Counting Matroids

Rudi Pendavingh

March 27, 2015

1. Introduction to counting matroids
   - A question
   - Some early answers
   - Conjectures
2. New upper bounds
   - Better matroid compression
   - An entropy method
3. Hard work
   - Counting stable sets
   - Matroid vs stable sets

How many matroids on a fixed ground set E exist?

Some names

$$[n] := \{1, \ldots, n\}$$

$$m_n := \#\{M \text{ matroid } : E(M) = [n]\}$$

$$m_{n,r} := \#\{M \text{ matroid } : E(M) = [n], r(M) = r\}$$

A naive upper bound

A matroid on E is determined by the set of its independent sets $\mathcal{I} \subseteq 2^E$. Hence

$$m_n \leq 2^{2^n}$$

Therefore,

$$\log m_n \leq 2^n$$

Logarithms are to the base 2 today.

A slightly less naive upper bound

A matroid on E of rank r is determined by bases $\mathcal{B} \subseteq \binom{E}{r} := \{X \subseteq E : |X| = r\}$. Hence

$$m_{n,r} \leq 2^{\binom{n}{r}} \leq 2^{\binom{n}{\lfloor n/2 \rfloor}}$$

Therefore $m_n = \sum_r m_{n,r} \leq (n+1)2^{\binom{n}{\lfloor n/2 \rfloor}}$, hence

$$\log m_n \leq \log(n+1) + \binom{n}{\lfloor n/2 \rfloor} \leq O(2^n/\sqrt{n}) \text{ as } n \to \infty$$

Here we used that

$$\frac{2^n}{\sqrt{n}}\sqrt{\frac{2}{\pi}}(1 - o(1)) \ \leq \ \binom{n}{\lfloor n/2 \rfloor} \ \leq \ \frac{2^n}{\sqrt{n}}\sqrt{\frac{2}{\pi}} \text{ as } n \to \infty$$

A faithful description of matroids..

Lemma

The following are equivalent for a matroid $M = (E, \mathcal{B})$ and a set $X \subseteq E$

- *X is a dependent set of M*
- *there is a circuit C of M so that $C \subseteq X$*
- *there is a circuit C of M so that $|X \cap cl_M(C)| > r_M(C)$*

Each matroid M on a fixed ground set E is determined by the set

$$\mathcal{K}(M) := \{(\mathrm{cl}_M(C), r_M(C)) : C \text{ a circuit of } M\}$$

..which is more compressed

Each matroid M on a fixed ground set E is determined by the set

$$\mathcal{K}(M) := \{(\mathrm{cl}_M(C), r_M(C)) : C \text{ a circuit of } M\}$$

Lemma

If M is a matroid on n elements, then $|\mathcal{K}(M)| \leq 2^{n+1}/(n+1)$.

Proof.

For each circuit C, there are $|C|$ sets $Y \subseteq C$ of size $r_M(C)$ so that $\mathrm{cl}_M(Y) = \mathrm{cl}_M(C)$. Hence

$$\#\{(\mathrm{cl}_M(C), r_M(C)) : C \text{ a circuit of } M, r_M(C) = i\} \leq \binom{n}{i}/(i+1).$$

Hence $|\mathcal{K}(M)| = \sum_{i<n} \binom{n}{i}/(i+1) = \sum_{i<n} \binom{n+1}{i+1}/(n+1) = 2^{n+1}/(n+1)$. □

Piff's upper bound on the number of matroids

Theorem (Piff, 1973)

$$\log m_n \leq O(2^n \log(n)/n) \text{ as } n \to \infty$$

Proof.

Each matroid on n elements is determined by $\mathcal{K} \subseteq 2^E \times [n+1]$ of size $|\mathcal{K}| \leq 2^{n+1}/(n+1)$, so

$$m_n \leq \sum_{i \leq 2^{n+1}/(n+1)} \binom{2^n(n+1)}{i} \leq \left(\frac{e2^n(n+1)}{2^{n+1}/(n+1)}\right)^{2^{n+1}/(n+1)}$$

and hence $\log m_n \leq (2^{n+1}/(n+1)) \cdot \log(e(n+1)^2/2)$. □

Here we used

$$\sum_{i=0}^{k} \binom{n}{i} \leq \left(\frac{en}{k}\right)^k$$

Some bounds for well-known classes of matroids

What do we know about the number of matroids on n elements in various classes?

class	upper bound	log u.b.
graphic matroids	$\binom{n+1}{2}^n$	$O(n \log n)$
binary matroids	$(2^n)^n$	n^2
$\mathrm{GF(q)}$-representable matroids	$(q^n)^n$	$n^2 \log q$
transversal matroids	$(2^n)^n$	n^2
real-representable matroids	2^{n^3}	n^3

Theorem (Alon, 1986)

The number of real-representable matroids of rank r on n elements is between

$$n^{(r-1)^2 n - O(r^2 n(\log r + \log\log n)/(\log n))} \text{ and } n^{r(r-1)n + O(nr \log\log n/\log n)}.$$

No construction within these classes will yield a lower bound near Piff's upper bound.

