9. Extremal Set Theory

Extremal Combinatorics

set system (family) $\mathcal{F} \subseteq 2^{[n]}$ with ground set $[n]$.

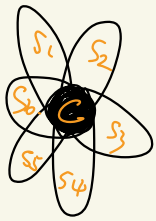
集族

Sunflowers

def $\mathcal{F} \subseteq 2^{[n]}$ is a sunflower of size r with center C

$$\begin{cases} |\mathcal{F}| = r \\ \forall S, T \in \mathcal{F}: \boxed{S \cap T = C.} \end{cases}$$

(can be ϕ)



Sunflower Lemma (Erdős–Rado, 1960)

$\mathcal{F} \subseteq \binom{[n]}{k}$. $\boxed{|\mathcal{F}| > k!(r-1)^k}$

given

$\exists$ a sunflower $G \subseteq \mathcal{F}$. s.t. $|G| = r$.

"hypergraph" 超图

$V: [n]$

$E: \mathcal{F}$.

Proof: induction

Basis: $k = 1$. $\mathcal{F} \subseteq \binom{[n]}{1}$ $|\mathcal{F}| > r-1$

$\exists\, r$ disjoint singletons

matching (匹配)

• $k \geq 2$:

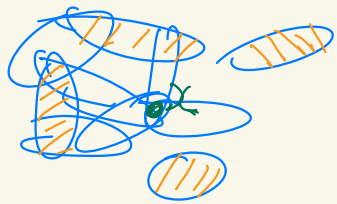
take largest $\boxed{G} \subseteq \mathcal{F}$. with disjoint members

$\forall S, T \in G$ that $S \neq T$ $S \cap T = \phi$

G中而那边没有公共顶点

Case 1: $|G| \geq r$. G is a sunflower of size r.

Case 2: $|G| \leq r-1$:



Goal: find a "popular" $x \in [n]$

$$\{S \in \mathcal{F} \mid x \in S\}$$

remove x $\longrightarrow$ $\boxed{\mathcal{H}} = \{S \setminus \{x\} \mid S \in \mathcal{F} \wedge x \in S\}$

$\mathcal{F} \subseteq \binom{[n]}{k}$ $\mathcal{H} \subseteq \binom{[n]}{k-1}$

归纳假设

If $|\mathcal{H}| > (k-1)!\,(r-1)^{k-1}$ $\Rightarrow$ $\boxed{\text{I.H.}}$

$\Rightarrow$ $\mathcal{H}$ contains a sunflower of size r.

Adding x back $\longrightarrow$ $\{S \cup \{x\} \mid S \in \mathcal{H}\} \subseteq \mathcal{F}$.

contains a sunflower.

$\Rightarrow$ $\mathcal{F}$ contains a sunflower.

$\mathcal{G}$: maximal matching.

$|\mathcal{G}| \le r-1$

$Y = \bigcup_{S \in \mathcal{G}} S$

"超边"

$\Rightarrow |Y| \le k(r-1)$

$\mathcal{F} \subseteq \binom{[n]}{3}$

$S_1 = \{1, 2, 4\}$.

$S_2 = \{2, 3, 4\}$.

$S_3 = \{3, 6, 7\}$.

$\Rightarrow Y = \{1, 2, 3, 4, 6, 7\}$

Claim: $\boxed{\forall S \in \mathcal{F} : Y \cap S \neq \phi}$

If otherwise:

$\exists T \in \mathcal{F} : Y \cap T = \phi$

$\Rightarrow \forall S \in \mathcal{G} : S \cap T = \phi$

(T is disjoint with all $S \in \mathcal{G}$)

Contradiction; ($\mathcal{G}$ is maximal)

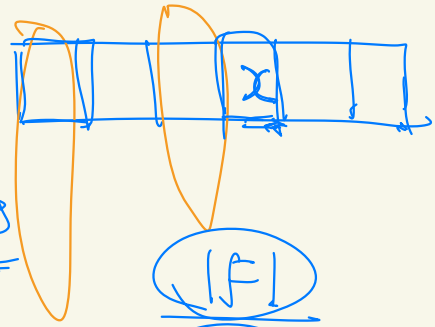
$\exists x \in Y$:

of $S \in \mathcal{F}$ containing x:

$|\{S \in \mathcal{F} \mid x \in S\}| \geq \frac{|\mathcal{F}|}{|Y|}$ (pigeonhole)

popular x.

$\mathcal{F} \subseteq \binom{[n]}{k}$, average



$|\mathcal{F}| > k!(r-1)^k$. 题目条件

$|Y| \le k\cdot(r-1)$

$|\mathcal{F}|$

$|Y|$

$$\Rightarrow |\{S\in\mathcal{F} \mid x\in S\}| \ge \frac{|\mathcal{F}|}{|Y|} > \frac{k!(r-1)^k}{k(r-1)}$$

$$= (k-1)!(r-1)^{k-1}.$$

第k个product.

$\|$

$|\mathcal{H}| > (k-1)!(r-1)^{k-1}$ $\Rightarrow$ I.H. $\square$

$|\mathcal{F}| > k!(r-1)^k$.

Sunflower Conjecture $\mathcal{F} \subseteq \binom{[n]}{k}$ $|\mathcal{F}| > (C(r))^k$.

$\$2500$ $\exists$ a sunflower $\mathcal{G} \subseteq \mathcal{F}$ s.t. $|\mathcal{G}| = r$

indep. on. k.

$O(n^3) \to O(n^{\lg 7})$ $\boxed{O(n^2)}$

Intersecting Family $\mathcal{F} \subseteq \binom{[n]}{k}$

所有都两两相交

Intersecting: $\forall S,T\in\mathcal{F}$, $S\cap T \ne \phi$

(trivially holds if $n < 2k$)

Erdős-Ko-Rado Theorem.

Ko Chao 柯召

$\boxed{\mathcal{F} \subseteq \binom{[n]}{k}}$ $n \geq 2k$.

$\forall S, T \in \mathcal{F}: S \cap T \neq \phi \Rightarrow \boxed{|\mathcal{F}| \leq \binom{n-1}{k-1}}$

($\mathcal{F}$ is intersecting)

Proof: (induction on n and k)

wlog. n.

