TTIC 31150/CMSC 31150 Mathematical Toolkit (Spring 2023)

Avrim Blum and Ali Vakilian

Lecture 15: Properties of unit ball in high dimensions

Recap

- Chernoff-Hoeffding bounds
- Use in randomized algorithm for routing to minimize congestion.
- Randomized complexity classes **RP** and **BPP**, connections to **P/poly**.
- Probability over uncountably-infinite spaces (σ-algebra, R.V. as measurable function).
- Gaussian Random Variables
- Dimensionality Reduction and the Johnson-Lindenstrauss Lemma

Volume in High-Dimensions

- Consider an object $O \in \mathbb{R}^d$. Shrink O by an ϵ-factor and call it O'.
  - $O' = \{(1-\epsilon)x \mid x \in O\}$

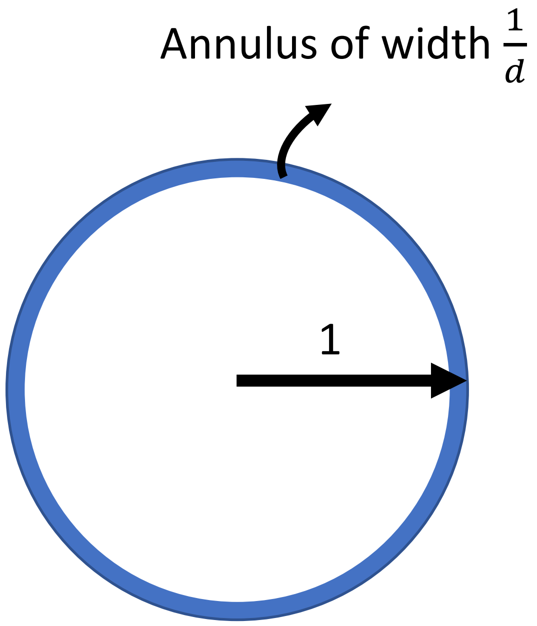
$$\text{volume}(O') = (1-\epsilon)^d \text{ volume}(O)$$

➢ Check it for cubes.

➢ Consider the partition of O into infinitesimal cubes.

- By $1 - x \leq e^{-x}$,

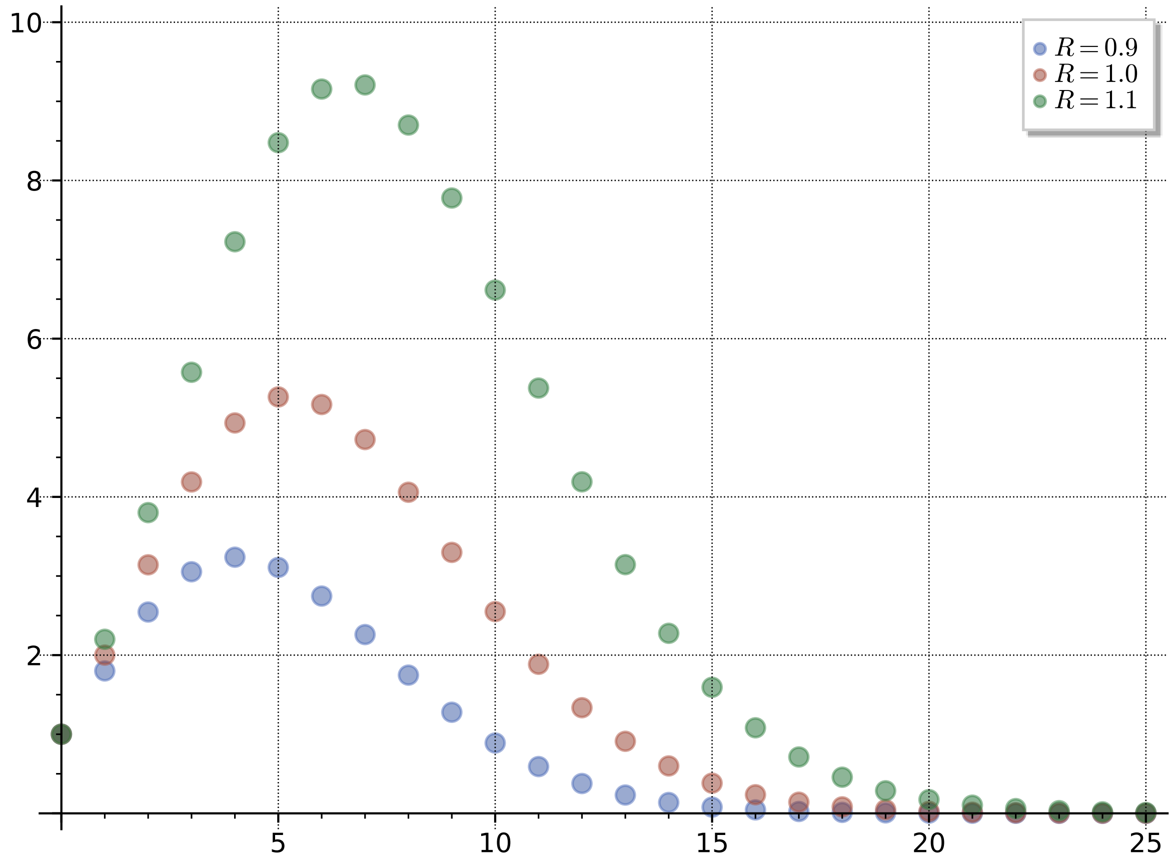
$$\text{volume}(O') \leq e^{-\epsilon d} \text{ volume}(O)$$