Paving matroids

Definition

A matroid M is *paving* if $|C| \geq r(M)$ for each circuit C of M.

Crapo and Rota (1970) consider it likely that paving matroids

> " ... *would actually predominate in any asymptotic enumeration of geometries*"

based on the enumeration of matroids up to 8 elements (Blackburn, Crapo, and Higgs).

Definition

A matroid M is *sparse paving* if both M and M^* are paving.

If almost all matroids are paving, then almost all matroids are sparse paving.

Sparse paving matroids

Lemma

Let $0 < r < |E|$ and let $\mathcal{B} \subseteq \binom{E}{r} := \{X \subseteq E : |X| = r\}$. The following are equivalent:

1. $M = (E, \mathcal{B})$ *is a sparse paving matroid*
2. $|X \triangle Y| > 2$ *for all distinct* $X, Y \in \binom{E}{r} \setminus \mathcal{B}$

Proof.

(1) $\Rightarrow$ (2): if $X, Y \in \binom{E}{r} \setminus \mathcal{B}$ and $|X \triangle Y| = 2$, then

$$2(r-1) \geq r(X) + r(Y) \geq r(X \cup Y) + r(X \cap Y)$$

so that either $r(X \cap Y) < r - 1$ or $r(X \cup Y) < r$.

In either case $M = (E, \mathcal{B})$ is not sparse paving. □

Sparse paving matroids

Lemma

Let $0 < r < |E|$ and let $\mathcal{B} \subseteq \binom{E}{r} := \{X \subseteq E : \ |X| = r\}$. The following are equivalent:

1. $M = (E, \mathcal{B})$ *is a sparse paving matroid*
2. $|X \triangle Y| > 2$ *for all distinct* $X, Y \in \binom{E}{r} \setminus \mathcal{B}$

Proof.

(2) $\Rightarrow$ (1): Suppose (2) holds. If M were not a matroid, then

$$\exists B, B' \in \mathcal{B}, \ \exists e \in B \setminus B', \ \forall f \in B' \setminus B : \ B - e + f \notin \mathcal{B}$$

Pick $f, f' \in B' \setminus B$ distinct, then $X := B - e + f$, $Y := B - e + f'$ violate (2), contradiction. If D is a dependent set of size $|D| = r - 1$, then $X := D + e$, $Y := D + f$ violate (2), contradiction. So M is paving and the dual argument shows that M^* is also paving. □

The Johnson graph

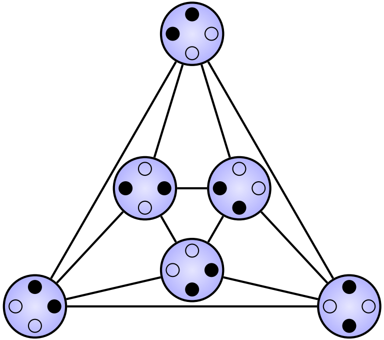
The *Johnson graph* $J(E,r)$ is the undirected graph with vertices

$$V(J(E,r)) = \binom{E}{r} := \{X \subseteq E : |X| = r\}$$

and edges

$$E(J(E,r)) = \{XY : \ |X \triangle Y| = 2\}$$

We put $J(n,r) := J([n],r)$.



$J(4,2)$

Lemma

$$S \text{ is a stable set of } J(E,r) \iff M = (E, \binom{E}{r} \setminus S) \text{ is a sparse paving matroid}$$

More names

$$s_n := \#\{M \text{ sparse paving matroid } : E(M) = [n]\}$$

$$s_{n,r} := \#\{M \text{ sparse paving matroid } : E(M) = [n], r(M) = r\}$$

$s_{n,r}$ equals the number of stable sets of $J(n, r)$

Knuth's lower bound on the number of matroids

Let $\alpha(G) := \max\{|S| : S \subseteq V(G) \text{ a stable set of } G\}$.

Theorem (Knuth, 1974)

$$\log s_{n,r} \geq \alpha(J(n,r))$$

Proof.

$J(n,r)$ has a stable set S_0 of size $\alpha(J(n,r))$. Each $S \subseteq S_0$ is a stable set of $J(n,r)$, so that

$$s_{n,r} \geq 2^{|S_0|} = 2^{\alpha(J(n,r))}$$

□

In 1974, this gave the lower bound

$$\log m_n \geq \log s_n \geq \log s_{n,n/2} \geq \alpha(J(n,n/2)) \geq \binom{n}{n/2}/(2n)$$

An improvement of Knuth's bound

Theorem (Graham & Sloane, 1980)

$J(n, r)$ has a stable set of size $\binom{n}{r}/n$.

