Introduction to VC-dimension and its applications in combinatorial problems

Hyoyoon Lee
2024.08.20

<u>Def.</u> $(X, \mathcal{F})$ is called a <u>set system</u> if X is a set and $\mathcal{F} \subseteq \mathcal{P}(X)$.
$\mathcal{P}(X) = \{Y : Y \subseteq X\}$

<u>Def.</u> Let $(X, \mathcal{F})$ be a set system.

(1) $A \subseteq X$ is <u>shattered</u> by $\mathcal{F}$ if $\forall A' \subseteq A$, $\exists F \in \mathcal{F}$ s.t. $A \cap F = A'$.

Equivalently, $A \subseteq X$ is shattered by $\mathcal{F}$ if $\{A \cap F : F \in \mathcal{F}\} = \mathcal{P}(A)$.

(2) $\mathcal{F}$ has <u>VC-dimension $\geq d$</u> if $\exists A \subseteq X$ shattered by $\mathcal{F}$ with $|A| = d$.

We write <u>$VC(\mathcal{F}) \geq d$</u>.

If $VC(\mathcal{F}) \geq d$ but not $VC(\mathcal{F}) \geq d+1$, then <u>$VC(\mathcal{F}) = d$</u>.

If $VC(\mathcal{F}) \geq d$ for all $d \in \mathbb{N}$, then <u>$VC(\mathcal{F}) = \infty$</u>.

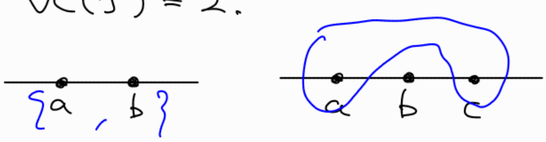
(3) For $A \subseteq X$, <u>$A \cap \mathcal{F}$</u> $:= \{A \cap F : F \in \mathcal{F}\}$.

<u>Example.</u>

(1) Let $X = \mathbb{R}$, $\mathcal{F} = \{(-\infty, a) : a \in \mathbb{R}\} \cup \{(b, \infty) : b \in \mathbb{R}\}$.

Then $VC(\mathcal{F}) = 2$.

$\because$



$\{a, b, c\}$ is not shattered.
$(\{a, b, c\} \cap \mathcal{F} \not\ni \{a, c\}.)$

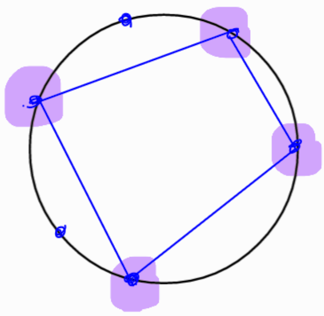
(2) Let $X = \mathbb{R}^2$, $\mathcal{F} = \{\text{convex polygons}\}$.

Then $VC(\mathcal{F}) = \infty$.

$\because$ Let $A \subset \{(x, y) : x^2 + y^2 = 1\}$, $|A| = n$.

$\forall A' \subseteq A$ with $|A'| = m \leq n$,

$\exists$ convex m-gon G s.t. $A \cap G = A'$.



<u>Def.</u> Let $(X, \mathcal{F})$ be a set system.

The <u>shatter function</u> $\pi_{\mathcal{F}}(n) := \max\{|A \cap \mathcal{F}| : A \subseteq X \text{ and } |A| = n\}$. $(\leq 2^n)$

<u>Lemma.</u> (Sauer-Shelah lemma)

Let $(X, \mathcal{F})$ be a set system. Assume $VC(\mathcal{F}) \leq d$.

