

# Random discrete matrices

Van H. Vu

Department of Mathematics

Rutgers

vanvu@math.rutgers.edu

( survey of the same title on [www.math.rutgers.edu/~vanvu](http://www.math.rutgers.edu/~vanvu))

**The models.**

$M_n$  (non-symmetric):  $n$  by  $n$  matrix with i.i.d (in many consideration this can be significantly weakened) entries:  $\xi_{ij}$ .

$Q_n$  (symmetric):  $\xi_{ij} = \xi_{ji}$ .

*Continuous models:*  $\xi_{ij}$  have continuous distribution. Representative example: Gaussian.

*Discrete models:*  $\xi_{ij}$  have discrete distribution. Representative example: Bernoulli ( $\pm 1$  with probability  $1/2$ ).

**The questions:**

- (1) Spectral norm ( $\|M_n\|$  = the largest singular value)
- (2) Rank of  $M_n$ .
- (3) Probability of Singularity.
- (4) Determinant.
- (5) Condition number ( $\|M_n\|\|M_n^{-1}\|$ ).
- (6) Limiting distribution of the spectra.

*Continuous models.* Precise answers. Established theory.

(1) Joint distribution of the eigenvalues (Ginibre 1962)

$$p(\lambda_1, \dots, \lambda_n) = c_n \prod_{[i < j]} |\lambda_i - \lambda_j|^2 \prod_{i=1}^n e^{-n|\lambda_i|^2}.$$

(2) Moment method (Wigner 1955)

$$\sum_{i=1}^n \lambda_i^k = \text{Trace} M_n^k.$$

(3) Stieltjes transform:

$$s_n(z) := \frac{1}{n} \text{Trace} \left( \frac{1}{\sqrt{n}} M_n - z I_n \right)^{-1} = \frac{1}{n} \sum_{i=1}^n \frac{1}{n^{-1/2} \lambda_i - z}.$$

*Discrete models.*

(1) Not available; Use of (2)(3) are more limited.

**New method**

(4) Concentration function:  $v = (a_1, \dots, a_n)$

$$P_v := \max_x \mathbf{P}\left(\sum_{i=1}^n a_i \xi_i = x\right).$$

Littlewood-Offord- Erdős (1943) If  $a_i \neq 0$ , then  $P_v = O(n^{-1/2})$ .

We developed a small theory around the notion of concentration function, focusing on Inverse statements. The main tool is [additive combinatorics](#).

*The rank problem.*

**Theorem.** (Komlós 1967) Almost surely  $M_n$  has full rank (or is non-singular).

**Conjecture.** (Weiss 80s) The same holds for symmetric matrices, i.e., almost surely  $Q_n$  has full rank.

**Theorem.** (Costello-Tao-V., 2004) Almost surely  $Q_n$  has full rank.

$Q_n$  can be seen as the adjacency matrix of the random graph  $G(n, 1/2)$  (switching  $-1$  to  $0$  does not impact the rank). So the above theorem can be rewritten as

**Theorem.** Almost surely  $G(n, 1/2)$  has full rank.

**Theorem.** (Costello-V. 2006) For  $p > (1 + o(1)) \log n/n$ ,  $G(n, p)$  a.s. has full rank.

**Theorem.** (Costello-V. 2006) For  $p > (1/2 + o(1)) \log n/n$ , a.s. the rank of  $G(n, p)$  equal  $n$  minus the number of isolated vertices.

*Quadratic Littlewood-Offord theorem.* Consider the quadratic form

$$Q(\xi_i) := \sum_{1 \leq i, j \leq n} a_{ij} \xi_i \xi_j.$$

**Theorem.** (Costello-Tao-V.) If  $a_{ij} \neq 0$ , then for any  $x$

$$P(Q = x) = O(n^{-1/4}).$$

**Conjecture.**  $O(n^{-1/4})$  can be replaced by  $O(n^{-1/2})$ .

If true, it is sharp, as one can take  $Q = (\xi_1 + \cdots + \xi_n)^2$ .

**Conjecture.** Almost surely a random regular graph with degree at least 3 has full rank.

*The singularity problem.* Estimate  $p_n$ , the probability that  $M_n$  is singular.

