VC-density in model theoretic structures

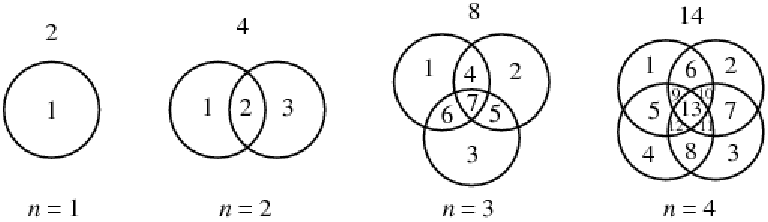
Anton Bobkov

June 3, 2015

Suppose we have an (infinite) collection of sets $\mathcal{F}$.
We define a shatter function $\pi_{\mathcal{F}}(n)$

$$\pi_{\mathcal{F}}(n) = \max\{\#\text{ of atoms in boolean algebra generated by } S \mid S \subset \mathcal{F} \text{ with } |S| = n\}$$

Example: Let $\mathcal{F}$ consist of all discs on a plane.



$$\pi_{\mathcal{F}}(1) = 2 \quad \pi_{\mathcal{F}}(2) = 4 \quad \pi_{\mathcal{F}}(3) = 8 \quad \pi_{\mathcal{F}}(4) = 14$$

$$\pi_{\mathcal{F}}(n) = n^2 - n + 2$$

More examples:

1. Lines on a plane $\pi_{\mathcal{F}}(n) = n^2/2 + n/2 + 1$
2. Disks on a plane $\pi_{\mathcal{F}}(n) = n^2 - n + 2$
3. Balls in $\mathbb{R}^3$ $\pi_{\mathcal{F}}(n) = n^3/3 - n^2 + 8n/3$
4. Intervals on a line $\pi_{\mathcal{F}}(n) = 2n$
5. Half-planes on a plane $\pi_{\mathcal{F}}(n) = n(n+1)/2 + 1$
6. Finite subsets of $\mathbb{N}$ $\pi_{\mathcal{F}}(n) = 2^n$
7. Polygons in a plane $\pi_{\mathcal{F}}(n) = 2^n$

Theorem (Sauer-Shelah '72)

Shatter function is either 2^n or bounded by a polynomial.

Definition

Suppose growth of shatter function for $\mathcal{F}$ is polynomial. Let r be the smallest real such that

$$\pi_{\mathcal{F}}(n) = O(n^r)$$

We define such r to be the vc-density of $\mathcal{F}$, $\mathrm{vc}(\mathcal{F})$. If shatter function grows exponentially, we let the vc-density to be infinite.

Applications

- NIP theories
- VC-Theorem in probability (VapnikChervonenkis 1971)
- Computational learning theory (PAC learning)
- Computational geometry
- Functional analysis (Bourgain-Fremlin-Talagrand theory)
- Abstract topological dynamics (tame dynamical systems)

History

- VapnikChervonenkis 1971 - introduce VC-dimension
- NIP theories (Shelah 1971, 1990)
- vc-density (Aschenbrenner, Dolich, Haskell, Macpherson, Starchenko '13)

Model Theory

Model Theory studies definable sets in first-order structures.

$$(\mathbb{Q}, 0, 1, +, \cdot, \leq)$$

$$\phi(x) = \exists y \; y \cdot y = x$$

In the structure above $\phi(x)$ defines a set of numbers that are a square.

$$(\mathbb{R}, 0, 1, +, \cdot, \leq)$$

$$\phi(x) = \exists y \; y \cdot y = x$$

In the structure above $\phi(x)$ defines the set $[0, \infty)$.

$$(\mathbb{R}, 0, 1, +, \cdot, \leq)$$

$$\psi(x_1, x_2) = (x_1 \cdot x_1 + x_2 \cdot x_2 \leq 1.5) \wedge (x_1^2 \leq x_2)$$

This defines a set in $\mathbb{R}^2$.

We work with families of uniformly definable sets. Fix a formula $\phi(x_1 \dots x_n, y_1, \dots y_m) = \phi(\vec{x}, \vec{y})$. Plug in elements from the model for y variables to get a family of definable sets in M^n.

$$\mathcal{F}_\phi^M = \{\phi(x_1, \dots, x_n, a_1, \dots a_n) \mid a_1, \dots a_n \in M\}$$

Define $\mathrm{vc}^M(\phi)$ to be the vc-density of the family $\mathcal{F}_\phi^M$
Open Question: it is possible for $\mathrm{vc}^M(\phi)$ to be irrational?

$$\phi(x_1, x_2, y_1, y_2, y_3) = (x_1 - y_1)^2 + (x_2 - y_2)^2 \leq y_3^2$$

In structure $(\mathbb{R}, +, \cdot, \leq)$ given $a, b, r \in \mathbb{R}$ the formula $\phi(x_1, x_2, a, b, r)$ defines a disk in $\mathbb{R}^2$ with radius r with center (a, b). Thus $\mathcal{F}_\phi^{\mathbb{R}}$ is a collection of all disks in $\mathbb{R}^2$.