- By fixing ϵ, as $d \to \infty$, $\frac{\text{volume}(O')}{\text{volume}(O)}$ approaches to zero
- Most of the volume is in the annulus of width $O(1/d)$ near boundary

Volume of the unit ball

- For fixed dimension d, the volume of sphere is a function of r and
- What about fixing r and increasing the dimension?



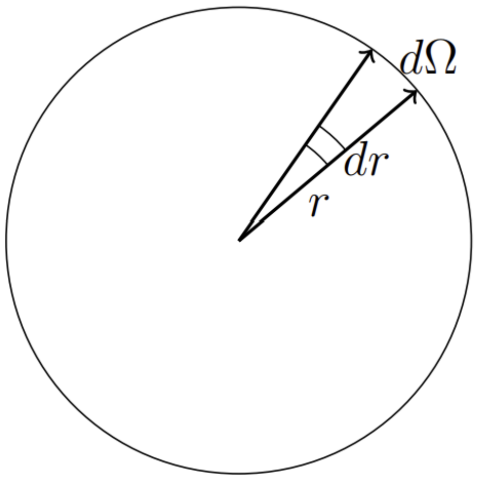
Dimension	Volume of a ball of radius R
0	1
1	$2R$
2	$\pi R^2 \approx 3.142 \times R^2$
3	$\frac{4\pi}{3} R^3 \approx 4.189 \times R^3$
4	$\frac{\pi^2}{2} R^4 \approx 4.935 \times R^4$
5	$\frac{8\pi^2}{15} R^5 \approx 5.264 \times R^5$
6	$\frac{\pi^3}{6} R^6 \approx 5.168 \times R^6$
7	$\frac{16\pi^3}{105} R^7 \approx 4.725 \times R^7$
8	$\frac{\pi^4}{24} R^8 \approx 4.059 \times R^8$
9	$\frac{32\pi^4}{945} R^9 \approx 3.299 \times R^9$
10	$\frac{\pi^5}{120} R^{10} \approx 2.550 \times R^{10}$
11	$\frac{64\pi^5}{10395} R^{11} \approx 1.884 \times R^{11}$

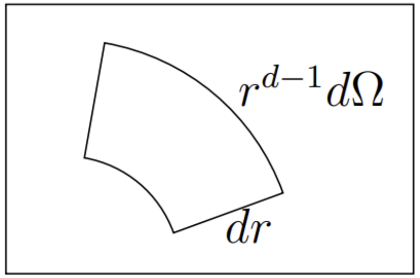
Volume of the unit ball

- For fixed dimension d, the volume of sphere is a function of r and grows as r^d
- What about fixing r and increasing the dimension?

