Przestrzeń Euklidesowa – cz. 2

Rzut na podprzestrzeń wektorową i metoda najmniejszych kwadratów

Def. Niech E = (V, <*,*>) będzie przestrzenią Euklidesa, zaś V_1 podprzestrzenią wektorową przestrzeni V. Wektor $h \in V$ nazywamy wektorem ortogonalnym do p.p.w. V_1, jeśli jest on ortogonalny do każdego wektora $y \in V_1$: $\langle h, y\rangle = 0$.

Wn. Aby wektor $h \in V$ był ortogonalny do p.p.w V_1 potrzeba i wystarcza, aby był ortogonalny do wszystkich wektorów dowolnej bazy p.w. V_1.
Istotnie, jeśli $\langle h, b_k\rangle = 0$ dla wektorów bazy $\{b_k\}_{k=1,2,\ldots,n}$ p.p.w. V_1 to również dla dowolnego wektora $y = c_1 b_1+c_2 b_2+\ldots+c_n b_n$, jako liniowej kombinacji wektorów bazy, zachodzi $\langle h, y\rangle = \langle h, c_1 b_1+c_2 b_2+\ldots+c_n b_n \rangle = 0$ wobec liniowości iloczynu skalarnego. Warunek konieczny jest oczywisty.

P1(a). Wektor $\boldsymbol{u} = [5, 2, -1]$ jest ortogonalny do p.p.w. $V_1 \subset \boldsymbol{R}^3$ danej równaniem $5x + 2y - z = 0$ (płaszczyzna π), gdyż $\boldsymbol{u}$ jest jej wektorem normalnym i równanie to możemy zapisać z użyciem standardowego iloczynu skalarnego w postaci: $\langle \boldsymbol{u}, [x, y, z]\rangle = 0$. Dlatego, każda baza p.p.w. V_1, np. układ $\{b_1 = [1, 0, 5], b_2 = [0, 1, 2]\}$, zawiera wektory ortogonalne do wektora $\boldsymbol{u}$.
(b) Wektor $w = [2, -3, 4]$ jest ortogonalny do p.p.w. $V_2 \subset \boldsymbol{R}^3$ danej równaniem parametrycznym l: $[x, y, z] = t\,[5, 2, -1]$, dla $t \in \boldsymbol{R}$ - prosta $l \perp \pi = V_1$ z P1(a).

Def. Podprzestrzenie wektorowe V_1 i V_2 przestrzeni Euklidesa E = (V, <*,*>) nazywamy ortogonalnymi, jeśli dowolne wektory $x \in V_1$ i $y \in V_2$ są ortogonalne.

Def. Zbiór wszystkich wektorów p.w. V ortogonalnych do p.p.w. $V_1 \subset V$ nazywamy dopełnieniem ortogonalnym p.p.w. V_1, ozn. $V_1^{\perp}$.

P2. W P1. prosta l, jako p.p.w. $V_2 \subset V = \boldsymbol{R}^3$, jest dopełnieniem ortogonalnym płaszczyzny π, czyli $V_2 = V_1^{\perp}$ oraz jednocześnie $V_1 = V_2^{\perp}$ – dlaczego ?
Uwaga: (*i*) $V^{\perp} = \mathbf{0}$, $\mathbf{0}^{\perp} = V$; (*ii*) $(V_1^{\perp})^{\perp} = V_1$; (*iii*) $V_1 \subset V_2 \Rightarrow V_2^{\perp} \subset V_1^{\perp}$.

Def. Niech V_1 będzie p.p.w. przestrzeni Euklidesa $E = (V, <*,*>)$.

Dla dowolnego wektora $y \notin V_1$ wektor $y_0 \in V_1$, taki że wektor $h = y - y_0$ jest ortogonalny do V_1 nazywamy rzutem wektora y na p.p.w. V_1.

Wn. Wektor $h = y - y_0 \perp V_1$ określa najmniejszą odległość $|h|$ wektora y od p.p.w. V_1, czyli $|y - y_0| = \min_{x \in V_1} d(y, x) < |y - y_1|$ dla $y_1 \in V_1$ jeśli tylko $y_1 \neq y_0$.

Istotnie, bo $y_0 - y_1 \in V_1$ i stąd $<h, y_0 - y_1> = 0$. Wówczas z tw. Pitagorasa mamy: $|y - y_0|^2 + |y_0 - y_1|^2 = |y - y_0 + y_0 - y_1|^2 = |y - y_1|^2$, czyli teza dla $y_1 \neq y_0$.

Tw (konstrukcja rzutu ortogonalnego wektora na p.p.w.)

