# Lecture 5

## 1. Distribution of matrix entries of $O(N)$

<u>Prop</u>: For random $g \in O(N)$ (chosen via Haar measure), the distribution of entries $g_{ij}$ is identical for any choice of $ij$

$$\begin{pmatrix} \boxed{g_{11}} & \cdots & & \\ \boxed{g_{21}} & \boxed{g_{22}} & & \\ \vdots & & \ddots & \\ & & & \boxed{g_{nn}} \end{pmatrix}$$

All identical dist.
(though not independent)

Proof: We prove a stronger claim.

For any unit vectors $u, v \in \mathbb{R}^N$, the random var.

$$\langle u, gv \rangle$$

has the same dist., no matter the choice of $u, v$

Let $u = e_i$, $v = e_j$ to imply Prop.

For $u, v, u', v' \in \mathbb{R}^N$ unit vectors, why is it that

$$\langle u, gv \rangle \underset{\text{dist.}}{\sim} \langle u', gv' \rangle \ ?$$

Find

$h, k \in O(N)$ s.t. $u' = hu$, $v' = kv$.

Have

$$\langle u', gv' \rangle = \langle hu, gkv \rangle = \langle u, (h^t g k) v \rangle$$

But

$$h^t g \sim g \quad \text{and} \quad h^t g k \sim g k \sim g$$

(as $g$ Haar dist.)

so

$$\langle u', gv' \rangle \sim \langle u, gv \rangle. \qquad \square$$

Observation: $\mathbb{E} g_{11} = 0$ for $g \in O(N)$ random.

Why? $-I \in O(N)$, so $-g \sim g \Rightarrow -g_{11} \sim g_{11}$

$$\Rightarrow \mathbb{E} g_{11} = 0.$$

<u>Observation</u>: $\mathbb{E}|g_{11}|^2 = \frac{1}{N}$.

Why? For $g \in O(N)$, $|g_{11}|^2 + |g_{21}|^2 + \dots + |g_{N1}|^2 = 1$

$\Rightarrow \mathbb{E}|g_{11}|^2 + \dots + \mathbb{E}|g_{N1}|^2 = 1 \Rightarrow N \cdot \mathbb{E}|g_{11}|^2 = 1$

<u>Thm</u>: For $g \in O(N)$ Haar dist.,

$$g_{11} \sim Y_1$$

for $Y = (Y_1, \dots, Y_N) \in \mathbb{R}^N$ a random vector unif. dist. on the unit sphere $S^{N-1}$.

(What is unif. dist. on $S^{N-1}$? Intuitively: $Y$ equally likely to be at any point of $S^{N-1}$. Rigorously: $Y$ distributed according to unique rotationally invariant measure on $S^{N-1}$

$\Leftrightarrow$ $hY$ has same dist as $Y$ for any fixed $h \in O(N)$)

## Proof of Thm:

**Explanation 1:** dist of $g$ invariant under $O(N)$ action

$\Downarrow$

dist. of first column of $g$ invariant under $O(N)$

$\Downarrow$

$$\begin{pmatrix} g_{11} \\ \vdots \\ g_{N1} \end{pmatrix} \sim \begin{pmatrix} Y_1 \\ \vdots \\ Y_N \end{pmatrix}$$ u.d. on $S^{N-1}$, invariant under $O(N)$. $\square$

**Explanation 2:** Have shown for $X = \begin{pmatrix} X_1 \\ \vdots \\ X_N \end{pmatrix} \in \mathbb{R}^N$ with $X_i \sim N_{\mathbb{R}}(0,1)$ iid,

$$\begin{pmatrix} g_{11} \\ \vdots \\ g_{N1} \end{pmatrix} \sim \frac{X}{\|X\|}$$

(first step of Gram-Schmidt, projects to unif. dist. on $S^{N-1}$)

