Inverse Theorems in Probability

Van H. Vu

Department of Mathematics
Yale University

Concentration and Anti-concentration

X: a random variable.

X: a random variable.

Concentration. If I is a long interval *far* from $\mathbf{E}X$, then $\mathbf{P}(X \in I)$ is small.

Concentration and Anti-concentration

X: a random variable.

Concentration. If I is a long interval *far* from $\mathbf{E}X$, then $\mathbf{P}(X \in I)$ is small.

Anti-concentration. If I is a short interval *anywhere*, then $\mathbf{P}(X \in I)$ is small.

The central limit theorem

$\xi_1, \ldots, \xi_n$ are iid copies of ξ with mean 0 and variance 1, then

$$\frac{\xi_1 + \cdots + \xi_n}{\sqrt{n}} \longrightarrow \mathbf{N}(0,1).$$

The central limit theorem

$\xi_1, \ldots, \xi_n$ are iid copies of ξ with mean 0 and variance 1, then

$$\frac{\xi_1 + \cdots + \xi_n}{\sqrt{n}} \longrightarrow \mathbf{N}(0,1).$$

In other words, for $X := \sum_{i=1}^n \xi_i/\sqrt{n}$, and any fixed $t > 0$

$$\mathbf{P}(X \in [t,\infty)) \to \frac{1}{\sqrt{2\pi}} \int_t^\infty e^{-t^2/2} dt = O(e^{-t^2/2}).$$

The central limit theorem

$\xi_1, \ldots, \xi_n$ are iid copies of ξ with mean 0 and variance 1, then

$$\frac{\xi_1 + \cdots + \xi_n}{\sqrt{n}} \longrightarrow \mathbf{N}(0,1).$$

In other words, for $X := \sum_{i=1}^n \xi_i/\sqrt{n}$, and any fixed $t > 0$

$$\mathbf{P}(X \in [t, \infty)) \to \frac{1}{\sqrt{2\pi}} \int_t^\infty e^{-t^2/2} dt = O(e^{-t^2/2}).$$

Concentration. Results of this type with general X (Chernoff, Bernstein, Azuma, Talagrand etc).

The central limit theorem

$\xi_1, \dots, \xi_n$ are iid copies of ξ with mean 0 and variance 1, then

$$\frac{\xi_1 + \dots + \xi_n}{\sqrt{n}} \longrightarrow \mathbf{N}(0,1).$$

In other words, for $X := \sum_{i=1}^n \xi_i / \sqrt{n}$, and any fixed $t > 0$

$$\mathbf{P}(X \in [t, \infty)) \to \frac{1}{\sqrt{2\pi}} \int_t^\infty e^{-t^2/2} dt = O(e^{-t^2/2}).$$

Concentration. Results of this type with general X (Chernoff, Bernstein, Azuma, Talagrand etc).

The central limit theorem

Berry-Esséen (1941): ξ has bounded third moment, then the rate of convergence is $O(n^{-1/2})$. For any t,

$$\mathbf{P}(X \in [t, \infty)) = \frac{1}{\sqrt{2\pi}} \int_t^\infty e^{-t^2/2} dt + O(n^{-1/2}).$$

Berry-Esséen (1941): ξ has bounded third moment, then the rate of convergence is $O(n^{-1/2})$. For any t,

$$\mathbf{P}(X \in [t,\infty)) = \frac{1}{\sqrt{2\pi}} \int_t^{\infty} e^{-t^2/2} dt + O(n^{-1/2}).$$

This implies that for any interval I

$$\mathbf{P}(X \in I) = \frac{1}{\sqrt{2\pi}} \int_I e^{-t^2/2} dt + O(n^{-1/2}).$$

The central limit theorem

Berry-Esséen (1941): ξ has bounded third moment, then the rate of convergence is $O(n^{-1/2})$. For any t,

$$\mathbf{P}(X \in [t, \infty)) = \frac{1}{\sqrt{2\pi}} \int_t^\infty e^{-t^2/2} dt + O(n^{-1/2}).$$

This implies that for any interval I

$$\mathbf{P}(X \in I) = \frac{1}{\sqrt{2\pi}} \int_I e^{-t^2/2} dt + O(n^{-1/2}).$$

The error term $n^{-1/2}$ is sharp: Take $\xi = \pm 1$ (Bernoulli) and n even, then $\mathbf{P}(X = 0) = \frac{\binom{n}{n/2}}{2^n} = \Theta(n^{-1/2})$.

The central limit theorem

Berry-Esséen (1941): ξ has bounded third moment, then the rate of convergence is $O(n^{-1/2})$. For any t,

$$\mathbf{P}(X \in [t, \infty)) = \frac{1}{\sqrt{2\pi}} \int_t^\infty e^{-t^2/2} dt + O(n^{-1/2}).$$

This implies that for any interval I

$$\mathbf{P}(X \in I) = \frac{1}{\sqrt{2\pi}} \int_I e^{-t^2/2} dt + O(n^{-1/2}).$$

The error term $n^{-1/2}$ is sharp: Take $\xi = \pm 1$ (Bernoulli) and n even, then $\mathbf{P}(X = 0) = \frac{\binom{n}{n/2}}{2^n} = \Theta(n^{-1/2})$.