$\mathcal{F}_0 = \{S \in \mathcal{F} \mid n \notin S\}$ ①

$\mathcal{F}_1 = \{S \in \mathcal{F} \mid n \in S\}$ ②

bipartition

① $\mathcal{F}_0 \subseteq \binom{[n-1]}{k}$ $S, T \in \mathcal{F}_0$ $S \cap T \neq \phi$

$\Rightarrow$ Induction Hypothesis: $|\mathcal{F}_0| \leq \binom{n-2}{k-1}$

② $\mathcal{F}_1 = \{S \in \mathcal{F} \mid n \in S\}$

$\Rightarrow \mathcal{F}_1' = \{S \setminus \{n\} \mid S \in \mathcal{F}_1\} \subseteq \binom{[n-1]}{k-1}$

△ If $\mathcal{F}_1'$ intersecting $\Rightarrow$ $|\mathcal{F}_1'| \overset{\text{I.H.}}{\leq} \binom{n-2}{k-2}$

$\Rightarrow |\mathcal{F}| = |\mathcal{F}_0| + |\mathcal{F}_1|$

$$= |\mathcal{F}_0| + |\mathcal{F}_1'|$$

$$\leq \binom{n-2}{k-1} + \binom{n-2}{k-2} = \binom{n-1}{k-1} \qquad \square$$

Key: $\mathcal{F}_1'$ intersecting? by shifting (compression)

Goal: find a special $\mathcal{F} \subseteq \binom{[n]}{k}$

s.t. $\mathcal{F}$ remains intersecting after deleting n.

---

def (Shift Operator) $\mathcal{F} \subseteq 2^{[n]}$. for $1 \leq i < j \leq n$,

$\forall T \in \mathcal{F}$: $\boxed{T_{ij} = (T \setminus \{j\}) \cup \{i\}}$

(i,j)-shift: $\boxed{S_{ij}(\cdot)}$:

i, j

$$\forall T \in \mathcal{F}: \boxed{S_{ij}(T)} = \begin{cases} \boxed{T_{ij}} & \text{if } \boxed{j \in T,\ i \notin T,\ T_{ij} \notin \mathcal{F}.} \\ T. & \text{otherwise} \end{cases}$$

$$S_{ij}(\mathcal{F}) = \{ S_{ij}(T) \mid T \in \mathcal{F} \}$$

Proposition: 1) $|S_{ij}(T)| = |T|$

$|S_{ij}(\mathcal{F})| = |\mathcal{F}|$ ✓

2) If $\mathcal{F}$ is intersecting, $S_{ij}(\mathcal{F})$ is intersecting.

$S_{ij}(A) \cap S_{ij}(B) \neq \phi$

Proof: $A, B \in \mathcal{F}$, $\boxed{A \cap B \neq \phi}$.

- $j \notin A \cap B$. $\Rightarrow$ safe.
- $j \in A \cap B$.
  - A successful B successful ✓
  - A failed. B failed. ✓
  - (wlog) $\boxed{\text{A failed B successful.}}$

$A_{ij} = (A \setminus \{j\}) \cup \{i\}$ $B_{ij} = (B \setminus \{j\}) \cup \{i\}$

Unsafe Case: $\boxed{A \cap B = \{j\}}$ $\boxed{\text{A failed}}$ B successful

① $i \in A$. $S_{ij}(A) = A$.

$S_{ij}(B) = B_{ij}$

$i \notin B$

$B_{ij} \notin \mathcal{F}$.

$S_{ij}(A) \cap S_{ij}(B) = A \cap B_{ij} = \{i\} \neq \phi.$

~~②~~ $\boxed{A_{ij} \in \mathcal{F}}$ $i \notin A.$ $S_{ij}(A) \cap S_{ij}(B)$

$= A \cap B_{ij} = \phi.$

$A_{ij} \cap B$

$= \underbrace{(A \setminus \{j\} \cap B)}_{\phi} \cup \underbrace{(\{i\} \cap B)}_{\phi} = \phi$

$\Rightarrow$ contradiction !

($\mathcal{F}$ is intersecting)

$\Rightarrow S_{ij}(\mathcal{F})$ is intersecting $\square$

Repeatedly apply (i,j)-shifting $S_{ij}(\mathcal{F})$

for $1 \leq i < j \leq n$.

$\mathcal{F} \to S_{ij}(\mathcal{F}) \to \cdots$

Eventually: $\mathcal{F} = S_{ij}(\mathcal{F})$ $\forall\, 1 \leq i < j \leq n$,

"fixed point" $\Rightarrow$ "$\mathcal{F}$ is shifted"

Wlog: $\boxed{\text{assume } \mathcal{F} \text{ is shifted}}$ ☆

Recall: $\mathcal{F} \subseteq \binom{[n]}{k}$, $n \geq 2k$, $\mathcal{F}$ intersecting

$$\Rightarrow |\mathcal{F}| \leq \binom{n-1}{k-1}$$

When $n=2k$: $\forall S \in \binom{[n]}{k}$

S, $[n] \setminus S$: at most one in $\mathcal{F}$.

$$\Rightarrow |\mathcal{F}| \leq \frac{1}{2}\binom{n}{k} = \frac{1}{2} \cdot \frac{n!}{k!(n-k)!}$$

$$= \frac{\cancel{n} \cdot (n-1)!}{(\cancel{2k}) \cdot (k-1)!(n-k)!} = \binom{n-1}{k-1}$$

When $n>2k$:

(i,j)

arbitrary intersecting $\mathcal{F}$ $\xrightarrow{\text{shifting}}$ shifted $\mathcal{F}'$!

$$|\mathcal{F}| \leq \binom{n-1}{k-1} \Longleftarrow |\mathcal{F}'| \leq \binom{n-1}{k-1}$$

$\mathcal{F}_1 = \{ S \in \mathcal{F} \mid n \in S \}$. $\mathcal{F}_1' = \{ S \setminus \{n\} \mid S \in \mathcal{F}_1 \}$

Now: $\mathcal{F}_1'$ is intersecting

Proof: Assume $\mathcal{F}_1'$ is non-intersecting.

$\exists\ A, B \in \mathcal{F}$ s.t. $\boxed{A \cap B = \{n\}}$

$\Rightarrow |A \cup B| = |A| + |B| - |A \cap B|$

$= 2k - 1 < n - 1$

$\exists\ i : 1 \le i \le n$ s.t. $i \in A \cup B$,

$C = A_{in} = (A \setminus \{n\}) \cup \{i\} \in \mathcal{F}$.

$\mathcal{F}$ is shifted

$\underset{\in \mathcal{F}}{A_{in}} \cap \underset{\in \mathcal{F}}{B} = \phi$

$\Rightarrow \mathcal{F}$ is NOT intersecting.