$$V(d) = \int_{x_1=-1}^{x_1=1} \int_{x_2=-\sqrt{1-x_1^2}}^{x_2=\sqrt{1-x_1^2}} \cdots \int_{x_d=-\sqrt{1-x_1^2-x_2^2-\cdots x_{d-1}^2}}^{x_d=\sqrt{1-x_1^2-x_2^2-\cdots x_{d-1}^2}} dx_1 dx_2 \cdots dx_d$$

$$V(d) = \int_{S^d} \int_{r=0}^{1} r^{d-1} d\Omega dr$$





Volume of the unit ball

- For fixed dimension d, the volume of sphere is a function of r and grows as r^d
- What about fixing r and increasing the dimension?

$$V(d) = \int_{x_1=-1}^{x_1=1} \int_{x_2=-\sqrt{1-x_1^2}}^{x_2=\sqrt{1-x_1^2}} \cdots \int_{x_d=-\sqrt{1-x_1^2-x_2^2-\cdots x_{d-1}^2}}^{x_d=\sqrt{1-x_1^2-x_2^2-\cdots x_{d-1}^2}} dx_1 dx_2 \cdots dx_d$$

$$V(d) = \int_{S^d} \int_{r=0}^{1} r^{d-1} d\Omega dr$$

$$V(d) = \frac{2\pi^{\frac{d}{2}}}{d\,\Gamma(\frac{d}{2})}$$

$\Gamma(x)$: generalization of factorial for non-integers

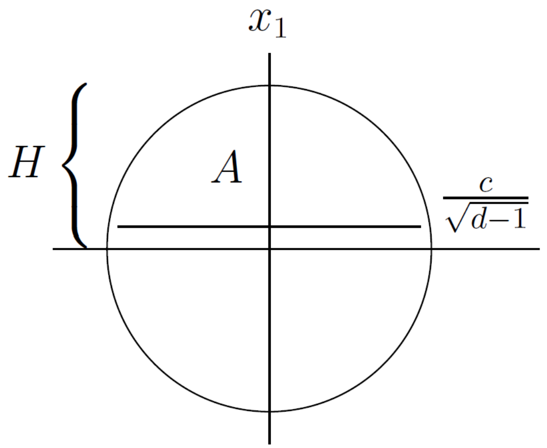
- $\Gamma(x) = (x-1)\Gamma(x-1)$
- $\Gamma(2) = \Gamma(1) = 1, \Gamma\left(\frac{1}{2}\right) = \sqrt{\pi}$
- For integer x, $\Gamma(x) = (x-1)!$

$$\lim_{d\to\infty} V(d) = \lim_{d\to\infty} \frac{2\,\pi^{d/2}}{d\,\Gamma(d/2)} = 0$$

Volume of the unit ball near Equator

- We saw that most of the volume of a unit ball is around its boundary.
- Now, we show similar property for near equator.

Theorem 2.2 *For $c \geq 1$ and $d \geq 3$, at least $1 - \frac{2}{c}e^{-c^2/2}$ fraction of the volume of the d-dimensional unit ball has $|x_1| \leq c/\sqrt{d-1}$.*



- ***H***: upper hemisphere
- ***A***: portion of unit ball with $x_1 \geq c/\sqrt{d-1}$

$$\text{volume}(A) = \int_{c/\sqrt{d-1}}^{1} (1 - x_1^2)^{\frac{d-1}{2}} V(d-1) dx_1$$

$$\text{volume}(A) \leq \int_{c/\sqrt{d-1}}^{\infty} \frac{x_1\sqrt{d-1}}{c} e^{-\frac{d-1}{2}x_1^2} V(d-1) dx_1 = V(d-1)\frac{\sqrt{d-1}}{c} \int_{c/\sqrt{d-1}}^{\infty} x_1 e^{-\frac{d-1}{2}x_1^2} dx_1$$