Anti-concentration. Results of this type with more general X.

The central limit theorem

Berry-Esséen (1941): ξ has bounded third moment, then the rate of convergence is $O(n^{-1/2})$. For any t,

$$\mathbf{P}(X \in [t, \infty)) = \frac{1}{\sqrt{2\pi}} \int_t^\infty e^{-t^2/2} dt + O(n^{-1/2}).$$

This implies that for any interval I

$$\mathbf{P}(X \in I) = \frac{1}{\sqrt{2\pi}} \int_I e^{-t^2/2} dt + O(n^{-1/2}).$$

The error term $n^{-1/2}$ is sharp: Take $\xi = \pm 1$ (Bernoulli) and n even, then $\mathbf{P}(X = 0) = \frac{\binom{n}{n/2}}{2^n} = \Theta(n^{-1/2})$.

Anti-concentration. Results of this type with more general X.

Littlewood-Offord-Erdös

$A = \{a_1, \dots, a_n\}$ (multi-) set of deterministic coefficients

$$S_A := a_1\xi_1 + \cdots + a_n\xi_n.$$

Theorem (Littlewood-Offord 1940)

If ξ is Bernoulli (taking values ± 1 with probability $1/2$) and a_i have absolute value at least 1, then for any open interval I of length 1,

$$\mathbf{P}(S_A \in I) = O(\frac{\log n}{n^{1/2}}).$$

Littlewood-Offord-Erdös

$A = \{a_1, \dots, a_n\}$ (multi-) set of deterministic coefficients

$$S_A := a_1\xi_1 + \cdots + a_n\xi_n.$$

Theorem (Littlewood-Offord 1940)

If ξ is Bernoulli (taking values ± 1 with probability $1/2$) and a_i have absolute value at least 1, then for any open interval I of length 1,

$$\mathbf{P}(S_A \in I) = O\left(\frac{\log n}{n^{1/2}}\right).$$

S_A may not satisfy the Central Limit Theorem.

Theorem (Erdös 1943)

$$\mathbf{P}(S_A \in I) \leq \frac{\binom{n}{\lfloor n/2 \rfloor}}{2^n} = O\left(\frac{1}{n^{1/2}}\right). \tag{1}$$

Kolmogorov-Rogozin

Levy's concentration function: $Q(\lambda, X) = \sup_{|I|=\lambda} \mathbf{P}(X \in I)$.

Theorem (Kolmogorov-Rogozin 1959-1961)

$S = X_1 + \cdots + X_n$ where X_i are independent. Then

$$Q(\lambda, S) = O\left(\frac{1}{\sqrt{\sum_{i=1}^n (1 - Q(\lambda, X_i))}}\right).$$

Kesten, Esseen, Halász (60s-70s).

Littlewood-Offord-Erdos: refinement

Recall $A := \{a_1, \ldots, a_n\}$

$$S_A := \xi_1\xi_1 + \cdots + a_n\xi_n.$$

Littlewood-Offord-Erdos: refinement

Recall $A := \{a_1, \ldots, a_n\}$

$$S_A := \xi_1\xi_1 + \cdots + a_n\xi_n.$$

One can improve anti-concentration bounds significantly under extra assumptions on the additive structure of A.

Littlewood-Offord-Erdos: refinement

Recall $A := \{a_1, \dots, a_n\}$

$$S_A := \xi_1\xi_1 + \cdots + a_n\xi_n.$$

One can improve anti-concentration bounds significantly under extra assumptions on the additive structure of A.

Discrete setting; ξ_i are iid ± 1; a_i are integers:

$$\rho(A) := \sup_x \mathbf{P}(S_A = x)$$

Littlewood-Offord-Erdos: refinement

Recall $A := \{a_1, \ldots, a_n\}$

$$S_A := \xi_1 \xi_1 + \cdots + a_n \xi_n.$$

One can improve anti-concentration bounds significantly under extra assumptions on the additive structure of A.

Discrete setting; ξ_i are iid ± 1; a_i are integers:

$$\rho(A) := \sup_x \mathbf{P}(S_A = x)$$

(instead of $\sup_{|I|=l} \mathbf{P}(X \in I)$).

Theorem (Erdös-Moser 1947)

Let a_i be distinct integers, then

$$\rho(A) = O(n^{-3/2} \log n).$$

Theorem (Sárkozy-Szemerédi 1965)

$$\rho(A) = O(n^{-3/2}).$$

Stanley's characterization

Theorem (Stanley 1980; Proctor 1982)

Let n be odd and $A_0 := \{ -\frac{n-1}{2}, \ldots, \frac{n-1}{2} \}$. Let A be any set of n distinct real numbers, then

$$\rho(A) \leq \rho(A_0).$$

The proofs are algebraic (hard Lepschetz theorem, Lie algebra).