CONTRADICTION ! $\square$

Erdös-Ko-Rado: $\mathcal{F}$ is intersecting $\Rightarrow |\mathcal{F}| \le \binom{n-1}{k-1}$

Katona's Proof (1972):

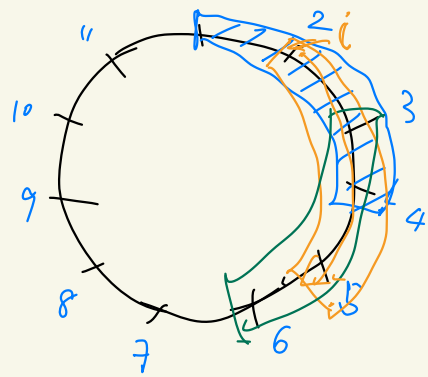
$$\begin{matrix} 1 & 2 & 3 & \cdots & n-1 & n \\ 2 & 3 & 4 & \cdots & n & 1 \end{matrix}$$

π: cyclic permutation. totally $(n-1)$!

k-arc: $G_\pi = \{ \{ \pi_{(i+j) \bmod n} \mid j \in [k] \} \mid i \in [n] \}$

Lemma $n \geq 2k$ $\forall \pi$.

$$|G_\pi \cap \mathcal{F}| \leq k.$$

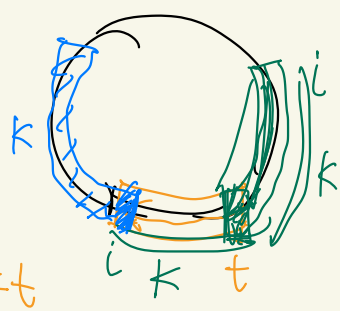


Proof: $A_i = \{ \pi_{(i+j) \bmod n} \mid j \in [k] \}$ k-arc.

$$\boxed{G_\pi = \{ A_i \mid i \in [n] \}}$$

Suppose $A_t \in \mathcal{F}$. $\mathcal{F}$ is intersecting

If $A_i \in \mathcal{F}$: $A_i \cap A_t \neq \phi$



$t-(k-1) \leq i \leq t+(k-1)$

$i \neq t$

$$\boxed{A_i \cap A_{i+k} = \phi}$$ $(n \geq 2k)$

$(k-1)$ pairs $\{A_i, A_{i+k}\}$ $\quad t-(k-1) \le i \le t$

$\mathcal{F}$ contains at most one of them.

$$|\mathcal{F} \cap \mathcal{G}_\pi| \le \underbrace{(k-1)}_{\text{one in pairs}} + \underbrace{1}_{A_t} = k. \quad \square$$

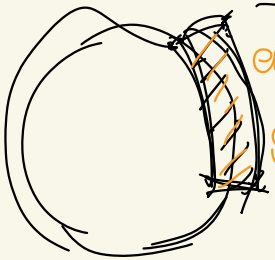
(double counting)

$$R = \{(S, \pi) \mid \pi \text{ is a cyclic permutation}, \ S \in \mathcal{F} \cap \mathcal{G}_\pi\}$$

Method 1: $|R| \le (n-1)! \cdot k$

Method 2: $\boxed{\forall S \in \mathcal{F}:}$ S is continuous, under π.

(S is a k-arc)



elements in S

$\#: k! \cdot (n-k)!$

$$\Rightarrow |R| = |\mathcal{F}| \cdot k! \cdot (n-k)!$$

$$\Rightarrow \quad |\mathcal{F}| \cdot k!(n-k)! \le (n-1)! \cdot k$$

$$\Rightarrow \quad |\mathcal{F}| \le \frac{(n-1)!}{(k-1)!(n-k)!} = \binom{n-1}{k-1} \qquad \square$$

Sperner System (Antichain)

$\boxed{\mathcal{F} \sqsubseteq 2^{[n]}}$

partial order 偏序 "$\sqsubseteq$" $\rightarrow$ POSET

chain: $S_1 \sqsubseteq S_2 \sqsubseteq \cdots \sqsubseteq S_k$

antichain 反链: $S, T \sqsubseteq \mathcal{F}$, $S \neq T$: $S \not\sqsubseteq T$.

k-uniform family: $\binom{[n]}{k}$ is an antichain.

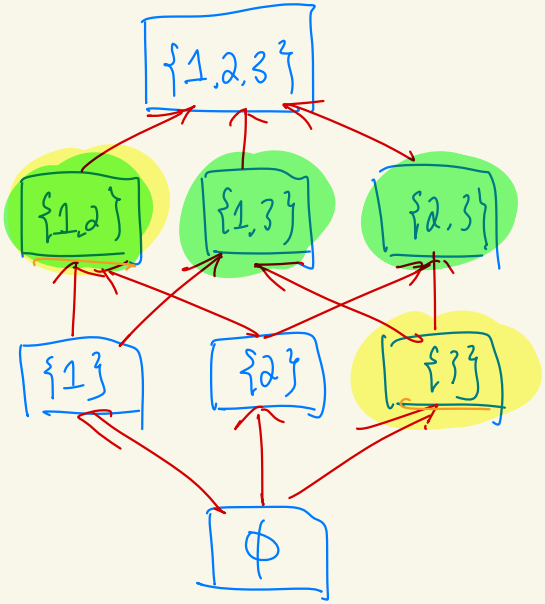
Q: How large $\mathcal{F}$ can be?

$\boxed{\binom{n}{k}}$ is maximized when $k = \lfloor \frac{n}{2} \rfloor$

Theorem (Sperner 1928) $\mathcal{F} \sqsubseteq 2^{[n]}$

$\mathcal{F}$ is an antichain $\Rightarrow$ $|\mathcal{F}| \le \binom{n}{\lfloor n/2 \rfloor}$

maximize



$\{3\}$

"shifting" $\to$ $\{1,3\}$, $\{2,3\}$

$\to$ shade

$\to$ shadow.