Proof.

For each $k \in \mathbb{Z}$, the set

$$S(n, r, k) := \{X \in \binom{[n]}{r} : \sum_{x \in X} x = k \mod n\}$$

is stable in $J(n, r)$. Since $\bigcup_{k=1}^{n} S(n, r, k) = \binom{[n]}{r}$, we have $|S(n, r, k)| > \binom{n}{r}/n$ for some k. □

And so we obtain

$$\log m_n \geq \log s_{n,n/2} \geq \alpha(J(n, n/2)) \geq \binom{n}{n/2}/n.$$

Conjectures on matroid asymptotics

We say that 'asymptotically almost all matroids have property $\mathcal{P}$' if

$$\lim_{n\to\infty} \frac{\#\{M \in \mathbb{M}_n : M \text{ has property } \mathcal{P}\}}{\#\mathbb{M}_n} = 1$$

Here $\mathbb{M}_n$ denotes the set of matroids with ground set $\{1, \ldots, n\}$.

Conjecture (Mayhew, Newman, Welsh and Whittle, 2011)

- *Asymptotically almost all matroids are sparse paving.*
- *If N is a fixed sparse paving matroid, then a.a.a. matroids have N as a minor.*
- *Asymptotically almost all matroids M on n elements have $\lfloor \frac{n}{2} \rfloor \leq r(M) \leq \lceil \frac{n}{2} \rceil$.*
- *Let $k \in \mathbb{N}$. Asymptotically almost all matroids are k-connected.*

Theorem (Oxley, Semple, Warshauer, Welsh, 2011)

Asymptotically almost all matroids are 3-connected.

Inspiration from matroid computation

Since 2012, the computer algebra system Sage contains a matroid package (P. & van Zwam). The development of this package raised the following problem:

> *How to store a general matroid on n elements as concise as possible, but so that the rank oracle takes poly(n) time?*

We settled on the `BasisMatroid`:

- stores a matroid on E elements of rank r as a the indicator function of its bases $\mathcal{B} \subseteq \binom{E}{r}$; the length of this description is $\binom{n}{r}$ bits.
- the rank oracle takes poly(n) time

We also still use the more human-friendly `CircuitClosuresMatroid`:

- stores a matroid M using Piff's compact description $\mathcal{K}(M) := \{(\mathrm{cl}_M(C), r_M(C)) : C \text{ a circuit of } M\}$; the length of the data is $O(2^n)$
- the rank oracle does not run in poly(n) time from this data

Break

A new upper bound

Theorem (Piff, 1973)

$$\log m_n \leq O\left(\frac{\log(n)}{n} 2^n\right) \text{ as } n \to \infty$$

Theorem (Bansal, P. , van der Pol, 2012)

$$\log m_n \leq O\left(\frac{\log^2(n)}{n} \binom{n}{n/2}\right) \text{ as } n \to \infty$$

Proof outline:

- a *cover* of a matroid M gives a compressed description of M
- each n-element matroid of rank r has a cover of 'size' $\leq k_{n,r}$
- # of matroids on n elements of rank r is $\leq$ the # of covers of size $\leq k_{n,r}$

We aim for $k_{n,r} \approx \binom{n}{r}/n \leq O(2^n/n^{3/2})$.

Matroid covers

Let $M = (E, \mathcal{B})$ be a matroid of rank r.

Definition

Let $X \subseteq E$. A flat F of M *covers* X if $|F \cap X| > r_M(F)$.

A flat that covers X certifies that X is dependent.

Definition

A collection $\mathcal{Z} \subseteq \mathcal{F}(M)$ *covers* M if

$$\text{for all } X \in \binom{E}{r} \setminus \mathcal{B}, \text{ there exists } F \in \mathcal{Z} \text{ covering } X$$

If $\mathcal{Z}$ covers M, then $\{(F, r_M(F)) : F \in \mathcal{Z}\}$ determines M.

Matroids have small local covers

Let $M = (E, \mathcal{B})$ be a matroid of rank r. Let $X \in \binom{E}{r}$.

Lemma

There exists a $\mathcal{Z}_X \subseteq \mathcal{F}(M)$ such that

- $|\mathcal{Z}_X| \leq r$ *and*
- $\mathcal{Z}_X$ *covers each dependent* $Y \in \binom{E}{r} \setminus \mathcal{B}$ *such that* $|Y \triangle X| \leq 2$

Proof.