Extensions

Stronger conditions, more dimensions etc: Beck, Katona, Kleitman, Griggs, Frank-Furedi, Halasz, Sali etc (1970s-1980s).

Extensions

Stronger conditions, more dimensions etc: Beck, Katona, Kleitman, Griggs, Frank-Furedi, Halasz, Sali etc (1970s-1980s).

Theorem (Halasz 1979)

Let k be a fixed integer and R_k be the number of solutions of the equation $a_{i_1} + \cdots + a_{i_k} = a_{j_1} + \cdots + a_{j_k}$. Then

$$\rho_A = O(n^{-2k-\frac{1}{2}} R_k).$$

What cause *large* anti-concentration probability ?

What cause *large* anti-concentration probability ?

Inverse Principle [Tao-V. 2005]

A set A with large ρ_A must have a strong additive structure.

A new look at anti-concentration

What cause *large* anti-concentration probability ?

Inverse Principle [Tao-V. 2005]

A set A with large ρ_A must have a strong additive structure.

Arak (1980s)

We will give many illustrations of this principle with applications.

Motivation: Additive Combinatorics

Freiman Inverse theorem: If $A + A = \{a + a' | a, a' \in A\}$ is small, then A has a strong additive structure.

Example. A is a dense subset (of density δ, say) of an interval J of length n/δ,

$$|A + A| \leq |J + J| \leq 2n/\delta \leq \frac{2}{\delta}|A|.$$

Motivation: Additive Combinatorics

Freiman Inverse theorem: If $A + A = \{a + a' | a, a' \in A\}$ is small, then A has a strong additive structure.

Example. A is a dense subset (of density δ, say) of an interval J of length n/δ,

$$|A + A| \leq |J + J| \leq 2n/\delta \leq \frac{2}{\delta}|A|.$$

Example. If A is a dense subset (of density δ, say) of a GAP of rank d then

$$|A + A| \leq |J + J| \leq 2^d n/\delta \leq \frac{2^d}{\delta}|A|.$$

Freiman Inverse Theorem

Theorem (Freiman Inverse Theorem 1975)

For any constant C there are constants d and $\delta > 0$ such that if $|A + A| \leq C|A|$, then A is a subset of density at least δ of a (generalized) arithmetic progression of rank at most d.

Freiman Inverse Theorem

Theorem (Freiman Inverse Theorem 1975)

For any constant C there are constants d and $\delta > 0$ such that if $|A + A| \le C|A|$, then A is a subset of density at least δ of a (generalized) arithmetic progression of rank at most d.

Collisions of pairs $a + a'$ vs collisions of subset sums $\sum_{a \in B; B \subset A} a$.

Inverse Littlewood-Offord theorems

Example. If A is a subset of a generalized arithmetic progression Q of rank d of cardinality n^C, then all numbers of the form $\pm a_1 \pm a_2 + \cdots \pm a_n$ belong to nQ, which has cardinality at most $n^d|Q| = n^{d+C}$; by pigeon hole principle

$$\rho_A := \ sup_x \mathbf{P}(S_A = x) \geq n^{-d-C}.$$

Inverse Littlewood-Offord theorems

Example. If A is a subset of a generalized arithmetic progression Q of rank d of cardinality n^C, then all numbers of the form $\pm a_1 \pm a_2 + \cdots \pm a_n$ belong to nQ, which has cardinality at most $n^d|Q| = n^{d+C}$; by pigeon hole principle

$$\rho_A := \ sup_x \mathbf{P}(S_A = x) \geq n^{-d-C}.$$

Theorem (First Inverse Littlewood-Offord theorem; Tao-V. 2006)

If $\rho_A \geq n^{-B}$ then there are constants $d, C > 0$ such that most of A belongs to a (generalize) arithmetic progression of cardinality n^C of rank at most d.

Inverse Littlewood-Offord theory

Extensions: Tao-V, Rudelson-Vershynin, Friedland-Sodin, Hoi Nguyen, Nguyen-V., Elliseeva-Zaitsev et al. etc

- Sharp relations between B, C, d.
- General ξ_i (not Bernoulli).
- Multi-dimensional versions $\mathbf{R}^d$; Abelian versions.
- Small probability version $\mathbf{P}(S_A \in I)$ (I interval in $\mathbf{R}$ or small ball in $\mathbf{R}^k$).
- Relaxing n^{-B} to $(1-c)^n$.
- Sum of not necessary independent random variables; etc.

