

Outline



2 Null vector asymptotics

▲□▶ ▲□▶ ▲□▶ ▲□▶ ▲□ ● のへぐ

3 Hypergraph 2-core

Outline



2 Null vector asymptotics

3 Hypergraph 2-core

◆□▶ ◆□▶ ◆臣▶ ◆臣▶ 臣 のへで

Create an $m \times n$ matrix M := M(n, m) over GF[2] by generating *m* i.i.d. rows each with *n* entries of 0s and 1s.

Each row has weight (number of 1s) independently distributed as some random variable $W_n \in \{1, 2, 3, ..., n\}$.

(ロ) (同) (三) (三) (三) (○) (○)

Given its weight, the row is chosen uniformly over all possibilities in $\{0,1\}^n$ with that many 1s.

Create an $m \times n$ matrix M := M(n, m) over GF[2] by generating *m* i.i.d. rows each with *n* entries of 0s and 1s.

Each row has weight (number of 1s) independently distributed as some random variable $W_n \in \{1, 2, 3, ..., n\}$.

Given its weight, the row is chosen uniformly over all possibilities in $\{0, 1\}^n$ with that many 1s.

A left null vector for *M* over GF[2] is a row vector $a \in \{0, 1\}^m$ such that $aM \equiv \mathbf{0}_n \mod 2$, where $\mathbf{0}_n$ is the row vector of *n* zeros. Trivially, $\mathbf{0}_m$ is always a null vector.

A non-trivial null vector corresponds to a non-empty subset of the row labels $\{1, \ldots, m\}$ such that the sum over the corresponding rows has no odd entries.



◆□> ◆□> ◆豆> ◆豆> ・豆 ・ 釣べ⊙





◆□▶ ◆□▶ ◆臣▶ ◆臣▶ ●臣 - の々で



















- 4 日 > 4 個 > 4 画 > 4 画 > - 画 - の Q ()



Outline



2 Null vector asymptotics

▲□▶ ▲□▶ ▲□▶ ▲□▶ ▲□ ● のへぐ

3 Hypergraph 2-core

Number of null vectors

We are interested in

N(n, m) = number of left null vectors for M(n, m)

(日) (日) (日) (日) (日) (日) (日)

in the regime where $n \to \infty$, $m = m_n$ with $m_n/n \to \alpha$, and $W_n \xrightarrow{d} W$ for some limiting weight distribution W.

Number of null vectors

We are interested in

N(n, m) = number of left null vectors for M(n, m)

in the regime where $n \to \infty$, $m = m_n$ with $m_n/n \to \alpha$, and $W_n \xrightarrow{d} W$ for some limiting weight distribution W.

Note:

no non-trivial null vectors $\Leftrightarrow N(n,m) = 1 \Leftrightarrow$ system $Mx \equiv y \mod 2$ has a solution $x \in \{0,1\}^n$ for every $y \in \{0,1\}^m$

Also related to XORSAT, spin-glass models, Ehrenfest urn, random walk on hypercube, switch-setting problems, ...

Expected number of null vectors

Let $\mathbf{1}_m$ denote the vector of m ones, and consider the event

$$A(n,m) = \{\mathbf{1}_m \text{ is null for } M(n,m)\}\$$

= {all column sums of $M(n,m)$ even}.

Then

$$\mathbb{E} N(n,m) = \sum_{k=0}^{m} \mathbb{E} \text{ number of null vectors of weight } k$$
$$= \sum_{k=0}^{m} \binom{m}{k} \mathbb{P} (A(n,k)).$$

◆□▶ ◆□▶ ◆□▶ ◆□▶ ● ● ● ●

We study asymptotics of $\mathbb{E}(n, \alpha n)$

Expected number of null vectors

Let $\mathbf{1}_m$ denote the vector of m ones, and consider the event

$$A(n,m) = \{\mathbf{1}_m \text{ is null for } M(n,m)\}\ = \{ all column sums of $M(n,m) \text{ even} \}.$$$

Then

$$\mathbb{E} N(n,m) = \sum_{k=0}^{m} \mathbb{E} \text{ number of null vectors of weight } k$$
$$= \sum_{k=0}^{m} {m \choose k} \mathbb{P} (A(n,k)).$$

We study asymptotics of $\mathbb{E}(n, \alpha n)$ in terms of

$$\rho(s) := \mathbb{E} s^{W}$$

< □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □

the generating function of the limiting weight distribution.