$\mathcal{F} \subseteq \binom{[n]}{k}$ k-uniform family

Shade $\nabla\mathcal{F} = \left\{ T \in \binom{[n]}{k+1} \;\middle|\; \exists S \in \mathcal{F} \text{ s.t. } S \subseteq T \right\}$

Shadow $\Delta\mathcal{F} = \left\{ T \in \binom{[n]}{k-1} \;\middle|\; \exists S \in \mathcal{F} \text{ s.t. } T \subseteq S \right\}$

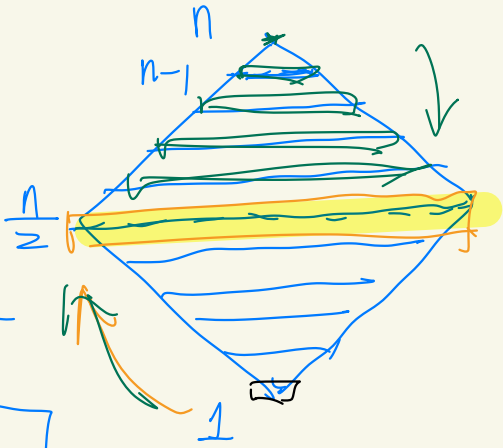
Lemma (Sperner) $|\nabla\mathcal{F}| \ge \frac{n-k}{k+1} \cdot |\mathcal{F}|$, $|\Delta\mathcal{F}| \ge \frac{k}{n-k+1} \cdot |\mathcal{F}|$.

北半球: shadow 变大 $\binom{[n]}{k}$

$$\boxed{|\Delta \mathcal{F}| \geq |\mathcal{F}| \quad \left(k \geq \frac{n+1}{2}\right)}$$

南半球: shade 变大

$$\boxed{|\nabla \mathcal{F}| \geq |\mathcal{F}| \quad \left(k \leq \frac{n-1}{2}\right)}$$



Proof: $R = \{(S,T) \mid \underset{k}{S} \in \mathcal{F},\ \underset{k+1}{T} \in \nabla\mathcal{F},\ S \subseteq T\}$

① For any $S \in \mathcal{F}$: $n-k$ $T \in \nabla\mathcal{F}$ s.t. $S \subseteq T$.

$\Rightarrow |R| = |\mathcal{F}| \cdot (n-k)$

② For any $T \in \mathcal{F}$: $\leq (k+1)$ $S \in \mathcal{F}$ s.t. $S \subseteq T$

deleting some of the elements might be not in $\mathcal{F}$ anymore.

$\Rightarrow |R| \leq |\nabla\mathcal{F}| \cdot (k+1)$

$\Rightarrow$ Double Counting: $|\nabla\mathcal{F}| \geq |\mathcal{F}| \cdot \frac{n-k}{k+1}$

$\Rightarrow$ When $k \leq \frac{n-1}{2}$, $|\nabla\mathcal{F}| \geq |\mathcal{F}|$. $\square$

Proposition: $\mathcal{F} \subseteq 2^{[n]}$, $\mathcal{F}_k = \mathcal{F} \cap \binom{[n]}{k}$

let k_{min} be the smallest k such that $\boxed{|\mathcal{F}_k| > 0}$.

$$\mathcal{F}' = \begin{cases} \mathcal{F} \setminus \mathcal{F}_{k_{min}} \cap \boxed{\nabla \mathcal{F}_{k_{min}}} & \text{if } k_{min} \leq \frac{n-1}{2} \\ \mathcal{F} & \text{otherwise.} \end{cases}$$

- k_{max} be the largest k such that $|\mathcal{F}_k| > 0$

$$\mathcal{F}'' = \begin{cases} \mathcal{F} \setminus \mathcal{F}_{k_{max}} \cap \Delta \mathcal{F}_{k_{max}} & \text{if } k_{max} \geq \frac{n+1}{2} \\ \mathcal{F} & \text{otherwise.} \end{cases}$$

$\Longrightarrow$ $\boxed{|\mathcal{F}'| \geq |\mathcal{F}|, \quad |\mathcal{F}''| \geq |\mathcal{F}|}$ (non-decreasing)

$\boxed{\text{If } \mathcal{F} \text{ is an antichain, then } \mathcal{F}', \mathcal{F}'' \text{ so.}}$

Proof: By contradiction, $\mathcal{F}'$ is NOT an antichain.

$\Rightarrow \exists\, S, T \in \mathcal{F}'$ s.t. $S \subseteq T$.

$\triangle$ $\boxed{\text{One of } S, T}$ must be in $\nabla \mathcal{F}_{k_{min}}$.

$\boxed{S \in \nabla \mathcal{F}_{k_{min}}}$, $T \in \mathcal{F}$

$\Rightarrow \exists\, R \in \mathcal{F}_{k_{min}} \subseteq \mathcal{F}$.

s.t. $R \subseteq S \subseteq T$. Contradiction

($\mathcal{F}$ is an antichain!)

Proof of Sperner's Theorem:

$$\mathcal{F}_k = \mathcal{F} \cap \binom{[n]}{k} = \{ S \in \mathcal{F} \mid |S| = k \}$$

- for the smallest k that $\begin{cases} |\mathcal{F}_k| > 0 \\ k \leq \frac{n-1}{2} \end{cases}$

  $\Rightarrow$ replace $\mathcal{F}_k$ by $\nabla \mathcal{F}_k$.

- for the largest k that $\begin{cases} |\mathcal{F}_k| > 0 \\ k \geq \frac{n+1}{2} \end{cases}$

  $\Rightarrow$ replace $\mathcal{F}_k$ by $\Delta \mathcal{F}_k$

$\Rightarrow \boxed{\mathcal{F} \subseteq \binom{[n]}{\lfloor n/2 \rfloor}}$

$\Downarrow$

$|\mathcal{F}| \leq \binom{n}{\lfloor n/2 \rfloor}$ $\square$

Proof by Lubell (1966):

Π : permutation of $[n]$

$S \subseteq [n]$ prefixes Π: $S = \{ \Pi_1, \Pi_2, \cdots, \Pi_{|S|} \}$

$1, 2, 3, \cdots, n$

$\rightarrow \boxed{\pi_1, \pi_2, \pi_3, \pi_4, \cdots, \pi_{|S|}} \cdots, \pi_n$

S

Given S. # of π prefixed by S: $|S|!\,(n-|S|)!$

$\mathcal{F}$ is an anti-chain: $\forall S, T \in \mathcal{F}$, $S \not\subseteq T$

$\Rightarrow$ No permutation π can be prefixed by ≥ 2 members in $\mathcal{F}$.