Dla p.p.w. V_1 z bazą $\{b_k\}_{k=1, 2, \ldots, m}$ określimy rzut wektora $y \notin V_1$ jako wektor $y_0 \in V_1$, a więc postaci: $y_0 = c_1 b_1 + c_2 b_2 + \ldots + c_m b_m$, gdzie współczynniki $c_k \in \boldsymbol{R}$ znajdziemy z warunków ortogonalności: $< y - y_0, b_k > = 0$. Warunki te tworzą układ *m* równań liniowych dla *m* niewiadomych $\{c_1, c_2, \ldots, c_m\}$:

$$\boxed{(*) \quad <y, b_k> = c_1 <b_1, b_k> + c_2 <b_2, b_k> + \ldots + c_m <b_m, b_k>, \ k = 1, 2, 3, \ldots, m.}$$

Macierz układu (*) nazywamy macierzą Gramma: $A = [<b_i, b_j>]_{i, j = 1, 2, \ldots, m}$.

Wn. Macierz Gramma jest nieosobliwa dla wektorów liniowo niezależnych.

Dowód: ćwiczenie.

Stąd, układ (*) jest układem Cramera i dlatego posiada zawsze jednoznaczne rozwiązanie niezerowe, jeśli tylko $y \neq \mathbf{0}$.

Wn. Dla układu ortogonalnych wektorów macierz Gramma jest diagonalna.

Dowód: ćwiczenie.

Stąd, gdy baza $\{b_k\}_{k=1, 2, \ldots, m}$ p.p.w. V_1 jest

(i) bazą ortogonalną, to $c_k = <y, b_k>/<b_k, b_k>$,

(ii) bazą ortonormalną to $c_k = <y, b_k>$ - rzut wektora y na k-ty wektor bazy.

Wskazanie współrzędnych $\{c_k\}$ wektora y_0 w bazie $\{b_k\}_{k=1, 2, \ldots, m}$ p.p.w. V_1 kończy konstrukcję rzutu wektora y na tę podprzestrzeń. Ozn. $y_0 = \mathrm{Proj}_{V1}(y)$.

P3. Wyznaczmy rzut wektora y = [1, 1, 1, -1]$\in \boldsymbol{R}^4$ na 1-wymiarową p.p.w. V_1 określoną przez układ 3 równań $x_1 + x_2 = x_2 + x_3 = x_3 + x_4 = 0$. Przechodząc do równania parametrycznego otrzymujemy V_1 = E([1, -1, 1, -1]), a stąd dla standardowego iloczynu skalarnego mamy y_0 = 1/2 [1, -1, 1, -1], co daje odległość wektora y od V_1: d(y, y_0) = $(1/4 + 9/4 + 1/4 + 1/4)^{1/2} = 3^{1/2}$. Zauważmy, że Proj_{V1}([1, 1, 1, 1]) = **0**, bo y' = [1, 1, 1, 1]$\in V_1^{\perp}$.

P4. Używając bazy ortogonalnej wielomianów Legendre'a {1, x, x^2-1/3, x^3-3/5x} w podprzestrzeni wielomianów stopnia co najwyżej 3, $\boldsymbol{R}_3$[x], łatwo znaleźć rzut wektora y = x^4 na tę p.p.w. przy iloczynie skalarnym danym przez całkę w przedziale [-1, 1]. Jest to tzw. aproksymacja wielomianowa. Ponieważ $\langle x^4, x\rangle = \langle x^4, x^3-3/5x\rangle = 0$ to rzut $y_0 = c_0 + c_2 (x^2-1/3) = 6/7\, x^2 – 3/35$, wobec $c_0 = \langle x^4, 1\rangle/\langle 1,1\rangle = 1/5$ i $c_2 = \langle x^4, x^2-1/3\rangle/\langle x^2-1/3, x^2-1/3\rangle = 6/7$.

P5. W p.w. funkcji ciągłych na przedziale [0, 2π] rozważmy (2n+1)-wymiarową p.p.w. V_1 z bazą ortogonalną funkcji {1, cos(x), sin(x), cos(2x), sin(2x),..., cos(nx), sin(nx)}. Rzut dowolnej funkcji ciągłej f(x) na tę p.p.w. jest optymalną aproksymacją tej funkcji przez tzw. wielomiany trygonometryczne postaci $P(x) = a_0+a_1\cos(x)+b_1\sin(x)+a_2\cos(2x)+b_2\sin(2x)+...+ a_n\cos(nx)+b_n\sin(nx)\in V_1$, gdzie współczynniki są znów postaci $a_k = \langle f(x), \cos(kx)\rangle/\langle \cos(kx), \cos(kx)\rangle$ oraz $b_k = \langle f(x), \sin(kx)\rangle/\langle \sin(kx), \sin(kx)\rangle$, gdzie $\langle f, g\rangle = \int_0^{2\pi} f(x)g(x)dx$.
Wówczas, np. x/2 = π/2 – (sin(x) + ½ sin(2x) + ...+1/n sin(nx)) dla x∈[0, 2π]. Rzut funkcji ciągłej na p.p.w. wielomianów trygonometrycznych jest zadaniem poszukiwania tzw. szeregu Fouriera funkcji ciągłej na przedziale. Współczynniki $\{a_0, a_k, b_k\}_{k=1, 2,..., n}$ są identyczne jak te znane z teorii szeregów Fouriera:

