Singularity of a matrix

Consider an $n \times n$ matrix $M$ as a collection of $n$ rows $r_i \in \mathbb{R}^n$. 

$$\text{rank}(M) = \dim(\text{span}\{r_i\}_{i \in [n]})$$

- Full rank: $\text{rank}(M) = n$
- Singular: $\text{rank}(M) < n$

A set of vectors $\{r_i\}_{i \in S}$ is a dependency if there exist non-zero coefficients $\{a_i\}_{i \in S}$ such that

$$\sum_{i \in S} a_i r_i = 0.$$ 

For continuous random matrices, $\mathbb{P}(\text{Singular}) = 0$.

In discrete random matrices, singularity comes from small dependencies.
Let $R_n = (x_{ij})_{1 \leq i, j \leq n}$ have iid $x_{ij}$ r.v.

**Thm** [Komlós, 1967] If $x_{ij}$ are uniform in $\{0, 1\}$ then

$$P(\text{rank}(R_n) < n) \leq n^{-1/2}.$$ 

Two equal rows form a dependency.

$$P(\text{rank}(R_n) < n) \geq P(r_1 = r_2) \geq \left(\frac{1}{2}\right)^n.$$ 

**Conjecture.** \(^1\) $$P(\text{rank}(R_n) < n) = \left(\frac{1}{2} + o(1)\right)^n.$$

\(^1\)Further results by Bourgain, Vu and Wood 2010
Symmetric \( \{0, 1\} \)-random matrices

Let \( Q_n = (x_{ij})_{1 \leq i, j \leq n} \) be a symmetric matrix.

- Symmetry introduces \( \binom{n}{2} \) non-trivial correlations.

**Thm** [Costello, Tao, Vu, 2006]

If \( x_{ij}, i \leq j \) are iid uniform in \( \{0, 1\} \), for any \( \beta > 0 \)

\[
P(\text{rank}(Q_n) < n) = O(n^{-1/8+\beta}).
\]
For an $n \times n$ matrix $M$ let

$$z(M) = \max(\#\{i : \text{all entries in row } i \text{ of } M \text{ equal zero}\}, \#\{i : \text{all entries in column } i \text{ of } M \text{ equal zero}\})$$

$$\text{rank}(M) \leq n - z(M).$$

**Question.** For which random matrices do we have

$$\text{rank}(M) = n - z(M)$$

with high probability?
Sparse \( \{0, 1\} \)-random matrices

Let \( Q_{n,p} = (x_{ij})_{1 \leq i,j \leq n} \) be symmetric with \( x_{ij}, i < j \) iid \( \text{Ber}(p) \) r.v.

\[ p < \frac{\ln n}{2n} : \text{2 equal rows with prob } 1-o(1) \]

**Thm** [Costello, Vu, 2006] For \( p = \frac{c \ln n}{n} \leq \frac{1}{2} \) and \( c > \frac{1}{2} \),

\[ \mathbb{P}(\text{rank}(Q_{n,p}) = n - z(Q_{n,p})) = 1 - o(1). \]
Sparse \(\{0, 1\}\)-random matrices

Let \(R_{n,p} = (x_{ij})_{1 \leq i,j \leq n}\) have iid \(x_{ij}\) \(\text{Ber}(p)\) r.v.

**Thm** [Addario-Berry, E., *CPC 2013*] For \(c > \frac{1}{2}\) and \(p \in (\frac{c \ln n}{n}, \frac{1}{2})\),

\[
P(\text{rank}(R_{n,p}) = n - z(R_{n,p})) = 1 - o(1).
\]

**Proof Sketch.**
(Similar technique used by Costello, Vu)

1. Prove that whp \(\text{rank}(R_{\alpha n,p}) \geq (1 - \varepsilon)n\).
2. Add row/column one-at-a-time. Dependency sets will be removed by the end of the exposure.
Lemma. Let $p \in \left(\frac{c \ln n}{n}, \frac{1}{2}\right)$ and $\varepsilon > 0$. There is a constant $\delta$

$$
\mathbb{P}(\text{rank}(R_{n,p}) < (1 - \varepsilon)n) = O(n^{-\delta n}).
$$

Proof. Consider a fixed $n \times n$ matrix $M$ and suppose

$$\text{rank}(M) = d \leq (1 - \varepsilon)n.$$

Suppose also that a reorder of columns gives

$$\text{rank}(A) = d.$$

$B \in \text{span}(A)$ so there is $G$ such that $AG = B$. In fact, $G$ is unique.
First step (non-symmetric case)

