VC-dimension in model theory, discrete geometry, and combinatorics

Gabriel Conant
University of Cambridge

April 23, 2021
Géométrie et Théorie des Modèles

VC-dimension

Weak Law of Large Numbers

Let Ω be a set. Given a tuple $\boldsymbol{a} = (a_1, \ldots, a_n) \in \Omega^n$, define the average measure along $\boldsymbol{a}$ on subsets of Ω via

$$\mathrm{Av}_{\boldsymbol{a}}(X) = \frac{|\{1 \leq i \leq n : a_i \in X\}|}{n}.$$

Suppose μ is a probability measure on Ω, and $X \subseteq \Omega$ is measurable.

Markov's Inequality: For any $\epsilon > 0$ and $n \geq 1$,

$$\mu^n(\{\boldsymbol{a} \in \Omega^n : |\mathrm{Av}_{\boldsymbol{a}}(X) - \mu(X)| < \epsilon\}) \geq 1 - \frac{\mu(X)(1 - \mu(X))}{\epsilon^2 n}.$$

So, with ϵ fixed and $n \to \infty$, the probability that a randomly chosen sequence in Ω^n approximates $\mu(X)$ with error ϵ approaches 1.

Goal: Replace the single set X by a collection of sets.

Uniform Weak Law of Large Numbers

Let Ω be a set, and fix a collection $\mathcal{S}$ of subsets of Ω.
Define $\pi_{\mathcal{S}} : \mathbb{N} \to \mathbb{N}$ such that

$$\pi_{\mathcal{S}}(n) = \max_{F \in [\Omega]^n} |\{X \cap F : X \in \mathcal{S}\}|.$$

This is called the shatter function for $\mathcal{S}$. Note: $\pi_{\mathcal{S}}(n) \leq 2^n$.

Theorem (Vapnik & Chervonenkis 1968)

Suppose Ω is finite and μ is a probability measure on Ω such that every set in $\mathcal{S}$ is measurable. Given $\epsilon > 0$, we have

$$\mu^n\big(\{\boldsymbol{a} \in \Omega^n : \sup_{X \in \mathcal{S}} |\operatorname{Av}_{\boldsymbol{a}}(X) - \mu(X)| < \epsilon\}\big) \geq 1 - \frac{8\pi_{\mathcal{S}}(n)}{1.03^{\epsilon^2 n}}.$$

This is only useful when $\pi_{\mathcal{S}}(n)$ grows sub-exponentially.

VC-dimension

Fix any set Ω, and $\mathcal{S} \subseteq \mathcal{P}(\Omega)$.

Note: If $\pi_{\mathcal{S}}(d) < 2^d$ then $\pi_{\mathcal{S}}(n) < 2^n$ for all $n \geq d$.

Lemma (Sauer; Shelah; Vapnik & Chervonenkis; Perles)

Given $d \geq 1$, if $\pi_{\mathcal{S}}(d) < 2^d$ then for any n,

$$\pi_{\mathcal{S}}(n) \leq \sum_{i=0}^{d-1} \binom{n}{i} = O(n^{d-1}).$$

Definition

The VC-dimension of $\mathcal{S}$ is $\mathrm{VC}(\mathcal{S}) = \sup\{d \in \mathbb{N} : \pi_{\mathcal{S}}(d) = 2^d\}$.

So the lemma says that if $\mathrm{VC}(\mathcal{S}) = d < \infty$ then $\pi_{\mathcal{S}}(n) = O(n^d)$.

Examples

* The VC-dimension of all intervals in $\mathbb{R}$ is 2.

* The VC-dimension of all connected convex sets in $\mathbb{R}^2$ is ∞.

* The VC-dimension of all balls in $\mathbb{R}^d$ is $d + 1$.

* The VC-dimension of all axis-parallel boxes in $\mathbb{R}^d$ is $2d$.

* (Despres 2014, unpublished) The VC-dimension of all axis-parallel cubes in $\mathbb{R}^d$ is $\lfloor (3d + 1)/2 \rfloor$.

Beyond $\mathbb{R}^d$

Theorem (Gillibert, Lachmann, Müllner 2020)

The VC-dimension of all axis-parallel boxes in $(\mathbb{R}/\mathbb{Z})^d$ is

$$(1 + o(1))d \log_2 d.$$

Let F_r be the free group on r generators $x_1, \ldots, x_r$. A box is a set obtained by first fixing "side lengths" $L_1, \ldots, L_d$, and then forming all words $w \in F_r$ in which x_i and x_i^{-1} appear at most L_i times.

