MATEMATICKÉ OBZORY 17/1981

PREDPOVEDANIE POMOCOU KOLMÍC V HILBERTOVÝCH PRIESTOROCH II

ANDREJ PÁZMAN, Bratislava

3. Podmienená pravdepodobnosť

Zo strednej školy vieme, že ak A, B sú dve náhodné udalosti, tak pravdepodobnosť toho, že nastane udalosť A za predpokladu, že nastala udalosť B, sa vypočíta podľa vzorca

$$p(A \mid B) = \frac{p(A \cap B)}{p(B)}, \tag{22}$$

ak $p(B) \neq 0$.

Podobne, ak ξ je náhodná veličina, ktorá nadobúda k hodnôt $x_1, \ldots, x_k$, a ak A_i je náhodná udalosť, že veličina ξ nadobudla hodnotu x_i, tak podmienená stredná hodnota náhodnej veličiny ξ, za predpokladu, že nastala udalosť B, sa vypočíta podľa vzorca

$$E(\xi \mid B) = \sum_{i=1}^{k} x_i \frac{p(A_i \cap B)}{p(B)} \tag{23}$$

Menej známe a trochu prekvapujúce je, že takáto podmienená pravdepodobnosť, resp. stredná hodnota, je vlastne projekcia na vhodne zvolený podpriestor v Hilbertovom priestore náhodných veličín s konečnými disperziami. To dokážeme v tomto paragrafe, ale len pre prípad, že množina $\mathcal{S}$ všetkých náhodných udalostí v danej situácii je konečná. Dôkaz vo všeobecnom prípade vyžaduje znalosti z teórie miery.

Zopakujme si niektoré pojmy z elementárnej pravdepodobnosti [4]. Množina $\mathcal{S}$ je algebra uzavretá na nasledujúce operácie: $A \cap B$, $A \cup B$, $A - B$. Pritom $A \cap B$ označuje udalosť „nastala udalosť A aj udalosť B", $A \cup B$ označuje udalosť „nastala udalosť A alebo udalosť B", $A - B$

označuje udalosť „nastala udalosť A, ale nenastala udalosť B". Nemožná udalosť, $\emptyset$, taktiež patrí do $\mathcal{S}$. Symbolom Ω označujeme istú udalosť (t. j. $\Omega = \cup\{A: A \in \mathcal{S}\}$), symbolom A^* opačnú udalosť k A (t. j. $A^* = \Omega - A$). Platí $\Omega \in \mathcal{S}$, $A \in \mathcal{S} \Rightarrow A^* \in \mathcal{S}$. Symbolom $A \subset B$ označujeme, že z udalosti A vyplýva udalosť B.

Elementárna udalosť je taká udalosť $C \in \mathcal{S}$, že platí: pre každé $A \in \mathcal{S}$ je alebo $A \cap C = C$, alebo $A \cap C = \emptyset$. Označme $\mathscr{C}$ množinu všetkých elementárnych udalostí patriacich do $\mathcal{S}$. Zrejme platí

$$
\begin{aligned}
&\text{a)}\ A = \cup\{C: C \in \mathscr{C}, C \subset A\}; \ (A \in \mathcal{S}), \\
&\text{b)}\ C_1 \cap C_2 = \emptyset; \ (C_1, C_2 \in \mathscr{C}, C_1 \neq C_2), \\
&\text{c)}\ \cup\{C: C \in \mathscr{C}\} = \Omega.
\end{aligned}
\tag{24}
$$

Pravdpodobnosť alebo *rozdelenie pravdepodobnosti* definujeme ako reálnu funkciu p na množine $\mathcal{S}$, ktorá má vlastnosti

$$
\begin{aligned}
&\text{a)}\ p(A) \in \langle 0, 1 \rangle; \ (A \in \mathcal{S}), \\
&\text{b)}\ p(\Omega) = 1, \\
&\text{c)}\ p(A \cup B) = p(A) + p(B), \quad \text{ak} \quad A \cap B = \emptyset,
\end{aligned}
$$

Z toho vyplývajú známe pravidlá:

$$p(A^*) = 1 - p(A)$$

$$p(\emptyset) = 0$$

$$p(A \cup B) \leq p(A) + p(B)$$

atď.

