Przestrzenie euklidesowe

Def. 1. *Iloczynem skalarnym* w dowolnej przestrzeni liniowej nad $\mathbb{R}$ nazywamy funkcję, która dowolnej parze wektorów x, y przyporządkowuje liczbę rzeczywistą (x, y) i spełnia dla dowolnych wektorów x, y, z i skalarów α, β warunki:

1. $(x, y) = (y, x)$ (*symetria*),
2. $(\alpha x + \beta y, z) = \alpha(x, z) + \beta(y, z)$ (*liniowość na pierwszej współrzędnej*),
3. $(x, x) \geq 0$ i $(x, x) = 0 \Leftrightarrow x = 0$ (*dodatnia określoność*).

Def. 2. Przestrzeń liniową, w której określony jest iloczyn skalarny nazywamy *przestrzenią euklidesową*. Przestrzeń euklidesową $\mathbb{R}^n$ ze standardowym iloczynem skalarnym oznaczamy E^n.

Fakt 1 (Liniowość na drugiej współrzędnej)**.** $(z, \alpha x + \beta y) = \alpha(z, x) + \beta(z, y)$.

Uwaga 1. Przekształcenie liniowe dowolnej przestrzeni liniowej w przestrzeń jej skalarów nazywamy *funkcjonałem liniowym.* Przekształcenie, które dowolnej parze wektorów przyporządkowuje skalar i jest liniowe na każdej współrzędnej nazywamy *funkcjonałem dwuliniowym.* Iloczyn skalarny definiuje się krótko jako symetryczny i dodatnio określony *funkcjonał dwuliniowy.*

Uwaga 2. Standardowy iloczyn skalarny określony wcześniej wzorem $\vec{x} \circ \vec{y} = x^T y$ jest szczególnym przypadkiem iloczynu (x, y). Będziemy pisać iloczyn skalarny w dowolnej przestrzeni euklidesowej (x, y), pamiętając, że w przestrzeni E^n $(x, y) = x^T y$. Inne przykłady iloczynów skalarnych pojawią się później.

Twierdzenie 2 (Nierówność Schwarza)**.** *Dla dowolnych wektorów x, y przestrzeni euklidesowej zachodzi:*

$$(x, y)^2 \leq (x, x)(y, y).$$

W przestrzeni E^n dla $x^T = [x_1, x_2, ..., x_n]$, $y^T = [y_1, y_2, ..., y_n]$:

$$\Big(\sum_{i=1}^{n} x_i y_i\Big)^2 \leq \Big(\sum_{i=1}^{n} x_i^2\Big)\Big(\sum_{i=1}^{n} y_i^2\Big).$$

Def. 3. *Normą* wektora x w przestrzeni euklidesowej nazywamy liczbę $||x|| = \sqrt{(x, x)}$. Wektor x nazywamy *unormowanym*, gdy $||x|| = 1$.

Norma wektora nazywana jest też jego *długością* i oznaczana $|\vec{x}|$.

Twierdzenie 3 (Nierówność trójkąta)**.** $||x + y|| \leq ||x|| + ||y||$.

Uwagi o normie

Uwaga 3. Normę wektora można wprowadzić w dowolnej przestrzeni liniowej V nad $\mathbb{R}$ jako dowolną funkcję, która spełnia warunki (dla $x, y \in V$, $\lambda \in \ R$):

1. $||\lambda x|| = |\lambda| ||x||$;
2. $||x + y|| \leq ||x|| + ||y||$;
3. $||x|| \geq 0$ i $||x|| = 0 \Leftrightarrow x = 0$.

Przestrzeń liniowa z określoną normą nazywa się *przestrzenią unormowaną*.

Uwaga 4. Norma wektora w przestrzeni euklidesowe oznaczana bywa w literaturze $||x||_2$ ponieważ jest szczególnym przypadkiem normy $||x||_p$ zdefiniowanej w $\mathbb{R}^n$ dla dowolnego $p \geq 1$ wzorem

$$||x|| = (|x_1|^p + |x_2|^p + ... + |x_n|^p)^{\frac{1}{p}}.$$

Wektory ortogonalne, kąty

Def. 4. Mówimy, że wektory x i y są *ortogonalne* (*prostopadłe*), gdy $(x, y) = 0$. Piszemy $x \perp y$.