Technical device

We can show that any model in our class $W_n \xrightarrow{d} W$ is sufficiently well-approximated by the binomial model in which

 $W_n = W_n^{\text{bin}} =$ number of odd components in multinomial $(W; \frac{1}{n}, \dots, \frac{1}{n})$

The binomial model is equivalent to:

Generate each row by distributing W units uniformly and independently among the n positions and then reducing occupancies mod 2.

Lemma

In the binomial model,

$$\mathbb{P}(A(n,m)) = 2^{-n} \sum_{j=0}^{n} \binom{n}{j} \left(\rho(1-\frac{2j}{n})\right)^{m}.$$

Exact formula

Proof.

Let Y_i = sum of column j and, for $J \subseteq \{1, 2, ..., n\}$, $S_{i,J}$ = sum of entries in column *i* and row subset *J*.

Then

$$\mathbb{P}(A(n,m)) = \mathbb{P}(\text{all } Y_j \text{ even}) = \mathbb{E}\prod_{j=1}^n \left(\frac{1+(-1)^{Y_j}}{2}\right)$$
$$= 2^{-n} \sum_J \mathbb{E}\left[(-1)^{\sum_{j \in J} Y_j}\right].$$

But $\sum_{i \in J} Y_i = \sum_{i=1}^m S_{i,J}$ (i.i.d. summands) so

$$\mathbb{P}(A(n,m)) = 2^{-n} \sum_{J} \left(\mathbb{E}\left[(-1)^{S_{1,J}} \right] \right)^{m}.$$

But $S_{1,J}$ is given by the sum of |J| components of a multinomial $(W; \frac{1}{n}, \dots, \frac{1}{n})$ vector reduced mod 2, so $(-1)^{S_{1,J}} \stackrel{d}{=} (-1)^Z$ where $Z \sim \text{Bin}(W, \frac{|J|}{n}).$ うつつ 川 へきゃくきゃくむゃ

Lemma

Suppose that $m_n/n \to \alpha \in (0, \infty)$ and that either (i) m_n is even for all n, or (ii) $\mathbb{P}[W \text{ is even}] > 0$. Then,

$$\lim_{n\to\infty}n^{-1}\log\mathbb{P}\left(A(n,m_n)\right)=-R_{\rho}(\alpha),$$

for a continuous and non-decreasing function R_{ρ} given by

$${\it R}_{
ho}(lpha) = -\log \sup_{\gamma \in [0,1/2]} \left(rac{(
ho(1-2\gamma))^lpha}{2\gamma^\gamma(1-\gamma)^{1-\gamma}}
ight).$$

◆□▶ ◆□▶ ◆□▶ ◆□▶ ● ● ● ●

Lemma

Suppose that $m_n/n \to \alpha \in (0, \infty)$ and that either (i) m_n is even for all n, or (ii) $\mathbb{P}[W \text{ is even}] > 0$. Then,

$$\lim_{n\to\infty}n^{-1}\log\mathbb{P}\left(A(n,m_n)\right)=-R_{\rho}(\alpha),$$

for a continuous and non-decreasing function R_{ρ} given by

$$\mathcal{R}_{
ho}(lpha) = -\log \sup_{\gamma \in [0,1/2]} \left(rac{(
ho(1-2\gamma))^lpha}{2\gamma^\gamma(1-\gamma)^{1-\gamma}}
ight).$$

Corollary (cf. Kolchin '94)

Let π_n denote the probability that all the *n* components of a multinomial $(m_n; \frac{1}{n}, ..., \frac{1}{n})$ vector are even. Suppose that m_n is even for all *n* and $m_n/n \to \alpha = \lambda \tanh \lambda \in (0, \infty)$. Then

$$\lim_{n\to\infty} n^{-1} \log \pi_n = \log \cosh \lambda - (\lambda \tanh \lambda)(1 - \log \tanh \lambda).$$

Define the continuous non-decreasing function F_{ρ} by

$$\mathcal{F}_{
ho}(lpha) = \log \sup_{\gamma \in [0,1/2]} \left(rac{(1+
ho(1-2\gamma))^{lpha}}{2\gamma^{\gamma}(1-\gamma)^{1-\gamma}}
ight), \;\; lpha \geq 0.$$

< □ > < 同 > < 三 > < 三 > < 三 > < ○ < ○ </p>

Set $\alpha_{\rho}^* := \inf\{\alpha \ge \mathbf{0} : F_{\rho}(\alpha) > \mathbf{0}\}.$

Define the continuous non-decreasing function F_{ρ} by

$$F_{\rho}(\alpha) = \log \sup_{\gamma \in [0, 1/2]} \left(\frac{(1 + \rho(1 - 2\gamma))^{\alpha}}{2\gamma^{\gamma}(1 - \gamma)^{1 - \gamma}} \right), \ \alpha \ge 0.$$

