O pewnym zastosowaniu analizy harmonicznej w rachunku prawdopodobieństwa

Krzysztof OLESZKIEWICZ, Warszawa

Jest to zapis odczytu wygłoszonego na XXVI Szkole Matematyki Poglądowej: *Twierdzenia z pogranicza.*

Związki analizy harmonicznej z rachunkiem prawdopodobieństwa dotyczą na ogół własności przekształcenia Fouriera i szeregów trygonometrycznych. Tym razem rozważymy jednak jej inne, znacznie bardziej elementarne zastosowanie probabilistyczne. Wśród wszystkich funkcji o wartościach rzeczywistych określonych na kostce dyskretnej $\{-1,1\}^n$ (czyli na zbiorze 2^n wierzchołków n-wymiarowej kostki $[-1,1]^n$) wyróżnijmy rodzinę 2^n funkcji $(w_A)_{A\subset\{1,2,\ldots,n\}}$, zwaną układem Walsha. Funkcja w_A przyjmuje w punkcie $x=(x_1,x_2,\ldots,x_n)$ wartość będącą iloczynem tych współrzędnych x_i, których indeks i należy do A. Innymi słowy

$$w_A(x)=\prod_{i\in A}x_i,$$

przy czym przyjmujemy, że $w_\emptyset$ przyjmuje na całym zbiorze $\{-1,1\}^n$ wartość 1. Jeśli zdefiniujemy iloczyn skalarny funkcji $f,g:\{-1,1\}^n\longrightarrow R$ wzorem

$$\langle f,g\rangle=\frac{1}{2^n}\sum_{x\in\{-1,1\}^n}f(x)g(x),$$

to łatwo będzie sprawdzić, że $\langle w_A,w_B\rangle=0$, jeśli $A\neq B$, oraz $\langle w_A,w_A\rangle=1$ (dowód pozostawiamy Czytelnikowi jako proste ćwiczenie). Wynika stąd, że układ Walsha jest układem ortonormalnym względem wprowadzonego iloczynu skalarnego. Zatem każda funkcja $f:\{-1,1\}^n\longrightarrow R$ daje się w co najwyżej jeden sposób zapisać jako kombinacja liniowa $\sum_{A\subset\{1,2,\ldots,n\}}a_Aw_A$, $(a_A\in R)$, bo wówczas

$$\langle f,w_B\rangle=\sum_{A\subset\{1,2,\ldots,n\}}a_A\langle w_A,w_B\rangle=a_B,$$

czyli z funkcji f można odczytać jej współczynniki a_A. Z drugiej strony każdą funkcję $f:\{-1,1\}^n\longrightarrow R$ można zapisać w postaci takiej kombinacji. Można to uzasadnić na dwa sposoby. Po pierwsze, można powołać się na fakt, że wymiar liniowy przestrzeni wszystkich funkcji rzeczywistych na $\{-1,1\}^n$ (nad R) jest równy 2^n i w związku z tym układ Walsha, jako układ 2^n funkcji niezależnych liniowo (bo ortogonalnych) jest bazą, a więc także rozpina liniowo tę przestrzeń. Po drugie, można uniknąć odwoływania się do zaawansowanej algebry liniowej – oczywiście (dlaczego?) każdą funkcję rzeczywistą na $\{-1,1\}^n$ można zapisać jako kombinację liniową funkcji $(f_y)_{y\in\{-1,1\}^n}$, przyjmujących wartości $f_y(y)=1$; $f_y(x)=0$ dla $x\neq y$. Zatem wystarczy przedstawić każdą z funkcji f_y jako kombinację liniową funkcji Walsha, to zaś jest proste:

$$f_y(x)=\prod_{i=1}^{n}\frac{x_i+y_i}{2}=\frac{1}{2^n}\sum_{A\subset\{1,2,\ldots,n\}}\Big(\prod_{i\in\{1,2,\ldots,n\}\setminus A}y_i\Big)\cdot w_A(x)$$

(zauważmy, że iloczyn jest zerem, jeśli którykolwiek z jego czynników jest zerem).