Def. 5. *Kątem* pomiędzy niezerowymi wektorami x i y nazywamy liczbę rzeczywistą $\varphi \in [0, \pi]$ taką, że

$$\cos\varphi = \frac{(x, y)}{||x|| \cdot ||y||}.$$

Uwaga 5. W przestrzeniach E^2 i E^3 tak zdefiniowany kąt jest odpowiednikiem rzeczywistej miary kąta pomiędzy wektorami.

Wn. 1. $(x, y) = ||x|| \cdot ||y|| \cdot \cos\varphi$.

Prostopadłość podprzestrzeni, dopełnienie ortogonalne

Def. 6. Mówimy, że dwie podprzestrzenie U i W przestrzeni euklidesowej E są *prostopadłe* (*ortogonalne*) co zapisujemy $U \perp W$, gdy $u \perp w$ dla wszystkich $u \in U$ i $w \in W$. Piszemy też $b \perp U$, gdy $lin\{b\} \perp U$.

Fakt 4. *Jeśli $\mathcal{B}_U$ i $\mathcal{B}_W$ są dowolnymi bazami podprzestrzeni U i W odpowiednio, to $U \perp W$ wtedy i tylko wtedy, gdy każdy wektor z bazy $\mathcal{B}_U$ jest prostopadły do każdego z $\mathcal{B}_W$.*

Def. 7. *Dopełnieniem ortogonalnym* podprzestrzeni U przestrzeni euklidesowej E nazywamy zbiór $U^\perp := \{x \in E : \ x \perp U\}$.

Fakt 5. *$U^\perp$ jest podprzestrzenią dla dowolnej podprzestrzeni U. Jeśli U jest podprzestrzenią przestrzeni E^n, to $\dim(U^\perp) = n - \dim U$.*

Iloczyn wektorowy w E^3

Def. 8. Mówimy, że układ wektorów (u, v, w) jest zorientowany zgodnie z bazą, gdy wyznacznik macierzy współrzędnych tych wektorów w tej bazie jest dodatni.

Def. 9. *Iloczynem wektorowym* wektorów a, b nazywamy:

1. wektor w ortogonalny do a, b, którego długość jest równa polu równoległoboku rozpiętego na a, b i taki że układ (a, b, w) jest zorientowany zgodnie z bazą, gdy a, b są liniowo niezależne,

2. wektor zerowy, gdy a, b są liniowo zależne.

   Iloczyn wektorowy oznaczamy $w = a \times b$.

Postać analityczna iloczynu wektorowego

Twierdzenie 6. $\begin{bmatrix} a_1 \\ a_2 \\ a_3 \end{bmatrix} \times \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix} = \left[\begin{vmatrix} a_2 & a_3 \\ b_2 & b_3 \end{vmatrix}, - \begin{vmatrix} a_1 & a_3 \\ b_1 & b_3 \end{vmatrix}, \begin{vmatrix} a_1 & a_2 \\ b_1 & b_2 \end{vmatrix} \right]^T =$ $\begin{vmatrix} i & j & k \\ a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \end{vmatrix}$, *gdzie i, j, k są wektorami bazy kanonicznej.*

Własności iloczynu wektorowego:

1. $a \times b = -b \times a$, (*antysymetria*)

2. $a \times (b + c) = a \times b + a \times c$,

3. $(\alpha a) \times b = \alpha(a \times b) = a \times (\alpha b)$ dla dowolnego $\alpha \in \mathbb{R}$,

4. $||a \times b|| = ||a|| \cdot ||b|| \cdot \sin \angle(a, b)$.

Przykład 1. Wyznaczyć ortogonalną bazę dopełnienia ortogonalnego wektora $[1, 2, 3]^T$.

Iloczyn mieszany, geometryczna interpretacja wyznacznika

Def. 10. *Iloczynem mieszanym* wektorów a, b, c nazywamy $(a, (b \times c))$.