Konečná reálna funkcia definovaná na množine elementárnych udalostí $\mathscr{C}$ sa nazýva *náhodná veličina*.

Nech p je pravdepodobnosť. Dve náhodné veličiny ξ, η nazývame p-ekvivalentné, ak platí

$$C \in \mathscr{C}, p(C) > 0 \Rightarrow \xi(C) = \eta(C)$$

Označme $\mathscr{F}$ množinu všetkých náhodných veličín. Relácia p-ekvivalencie na $\mathscr{F}$ je reflexívna, symetrická a tranzitívna (ako sa môžete sami presvedčiť) a definuje rozklad množiny $\mathscr{F}$ na triedy p-ekvivalentných náhodných veličín. Označme symbolom $\{\xi\}$ triedu obsahujúcu náhodnú veličinu ξ. Náhodnú veličinu ξ nazveme *reprezentantom* triedy $\{\xi\}$.

Označme $\mathscr{F}_p$ množinu všetkých tried p-ekvivalencie. Na $\mathscr{F}_p$ možno definovať súčet vzťahom

$$\{\xi\}+\{\eta\}\equiv\{\xi+\eta\}$$

a násobenie skalárom $\alpha \in T$

$$\alpha\{\xi\}\equiv\{\alpha\xi\}.$$

Presvedčte sa sami, že definície sú korektné, t. j., že nezávisia od výberu reprezentanta z triedy. Takisto sa presvedčte, že množina $\mathscr{F}_p$ s takýmito operáciami je lineárny priestor (opäť označovaný $\mathscr{F}_p$).

Ak $A \in \mathscr{S}$, symbolom 1_A označme náhodnú veličinu definovanú vzťahmi

$$\begin{aligned} 1_A(C)&=1, \quad \text{ak} \quad C\subset A \\ 1_A(C)&=0, \quad \text{ak} \quad C\not\subset A \end{aligned} \tag{25}$$

Tvrdenie 4. Lineárny priestor $\mathscr{F}_p$ so skalárnym súčinom

$$\langle\{\xi\},\{\eta\}\rangle=\sum_{C\in\mathscr{C}}\xi(C)\eta(C)p(C) \tag{26}$$

je Hilbertov priestor. Stredná hodnota náhodnej veličiny $\xi\in\mathscr{F}$ je

$$E_p(\xi)=\langle\{\xi\},\{1_\Omega\}\rangle \tag{27}$$

a jej disperzia je

$$D_p(\xi)=\|\{\xi\}\|^2-(\langle\{\xi\},\{1_\Omega\}\rangle)^2 \tag{28}$$

Dôkaz. Ak $\xi_1\in\{\xi_2\}$, $\eta_1\in\{\eta_2\}$, zrejme platí

$$\sum_{C\in\mathscr{C}}\xi_1(C)\eta_1(C)p(C)=\sum_{C\in\mathscr{C}}\xi_2(C)\eta_2(C)p(C)$$

teda skalárny súčin je vzťahom (26) určený jednoznačne. Overenie vlastností a, b, c skalárneho súčinu je elementárne. Úplnosť dokážeme podobne ako v príklade 4.

Platí

$$E_p(\xi)=\sum_{C\in\mathscr{C}}\xi(C)p(C)=\langle\{\xi\},\{1_\Omega\}\rangle$$

$$\begin{aligned} D_p(\xi)&\equiv E_p[(\xi-E_p(\xi))^2]=E_p(\xi^2)-E_p^2(\xi)= \\ &=\|\{\xi\}\|^2-[\langle\{\xi\},\{1_\Omega\}\rangle]^2. \quad \square \end{aligned}$$

Poznámka. Ak predpokladáme, že $p(C) > 0$ pre každú elementárnu udalosť $C \in \mathscr{C}$, tak každá trieda $\{\xi\}$ obsahuje jedinú náhodnú veličinu ξ. Kvôli zjednodušeniu symboliky prijmeme tento predpoklad v ďalších úvahách.