Little bit about the proof

Toy case. a_i are elements of F_p for some large prime p, viewed as integers between 0 and $p-1$, and

$$\rho = \rho(A) = \mathbf{P}(S = 0).$$

Notation. $e_p(x)$ for $\exp(2\pi\sqrt{-1}x/p)$.

$$\rho = \mathbf{P}(S = 0) = \mathbf{E}\mathbf{I}_{S=0} = \mathbf{E}\frac{1}{p}\sum_{t \in F_p} e_p(tS).$$

By independence

$$\mathbf{E}e_p(tS) = \prod_{i=1}^{n} \mathbf{E}e_p(t\xi_i a_i) = \prod_{i=1}^{n} \cos\frac{\pi t a_i}{p}.$$

Thus

$$\rho \leq \frac{1}{p} \sum_{t \in \mathbb{F}_p} \prod_i |\frac{\cos \pi a_i t}{p}|.$$

Facts. $|\sin \pi z| \geq 2\|z\|$ where $\|z\|$ is the distance of z to the nearest integer.

$$|\cos \frac{\pi x}{p}| \leq 1 - \frac{1}{2} \sin^2 \frac{\pi x}{p} \leq 1 - 2\|\frac{x}{p}\|^2 \leq \exp(-2\|\frac{x}{p}\|^2).$$

Key inequality

$$\rho \leq \frac{1}{p} \sum_{t \in \mathbb{F}_p} \prod_i |\cos \frac{\pi a_i t}{p}| \leq \frac{1}{p} \sum_{t \in F_p} \exp(-2 \sum_{i=1}^{n} \|\frac{a_i t}{p}\|^2).$$

Thus

$$\rho \leq \frac{1}{p} \sum_{t \in \mathbb{F}_p} \prod_i |\frac{\cos \pi a_i t}{p}|.$$

Facts. $|\sin \pi z| \geq 2\|z\|$ where $\|z\|$ is the distance of z to the nearest integer.

$$|\cos \frac{\pi x}{p}| \leq 1 - \frac{1}{2} \sin^2 \frac{\pi x}{p} \leq 1 - 2\|\frac{x}{p}\|^2 \leq \exp(-2\|\frac{x}{p}\|^2).$$

Key inequality

$$\rho \leq \frac{1}{p} \sum_{t \in \mathbb{F}_p} \prod_i |\cos \frac{\pi a_i t}{p}| \leq \frac{1}{p} \sum_{t \in F_p} \exp(-2 \sum_{i=1}^n \|\frac{a_i t}{p}\|^2).$$

If a_i, t were vectors in a vector space, the key inequality suggests that $a_i \cdot t$ is close to zero very often. Thus, most a_i are close to a small dimensional subspace.

Second step: Large level sets

Consider the level sets $S_m := \{t | \sum_{i=1}^n \|a_i t/p\|^2 \leq m\}$.

$$n^{-C} \leq \rho \leq \frac{1}{p} \sum_{t \in \mathbb{F}_p} \exp(-2 \sum_{i=1}^n \|\frac{a_i t}{p}\|^2) \leq \frac{1}{p} + \frac{1}{p} \sum_{m \geq 1} \exp(-2(m-1)) |S_m|.$$

Since $\sum_{m \geq 1} \exp(-m) < 1$, there must be is a large level set S_m such that

$$|S_m| \exp(-m+2) \geq \rho p. \tag{2}$$

In fact, since $\rho \geq n^{-C}$, we can assume that $m = O(\log n)$.

Double counting the the triangle inequality

By double counting we have

$$\sum_{i=1}^{n}\sum_{t\in S_m}\|\frac{a_i t}{p}\|^2=\sum_{t\in S_m}\sum_{i=1}^{n}\|\frac{a_i t}{p}\|^2\leq m|S_m|.$$

So, for most a_i

$$\sum_{t\in S_m}\|\frac{a_i t}{p}\|^2\leq\frac{m}{n'}|S_m|. \tag{3}$$

By averaging, the set of a_i satisfying (3) has size at least $n-n'$.
We are going to show that A' is a large subset of a GAP.
Since $\|\cdot\|$ is a norm, by the triangle inequality, we have for any $a\in kA'$

$$\sum_{t\in S_m}\|\frac{at}{p}\|^2\leq k^2\frac{m}{n'}|S_m|. \tag{4}$$

More generally, for any $l\leq k$ and $a\in lA'$

$$\sum \|\frac{at}{p}\|^2 \leq l^2\frac{m}{n'}|S_m|. \tag{5}$$

Duality

Define $S_m^* := \{a | \sum_{t \in S_m} \| \frac{at}{p} \|^2 \leq \frac{1}{200} |S_m| \}$; S_m^* can be viewed as some sort of a *dual* set of S_m. In fact,

$$|S_m^*| \leq \frac{8p}{|S_m|}. \tag{6}$$

To see this, define $T_a := \sum_{t \in S_m} \cos \frac{2\pi a t}{p}$. Using the fact that $\cos 2\pi z \geq 1 - 100 \|z\|^2$ for any $z \in \mathbf{R}$, we have, for any $a \in S_m^*$

$$T_a \geq \sum_{t \in S_m} (1 - 100 \| \frac{at}{p} \|^2) \geq \frac{1}{2} |S_m|.$$

One the other hand, using the basic identity $\sum_{a \in \mathbb{F}_p} \cos \frac{2\pi a x}{p} = p \mathbf{I}_{x=0}$, we have

$$\sum_{a \in \mathbb{F}_p} T_a^2 \leq 2p |S_m|.$$

(6) follows from the last two estimates and averaging.

