Schedule of the talk

- Linear Algebra: Reminder
- Fisher's inequality
- Vapnik-Chervonenkis dimension
- Sets with few intersection sizes
- Constructive Ramsey graphs
- The flipping cards game

Linear Algebra: Reminder (1)

Let V be a vector space over a field $\mathbb{F}$, $v_1, \ldots, v_n \in V$.

The vectors $v_1, \ldots, v_n$ are **linearly independent** if there is no linear relation

$$\lambda_1 v_1 + \cdots + \lambda_n v_n = 0$$

with $\lambda_i \neq 0$ for at least one i.

The set

$$\text{span}\{v_1, \ldots, v_n\} = \{\lambda_1 v_1 + \cdots + \lambda_n v_n : \lambda_1, \ldots, \lambda_n \in \mathbb{F}\}$$

is called the **span** of $v_1, \ldots, v_n$ and a subspace of V.

A **basis** of V is a set of linearly independent vectors whose span is V. The cardinality of each basis equals the **dimension** of V.

Proposition 1 (linear algebra bound) *If $v_1, \ldots, v_k$ are linearly independent vectors in V, then $k \leq \dim V$.*

Examples:

$\mathbb{F}^n$ is a vector space over $\mathbb{F}$ of dimension n with basis $e_1, \ldots, e_n$, where $e_i = (0, \ldots, 0, \underset{\substack{\uparrow \\ i}}{1}, 0, \ldots, 0)$.

The subspace $\text{span}\{v_1, \ldots, v_n\} \subseteq V$ has dimension at most n and its dimension is exactly n iff $v_1, \ldots, v_n$ are linearly independent.

Linear Algebra: Reminder (2)

In the vector space $V = \mathbb{R}^n$ we use the Euclidean standard **scalar product** defined by

$$\langle u, v\rangle = u_1v_1 + \cdots + u_nv_n = \sum_{i=1}^{n} u_iv_i$$

for two vectors $u = (u_1, \ldots, u_n)$ and $v = (v_1, \ldots, v_n)$.

Properties:

- $\langle u, v\rangle = \langle v, u\rangle$
- $\langle \lambda u + \mu v, w\rangle = \lambda\langle u, w\rangle + \mu\langle v, w\rangle$
- $\langle v, v\rangle \geq 0$ and $\langle v, v\rangle = 0 \Leftrightarrow v = 0$.

for all $u, v, w \in \mathbb{R}^n$ and $\lambda, \mu \in \mathbb{R}$

Example:

$A_1, A_2 \subseteq \{1, \ldots, n\} \rightsquigarrow v_1, v_2 \in \{0,1\}^n \subseteq \mathbb{R}^n$ (**incidence vectors**):

$$v_i = (v_{i1}, \ldots, v_{in}) \text{ where } v_{ij} = \begin{cases} 1 \ , & j \in A_i \\ 0 \ , & j \notin A_i \end{cases} \ , \quad i = 1, 2$$

$$\begin{aligned} \langle v_1, v_2\rangle &= |A_1 \cap A_2| \\ \langle v_i, v_i\rangle &= |A_i|, \ i = 1, 2 \end{aligned}$$

Fisher's inequality

Theorem 1 *(Majumdar 1953) Let $A_1, \ldots, A_m$ be distinct subsets of $\{1, \ldots, n\}$ such that $|A_i \cap A_j| = k$ for some fixed $1 \leq k \leq n$ for every $i \neq j$. Then $m \leq n$.*

Proof. (Babai and Frankl 1992)

$v_i \in \mathbb{R}^n,\ 1 \leq i \leq m$: incidence vectors of A_i

Goal: $v_1, \ldots, v_m$ are linearly independent in $\mathbb{R}^n$.

Assume $\sum_{i=1}^m \lambda_i v_i = 0$ for some $\lambda_1, \ldots, \lambda_m \in \mathbb{R}$.

Using

$$\langle v_i, v_j \rangle = |A_i \cap A_j| = \begin{cases} |A_i| & , \quad i = j \\ k & , \quad i \neq j \end{cases}$$

we conclude

$$\begin{aligned} 0 &= \left\langle \sum_i \lambda_i v_i, \sum_j \lambda_j v_j \right\rangle = \sum_i \lambda_i^2 |A_i| + \sum_{i \neq j} \lambda_i \lambda_j k \\ &= \sum_i \lambda_i^2 (|A_i| - k) + k \left(\sum_i \lambda_i^2 + \sum_{i \neq j} \lambda_i \lambda_j \right) \\ &= \sum_i \lambda_i^2 (|A_i| - k) + \left(\sum_i \lambda_i \right)^2 . \end{aligned}$$

Since $|A_i| \geq k$ for all i we must have

$$\begin{aligned} \lambda_1 + \cdots + \lambda_m &= 0 && (1) \\ \lambda_i^2(|A_i| - k) &= 0 \text{ for } i = 1, \ldots, m && (2) \end{aligned}$$

Assume there is an i with $\lambda_i \neq 0$.