$$= \frac{V(d-1)}{c\sqrt{d-1}} e^{-\frac{c^2}{2}}$$

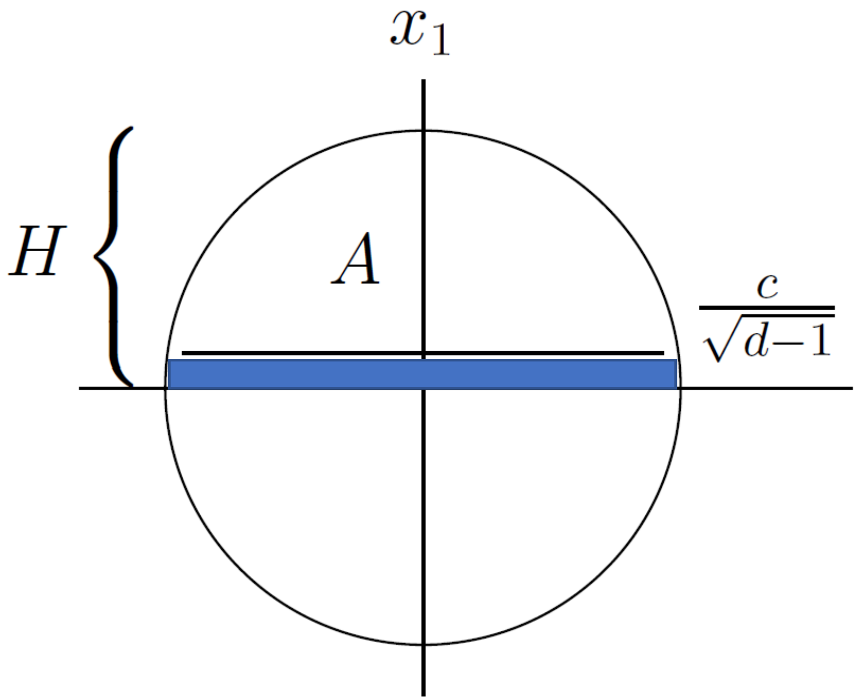
- $1 - x \leq e^{-x}$
- $\frac{x_1\sqrt{d-1}}{c} \geq 1$ for every $x_1 \geq c/\sqrt{d-1}$

Volume of the unit ball near Equator

- We saw that most of the volume of a unit ball is around its boundary.
- Now, we show similar property for near equator.

Theorem 2.2 *For $c \geq 1$ and $d \geq 3$, at least $1 - \frac{2}{c}e^{-c^2/2}$ fraction of the volume of the d-dimensional unit ball has $|x_1| \leq c/\sqrt{d-1}$.*

- ***H***: upper hemisphere
- ***A***: portion of unit ball with $x_1 \geq c/\sqrt{d-1}$



Lowerboud on volume of ***H***

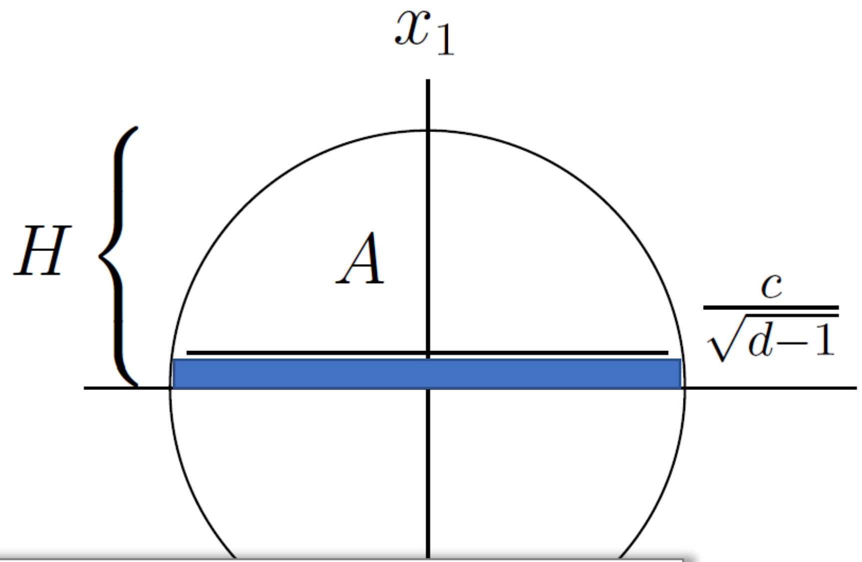
- Cylinder of height $1/\sqrt{d-1}$ and radius $\sqrt{1-\frac{1}{d-1}}$
- $V(d-1)\left(1-\frac{1}{d-1}\right)^{\frac{d-1}{2}}\frac{1}{\sqrt{d-1}} \xrightarrow{\text{by } (1-x)^a \geq 1-ax} \frac{V(d-1)}{2\sqrt{d-1}}$

Volume of the unit ball near Equator

- We saw that most of the volume of a unit ball is around its boundary.
- Now, we show similar property for near equator.