Funkcji $f:\{-1,1\}^n\longrightarrow R$ przyporządkujmy nową funkcję $Lf:\{-1,1\}^n\longrightarrow R$ wzorem

$$(Lf)(x)=\sum_{y\in\{-1,1\}^n:d(x,y)=1}f(y),$$

gdzie $d(x,y)$ oznacza liczbę współrzędnych, którymi różnią się punkty $x,y\in\{-1,1\}^n$, tzn. liczbę tych $1\le i\le n$, dla których $x_i\neq y_i$ (jest to tzw. metryka Hamminga). Innymi słowy wartość funkcji Lf w punkcie x definiujemy jako sumę wartości funkcji f we wszystkich wierzchołkach sąsiadujących z x. Łatwo zauważyć, że operator L jest liniowy, czyli

$$L\Big(\sum_{k=1}^{m}\alpha_kf_k\Big)=\sum_{k=1}^{m}\alpha_k\cdot Lf_k;\quad \alpha_k\in R.$$

Pozostawmy Czytelnikowi sprawdzenie, że $Lw_A = (n - 2|A|) \cdot w_A$, gdzie $|A|$ oznacza liczbę elementów zbioru A (co w języku algebry liniowej oznacza, że funkcje Walsha są wektorami własnymi operatora L).

Czas na oczekiwane zastosowanie w rachunku prawdopodobieństwa. Rozważmy cząstkę błądzącą losowo po wierzchołkach n–wymiarowej kostki. Błądzenie zaczyna się w chwili $t = 0$ w wierzchołku $(1, 1, \ldots, 1)$, a potem błądzenie odbywa się według następujących reguł: jeśli cząstka w chwili t znajduje się w punkcie y, to w chwili $t + 1$ z prawdopodobieństwem $1/n$ będzie w wierzchołku z sąsiadującym z y (ponieważ każdy wierzchołek ma dokładnie n sąsiadów, znaczy to, że do każdego z nich cząstka przejdzie z jednakowym prawdopodobieństwem). Spróbujmy odpowiedzieć na pytanie, z jakim prawdopodobieństwem cząstka w chwili t będzie się znajdowała w punkcie $x \in \{-1, 1\}^n$.

Oznaczmy to prawdopodobieństwo przez $p_t(x)$. Z zasad błądzenia wynika, że $p_{t+1} = Lp_t/n$, a zatem $p_t = n^{-t} \cdot L^t p_0$, gdzie L^t oznacza t–krotne zastosowanie operatora L. Skoro na mocy założeń

$$p_0 = f_{(1,1,\ldots,1)} = \frac{1}{2^n} \sum_{A \subset \{1,2,\ldots,n\}} w_A,$$

z liniowości operatora L wynika, że

$$L^t p_0 = \frac{1}{2^n} \sum_{A \subset \{1,2,\ldots n\}} (n - 2|A|)^t \cdot w_A.$$

Otrzymujemy zatem jawny wzór

$$p_t(x) = \frac{1}{2^n} \sum_{A \subset \{1,2,\ldots n\}} \left(1 - \frac{2|A|}{n}\right)^t w_A(x).$$

Ponieważ $|A| \in \{0, 1, \ldots n\}$, więc $|1 - \frac{2|A|}{n}| < 1$, o ile $A \neq \emptyset$ i $A \neq \{1, 2, \ldots, n\}$. Wynika stąd, że dla dużych wartości czasu t prawdopodobieństwo $p_t(x)$ można z bardzo dużą dokładnością przybliżyć przez $\frac{1}{2^{n-1}}$, jeśli liczba współrzędnych, którymi x różni się od $(1, 1, \ldots, 1)$ ma inną parzystość niż t, przez 0 zaś, jeśli ma tę samą parzystość.

Pozostawmy Dociekliwemu Czytelnikowi przemyślenie, jak zmieni się sytuacja, jeśli błądząca cząstka będzie z prawdopodobieństwem $\frac{1}{n+1}$ przechodzić do sąsiedniego wierzchołka, a z prawdopodobieństwem $\frac{1}{n+1}$ – pozostawać w tym samym wierzchołku.