$$\boxed{\sum_{S \in \mathcal{F}} |S|!\,(n-|S|)!} \leq n!$$

counting permutation π

$$\sum_{S \in \mathcal{F}} \frac{|S|!\,(n-|S|)!}{n!} = \boxed{\sum_{S \in \mathcal{F}} \frac{1}{\binom{n}{|S|}} \leq 1}$$

$$\binom{n}{|S|} \leq \binom{n}{\lfloor n/2 \rfloor}$$

LYM inequality

$$\Rightarrow \sum_{S\in\mathcal{F}} \frac{1}{\binom{n}{|S|}} \le |\mathcal{F}| \cdot \frac{1}{\binom{n}{\lfloor n/2 \rfloor}} \le 1 \Rightarrow \boxed{|\mathcal{F}| \le \binom{n}{\lfloor \frac{n}{2} \rfloor}} \quad \square$$

Lubell, Yamamoto, Maschalkin

↳ LYM inequality $\quad \mathcal{F} \subseteq 2^{[n]}$ is an antichain

then $\displaystyle\sum_{S\in\mathcal{F}} \frac{1}{\binom{n}{|S|}} \le 1$

Proof: (Alon) Probabilistic Method

π : uniform random permutation on $[n]$

$$\boxed{C_\pi = \{ \{\pi_1\}, \{\pi_1, \pi_2\}, \cdots, \{\pi_1, \pi_2, \cdots, \pi_n\} \}}$$

(maximal-chain)

$\forall S \in \mathcal{F}$: $\quad X_S = \mathbb{1}\{S \in C_\pi\}$

$$\Rightarrow \mathbb{E}[X_S] = \Pr(S \in C_\pi) = \frac{1}{\binom{n}{|S|}} \quad \text{(uniform)}$$

$$X = \sum_{S\in\mathcal{F}} X_S \Rightarrow \mathbb{E}[X] = \sum_{S\in\mathcal{F}} \frac{1}{\binom{n}{|S|}} \le 1.$$

$$X = |\mathcal{F} \cap G_\pi| \le 1. \qquad \square$$

anti-chain

Sperner's Theorem

Sperner's Proof (shade & shadow)

LYM inequality

- Lubell's Proof (counting)
- Alon's Proof. (probabilistic method)

Sauer's Lemma & VC-Dimension

Learning Theory

$\mathcal{F} \subseteq 2^X$ set family

def (Shattering) $R \subseteq X$

"打散"

trace of $\mathcal{F}$ on R: $\mathcal{F}|_R = \{ S \cap R \mid S \in \mathcal{F} \}$

size

$\mathcal{F}$ shatters R: $\boxed{\mathcal{F}|_R = 2^R}$ $\Leftrightarrow |\mathcal{F}|_R| = 2^{|R|}$

i.e. $\forall T \subseteq R, \ \exists S \in \mathcal{F}$ s.t. $T = S \cap R$

Machine Learning

Learning Theory

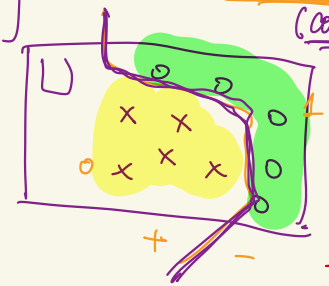
instance $\boxed{x}$ → model f → prediction $f(x)$ ~ label $y \in \{0,1\}$

$f \in \mathcal{F}$

hypothesis space (concept)

binary Classification

R: $\boxed{n \text{ instances}}$



Set family

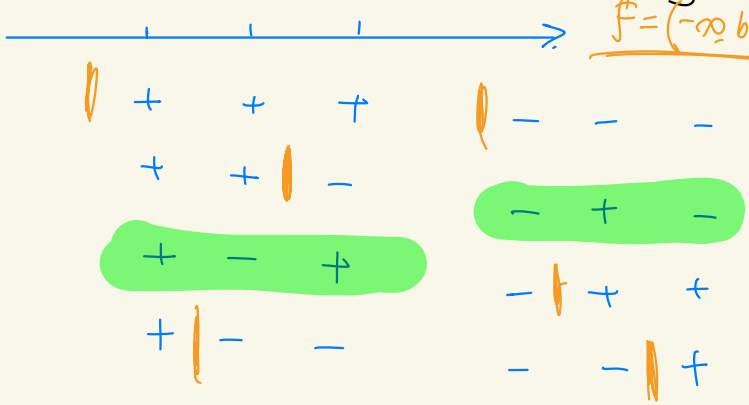
dichotomy 对分

$S \cap R = T$

e.g. $n=3$ in $(-\infty, +\infty)$

$\boxed{\mathcal{F} := [a, \infty)}$

$\mathcal{F} = (-\infty, b] \cup [a, +\infty)$



Complexity Measure 复杂性度量

$\|f(x) - y\| \le C(\mathcal{F}) + \text{loss}$

$\boxed{\text{generalization bound}}$

PAC Learning — Leslie Valient

$\Pr(|f(x)-y|<\varepsilon) > 1-\delta$

def (VC-Dimension)

Vapnik Chervonenkis

$\mathcal{F} \subseteq 2^X$. VC-dim$(\mathcal{F})$:

Size of largest $R \subseteq X$

s.t. $\mathcal{F}$ can shatter R.

$\boxed{S} \subseteq X$ → Characteristic function

反例 样本

$\boxed{f_S}: X \to \{0,1\}$ $\quad f_S(x) = \mathbb{1}\{x \in S\}$

model / concept / hypothesis

$\mathcal{F} = \{\boxed{f_S} \mid S \in \mathcal{F}\}$

$R \uparrow \quad \mathcal{F} \uparrow$ (complexity)

Sauer's Lemma:

Thm Let $\mathcal{F} \subseteq 2^{[n]}$. If $|\mathcal{F}| > \boxed{\sum_{1 \le i < k} \binom{n}{i}}$

$\Rightarrow \exists R \in \binom{[n]}{k}$ such that $\boxed{\mathcal{F} \text{ shatters } R}$.

$$\boxed{\Rightarrow |\mathcal{F}| > \sum_{1 \le i < k} \binom{n}{i} \Rightarrow \text{VC-dim}(\mathcal{F}) \ge k}$$

growth function

Sauer's Lemma

Shelah-Perles

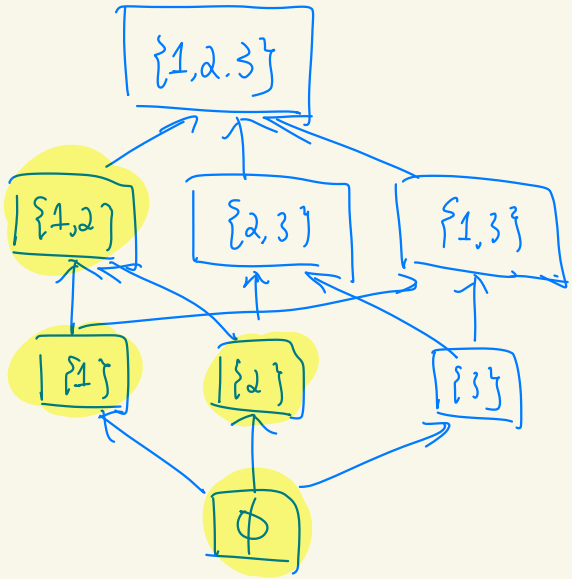
Vapnik-Chervonenkis

Proof by induction (See FML Chapter 3.3)

Proof by shifting

Heredity 遗传性 继承 (ideal. abstract simplicial complex.)

