Non-Asymptotic Theory of Random Matrices
Lecture 20: The recurrence set (Ergodic approach)

Lecturer: Roman Vershynin Scribe: Igor Rumanov

Tuesday, March 13, 2007

Consider again the sum (see previous lectures)
$S = \sum_1^n = a_k \xi_k$ $\qquad k1 \le a_k \le k2$

Then Small Ball Probability:
$p_\varepsilon(a) = \sup_v \mathbb{P}(|S - v| \le \varepsilon)$

GOAL : <u>S. B. P. Theorem:</u>

$$p_\varepsilon(a) \le \frac{1}{\sqrt{k}} \left(\varepsilon + \frac{1}{D(a)} \right)$$

where $D(a) = D_{\alpha,k}(a)$ - the essential Least Common Denominator (LCD) of a
(if coefficients of $a = (\frac{p_1}{q_1}, \frac{p_2}{q_2}, \dots, \frac{p_n}{q_n})$

For $\inf t > 0$: all except k coefficients of a are of distance α from nonzero integers

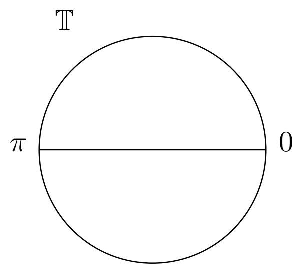
<u>Example:</u> $(1, 1 + \frac{1}{n}, \dots, 1 + \frac{n}{n})$ and $(1, 1, \dots, 1)$
LCD $= n$ - fewer cancellations $\qquad$ LCD $= 1$

Argument of Halasz : S.B.P. $\le ||char.\ function||_{L^1}$
(by Esseen) :

$$p_\varepsilon(a) \le \int_{-\pi/2}^{\pi/2} |\phi(t/\varepsilon)| dt = \int_{-\pi/2}^{\pi/2} | \prod_1^n \cos(\frac{a_k t}{\varepsilon})| dt$$

(we want the integral to be small)

Consider



moving n particles
on 1-torus (circle),
speeds $a_1, \ldots, a_n$

If speeds were random they would form really good incommesurate set. But they are not random.

1 Ergodic approach:

$$p_\varepsilon(a) = \int_{-\pi/2}^{\pi/2} e^{-\frac{1}{2}f(t/\varepsilon)} dt,$$

where $f(t) = \sum_1^n \sin^2(a_k t)$.
We want $f(t)$ to be big. We will apply

Sum-sets argument :
Last time we got for the small ball probability (S. B. P.):

$$p_\varepsilon(a) \lesssim \frac{1}{\sqrt{\eta n}} |T(\eta n, \pi)| + e^{-c\eta n}, \qquad \forall \eta > 0$$

Level set : $T(\eta n, \pi) = \{t \in [-\pi; \pi] : f(t/\varepsilon) \le \eta n\}$
$\uparrow$
Bad set We want it to be small

Let $t \in T(\eta n, \pi)$. Then

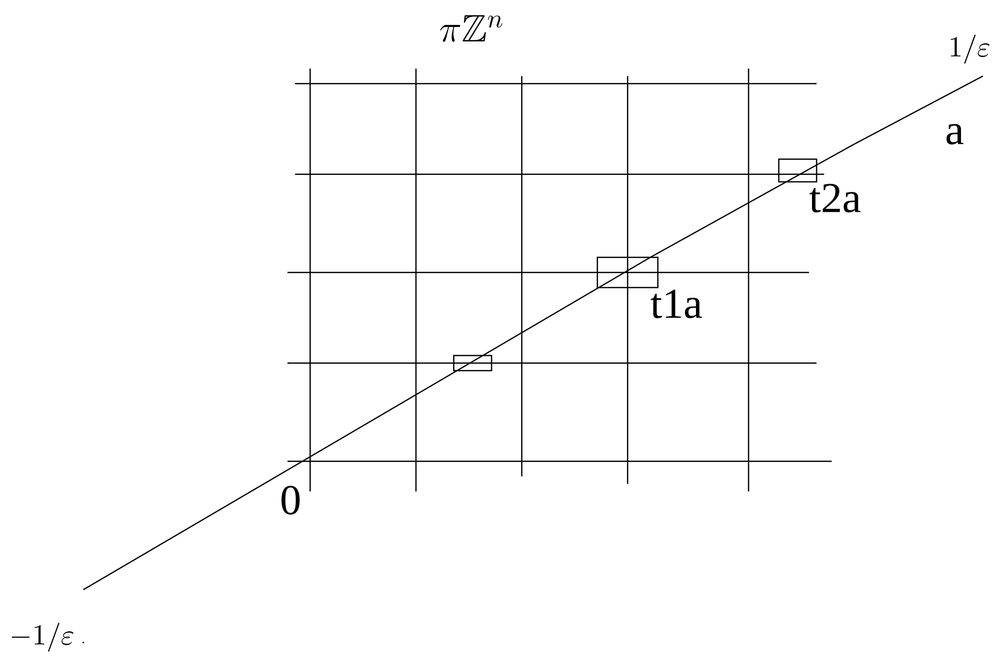
$$f(t) = \sum_1^n \sin^2(a_k t) \le \eta n$$

