vec{v} + \vec{v}') = F(\vec{u}, \vec{v}) + F(\vec{u}, \vec{v}')$.

Lemat 13.4. Niech F będzie funkcjonałem dwuliniowym a $B = \vec{v}_1, \dots, \vec{v}_n$ bazą przestrzeni $\mathbb{V}$. Wtedy F jest jednoznacznie zadany przez macierz

$$M^B(F) = (F(\vec{v}_i, \vec{v}_j))_{i,j=1,\dots,n}.$$

Co więcej,

$$[F(\vec{u}, \vec{v})] = [\vec{u}]_B^T M^B(F) [\vec{v}]_B$$

W szczególności, dla dwóch funkcjonałów dwuliniowych F, G zachodzi $F = G$ wtedy i tylko wtedy, gdy $M^B(F) = M^B(G)$.

$$\begin{aligned} u &= \sum \alpha_i v_i \\ v &= \sum \beta_j v_j \end{aligned}$$

$$F(u, v) = F\left(\sum \alpha_i v_i, \sum \beta_j v_j\right) = \sum \alpha_i \beta_j F(v_i, v_j)$$

$$(\alpha^T)_B \cdot M^B \cdot (\beta)_B = \left(\sum \alpha_i \cdot E_i^T\right)_B \cdot M^B \cdot \left(\sum \beta_j \cdot E_j\right)$$

$$= \sum \alpha_i \beta_j (E_i^T \cdot M^B \cdot E_j)$$

$$\begin{aligned} & \parallel \\ & F(u, v) \\ & (\alpha)_B^T M^B(F) (\beta)_B \end{aligned}$$

$$\begin{aligned} & \left[\begin{array}{c} E_i^T \\ 1 \\ \vdots \end{array} \right] \left[\begin{array}{c} \vdots \\ v_j \\ \vdots \end{array} \right] = (M^B)_{ij} \\ & \parallel \\ & F(v_i, v_j) \end{aligned}$$

$$= \sum \alpha_i \beta_j F(v_i, v_j)$$

Lemat 13.5. Jeśli F jest funkcjonałem dwuliniowym, zaś A, B są dwiema bazami, to

$$M^B(F) = M_{BA}^T M^A(F) M_{BA} ,$$

gdzie M_{BA} to macierz zmiany bazy z B do A .

dwa funkcjonały

$$(v)_B^T (u)_B$$

$$F(u, v) = (u)_B^T M^B (v)_B \stackrel{?}{=} F'(u, v) = (v)_B^T (u)_B$$

$b_i, b_j \in B$

$$(b_i)_B^T M^B (b_j)_B$$

$$F'(b_i, b_j) = \underbrace{(b_i)_B^T M_{BA}^T M^A(F) M_{BA} (b_j)_B}_{M_{BA} (b_j)_B^T = (b_j)_A^T} = (b_i)_A^T M^A(F) (b_j)_A = F(b_i, b_j)$$

Jako wniosek otrzymujemy:

Wniosek 13.6. Niech A, B to dwie bazy przestrzeni Euklidesowej. Wtedy

$$M^B = M_{BA}^T M^A M_{BA} ,$$

gdzie M_{BA} to macierz zmiany bazy.

Fakt 13.7. Dla bazy ortonormalnej B dla iloczynu skalarnego $\langle \cdot, \cdot \rangle$ mamy $M^B = \text{Id}$.

$$(M^B)_{ij} = \langle b_i, b_j \rangle = \begin{cases} 0 & i \neq j \\ 1 & i = j \end{cases}$$

$$M^B = \text{Id}$$

Definicja 13.8 (Macierz dodatnio określona). Macierz M wymiaru $n \times n$ jest *dodatnio określona*, jeśli funkcja $\langle \cdot, \cdot \rangle : (\mathbb{R}^n)^2 \rightarrow \mathbb{R}$ określona jako

$$(\vec{u}, \vec{v}) \mapsto \vec{u}^T M \vec{v}$$

jest iloczynem skalarnym na $\mathbb{R}^n$.

Dla jakich macierzy to jest dobra definicja?

$$\begin{array}{ccc} \langle u, v \rangle = \langle v, u \rangle & & \\ \langle E_i, E_j \rangle & \quad \quad & \langle E_j, E_i \rangle \\ \downarrow & & \uparrow \\ m_{ij} & & m_{ji} \end{array}$$

Fakt 13.9. Macierz M jest dodatnio określona wtedy i tylko wtedy gdy:

1. jest symetryczna oraz
2. dla każdego wektora $\vec{v} \neq 0$ zachodzi

$$\vec{v}^T M \vec{v} > 0.$$

Lemat 13.10. M jest dodatnio określona $\iff M = A^T A$ dla pewnej odwracalnej macierzy A .

Takie A można efektywnie uzyskać z dodatnio określonej symetrycznej macierzy M .

Co więcej, istnieje takie A , które jest górno-trójkątne i je również można efektywnie uzyskać.