集族 $\mathcal{F} \subseteq 2^X$. $\forall S \subseteq T \in \mathcal{F} \Rightarrow S \in \mathcal{F}$.



Thm $\mathcal{F}$ is a hereditary family.

$R \in \mathcal{F} \Rightarrow \mathcal{F}$ shatters R

existence

Contradiction ?

$$|\mathcal{F}| \le \left| \bigcup_{1 \le i \le k} \binom{X}{i} \right|$$

$$= \sum_{1 \le i \le k} \binom{n}{i}$$

For hereditary families, we only need to show that

$\exists$ an $R \in \mathcal{F}$ of size $|R| \ge k$.

Sauer's Lemma ✓ for hereditary family

? for non-hereditary family

reduction (shifting)

Pipeline :

$$\boxed{\text{arbitrary } \mathcal{F}} \xrightarrow[|\mathcal{F}'| \geq |\mathcal{F}|]{\text{shifting}} \text{hereditary } \mathcal{F}'$$

$$\text{hereditary } \mathcal{F}' \xrightarrow{\text{VC-dim}(\mathcal{F}) \geq \text{VC-dim}(\mathcal{F}')} \mathcal{F} \text{ shatters a } k\text{-set}$$

$$\mathcal{F}' \text{ shatters a } k\text{-set} \longrightarrow \mathcal{F} \text{ shatters a } k\text{-set}$$

Down-Shift operator S_i

- for any $T \in \mathcal{F}$:

$$S_i(T) = \begin{cases} T \setminus \{i\} & \text{if } i \in T,\ T \setminus \{i\} \notin \mathcal{F} \\ T & \text{otherwise} \end{cases}$$

- $S_i(\mathcal{F}) = \{ S_i(T) \mid T \in \mathcal{F} \}$

$\Rightarrow$ Repeatedly Applying S_i to $\mathcal{F}$ for all $i \in [n]$

finiteness $\Rightarrow$ eventually : down-shifted

$S_i(\mathcal{F}) = \mathcal{F}$.

Proposition : 1) If $\mathcal{F}$ is down-shifted $\Rightarrow$ $\mathcal{F}$ is hereditary.

2) $|S_i(\mathcal{F})| = |\mathcal{F}|$

3) $|S_i(\mathcal{F})|_R| \leq |\mathcal{F}|_R|$

If $S_i(\mathcal{F})$ shatters R
$\Rightarrow$ $\mathcal{F}$ shatters R.

(VC-dimension)

$\hookrightarrow$ $A \in S_i(\mathcal{F})$ $\Rightarrow \begin{cases} \text{Case 1: } A = S_i(A \cup \{i\}) \\ \text{Case 2: } A = S_i(A) \end{cases}$

$\Rightarrow A \cap R \in S_i(\mathcal{F}|_R)$

After down-shifting

$\mathcal{F}'$ is hereditary. $\Rightarrow \exists S \in \binom{[n]}{\ell}$ with $\ell \geq k$, $2^S \subseteq \mathcal{F}$.

$|\mathcal{F}| > \sum_{1 \leq i < k} \binom{n}{i}$

$\parallel$

$|\mathcal{F}|$

$R \in \binom{S}{k} \Rightarrow \mathcal{F}$ shatters R.

Kruskal-Katona Theorem

k-uniform family

def (shadow) Let $\mathcal{F} \subseteq \binom{[n]}{k}$. the shadow of $\mathcal{F}$ is defined

as: $\boxed{\Delta \mathcal{F}} = \left\{ T \in \binom{[n]}{k-1} \;\middle|\; \exists S \in \mathcal{F} \text{ s.t. } T \subseteq S \right\}$

Compression : How small can the shadow $\Delta \mathcal{F}$ be ?

k-cascade representation of m

$\exists^{!}$ unique representation :

$$m = \boxed{\binom{m_k}{k}} + \binom{m_{k-1}}{k-1} + \cdots + \binom{m_t}{t} = \sum_{l=t}^{k} \binom{m_l}{l}$$

where ① $m_k > m_{k-1} > \cdots > m_t \geq t \geq 1$.

e.g. $m = 8$, $k = 3$.

$$8 = \underset{\downarrow \atop 4}{\binom{4}{3}} + \underset{\downarrow \atop 3}{\binom{3}{2}} + \underset{\downarrow \atop 1}{\binom{1}{1}}$$

$l = k$.

while $m > 0$ do:

Greedy

find m_l : largest number s.t. $\binom{m_l}{l} \leq m$

$m = m - m_l$.

$l = l - 1$

① If $m_{l-1} \geq m_l$:

$$m \geq \binom{m_l}{l} + \binom{m_{l-1}}{l-1} \geq \binom{m_l}{l} + \binom{m_l}{l-1} = \binom{m_l+1}{l}$$

Contradiction !

② Suppose $m = \sum_{l=t}^{k} \binom{a_l}{l} = \sum_{l=r}^{k} \binom{b_l}{l}$.

WTS : $a_k = b_k = m_k$.

If $a_k < b_k$: $m = \binom{a_k}{k} + \binom{a_{k-1}}{k-1} + \cdots + \binom{a_t}{t}$

$$\le \binom{a_k}{k} + \binom{a_k-1}{k-1} + \cdots + \binom{a_k-k+1}{1} + 1 - 1$$

$$\boxed{\binom{n}{k} + \binom{n}{k-1} = \binom{n+1}{k}}$$

$$\binom{a_k-k+1}{0}$$

$$\le \binom{a_k+1}{k} - 1 < \binom{b_k}{k} \le m$$

$\Rightarrow$ Contradiction !