Then $\forall n \geq d$, $\pi_{\mathcal{F}}(n) \leq \sum_{i=0}^{d} \binom{n}{i}$. $(= O(n^d)$, i.e. polynomial growth.$)$

Notation.

$$Av(x_1, \dots, x_n ; A) := \frac{(1_A(x_1) + \cdots + 1_A(x_n))}{n} = \text{relative frequency of } A$$

Theorem. (Weak law of large numbers) (WLLN)

$\forall$ event F, $\forall \varepsilon > 0$, $\forall n > 0$,

$$\mathbb{P}\left[\left| Av(x_1, \dots, x_n ; F) - \mathbb{P}[F] \right| \geq \varepsilon \right] \leq \frac{1}{4n\varepsilon^2} \quad (\to 0 \text{ as } n \to \infty).$$

Remark. $P(A \cup B) \leq P(A) + P(B)$

Let $\mathcal{F}$ be some family of events. By WLLN and union bound,

$$\mathbb{P}\left[\exists \text{ event } F \in \mathcal{F} \text{ s.t. } \left| Av(x_1, \dots, x_n ; F) - \mathbb{P}[F] \right| \geq \varepsilon \right]$$

$$= \mathbb{P}\left[\sup \left\{ \left| Av(x_1, \dots, x_n ; F) - \mathbb{P}[F] \right| : F \in \mathcal{F} \right\} \geq \varepsilon \right] \leq \frac{1}{4n\varepsilon^2} \cdot |\mathcal{F}|.$$

This "uniform version" w.r.t. $\mathcal{F}$ has dependency on $\mathcal{F}$.

Theorem. (VC-theorem) (Uniform version of the WLLN)

Let $(X, \mathcal{F})$ be a set system with finite X.

Regard X as a probability space with some $\mathbb{P}$. $\forall \varepsilon > 0$,

(e.g. $\mathbb{P}(A) = |A| / |X|$. Counting probability.)

$$\mathbb{P}\left[\sup \left\{ \left| Av(x_1, \dots, x_n ; F) - \mathbb{P}[F] \right| : F \in \mathcal{F} \right\} > \varepsilon \right] \leq 8 \pi_{\mathcal{F}}(n) \exp\left(-\frac{n\varepsilon^2}{32} \right).$$

($\to 0$ as $n \to \infty$ if $\pi_{\mathcal{F}}(n)$ has polynomial growth.)

Corollary.

(no dependency on $\mathcal{F}$)

$\forall d \in \mathbb{N}$, $\forall \varepsilon > 0$, $\exists N = N(d, \varepsilon)$ s.t.

$\forall$ set system $(X, \mathcal{F})$ with some $\mathbb{P}$ on finite X,

$VC(\mathcal{F}) \leq d \Rightarrow \exists$ ε-approximation for $\mathcal{F}$ of size $\leq N$.

($\exists \{x_1, \dots, x_N\} \subseteq X$ s.t. $\forall F \in \mathcal{F}$, $\left| Av(x_1, \dots, x_N ; F) - \mathbb{P}[F] \right| < \varepsilon$.

(possibly with repetitions))

<u>Proof.</u> By VC-theorem,

$$\mathbb{P}\left[\sup\left\{\left|Av(x_1,\dots,x_n;F)-\mathbb{P}[F]\right| : F\in\mathcal{F}\right\}>\varepsilon\right]\le 8\pi_{\mathcal{F}}(n)\exp\left(-\frac{n\varepsilon^2}{32}\right).$$

$$\Rightarrow \mathbb{P}\left[\forall F\in\mathcal{F},\ \left|Av(x_1,\dots,x_n;F)-\mathbb{P}[F]\right|\le\varepsilon\right]\ge 1-8\pi_{\mathcal{F}}(n)\exp\left(-\frac{n\varepsilon^2}{32}\right).$$

By SS lemma, $8\pi_{\mathcal{F}}(n)\exp\left(-\frac{n\varepsilon^2}{32}\right)\to 0$ as $n\to\infty$.

In particular, with sufficiently large N depending only on d and ε,

$$\mathbb{P}\left[\forall F\in\mathcal{F},\ \left|Av(x_1,\dots,x_N;F)-\mathbb{P}[F]\right|<\varepsilon\right]\ne 0.$$

Thus $\exists$ at least one N-tuple $\{x_1,\dots,x_N\}$ s.t.