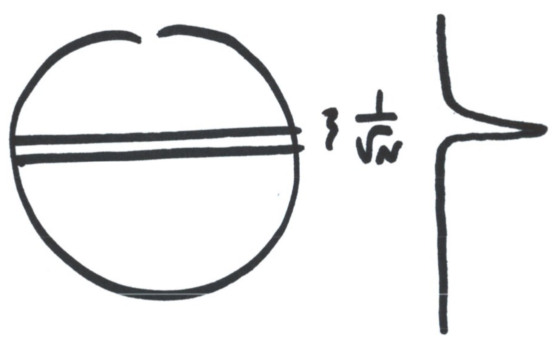
$\square$

## 2. Distribution of coordinates of $S^{N-1}$

__Thm__ (Borel's lemma 1906): For $Y = \begin{pmatrix} Y_1 \\ \vdots \\ Y_N \end{pmatrix}$ unif. dist. on $S^{N-1}$

$$\mathbb{P}\left( Y_1 \in \left[ \frac{\alpha}{\sqrt{N}}, \frac{\beta}{\sqrt{N}} \right] \right) \to \frac{1}{\sqrt{2\pi}} \int_\alpha^\beta e^{-t^2/2}\, dt.$$

(That is $\sqrt{N} \cdot Y_1 \xrightarrow{\text{dist.}} N_{\mathbb{R}}(0,1)$)



<u>Proof</u>: For $X = \begin{pmatrix} x_1 \\ \vdots \\ x_N \end{pmatrix}$ with iid $X_i \sim N_{\mathbb{R}}(0,1)$

$$Y_1 \sim \frac{X_1}{\|X\|} = \frac{X_1}{\sqrt{X_1^2 + \cdots + X_N^2}}$$

Observe:

$$\mathbb{E}\left(X_1^2 + \cdots + X_N^2\right) = N$$

So by law of large numbers $(X_1^2 + \cdots + X_N^2)/N \to 1$ in prob.

So for any $\varepsilon > 0$,

$$\mathbb{P}\left( \frac{\sqrt{X_1^2 + \cdots + X_N^2}}{\sqrt{N}} \text{ is } \geq (1+\varepsilon) \text{ or } \leq (1-\varepsilon) \right) = o(1)$$

as $N \to \infty$

Hence for any $\varepsilon > 0$



$$\mathbb{P}\left(\frac{X_1}{\sqrt{X_1^2 + \cdots + X_N^2}} \in \left[\frac{\alpha}{\sqrt{N}}, \frac{\beta}{\sqrt{N}}\right]\right) \leq \mathbb{P}\left(X_1 \in [(1-\varepsilon)\alpha, (1+\varepsilon)\beta]\right) + o(1)$$

(for $\alpha, \beta > 0$)

$$\to \int_{(1-\varepsilon)\alpha}^{(1+\varepsilon)\beta} \frac{e^{-t^2/2}}{\sqrt{2\pi}} \, dt$$

As $\varepsilon > 0$ arbitrary

$$\limsup_{N \to \infty} \mathbb{P}\left(\frac{X_1}{\|X\|} \in \left[\frac{\alpha}{\sqrt{N}}, \frac{\beta}{\sqrt{N}}\right]\right) \leq \int_{\alpha}^{\beta} \frac{e^{-t^2/2}}{\sqrt{2\pi}} \, dt .$$

But same arg. shows

$$\limsup_{N \to \infty} \mathbb{P}\left(\frac{X_1}{\|X\|} \notin \left[\frac{\alpha}{\sqrt{N}}, \frac{\beta}{\sqrt{N}}\right]\right) \leq \int_{t \notin [\alpha, \beta]} \frac{e^{-t^2/2}}{\sqrt{2\pi}} \, dt$$

$$\Rightarrow \mathbb{P}\left(\frac{X_1}{\|X\|} \in \left[\frac{\alpha}{\sqrt{N}}, \frac{\beta}{\sqrt{N}}\right]\right) \to \int_\alpha^\beta \frac{e^{-t^2/2}}{\sqrt{2\pi}}\, dt.$$

$\square$

<u>Cor</u>: For $g \in O(N)$,

$$\sqrt{N}\, g_{11} \longrightarrow N_{\mathbb{R}}(0,1)$$

That is

$$g_{11} \approx \frac{N_{\mathbb{R}}(0,1)}{\sqrt{N}} \Rightarrow g_{ij} \approx \frac{N_{\mathbb{R}}(0,1)}{\sqrt{N}} \quad \forall ij$$