By considering the probability that there are two equal rows

$$p_n \geq (1/2 + o(1))^n.$$

**Conjecture.**  $p_n = (1/2 + o(1))^n$ .

**Theorem.** (Komlós 1967)  $p_n = o(1)$ .

**Theorem.** (Komlós 1976)  $p(n) = O(n^{-1/2})$ .

**Theorem.** (Kahn-Komlós-Szemerédi 1995)  $p_n = O(.999^n)$ .

**Theorem.** (Tao-V. 2004)  $p_n = O(.96^n)$ .

**Theorem.** (Tao-V. 2005)  $p_n = (3/4 + o(1))^n$ .

**Theorem.** (Bourgain-V.-Wood 2007)  $p_n = (\sqrt{1/2} + o(1))^n$ .

*Some other models.*

Instead of  $M_n$  consider the following "lazy" model  $M_n^{lazy}$ . The entries of  $M_n^{lazy}$  are i.i.d random variables which equal zero with probability half and 1 and  $-1$  with probability one quarter. (If one thinks of the entries of  $M_n$  as fair coin flips, then in the "lazy" model about half of the time we are lazy and simply write zero instead flipping a coin.) It is clear that for the lazy model the singular probability  $p_n^{lazy}$  is again at least  $(1/2 + o(1))^n$  (which is the probability that there is a zero row).

**Theorem.** (Bourgain-V.-Wood 2007)  $p_n^{lazy} = (1/2 + o(1))^n$ .

For a general result that covers the last two theorems, see Friday's lecture.

*Random walks and Lazy random walks.*

Let

$$S := \sum_{i=1}^n a_i \xi_i \text{ and } S^\mu := \sum_{i=1}^n a_i \xi_i^\mu.$$

Consider  $P(S = 0)$  and  $P(S^\mu = 0)$ . Intuitively, the second probability is much larger. However, there are cases where the two probabilities are comparable. For example, if all  $a_i = 1$ , then both probabilities are  $\Theta(n^{-1/2})$ .

The core of our method for the singularity problem is a theorem that **characterizes all sets  $a_i$  where the two probabilities are comparable.** (Even when both of them are **exponentially small.**) It relies on a method of Halasz (1975) and many arguments from additive combinatorics.

*Determinant.*  $|DetM_n|$ .

**Fact 1.** Komlós (1967) result implies that a.s.  $|DetM_n|$  is positive. In fact, since  $|DetM_n|$  is divisible by  $2^{n-1}$ , it is at least  $2^{n-1}$ .

**Fact 2.** (Hadamard bound)  $|DetM_n| \leq n^{n/2}$ .

**Fact 3.** Turán (1940s) observed that  $E(Det^2M_n) = n! = n^{(1+o(1))n}$ .

**Proof.** Linearity of expectation.

**Conjecture.** A.s.  $|DetM_n| = n^{(1/2+o(1))n}$ .

**Theorem.** (Tao-V. 2003) A.s.  $|DetM_n| = n^{(1/2+o(1))n}$ . In fact, a.s.

$$|DetM_n| \geq \sqrt{n!} \exp(-29\sqrt{n \log n}).$$

The main idea here is to view the determinant as the volume of the parallelepiped spanned by the row vectors of the matrix. Now expose the matrix row by row and compute the volume as the product of the distances from the  $(i+1)$ st vector to the plane spanned by the first  $i$  vectors.

**Lemma.** With very high probability  $d_i \approx \sqrt{n-i}$ .

**Conjecture.** A.s.  $|DetQ_n| = n^{(1/2+o(1))n}$ .

The main difficulty with symmetric matrices is that the rows are **no longer dependent**.

*The condition number problem.*

Let  $M$  be an  $n \times n$  matrix, the *condition number*  $\kappa(M)$  is defined as

$$\kappa(M) := \|M\| \|M^{-1}\|.$$

where  $\|\cdot\|$  is the spectral norm. If  $M$  is singular,  $\kappa(M) = \infty$ .

The condition number plays a crucial role in numerical linear algebra. For example, the accuracy and stability of most algorithms used to solve the equation  $Mx = b$  depend on  $\kappa(M)$ .

$\|M_n\|$  is, with very high probability,  $\Theta(\sqrt{n})$ . It is much harder to estimate  $\|M_n^{-1}\|$ .

A more general and more practical problem is to estimate  $\kappa(A + M_n)$ , where  $A$  is a fixed matrix.