Twierdzenie 7. $a \circ (b \times c) = \begin{vmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{vmatrix}$

Wn. 2. *Objętość równoległościanu rozpiętego na wektorach a, b, c:* $V(a, b, c) = |(a, (b \times c))| = \left| \begin{vmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{vmatrix} \right|$.

Rzuty ortogonalne

Def. 11. *Rzutem prostokątnym* (*ortogonalnym*) wektora na niezerową podprzestrzeń U przestrzeni euklidesowej E nazywamy przekształcenie π_U, które każdemu wektorowi x przyporządkowuje taki wektor $x' = \pi_U(x)$, że wektor $(x - x') \perp U$. Rzut na podprzestrzeń jednowymiarową $lin\{b\}$ oznaczamy π_b.

Fakt 8. *π_u jest przekształceniem liniowym dla dowolnej podprzestrzeni U.*

Fakt 9. *Dla dowolnego niezerowego $b \in E^n$, $\pi_b(x) = \dfrac{bb^T}{||b||^2}x$.*

Przykład 2. Znaleźć macierz rzutu π_b dla $b^T = [1, 2, 2]$ oraz $\pi_b(x)$, $\pi_b(y)$ dla $x^T = [2, 1, 1]$ i $y^T = [2, -2, 1]$.

Uwaga 6. $\pi_b(x)$ można znaleźć mnożąc macierz $\dfrac{bb^T}{||b||^2}$ przez wektor x lub skalar $\dfrac{b^Tx}{||b||^2}$ przez wektor b.

Fakt 10. *Jeśli kolumny macierzy B tworzą bazę podprzestrzeni U przestrzeni E^n, to*

$$\pi_U(x) = B(B^TB)^{-1}B^Tx$$

.

Przykład 3. Znaleźć macierz rzutu π_U dla $U = lin\{[1, 1, 0]^T, [1, 1, 1]^T\}$ i $\pi_U(x)$ dla $x^T = [2, 0, 2]$.

Uwaga 7. Macierz $(B^TB)^{-1}B^T$ nazywana jest *pseudoodwrotnością* macierzy B. Taka macierz istnieje dla dowolnej (niekoniecznie kwadratowej!) macierzy B o liniowo niezależnych kolumnach.

Uwaga 8. Macierz rzutu π_b jest szczególnym przypadkiem macierzy rzutu π_U, ponieważ dla dowolnego niezerowego wektora b iloczyn b^Tb jest niezerowym skalarem i $(b^Tb)^{-1} = \dfrac{1}{||b||^2}$.

Def. 12. Macierz P nazywamy *idempotentną*, gdy $P^2 = P$.

Fakt 11. *Macierze rzutów są symetryczne i idempotentne.*

Twierdzenie 12. *Każda symetryczna i idempotentna macierz P jest macierzą rzutu na swoją przestrzeń kolumn.*

Fakt 13. *Dla dowolnego $x \in E^n$ i y należącego do podprzestrzeni U zachodzi*

$$||x - \pi_U(x)|| \leq ||x - y||.$$

Błąd przybliżenia, najlpepsze przybliżone rozwiązanie

Def. 13. Wektor $x-y$ nazywamy *wektorem błędu* aproksymacji wektora x wektorem $y \in U$, a jego długość $||x-y)||$ *wielkością błędu* tej aproksymacji.

Wn. 3. *Najlepszą aproksymacją dowolnego wektora $x \in E^n$ wektorem z podprzestrzeni W jest jego rzut ortoganlny na tę podprzestrzeń.*

Def. 14. *Najlepszym rozwiązaniem* układu równań $Ax = b$, gdzie $A \in \mathbf{M}_{m\times n}$ nazywamy taki wektor $\overline{x}$, że

$$||A\overline{x} - b|| \leq ||Ax - b||,$$

dla wszystkich $x \in E^n$.