$\stackrel{(2)}{\Rightarrow} |A_i| = k$

$\Rightarrow |A_j| > k$ for all $j \neq i$

$\stackrel{(2)}{\Rightarrow} \lambda_j = 0$ for all $j \neq i$

We end up with $\lambda_1 + \cdots + \lambda_m = \lambda_i \neq 0$, a contradiction to (1). Consequently,

$$\lambda_1 = \cdots = \lambda_m = 0.$$

□

Vapnik-Chervonenkis dimension

$\mathcal{F}$: family of subsets of an n-element set X

$Y \subseteq X$ is **shattered** by $\mathcal{F}$ if $\{E \cap Y : E \in \mathcal{F}\} = \mathcal{P}(Y)$

$\mathcal{F}$ is (n, k)-**dense** if there is a $Y \subseteq X$ with $|Y| = k$ such that Y is shattered by $\mathcal{F}$.

Remark: $\mathcal{F}$ (n, k)-dense $\Rightarrow \mathcal{F}$ (n, l)-dense for all $l < k$

The **Vapnik-Chervonenkis dimension** of $\mathcal{F}$ is the largest k for which $\mathcal{F}$ is (n, k)-dense.

Theorem 2 *If* $|\mathcal{F}| > \sum_{i=0}^{k-1} \binom{n}{i}$ *then the Vapnik-Chervonenkis dimension of* $\mathcal{F}$ *is at least k.*

Proof. (Frankl and Pach 1983)

$\mathcal{F} = \{E_1, \ldots, E_s\}$
$S_1, \ldots, S_r$: All subsets of X of size at most $k-1$

Define the $s \times r$ matrix $M = (m_{ij})$ by

$$m_{ij} = \begin{cases} 1 & , \quad E_i \supseteq S_j \\ 0 & , \quad \text{otherwise} \end{cases}$$

Since $s > r$, the rows $m_i = (m_{i1}, \dots, m_{ir})$, $1 \le i \le s$ cannot be linearly independent as elements of $\mathbb{R}^r$, i.e. there are $\lambda_1, \dots, \lambda_s$ not all zero such that

$$\sum_{i=1}^{s} \lambda_i m_i = 0. \tag{3}$$

Set for $T \subseteq X$

$$g(T) := \sum_{i\colon E_i \supseteq T} \lambda_i.$$

For $j = 1, \dots, r$ we have

$$g(S_j) = \sum_{i\colon E_i \supseteq S_j} \lambda_i = \sum_{i=1}^{s} \lambda_i m_{ij} \overset{(3)}{=} 0 \tag{4}$$

There is a $T \subseteq X$ with $g(T) \neq 0$.

Choose a subset $Y \subseteq X$ of minimum cardinality such that $g(Y) \neq 0$. By (4), $|Y| \geq k$.

Goal: Y is shattered by $\mathcal{F}$.

Let $Z \subseteq Y$. Because of the minimality of Y we get

$$\begin{aligned} 0 &\neq (-1)^{|Y \setminus Z|} g(Y) \\ &= \sum_{T\colon Z \subseteq T \subseteq Y} (-1)^{|T \setminus Z|} g(T) \\ &= \sum_{T\colon Z \subseteq T \subseteq Y} (-1)^{|T \setminus Z|} \sum_{i\colon E_i \supseteq T} \lambda_i \\ &= \sum_{i\colon E_i \supseteq Z} \lambda_i \sum_{T\colon Z \subseteq T \subseteq Y \cap E_i} (-1)^{|T \setminus Z|} \\ &= \sum_{i\colon E_i \cap Y = Z} \lambda_i \end{aligned}$$

since for $A \subseteq B$ with $n = |B \setminus A|$

$$\sum_{T\colon A \subseteq T \subseteq B} (-1)^{|T \setminus A|} = \sum_{k=0}^{n} \binom{n}{k} (-1)^k = \begin{cases} 1 & , \quad n = 0 \\ 0 & , \quad n \geq 1 \end{cases} .$$

Therefore there must be a member E_i of $\mathcal{F}$ such that $E_i \cap Y = Z$. □

Function spaces

$\mathbb{F}$ arbitrary field, $\Omega \subseteq \mathbb{F}^n$

$\mathbb{F}^\Omega := \{f \mid f : \Omega \to \mathbb{F}\}$: vector space over $\mathbb{F}$

Lemma 1 *For $i = 1, \ldots, m$ let $f_i \in \mathbb{F}^\Omega$ and $v_i \in \Omega$ such that*

- $f_i(v_i) \neq 0$ *for all* i
- $f_i(v_j) = 0$ *for all* $j < i$.