$$a_0 = \frac{1}{\sqrt{2\pi}} \int_0^{2\pi} f(x)dx;\ a_k = \frac{1}{\sqrt{\pi}} \int_0^{2\pi} f(x)\cos(kx)dx;\ b_k = \frac{1}{\sqrt{\pi}} \int_0^{2\pi} f(x)\sin(kx)dx\,.$$

Powiększając wymiar 2n+1 = dim V_1 dokonujemy najczęściej lepszego przybliżenia rzutowanej funkcji, tzn. odległość d(f, V_1) staje się coraz mniejsza.

Metoda najmniejszych kwadratów – rozwiązywanie układów sprzecznych.

Niech teoretycznie wielkość y zależy liniowo od zmiennych $\{x_1, x_2, ..., x_m\}$, czyli $y = c_1x_1 + c_2x_2 + ... + c_mx_m$, gdzie parametry c_k są szukanymi niewiadomymi. W pewnym doświadczeniu wyznaczamy n-krotnie ciąg $\{y, x_1, ..., x_m\}$ z błędami pomiarowymi, tak że układ

$$
\begin{aligned}
&y_1 \approx c_1x_{11} + c_2x_{21} + ... + c_mx_{m1}, \\
&y_2 \approx c_1x_{12} + c_2x_{22} + ... + c_{m2}x_{m2}, \qquad \Leftrightarrow \qquad y \approx c_1\, b_1 + c_2\, b_2 + ... + c_m\, b_m \\
&.................................... \\
&y_n \approx c_1x_{1n} + c_2x_{2n} + ... + c_{mn}x_{mn},
\end{aligned}
$$

jest sprzeczny, co znaczy że wektor kolumnowy $y = [y_1, y_2,..., y_n]^T$ <u>nie jest liniową kombinacją</u> wektorów kolumnowych:

$b_1 = [x_{11}, x_{12}, ... , x_{1n}]^T$, $b_2 = [x_{21}, x_{22}, ... , x_{2n}]^T$, ... , $b_m = [x_{m1}, x_{m2}, ... , x_{mn}]^T$

i możemy jedynie poszukiwać optymalnego wektora $y_0 \in V_1 = E(b_1, b_2 , ... , b_m)$, tak żeby $d(y, y_0) = \min$. (tu występują „najmniejsze kwadraty”).

Inaczej, szukamy rzutu kolumny wyrazów wolnych y na V_1, jako p.p.w. wymiaru m (jeśli, co zakładamy, wektory b_k są liniowo niezależne).

Wówczas, ciąg szukanych współczynników c_k jest rozwiązaniem układu równań

$$<y, b_k> = c_1 <b_1, b_k> + c_2 <b_2, b_k> + ... + c_m <b_m, b_k>,$$

z nieosobliwą macierzą Gramma $G = <b_i, b_j>$ (patrz konstrukcja rzutu na p.p.w.).

P6. Rozwiążmy sprzeczny układ równań: $u + w = 1$, $u – w = 1$, $u = -1$, stosując metodę najmniejszych kwadratów, jako rzut wektora $y = [1, 1, -1]$ na 2-wymiarową (!) p.p.w. $V_1 = E(b_1, b_2)$ dla $b_1 = [1,1,1]$, $b_2 = [1, -1, 0]$). Zauważmy, że baza jest układem ortogonalnym, więc rzut $y_0 = c_1 [1,1,1] + c_2 [1, -1, 0]$, gdzie $c_1 = <y, b_1>/< b_1, b_1> = 1/3$ oraz $c_2 = <y, b_2>/< b_2, b_2> = 0$. Ostatecznie, mamy $y_0 = 1/3b_1 = [1/3, 1/3, 1/3]$, jako kolumnę wyrazów wolnych niesprzecznego już układu równań dobranego optymalnie do układu danego:

$u' + w' = 1/3$, $u' – w' = 1/3$, $u' = 1/3$ z optymalnym rozwiązaniem postaci:

$[u', w'] = [1/3, 0] = [c_1, c_2] \approx [u, w]$.