Take

$$\mathcal{Z}_X := \{\mathrm{cl}_M(X - x) : x \in X\}$$

If $Y = X - x + y$ and $\mathrm{cl}_M(X - x)$ does not cover Y, then

$$r - 1 = |X - x| \leq |\mathrm{cl}_M(X - x) \cap Y| \leq r_M(\mathrm{cl}_M(X - x)) \leq r - 1$$

So $r_M(X - x) = r - 1$ and $y \notin \mathrm{cl}_M(X - x)$, hence $r_M(X - x + y) = r$, i.e. Y independent. □

$J(n,r)$ has a small dominating set

Let $G=(V,E)$ be a graph. A set $D\subseteq V$ is *dominating* if

$$D\cup N(D)=V$$

where $N(D) := \{v\in V\setminus D : d\in D, dv\in E\}$

Theorem (Lovász, 1975)

If $G=(V,E)$ is d-regular, then G has a dominating set $D\subseteq V$ with

$$|D|\leq \frac{\ln(d+1)+1}{d+1}|V|$$

Corollary

The Johnson graph $J(n,r)$ has a dominating set D with

$$|D|\leq \frac{\ln(r(n-r)+1)+1}{r(n-r)+1}\binom{n}{r}$$

Matroids have small covers

Theorem

Let $M = (E, \mathcal{B})$ be a matroid of rank r, with $n = |E|$. Then M has a cover $\mathcal{Z}$ such that

$$|\mathcal{Z}| \leq \frac{\ln(r(n-r)+1)+1}{n-r}\binom{n}{r} =: k_{n,r}$$

Proof.

Take

$$\mathcal{Z} := \bigcup_{X \in D} \mathcal{Z}_X$$

where

- D is a dominating set of $J(E, r)$ with $|D| \leq \frac{\ln(r(n-r)+1)+1}{r(n-r)+1}\binom{n}{r}$
- each $\mathcal{Z}_X$ covers the non-bases in $N(\{X\}) \cup \{X\}$, with $|\mathcal{Z}_X| \leq r$

□

Finishing up

Theorem (Bansal, P. , van der Pol 2012)

$$\log m_n \leq O\left(\frac{\log^2(n)}{n}\binom{n}{n/2}\right) \textit{ as } n \to \infty$$

Proof.

Each matroid $M = (E, \mathcal{B})$ on n elements of rank r is determined by the set

$$\{(F, r_M(F)) : F \in \mathcal{Z}\} \subseteq 2^E \times \{0, \ldots, r-1\}$$

for some cover $\mathcal{Z}$ with $|\mathcal{Z}| \leq k_{n,r}$. Hence

$$m_{n,r} \leq \sum_{j \leq \min\{k_{n,r}, k_{n,n-r}\}} \binom{2^n n}{j} \leq \left(\frac{2^n n}{k_{n,n/2}}\right)^{k_{n,n/2}}$$

where $k_{n,n/2} \approx \frac{8\ln(n)}{n^2}\binom{n}{n/2}$. □

A newer upper bound

Theorem (Bansal, P. , van der Pol, 2012)

$$\log m_n \leq O\left(\frac{\log^2(n)}{n}\binom{n}{n/2}\right) \text{ as } n \to \infty$$

Theorem (Bansal, P. , van der Pol, 2013)

$$\log m_n \leq O\left(\frac{\log(n)}{n}\binom{n}{n/2}\right) \text{ as } n \to \infty$$

Proof outline:

- we use entropy to bound $m_{n,r}$ in terms of $m_{n-t,r-t}$
- we derive a bound on $m_{n,2}$
- putting together the two, we get a sufficient bound

Further applications: counting minor-closed classes and counting oriented matroids.

Entropy and counting

Let X be a random variable drawn from a finite set S with probability $p \in \mathbb{R}_+^S$

The *entropy* of X is

$$H(X) := \sum_{s \in S} p_s \log(1/p_s)$$

Lemma

$$\max\{\sum_{s \in S} p_s \log(1/p_s) : p \text{ a probability distribution on } S\} = \log |S|$$

The maximum is attained by the uniform distribution, i.e. $p_s = |S|^{-1}$ for all $s \in S$; then

$$H(X) = \sum_{s \in S} p_s \log(1/p_s) = \log |S|$$

To bound $|S|$ is to bound the entropy of the random variable X drawn uniformly from S

Shearer's Lemma

- If $S = S_1 \times \cdots \times S_m$, and X is drawn from S, then X is a vector with entries $X_i \in S_i$.
- For a set $A \subseteq [m]$, we denote the restriction of X to A by $X_A := (X_i)_{i \in A}$
- If X is drawn according to p, then X_A is drawn according to p_A where

$$p_A(Y) = \sum_{X_A = Y} p(X)$$

Definition

A collection of sets $\mathcal{A} \subseteq 2^{[m]}$ is a *k-cover* if each $i \in [m]$ is in $\geq k$ sets from $\mathcal{A}$.