Questions

* What is the VC-dimension of all left translates of a fixed box in F_r?
* What is the VC-dimension of all left translates of all boxes in F_r?
* Same questions for the free nilpotent group of rank r and step s.

Finitely approximated measures

Approximations

Let $(\Omega, \mathcal{B}, \mu)$ be a finitely additive probability space.

Definition

An **ϵ-approximation** for $\mathcal{S} \subseteq \mathcal{B}$ is a finite tuple $\boldsymbol{a} \in \Omega^n$ such that, for any $X \in \mathcal{S}$, we have $|\operatorname{Av}_{\boldsymbol{a}}(X) - \mu(X)| < \epsilon$.

Corollary (of the VC Theorem and Sauer-Shelah Lemma)

Assume Ω is finite and suppose $\mathcal{S} \subseteq \mathcal{B}$ has $\mathrm{VC}(\mathcal{S}) = d$. Then for any $\epsilon > 0$, $\mathcal{S}$ has an ϵ-approximation of length $n \leq O(d\epsilon^{-2}\log(d\epsilon^{-1}))$.

Counterpoint: Let μ be a translation invariant finitely additive probability measure μ on subsets of $\mathbb{Z}$, and let $\mathcal{S}$ be the set of intervals in $\mathbb{Z}$. Then $\mathrm{VC}(\mathcal{S}) = 2$, but $\mathcal{S}$ has no ϵ-approximations for any $\epsilon < \frac{1}{2}$.

Finitely approximated measures

Let $(\Omega, \mathcal{B}, \mu)$ be as before.

We say that μ is finitely approximated with respect to $\mathcal{S} \subseteq \mathcal{B}$ if for all $\epsilon > 0$, there is an ϵ-approximation for $\mathcal{S}$.

Note: If Ω is finite then μ is finitely approximated with respect to $\mathcal{B}$.

Remark. If μ is finitely approximated with respect to $\mathcal{S} \subseteq \mathcal{B}$, and $\mathsf{VC}(\mathcal{S}) = d$, then for all $\epsilon > 0$, there is an ϵ-approximation for $\mathcal{S}$ of length $n \leq O(d\epsilon^{-2} \log(d\epsilon^{-1}))$.

Model Theory & NIP

Model Theory

Let T be a complete theory with infinite models. Fix a sufficiently saturated and strongly homogeneous monster model $\mathcal{U}$.

Given a tuple $\boldsymbol{x}$ of variables, $\mathrm{Def}_{\boldsymbol{x}}(\mathcal{U})$ denotes the Boolean algebra of definable (with parameters) subsets of $\mathcal{U}^{\boldsymbol{x}}$.

A Keisler measure (in $\boldsymbol{x}$) is a finitely additive probability measure on $\mathrm{Def}_{\boldsymbol{x}}(\mathcal{U})$. E.g., a type $p \in S_{\boldsymbol{x}}(\mathcal{U})$ is a $\{0, 1\}$-valued Keisler measure.

Definition (Chernikov-Starchenko)

A Keisler measure μ in $\boldsymbol{x}$ is finitely approximated if, for any formula $\varphi(\boldsymbol{x}, \boldsymbol{y})$, μ is finitely approximated with respect to the collection of subsets of $\mathcal{U}^{\boldsymbol{x}}$ defined by instances $\varphi(\boldsymbol{x}, \boldsymbol{b})$ for all $\boldsymbol{b} \in \mathcal{U}^{\boldsymbol{y}}$.

NIP Theories

Definition

T is NIP if for any formula $\varphi(\boldsymbol{x}, \boldsymbol{y})$, the collection of subsets of $\mathcal{U}^{\boldsymbol{x}}$ defined by instances $\varphi(\boldsymbol{x}, \boldsymbol{b})$ for $\boldsymbol{b} \in \mathcal{U}^{\boldsymbol{y}}$ has finite VC-dimension.

Main Focus: Finitely approximated Keisler measures in NIP theories

Theorem (Hrushovski, Pillay, Simon 2010)

Assume T is NIP and let μ be a Keisler measure in $\boldsymbol{x}$. Then μ is finitely approximated if and only if there is a small model $M \prec \mathcal{U}$ such that:

(*i*) *μ is definable over M, i.e., for any $\varphi(\boldsymbol{x}, \boldsymbol{y})$ and $r < s$ in $[0, 1]$, there is an M-definable set $X \subseteq \mathcal{U}^{\boldsymbol{y}}$ satisfying*

$$\{\boldsymbol{b} \in \mathcal{U}^{\boldsymbol{y}} : \mu(\varphi(\boldsymbol{x}, \boldsymbol{b})) \leq r\} \subseteq X \subseteq \{\boldsymbol{b} \in \mathcal{U}^{\boldsymbol{y}} : \mu(\varphi(\boldsymbol{x}, \boldsymbol{b})) < s\}.$$

(*ii*) *μ is finitely satisfiable in M, i.e., any definable subset of $\mathcal{U}^{\boldsymbol{x}}$ with positive measure has a solution in $M^{\boldsymbol{x}}$.*

Examples of finitely approximated measures

HPS: If T is NIP then μ is finitely approximated if and only if it is definable and finitely satisfiable in a small model.