$\Rightarrow$ at most k terms in the above formula can be bigger than $\frac{\eta n}{k} =: \frac{\alpha^2}{4}$ or

$$\eta := \frac{\alpha^2 k}{4n}$$

All except k terms satisfy:

$$|\sin\left(\frac{a_k t}{\varepsilon}\right)| \leq \frac{\alpha}{2} \qquad \Rightarrow \qquad dist\left(\frac{a_k t}{\varepsilon}, \pi\mathbb{Z}\right) < \alpha$$



Definition 1. *(Recurrence Set) : The recurrence set $I(a) = I_{\alpha,k}(a)$ of a vector $a \in \mathbb{R}^n$ is the set of all $t \in \mathbb{R}$ s.t. all except k of the coordinates of a are α-close to $\mathbb{Z}$.*

So one has

$$\frac{t}{\pi\varepsilon} \in I(a) \qquad \Rightarrow \qquad \boxed{\mathrm{T}(\eta n, \pi) \ \subseteq \ \pi\varepsilon I(a)}$$

$$\underset{\textit{Bad set}}{\uparrow} \qquad\qquad \underset{\textit{Recurrence set}}{\uparrow}$$

and

$$|T(\eta n, \pi)| \ \leq \ |\pi\varepsilon I(a) \cap [-\pi, \pi)| \ = \ \pi\varepsilon |I(a) \cap [-\frac{1}{\varepsilon}, \frac{1}{\varepsilon}]|$$

Let us define density of the recurrence set in an interval symmetric w.r.t. the origin:

$$dens(I, y) := \frac{|I \cap [-y, y)|}{2y}$$

We thus proved :

$$|T(\eta n, \pi)| \leq 2\pi \cdot dens\left(I(a), \frac{1}{\varepsilon}\right)$$

It follows that

$$\boxed{\mathrm{p}_\varepsilon(a) \;\lesssim\; \tfrac{1}{\alpha\sqrt{k}} \cdot dens(I(a), \tfrac{1}{\varepsilon}) \;+\; e^{-c\alpha^2 k}}$$

The question arises about
<u>Density of the recurrence set</u> $\left(\lesssim \frac{1}{D(a)}\right)$?

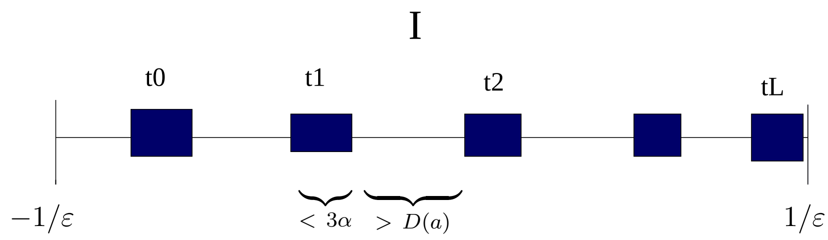
Suppose : density is big.
<u>Strategy:</u> 1) Recurrence set has lots of gaps ($\Rightarrow$ since $I \subset [-\frac{1}{\varepsilon}, \frac{1}{\varepsilon}]$ there exist small gaps)
2) Each gap bounds L.C.D. above ($\Rightarrow$ LCD small)
(then a may indeed be embedded into arithmetic progression)
$\uparrow$
Suppose $t_0 a, t_1 a$ (see picture above) are close to the integers $\mathbb{Z}^n \Rightarrow (t_1 - t_0)a$ is close to $\mathbb{Z}^n \Rightarrow D(a) \leq t_1 - t_0$