$\forall F\in\mathcal{F},\ \left|Av(x_1,\dots,x_N;F)-\mathbb{P}[F]\right|<\varepsilon.$ ■

Application to bounding transversal numbers

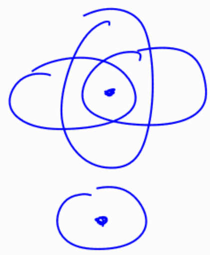
<u>Def.</u>

Let $(X,\mathcal{F})$ be a set system.

(1) $T\subseteq X$ is a <u>transversal</u> of $\mathcal{F}$ if $\forall F\in\mathcal{F}$, $T\cap F\ne\emptyset$.

The <u>transversal number</u> of $\mathcal{F}$,

<u>$\tau(\mathcal{F})$</u> is the smallest size of a transversal.

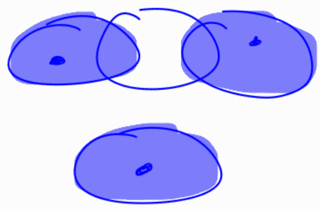


$\tau=2$

(2) $\mathcal{G}\subseteq\mathcal{F}$ is a <u>packing</u> if $\forall F_1\ne F_2\in\mathcal{G}$, $F_1\cap F_2=\emptyset$.

The <u>packing number</u> of $\mathcal{F}$,

<u>$\nu(\mathcal{F})$</u> is the largest size of a packing.



$\nu=3$

<u>Remark.</u>

(1) $\nu(\mathcal{F})\le\tau(\mathcal{F})$.

($\because$ Let $\mathcal{G}\subseteq\mathcal{F}$ be a packing with the largest size.

Then $\forall F\in\mathcal{G}$, any transversal $T\subseteq X$ should

contain one element from F.)

(2) It is difficult to calculate the actual transversal/packing number.

(integer programming in NP-hard.)

(3) Very little is known about the reverse direction:

$\tau \le f(\nu)$ for some fct f under "reasonable" conditions?

(4) If H is a hypergraph, then we have a set system

$X = V(H)$, $\mathcal{F} = E(H) \subseteq \mathcal{P}(V(H))$,

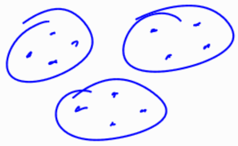
and a transversal of $\mathcal{F}$ is called a <u>vertex cover</u> of H, <u>$\tau(H)$</u>,

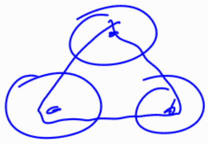
a packing of $\mathcal{F}$ is called a <u>matching</u> of H, <u>$\nu(H)$</u>.

(5) If H is an r-uniform hypergraph, i.e. $\forall e \in E(H)$, $|e| = r$, then

$\tau(H) \le r \cdot \nu(H)$.

($\because$ The union of edges from a maximal matching is a vertex cover.)





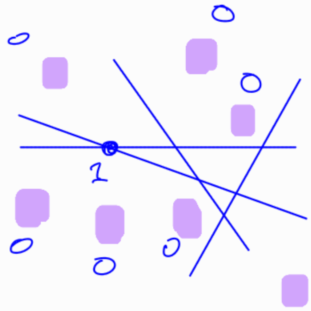
Ryser's Conjecture (1971)

If H is an r-uniform r-partite hypergraph, then $\tau(H) \le (r-1) \cdot \nu(H)$.

(Solved ~~only~~ for $r = 2$: König's theorem (1931).)

and $r = 3$ by Ron Aharoni (2001)

<u>Example.</u>



Let $X = \mathbb{R}^2$ and $\mathcal{F}$ be a set of n lines $\ell_1, \dots, \ell_n$ s.t.

$\forall i < j < k$, $|\ell_i \cap \ell_j| = 1$ but $\ell_i \cap \ell_j \cap \ell_k = \emptyset$.

Then $\nu(\mathcal{F}) = 1$ ($\because$ every two lines intersect),

but $\tau(\mathcal{F}) \ge \frac{n}{2}$ ($\because$ any point is contained in at most 2 lines).