Then $f_1, \ldots, f_m$ are linearly independent in $\mathbb{F}^\Omega$.

Proof. Assume there is a linear relation

$$\lambda_1 f_1 + \cdots + \lambda_m f_m = 0 \tag{5}$$

with not all $\lambda_i = 0$. Take the smallest j for which $\lambda_j \neq 0$. Then (5) evaluated at v_j yields

$$0 = \lambda_1 f_1(v_j) + \cdots + \lambda_j f_j(v_j) + \cdots + \lambda_n f_n(v_j) = \lambda_j f_j(v_j)$$

and hence $\lambda_j = 0$ because $f_j(v_j) \neq 0$, a contradiction.

□

Sets with few intersection sizes (1)

$\mathcal{F}$: family of subsets of $\{1, \ldots, n\}$, $L \subseteq \{0, \ldots, n\}$

$\mathcal{F}$ is L-**intersecting** if $|A \cap B| \in L$ for all distinct members A, B of $\mathcal{F}$.

Theorem 3 *(Frankl and Wilson 1981) If $\mathcal{F}$ is L-intersecting, then $|\mathcal{F}| \leq \sum_{i=0}^{|L|} \binom{n}{i}$.*

Proof.

$\mathcal{F} = \{A_1, \ldots, A_m\}$, $|A_1| \leq |A_2| \leq \cdots \leq |A_m|$
$v_1, \ldots, v_m \in \{0, 1\}^n$ incidence vectors of $A_1, \ldots, A_m$

Define for $\Omega = \{0, 1\}^n$ in $\mathbb{R}^\Omega$ for $i = 1, \ldots, m$ the polynomial functions

$$f_i(x) = \prod_{l \in L\colon l < |A_i|} \left(\langle v_i, x \rangle - l\right), x \in \Omega.$$

Note that

- $f_i(v_i) \neq 0$ for all i (since $\langle v_i, v_i \rangle = |A_i|$)
- $f_i(v_j) = 0$ for all $j < i$ ($\langle v_i, v_j \rangle = \underbrace{|A_i \cap A_j|}_{\in L} < |A_i|$)

Hence, by lemma 1, $f_1, \ldots, f_m$ are linearly independent in $\mathbb{R}^{\Omega}$.

Each f_i is in the span of pure monomials $x_{i_1}x_{i_2}\ldots x_{i_s}$ with $i_1 < i_2 < \cdots < i_s$ and degree $s \leq |L|$ because $y^2 = y$ for $y \in \{0, 1\}$.

Since the dimension of this span is at most $\sum_{s=0}^{|L|} \binom{n}{s}$, the theorem follows. $\square$

Theorem 4 *(Deza, Frankl and Singhi 1983) Let p be a prime number and L and $\mathcal{F}$ as above. If*

- $|A_i| \notin L \pmod{p}$ *for all* i
- $|A_i \cap A_j| \in L \pmod{p}$ *for all* $i \neq j$

then $|\mathcal{F}| \leq \sum_{i=0}^{|L|} \binom{n}{i}$.

Proof.

$v_1, \ldots, v_m \in \{0, 1\}^n$ incidence vectors of $A_1, \ldots, A_m$

Define for $\Omega = \{0, 1\}^n$ in $(\mathbb{F}_p)^\Omega$ for $i = 1, \ldots, m$ the polynomial functions

$$f_i(x) = \prod_{l \in L} (\langle v_i, x \rangle - l) \ , x \in \Omega.$$

Note that

- $f_i(v_i) \neq 0$ for all i (since $\langle v_i, v_i \rangle = |A_i|$)
- $f_i(v_j) = 0$ for all $j \neq i$ (since $\langle v_i, v_j \rangle = |A_i \cap A_j|$)

Hence, by lemma 1, $f_1, \ldots, f_m$ are linearly independent in $(\mathbb{F}_p)^\Omega$.

The theorem follows as in the previous proof. □

Constructive Ramsey graphs

A **clique** is a set of pairwise adjacent vertices in a graph.

An **independent set** is a set of pairwise non-adjacent vertices in a graph.