Shelah ('90) classified number of isomorphic classes for non-standard models. Important groups of structures included: stable, NIP, simple. A model M is said to be NIP if all uniformly definable families in it have finite vc-density.

- $(\mathbb{C}, 0, 1, +, \cdot)$ is stable (so both NIP and simple)
- $(\mathbb{R}, 0, 1, +, \cdot, \leq)$ is NIP and not stable
- $(\mathbb{Q}_p, 0, 1, +, \cdot, |)$ is NIP and not stable
- Random graph (V, R) is simple and not stable.
- Pseudo-finite fields are simple and not stable.
- $(\mathbb{Q}, 0, 1, +, \cdot)$ is in neither of those categories.

Given an NIP structure M we define a vc-function of n to be the largest vc-density achieved by families of uniformly definable sets in M^n.

$$\mathrm{vc}^M(n) = max\left\{\mathrm{vc}(\phi) \mid \phi(\vec{x}, \vec{y}) \text{ with } |\vec{x}| = n\right\}$$

Easy to show $\mathrm{vc}_M(n) \geq n\,\mathrm{vc}(1)$, $\mathrm{vc}(1) \geq 1$
Open question: Is $\mathrm{vc}_M(n) = n\,\mathrm{vc}_M(1)$? If not, is there a linear relationship?

Examples

- $(\mathbb{R}, 0, 1, +, \cdot, \leq)$ has $\mathrm{vc}(n) = n$ (true for all quasi o-minimal structures)
- $(\mathbb{C}, 0, 1, +, \cdot)$ has $\mathrm{vc}(n) = n$
- $(\mathbb{Q}_p, 0, 1, +, \cdot)$ has $\mathrm{vc}(n) \leq 2n - 1$
- ACVF has $\mathrm{vc}(n) \leq 2n$.

vc-density in trees

Consider structure $(T, \leq)$ where elements of T are vertices of a rooted tree and we say that $a \leq b$ if a is below b in the tree.

- Trees are NIP (Parigot '82)
- Trees are dp-minimal (Simon '11)
- Trees have $vc(n) = n$ (B. '13)

$\mathrm{tp}(a)$, a type of an element a is a set of all the formulas that that are true about a. Parigot's observation: there is a natural way to split a tree into parts A, B such that for $a \in A$ and $b \in B$ we have

$$\mathrm{tp}(a), \mathrm{tp}(b) \vdash \mathrm{tp}(ab)$$

This allows us to decompose complex types into simple parts, which we can use to compute vc-density.

Further applications

- $(\mathbb{Q}_p, 0, 1, +, \cdot, |)$
- other partial orderings, lattices

vc-density in Shelah-Spencer graphs

Consider a random graph on n vertices where the probability of the given two vertices having an edge is $n^{-\alpha}$. Shelah-Spencer graph is a limit of such graphs for α irrational in $(0,1)$. We view it in a language with a single binary relation.

- Shelah-Spencer graphs can be axiomatized (Shelah-Spencer '88)
- Shelah-Spencer graphs are stable (Baldwin-Shi '96, Baldwin-Shelah '97)

We show that $\mathrm{vc}^V(1) = \infty$, so vc-function is infinite. However to any formula $\phi(\vec{x}, \vec{y})$ we can prescribe a natural notion of dimension ϵ, and we have

$$\mathrm{vc}^V(\phi) < \frac{|x|}{\epsilon}$$

So even though vc-function is not well-behaved, there is still a linear structure on vc-density.

To a finite graph A assign a dimension $\delta(A) = |V| - \alpha|E|$. B/A is an extension. $\delta(B/A)$ is $\delta(A) = |V_B/V_A| - \alpha|E_B/E_A|$. B/A is called minimal if its dimension is negative, but every subextension is positive. $(A_0, \ldots A_n)$ is a n-minimal chain if A_{i+1}/A_i is minimal and adds no more than n new veritces. For any A_0 and n there exists a maximal n-minimal chain, moreover the largest set in that chain is unique. Such sets are called n-strong as every extension with no more than n new veritces