<u>Def.</u>

(1) Let $(X, \mathcal{F})$ be a set system with X finite.

A function $\phi : X \to [0,1]$ is a <u>fractional transversal</u> for $\mathcal{F}$

if $\forall F \in \mathcal{F}$, $\sum_{x \in F} \phi(x) \ge 1$.

The <u>size</u> of ϕ is $\sum_{x \in X} \phi(x)$.

The <u>fractional transversal number $\tau^*(\mathcal{F})$</u> is

the infimum of the sizes of fractional transversals for $\mathcal{F}$.

(2) Let $(X, \mathcal{F})$ be a set system with X finite.

A function $\psi : \mathcal{F} \to [0,1]$ is a <u>fractional packing</u> for $\mathcal{F}$

if $\forall x \in X$, $\sum_{\{F \in \mathcal{F} : x \in F\}} \psi(F) \leq 1$.

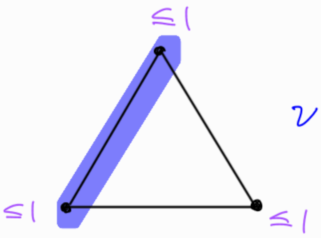
The <u>size</u> of ψ is $\sum_{F \in \mathcal{F}} \psi(F)$.

The <u>fractional packing number $\nu^*(\mathcal{F})$ is</u>

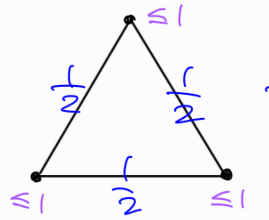
the supremum of the sizes of fractional packings for $\mathcal{F}$.

<u>Example.</u>

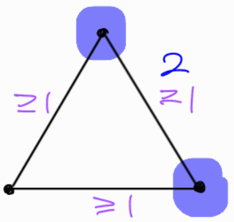
Triangle. $X = \{a_1, a_2, a_3\}$, $\mathcal{F} = \{\{a_1,a_2\}, \{a_2,a_3\}, \{a_3,a_1\}\}$.



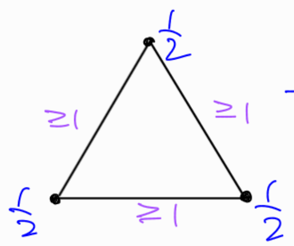
$\nu = 1$ matching number



$\nu^* = \frac{3}{2}$



$\tau =$ min. vertex cover

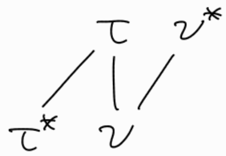


$\tau^* = \frac{3}{2}$

<u>Remark.</u>

$\nu^* \geq \nu \leq \tau \geq \tau^*$.

max min



<u>Remark.</u>

(1) $X = [2n]$, $\mathcal{F} = \binom{X}{n}$. Then $\tau = n+1$ and $\tau^* = 2$.

($\because$ If n points $a_1, \dots, a_n$ are chosen, then $X \setminus \{a_1, \dots, a_n\} \in \binom{X}{n}$ and

$\{a_1, \dots, a_n\} \cap (X \setminus \{a_1, \dots, a_n\}) = \emptyset \Rightarrow \tau > n$.

Assign $\phi(x) = \frac{1}{n}$ for all $x \in X \Rightarrow \tau^* \leq 2$.

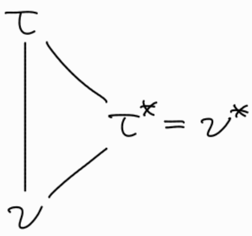
$\sum_{x \in X} \phi(x) = \sum_{x \in F} \phi(x) + \sum_{x \in X \setminus F} \phi(x) \geq 1 + 1 = 2$.)

(2) $X = \binom{[n]}{2}$, $\mathcal{F} = \{\{\{1,2\}, \{1,3\}, \{1,4\}, \dots\},$

$\{\{2,1\}, \{2,3\}, \{2,4\}, \dots\}, \dots\}$

Then $\nu = 1$, $\nu^* = \frac{n}{2}$.