Theorem (Shearer)

Suppose X is a random variable drawn from $S = S_1 \times \cdots \times S_m$. Let $\mathcal{A}$ be a k-cover of $[m]$. Then

$$kH(X) \leq \sum_{A \in \mathcal{A}} H(X_A)$$

An application

Theorem (Shearer)

Suppose X is a random variable drawn from $S = S_1 \times \cdots \times S_m$. Let $\mathcal{A}$ be a k-cover of $[m]$. Then

$$kH(X) \leq \sum_{A \in \mathcal{A}} H(X_A)$$

Theorem

Let $T \subseteq \mathbb{Z}^3$ be a finite set of points. Let $T_i := \{\pi_i(t) : t \in T\}$, where π_i is the orthogonal projection parallel to e_i. Then $|T|^2 \leq |T_1| \cdot |T_2| \cdot |T_3|$.

Proof.

Let X be the random variable drawn uniformly from T. Then by Shearer's Lemma

$$2 \log |T| = 2H(X) \leq H(X_{\{2,3\}}) + H(X_{\{1,3\}}) + H(X_{\{1,2\}}) \leq \log |T_1| + \log |T_2| + \log |T_3|.$$

□

Application to counting matroids

Lemma

$$\log(1 + m_{n,r}) \leq \frac{n}{n-r} \log(1 + m_{n-1,r})$$

Proof.

Let X be the random variable drawn uniformly from sets $\mathcal{B} \subseteq \binom{E}{r}$ satisfying base exchange. We identify X with its indicator vector in $\{0,1\}^{\binom{E}{r}}$.

- Let $A_e := \{Y \in \binom{E}{r} : e \notin Y\}$. Then $\mathcal{A} := \{A_e : e \in E\}$ is an $(n-r)$-cover of $\binom{E}{r}$
- X_{A_e} is a set of subsets from $\binom{E-e}{r}$ satisfying base exchange
- By Shearer's Lemma,

$$(n-r)\log(1 + m_{n,r}) = (n-r)H(X) \leq \sum_{e \in E} H(X_{A_e}) \leq n \log(1 + m_{n-1,r})$$

□

Finishing up

Theorem

$$\log(1+m_{n,r})/\binom{n}{r} \leq \log(1+m_{n-r+t,t})/\binom{n-r+t}{t}$$

Lemma

$$1+m_{n,2} \leq (n+1)^n$$

Theorem (Bansal, P. , van der Pol, 2013)

$$\log m_n \leq O\left(\frac{\log(n)}{n}\binom{n}{n/2}\right) \text{ as } n \to \infty$$

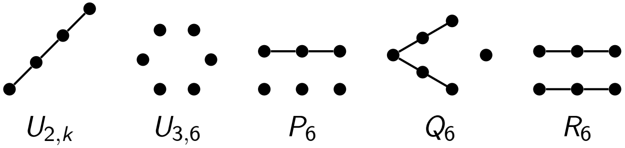
Proof.

$$\frac{\log(1+m_{n,r})}{\binom{n}{r}} \leq \frac{\log(1+m_{n-r+2,2})}{\binom{n-r+2}{2}} \leq \frac{(n-r+2)\log(n-r+3)}{(n-r+2)(n-r+1)/2} = \frac{2\log(n-r+3)}{n-r+1}$$

Matroids without a sparse paving minor

Conjecture (Mayhew, Newman, Welsh and Whittle, 2011)

If N is a fixed sparse paving matroid, then asymptotically almost all matroids have N as a minor.



Theorem (P., van der Pol, 2013)

If $N = U_{2,k}$ for some $k \geq 2$, or if N is one of $U_{3,6}$, P_6, Q_6 or R_6, then

$$\lim_{n\to\infty} \frac{\#\{M \in \mathbb{M}_n : M \text{ does not have } N \text{ as a minor}\}}{\#\mathbb{M}_n} = 0.$$

Counting matroids in minor-closed classes

Theorem (P., van der Pol, 2013)

If $N = U_{2,k}$ for some $k \geq 2$, or if N is one of $U_{3,6}$, P_6, Q_6 or R_6, then

$$\lim_{n\to\infty} \frac{\#\{M \in \mathbb{M}_n : M \text{ does not have } N \text{ as a minor}\}}{\#\mathbb{M}_n} = 0.$$

Lemma

If M is a simple matroid of rank 3 without $U_{2,k}$ minor, then $|E(M)| \leq k^2$.

Hence there are at most $O(k^{2n})$ matroids on n elements of rank 3 without $U_{2,k}$, hence

$$\frac{\log(m'_{n,r}+1)}{\binom{n}{r}} \leq \frac{\log(m'_{n-r+3,3}+1)}{\binom{n-r+3}{3}} \leq O(1/n^2)$$

which implies $\log m'_n \leq O(\binom{n}{n/2}/n^2)$ for the number of matroids without $U_{2,k}$.

