Notes for talk by Elizabeth Meckes given at the MSRI
Connections for Women: Geometry and probability in high dimensions
workshop on August 17, 2017

1 Introduction

1.1 Concentration on $O(n)$

Orthogonal group of matrices: $O(n) = \{U : U \text{ is an } n \times n \text{ matrix}, UU^T = U^T U = I_n\}$

1.2 Haar measure

There exists a unique translation-invariant probability measure (p.m.) on $O(n)$:

$$U \sim \text{Haar}(O(n)) : \qquad\qquad U \overset{d}{=} AU \overset{d}{=} UA$$

where A is fixed in $O(n)$.

- Build one:
  - (i) fill $n \times n$ matrix with i.i.d. $\mathcal{N}(0,1)$
  - (ii) perform Gram-Schmidt to get U

2 Concentration of Measure

Idea: In some contexts, functions with small local fluctuations are "essentially constant".

Lemma 1 (Lévy)**.** *If $X \sim \text{unif}(\mathbb{S}^{n-1})$ and $f : \mathbb{R}^n \to \mathbb{R}$ is L-Lipschitz, then*

$$\mathbb{P}(|f(X) - M| > Lt) \leq 2e^{-(n-2)t^2}.$$

- Think $t^2 = \frac{C}{n-2}$ where C is a large constant.
- We cannot have concentration on $O(n)$. But we have the next best thing:

Theorem 1 (Milman-Schectman)**.** *If $F : O(n) \to \mathbb{R}$ is L-Lipschitz with respect to the Hilbert-Schmidt norm $\|\cdot\|_{HS}$ (i.e. $\langle A, B\rangle_{HS} = \text{tr}(AB^*)$) and $U \sim \text{Haar}(SO(n))$ or $U \sim \text{Haar}(SO^-(n))$, then*

$$\mathbb{P}(|F(U) - \mathbb{E}F(U)| > t) \leq 2e^{-\frac{(n-2)t^2}{4L^2}}.$$

Johnson-Lindenstrauss Lemma: If you project n points in $\mathbb{R}^d$ onto a random $\sim \log n$-dimensional subspace, then the pairwise distances between poitns are basically preserved. The formal statement of the J-L lemma follows.

Lemma 2 (Johnson-Linderstrauss)**.** *There exist universal constants c, C such that the following holds: Let $\{x_j\}_{j=1}^n \subset \mathbb{R}^d$. Let P be a random $k \times d$ matrix given by the first k rows of $U \sim$ Haar$(O(d))$. Fix $\varepsilon > 0$, and let $k := \frac{a \log n}{\varepsilon^2}$. With probability $\geq 1 - Cn^{2-ac/4}$,*

$$(1-\varepsilon)\|x_i - x_j\|^2 \leq \frac{d}{k}\|Px_j - Px_i\|^2 \leq (1+\varepsilon)\|x_i - x_j\|^2.$$

Proof. Fix i, j and let $x := \frac{x_i - x_j}{\|x_i - x_j\|}$. We want: $\sqrt{1-\varepsilon} \leq \sqrt{\frac{d}{k}}\|Px\| \leq \sqrt{1+\varepsilon}$ with high probability. Let $F_x(U) := \sqrt{\frac{d}{k}}\|Px\|$. Then $F_x : O(d) \to \mathbb{R}$ is $\sqrt{\frac{d}{k}}$-Lipschitz:

$$\begin{aligned} |F_x(U) - F_x(U')| &= \sqrt{\frac{d}{k}}\Big|\|Px\| - \|P'x\|\Big| \leq \sqrt{\frac{d}{k}}\|(P-P')x\| \leq \sqrt{\frac{d}{k}}\|P - P'\|_{op} \\ &\overset{CS}{\leq} \sqrt{\frac{d}{k}}\|P' - P\|_{HS} \leq \sqrt{\frac{d}{k}} d_{HS}(U, U'). \end{aligned}$$

This implies

$$\mathbb{P}(|F_x(U) - \mathbb{E}F_x(U)| > t) \leq 2e^{-\frac{cdt^2}{d/k}} = e^{-ckt^2} = e^{-\frac{ca\log n}{\varepsilon^2}t^2}.$$

- What is the value of $\mathbb{E}F_x(U)$?

  $$\mathbb{E}F_x(U)^2 = \frac{d}{k}\mathbb{E}\|Px\|^2 = \frac{d}{k}\mathbb{E}(\sqrt{u_{11}^2 + \cdots + u_{ik}^2})^2 = 1$$

  where the last equality follows by symmetry.

