VC-density in model theoretic structures

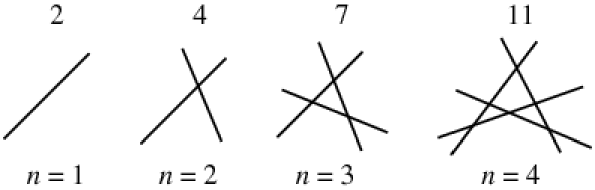
Anton Bobkov

June 3, 2015

Suppose we have an (infinite) collection of sets $\mathcal{F}$.
We define the shatter function $\pi_{\mathcal{F}} \colon \mathbb{N} \longrightarrow \mathbb{N}$ of $\mathcal{F}$

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

Example: Let $\mathcal{F}$ consist of all half-planes in the plane.



$$\pi_{\mathcal{F}}(1) = 2 \quad \pi_{\mathcal{F}}(2) = 4 \quad \pi_{\mathcal{F}}(3) = 7 \quad \pi_{\mathcal{F}}(4) = 11$$

$$\pi_{\mathcal{F}}(n) = n^2/2 + n/2 + 1$$

More examples:

1. Disks in the plane: $\pi_{\mathcal{F}}(n) = n^2 - n + 2$
2. Balls in $\mathbb{R}^3$: $\pi_{\mathcal{F}}(n) = n^3/3 - n^2 + 8n/3$
3. Intervals in the line: $\pi_{\mathcal{F}}(n) = 2n$
4. Finite subsets of $\mathbb{N}$: $\pi_{\mathcal{F}}(n) = 2^n$
5. Convex polygons in the plane: $\pi_{\mathcal{F}}(n) = 2^n$

Theorem (Sauer-Shelah '72)

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

Definition

Suppose the growth of the shatter function of $\mathcal{F}$ is polynomial. Let $\mathrm{vc}(\mathcal{F})$ be the infimum of all positive reals r such that

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

Call $\mathrm{vc}(\mathcal{F})$ the vc-density of $\mathcal{F}$. If the shatter function grows exponentially, we let $\mathrm{vc}(\mathcal{F}) := \infty$.

Applications

- Model Theory (NIP theories)
- VC-Theorem in probability (Vapnik-Chervonenkis '71)
- Computational learning theory (PAC learning, Warmuth conjecture)
- Computational geometry
- Functional analysis (Bourgain-Fremlin-Talagrand theory)
- Abstract topological dynamics (tame dynamical systems)

History

- VC-dimension defined by Vapnik-Chervonenkis '71
- NIP theories studied by Shelah '71
- vc-density in model theoretic context introduced by Aschenbrenner, Dolich, Haskell, Macpherson, Starchenko '13

Model Theory

Model Theory studies definable sets in first-order structures.

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

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

$\phi(\mathbb{Q})$ defines the set of rationals that are a square.

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

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

$\phi(\mathbb{R})$ defines the set $[0, \infty)$.

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

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

$\psi(\mathbb{R}^2)$ defines the set in $\mathbb{R}^2$ that is an intersection of a disc with an inside of a parabola.

Definition

Fix a formula $\phi(x_1 \dots x_m, y_1, \dots y_n) = \phi(\vec{x}, \vec{y})$ and structure M. Plug in elements from M for y variables to get a family of definable sets in M^m.

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

$\mathcal{F}_\phi^M$ is a <u>uniformly definable family</u>.
Define $\mathrm{vc}^M(\phi)$ to be the vc-density of the family $\mathcal{F}_\phi^M$.

Example

Consider the following formula in structure $(\mathbb{R}, 0, 1+, \cdot, \leq)$

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

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

For a given structure M, possible numbers of isomorphism classes for structures elementarily equivalent to it were classified by Shelah ('78). This classification involved defining important dividing lines describing complexity of structures. One of those diving lines is whether the structure is NIP or not.

Definition

Structure M is said to be <u>NIP</u> (no independence property) if all uniformly definable families in it have finite vc-density.

Examples:

- $(\mathbb{C}, 0, 1, +, \cdot)$ is NIP
- $(\mathbb{R}, 0, 1, +, \cdot, \leq)$ is NIP
- $(\mathbb{Q}_p, 0, 1, +, \cdot, |)$ is NIP
- Random graph (V, R) is not NIP
- $(\mathbb{Q}, 0, 1, +, \cdot)$ is not NIP.

Given an NIP structure M we define $\mathrm{vc}^M(n)$ to be the largest vc-density achieved by uniformly definable families in M^n.

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

Easy to show:

$$\mathrm{vc}^M(n) \geq n \cdot \mathrm{vc}^M(1) \geq n$$

Open Question: If M is NIP, is it possible for $\mathrm{vc}^M(\phi)$ to be irrational?
Open Question: Is $\mathrm{vc}^M(n) = n\,\mathrm{vc}^M(1)$? If not, is there a linear relationship? If $\mathrm{vc}(1) < \infty$ do we have $\mathrm{vc}(n) < \infty$?

Examples:

- $(\mathbb{R}, 0, 1, +, \cdot, \leq)$ has $\mathrm{vc}(n) = n$
- $(\mathbb{C}, 0, 1, +, \cdot)$ has $\mathrm{vc}(n) = n$
- $(\mathbb{Q}_p, 0, 1, +, \cdot)$ has $\mathrm{vc}(n) \leq 2n - 1$

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. '14)

proof background

$\text{tp}(a)$, 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

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

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

Shelah-Spencer graphs

Let α irrational $\in (0,1)$. Consider a random graph on n vertices where the probability of any given two vertices having an edge is $n^{-\alpha}$. Shelah-Spencer ('88) showed that 0-1 law holds for first-order sentences. An (infinite) structure satisfying those axioms is called a Shelah-Spencer graph.

- Shelah-Spencer graphs are stable (Baldwin-Shi '96, Baldwin-Shelah '97)
- Quantifier elimination (Laskowski '06)

Background

- B/A is called an <u>extension</u> if A is a subgraph of B.
- For B/A we define a <u>dimension</u>

$$\delta(B/A) = |V_B/V_A| - \alpha|E_B/E_A|$$

- B/A is called <u>minimal</u> if it has negative dimension, but its every subextension has a positive dimension.
- $(A_0, \dots A_n)$ is a <u>minimal chain</u> if each A_{i+1}/A_i is minimal.
- B/A is a <u>chain-minimal</u> extension if there exists a minimal chain $(A_0, \dots A_n)$ such with $A_0 = A, A_n = B$.

For B/A chain-minimal define

$$\phi_{A,B}(\vec{x}) = \exists \vec{x}^* \text{ such that } \vec{x}^*/\vec{x} \text{ is isomorphic to } B/A$$

Theorem (quantifier elimination, Laskowski '06)

In Shelah-Spencer graph every definable set can be defined by a boolean combination of formulas $\phi_{A_i,B_i(\vec{x})}$.

vc-density in Shelah-Spencer graphs

Theorem (B. '15)

For a formula $\phi(\vec{x}, \vec{y})$ we can define ϵ_L, ϵ_U explicitly computable from $\delta(B_i/A_i)$ such that

$$\epsilon_L|\vec{x}| \leq \mathrm{vc}(\phi) \leq \epsilon_U|\vec{x}|$$

Corollary

$\mathrm{vc}(1) = \infty$, *so vc-function is not well-behaved for this structure.*

Future work

- Other partial orderings, lattices
- Other graph structures, in particular flat graphs
- $(\mathbb{Q}_p, 0, 1, +, \cdot, |)$