$A \leftarrow$ górno-trójkątna
 $i > 0$ na przek.

$$\Leftarrow M = A^T A \quad M^T = (A^T A)^T = A^T A$$

$$(\vec{v}^T A^T A \vec{v}) = (A \vec{v})^T (A \vec{v})$$

$$A - \text{odwracalna}, \vec{v} \neq 0 \quad A \vec{v} \neq 0$$

$$(A \vec{v})^T (A \vec{v}) = (A \vec{v}) \cdot (A \vec{v}) = \|A \vec{v}\|^2 > 0$$

$\Rightarrow$

M - dodatnio określone, $\exists A$ $M = A^T A$

M wyznacza iloczyn skalarny
 $B \leftarrow$ baza ortonormalna

$$(F_M(\vec{u}, \vec{v})) = (\vec{u}^T \cdot M \vec{v})$$

$$M^E = \bar{M}$$

$$E_i^T M E_j^T = M_{ij}$$

$$M^B(F) = Id$$

$\langle b_i, b_j \rangle$

$$M = M^E(F) = M_{EB}^T \underbrace{M^B(F)}_{Id} M_{EB} = M_{EB}^T M_{EB}$$

Ortogonalizacja GS

$E \rightarrow A$

$M_{EA} \leftarrow$ górna trójkątna
 > 0 na przek.

Twierdzenie 13.11 (Rozkład Cholesky'ego). Dla macierzy dodatnio określonej M istnieje jedyne A górnotrójkątne o dodatnich elementach na przekątnej, takie że $M = A^T A$.

Taką reprezentację nazywamy rozkładem Cholesky'ego, co więcej, można ją efektywnie wyliczyć.

Algorytm wyliczający taką macierz jest zadaniem na ćwiczenia. Algorytm ten pokazuje też jedyność takiej macierzy.

Istnieją też inne wydajne sposoby sprawdzania, czy macierz jest dodatnio określona.

Będziemy korzystać z prostego faktu (Lemat i dowód pominięty na wykładzie):

Lemat 13.12. Macierz M jest dodatnio określona wtedy i tylko wtedy, gdy istnieje baza $\vec{B}_1, \dots, \vec{B}_n$ taka że

- dla $i \neq j$ mamy $\vec{B}_i^T M \vec{B}_j = 0$;
- dla każdego i mamy $\vec{B}_i^T M \vec{B}_i = 1$.

Innymi słowy, istnieje „baza ortonormalna” dla „iloczynu skalarowego” zadanego przez M .

Dowód. $\Rightarrow$ Jeśli M jest dodatnio określona, to zadane iloczyn skalarny, dla którego istnieje baza ortonormalna $\vec{B}_1, \dots, \vec{B}_n$; ta baza spełnia warunki zadania.

$\Leftarrow$ Niech F_M będzie formą dwuliniową zadaną przez M i M^A oznacza macierz tej formy w bazie A . Oczywiście $M^E = M$, gdzie E to baza standardowa.

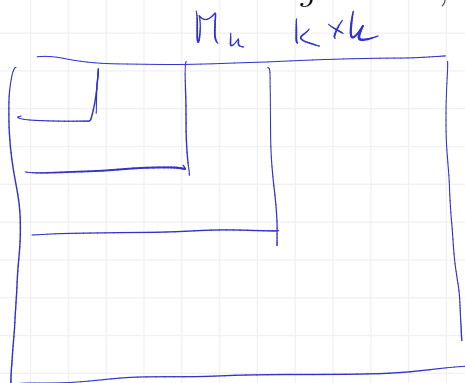
Niech $B = (\vec{B}_1, \dots, \vec{B}_n)$ będzie bazą jak w treści zadania. Wtedy $M^B = \text{Id}$ i w takim razie

$$M = M^E = M_{EB}^T M^B M_{EB} = M_{EB}^T M_{EB}$$

i tym samym M jest dodatnio określona. $\square$

Dla macierzy M niech M_k oznacza macierz $k \times k$ która jest „w lewym górnym rogu” macierzy M .

Twierdzenie 13.13 (Kryterium Sylwestera). Symetryczna macierz M jest dodatnio określona $\iff$ dla każdego $k = 1, 2, \dots, n$ macierz M_k spełnia $\det(M_k) > 0$.



$$\langle \vec{B}_i, \vec{B}_i \rangle \leq 0$$

$$\det(M_k) > 0$$

M

$M_k \leftarrow$ macierz il. sk. obciętego do $\mathbb{R}^k$

$$|M_k| \leq |A^T A| = |A^T| \cdot |A| = |A|^2 > 0$$

Do charakteryzacji macierzy dodatnio określonych można też podejść inaczej — od strony wartości własnych:

Twierdzenie 13.14. Symetryczna macierz liczb rzeczywistych M jest dodatnio określona wtedy i tylko wtedy, gdy ma same dodatnie wartości własne.

Dodatnio-określoną macierz (liczb rzeczywistych) można „spierwiastkować”:

Wniosek 13.15. Dla dodatnio określonej macierzy liczb rzeczywistych M istnieje macierz N , taka że $M = N^2$.