Corollary (via Shelah, Keisler, etc.)

If T is stable then every Keisler measure is finitely approximated.

Other examples:

* Suppose T is NIP and pseudofinite. Then the normalized pseudofinite counting measure is finitely approximated.
* Suppose T is an *o*-minimal expansion of RCF. Let μ be the Lebesgue measure relativized to some set of finite measure. Then μ is finitely approximated.

fsg groups

Definable groups in NIP theories

Let $G = G(\mathcal{U})$ be a group definable in T.
G is definably amenable if it admits a left-invariant Keisler measure.

Theorem (Hrushovski & Pillay 2007)

Assume T is NIP and suppose G admits a left-invariant finitely approximated Keisler measure μ. Then μ is the unique left-invariant Keisler measure on G. Moreover, μ is right-invariant and is the unique right-invariant Keisler measure on G.

In the NIP context, a group G admitting such a measure is called *fsg*.

(Hrushovski, Peterzil, & Pillay 2006)

* If T is stable then G is definably amenable, and thus is an *fsg* group. (Proved independently by Newelski & Petrykowski.)
* If T is an *o*-minimal expansion of RCF, then any definably compact definable group (e.g., $SO_3(R)$) is an *fsg* group.

Definable sets in *fsg* groups

Suppose T is NIP and G is a definable *fsg* group.
Let μ denote the unique left-invariant Keisler measure on G.

Proposition

Given a definable set $X \subseteq G$, the following are equivalent:

1. G is the union of finitely many left translates of X.
2. G is the union of finitely many right translates of X.
3. $\mu(X) > 0$.

Proof. $(1 \Rightarrow 3)$. μ is finitely additive and left-invariant.

$(3 \Rightarrow 2)$. Choose a finite $\mu(X)$-approximation $\boldsymbol{a} \in G^n$ for the collection of left translates of X. Fix $g \in G$. Then $|\operatorname{Av}_{\boldsymbol{a}}(g^{-1}X) - \mu(g^{-1}X)| < \mu(X)$.

Since $\mu(g^{-1}X) = \mu(X)$, it follows that $\operatorname{Av}_{\boldsymbol{a}}(g^{-1}X) > 0$.

So some $a_i \in \boldsymbol{a}$ is in $g^{-1}X$, i.e., $g \in Xa_i^{-1}$.

$(2 \Rightarrow 3)$ and $(3 \Rightarrow 1)$ follow similarly, using right-invariance of μ. □

Approximating definable sets by subgroups

T is NIP and G is a definable *fsg* group witnessed by μ.

A definable set $X \subseteq G$ is generic if $\mu(X) > 0$.

Note: If X is generic, then the number of left/right translates of X required to cover G is bounded by $O(d\mu(X)^{-2}\log(d\mu(X)^{-1}))$, where d is the VC-dimension of the collection of right/left translates of X.

Theorem

Assume T is stable. Given a definable set $X \subseteq G$, there is a definable finite-index normal subgroup $H \leq G$ such that for any coset C of H, either $C \cap X$ or $C \backslash X$ is non-generic.

Remark: This theorem follows from classical stable group theory. But the above formulation is motivated by more recent work on "stable arithmetic regularity" (Terry & Wolf 2017/2018; C., Pillay, & Terry 2017).

Generic compact domination

A subset $X \subseteq G$ is type-definable if it is an intersection of a bounded number of definable subsets of G. If each of these definable sets is generic, then X is called wide.

Let G^{00} be the smallest type-definable bounded-index subgroup of G. (This exists for any group definable in an NIP theory by a result of Shelah.)

G/G^{00} is a compact Hausdorff group under the "logic topology".

Theorem (Hrushovski, Pillay, & Simon 2010; Simon 2015)

(T NIP and G *fsg*.) *Fix a definable set $X \subseteq G$, and let K be the set of cosets C of G^{00} such that both $C \cap X$ and $C \backslash X$ are wide. Then K is a closed set in G/G^{00}, and is Haar null.*

If T is *o*-minimal (or, more generally, distal) then a stronger version of this holds with "wide" replaced by "nonempty".