Lemma 2. *(Gaps) : Let* $t_0 \in I(a)$*. Then*
(1) $t_0 + 3\alpha \notin I(a)$
(2) Let $t_1 \in I(a)$ *be s.t.* $t_1 > t_0 + 3\alpha$*. Then* $t_1 - t_0 \geq D(a)$

<u>Proof:</u> *Exercise* (1): If $t_0 a$ is α-close to $\mathbb{Z}^n$, then $(t_0 + 3\alpha)a$ is not α-close anymore.
(2) See above : subtract.

Lemma 3. *(Recurrence set via LCD) :*

$$dens\left(I(a), \frac{1}{\varepsilon}\right) \lesssim \alpha \left(\varepsilon + \frac{1}{D(a)}\right)$$



<u>Proof:</u>

$$I = I(a) \cap [-\frac{1}{\varepsilon}, \frac{1}{\varepsilon}]$$

$$t_0 := \min\{t \in I\}$$

If $I \subseteq [t_0, t_0 + 3\alpha]$, then

$$dens\,(I(a), 1\varepsilon) \;\leq\; \frac{3\alpha}{2/\varepsilon} \;\lesssim\; \alpha\varepsilon \qquad\qquad QED$$

If not, define $t_1, \ldots, t_L \in I$:

$$t_l = \min t \in I : t_l > t_{l-1} + 3\alpha$$

By Part 1 of Lemma,

$$I \subset \cup_0^L \, [t_l, t_l + 3\alpha]$$

By Part 2,

$$t_L - t_0 = \sum_{l=1}^{L} (t_l - t_{l-1}) \geq L \cdot D(a)$$

(but $t_L - t_0 \leq 2/\varepsilon$).

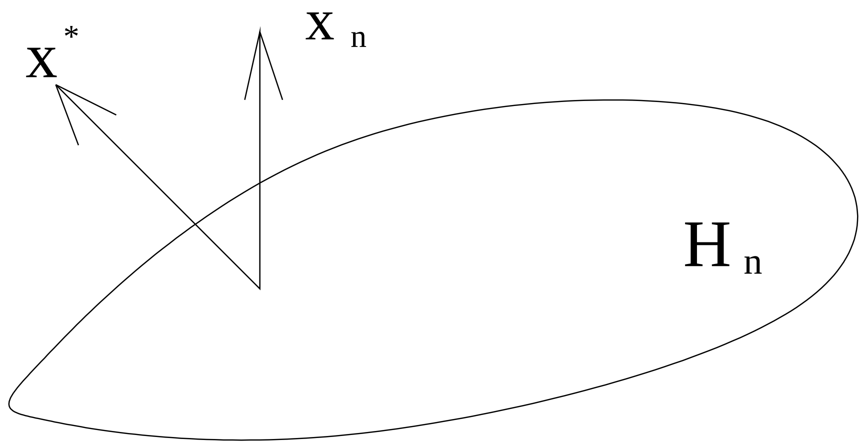
Therefore, combining the above statements, one obtains

$$dens\left(I(a), \frac{1}{\varepsilon}\right) \;\leq\; \frac{|\cup_0^L \, [t_l, t_l + 3\alpha]|}{L \cdot D(a)} \;=\; \frac{(L+1) \cdot 3\alpha}{L \cdot D(a)} \;\lesssim\; \frac{\alpha}{D(\alpha)} \qquad \blacksquare$$

Then finally the theorem follows:

$$p_\varepsilon(a) \;\leq\; \frac{1}{\sqrt{k}}\left(\varepsilon + \frac{1}{D(a)}\right) \;+\; e^{-c\alpha^2 k} \qquad \blacksquare \; S.B.P.T.$$