$\Rightarrow a_k = b_k. \quad \Rightarrow \quad a_i = b_i \quad \forall i \in [k].$

Co-lexicographic order of subsets

(colex)

$A, B \subseteq [n], \ A \neq B$

$A < B$ if $\max A \setminus B < \max B \setminus A$

e.g. $\binom{[5]}{3}$

$\{3, 2, 1\}$

$\{4, 2, 1\}$

$\{4, 3, 1\}$

$\{4, 3, 2\}$

$\{5, 2, 1\}$

$\{5, 3, 1\}$

$\{5, 3, 2\}$

$\binom{[n]}{k} = \binom{[n-1]}{k}$ + prefix

$\binom{[3]}{3} = \binom{[4]}{3}$ prefix

$R(m,k)$: first m members of $\binom{\mathbb{N}}{k}$ in colex order

set family

$\{5,4,1\}$

$\{5,4,2\}$

$\{5,4,3\}$

$\vdots$

Claim: $R\left(\binom{n}{k}, k\right) = \binom{[n]}{k}$

$\Rightarrow R(m,k)$ highly overlapped !

extremal set ! 极值集合 ☆

Thm $m = \binom{m_k}{k} + \binom{m_{k-1}}{k-1} + \cdots + \binom{m_t}{t}$ cascade

$$|\Delta R(m,k)| = \binom{m_k}{k-1} + \binom{m_{k-1}}{k-2} + \cdots + \binom{m_t}{t-1}$$

Proof: construct $R(m,k)$

- $\binom{[m_k]}{k}$

$$8 = \overset{m_k}{\binom{4}{3}} + \overset{m_{k-1}}{\binom{3}{2}} + \binom{1}{1}$$

- $\binom{[m_{k-1}]}{k-1} \cup \{m_k+1\}$

$\vdots$

- $\binom{[m_t]}{t} \cup \{m_k+1, m_{k-1}+1, \dots, m_{t+1}+1\}$

$m=8$
$k=3$

$\binom{[4]}{3}$

$\binom{[3]}{2}$

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

$\{5,\{2,1\}\}$
$\{5,\{3,1\}\}$
$\{5,\{3,2\}\}$
$\{5,4,1\}$
$\{5,4,2\}$
$\{5,4,3\}$

Construct $\Delta R(m,k)$:

- $\binom{[m_k]}{k-1}$ ← removing an element.
- $\binom{[m_{k-1}]}{k-2} \cup \{m_k+1\}$

$\vdots$

- $\binom{[m_t]}{t-1} \cup \{m_k+1, \dots, m_{t+1}+1\}$

$$\Rightarrow |\Delta R(m,k)| = \sum_{l=t}^{k} \binom{[m_l]}{l-1}$$

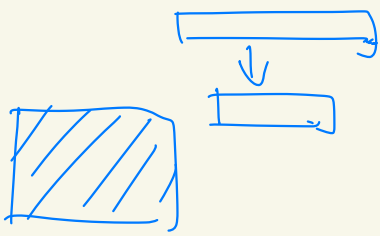
$\square$

Kruskal-Katona Theorem. (Fundamental Theorem of Extremal Set Theory)

Thm $\mathcal{F} \subseteq \binom{[n]}{k}$ with $|\mathcal{F}| = m = \sum_{l=t}^{k} \binom{m_l}{l}$

$$\Rightarrow |\Delta \mathcal{F}| \geq \binom{m_k}{k-1} + \binom{m_{k-1}}{k-2} + \cdots + \binom{m_t}{t-1}$$

* Δ $\mathcal{F} \subseteq 2^{[n]}$: vector $f = (|\mathcal{F}_0|, |\mathcal{F}_1|, \cdots, |\mathcal{F}_n|)$



$\mathcal{F}_k = \mathcal{F} \cap \binom{[n]}{k}$

$\Delta \mathcal{F}_{k+1}$ + additional k-sets

$\rightarrow$ given $|\mathcal{F}_{i-1}| \Rightarrow |\mathcal{F}_i|$ lower bound

def (Shift Operator) $\mathcal{F} \subseteq 2^{[n]}$ for $1 \leq i < j \leq n$,

$\forall T \in \mathcal{F}$: $\boxed{T_{ij} = (T \setminus \{j\}) \cup \{i\}}$ $\quad$ $\neq j$

(i,j)-shift : $\boxed{S_{ij}(\cdot)}$:

$\forall T \in \mathcal{F}$: $\boxed{S_{ij}(T)} = \begin{cases} T_{ij} & \text{if } j \in T,\ i \notin T,\ T_{ij} \notin \mathcal{F}. \\ T. & \text{otherwise} \end{cases}$

$S_{ij}(\mathcal{F}) = \{ S_{ij}(T) \mid T \in \mathcal{F} \}$

(i,j)-shift:

① $|S_{ij}(T)| = |T|$

$\Downarrow$

$|S_{ij}(\mathcal{F})| = |\mathcal{F}|$

② Property-maintaineane.

Proposition: $\Delta(S_{ij}(\mathcal{F})) \sqsubseteq S_{ij}(\Delta\mathcal{F})$

Cor $|\Delta S_{ij}(\mathcal{F})| \leq |S_{ij}(\Delta\mathcal{F})| = |\Delta\mathcal{F}|$

→ shifting does NOT increase the size of shadow!

$|\Delta S_{ij}(\mathcal{F})| \leq |\Delta\mathcal{F}|$ for $\forall\,(i,j)$-shift S_{ij}

Now we assume that $\mathcal{F}$ is shifted.