Set $k := c_1 \sqrt{\frac{n'}{m}}$, for a properly chosen constant c_1. By (5) we

The role of $\mathbb{F}_p$ is now no longer important, so we can view the a_i as integers. Notice that (7) leads us to a situation similar to that of Freiman's inverse result. In that theorem, we have a bound on $|2A|$ and conclude that A has a strong additive structure. In the current situation, 2 is replaced by k, which can depend on $|A|$.

Theorem (Long range inverse theorem)

Let $\gamma > 0$ be constant. Assume that X is a subset of a torsion-free group such that $0 \in X$ and $|kX| \leq k^\gamma |X|$ for some integer $k \geq 2$ that may depend on $|X|$. Then there is proper symmetric GAP Q of rank $r = O(\gamma)$ and cardinality $O_\gamma(k^{-r}|kX|)$ such that $X \subset Q$.

Applications: Quick proofs of forward theorems

Example. Sárközy-Szemerédi 1965. If a_i are different integers, then

$$\rho_A = O(n^{-3/2}).$$

Applications: Quick proofs of forward theorems

Example. Sárközy-Szemerédi 1965. If a_i are different integers, then

$$\rho_A = O(n^{-3/2}).$$

Assume $\rho_A \geq Cn^{-3/2}$, say, then the optimal inverse theorem implies that most of a_i belong to a GAP of cardinality at most cn, with $c \to 0$ as $C \to \infty$. So for large C we obtain a contradiction.

Applications: Quick proofs of forward theorems

Example. Sárközy-Szemerédi 1965. If a_i are different integers, then

$$\rho_A = O(n^{-3/2}).$$

Assume $\rho_A \geq Cn^{-3/2}$, say, then the optimal inverse theorem implies that most of a_i belong to a GAP of cardinality at most cn, with $c \to 0$ as $C \to \infty$. So for large C we obtain a contradiction.

Example. A stable version of Stanley's result.

Theorem (H. Nguyen 2010)

If $\rho_A \geq (C_0 - \epsilon)n^{-3/2}$ for an optimal constant C_0, then A is δ-close to $\{-\lfloor n/2 \rfloor, \dots, \lfloor n/2 \rfloor\}$.

Applications: Quick proofs of forward theorems

Example. Sárközy-Szemerédi 1965. If a_i are different integers, then

$$\rho_A = O(n^{-3/2}).$$

Assume $\rho_A \geq Cn^{-3/2}$, say, then the optimal inverse theorem implies that most of a_i belong to a GAP of cardinality at most cn, with $c \to 0$ as $C \to \infty$. So for large C we obtain a contradiction.

Example. A stable version of Stanley's result.

Theorem (H. Nguyen 2010)

If $\rho_A \geq (C_0 - \epsilon)n^{-3/2}$ for an optimal constant C_0, then A is δ-close to $\{-\lfloor n/2 \rfloor, \dots, \lfloor n/2 \rfloor\}$.

Example. Frankl-Füredi 1988 conjecture on Erdös' type (sharp) bound in high dimensions (Kleitman $d = 2$, Tao-V. 2010, $d \geq 3$).

Applications: Random matrices and the singular probability problem

Let M_n be a random matrix whose entries a random Bernoulli variables (± 1).

Let M_n be a random matrix whose entries a random Bernoulli variables (± 1).

Problem. Estimate $p_n := \mathbf{P}(M_n \textit{ singular}) = \mathbf{P}(\det M_n = 0)$.

Let M_n be a random matrix whose entries a random Bernoulli variables (± 1).

Problem. Estimate $p_n := \mathbf{P}(M_n \textit{ singular}) = \mathbf{P}(\det M_n = 0)$.

- Komlos 1967: $p_n = o(1)$
- Komlos 1975: $p_n \leq n^{-1/2}$.
- Kahn-Komlos-Szemeredi 1995: $p_n \leq .999^n$
- Tao-V. 2004: $p_n \leq .952^n$.
- Tao-V. 2005 $p_n \leq (3/4 + o(1))^n$.
- Bourgain-V.-Wood (2009) $p_n \leq (\frac{1}{\sqrt{2}} + o(1))^n$.

