High Dimensional Space

Jayati Kaushik

St. Joseph's University, Bengaluru

Foundations of Data Science
BDA2121

Law of Large Numbers

$$\text{Prob}\left(\left|\frac{x_1 + x_2 + \cdots + x_n}{n} - E(x)\right| \geq \epsilon\right) \leq \frac{Var(x)}{n\epsilon^2}$$

The Geometry of High Dimension

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

The Geometry of High Dimension

If S is the unit ball in d-dimensions.
Then, atleast $1 - e^{-\epsilon d}$ fraction of the ball is concentrated in $S \;/\; (1-\epsilon)S$

Properties of a Unit Ball

Volume of the unit ball

Lemma

The surface area $A(d)$ and the volume $V(d)$ of a unit-radius ball in d-dimensions are given by

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

Properties of the Unit Ball

Volume Near the equator

Theorem

For $c \geq 1$ and $d \geq 3$, at least a $1 - \frac{2}{c}e^{-\frac{c^2}{2}}$ fraction of the d-dimensional unit has $|x_1| \leq \frac{c}{\sqrt{d-1}}$

Properties of the Unit Ball

Near Orthogonality

Theroem

Consider drawing n points $\mathbf{x_1}, \mathbf{x_2}, \ldots, \mathbf{x_n}$ at random from the unit ball. With probability $1 - \mathcal{O}(1/n)$

1. $|\mathbf{x_i}| \leq 1 - \frac{1 \ln n}{d} \quad \forall i$
2. $|\mathbf{x_i} \cdot \mathbf{x_j}| \leq \frac{\sqrt{6 \ln n}}{\sqrt{d-1}} \quad \forall i \neq j$