Application to Strong Invertibility Theorem



Apply S.B.P.T. for X^* = random normal (to span $(x_1, \ldots, x_n)$)

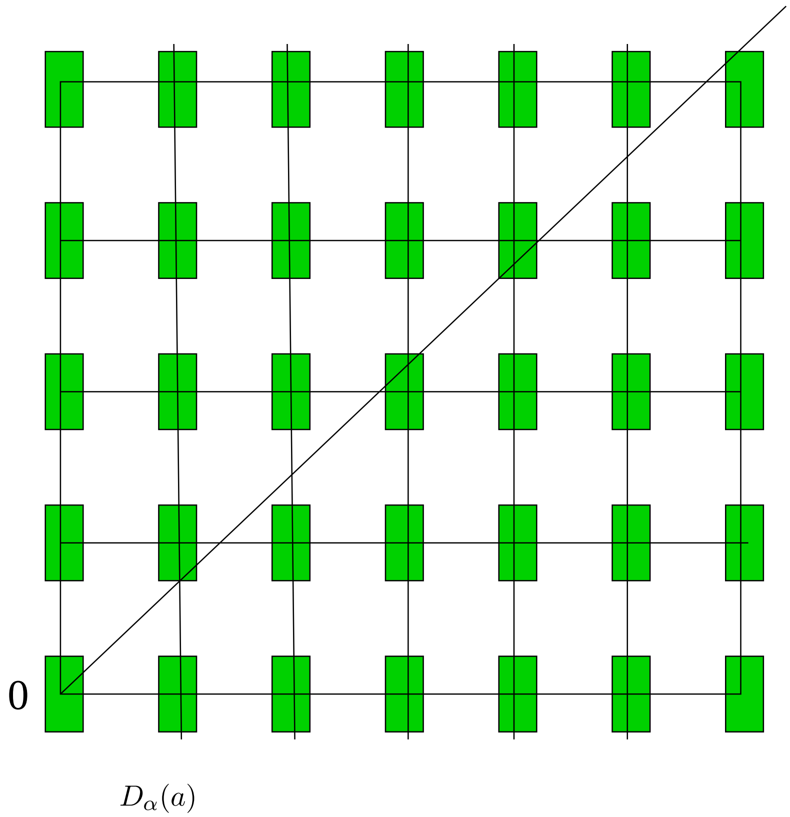
(For example, consider vector $(1, 1, \ldots, 1)$: then probability $\mathbb{P}(\sum \pm 1 = 0) \sim \frac{1}{\sqrt{n}}$. Now taking into account the factor $D(a)$ one gets that the coefficients are incomparable).

LCD of X^* is exponentially big ?

Ex. (Problem) : If $a = (a_1, \ldots, a_n)$ with i.i.d. gaussian a_k then $D(a) \geq e^{cn}$ with probability $1 - e^{-cn}$.

But X^* has no independent components.
We come to

Orchard model for LCD :



$D(a)$ = how far one can see in an orchard

Theorem 4 ([1]). : *With probability* $1 - e^{-cn}$, $D(X^*) \ \geq \ e^{cn}$

Implies **Distance bound**

$$dist(x_n, H_n) \ \geq \ |\langle X^*, x_n \rangle| \ = \ |\sum a_k \xi_k|$$

$$\boxed{\mathrm{P}(\mathrm{dist}(\mathrm{x}_n, H_n) < \varepsilon) \ \leq \ \varepsilon \ + \ c^n}$$

(Before we proved this with $\frac{1}{\sqrt{n}}$ instead of c^n).
This implies Strong Invertibility Theorem :

$$\boxed{\mathrm{P}\Big(s_n(A) \ \leq \ \frac{\varepsilon}{\sqrt{n}}\Big) \ \leq \ C\varepsilon \ + \ c^n}$$

(the last formula is true for all subgaussian matrices)

Conjecture: In the last formula

$c = \frac{1}{2} + o(1)$ (recall Bernoulli case).
C = 1

References

[1] Mark Rudelson and Roman Vershynin. The Littlewood-Offord problem and invertibility of random matrices. *http://www.math.ucdavis.edu/˜vershynin/papers/rv-invertibility.pdf preprint - submitted*, 2007.