Applications: Bounding the singular probability

Insight. Let X_i be the row vectors and $v = (a_1, \ldots, a_n)$ be the normal vector of $\mathrm{Span}(X_1, \ldots, X_{n-1})$

$$\mathbf{P}(X_n \in \mathrm{Span}(X_1, \ldots, X_{n-1}) = \mathbf{P}(X_n \cdot v = 0) = \mathbf{P}(a_1\xi_1 + \cdots + a_n\xi_n = 0).$$

Applications: Bounding the singular probability

Insight. Let X_i be the row vectors and $v = (a_1, \dots, a_n)$ be the normal vector of $\mathrm{Span}(X_1, \dots, X_{n-1})$

$$\mathbf{P}(X_n \in \mathrm{Span}(X_1, \dots, X_{n-1}) = \mathbf{P}(X_n \cdot v = 0) = \mathbf{P}(a_1\xi_1 + \dots + a_n\xi_n = 0).$$

By Inverse Theorems this probability is either very small, or $A = \{a_1, \dots, a_n\}$ has a strong structure, which is also unlikely as it forms a normal vector of a random hyperplane.

Replacing $\mathbf{P}(X_n \cdot v = 0)$ by

$$\mathbf{P}(|X_n \cdot v| \leq \epsilon) = \mathbf{P}(a_1\xi_1 + \cdots + a_n\xi_n \in [-\epsilon, \epsilon]),$$

one can show that with high probability $|X_n \cdot v|$ is not very small. This, in turn, bounds the least singular value from below.

- Tao-V 2006: For any C, there is B such that $\mathbf{P}(\sigma_{min} M_n \leq n^{-B}) \leq n^{-C}$.
- Rudelson-Vershynin 2007: $\mathbf{P}(\sigma_{min} M_n \leq \epsilon n^{-1/2}) \leq C(\epsilon + .9999^n)$ for any $\epsilon > 0$.

Applications: Random matrices and the Circular Law

Conjecture (Circular Law 1960s)

Let $M_n(\xi)$ be a random matrix whose entries are iid collies of a random variable ξ with mean 0 and variance 1. Then the distribution of the eigenvalues of $\frac{1}{\sqrt{n}}M_n$ tends to the uniform distribution on the unit circle.

Mehta (1960s), Edelman (1980s), Girko (1980s), Bai (1990s), Gotze-Tykhomirov, Pan-Zhu (2000s); Tao-V (2007); (Tao-V: Bullentin AMS; Chafai et al.: Surveys in Probability).

Laws for matrices with dependent entries.

- Chafai et. al (2008): Markov matrices.
- Hoi Nguyen (2011): proving Chatterjee-Diaconnis conjecture concerning random double stochastic matrices.
- Gotze-Tykhomirov; Sosnyikov et. al. (2011): law for product of random matrices.
- Adamczak et. al. (2010): law for matrices with independent rows
- Naumov, Nguyen-O'rourke (2013): Elliptic Law.

Applications: Random walks

In 1921, Polya proved his famous drunkard's walk theorem on $\mathbf{Z}^d$.

$$S_n := \sum_{j=1}^{n} \xi_j f_j$$

where f_j is chosen uniformly from $E := \{e_1, \dots, e_d\}$.

Theorem (Drunkard walk's theorem; Polya 1921)

For any $d \geq 1$, $\mathbf{P}(S_n = 0) = \Theta(n^{-d/2})$. In particular, the walk is recurrent only if $d = 1, 2$.

What happens if $f_1, \dots, f_n$ are n different unit vectors ?

Theorem (Suburban drunkard walk's theorem; Herdade-V. 2014)

Consider a set V of n different unit vectors which is effectively d-dimensional. Then

- *For $d \geq 4$, $\mathbf{P}(S_{n,V} = 0) \leq n^{-\frac{d}{2} - \frac{d}{d-2} + o(1)}$.*
- *For $d = 3$, $\mathbf{P}(S_{n,V} = 0) \leq n^{-4+o(1)}$.*

Theorem (Suburban drunkard walk's theorem; Herdade-V. 2014)

Consider a set V of n different unit vectors which is effectively d-dimensional. Then

- *For* $d \geq 4$, $\mathbf{P}(S_{n,V} = 0) \leq n^{-\frac{d}{2} - \frac{d}{d-2} + o(1)}$.
- *For* $d = 3$, $\mathbf{P}(S_{n,V} = 0) \leq n^{-4+o(1)}$.
- *For* $d = 2$, $\mathbf{P}(S_{n,V} = 0) \leq n^{-\omega(1)}$.

Case $d = 2$**.** If $\mathbf{P}(S_{n,V} = 0) \geq n^{-C}$, then V belongs to a small GAP by Inverse theorems. But it also belongs to the unit circle.

Theorem (Suburban drunkard walk's theorem; Herdade-V. 2014)

Consider a set V of n different unit vectors which is effectively d-dimensional. Then

- *For $d \geq 4$,* $\mathbf{P}(S_{n,V} = 0) \leq n^{-\frac{d}{2} - \frac{d}{d-2} + o(1)}$.
- *For $d = 3$,* $\mathbf{P}(S_{n,V} = 0) \leq n^{-4+o(1)}$.
- *For $d = 2$,* $\mathbf{P}(S_{n,V} = 0) \leq n^{-\omega(1)}$.