$\mathcal{F}$: finite



Fact.

(1) $\mathcal{F}$: finite $\Rightarrow \tau^*(\mathcal{F}) = \nu^*(\mathcal{F})$.

(2) It is easy to calculate or bound the fractional transversal number (linear programming, P.)

Thus it is very useful to bound τ in terms of τ^*.

Def.

Let $(X, \mathcal{F})$ be a set system with some $\mathbb{P}$ on X.
$A \subseteq X$ is called an <u>ε-net</u> if $\forall F \in \mathcal{F}$ s.t. $\mathbb{P}[F] \geq \varepsilon$, $A \cap F \neq \emptyset$.

Remark.

Note that every ε-approximation is an ε-net.

($\because$ Let $\{x_1, \dots, x_n\}$ be an ε-approximation.

$\forall F \in \mathcal{F}$, $|Av(x_1, \dots, x_n ; F) - \mathbb{P}[F]| < \varepsilon$

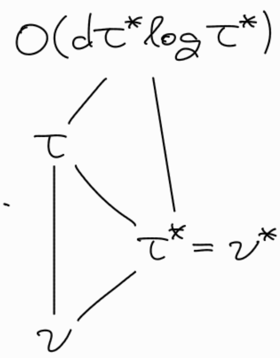
$\Rightarrow$ If $\mathbb{P}[F] \geq \varepsilon$ but $\{x_1, \dots, x_n\} \cap F = \emptyset$, then $Av(x_1, \dots, x_n ; F) = 0$

and $|Av(x_1, \dots, x_n ; F) - \mathbb{P}[F]| \geq \varepsilon$, ↯.)

Fact.

$VC(\mathcal{F}) \leq d \Rightarrow \exists\, \varepsilon$-net for $\mathcal{F}$ of size $\leq Cd \frac{1}{\varepsilon} \log \frac{1}{\varepsilon}$.

$\mathcal{F}$: finite, $VC(\mathcal{F}) \leq d$



Proposition.

Let $(X, \mathcal{F})$ be a set system with finite X and $VC(\mathcal{F}) \leq d$.
Then $\tau(\mathcal{F}) \leq Cd\,\tau^*(\mathcal{F}) \log \tau^*(\mathcal{F})$, where C : constant.

Proof.

Let $\phi : X \to [0,1]$ be an optimal fractional transversal, so that

$\tau^*(\mathcal{F}) = \sum_{x \in X} \phi(x)$.

A function $\mathbb{P}$ defined by $\mathbb{P}(\{x\}) = \dfrac{\phi(x)}{\tau^*(\mathcal{F})}$ for all $x \in X$ defines a probability on X.

Note that $\forall F \in \mathcal{F}$, $\mathbb{P}(F) = \dfrac{\sum_{x\in F} \phi(x)}{\tau^*(\mathcal{F})} \geq \dfrac{1}{\tau^*(\mathcal{F})}$.

Let $A \subseteq X$ be a $\dfrac{1}{\tau^*(\mathcal{F})}$-net for $\mathcal{F}$ of size $\leq C d \tau^*(\mathcal{F}) \log \tau^*(\mathcal{F})$.

Then $\forall F \in \mathcal{F}$, $A \cap F \neq \emptyset$, i.e., A is a transversal for $\mathcal{F}$ of desired size. ∎

<u>Def.</u>

Let $(X, \mathcal{F})$ be a set system.

$(X^*, \mathcal{F}^*)$ is called the <u>dual set system</u>, where

$X^* = \mathcal{F}$ and $\mathcal{F}^* = \{\mathcal{F}_x : x \in X\}$ with $\mathcal{F}_x := \{F \in \mathcal{F} : x \in F\}$.

The <u>dual VC-dimension</u> of $\mathcal{F}$, <u>$VC^*(\mathcal{F})$</u> is $VC(\mathcal{F}^*)$.