The “local” case: NIP formulas

NIP formulas

Fix a formula $\varphi(\boldsymbol{x}, \boldsymbol{y})$.

- $\mathcal{S}_\varphi$ denotes the collection of subsets of $\mathcal{U}^{\boldsymbol{x}}$ defined by instances $\varphi(\boldsymbol{x}, \boldsymbol{b})$ for $\boldsymbol{b} \in \mathcal{U}^{\boldsymbol{y}}$.
- $\mathrm{Def}_\varphi(\mathcal{U})$ denotes the Boolean algebra on $\mathcal{U}^{\boldsymbol{x}}$ generated by $\mathcal{S}_\varphi$.

We say $\varphi(\boldsymbol{x}, \boldsymbol{y})$ is NIP if $\mathcal{S}_\varphi$ has finite VC-dimension. A φ-Keisler measure is a finitely additive probability measure on $\mathrm{Def}_\varphi(\mathcal{U})$.

Theorem (Gannon 2018)

Assume $\varphi(\boldsymbol{x}, \boldsymbol{y})$ is NIP. Then a φ-Keisler measure μ is finitely approximated if and only if there is some $M \prec \mathcal{U}$ such that μ is definable over M and finitely satisfiable in M.

Graph regularity

Let (V, E) be a finite graph and fix $\epsilon > 0$.

Given $X, Y \subseteq V$, call (X, Y) ϵ-regular if there is some $\delta \in [0, 1]$ such that for any $A \subseteq X$ and $B \subseteq Y$, we have

$$\big||E \cap (A \times B)| - \delta|A \times B|\big| < \epsilon|X \times Y|$$

Szemerédi's Regularity Lemma (weak version)

Let (V, E) be a finite graph. Then for any $\epsilon > 0$ there is a partition $V = X_1 \cup \ldots \cup X_n$, with $n \leq O_\epsilon(1)$, and a set $P \subseteq \{1, \ldots, n\}^2$ such that:

- $|\bigcup_{(i,j)\in P} X_i \times X_j| > (1 - \epsilon)|V|^2$ and
- if $(i, j) \in P$ then (X_i, X_j) is ϵ-regular.

In general, the best upper bound on m is $\exp^{(1/\epsilon)^{O(1)}}(1)$.

Graph regularity and VC-dimension

The VC-dimension of a graph (V, E) is the VC-dimension of the collection of subsets of V defined by $E(x, b)$ for $b \in V$.

Theorem (Alon, Fisher, Newman 2007; Lovász-Szegedy 2010)

Let (V, E) be a finite graph of VC-dimension d. Then for any $\epsilon > 0$ there is a partition $V = X_1 \cup \ldots \cup X_n$, with $n \leq O_d((1/\epsilon)^{O_d(1)})$, and a set $P \subseteq \{1, \ldots, n\}^2$ such that:

- $|\bigcup_{(i,j) \in P} X_i \times X_j| > (1 - \epsilon)|V|^2$ *and*
- *if $(i, j) \in P$ then (X_i, X_j) is ϵ-regular with density in $[0, \epsilon) \cup (1 - \epsilon, 1]$.*

Graph regularity and VC-dimension, continued

Theorem (Chernikov & Starchenko 2016)

Suppose (V, E) is an arbitrary graph of VC-dimension d, and μ is a finitely approximated E-Keisler measure on V.

Then for any $\epsilon > 0$ there is a partition $V = X_1 \cup \ldots \cup X_n$, with $n \leq O_d((1/\epsilon)^{O_d(1)})$, and a set $P \subseteq \{1, \ldots, n\}^2$ such that:

- $\mu^2(\bigcup_{(i,j)\in P}(X_i \times X_j)) > 1 - \epsilon$.
- *If $(i, j) \in P$ then (X_i, X_j) is ϵ-regular with density in $[0, \epsilon) \cup (1 - \epsilon, 1]$.*
- *Each X_i is a Boolean combination of neighborhoods of E.*

Idea: Fix an ϵ-approximation $\boldsymbol{a} \in V^m$ for E with $m \leq O_d((1/\epsilon)^{O(1)})$. The X_i's are atoms of the Boolean algebra generated by E-neighborhoods of vertices in $\boldsymbol{a}$. Use Sauer-Shelah to bound the number of such atoms by $O_d((1/\epsilon)^{O_d(1)})$.

Approximate subgroups

Let G be a group.
A k-approximate subgroup of G is a subset $A \subseteq G$ such that $1 \in A$, $A = A^{-1}$, and $A^2 \subseteq FA$ for some $F \subseteq G$ of size k.