Metoda najmniejszych kwadratów

Twierdzenie 14. *Zbiór najlepszych rozwiązań układu równań $Ax = b$ pokrywa się ze zbiorem rozwiązań układu:*

$$A^T Ax = A^T b.$$

Przedstawiony w Twierdzeniu 14 sposób znalezienia najlepszego rozwiązania układu $Ax = b$ nazywamy *metodą najmniejszych kwadratów*, a układ $A^T Ax = A^T b$ jego *układem normalnym.*

Przykład 4. Znaleźć najlepsze rowiązanie układu równań: $\left\{\begin{array}{rcrcr} x & + & y & = & 2 \\ x & + & 2y & = & 3 \\ x & + & 2y & = & 2 \end{array}\right.$.

Prosta regresji

Przykład 5. Znaleźć prostą $y = ax + b$ o tej własności, że dla danych punktów $(x_1, y_1), (x_2, y_2), ..., (x_k, y_k)$ suma

$$\sum_{i=1}^{k} [(ax_i + b) - y_i]^2$$

jest najmniejsza. Wyznaczyć taką prostą dla punktów $(-2, 1)$, $(0, 0)$, $(2, 1)$, $(4, 2)$.

- Szukamy najlepszego rozwiązania układu równań liniowaych. $\begin{bmatrix} x_1 & 1 \\ x_2 & 1 \\ \vdots & \vdots \\ x_k & 1 \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix} =$

$\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_k \end{bmatrix}$ z niewiadomymi a i b.

- Układ normalny $A^TAx = A^Ty$ przyjmuje postać:

$$\begin{bmatrix} \sum_{i=1}^k x_i^2 & \sum_{i=1}^k x_i \\ \sum_{i=1}^k x_i & k \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} \sum_{i=1}^k x_i y_i \\ \sum_{i=1}^k y_i \end{bmatrix}$$

- $det(A^TA) = \sum_{1 \le i < j \le k} (x_i - x_j)^2 \neq 0$ jeśli tylko nie wszystkie x_i są równe
- i wtedy układ ma dokładnie jedno rozwiązanie.
- Znaleziona w ten sposób prosta $y = ax + b$ nazywana jest *prostą regresji.*

Uwaga 9. Jeśli kolumny macierzy B są parami ortogonalne, to macierz B^TB jest diagonalna (i łatwo ją odwrócić!). Jeśli dodatkowo każda kolumna ma długość 1, to B^TB jest macierzą jednostkową i $\pi_U(x) = BB^Tx$.

Przykład 6. Znaleźć macierz π_U dla $U = lin\{[1, 2, 2]^T, [2, -2, 1]^T\}$.

Def. 15. Bazę nazywamy *ortogonalną, gdy każde jej dwa wektory są ortogonalne.* Jeśli dodatkowo jej wektory są unormowane, to nazywamy ją *bazą ortonormalną.*

Iloczyn skalarny w bazie ortonormalnej

Fakt 15. *Jeśli $\mathcal{B} = \{b_1, b_2, ..., b_k\}$ jest ortonormalną bazą k-wymiarowej przestrzeni euklidesowej z dowolnym iloczynem skalarnym,* $x' = \begin{bmatrix} x'_1 \\ x'_2 \\ \vdots \\ x'_k \end{bmatrix}$ *i* $y = \begin{bmatrix} y'_1 \\ y'_2 \\ \vdots \\ y'_k \end{bmatrix}$ *są odpowiednio wektorami współrzędnych wektorów x, y w bazie $\mathcal{B}$, to:*

$$(x, y) = (x')^T y'.$$

W szczególności $||x|| = \sqrt{(x'_1)^2 + (x'_2)^2 + ... + (x'_k)^2}$.

(Dowolny iloczyn skalarny i norma wektora w dowolnej bazie ortonormalnej wyrażają się tym samymi wzorami, co standardowy iloczyn skalarny i długość wektora w E^k.)