Theorem 2.2 *For $c \geq 1$ and $d \geq 3$, at least $1 - \frac{2}{c}e^{-c^2/2}$ fraction of the volume of the d-dimensional unit ball has $|x_1| \leq c/\sqrt{d-1}$.*

- ***H***: upper hemisphere
- ***A***: portion of unit ball with $x_1 \geq c/\sqrt{d-1}$



- $\text{volume}(A) \leq \frac{V(d-1)}{c\sqrt{d-1}} e^{-c^2/2}$
- $\text{volume}(H) \geq \frac{V(d-1)}{2\sqrt{d-1}}$

$$\frac{\text{volume}(A)}{\text{volume}(H)} \leq \frac{2}{c} e^{-c^2/2}$$

Another proof that the volume of the ball approaches to zero as $d \to \infty$?

Revisiting near orthogonality

- If we draw two points at random from the unit ball, w.h.p., they will be nearly orthogonal.
  - W.h.p., both will be close to the surface,
  - W.h.p., both will have length $1 - O(1/d)$
  - $\Longrightarrow$ w.h.p., their inner product is $\pm O(1/\sqrt{d})$

Theorem 2.4 *Consider drawing n points $\mathbf{x}_1, \cdots, \mathbf{x}_n$ at random from the unit ball. With probability $1 - O(1/n)$,*

- $|\mathbf{x}_i| \geq 1 - \frac{2 \ln n}{d}$ *for all i, and*
- $|\mathbf{x}_i \cdot \mathbf{x}_j| \leq \frac{\sqrt{6 \ln n}}{\sqrt{d-1}}$ *for all $i \neq j$.*

How does a typical random vector on the sphere look like?

Revisiting near orthogonality

Theorem 2.4 *Consider drawing n points* $\mathbf{x}_1, \cdots, \mathbf{x}_n$ *at random from the unit ball. With probability* $1 - O(1/n)$,

- $|\mathbf{x}_i| \geq 1 - \frac{2\ln n}{d}$ *for all i, and*
- $|\mathbf{x}_i \cdot \mathbf{x}_j| \leq \frac{\sqrt{6\ln n}}{\sqrt{d-1}}$ *for all* $i \neq j$.

- We saw that for a vector $\mathbf{x}$ sampled from the unit ball $\Pr[|\mathbf{x}| < 1 - \epsilon] \leq e^{-\epsilon d}$
  - Setting $\epsilon = \frac{2\ln n}{d}$, with probability $1 - 1/n$, all vectors $\mathbf{x}_1, \cdots, \mathbf{x}_n$ have length at least $1 - \frac{2\ln n}{d}$
- For any pair i, j: consider $\mathbf{x}_i$ as "north", then projection of $\mathbf{x}_j$ onto the north direction is more than $\frac{c}{\sqrt{d-1}}$ with probability at most $\frac{2}{c}e^{-c^2/2}$
  - Setting $c = \sqrt{6\ln n}$, with probability $1 - 1/n$, all inner products are at most $\frac{\sqrt{6\ln n}}{\sqrt{d-1}}$

Generating points uniformly at random from a ball

- First, consider the case of two-dimensional space:
  - Sample from $[-1,1]^2$ and project onto the ball
  - Sample from $[-1,1]^2$ and reject those outside t
- Higher dimensions?
  - Volume of the unit ball approaches to zero.

For generating from the entire unit ball:

- Instead of normalizing, $\rho \frac{\mathbf{x}}{\|\mathbf{x}\|}$
- density of ρ is proportional to r^{d-1}
- $\rho(r) = dr^{d-1}$

Solution: generate a point each of whose coordinates is an independent Gaussian variable. The probability density of $\mathbf{x}$ is $p(\mathbf{x}) = \frac{1}{(2\pi)^{\frac{d}{2}}} e^{-\frac{x_1^2 + x_2^2 + \cdots + x_d^2}{2}}$

- Spherically symmetric
- Normalizing the vector $\mathbf{x} = (x_1, \cdots, x_d)$ to a unit vector