Proof: (induction on m) → Suppose $\underset{|\mathcal{F}'|}{\underset{\parallel}{m'}} < m$. $|\Delta\mathcal{F}'| \geq \sum_{l=t}^{k} \binom{m'_l - 1}{l-1}$

$\mathcal{F}_0 = \{ A \in \mathcal{F} \mid 1 \notin A \} \in \binom{[n]}{k}$

$\mathcal{F}_1 = \{ A \in \mathcal{F} \mid 1 \in A \} \in \binom{[n]}{k}$

↳ $\mathcal{F}_1' = \{ A \setminus \{1\} \mid A \in \mathcal{F}_1 \} \in \binom{[n]}{k-1}$

$\Rightarrow \boxed{|\mathcal{F}| = |\mathcal{F}_0| + |\mathcal{F}_1| = |\mathcal{F}_0| + |\mathcal{F}_1'|}$

Lemma 1. $|\Delta \mathcal{F}| \geq |\Delta \mathcal{F}_1| + |\mathcal{F}_1'|$

Proof: $\boxed{\Delta \mathcal{F}_1} = \left\{ A \in \binom{[n]}{k-1} \;\middle|\; \exists B \in \mathcal{F}_1.\ A \subseteq B \right\}$

$$= \left\{ A \in \binom{[n]}{k-1} \;\middle|\; 1 \in A.\ \exists B \in \mathcal{F}_1.\ A \subseteq B \right\}$$

$$\cup \underbrace{\left\{ A \in \binom{[n]}{k-1} \;\middle|\; 1 \notin A.\ \exists B \in \mathcal{F}_1.\ A \subseteq B \right\}}_{\mathcal{F}_1'}$$

$$= \left\{ S \cup \{1\} \;\middle|\; S \in \Delta \mathcal{F}_1' \right\} \cup \mathcal{F}_1'$$

$$\Rightarrow |\Delta \mathcal{F}_1| = |\Delta \mathcal{F}_1'| + |\mathcal{F}_1'|$$

$$\Rightarrow |\Delta \mathcal{F}| \geq |\Delta \mathcal{F}_1| \geq |\Delta \mathcal{F}_1'| + |\mathcal{F}_1'|. \qquad \square$$

Lemma 2 for shifted $\mathcal{F}$: $\Delta \mathcal{F}_0 \subseteq \mathcal{F}_1'$.

Proof: $A \in \Delta \mathcal{F}_0 \subseteq \binom{[n]}{k-1}$, $A \cup \{j\} \in \mathcal{F}_0 \subseteq \mathcal{F}$. $\exists j > 1$

$\rightarrow$ $(1, j)$-shift: $\underbrace{A \cup \{1\}}_{\mathcal{F}_1} \in \mathcal{F}. \Rightarrow A \in \mathcal{F}_1'$ $\square$

Lemma 3. If $\mathcal{F}$ is shifted,

$$|\mathcal{F}_1| \geq \binom{m_k-1}{k-1} + \binom{m_{k-1}-1}{k-2} + \cdots + \binom{m_t-1}{t-1}$$

Proof: $\boxed{|\mathcal{F}_1| < \sum_{\ell=t}^{k} \binom{m_\ell-1}{\ell-1}}$ $\qquad |\mathcal{F}| = |\mathcal{F}_0| + |\mathcal{F}_1|$

$$\Rightarrow |\mathcal{F}_0| = m - |\mathcal{F}_1| \qquad \sum_{\ell=t}^{k} \binom{m_\ell}{\ell}$$

$$\geq \sum_{\ell=t}^{k} \left(\binom{m_\ell}{\ell} - \binom{m_\ell-1}{\ell-1} \right) = \sum_{\ell=t}^{k} \binom{m_\ell-1}{\ell}$$

$$\Rightarrow |\Delta \mathcal{F}_0| \geq \binom{m_k-1}{k-1} + \cdots + \binom{m_t-1}{t-1} > |\mathcal{F}_1|$$

By Lemma 2. $|\mathcal{F}_1| \geq |\Delta \mathcal{F}_0|$. $\Rightarrow$ Contradiction! $\square$

$$|\mathcal{F}| = |\mathcal{F}_0| + |\mathcal{F}_1|$$

$$\begin{cases} ① \ |\Delta \mathcal{F}| \geq |\Delta \mathcal{F}_1| + |\mathcal{F}_1| \\ ② \ |\Delta \mathcal{F}_0| \leq |\mathcal{F}_1| \\ ③ \ |\mathcal{F}_1| \geq \sum_{\ell=t}^{k} \binom{m_\ell-1}{\ell-1} \end{cases}$$

$$|\Delta \mathcal{F}| \geq |\Delta \mathcal{F}_1| + |\mathcal{F}_1|$$

$$\geq \binom{m_k-1}{k-2} + \binom{m_{k-1}-1}{k-3} + \cdots + \binom{m_t-1}{t-2} \qquad \text{I.H.}$$

$$+ \binom{m_k-1}{k-1} + \binom{m_{k-1}-1}{k-2} + \cdots + \binom{m_t-1}{t-1} \qquad \text{Lemma 3}$$

$$= \binom{m_k}{k-1} + \binom{m_{k-1}}{k-2} + \cdots + \binom{m_t}{t-1} \qquad \square$$

---

def (r-shadow)

$$\Delta_r \mathcal{F} = \{ S \mid |S| = r, \ \exists T \in \mathcal{F} \text{ s.t. } S \subseteq T \}$$

$$\Delta_r \mathcal{F} = (\underbrace{\Delta \cdot \Delta \cdots \Delta}_{k-r}) \mathcal{F}.$$

$$|\Delta \mathcal{F}| \geq \binom{m_k}{k-1} + \binom{m_{k-1}}{k-2} + \cdots + \binom{m_t}{t-1}$$

$$\Rightarrow |\Delta_r \mathcal{F}| \geq \boxed{\binom{m_k}{r}} + \binom{m_{k-1}}{r-1} + \cdots + \binom{m_t}{t-k+r}$$

Application: (Erdős-Ko-Rado)

$\mathcal{F} \subseteq \binom{[n]}{k}$ $\quad S \cap T \neq \emptyset \ \forall S, T \in \mathcal{F}$. $\quad |\mathcal{F}| \leq \binom{n-1}{k-1}$

non-intersecting

$S \not\subseteq T$

Proof: (Daykin 1974, Clements 1976)

Suppose $|\mathcal{F}| > \binom{n-1}{k-1}$

dual system $\mathcal{G} = \{\bar{S} \mid S \in \mathcal{F}\}$. $\bar{S} = [n] \setminus S$

$\Rightarrow$ $\mathcal{F}$ and $\Delta_k \mathcal{G}$ disjoint.

$\Rightarrow |\mathcal{F}| + |\Delta_k \mathcal{G}| \leq \binom{n}{k}$ $\quad \subseteq \binom{[n]}{n-k}$

$|\mathcal{G}| = |\mathcal{F}| > \binom{n-1}{k-1} = \binom{n-1}{n-k}$

By Kruskal-Katona, $|\Delta_k \mathcal{G}| \geq \binom{n-1}{k}$

$\Rightarrow |\mathcal{F}| + |\Delta_k \mathcal{G}| > \binom{n-1}{k-1} + \binom{n-1}{k} = \binom{n}{k}$ $\quad \square$

Contradiction!