Nech $\mathscr{S}_0 \subset \mathscr{S}$ je podalgebra algebry náhodných udalostí $\mathscr{S}$. (To znamená $\emptyset \in \mathscr{S}_0$, $\Omega \in \mathscr{S}_0$; $A, B \in \mathscr{S}_0 \Rightarrow A \cup B \in \mathscr{S}_0$ $A - B \in \mathscr{S}_0$). Označme $\mathscr{C}_0$ množinu elementárnych udalostí v $\mathscr{S}_0$. Náhodnú veličinu definovaná na $\mathscr{C}$, ktorá je konštantná na množinách patriacich do $\mathscr{C}_0$ (t. j. $C_1 \subset C$, $C_2 \subset C$, $C \in \mathscr{C}_0$, $C_1, C_2 \in \mathscr{C} \Rightarrow \xi(C_1) = \xi(C_2)$) nazývame $\mathscr{S}_0$-merateľnou.

Cvičenie 5. Dokážte, že množina M_0 všetkých $\mathscr{S}_0$-merateľných funkcií je podpriestor Hilbertovho priestoru $\mathscr{F}_p$.

Označme symbolom P_0 projekciu na M_0 v Hilbertovom priestore $\mathscr{F}_p$. Je zrejmé (z definície M_0), že $P_0\xi$ je $\mathscr{S}_0$-merateľná funkcia pre každé $\xi \in \mathscr{F}_p$.

Projekcia P_0 je určená podmienenými strednými hodnotami náhodných veličín, tak ako je to ukázané v nasledujúcom tvrdení.

Tvrdenie 5. Nech $\xi \in \mathscr{F}_p$. Pre každé $C \in \mathscr{C}$ platí

$$(P_0\xi)(C) = E_p(\xi \mid C_0) \tag{29}$$

kde C_0 je tá udalosť patriaca do $\mathscr{C}_0$, ktorá obsahuje C (t. j. $C \subset C_0$), a kde podmienená stredná hodnota je (porovnaj (23))

$$E_p(\xi \mid C_0) = \sum_{\substack{C \subset C_0 \\ C \in \mathscr{C}}} \xi(C) \frac{p(C)}{p(C_0)} \tag{30}$$

Dôkaz. Označme ζ tú $\mathscr{S}_0$-merateľnú náhodnú veličinu, pre ktorú platí

$$\zeta(C) = E_p(\xi \mid C_0)$$

pre každé $C \in \mathscr{C}$ a pre $C_0 \supset C$, $C_0 \in \mathscr{C}_0$. Pre ľubovoľné $\eta \in M_0$ vypočítajme skalárny súčin

$$\langle \xi - \zeta, \eta \rangle = \sum_{C_0 \in \mathscr{C}_0} \left[\sum_{C \subset C_0} (\xi(C) - E_p(\xi \mid C_0))\eta(C)p(C) \right] \tag{31}$$

Zvoľme $C^{(0)} \in \mathscr{C}$, $C^{(0)} \subset C_0$. Keďže η *je* $\mathscr{S}_0$-merateľná, platí $\eta(C) = \eta(C^{(0)})$ pre každé $C \subset C_0$. Môžeme teda písať na základe (30)

$$\sum_{\substack{C \subset C_0 \\ C \in \mathscr{C}}} \xi(C)\eta(C)p(C) = \eta(C^{(0)})p(C_0)E_p(\xi \mid C_0)$$

