Introduction to Model Theory

Charles Steinhorn, Vassar College
Katrin Tent, University of Münster

CIRM, January 9, 2018

Third lecture, part 2

- More advanced topics: Stability, forking, ranks, NIP and VC

V-C classes

Let $\mathcal{C}$ be a set of subsets of a(n infinite) set X. For $F \subseteq X$ write

$$\mathcal{C} \cap F := \{C \cap F : C \in \mathcal{C}\}.$$

Focus first on finite F. So $|\mathcal{C} \cap F| \leq 2^{|F|}$.

We say that $\mathcal{C}$ shatters F if we have equality.

Define $\pi_{\mathcal{C}} : \mathbb{N} \to \mathbb{N}$ by

$$\pi_{\mathcal{C}}(n) := \max\{|\mathcal{C} \cap F| : F \subset X, |F| = n\}.$$

So $0 \leq \pi_{\mathcal{C}}(n) \leq 2^n$.

Theorem (Sauer; Shelah; Vapnik-Chervonenkis; all around 1971)

Either $\pi_{\mathcal{C}}(n) = 2^n$ for all n or there is some $d \in \mathbb{N}$ such that $\pi_{\mathcal{C}}(n) \leq n^d$ for all sufficiently large n.

Call $\mathcal{C}$ a Vapnik-Chervonenkis, or just V-C class if there is some m such that $\pi_{\mathcal{C}}(n) < 2^n$ for all $n > m$. The least such m is called the V-C dimension of $\mathcal{C}$, and is ∞ otherwise.

A Better (monomial) Bound (Blumer, et.al., 1989) If $\mathcal{C}$ is a V-C class, then $\pi_{\mathcal{C}}(n) < (en/d)^d$, where d is the V-C dimension of $\mathcal{C}$.

Some Examples

- Let $\mathcal{H} := \{ax + by + c > 0 : a, b, c \in \mathbb{R}\}$. Then VC-dim $(\mathcal{H}) = 3$.
- (Even simpler) Let $\mathcal{L} := \{x \geq c : c \in \mathbb{R}\}$. Then VC-dim $(\mathcal{L}) = 1$.
- Let $\mathcal{C}_{Con} := \{\text{convex } C \subseteq \mathbb{R}^2\}$. Then $\mathcal{C}_{Con}$ is not a V-C class.

Back to model theory

We always work within an $\mathcal{L}$-structure $\mathcal{M}$ in the background.

Let $\varphi(\bar{x}; \bar{y})$ be a (partitioned) formula, where the length of $\bar{x}$ is n and of $\bar{y}$ is m.
A set $A \subseteq M^n$ is shattered by φ if for every $C \subseteq A$ there is some $b_C \in M^m$ such that

$$\mathcal{M} \models \varphi(\bar{c}; \bar{b}_C) \text{ if and only if } \bar{c} \in C.$$

By Compactness: this is equivalent to saying that φ shatters every finite subset of A.

A formula $\varphi(\bar{x}; \bar{y})$ is NIP or dependent if no infinite set of n-tuples is shattered by φ.

If $\varphi(\bar{x}; \bar{y})$ is NIP, then the V-C dimension of φ is the largest m such that φ shatters a set of size m.

A theory T is NIP every formula $\varphi(\bar{x}; \bar{y})$ is NIP.

More about NIP formulas

Let $\varphi^{opp}(\bar{y};\bar{x}) = \varphi(\bar{x};\bar{y})$.

Proposition

The formula $\varphi(\bar{x};\bar{y})$ is NIP iff $\varphi^{opp}(\bar{y};\bar{x})$ is NIP.

Proof.

Suppose that $\varphi(\bar{x};\bar{y})$ is not NIP (has IP). Apply compactness to find a set $A := \{a_S : S \subseteq \mathbb{N}\}$ such that φ shatters A. Let $b_I \in M^m$ be a parameter that corresponds to a set I of subsets of $\mathbb{N}$. For each $k \in \mathbb{N}$ let $W_k := \{S \subseteq \mathbb{N} : k \in S\}$, and put $\bar{b}_k = \bar{b}_{W_k}$. Then, for every $S \subset \mathbb{N}$,

$$\mathcal{M} \models \varphi^{opp}(\bar{b}_k, \bar{a}_S) \;\Leftrightarrow\; \mathcal{M} \models \varphi(\bar{a}_S, \bar{b}_k) \;\Leftrightarrow\; k \in S.$$

□

Lemma

Let T be a theory. If all formulas $\varphi(x; \bar{y})$ where x is a single variable are NIP, then T is NIP.