*Motivation.* **Why the simplex algorithm runs fast ?** Smooth Analysis (Spielman-Teng 2000).

Key point: **Noise helps !!**  $\kappa(A)$  may be large, but  $\kappa(A + Noise)$  is almost surely small.

Spielman-Teng assumed that Noise is random Gaussian and proved that

**Theorem.** (Spielman-Teng 2000: Gaussian Noise is good) **Assuming that**  $\|A\| = n^{O(1)}$ , then a.s.  $\kappa(A + Noise) = n^{O(1)}$ .

**Question.** (S-T) **What happens with discrete noise ?**

*Special case.*  $A = 0$ ,  $Noise = M_n$ . Rudelson (2005), Tao-V. (2005) proved Spielman-Teng statement for this case. Recently Rudelson-Vershynin (2007) obtained a very precise estimate.

**Theorem.** (Rudelson-Vershynin 2007)  $P(\kappa(M_n) \geq Cn) = O(C^{-1})$ .

*General case.* Any  $A$ , any discrete noise:

**Theorem.** (Tao-V. 2006: Discrete Noise is good) Assuming that  $\|A\| = n^{O(1)}$ , then with very high probability  $\kappa(A + Noise) = n^{O(1)}$ .

In fact, the assumptions are quite general and address well some real life situations. For instance, the entries of Noise do not need to be i.i.d ( this corresponds to: **Noise occurring to a large entry usually has larger variance**). Furthermore, one can allow lots of entries of Noise to be zero (this corresponds to: **Certain entries in  $A$  are noise-free**).

*Main tool.* **Inverse Littlewood-Offord theorem.** Recall the definition

$$P_v := \max_x \mathbf{P}\left(\sum_{i=1}^n \xi_i a_i = x\right).$$

Littlewood-Offord (1943) (Erdős ) If  $a_i \neq 0$ , then  $P_v = O(n^{-1/2})$ .

The bound is sharp. One can take  $a_i = 1$ .

If one forbids the  $a_i$  be the same, then the bound jumps quite a bit

**Theorem.** (Erdős-Moser, Sárközi-Szemerédi, Stanley 80s) **If the  $a_i$  are different, then  $P_v = O(n^{-3/2})$ .**

Again, this bound is sharp. One can take  $a_i = i$

**Inverse Question.** **When is  $P_v$  large ?**

For instance, if we know  $P_v \geq n^{-5}$ , what can we say about the  $a_i$ . We obtained a complete answer for this question.

Example:  $v = \{a_i\}$  is a subset of integers in a (symmetric) interval  $I$  of length  $M$ . The the random sum  $\sum_i a_i \xi_i$  takes value in the interval  $nI$ . Thus by the pigeon hole principle:

$$P_v \geq (nM)^{-1}.$$

So if  $v$  is a subset of a short interval (arithmetic progression), then  $P_v$  is large.

(Inverse) Question: **Given that  $P_v$  is large, is it true that  $v$  is a subset of a short arithmetic progression ?**

This has the spirit of Freiman's theorem from additive combinatorics.  
For the precise answer: Sunday's lecture.

*Limiting distribution of the spectra.*

**Theorem.** (Wigner 1955) The distribution of the eigenvalues of  $Q_n$  follows the semi-circle law.

Wigner proof introduced the Trace method.

**Conjecture.** The distribution of the eigenvalues of  $M_n$  follows the circular law.

Let  $A_n$  be a random matrix with i.i.d. entries having mean 0 and variance 1.

**Conjecture.** (Circular Law Conjecture) The distribution of the eigenvalues of  $A_n$  follows the circular law.

The statement is true for Gaussian (Ginibre, Mehta). For general continuous models, Girko (1984), Bai (1997) (with assumption on the  $(2 + \epsilon)$ th moment) .

The symmetric (semi-circle) version of the last conjecture was proved by Pastur in the 1960s.

Tao-V. (2007), Gotze-Tikhomirov (2007)

**Theorem.** The distribution of the eigenvalues of  $M_n$  follows the circular law.

In fact this follows rather quickly from the Noise-is-Good theorem.

**Extension.** Gotze-Tikhomirov proved CL for entries with mean 0 having sub-gaussian tails (exponential decay, in particular all moments should be bounded).

Tao-V. needs to assume bounded  $(2 + \epsilon)$ th moment (so very close to the main conjecture; plus a “possible” additional condition).