Po dosadení do (31) dostávame

$$\langle \xi - \zeta, \eta \rangle = 0$$

Teda z Pytagorovej vety (cvičenie 2) vyplýva, že pre každé $\varkappa \in M_0$ je

$$\|\xi - \varkappa\|^2 = \|\xi - \zeta\|^2 + \|\zeta - \varkappa\|^2$$

Teda

$$\|\xi - \zeta\|^2 = \min \{\|\xi - \varkappa\|^2 : \varkappa \in M_0\}$$

čo značí (rovnosť (11)), že

$$P_0\xi = \zeta. \quad \square$$

Podobne aj podmienenú pravdepodobnosť počítame pomocou projekcie. Skutočne, nech $A \subset \mathscr{S}$ je udalosť. Potom platí

$$E_p(1_A \mid C_0) = \sum_{\substack{C \in \mathscr{C} \\ C \subset C_0}} 1_A(C) \frac{p(C)}{p(C_0)}$$

$$= p(A \mid C_0)$$

Teda

$$(P_0 1_A)(C) = p(A \mid C_0) \tag{32}$$

pre každé $C \subset \mathscr{C}$ a pre $C_0 \supset C$, $C_0 \in \mathscr{C}_0$.

Vráťme sa teraz k príkladu 1. Za množinu všetkých elementárnych udalostí $\mathscr{C}$ zvolíme množinu všetkých možných dráh zajaca od prvého do $(n+i)$-tého skoku. Takýchto dráh je 2^{n+i}. Udalosťou je napríklad niektorá možná poloha zajaca po $(n+i)$-tom skoku. Skladá sa z tých možných dráh zajaca (z tých elementárnych udalostí), ktoré sa končia v uvažovanej polohe.

Množinu $\mathscr{C}_0$ utvorme takto. Každá $C_0 \in \mathscr{C}_0$ je zostavená z tých dráh (z tých elementárnych udalostí), ktoré sa zhodujú na 1-vom až n-tom skoku. Teda udalosti $C_0 \in \mathscr{C}_0$ môžeme vzájomne jednoznačne priradiť dráhu zajaca od 1-vého do n-tého skoku.

Množinou $\mathscr{C}$ je určená algebra udalostí $\mathscr{S}$. Pravdepodobnosť p na $\mathscr{C}$ (a tým aj na $\mathscr{S}$) určíme z pravidla o pravdepodobnosti zmeny skoku uvedeného v príklade 1. Je zrejmé, že $p(C) > 0$ pre každú dráhu $C \in \mathscr{C}$.

Tak napríklad dráha, v ktorej zajac zmení smer pri každom skoku má pravdepodobnosť

$$\left(\frac{1}{2}\right)^{n+i}$$

Množinou $\mathscr{C}_0$ je stanovená algebra $\mathscr{S}_0$, a teda aj podpriestor M_0, v súlade s cvičením 5. V súlade s tvrdením 5, z projekcie P_0 na M_0 môžeme určiť všetky podmienené pravdepodobnosti a podmienené stredné hodnoty za podmienky, že je známa dráha zajaca v 1-vom až n-tom skoku (t. j., že je daná udalosť $C_0 \in \mathscr{C}_0$). Špeciálne je takto určená podmienená pravdepodobnosť rôznych polôh zajaca po $(n+i)$-tom skoku, čo je predpoveď požadovaná v príklade 1.

4. Odhady parametrov

V príklade s kľučkujúcim zajacom predpoveď jeho polohy sa zakladá na určení podmienenej pravdepodobnosti. Bolo to preto, lebo pravdepodobnosťami bolo opísané celé správanie sa zajaca.

Naproti tomu v príklade 2 informácie o páde kameňa nám v prvom rade dáva fyzika, v súlade s ktorou poloha kameňa $(x(t),\ y(t))$ v čase t, pri zanedbaní odporu vzduchu, je

$$\begin{aligned} x(t) &= vt \\ y(t) &= h - \frac{1}{2}gt^2 \end{aligned} \tag{33}$$