Counting oriented matroids

Let $p_{n,r}$ denote the number of *oriented* matroids on $E = [n]$ of rank r. By entropy counting

$$\log(1 + p_{n,r})/\binom{n}{r} \leq \log(1 + p_{n-r+t,t})/\binom{n-r+t}{t}$$

Theorem (Felsner& Valtr; Bern, Eppstein, Plasman &Yao)

$$0.1887n^2 \leq \log p_{n,3} \leq 1.085n^2$$

So for each $t \geq 3$ there is a c_t such that $\log(1+p_{n,t})/\binom{n}{t} \leq c_t/n$, indeed $c_t \leq c_{t-1} \leq \cdots \leq c_3$.

Conjecture

Asymptotically almost all matroids are not orientable.

If we can show $c_t < \frac{1}{2}$ for some t, then the conjecture is proven. Perhaps $t = 4$ or $t = 5$?

Break

Where are we?

So far, we have

$$\frac{1}{n}\binom{n}{n/2} \leq \log s_n \leq \log m_n \leq O\left(\frac{\log n}{n}\binom{n}{n/2}\right) \text{ as } n \to \infty$$

Where is the gap?

In what follows, we show:

Theorem (Bansal, P., van der Pol, 2012)

$$\log m_n \leq \frac{2}{n}\binom{n}{n/2}(1+o(1)) \textit{ as } n \to \infty$$

Theorem (P., van der Pol, 2014)

$$\log m_n = (1+o(1)) \log s_n \textit{ as } n \to \infty$$

A first bound on $s_{n,r}$

For any graph G on N vertices, the number of stable sets $i(G)$ satisfies

$$2^{\alpha(G)} \leq i(G) \leq \sum_{k=0}^{\alpha(G)} \binom{N}{k} \leq \left(\frac{eN}{\alpha(G)}\right)^{\alpha(G)}$$

Theorem (Hoffman)

Let $G = (V, E)$ be d-regular, on N vertices, with smallest eigenvalue $-\lambda$. Then

$$\alpha(G) \leq \frac{\lambda}{d+\lambda} N$$

For $G = J(n, r)$, we have $d = r(n-r)$ and $\lambda = r$ if $r \leq n/2$. Hence

$$\frac{1}{n}\binom{n}{r} \leq \log s_{n,r} = \log i(J(n,r)) \leq \frac{1}{n-r+1}\binom{n}{r}\log(e(n-r+1))$$

Counting stable sets in regular graphs

Let $G = (V, E)$ be d-regular, on N vertices, with smallest eigenvalue $-\lambda$. Put

$$\alpha := \frac{\lambda}{d+\lambda}, \sigma := \frac{\ln(d+1)}{d+\lambda}$$

Theorem

$$i(G) \leq (\frac{e}{\sigma})^{\sigma N} \cdot 2^{\alpha N}$$

Lemma (Alon, Balogh, Morris, and Samotij)

If $U \subseteq V$, then there exist sets $S, A \subseteq V$ such that

- $S \subseteq U \subseteq S \cup N(S) \cup A$
- S *uniquely determines* A
- $|S| \leq \sigma N$ *and* $|A| \leq \alpha N$

Proof of the Lemma

Lemma (Alon, Balogh, Morris, and Samotij)

If $U \subseteq V$, then there exist sets $S, A \subseteq V$ such that

- $S \subseteq U \subseteq S \cup N(S) \cup A$
- *S uniquely determines A*
- $|S| \leq \sigma N$ *and* $|A| \leq \alpha N$

Fix linear ordering $<$ of V.

Construction of S, A

- Put $A \leftarrow V$, $S \leftarrow \emptyset$
- while $|A| > \alpha N$:
  - let v be the $<$-maximal vertex among max. degree vertices in $G[A]$
  - if $v \in U$, put $S \leftarrow S + v$ and $A \leftarrow A \setminus (N(v) + v)$
  - if $v \notin U$, put $A \leftarrow A - v$

The lemma is direct from this construction, except that $|S| \leq \sigma N$ at termination.

Why S is small..

Fix linear ordering $<$ of V.

Construction of S, A

- Put $A \leftarrow V$, $S \leftarrow \emptyset$
- while $|A| > \alpha N$:
  - let v be the $<$-maximal vertex among max. degree vertices in $G[A]$
  - if $v \in U$, put $S \leftarrow S + v$ and $A \leftarrow A \setminus (N(v) + v)$
  - if $v \notin U$, put $A \leftarrow A - v$

Lemma (Alon&Chung; Haemers)

For any $\varepsilon > 0$, if $|A| = (\alpha + \varepsilon)N$, then $G[A]$ contains a vertex of degree at least $\varepsilon(d + \lambda)$.

Corollary (Bansal, P., van der Pol)

At termination, $|S| \leq \sigma N$.