Set $\alpha_{\rho}^* := \inf\{\alpha \ge \mathbf{0} : F_{\rho}(\alpha) > \mathbf{0}\}.$

Theorem Suppose that $m_n/n \to \alpha \in (0, \infty)$. Then, $\lim_{n \to \infty} n^{-1} \log \mathbb{E} N(n, m_n) = \begin{cases} 0 & \text{if } \alpha \le \alpha_{\rho}^*; \\ F_{\rho}(\alpha) > 0 & \text{if } \alpha > \alpha_{\rho}^*. \end{cases}$

◆□▶ ◆□▶ ◆□▶ ◆□▶ ● ● ● ●

Define the continuous non-decreasing function F_{ρ} by

$$F_{\rho}(\alpha) = \log \sup_{\gamma \in [0, 1/2]} \left(\frac{(1 + \rho(1 - 2\gamma))^{\alpha}}{2\gamma^{\gamma}(1 - \gamma)^{1 - \gamma}} \right), \ \alpha \ge 0.$$

Set $\alpha_{\rho}^* := \inf\{\alpha \ge \mathbf{0} : F_{\rho}(\alpha) > \mathbf{0}\}.$



Here, if $\mathbb{P}(W \ge 2) = 1$ and $\mathbb{E} W < \infty$, then $1/2 \le \alpha_{\rho}^* < 1$.

・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・
 ・

Define the continuous non-decreasing function F_{ρ} by

$$F_{\rho}(\alpha) = \log \sup_{\gamma \in [0, 1/2]} \left(\frac{(1 + \rho(1 - 2\gamma))^{\alpha}}{2\gamma^{\gamma}(1 - \gamma)^{1 - \gamma}} \right), \ \alpha \ge 0.$$

Set $\alpha_{\rho}^* := \inf\{\alpha \ge \mathbf{0} : F_{\rho}(\alpha) > \mathbf{0}\}.$



Here, if $\mathbb{P}(W \ge 2) = 1$ and $\mathbb{E} W < \infty$, then $1/2 \le \alpha_{\rho}^* < 1$.

In the fixed-weight case $\mathbb{P}(W = r) = 1$ versions of this result were obtained by BALAKIN *et al.* '92, KOLCHIN '94, CALKIN '97, COOPER '99.

First non-trivial null vector

Consider building M(n, m) one row at a time.

Let T_n be the first *m* for which a non-zero null vector appears, i.e., the first *m* for which row *m* is in the linear span of rows $1, \ldots, m-1$.

Theorem

Suppose $3 \le W_n \le B$ for some $B < \infty$. Then, with probability tending to 1, $\alpha_{\rho}^* < T_n/n < \underline{\alpha}_{\rho}$.

Here $\underline{\alpha}_{\rho} \leq 1$ is another threshold defined in terms of the 2-core of the hypergraph.

◆□▶ ◆□▶ ◆□▶ ◆□▶ ● ● ● ●

First non-trivial null vector

Consider building M(n, m) one row at a time.

Let T_n be the first *m* for which a non-zero null vector appears, i.e., the first *m* for which row *m* is in the linear span of rows $1, \ldots, m-1$.

Theorem

Suppose $3 \le W_n \le B$ for some $B < \infty$. Then, with probability tending to 1, $\alpha_{\rho}^* < T_n/n < \underline{\alpha}_{\rho}$.

Here $\underline{\alpha}_{\rho} \leq 1$ is another threshold defined in terms of the 2-core of the hypergraph.

Conjecture

Under the above conditions, $n^{-1}T_n \rightarrow \underline{\alpha}_{\rho}$ in probability.

An equivalent version of this conjecture in the fixed weight case has been proved by DUBOIS & MANDLER '02, DIETZFELBINGER *et al.* '10.

Fixed weight case

When $\mathbb{P}(W = r) = 1$, i.e., $\rho(s) = s^r$.

Write α_r and $\underline{\alpha}_r$ for α_ρ and $\underline{\alpha}_\rho$ in this case.

Following BALAKIN *et al.* '92, KOLCHIN '94, CALKIN '97, COOPER '99, '04

r	1	2	3	4	5	6	7
α_r^*	0	0.5	0.889493	0.967147	0.989162	0.996228	0.998650
$\underline{\alpha}_r$	—	—	0.917935	0.976770	0.992438	0.997380	0.999064

< □ > < 同 > < 三 > < 三 > < 三 > < ○ < ○ </p>

Fixed weight case

When $\mathbb{P}(W = r) = 1$, i.e., $\rho(s) = s^r$.