<u>Remark.</u>

(1) Given $(X, \mathcal{F})$, the <u>incidence matrix</u> M is defined as an $|X| \times |\mathcal{F}|$-matrix s.t. $M_{x,F} = \begin{cases} 1 & \text{if } x \in F, \\ 0 & \text{if } x \notin F. \end{cases}$

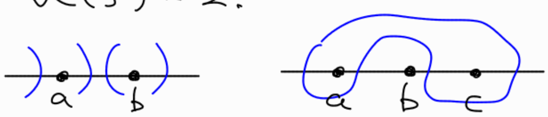
(2) If the incidence matrix of $(X', \mathcal{F}')$ is the same with M of $(X, \mathcal{F})$,
(up to permutations of rows/columns),
then $(X, \mathcal{F})$ and $(X', \mathcal{F}')$ are "essentially" the same.

(3) Given $(X, \mathcal{F})$ with incidence matrix M,
the incidence matrix of $(X^*, \mathcal{F}^*)$ is M^T.
($\because$ $x \in F$ iff $F \in \mathcal{F}_x$.)

<u>Example.</u>

Let $X = \mathbb{R}$, $\mathcal{F} = \{(-\infty, a) : a \in \mathbb{R}\} \cup \{(b, \infty) : b \in \mathbb{R}\}$.

Then $VC(\mathcal{F}) = 2$.

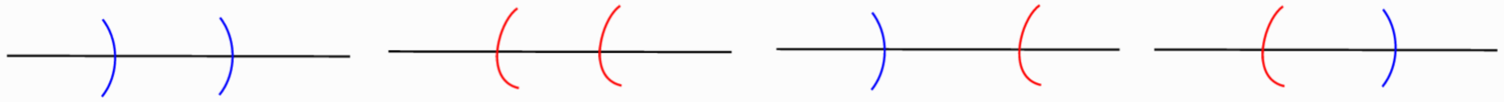
$\because$ 

$X^* = \{(-\infty, a) : a \in \mathbb{R}\} \cup \{(b, \infty) : b \in \mathbb{R}\}$,

$\mathcal{F}^* = \Big\{ \{(-\infty, a) : c \in (-\infty, a)\} \cup \{(b, \infty) : c \in (b, \infty)\} : c \in \mathbb{R} \Big\}$.

$\ni \mathcal{F}_1 = \{ \ldots, (-\infty, 10), \ldots, (-\infty, 3\sqrt{2}), \ldots, (-\infty, 1.01), \ldots,$
$\ldots, (-10^{10^{10}}, \infty), \ldots, (0.9999, \infty), \ldots \}$

$VC^*(\mathcal{F}) = 1$.



Proposition.

Let $(X, \mathcal{F})$ be a set system.

Then $VC^*(\mathcal{F}) < 2^{VC(\mathcal{F})+1}$ and $VC(\mathcal{F}) < 2^{VC^*(\mathcal{F})+1}$.

(In particular, $VC(\mathcal{F}) < \infty$ iff $VC^*(\mathcal{F}) < \infty$.)

Proof.

Assume $VC(\mathcal{F}) \geq 2^n$. Then $\underbrace{\exists A \subseteq X \text{ shattered by } \mathcal{F}}_{\text{possible}}$ and $|A| = 2^n$.

Write $A = \{a_C : C \subseteq [n]\}$.

$\forall k \in [n]$, let $F_k \in \mathcal{F} = X^*$ s.t. $A \cap F_k = \{a_C : k \in C \subseteq [n]\}$.

Then $\{F_1, \ldots, F_n\} \subseteq X^*$ is shattered by $\mathcal{F}^*$.

$(\because \forall 1 \leq i_1 < \cdots < i_\ell \leq n,\ \{F_1, \ldots, F_n\} \cap \underbrace{\mathcal{F}_{a_{\{i_1, \ldots, i_\ell\}}}}_{\text{collection of } F\text{'s s.t. } F \ni a_{\{i_1, \ldots, i_\ell\}}} = \{F_{i_1}, \ldots, F_{i_\ell}\}.)$ ∎

Fact. (Helly's theorem)

Let $X = \mathbb{R}^d$, $\mathcal{F}$ be some finite family of convex subsets of $\mathbb{R}^d$.