Predpovedať miesto dopadu kameňa znamená teda odhadnúť neznáme parametre v a g. Nech $(\xi(t_i),\ \eta(t_i))$ je približne určená poloha kameňa v čase t_i; $(i = 1, \ldots, 5)$. Hodnoty $\xi(t_i)$, $\eta(t_i)$ sú náhodné vďaka náhodným chybám pri určovaní polohy kameňa počas jeho letu. Nájsť postup na výpočet odhadov parametrov v a g znamená vyjadriť tieto odhady ako funkcie pozorovaných náhodných veličín $\xi(t_1), \ldots, \xi(t_5), \eta(t_1), \ldots, \eta(t_5)$. Takéto funkcie sa spravidla určujú riešením rovnice

$$\sum_{i=1}^{5}\left[(\xi(t_i) - v(\xi,\eta)t_i)^2 + (\eta(t_i) - \frac{1}{2}g(\xi,\eta)t_i^2)^2\right] = \tag{34}$$

$$= \min_{v, g} \sum_{i=1}^{5} \left[(\xi(t_i) - vt_i)^2 + (\eta(t_i) - \frac{1}{2} g t_i^2)^2 \right]$$

(metóda najmenších štvorcov).

Porovnanie rovnice (34) s rovnosťou (12) ukazuje, že opäť vzniká úloha projekcie na podpriestor Hilbertovho priestoru. Že je to skutočne tak, ukážeme vo všeobecnejšie formulovanej úlohe.

Nech $\xi_1, \ldots, \xi_n$ sú náhodné veličiny, ktoré

a) sú nekorelované, t. j.

$$E[(\xi_i - E(\xi_i))(\xi_j - E(\xi_j))] = 0; \qquad (i \neq j)$$

b) majú disperzie úmerné zadaným číslom $q_1, \ldots, q_n$, t. j.

$$E[\xi_i - E(\xi_i)]^2 = kq_i$$

c) majú stredné hodnoty tvaru

$$E(\xi_i) = \sum_{j=1}^{m} \alpha_j F_{ij}; \qquad (i = 1, \ldots, n) \tag{35}$$

kde $m \leq n$ a kde $\alpha_1, \ldots, \alpha_m$ sú neznáme (reálne) parametre ($\boldsymbol{\alpha}' \equiv (\alpha_1, \ldots, \alpha_m) \in R^m$) a matica $\mathbf{F}$ koeficientov F_{ij} je známa. Predpokladáme, že hodnosť (počet lineárne nezávislých stĺpcov) matice $\mathbf{F}$ rovná sa m.

Ak označíme $\boldsymbol{\xi} \equiv (\xi_1, \ldots, \xi_n)$ vektor pozorovaných náhodných veličín, tak rovnosti (35) môžeme zapísať v tvare

$$E(\boldsymbol{\xi}) = \mathbf{F}\boldsymbol{\alpha}$$

Funkcie $\alpha_1(\boldsymbol{\xi}), \ldots, \alpha_m(\boldsymbol{\xi})$ premenných $\xi_1, \ldots, \xi_n$, ktorými odhadujeme parametre $\alpha_1, \ldots, \alpha_m$, určujeme riešením rovnice

$$\begin{aligned} &\sum_{i=1}^{n} \frac{1}{q_i} \left(\xi_i - \sum_{j=1}^{m} F_{ij} \alpha_j(\boldsymbol{\xi}) \right)^2 = \\ &= \min_{\alpha \in R^m} \sum_{i=1}^{n} \frac{1}{q_i} \left(\xi_i - \sum_{j=1}^{m} F_{ij} \alpha_j \right)^2 \end{aligned} \tag{36}$$

Označme H_q Hilbertov priestor $R^{n'}$ so skalárnym súčinom

$$\langle \mathbf{x}, \mathbf{y} \rangle_q \equiv \sum_{i=1}^{n} \frac{x_i y_i}{q_i}$$

(porovnaj príklady 4 a 5) a označme M jeho podpriestor

$$M \equiv \{\mathbf{x} \in R^n : \mathbf{x} = \mathbf{F}\boldsymbol{\alpha};\ (\boldsymbol{\alpha} \in R^m)\} \tag{37}$$