Lemma (Alon, Balogh, Morris, and Samotij)

If $U \subseteq V$, then there exist sets $S, A \subseteq V$ such that

- $S \subseteq U \subseteq S \cup N(S) \cup A$
- *S uniquely determines A*
- $|S| \leq \sigma N$ *and* $|A| \leq \alpha N$

Theorem

$$i(G) \leq (\frac{e}{\sigma})^{\sigma N} \cdot 2^{\alpha N}$$

Proof.

- Let $U \subseteq V(G)$ be a stable set of G
- pick S, A as in the lemma; U is stable, so $S \subseteq U \subseteq S \cup A$

$$i(G) \leq (\# \text{ of possible } S) \cdot (\# \text{ of possible } A \cap U \text{ given } S) \leq \left(\sum_{k=0}^{\sigma N} \binom{N}{k} \right) \cdot 2^{\alpha N}$$

□

Bound on the number of sparse paving matroids

Theorem

$$\log s_n \leq \frac{2}{n}\binom{n}{n/2}(1+o(1)) \text{ as } n \to \infty$$

Proof.

By the Theorem

$$s_{n,r} = i(J(n,r)) \leq (\frac{e}{\sigma})^{\sigma N} \cdot 2^{\alpha N}$$

where (for $r \leq n/2$):

$$\alpha = \frac{1}{n-r+1}, \ \sigma = \frac{\ln(r(n-r)+1)}{r(n-r+1)}, \ N = \binom{n}{r}$$

The bound on $s_n = \sum s_{n,r}$ is dominated by $r \approx n/2$. Then $\alpha \approx \frac{2}{n}$, $\sigma \approx \frac{8 \ln n}{n^2}$. □

How about general matroids?

Lemma (Alon, Balogh, Morris, and Samotij)

If $U \subseteq V$, then there exist sets $S, A \subseteq V$ such that

- $S \subseteq U \subseteq S \cup N(S) \cup A$
- *S uniquely determines A*
- $|S| \leq \sigma N$ *and* $|A| \leq \alpha N$

Consider a matroid $M = (E, \mathcal{B})$ of rank r.

To encode M, we put $U = \binom{E}{r} \setminus \mathcal{B}$ and apply the Lemma to get S, A. Then

$$S \subseteq U \subseteq S \cup N(S) \cup A$$

We have

$$m_{n,r} \leq (\#\text{ of possible } S) \cdot (\#\text{ of possible } N(S) \cap U) \cdot (\#\text{ of possible } A \cap U \text{ given } S)$$

Matroids have small local covers II

Let $M = (E, \mathcal{B})$ be a matroid of rank r. Let $X \in \binom{E}{r}$.

Lemma

There exists a set of flats $\mathcal{Z}_X$ of M covering each dependent $Y \in N(X)$, such that $|\mathcal{Z}_X| \leq r$.

Lemma

Suppose X is dependent.
There exists a set of flats $\mathcal{Z}_X$ of M covering each dependent $Y \in N(X)$, such that $|\mathcal{Z}_X| \leq 2$.

Proof.

- If $r_M(X) < r - 1$, put $\mathcal{Z}_X = \{\text{cl}(X)\}$.
- If $r_M(X) = r - 1$, X contains a unique circuit C, is disjoint from a unique cocircuit D. Put $\mathcal{Z}_X = \{\text{cl}(C), E \setminus D\}$

□

Bound on the number of matroids

Theorem (Bansal, P., van der Pol , 2012)

$$\log m_n \leq \frac{2}{n}\binom{n}{n/2}(1 + o(1)) \text{ as } n \to \infty$$

Proof.

To encode M, we put $U = \binom{E}{r} \setminus \mathcal{B}$ and apply the Lemma to get S, A. Then

$$S \subseteq U \subseteq S \cup N(S) \cup A$$

The number of possible $(S, A \cap U)$ is bounded as before.
For each $X \in S$, we make a local cover $\mathcal{Z}_X$ s.t. $|\mathcal{Z}_X| \leq 2$ and put

$$\mathcal{Z} := \bigcup_{X \in S} \mathcal{Z}_X$$

Then $\mathcal{Z}$ determines $N(S) \cap U$, and $|\mathcal{Z}| \leq 2|S| \leq 2\sigma N$, bounding the number of $N(S) \cap U$. □

In the proof, we encode U as a triple

$$S, N(S) \cap U, A \cap U$$

- there are $\leq \sum_{k=0}^{\sigma N} \binom{N}{k} \leq (e/\sigma)^{\sigma N}$ possibilities for S
- there are $\leq \sum_{k=0}^{2\sigma N} \binom{Nn}{k} \leq (en/\sigma)^{2\sigma N}$ possibilities for $N(S) \cap U$
- there are $\leq 2^{\alpha N}$ possibilities for $A \cap U$ (given S)