Assume that $\forall \mathcal{F}' \subseteq \mathcal{F}$ s.t. $|\mathcal{F}'| = d+1$, $\bigcap \mathcal{F}' \neq \emptyset$.

Then $\bigcap \mathcal{F} \neq \emptyset$ $(\Leftrightarrow \tau(\mathcal{F}) = 1)$.

Def.

Let $(X, \mathcal{F})$ be a set system. $\mathcal{F}$ has fractional Helly number k if

$\forall \alpha > 0, \exists \beta > 0$ s.t. $\forall F_1, \ldots, F_n \in \mathcal{F}$,

if at least α-fraction of elements $I \in \binom{[n]}{k}$ satisfy $\bigcap_{i \in I} F_i \neq \emptyset$,

then $\exists J \subseteq [n]$ s.t. $|J| \geq \beta n$ and $\bigcap_{j \in J} F_j \neq \emptyset$.

Remark.

(1) Helly's theorem does not hold if convexity is replaced with $VC(\mathcal{F}) \leq 2$.

(2) If $X = \mathbb{R}^d$, $\mathcal{F}$ is some family of convex subsets of $\mathbb{R}^d$, then $\mathcal{F}$ has fractional Helly number $d+1$ (Katchalski, Liu).

Theorem. (Matousek)

Let $(X, \mathcal{F})$ be a set system with $VC^*(\mathcal{F}) \leq k-1$.

Then $\mathcal{F}$ has fractional Helly number k.

Def.

The VC-density of $\mathcal{F}$, $vc(\mathcal{F}) = \limsup_{n\to\infty} \dfrac{\log(\pi_{\mathcal{F}}(n))}{\log n}$.

That is, infimum of r's s.t. $\pi_{\mathcal{F}}(n) = O(n^r)$.

(By SS lemma, $vc(\mathcal{F}) \leq VC(\mathcal{F})$.)

VC-codensity $vc^*(\mathcal{F}) = vc(\mathcal{F}^*)$.

Def.

Let $p \geq q$ be natural numbers and $(X, \mathcal{F})$ be a set system.

$(X, \mathcal{F})$ satisfies (p,q)-property if

$\forall \mathcal{F}_0 \subseteq \mathcal{F}$ s.t. $|\mathcal{F}_0| = p$, $\exists \mathcal{F}_1 \subseteq \mathcal{F}_0$ s.t. $|\mathcal{F}_1| = q$ and $\bigcap \mathcal{F}_1 \neq \emptyset$.

Remark.

Helly's theorem is equivalent to

$\left(\begin{array}{l} X = \mathbb{R}^d,\ \mathcal{F}: \text{finite family of convex subsets satisfies } (d+1, d+1)\text{-property} \\ \Rightarrow \bigcap \mathcal{F} \neq \emptyset \ (\Leftrightarrow \tau(\mathcal{F}) = 1). \end{array} \right)$

Fact. (Alon, Kleitman)

Let $p \geq q \geq d+1$ be natural numbers. Then $\exists N$ s.t.

$\mathcal{F}$: finite family of convex subsets of $\mathbb{R}^d$ satisfying (p,q)-property

$\Rightarrow \tau(\mathcal{F}) \leq N$.

<u>Theorem.</u> (Alon, Kleitman + Matousek)

Let $p \geq q \geq d+1$ be natural numbers, $\mathcal{F}$ satisfy (p,q)-property.

Then $\exists N = N(p,q,d)$ s.t. $\left(\mathcal{F} : \text{finite and } vc^*(\mathcal{F}) \leq d \Rightarrow \tau(\mathcal{F}) \leq N \right)$.

Main references:

- Artem Chernikov, Model theory and Combinatorics (draft)
- Pierre Simon, A Guide to NIP Theories