Dowód. Niech $M = A^{-1}DA$, gdzie D jest macierzą przekątniową, ma ona same dodatnie wartości $\lambda_1, \dots, \lambda_n$ na przekątnej. Weźmy D' : przekątniową o elementach $\sqrt{\lambda_1}, \dots, \sqrt{\lambda_n}$ na przekątnej i macierz $N = A^{-1}D'A$. Wtedy $N^2 = M$. $\square$

$$A^{-1}D$$

$$A = M$$

13.1 PCA raz jeszcze

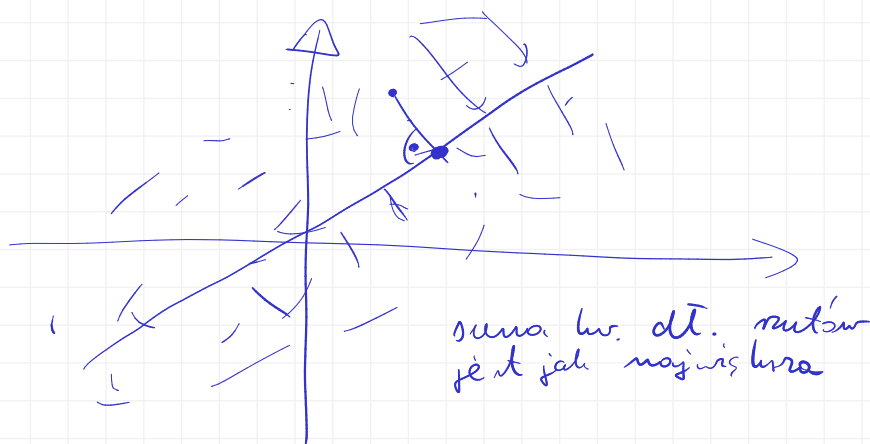
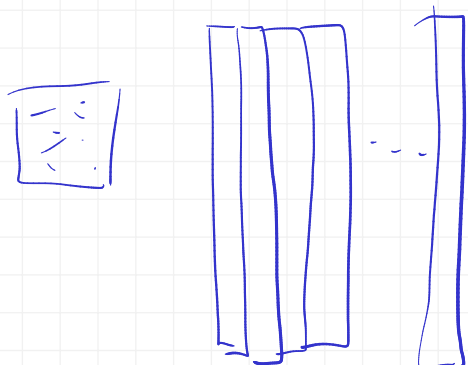
Przypomnijmy:

Na wejściu otrzymujemy kolekcję (m) zdjęć twarzy ludzi, traktujemy je jako wektory wymiaru $n = 256 \cdot 256$ z $\mathbb{R}^n$. Na wszystkie razem traktujemy jako macierz M wymiaru $n \times m$. Obrazy są uśrednione, tzn. dla każdego wektora zakładamy, że $\sum_i v_i = 0$.

$M \leftarrow$ wektory twarzy

$$A_j \sum_i m_{ij} = 0$$

$$(\sum_i M_{ij} = 0)$$



chcemy: v , $\|v\| = 1$

$$\sum_i \|P_v M_i\|^2 \leftarrow \text{maksymalne}$$

$$\sum_i (M_i \cdot v)^2 \cdot (\underbrace{v \cdot v}_1) \leftarrow \text{suma kw.}$$

$$= \sum_i (M_i \cdot v)^2 = \sum_i (M_i^T v)^2 = \sum_i (M_i^T v) \cdot (M_i^T v)$$

$$\begin{pmatrix} M_1^T \\ \vdots \\ M_m^T \end{pmatrix} \begin{pmatrix} v \\ \vdots \\ v \end{pmatrix} = \dots$$

$$M_i^T V$$

$$M^T$$

$$V$$

$$(M^T V) \cdot (M^T V)$$

$$\begin{pmatrix} \vdots \\ \vdots \\ \vdots \end{pmatrix} \cdot \begin{pmatrix} V \\ V \\ V \end{pmatrix} = \begin{pmatrix} M_1^T V \\ M_2^T V \\ \vdots \\ M_k^T V \end{pmatrix} \cdot \begin{pmatrix} M_1^T V \\ M_2^T V \\ \vdots \\ M_k^T V \end{pmatrix}$$

$$(M^T V)^T (M^T V)$$

$$= V^T (M M^T) V$$

$V = \text{matrix } 1 \times n$
 $\text{table } V, \text{ i.e. } \max$

$$V = \sum \alpha_i V_i \quad \sum \alpha_i^2 = 1 \quad (M M^T) - \text{kwadratowa, sym,}$$

„prawie” dod. obr.

$$\sum_{i,j} \alpha_i \alpha_j V_i^T (M M^T) V_j \geq 0$$

$$= \sum_{i,j} \alpha_i \alpha_j \lambda_j V_i^T V_j$$

λ_j
 $0 \leq \alpha_j$

$$= \sum_i \alpha_i^2 \lambda_i$$

$$\sum \alpha_i^2 = 1$$

$$\lambda_1 \geq \lambda_2 \geq \lambda_3 \geq \dots \geq \lambda_k$$

$\sum \alpha_i^2$

~~0~~

$$\lambda_1 = 1$$

$$\lambda_2 = \dots = \lambda_n = 0$$