VC DIMENSION IN GROUPS

BASED ON JOINT WORKS W/

a) B. MCDONALD & E. WYMAN

b) B. RODGERS

§1 DEFINITION

$(G, \cdot)$: FINITE GROUP OF ORDER N.

$S \subseteq G$

S SHATTERS $Y = \{y_1, \cdots, y_k\}$

$\Leftrightarrow$

$\forall \Phi \subseteq A \subseteq Y$, $\exists\, x \in G$ s.t. $xS \cap Y = A$

THE VAPNIK-CHERVONENKIS DIMENSION OF $S \subseteq G$ ($\text{VC-dim}_G(S)$) IS THE SIZE OF THE LARGEST SHATTERED BY S.

N.B.: $\text{VC-dim}_G(S) \leq \log_2 N$. $(N := |G|)$

e.g. $G = \mathbb{Z}/11\mathbb{Z}$ $S = \{1^2, 2^2, 3^2, 4^2, 5^2\}$

$= \{1, 3, 4, 5, 9\}$

e.g. $G = \mathbb{Z}/11\mathbb{Z}$

$S = \{1^2, 2^2, 3^2, 4^2, 5^2\}$

$= \{1, 3, 4, 5, 9\}$

$S - 1 = \{0, 2, 3, 4, 8\}$

$11 < 2^4$

$Y = \{0, 1, 2\}$

$Y = \{0, 1, 2, 3\}$

$VC\dim_G(S) = 3$

0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10

$(S-2) \cap Y = \{0\}$

NOTE:

$$\mathcal{F}(S) = \{ xS : x \in G \}$$ SET-SYSTEMS

$$\mathcal{H}(S) = \{ \mathbb{1}_{xS} : x \in G \}$$ CLASSIFIERS

$$\mathbb{1}_A(x) = \begin{cases} 1 & x \in S \\ 0 & x \notin S \end{cases}$$

§2 MAIN QUESTION

Q. $G_N = (\mathbb{Z}/N\mathbb{Z}, +)$ (N PRIME)

$\mathcal{A} = \{ x^2 : x \in \mathbb{Z}/N\mathbb{Z} \}$

WHAT IS $VC\text{-}\dim_{G_N}(\mathcal{A})$ AS $N \longrightarrow \infty$?

Thm (MCDONALD – S. – WYMAN, 2025 IN DISCRETE MATH.)

LET $G = (\mathbb{Z}/N\mathbb{Z}, +)$ & $\mathcal{A} = \{ x^2 : x \in (\mathbb{Z}/N\mathbb{Z}) \}$.

THEN, AS $N \to \infty$ THROUGH THE PRIMES,

$$VC\text{-}\dim_G(\mathcal{A}) \geq \left(\frac{1}{2} - o(1)\right) \log_2 N$$

NOTE: GENERALIZES TO $G = (\mathbb{F}_q, +)$ &

$$\mathcal{A}_r = \{ x^r : x \in \mathbb{F}_q \}$$

(q PRIME POWER
$r \mid q-1$)

SKETCH OF PROOF

χ : LEGENDRE SYMBOL MOD N.

$$\chi(n) = \begin{cases} 1 & n = \square \\ -1 & n \neq \square \\ 0 & n = 0 \end{cases}$$

$$\mathbb{1}_{\text{sq}} \approx \frac{1+\chi}{2}$$

S SHATTERS $Y = \{y_1, \dots, y_k\}$

$\Updownarrow$

$\forall\, 0 \subseteq A \subseteq Y, \quad \exists\, x \in \mathbb{Z}/N\mathbb{Z} \quad$ s.t. $\quad (x + \mathcal{A}) \cap Y = A$

FOR SIMPLICITY TAKE $A = Y$.

NEED TO SHOW: $\exists\, x$ s.t. $Y - x \subseteq \mathcal{A}$ $\qquad Y = \{y_1, \dots, y_k\}$

$$\sum_x \prod_{j=1}^{k} \mathbb{1}_{\mathcal{A}}(y_k - x) > 0$$

$$\sum_{x} \prod_{j=1}^{k} \mathbb{1}_{\square}(\gamma_k - x)$$

$$\mathbb{1} \approx \frac{1 + \chi}{2}$$

$$\sum_{x} \left[\frac{1}{2^k} + \sum \chi(\ldots) \right]$$

$\frac{1}{2^k} > 0$

SMALL

$$\sum_{x=0}^{N-1} \chi(f(x))$$

WEIL'S BOUND FOR CHARACTER.

CONJ. (MCDONALD – S. – WYMAN, CORRECTED FORM BY RODGERS – S.)

FOR $G = (\mathbb{Z}/N\mathbb{Z}, +)$ & $\mathcal{A}_r = \{x^r : x \in \mathbb{Z}/N\mathbb{Z}\}$,

$$\mathrm{VCdim}_G(\mathcal{A}_r) = (1 + o(1)) \log_r N$$

AS $N \to \infty$ ALONG $N \equiv 1 \pmod r$. N PRIME.

$0 < p < 1$

§3 RANDOM MODEL

$(G, \cdot)$ $\longrightarrow$ NOT NECESSARILY ABELIAN.

S $\longrightarrow$ RANDOMLY SELECTED

$S \sim \text{BERN}(G, p)$

FIX $0 < p < 1$,

$$\Pr[x \in S] = p \qquad \text{(i.i.d. FOR } x \in G\text{)}$$

$$\Leftrightarrow \quad \Pr[S = W] = p^{|W|}(1-p)^{N-|W|}$$

$W \to$ DETERMINISTIC

HEURISTIC : SQUARES ARE PSEUDORANDOM

$$\mathcal{L} \simeq \mathrm{BERN}\left(\mathbb{Z}/N\mathbb{Z}, 1/2\right)$$

MORE GENERALLY

$$\mathcal{L}_r \simeq \mathrm{BERN}\left(\mathbb{Z}/N\mathbb{Z}, 1/r\right).$$

Thm: (RODGERS – S., 2025+)

FIX $0 < p < 1$. SET $1/r = \min(p, 1-p)$

LET $(G, \cdot)$ BE ANY FINITE GROUP OF ORDER N. LET $S \subseteq G$ BE A BERNOULLI SAMPLED RANDOM SUBSET WITH PROB p.

THEN ASYMPTOTICALLY ALMOST SURELY, AS $N \to \infty$, (PROB OF FAILURE $O(N^{-\eta})$)

$$VC\dim_G(S) = (1+o(1)) \log_r N.$$