A graph is a **Ramsey graph** with respect to t if it has no clique and no independent set of size t.

Erdős (1947): Proved the existence of Ramsey graphs of order $n = \lfloor 2^{t/2} \rfloor$ using the probabilistic method.

Aim: **explicitely** construct Ramsey graphs with respect to a fixed t of large order

Order $n = (t-1)^2$: disjoint union of $t-1$ cliques of size $t-1$ each (Turán)

Ramsey graph of order $n = \Omega(t^3)$

Construction by Zsigmond Nagy (1972)

vertex set: all subsets of $\{1, \ldots, t-1\}$ of size 3
edge set E: $\{A, B\} \in E \Leftrightarrow |A \cap B| = 1$

Verification. Let $A_1, \ldots, A_m$ be a clique. We have $|A_i \cap A_j| = 1$ for every $i \neq j$. Hence, $m \leq t-1$ by Fisher's inequality.

Let $A_1, \ldots, A_m$ be an independent set with incidence vectors $v_1, \ldots, v_m \in (\mathbb{F}_2)^{t-1}$. We have $|A_i \cap A_j| \in \{0, 2\}$ for all $i \neq j$. Therefore we get in $\mathbb{F}_2$

$$\langle v_i, v_j \rangle = |A_i \cap A_j| = \begin{cases} 0 & , \quad i \neq j \\ 1 & , \quad i = j \end{cases}$$

Assume there is a linear relation

$$\lambda_1 v_1 + \cdots + \lambda_m v_m = 0$$

over $\mathbb{F}_2$. Then we get for every i

$$\lambda_i = \langle \lambda_1 v_1 + \cdots + \lambda_m v_m, v_i \rangle = \langle 0, v_i \rangle = 0.$$

Therefore $v_1, \ldots, v_m$ are linearly independent in $(\mathbb{F}_2)^{t-1}$ and hence $m \leq t-1$.

Ramsey graph of order $t^{\Omega(\ln t/ \ln\ln t)}$

Construction by Frankl (1977)

Define for a prime number p the graph G_p by

Vertex set: all subsets of $\{1, \ldots, p^3\}$ of size $p^2 - 1$
Edge set E: $\{A, B\} \in E \Leftrightarrow |A \cap B| \not\equiv -1 \pmod{p}$

Theorem 5 *The graph G_p is a Ramsey graph with respect to $\sum_{i=0}^{p-1} \binom{p^3}{i} + 1$.*

Remark. The theorem yields for a fixed t a Ramsey graph of order $t^{\Omega(\ln t/ \ln\ln t)}$.

Proof. Let $A_1, \ldots, A_m$ be an independent set. We have

$$|A_i \cap A_j| \in \{p-1, 2p-1, \ldots, p^2-p-1\}$$

for every $i \neq j$. By Theorem 3,

$$m \leq \sum_{i=0}^{p-1} \binom{p^3}{i}.$$

A clique of size m consists of sets $A_1, \ldots, A_m$ with

- $|A_i \cap A_j| \not\equiv -1 \pmod p$ for every $i \neq j$
- $|A_i| \equiv -1 \pmod p$ for all i.

Applying Theorem 4 with $L = \{0, \ldots, p-2\}$, we conclude

$$m \leq \sum_{i=0}^{p-1} \binom{p^3}{i}.$$

□

Verification of the remark. By the theorem, G_p for $p = \max\{q \text{ prime: } \sum_{i=0}^{q-1} \binom{q^3}{i} < t\}$ is a Ramsey graph with respect to t for any t.

It is of order

$$n = \binom{p^3}{p^2-1} \geq \left(\frac{p^3}{p^2-1}\right)^{p^2-1} = p^{\Omega(p^2)}.$$

Furthermore, we have, since there is a prime between N and $2N$ for any integer N

$$t \leq \sum_{i=0}^{2p-1} \binom{(2p)^3}{i} \leq 2p\binom{(2p)^3}{2p-1} \leq (2p)^{6p-2} = p^{O(p)}.$$

This yields

$$p = \Omega(\ln t/\ln\ln t)$$

since for sufficiently large t

$$\left(\frac{\ln t}{\ln\ln t}\right)^{\frac{\ln t}{\ln\ln t}} < t$$

and we get

$$n = p^{\Omega(p^2)} = t^{\Omega(p)} = t^{\Omega(\ln t/\ln\ln t)}.$$

The flipping cards game

$u = (u_1, \ldots, u_n)$, $v = (v_1, \ldots, v_n) \in \{0, 1\}^n$

Probe: 0-1 vector of length n containing exactly one bit of each pair u_i, v_i, e.g. $(v_1, u_2, u_3, \ldots, u_n)$

Goal: Decide whether $u = v$ with probes and using as little non-reusable memory as possible

Theorem 6 *(J. Edmonds, R. Impagliazzo) For $n = r^2$ it is possible to test the equality of u, v using only $r + 1$ probes and writing down only r bits in the memory.*

Proof.