- $\mathrm{var}(F_x(U)) = 1 - (\mathbb{E}F_x(U))^2$, so $(\mathbb{E}F_x(U))^2 = 1 - \mathrm{var}(F_x(U))$.
- $\mathrm{var}(F_x(U)) = \int_0^\infty \mathbb{P}(|F_x(U) - \mathbb{E}F_x(U)|^2 > t)\,dt \leq \int_0^\infty Ce^{-ckt}\,dt = \frac{C}{ck}$ and $k = \frac{a\log n}{\varepsilon^2}$, so take $\varepsilon < \frac{ca\log n}{C + cn\log n}$ to get $1 - \frac{C}{ck} \geq 1 - \varepsilon/2$.
- The "good" set has probability $1 - Ce^{-ac\log n}$. We have n^2 good sets; intersect them all.

□

3 Dvoretzky's Theorem

Theorem 2 (Dvoretzky)**.** *Let $\|\cdot\|$ be an arbitrary norm on $\mathbb{C}^n$. There is an invertible linear map $T : \mathbb{C}^n \to \mathbb{C}^n$ such that for $\varepsilon > 0$, if $k \leq C\varepsilon^2 \log n$ and $E \subset \mathbb{C}^n$ is random subspace with $\dim(E) = k$, then with probability $\geq 1 - Ce^{-ck}$,*

$$1 - \varepsilon \leq \frac{\|Tv\|}{\|v\|} \leq 1 + \varepsilon,\ \forall v \in E.$$

3.1 Outline of a proof

0.) Assume the maximal volume ellipsoid in $\{\|\cdot\| = 1\}$ is $\mathbb{S}^n_{\mathbb{C}}$.
1.) Consider $X_v(U) := \|U_v\| - \mathbb{E}_U\|U_v\|$ where U is random and v is fixed. Consider

$$F_E(U) = \sup_{\substack{v\in E\\ \|v\|=1}} |X_v(U)|.$$

Then $F_E(U)$ is 1-Lipschitz, so it concentrates at its mean $\mathbb{E}F_E(U)$.

- $\mathbb{E}F_E(U) = \mathbb{E}\left(\sup_{\substack{v\in E\\ \|v\|=1}} \left|\|U_v\| - \underbrace{\mathbb{E}\|U_v\|}_{M}\right|\right)$
- $\mathbb{P}(\|U_v\| - \|U_w\| > t) \le Ce^{-\frac{cnt^2}{\|v-w\|^2}}$
- "Sub-Gaussian increment" $d(v,w) = \frac{\|v-w\|}{\sqrt{n}}$
- Dudley's entropy bound $\mathbb{E}\left[\sup_{\substack{v\in E\\ \|v\|=1}} |X_v(U)|\right] \le C\sqrt{\frac{k}{n}}$ implies that for all $v \in E$,

  $$\left|\|U_v\| - M\right| \le t + C\sqrt{\frac{k}{n}}$$

  with high probability (here $c\sqrt{\frac{\log n}{n}} \le M \le 1$).
- Now fix $\varepsilon > 0$, $k = cnM^2\varepsilon^2$. For $t = M\varepsilon/2$ we have with high probability $\ge 1 - Ce^{-c\varepsilon^2 \log n}$, $M(1-\varepsilon) \le \|U_v\| \le M(1+\varepsilon)$ for all $v \in E$.
- Finally, choose $E = \text{span}\{e_1, \dots, e_k\}$. □

Theorem 3 (Meckes, '13)**.** *Let X be a random vector in $\mathbb{R}^d$ with $\mathbb{E}\|X\|^2 = d$, $\mathbb{E}\left|\|X\|^2 - d\right| \le \frac{Ld}{\sqrt[3]{\log d}}$ and $\sup_{\xi\in\mathbb{S}^{n-1}} \mathbb{E}[\langle \xi, X\rangle^2] = 1$. Also let $X_v^{(k)} = (\langle x, v_1\rangle, \dots, \langle x, v_k\rangle)$ and $V = \begin{pmatrix} -v_1- \\ \vdots \\ -v_d- \end{pmatrix} \in O(d)$. Fix $\delta > 0$ and let $k := \frac{\delta \log d}{\log\log d}$. Then $\exists c = c(\delta) > 0$ such that with high probability, $d_{BL}(X_v^{(k)}, Z) \le Ce^{-c\log\log d}$ where Z is a Gaussian.*

[Here $d_{BL}(X,Y) = \sup_{\substack{f:\mathbb{R}^k\to\mathbb{R}\\ \|f\|_\infty \le 1\\ |f|_C \le 1}} |\mathbb{E}f(X) - \mathbb{E}f(Y)|$ is the bounded Lipschitz distance.]