Ortogonalizacja Grama-Schmidta

Twierdzenie 16. *Niech $\{u_1, u_2, ..., u_k\}$ będzie bazą podprzestrzeni W przestrzeni euklidesowej E^n. Wówczas układ wektorów:*

$$\begin{cases} v_1 = u_1, \\ v_2 = u_2 - \frac{u_2^T v_1}{|v_1|^2} v_1, \\ v_3 = u_3 - \left[\frac{u_3^T v_1}{|v_1|^2} v_1 + \frac{u_3^T v_2}{|v_2|^2} v_2\right], \\ \vdots \\ v_k = u_k - \left[\frac{u_k^T v_1}{|v_1|^2} v_1 + ... + \frac{u_k^T v_{k-1}}{|v_{k-1}|^2} v_{k-1}\right], \end{cases} ;$$

jest ortogonalną bazą podprzestrzeni W.

Uwaga 10. Wektory $\frac{v_1}{||v_1||}, \frac{v_2}{||v_2||}, ..., \frac{v_k}{||v_k||}$ tworzą ortonormalną bazę podprzestrzeni W.

Uwaga 11. Zamiast $\frac{u^T v}{||v||^2} v$ można brać $\frac{v v^T}{||v||^2} u$. Jednak wyznaczanie iloczynów skalarnych jest wygodniejsze od wyznaczania macierzy rzutów.

Uwaga 12. Ortogonalizacja Grama-Schmidta polega na odejmowaniu od kolejnych wyjściowych wektorów ich rzutów prostokątnych na podprzestrzenie generowane przez poprzednie wektory.

Uwaga 13. Wektory v_i w ortogonalizacji można zastąpić dowolnymi wektorami z nimi proporcjonalnymi.

Przykład 7. Zortogonalizować metodą Grama-Schmidta układy wektorów:

1. $\{[1, 1, 1]^T, [1, 2, 3]^T\}$.
2. $\{[1, 2, 2]^T, [4, 5, 2]^T, [5, 1, 1]^T\}$

Macierze ortogonalne

Def. 16. Macierz kwadratową A nazywamy *ortogonalną*, gdy $A^{-1} = A^T$.

Fakt 17. *Dla macierzy kwadratowej A stopnia n następujące warunki są równoważne:*

1. *A jest ortogonalna;*
2. *$AA^T = I$;*
3. *$A^T A = I$;*
4. *wiersze macierzy A tworzą bazę ortonormalną przestrzeni $\mathbb{R}^n$;*
5. *kolumny macierzy A tworzą bazę ortonormalną przestrzeni $\mathbb{R}^n$.*

Przykłady macierzy ortogonalnych

- Macierz obrotu o kąt α na płaszczyźnie $\begin{bmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{bmatrix}$,
- Macierze symetrii względem Ox - $\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$, względem Oy - $\begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix}$ i względem prostej $y = x$ - $\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$;
- $-\mathcal{I}_n$ - macierz symetrii względem początku układu w $\mathbb{R}^n$;
- Macierze permutacji P_σ, np. $P_{(1,2,3)} = \begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}$

  Stąd wniosek, że $P_\sigma^{-1} = P_\sigma^T$;
- $\begin{bmatrix} \cos\alpha & -\sin\alpha & 0 \\ \sin\alpha & \cos\alpha & 0 \\ 0 & 0 & 1 \end{bmatrix}$ jest macierzą obrotu wokół Oz;
- $\frac{1}{3}\begin{bmatrix} 2 & -2 & 1 \\ 2 & 1 & -2 \\ 1 & 2 & 2 \end{bmatrix}$, $\frac{1}{9}\begin{bmatrix} 1 & -8 & -4 \\ -8 & 1 & -4 \\ -4 & -4 & 7 \end{bmatrix}$.

Endomorfizmy ortogonalne

Def. 17. Endomorfizm liniowy L przestrzeni E^n, który zachowuje iloczyn skalarny nazywamy *Endomorfizmem ortogonalnym.*

$$(x, y) = (L(x), L(y))$$

dla dowolnych x, y.