Likewise, $D$ is determined by $A, B, C$. Thus, given $A, B, C$ there is a matrix $F$ such that

$$\text{rank}(M) = d \iff D = M.$$ 

Therefore,

$$\mathbb{P}(\text{rank}(R_{n,p}) \leq (1 - \varepsilon)n) \leq \binom{n}{(1-\varepsilon)n} \mathbb{P}(\text{rank}(R_{n,p} = d \mid \square)) \leq 2^n \max(p, 1-p)^{(\varepsilon n)^2}.$$ 

If $p \in \left(\frac{c\ln n}{n}, \frac{1}{2}\right)$, then for some $\delta > 0$

$$\mathbb{P}(\text{rank}(R_{n,p}) \leq (1 - \varepsilon)n) = O(n^{-\delta n}).$$ 

□
Symmetric matrix process

Let \( \{ U_{ij} \}_{1 \leq i < j \leq n} \) be independent r.v. uniform in \((0, 1)\).
Let \( \{ Q_{n,p} \}_{p \in (0, 1)} \) be the symmetric matrix-valued markov process given by

\[
Q_{n,p}(i, j) = Q_{n,p}(j, i) = 1_{[U_{ij} \leq p]}.
\]

Hitting times:

\[
\tau = \inf \{ p : \text{rank}(Q_{n,p}) = n \}
\]

\[
\tau_c = \inf \{ p : z(Q_{n,p}) = 0 \}.
\]

\( \blacktriangleright \) **Fact:** \( \tau \geq \tau_c \)

Proof. \( \text{rank}(Q_{n,p} \leq n - z(Q_{n,p}) \).

**Thm**[Addario-Berry, E., *CPC 2013*]

\[
\mathbb{P}(\tau = \tau_c) \to 1, \text{ as } n \to \infty,
\]
We want to understand

\[ \mathbb{P}(\text{rank}(Q_{n,c}) = n) \]

- **No monotonicity** in the rank process.
- **Structure of sparse matrices** implies dependencies involve a large number of rows:

If \( S \subset [n] \) is small, then some column \( j \) has exactly one 1 whp. Then for any non-zero coefficients \( \{a_i\}_{i \in S} \)

\[
\sum_{i \in S} a_i r_{ij} \neq 0.
\]
Proof idea

Take $z_p = z(Q_{n,p})$.

1. Set $p'$ with $\mathbb{E}(z_{p'}) = 100$.

   It is unlikely that $\tau_c \leq p'$.

2. Condition on $z_{p'}$.

   Eg. suppose $z_{p'} = 100$, and relabel rows/col. for $Q_{n,p'}$.

$S_{n,p}$: the random submatrix of the relabelled $Q_{n,p}$.

- $S_{n,p}$ evolves indep. from zero rows/col.
- $\tau_c$ is conditionally indep. of structure of $S_{n,p}$ given $z_{p'}$. 
Proof idea

- By time $\tau_c$:
  - Expected numbers of ones in rows 1, ..., 100 is $O(1)$.
  - Whp no dependencies in the rows 1, ..., 100.

3. Condition on the values of rows 1, ..., 100 at time $\tau_c$.

We get a partially deterministic matrix $S$, and remains to strengthen Costello-Vu thm to $S$:

$$\mathbb{P}(\text{rank}(S) = n - z(S)) = 1 - o(1).$$

It is not difficult to do. :) □
In discrete random matrices, singularity comes from small dependencies.

- All-zeros rows
- Equal rows

For the matrix process $Q_{n,p}$ with probability $1 - o(1)$:

- Equal rows are not present for $\frac{\ln n}{2n} < p < 1/2$,
- singularity disappears when the last zero row disappears.
THANKS !