$u = (u_1, \ldots, u_r)$, $u_i \in \{0, 1\}^r$
$v = (v_1, \ldots, v_r)$, $v_i \in \{0, 1\}^r$

Consider the following protocol:

Probe 0: $(u_1, \ldots, u_r)$
$\rightsquigarrow$ write down $w_0 := u_1 + \cdots + u_r \bmod 2$ in the memory

For $1 \leq i \leq r$:

Probe i: $(u_1, \dots, u_{i-1}, v_i, u_{i+1}, \dots, u_r)$
$\rightsquigarrow w_i := u_1 + \cdots + u_{i-1} + v_i + u_{i+1} + \cdots + u_r \bmod 2$
Stop and report $u \neq v$ if $w_0 \neq w_i$

Answer $u = v$ at the end if not $u \neq v$ reported

This protocol reports $u \neq v \Leftrightarrow u_i \neq v_i$ for some $1 \leq i \leq r \Leftrightarrow u \neq v$ $\square$

Theorem 7 *(Pudlák, Sgall 1997) It is possible to test the equality of u, v using only $O(\log n)$ probes and writing down only $O((\log n)^2)$ bits in the memory.*

Proof. Note that

$$u = v \Leftrightarrow 0 = \langle u - v,\, u - v\rangle = \langle u,\, u\rangle + \langle v,\, v\rangle - 2\langle u,\, v\rangle$$
$$= \sum_{i=1}^{n} u_i^2 + \sum_{i=1}^{n} v_i^2 - 2\left(\sum_{i=1}^{n} u_i \sum_{i=1}^{n} v_i - \sum_{i\neq j} u_i v_j\right)$$

Probe 1: $(u_1, \ldots, u_n)$
$\rightsquigarrow \sum_{i=1}^{n} u_i^2,\ \sum_{i=1}^{n} u_i$

Probe 2: $(v_1, \ldots, v_n)$
$\rightsquigarrow \sum_{i=1}^{n} v_i^2,\ \sum_{i=1}^{n} v_i$

Probe 3: $(u_1, \ldots, u_{\lfloor \frac{n}{2} \rfloor}, v_{\lfloor \frac{n}{2} \rfloor + 1}, \ldots, v_n)$

$\rightsquigarrow \sum_{i=1}^{\lfloor \frac{n}{2} \rfloor} \sum_{j=\lfloor \frac{n}{2} \rfloor + 1}^{n} u_i v_j$

Probe 4: $(v_1, \ldots, v_{\lfloor \frac{n}{2} \rfloor}, u_{\lfloor \frac{n}{2} \rfloor + 1}, \ldots, u_n)$

$\rightsquigarrow \sum_{i=\lfloor \frac{n}{2} \rfloor + 1}^{n} \sum_{j=1}^{\lfloor \frac{n}{2} \rfloor} u_i v_j$

Probe 5: $(u_1, \ldots, u_{\lfloor \frac{n}{4} \rfloor}, v_{\lfloor \frac{n}{4} \rfloor + 1}, \ldots, v_{\lfloor \frac{n}{2} \rfloor}, u_{\lfloor \frac{n}{2} \rfloor + 1}, \ldots, u_{\lfloor \frac{3n}{4} \rfloor},$
$v_{\lfloor \frac{3n}{4} \rfloor + 1}, \ldots, v_n)$
$\rightsquigarrow \sum_{i=1}^{\lfloor \frac{n}{4} \rfloor} \sum_{j=\lfloor \frac{n}{4} \rfloor + 1}^{\lfloor \frac{n}{2} \rfloor} u_i v_j + \sum_{i=\lfloor \frac{n}{2} \rfloor + 1}^{\lfloor \frac{3n}{4} \rfloor} \sum_{j=\lfloor \frac{3n}{4} \rfloor + 1}^{n} u_i v_j$

Continue like that until you have considered all products $u_i v_j$ for $i \neq j$ and finally sum all values stored in the memory up to get $\langle u - v,\, u - v \rangle$.

This protocol needs $2\lceil \log n \rceil + 2$ probes and for each memorized number $2\lceil \log(n+1) \rceil$ bits of memory (since all these numbers lie between 0 and n^2). □