Write α_r and $\underline{\alpha}_r$ for α_ρ and $\underline{\alpha}_\rho$ in this case.

Following BALAKIN *et al.* '92, KOLCHIN '94, CALKIN '97, COOPER '99, '04

r	1	2	3	4	5	6	7
α_r^*	0	0.5	0.889493	0.967147	0.989162	0.996228	0.998650
Ωr	—	—	0.917935	0.976770	0.992438	0.997380	0.999064

Lemma

•
$$1 - \alpha_r^* \sim \frac{e^{-r}}{\log 2}$$
 (Calkin '97);

•
$$1 - \underline{\alpha}_r \sim e^{-r}$$
.

• So
$$\alpha_r^* < \underline{\alpha}_r < 1$$
 for large enough r.

Outline

1 Introduction

2 Null vector asymptotics

3 Hypergraph 2-core

Hypergraph 2-core

Starting from the hypergraph represented by M(m, n), iterate the following:

Identify a column (vertex) with a single non-zero. Delete the corresponding row (hyperedge).

When no column (vertex) with a single non-zero remains, the resulting hypergraph is the 2-core of the original hypergraph.

< □ > < 同 > < 三 > < 三 > < 三 > < ○ < ○ </p>

Hypergraph 2-core

Starting from the hypergraph represented by M(m, n), iterate the following:

Identify a column (vertex) with a single non-zero. Delete the corresponding row (hyperedge).

When no column (vertex) with a single non-zero remains, the resulting hypergraph is the 2-core of the original hypergraph.

Lemma (Molloy, Cooper)

If the 2-core is empty, the hypergraph has no hypercycle. If the 2-core has more rows (hyperedges) than columns (vertices), the hypergraph contains a hypercycle.

Hypergraph 2-core

By a (minor extension of a) theorem of DARLING & NORRIS '08 one can characterize

- those α for which the 2-core asymptotically has more rows than columns,
- or vice versa,
- and hence $\underline{\alpha}_{\rho}$,

depending on whether

$$\psi(\boldsymbol{g}^*(\alpha)) < \mathsf{0} \quad \text{or} \quad \psi(\boldsymbol{g}^*(\alpha)) > \mathsf{0}.$$

Here

$$egin{aligned} egin{split} egin{split} egin{aligned} g^*(lpha) &= \sup\left\{x\in(0,1):-rac{\log(1-x)}{
ho'(x)}\leqlpha
ight\};\ \psi(x) &= x+\left(1+rac{
ho(x)}{
ho'(x)}-x
ight)\log(1-x). \end{split}$$

◆□▶ ◆□▶ ◆臣▶ ◆臣▶ 三臣 - 釣��

Hypergraph 2-core: non-monotonicity

In the fixed weight case, COOPER '04 showed that the core aspect ratio transition is monotone. In the random setting, more can occur.



Hence as α ranges in (0, 1), $\psi(g^*(\alpha))$ changes sign from +ve, to -ve, to +ve, and finally to -ve again.

References

- G.V. BALAKIN, V.F. KOLCHIN & V.I. KHOKHLOV, Hypercycles in a random hypergraph. *Discrete Math. Appl.* (1992).
- N.J. CALKIN, Dependent sets of constant weight binary vectors. *Combinat. Probab. Comput.* (1997).
- C. COOPER, Asymptotics for dependent sums of random vectors. *Random Struct. Alg.* (1999).
- C. COOPER, The cores of random hypergraphs with a given degree sequence. *Random Struct. Alg.* (2004).
- R.W.R. DARLING, D.A. LEVIN & J.R. NORRIS, Continuous and discontinuous phase transitions in hypergraph processes. *Random Struct. Alg.* (2004).
- R.W.R. DARLING & J.R. NORRIS, Structure of large random hypergraphs. *Ann. Appl. Probab.* (2005).
- R.W.R. DARLING & J.R. NORRIS, Differential equation approximations for Markov chains. *Probab. Surv.* (2008).
- R.W.R. DARLING, M.D. PENROSE, A.R. WADE & S.L. ZABELL, Rank deficiency in sparse random GF[2] matrices. *Electr. J. Probab.* (2012).
- M. DIETZFELBINGER *et al.*, Tight thresholds for cuckoo hashing via XORSAT. *Proc. 37th Inl. Collog. Autom. Lang. Prog.*, 2010.
- O. DUBOIS & J. MANDLER, The 3-XORSAT threshold. *Proc. 43rd Annual IEEE Symp. Foun. Comp. Sci.*, 2002.
- V.F. KOLCHIN, Random graphs and systems of linear equations in finite fields. *Random Struct. Alg.* (1994).