High Dimensional Space

Jayati Kaushik

St. Joseph's University, Bengaluru

Foundations of Data Science
BDA2121

Generating Points Uniformly at Random from a Ball

1. Generate $x_1, x_2, \ldots, x_d$ usinf a zero mean, unit variance Gaussian;i.e.
$$\frac{1}{\sqrt{2\pi}} exp(-x^2/2)$$
on the real line.
2. Then, the pdf of $\mathbf{x}$ is
$$p(\mathbf{x}) = \frac{1}{(2\pi)^{\frac{d}{2}}} exp(-\frac{x_1^2 + x_2^2 + \ldots, +x_d^2}{2})$$
3. Normalize the vector $\mathbf{x} = (x_1, x_2, \ldots, x_d)$ to a unit vector
$$\frac{\mathbf{x}}{|\mathbf{x}|}$$
4. To generate a point $\mathbf{y}$ uniformly over the ball, scale the point $\frac{\mathbf{x}}{|\mathbf{x}|}$ by a scalar $\rho \in [0, 1]$.

Gaussians in High Dimensions

The d dimensional spherical Gaussian with 0 mean and σ^2 variance in each coordinate has the density function

$$p(\mathbf{x}) = \frac{1}{2\pi^{d/2}\sigma^d} exp(-\frac{|\mathbf{x}|^2}{2\sigma^2})$$

Gaussians in High Dimension

Gaussian Annulus Theorem

For a d-dimensional spherical Gaussian with unit variance in each direction, for any $\beta \leq \sqrt{d}$, all but at most $3e^{c\beta^2}$ of the probability mass mass lies within the annulus $\sqrt{d} - \beta \leq |\mathbf{x}| \leq \sqrt{d} + \beta$ where c is a fixed positive constant.