Rovnicu (36) môžeme prepísať do tvaru

$$\|\boldsymbol{\xi} - \mathbf{F}\boldsymbol{\alpha}(\boldsymbol{\xi})\|_q^2 = \min_{\alpha \in R^m} \|\boldsymbol{\xi} - \mathbf{F}\boldsymbol{\alpha}\|_q^2 \tag{38}$$

Tvrdenie 6. Rovnice (36) resp. (38) majú práve jedno riešenie $\boldsymbol{\alpha}(\boldsymbol{\xi})$. Projekcia P na podpriestor M v Hilbertovom priestore H_q má tvar

$$P\mathbf{y} = \mathbf{F}\boldsymbol{\alpha}(\mathbf{y});\ (\mathbf{y} \in H_q).$$

Dôkaz vyplýva bezprostredne z tvrdenia 2 a z definície projekcie. Jednoznačnosť riešenia $\boldsymbol{\alpha}(\mathbf{y})$ je dôsledkom jednoznačnosti prvku $P\mathbf{y}$ (tvrdenie 2) a toho, že matica $\mathbf{F}$ má hodnosť m. □.

Riešenie rovníc (36), resp. (38) možno nájsť aj v explicitnom tvare (a tým vlastne explicitne riešiť úlohu hľadania projekcie na M). Pritom stačí použiť štandardný postup hľadania extrému funkcie m premenných:

a) Vypočítame riešenie rovníc

$$\frac{\partial}{\partial \alpha_j} \|\boldsymbol{\xi} - \mathbf{F}\boldsymbol{\alpha}\|_q^2 = 0; \qquad (j = 1, \ldots, m)$$

t. j. rovníc

$$\sum_{i=1}^{n} \frac{1}{q_i}\left(\xi_i - \sum_{k=1}^{m} F_{ik}\alpha_k\right)F_{ij} = 0; \qquad (j = 1, \ldots, m)$$

Usilovný čitateľ sa môže presvedčiť, že riešenie je v tvare

$$\boldsymbol{\alpha}(\boldsymbol{\xi}) = (\mathbf{F}'\mathbf{W}\mathbf{F})^{-1}\mathbf{F}'\mathbf{Q}\boldsymbol{\xi} \tag{39}$$

kde

$$\mathbf{W} = \begin{pmatrix} q_1^{-1} & & 0 \\ & \ddots & \\ 0 & & q_m^{-1} \end{pmatrix}, \qquad \mathbf{Q} = \begin{pmatrix} q_1 & & 0 \\ & \ddots & \\ 0 & & q_m \end{pmatrix}$$

b) Vypočítame maticu druhých derivácií $\mathbf{V}$

$$\{V\}_{ij} \equiv \frac{\partial^2}{\partial \alpha_i\, \partial \alpha_j} \|\boldsymbol{\xi} - \mathbf{F}\boldsymbol{\alpha}\|_q^2$$

a presvedčíme sa, že matica $\mathbf{V}$ je pozitívne definitná. Teda riešenie (39) skutočne zodpovedá minimu formy $\|\xi - \mathbf{F}\alpha\|_q^2$.

Cvičenie 6. Pomocou tvrdenia 3 dokážte, že vzťahom (39) je skutočne určená projekcia, t. j. že zobrazenie

$$P: \mathbf{x} \in H_q \mapsto (\mathbf{F}'\mathbf{W}\mathbf{F})^{-1}\mathbf{F}'\mathbf{Q}\mathbf{x} \in M$$

má vlastnosti a, b, c, d uvedené pred tvrdením 3.

Tým sme sa presvedčili, že aj v tomto prípade odhadovania parametrov, a tým aj predpovedanie polohy kameňa treba vykonať pomocou projekcie v Hilbertovom priestore.

Literatúra

[1] Gedeonová, E.: Vektory očami algebraika. Matematické obzory, 8, str. 11—31.

[2] Kluvánek, I.—Mišík, L.—Švec, M.: Matematika I. SVTL Bratislava 1959.

[3] Kufner, A.: Geometrie Hilbertova prostoru. SNTL Praha 1973.

[4] Riečan, B.: O pravdepodobnosti a miere. Alfa Bratislava 1972.