Wn. 4. *Endomorfizm ortogonalny L zachowuje normę wektora, ortogonalność wektorów i kąty pomiędzy wektorami x, y:*

1. $||x|| = ||L(x)||$ *dla dowolnego x, (czyli L jest izometrią),*
2. $x \perp y \Rightarrow L(x) \perp L(y)$,
3. $\angle(x, y) = \angle(L(x), L(y))$.
4. *Macierz A endomorfizmu L w bazie standardowej jest ortogonalna wtedy i tylko wtedy, gdy L jest ortogonalny.*

Grupy $O(n)$ i $SO(n)$

Twierdzenie 18. *1. Macierze ortogonalne stopnia n z działaniem mnożenia tworzą grupę przekształceń.*

2. Wyznacznik macierzy ortogonalnej jest równy 1 lub -1.

3. Macierze ortogonalne stopnia n o wyznaczniku jeden są podgrupą grupy wszystkich macierzy ortogonalnych stopnia n.

Def. 18. Grupę macierzy ortogonalnych stopnia n nazywamy *grupą ortogonalną stopnia n* i oznaczamy $O(n)$, a jej podgrupę macierzy o wyznaczniku 1 nazywamy *specjalną grupą ortogonalną* i oznaczamy $SO(n)$.

Diagonalizacja macierzy symetrycznych za pomocą macierzy ortogonalnych

Twierdzenie 19. *Macierz endomorfizmu L w dowolnej bazie ortonormalnej jest symetryczna wtedy i tylko wtedy, gdy $(L(x), y) = (x, L(y))$ dla dowolnych wektorów x i y.*

Twierdzenie 20 (Spektralne dla rzeczywistych macierzy symetrycznych)**.** *Niech S będzie macierzą symetryczną stopnia n. Wówczas:*

1. *Wartości własne macierzy S są liczbami rzeczywistymi.*
2. *Wektory własne odpowiadające różnym wartościom własnym są ortogonalne.*
3. *Istnieje baza przestrzeni E^n złożona z unormowanych i ortogonalnych wektorów własnych macierzy S. Macierz Q przejścia z bazy kanonicznej do bazy wektorów własnych jest ortogonalna i macierz $Q^T SQ$ jest diagonalna.*

Twierdzenie 21. *Macierz rzeczywista kwadratowa jest diagonalizowalna za pomocą macierzy ortogonalnej wtedy i tylko wtedy, gdy jest symetryczna.*

Przykład 8. Zdiagonalizować ortogonalnie dane macierze symetryczne. Wskazać odpowiednie macierze diagonalne i ortogonalne.

1. $\begin{bmatrix} 3 & -1 \\ -1 & 3 \end{bmatrix}$,
2. $\begin{bmatrix} 5 & -4 & 2 \\ -4 & 5 & 2 \\ 2 & 2 & 8 \end{bmatrix}$.

Rozkład $S = Q\Lambda Q^T$

Wn. 5. *Każda rzeczywista macierz symetryczna* S *ma rozkład*

$$S = Q\Lambda Q^T.$$

$Q = \begin{bmatrix} | & | & | & | \\ v_1 & v_2 & ... & v_n \\ | & | & | & | \end{bmatrix}$ *jest ortogonalną macierzą unormowanych wektorów własnych macierzy* S, *a* $\Lambda = \begin{bmatrix} \lambda_1 & 0 & ... & 0 \\ 0 & \lambda_2 & ... & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & ... & \lambda_n \end{bmatrix}$, *gdzie* $\lambda_1, \lambda_2, ..., \lambda_n$ *są wartościami własnymi wektorów* $v_1, v_2, ..., v_n$ *odpowiednio.*

Rozkład spektralny macierzy symetrycznej

Wn. 6. *Jeżeli* $v_1, v_2, ..., v_n$ *są unormowanymi wektorami własnymi symetrycznej macierzy* S, *a* $\lambda_1, \lambda_2, ..., \lambda_n$ *są odpowiadającymi im wartościami własnymi, to*

$$S = \lambda_1 v_1 v_1^T + \lambda_2 v_2 v_2^T + ... + \lambda_n v_n v_n^T.$$

Def. 19. Opisany we wniosku rozkład macierzy symetrycznej na sumę macierzy rzędu 1 nazywamy *rokładem spektralnym.*

Przykład 9. Znaleźć rokłady spektralne macierzy symetrycznych z przykładu 8.