Freiman's Theorem (1973) If A is a finite k-approximate subgroup of $\mathbb{Z}$ then there is a generalized arithmetic progression $P \subseteq 4A$ of rank $r = O_k(1)$ such that $A \subseteq F + P$ for some $F \subseteq \mathbb{Z}$ of size $O_k(1)$.

Theorem (Breuillard, Green, & Tao 2010)

If A is a finite k-approximate subgroup of an arbitrary group G, then there is a coset nilprogression $P \subseteq A^4$ of rank and step $O_k(1)$ such that $A \subseteq FP$ for some $F \subseteq G$ of size $O_k(1)$.

A coset nilprogression of rank r and step s is a set of the form XH where X is a homomorphic image of a box in the free nilpotent group of rank r and step s, and H is a finite subgroup normalized by X.

VC-dimension for subsets of groups

Given a group G, the VC-dimension of a subset $A \subseteq G$ is the VC-dimension of the set of left translates of A.

Examples:

1. If H is a proper subgroup of G, then the VC-dimension of H is 1.
2. If G is abelian and $P \subseteq G$ is a generalized arithmetic progression of rank r, then the VC-dimension of A is at most $2r$.

Question: Is the VC-dimension of a coset nilprogression of rank r and step s (in an arbitrary group) bounded in terms of r and s?

Approximate subgroups of bounded VC-dimension

Theorem (C.-Pillay 2020)

Suppose G is a group and $A \subseteq G$ is a finite k-approximate subgroup of VC-dimension d. Then, for any $\epsilon > 0$, there is a coset nilprogression $P \subseteq G$ of rank and step $O_{k,d,\epsilon}(1)$ such that:

* *$P \subseteq A^2$ and $A \subseteq FP$ for some $F \subseteq G$ of size $O_{k,d,\epsilon}(1)$.*
* *There is a set $Z \subseteq AP$ of size at most $\epsilon|A|$ such that for any $g \in G \backslash Z$, either $|gP \cap A| < \epsilon|P|$ or $|gP \backslash A| < \epsilon|P|$.*
* *There is some $D \subseteq F$ such that $|A \triangle DP| < \epsilon|A|$.*

This result unifies the BGT theory of approximate groups with the study of "tame arithmetic regularity", c.f., Terry & Wolf and C., Pillay, & Terry.

Proof Strategy

Let G be a sufficiently saturated expansion of a group, and $A \subseteq G$ a pseudofinite k-approximate subgroup of VC-dimension d.

Let μ be the $|A|$-normalized pseudofinite counting measure on definable subsets of G, i.e., $\mu(X) = \mathrm{st}(|X|/|A|)$.

Then μ is $\mathbb{R}$-valued on definable subsets of $\langle A \rangle = \bigcup_{n \geq 1} A^n$.

Let $\varphi(x; y, z)$ be the formula $x \in y \cdot A \cdot z$. Since A is pseudofinite and of finite VC-dimension, it follows that μ is finitely approximated with respect to any φ-definable set.

Step 1: (Connected component)
There is a smallest φ-type-definable bounded-index subgroup of $\langle A \rangle$, which we denote H. Moreover, H is contained in A^2. In fact,

$$H = \{x \in \langle A \rangle : \mu(xgA \triangle gA) = 0 \text{ for all } g \in G\}.$$

Proof Strategy

$\langle A \rangle / H$ is a locally compact Hausdorff group under the logic topology.

Step 2: (Local *fsg*)
If $K \subseteq \langle A \rangle / H$ is the set of cosets C of H such that neither $C \cap A$ nor $C \backslash A$ is contained in a definable set of μ-measure 0, then K is compact and Haar null.

Corollary: For any $\epsilon > 0$, there is a definable set $X \subseteq A^2$ (containing H) and a definable set $Z \subseteq AX$ with $\mu(Z) < \epsilon$, such that for any $g \in G \backslash Z$, either $\mu(gX \cap A) = 0$ or $\mu(gX \backslash A) = 0$.

Step 3 (Multiplicative combinatorics)
Use BGT and ideas around Ruzsa's Covering Lemma to replace X by a pseudofinite coset nilprogression.

Recap

Graph Regularity

* Arbitrary graphs admit regular partitions with uniform densities.
* Graphs of bounded VC-dimension admit regular partitions of polynomial size with almost 0 or 1 densities.

Approximate subgroups

* Arbitrary approximate subgroups are commensurable to coset nilprogressions.
* Approximate subgroups of bounded VC-dimension are approximated by translates of coset nilprogressions.

Question: Is there a proof of the result on approximate groups of bounded VC-dimension that avoids the appeal to BGT?