Размерность Вапника–Червоненкиса

ДА 28.4.2018

Пусть $\mathcal{R} \subset \mathrm{Subsets}(X)$ — семейство подмножеств произвольного множества X. Множество $A \subset X$ называется *дробящимся* системой $\mathcal{R}$, если пересечения A с множествами из $\mathcal{R}$ образуют все подмножества A. *Размерностью Вапника–Червоненкиса* $\mathrm{VC}(X, \mathcal{R})$ (или VC-размерностью) пары $(X, \mathcal{R})$ называется размер максимального (по мощности) подмножества $A \subset X$, дробящегося $\mathcal{R}$. Если максимального подмножества нет, то полагают $\mathrm{VC}(X, \mathcal{R}) := \infty$.

Теорема Радона. Любые $n+2$ точки в $\mathbb{R}^n$ можно разбить на два множества, выпуклые оболочки которых пересекаются.

1. (a) С помощью теоремы Радона докажите, что VC-размерность семейства всех полупространств в $\mathbb{R}^n$ равна $n+1$.
(b) Докажите теорему Радона.

2. Возможно ли равенство $\mathrm{VC}(\mathbb{R}^2, \mathcal{R}) = \infty$ для некоторого набора $\mathcal{R} \subset 2^{\mathbb{R}^2} (= \mathrm{Subsets}(\mathbb{R}^2))$?

3. Докажите, что в любом семействе VC-размерности d, в каждом множестве которого не более r элементов, найдутся такие подмножества X и Y, что

(a) $|X \cap Y| \leqslant r - d$;
(b) $|X \cap Y| \geqslant d - 1$.

4. Докажите, что если $\mathcal{R} \subset 2^{[n]}$, то $|\mathcal{R}| \leqslant C_n^0 + C_n^1 + C_n^2 + \ldots + C_n^{\mathrm{VC}([n], \mathcal{R})}$.

Домашнее задание.

1. Докажите, что если $\mathcal{R} \subset 2^{[n]}$ и $|\mathcal{R}| = n$, то для любого $k = 1, 2, \ldots, n$ найдётся такое множество A, что $|\{R \cap A \mid R \in \mathcal{R}\}| \geqslant k = |A| + 1$.

2. Докажите, что если $\mathcal{R} \subset 2^{[n]}$ — семейство VC-размерности d, то существует *наследственное* (т. е. содержащее с каждым множеством все его подмножества) семейство $\mathcal{R}' \subset 2^{[n]}$ VC-размерности d, для которого

(a) $|\mathcal{R}'| \leqslant |\mathcal{R}|$;
(b) $|\mathcal{R}'| \geqslant |\mathcal{R}|$.