Volumes in high-dimensional space

Daniel Hsu

COMS 4772

Simple volumes

- In $\mathbb{R}^1$, line segment

$$[a, b] = \{x \in \mathbb{R} : a \leq x \leq b\}$$

has one-dimensional volume (a.k.a. *length*) $b - a$.

- In $\mathbb{R}^2$, square

$$[a, b]^2 = \{(x_1, x_2) \in \mathbb{R}^2 : x_1, x_2 \in [a, b]\}$$

has two-dimensional volume (a.k.a. *area*) $(b - a)^2$.

- In $\mathbb{R}^3$, cube

$$[a, b]^3 = \{(x_1, x_2, x_3) \in \mathbb{R}^3 : x_1, x_2, x_3 \in [a, b]\}$$

has three-dimensional volume (a.k.a. *volume*) $(b - a)^3$.

d-dimensional volumes

- Hypercube

$$[a, b]^d = \{(x_1, x_2, \dots, x_d) \in \mathbb{R}^d : x_1, x_2, \dots, x_d \in [a, b]\}$$

  has d-dimensional volume $(b - a)^d$.
- Use $\mathrm{vol}(A)$ to denote d-dimensional volume of $A \subseteq \mathbb{R}^d$.
- For $A \subseteq \mathbb{R}^d$ and $c \geq 0$, let

$$cA := \{c\boldsymbol{x} : \boldsymbol{x} \in A\}.$$

  - Example: if $A = [0, 1]^d$, then $cA = [0, c]^d$ and $\mathrm{vol}(cA) = c^d$.
- In general,

$$\mathrm{vol}(cA) = c^d \, \mathrm{vol}(A).$$

Weird facts about the unit ball

Unit ball $B^d := \{\boldsymbol{x} \in \mathbb{R}^d : \|\boldsymbol{x}\|_2 \leq 1\}$.

1. Lengths of most points in B^d are close to one.
2. Most points in B^d are near the "equator".
3. $\lim_{d \to \infty} \mathrm{vol}(B^d) = 0$.
   - By contrast, hypercube $[-1, 1]^d$ has volume 2^d.

Length of most points in the unit ball

- For $\varepsilon \in (0,1)$, consider $(1-\varepsilon)B^d$ (i.e., ball of radius $1-\varepsilon$).
- $\mathrm{vol}((1-\varepsilon)B^d) = (1-\varepsilon)^d \,\mathrm{vol}(B^d)$
- Therefore

  $$(1-\varepsilon)^d \;\le\; e^{-\varepsilon d}$$

  fraction of points in B^d have length at most $1-\varepsilon$.

Most points in unit ball are near the "equator"

- Let $\boldsymbol{u}$ be a unit vector ("north pole"), and $\varepsilon \in (0,1)$.
- "Equator": $\{\boldsymbol{x} \in B^d : \langle \boldsymbol{u}, \boldsymbol{x} \rangle = 0\}$
- "Tropics": $\{\boldsymbol{x} \in B^d : -\varepsilon \le \langle \boldsymbol{u}, \boldsymbol{x} \rangle \le \varepsilon\}$
- Points north of the tropics, $\{\boldsymbol{x} \in B^d : \langle \boldsymbol{u}, \boldsymbol{x} \rangle > \varepsilon\}$, are within distance $\sqrt{1-\varepsilon^2}$ of $\varepsilon\boldsymbol{u}$.
  - Hence contained in ball of radius $\sqrt{1-\varepsilon^2}$.
  - Volume is at most $(1-\varepsilon^2)^{d/2}\,\mathrm{vol}(B^d)$.
- Similarly, points south of tropics have volume at most $(1-\varepsilon^2)^{d/2}\,\mathrm{vol}(B^d)$.
- So volume outside tropics is at most

  $$2(1-\varepsilon^2)^{d/2}\,\mathrm{vol}(B^d) \;\le\; 2e^{-\varepsilon^2 d/2}\,\mathrm{vol}(B^d)\,.$$

Volume of unit ball

- Consider an orthonormal basis $\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_d$ of $\mathbb{R}^d$.
- Let T_i be the "tropics" when $\boldsymbol{u}_i$ is the "north pole".
- Volume of points in $\bigcap_{i=1}^d T_i$ is

$$\operatorname{vol}\left(\bigcap_{i=1}^{d} T_i\right) \;\geq\; \operatorname{vol}(B^d) - \sum_{i=1}^{d} \operatorname{vol}(T_i^c) \;\geq\; \left(1 - 2de^{-\varepsilon^2 d/2}\right) \operatorname{vol}(B^d)\,.$$

- But $\operatorname{vol}\left(\bigcap_{i=1}^d T_i\right) = \operatorname{vol}([-\varepsilon, \varepsilon]^d) = (2\varepsilon)^d$.
- If $2de^{-\varepsilon^2 d/2} \leq 1$, then

$$\operatorname{vol}(B^d) \;\leq\; \frac{(2\varepsilon)^d}{1 - 2de^{-\varepsilon^2 d/2}}\,.$$

- For $\varepsilon = \sqrt{2\ln(4d)/d}$, bound is

$$\operatorname{vol}(B^d) \;\leq\; 2\left(\frac{8\ln(4d)}{d}\right)^{d/2} \;\xrightarrow{d\to\infty}\; 0\,.$$