Case $d = 2$**.** If $\mathbf{P}(S_{n,V} = 0) \geq n^{-C}$, then V belongs to a small GAP by Inverse theorems. But it also belongs to the unit circle. These two cannot occur at the same time due to number theoretic reasons.

Theorem (Suburban drunkard walk's theorem; Herdade-V. 2014)

Consider a set V of n different unit vectors which is effectively d-dimensional. Then

- *For* $d \geq 4$, $\mathbf{P}(S_{n,V} = 0) \leq n^{-\frac{d}{2} - \frac{d}{d-2} + o(1)}$.
- *For* $d = 3$, $\mathbf{P}(S_{n,V} = 0) \leq n^{-4+o(1)}$.
- *For* $d = 2$, $\mathbf{P}(S_{n,V} = 0) \leq n^{-\omega(1)}$.

Case $d = 2$**.** If $\mathbf{P}(S_{n,V} = 0) \geq n^{-C}$, then V belongs to a small GAP by Inverse theorems. But it also belongs to the unit circle. These two cannot occur at the same time due to number theoretic reasons.

Toy example. For any R, the square grid has only $R^{o(1)}$ points on $C(0, R)$ (Sum of two squares problem).

$$P_n(x) = \xi_n x^n + \cdots + \xi_1 x + \xi_0.$$

ξ_i are iid copies of ξ having mean 0 and variance 1.

Applications: Random polynomials

$$P_n(x) = \xi_n x^n + \cdots + \xi_1 x + \xi_0.$$

ξ_i are iid copies of ξ having mean 0 and variance 1.

How many real roots does P_n have ?

$$P_n(x) = \xi_n x^n + \cdots + \xi_1 x + \xi_0.$$

ξ_i are iid copies of ξ having mean 0 and variance 1.

How many real roots does P_n have ?

This leads the development of the theory of random functions.

Number of real roots of a random polynomials

- Waring (1782): $n = 3$, Sylvester.
- Bloch-Polya (1930s): ξ Bernoulli, $\mathbf{E}N_n = O(\sqrt{n})$.
- Littlewood-Offord (1939-1943) General ξ,

$$\frac{\log n}{\log \log n} \leq \mathbf{E}N_n \leq \log^2 n.$$

- Kac (1943) ξ Gaussian

$$\mathbf{E}N_n = \frac{1}{\pi}\int_{-\infty}^{\infty}\sqrt{\frac{1}{(t^2-1)^2} + \frac{(n+1)^2 t^{2n}}{(t^{2n+2}-1)^2}}dt = (\frac{2}{\pi}+o(1))\log n.$$

- Kac (1949) ξ uniform on $[-1,1]$, $\mathbf{E}N_n = (\frac{2}{\pi} + o(1))\log n$.
- Stevens (1967) $\mathbf{E}N_n = (\frac{2}{\pi} + o(1))\log n$, ξ smooth.
- Erdös-Offord (1956) $\mathbf{E}N_n = (\frac{2}{\pi} + o(1))\log n$, ξ Bernoulli.
- Ibragimov-Maslova (1969) $\mathbf{E}N_n = (\frac{2}{\pi} + o(1))\log n$, general ξ.

Number of real roots: The error term

Willis (1980s), Edelman-Kostlan (1995): If ξ is Gaussian

$$\mathbf{E}N_n - \frac{2}{\pi}\log n \to C_{Gauss} \approx .625738.$$

Number of real roots: The error term

Willis (1980s), Edelman-Kostlan (1995): If ξ is Gaussian

$$\mathbf{E} N_n - \frac{2}{\pi} \log n \to C_{Gauss} \approx .625738.$$

This was done by carefully evaluating Kac's integral formula.

Number of real roots: The error term

Willis (1980s), Edelman-Kostlan (1995): If ξ is Gaussian

$$\mathbf{E}N_n - \frac{2}{\pi}\log n \to C_{Gauss} \approx .625738.$$

This was done by carefully evaluating Kac's integral formula.

Tao-V. 2013, Hoi Nguyen-Oanh Nguyen-V. (2014)

Theorem (Yen Do-Hoi Nguyen-V 2015)

There is a constant C_ξ depending on ξ such that

$$\mathbf{E}N_n - \frac{2}{\pi}\log n \to C_\xi.$$

Number of real roots: The error term

Willis (1980s), Edelman-Kostlan (1995): If ξ is Gaussian

$$\mathbf{E} N_n - \frac{2}{\pi} \log n \to C_{Gauss} \approx .625738.$$

This was done by carefully evaluating Kac's integral formula.

Tao-V. 2013, Hoi Nguyen-Oanh Nguyen-V. (2014)

Theorem (Yen Do-Hoi Nguyen-V 2015)

There is a constant C_ξ depending on ξ such that

$$\mathbf{E} N_n - \frac{2}{\pi} \log n \to C_\xi .$$

The value of C_ξ depends on ξ and is not known in general, even for $\xi = \pm 1$.