The case $r \approx n/2$ again dominates the bound, and then

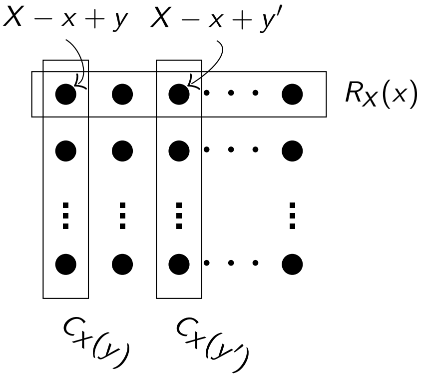
$$\alpha \approx \frac{2}{n}, \ \sigma \approx \frac{8 \ln n}{n^2}$$

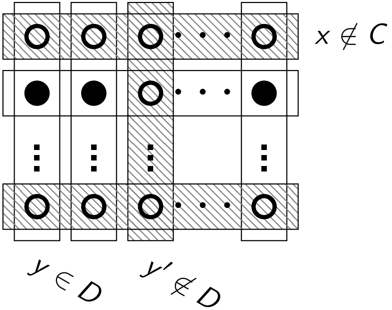
So the bottleneck is the bound on the number of possible $A \cap U$.

Can we further compress $A \cap U$ to get a better bound?

The neighborhood $N(X)$ in the Johnson graph

The sets $R_X(x) := \{X - x + y : y \in E \setminus X\}$ and $C_X(y) := \{X - x + y : x \in X\}$ are cliques.





Lemma

Let $M = (E, \mathcal{B})$ be a matroid, let $X \in U = \binom{E}{r} \setminus \mathcal{B}$ and $x \in X \not\ni y$. If $R_X(x) \setminus U \neq \emptyset$ and $C_X(y) \setminus U \neq \emptyset$, then the set $U \cap N(X)$ is determined by $R_X(x) \cap U$ and $C_X(y) \cap U$.

A better Lemma

Lemma (Alon, Balogh, Morris, and Samotij)

If $U \subseteq V$, then there exist sets $S, A \subseteq V$ such that

- $S \subseteq U \subseteq S \cup N(S) \cup A$
- S *uniquely determines* A
- $|S| \leq \sigma N$ *and* $|A| \leq \alpha N$

For any fixed k we have this variant:

Lemma

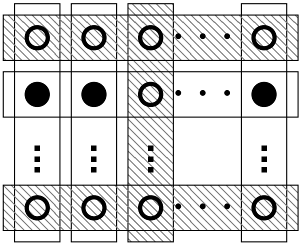
If $U \subseteq V$, then there exist sets $S, A \subseteq V$ such that

- $S \subseteq U \subseteq S \cup N(S) \cup A$
- S *uniquely determines* A
- $|S| \leq (\sigma + \alpha/k)N$, $|A| \leq \alpha N$, *max. degree in* $G[A]$ *is* $< k$

Lemma

If $U \subseteq V$, then there exist sets $S, A \subseteq V$ such that

- $S \subseteq U \subseteq S \cup N(S) \cup A$
- *S uniquely determines A*
- $|S| \leq (\sigma + \alpha/k)N$, $|A| \leq \alpha N$, $\Delta(G[A]) < k$



$N(X)$, X dependent

Theorem (P., van der Pol, 2014)

$\log m_n = (1 + o(1)) \log s_n$ *as* $n \to \infty$

Proof.

- To encode a matroid $M = (E, \mathcal{B})$ on n of rank r , we put $U = \binom{E}{r} \setminus \mathcal{B}$.
- Apply the Lemma with $k = \min\{r, n - r\}$. Then $S \subseteq U \subseteq S \cup N(S) \cup A$
- As before a cover $\mathcal{Z}$ of size at most $2|S|$ determines $N(S) \cap U$.
- *Pick a certain stable set $T \subseteq A \cap U$. Then $A \cap U$ is determined by $(S, T, \mathcal{Z})$.*
- Now $\log m_{n,r} \leq \log(\#\text{ of } T) + \log(\#\text{ of } (S, \mathcal{Z})) \leq \log s_{n,r} + $ (relatively small)

□

Last words

Conjecture

Asymptotically all matroids M on n elements have $\lfloor n/2 \rfloor \leq r(M) \leq \lceil n/2 \rceil$.

Theorem

There is a β so that asymptotically almost all sparse paving matroids M on n elements have

$$n/2 - \beta\sqrt{n} \leq r(M) \leq n/2 + \beta\sqrt{n}$$

Corollary

There is a β' so that asymptotically almost all matroids M on n elements have

$$n/2 - \beta'\sqrt{n} \leq r(M) \leq n/2 + \beta'\sqrt{n}$$

We know virtually nothing about the similarity of adjacent Johnson graphs. E.g.

Conjecture

For any n and $r < r' \leq n/2$, we have $s_{n,r} \leq s_{n,r'}$.

Thank you