Tight lower bounds for the size of epsilon-nets

János Pach, Gábor Tardos

Introduction

Couple $(X, \mathcal{R})$ is called *range space* in universe $\mathcal{U}$ if $X \subset \mathcal{U}$ is a finite set and $\mathcal{R} \subset 2^{\mathcal{U}}$ is a system of sets. The set $A \subset X$ is called *shattered* if for every $B \subset A$ there is a set $R_B \in \mathcal{R}$ such that $R_B \cap A = B$. The size of the largest shattered subset of X is called the *dimension* of the range space $(X, \mathcal{R})$.

For every $\varepsilon > 0$, the set $S \subset X$ is called the *ε-net* for the range space $(X, \mathcal{R})$ if every range $R \in \mathcal{R}$ with $|R \cap X| \geq \varepsilon |X|$ contains at least one element of S.

Theorem (Matoušek, Seidl, Welzl, 1990-2) All range spaces $(X, \mathcal{R})$, where X is a finite set of points in $\mathbb{R}^3$ and $\mathcal{R}$ consists of half-spaces, admit ε-nets of size $O(\frac{1}{\varepsilon})$.

Theorem (Aronov, Ezra, Sharir, 2010) All range spaces $(X, \mathcal{R})$, where X is a finite set of points in $\mathbb{R}^2$ (or $\mathbb{R}^3$) and $\mathcal{R}$ consists of axis-parallel rectangles (boxes), admit ε-nets of size $O(\frac{1}{\varepsilon} \log \log \frac{1}{\varepsilon})$.

Let $(X, \mathcal{R})$ be a range space with ranges from $\mathbb{R}^m$. The *dual range space* $(\mathbb{R}, \cup_{x \in \mathbb{R}^m} \mathcal{R}_x)$ is defined as a system on the underlying set $\mathcal{R}$ consisting of the sets $\mathcal{R}_x = \{R \mid x \in R \in \mathcal{R}\}$, for all $x \in \mathbb{R}^m$.

Main results

Theorem 1 For any $\varepsilon > 0$ and for any sufficiently large integer $n \geq n_0(\varepsilon)$, there exists a dual range space $\sum^*$ of VC-dimension 2, induced by n axis-parallel rectangles in $\mathbb{R}^2$, in which the minimum size of an ε-net is at least $C\frac{1}{\varepsilon} \log \frac{1}{\varepsilon}$. Here $C > 0$ is an absolute constant.

Theorem 2 For any $\varepsilon > 0$ and for any sufficiently large integer $n \geq n_0(\varepsilon)$, there exists a (primal) range space $\sum = (X, \mathcal{R})$ of VC-dimension 2, where X is a set of n points of $\mathbb{R}^4$, $\mathcal{R}$ consists of axis-parallel boxes with one of their vertices at the origin, and in which the size of the smallest ε-net is at least $C\frac{1}{\varepsilon} \log \frac{1}{\varepsilon}$. Here $C > 0$ is an absolute constant.

Theorem 3 For any $\varepsilon > 0$ and for any sufficiently large integer $n \geq n_0(\varepsilon)$, there exists a (primal) range space $\sum = (X, \mathcal{R})$ of VC-dimension 2, where X is a set of n points of $\mathbb{R}^4$, $\mathcal{R}$ consists of half-spaces, and in which the size of the smallest ε-net is at least $C\frac{1}{\varepsilon} \log \frac{1}{\varepsilon}$. Here $C > 0$ is an absolute constant.

Theorem 4 For any $\varepsilon > 0$ and for any sufficiently large integer $n \geq n_0(\varepsilon)$, there exists a (primal) range space $\sum = (X, \mathcal{R})$, where X is a set of n points in the plane, $\mathcal{R}$ consists of axis-parallel rectangles, and in which the size of the smallest ε-net is at least $C\frac{1}{\varepsilon} \log \log \frac{1}{\varepsilon}$. Here $C > 0$ is an absolute constant.

The structure of the proofs in the paper is the following:

$$\left.\begin{array}{l}\text{Lemma 1}\\ \text{Lemma 2}\end{array}\right\} \Longrightarrow \text{Theorem 1} \Longrightarrow \left.\begin{array}{l}\text{Theorem 2}\\ \\ \text{Lemma 3}\end{array}\right\} \Longrightarrow \text{Theorem 3}$$

$$\text{Lemma 4} \Longrightarrow \text{Theorem 4}$$

Useful tools

Let $c \geq 2$ and $d \geq 1$ be integers. Let $x \in [c]^k = \{0, 1, \ldots, c-1\}^k$, that is $x = x_1 x_2 \ldots x_k$, $k \in [d]$. Expanding x as a c-ary fraction we define $\bar{x} = \sum_{j=1}^{k} x_j / c^k$. For any $0 \leq k \leq d$, $u \in [c]^k$ and $v \in [c]^{d-k}$ we define an open axis-parallel rectangle $R^k_{u,v}$ in the unit square as

$$R^k_{u,v} = (\bar{u}, \bar{u} + c^{-k}) \times (\bar{v}, \bar{v} + c^{k-d})$$

and consider the family

$$\mathcal{R} = \mathcal{R}(c, d) = \{R^k_{u,v} \mid 0 \leq k \leq d, u \in [c]^k, v \in [c]^{d-k}, u_k = v_{d-k}\}$$

Clearly $|\mathcal{R}| = (d+1)c^{d-1}$. Finally $\sum = \sum(c, d)$ be the infinite range space $(\mathbb{R}^2, \mathcal{R})$ and let $\sum^* = \sum^*(c, d)$ denote its dual range space.

Lemma 1 Let $d \geq 1$, $r \geq 2$, $c \geq 3$ and let $\sum^* = \sum^*(c, d)$. If a subset $I \subset \mathcal{R}(c, d)$ contains no r-element range of $\sum^*$ then

$$|I| \leq (r-1)\frac{c-1}{c-2}c^{d-1}.$$

Lemma 2 Both $\sum$ and $\sum^*$ have VC-dimension 2.

Lemma 3 Let P be a finite set of points in the positive orthant of $\mathbb{R}^d$. To each $p \in P$ we can assign a point p' in the positive orthant of $\mathbb{R}^d$ so that the set $P' = \{p' \mid p \in P\}$ satisfies the following condition. For any axis-parallel box $B \subset \mathbb{R}^d$ that contains the origin, there is a half-space $H_B \subset \mathbb{R}^d$ which contains the origin and for which $\{p' \mid p \in B \cap P\} = P' \cap H_B$.

Lemma 4 Let $n \geq 2$, $r = \lceil \log\log n / 5 \rceil$ be integers and $\varepsilon = r/n$. Let X be a set of n randomly uniformly selected points of unit square, and $\mathcal{R}$ denote the family of all axis-parallel rectngles of the form $[j/2^t, (j+1)/2^t) \times [a, b]$, where $j, t \in \mathbb{N}_0$ and $a < b$ are reals. Then, with probability tending to 1, the range space $(X, \mathcal{R})$ does not admit an ε-net of size at most $n/2$.