Lemma

Boolean combinations of NIP formulas are NIP.

Examples of NIP theories

- Every stable theory is NIP
- Every o-minimal theory is NIP
- ACVF is NIP
- The theory of $\mathbb{Q}_p$ is NIP.
- The theory of the countable Random Graph is not NIP.

See forkinganddividing.com

PAC Learning

Probably Approximately Correct Learning was developed by L. Valiant (1984). Uniform versions of the Law of Large Numbers such as the following are key.

Theorem (Vapnik-Chervonenkis (1971))

Let (X, μ) be a probability space and $\mathcal{S}$ a family of (measurable) subsets of X. For $S \in \mathcal{S}$ and $x_1, \ldots, x_n \in X$ let

$$Av(x_1, \ldots, x_n; S) = \frac{1}{n}|S \cap \{x_1, \ldots, x_n\}|.$$

Then for every $\epsilon > 0$ we have

$$\mu^n \left(\sup_{S \in \mathcal{S}} |Av(x_1, \ldots x_n; S) - \mu(S)| > \epsilon \right) \leq 8\, \pi_{\mathcal{S}}(n) \exp\left(-\frac{n\epsilon^2}{32} \right).$$

Apply these theorems to $\mathcal{S}$ that are VC classes, e.g., uniformly definable families of definable sets in o-minimal expansions of the real field (e.g., neural networks).

A brief bit about pseudofiniteness

A theory T is pseudofinite if it is elementarily equivalent to an ultraproduct of finite structures.
I want to focus on a particular class of finite structures, finite fields.
Ax characterized pseudofinite fields as exactly the infinite models of the theory of finite fields. In particular, *ultraproducts* of finite fields are pseudofinite.

Pseudofinite fields have IP. This can be seen by looking at *Paley graphs* on fields $\mathbb{F}_q$ where q is congruent to 1 mod 4. In the ultraproduct, the graph is elementarily equivalent to the countable Random graph.

Theorem (Chatzidakis-van den Dries-Macintyre ('92))

Let $\varphi(x_1, \ldots, x_n; y_1, \ldots, y_m)$ be a formula in the language of rings. Then there is a $C > 0$ and finitely many pairs (d_i, μ_i), $i \leq K$, with $d_i \in \{0, 1, \ldots, n\}$ (dimension) and $\mu_i \in \mathbb{Q}^{>0}$ (measure) such that for each finite field $\mathbb{F}_q$ and each $\bar{b} \in \mathbb{F}_q^m$,

i. *there is some $i \leq K$ satisfying*

$$\left| |\varphi(\mathbb{F}_q^n, \bar{b})| - \mu_i q^{d_i} \right| < C q^{d_i - \frac{1}{2}} \qquad (*)$$

ii. *(Definability of dimension and measure) For each (d_i, μ_i) there is a formula $\psi_i(\bar{y})$, such that for all $\mathbb{F}_q$ we have $\psi_i(\mathbb{F}_q^m) = \{\bar{b} \in \mathbb{F}_q^m : \bar{b}$ satisfies $(*)$ for $(d_i, \mu_i)\}$.*

Theorem (Tao's Algebraic Regularity Lemma)

For every $N \in \mathbb{N}^{>0}$ there is $C = C_N \in \mathbb{N}^{>0}$ such that whenever F is a finite field of cardinality greater than C, V and W are non-empty sets in cartesian powers of F, and $E \subseteq V \times W$, with V, W and E all definable of 'complexity' $\leq N$, then there are partitions $V = V_1 \cup \ldots \cup V_a$ and $W = W_1 \cup \ldots \cup W_b$ into definable sets of complexity $\leq C$, with:

1. *for all $i = 1, \ldots, a$ and $j = 1, \ldots, b$, we have $|V_i| \geq |V|/C$ and $|W_j| \geq |W|/C$, and*
2. *for all i, j, and sets $A \subset V_i$ and $B \subset W_j$, we have*

$$\big||E \cap (A \times B)| - d_{ij}|A||B|\big| \leq C|F|^{-1/4}|V_i|\,|W_j|,$$

where $d_{ij} = |E \cap (V_i \times W_j)|/|V_i||W_j|$.

The result holds for asymptotic classes of finite structures as well.