Random polynomials: Double roots and Common roots

$$S := x^n \xi_n + \cdots + x\xi_1 + \xi_0.$$

Random polynomials: Double roots and Common roots

$$S := x^n \xi_n + \cdots + x\xi_1 + \xi_0.$$

Theorem (Yen Do-Hoi Nguyen-V 2014+)

For general ξ, the probability that P_n has a double root is essentially the probability that it has a double root at 1 or -1. (This probability is $O(n^{-2})$).

Theorem (Kozma- Zeitouni 2012)

A system of $d+1$ random Bernoulli polynomials in d variables does not have common roots ***whp****.*

Applications: Complexity theory

Scott Aaronson-H. Nguyen (2014)
Let $C_n := \{-1, 1\}^n$. For a matrix M, define the *score* of M

$$s_0(M) := \mathbf{P}_{x \in C_n}(Mx \in C_n).$$

If M is a product of permutation and reflection matrices, then $s_0 = 1$.

Does one have an inverse statement in some sense ?

Theorem (H. Nguyen-Aaronson 2014+)

If M is orthogonal and has score at least n^{-C}, then most rows contain an entry of absolute value at least $1 - n^{-1+\epsilon}$.

Extensions: Higher degree Littlewood-Offord

Instead of $S_A = \sum_i^n a_i \xi_i$ consider a quadratic form

$$Q_A = \sum_{1 \le i,j \le n} a_{ij} \xi_i \xi_j .$$

Theorem (Costello-Tao-V. 2005)

Let $A = \{a_{ij}\}$ be a set of non-zero real numbers, then

$$\mathbf{P}(Q_A = 0) \le n^{-1/4}.$$

Extensions: Higher degree Littlewood-Offord

Instead of $S_A = \sum_i^n a_i \xi_i$ consider a quadratic form

$$Q_A = \sum_{1 \le i,j \le n} a_{ij} \xi_i \xi_j .$$

Theorem (Costello-Tao-V. 2005)

Let $A = \{a_{ij}\}$ be a set of non-zero real numbers, then

$$\mathbf{P}(Q_A = 0) \le n^{-1/4}.$$

Costello (2009) improve to bound to $n^{-1/2+o(1)}$ which is best possible.

Theorem (Quadratic Littewood-Offord)

Let $A = \{a_{ij}\}$ be a set of non-zero real numbers, then

$$\sup_x \mathbf{P}(Q_A = x) \le n^{-1/2+o(1)}.$$

Extensions: Higher degree Littlewood-Offord

Theorem (Costello-Tao-V. 2005)

Let P be a polynomial of degree d with non-zero coefficients in $\xi_1, dots, \xi_n$, then

$$\mathbf{P}(P = 0) \leq n^{-c_d}.$$

Extensions: Higher degree Littlewood-Offord

Theorem (Costello-Tao-V. 2005)

Let P be a polynomial of degree d with non-zero coefficients in $\xi_1, dots, \xi_n$, then

$$\mathbf{P}(P = 0) \leq n^{-c_d}.$$

$c_d = 2^{-d^2}$;

Extensions: Higher degree Littlewood-Offord

Theorem (Costello-Tao-V. 2005)

Let P be a polynomial of degree d with non-zero coefficients in $\xi_1, dots, \xi_n$, then

$$\mathbf{P}(P = 0) \leq n^{-c_d}.$$

$c_d = 2^{-d^2}$;
Razborov-Viola 2013 (complexity theory): $c_d = 2^{-d}$.

Extensions: Higher degree Littlewood-Offord

Theorem (Costello-Tao-V. 2005)

Let P be a polynomial of degree d with non-zero coefficients in $\xi_1, dots, \xi_n$, then

$$\mathbf{P}(P = 0) \leq n^{-c_d}.$$

$c_d = 2^{-d^2}$;
Razborov-Viola 2013 (complexity theory): $c_d = 2^{-d}$.
Meka-Oanh Nguyen-V. (2015): $c_d = 1/2 + o(1)$.

Applications of higher degree Littlewood-Offord inequalities

Bounding the singular probability of random symmetric matrix.

- Costello-Tao-V 2005: $p_n^{sym} = o(1)$ (establishing a conjecture of B. Weiss 1980s).
- Costello 2009: $p_n^{sym} \leq n^{-1/2+o(1)}$.
- H. Nguyen 2011: $p_n^{sum} \leq n^{-\omega(1)}$.
- Vershynin 2011: $p_n^{sym} \leq \exp(-n^{\epsilon})$.

Bounding the least singular value: H. Nguyen, Vershynin (2011).

Directions of research

- Sharp bound for high degree polynomials (Meka et al. 2015)
- Inverse theorems for high degree polynomials (Hoi Nguyen 2012, H. Nguyen-O'rourke 2013).
- Dependent models (Pham et al., Nguyen, Tao 2015).
- Further applications.