Lecture 2 - Math 4850 - Spring 2009 - Professor John Hopcroft

Notes by Ramu Nachiappan rn54 & Grace Ross grr23.

The lecture will review topics from last class since many students thought it went too fast. The goal of this lecture is to develop an intuition for geometry in higher dimensions.

What is a hypersphere?

In 5 dimensions for example it can be represented by the formula:

$$x_1^2 + x_2^2 + x_3^2 + x_4^2 + x_5^2 \leq 1$$

It is the object in 5D Euclidean space of all points within 1 unit of the origin.

Hypercube

Volume of unit hypercube remains 1 as the dimension increases. This is shown in the integration below:

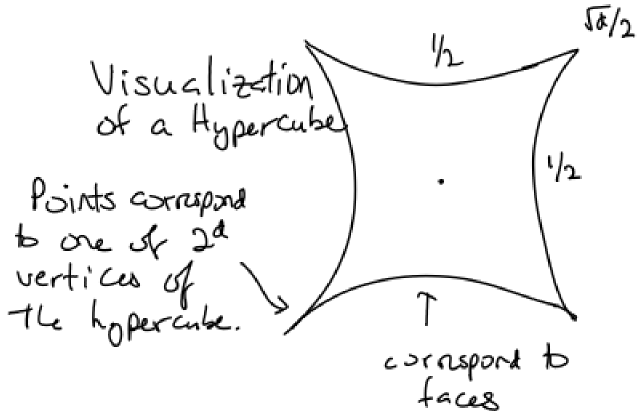
$$\int_{-.5}^{.5}\int_{-.5}^{.5}...\int_{-.5}^{.5} dx_1 dx_2 ... dx_d$$

maximum distance: $\sqrt{d}$

number of vertices: 2^d

a face of the hypercube lies in dimension $d-1$.

Two faces intersect in linear subspaces of dim $d-2$ which are correspond to edges in a cube.



Hybersphere

Volume goes to zero as d approaches infinity.

maximum distance: 1

As the maximum distance between two points in the hypersphere does not increase as the dimension increases, it makes intuitive sense that the volume goes to zero.

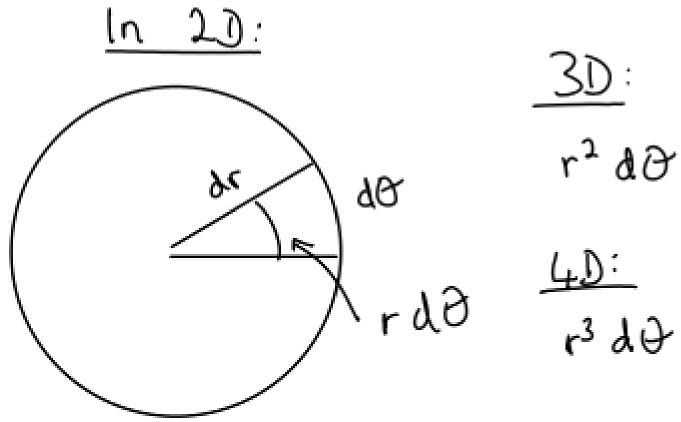
The integral for the volume of a sphere is difficult to compute in Cartesian coordinates as is clear from the formula below:

$$\int_{-1}^{1} \int_{\sqrt{1-x_1^2}}^{\sqrt{1-x_1^2}} \dots \int_{-\sqrt{1-x_1^2-x_2^2\dots-x_d^2}}^{\sqrt{1-x_1^2-x_2^2\dots-x_d^2}} dx_1 dx_2 \dots dx_d$$

So instead we can convert it to polar (look to previous lecture for the integration details):

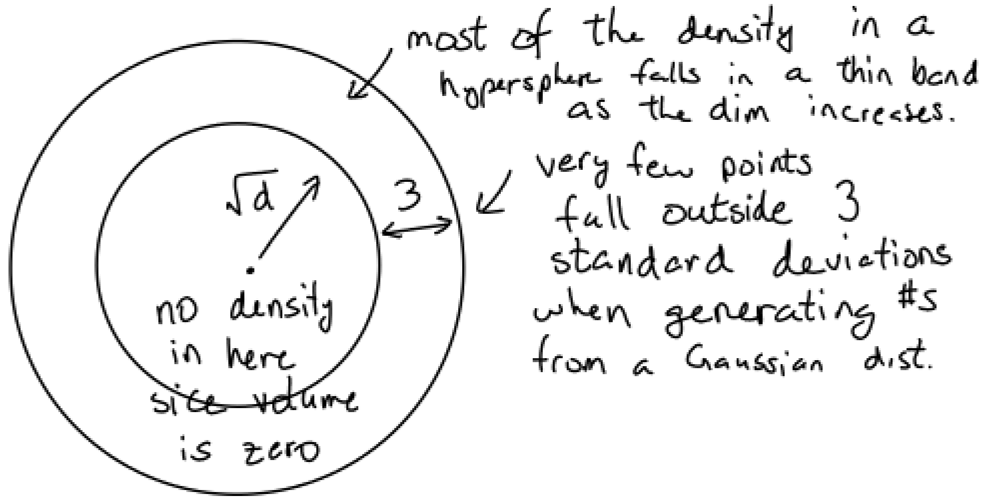
$$\int_{S_d} \int_{r=0}^{1} r^{d-1} dr \, d\Omega$$

where Ω is the interior angle.



Sampling points randomly on a hypersphere:

If we sample points in a hypersphere, we will find that most fall in a narrow band on the edge since that is where the majority of the volume lies.



In lower dimensions it is possible to generate points on the hypersphere by sampling points on the corresponding the hypercube and then accepting only the points that lie within the sphere. However in higher dimensions, most the points generated on the hypercube will lie on one of the 2^d vertices which will be outside the sphere. Almost none of the samples in the hypercube will fall within the sphere of interest.

Formula for equator of a hypersphere:

We first pick one point to be the north pole of the sphere arbitrarily with the coordinates:

$$p = (1, 0, 0...0, 0)$$

Then the equator for the hypersphere can be determined. It will have the formula

$$x_1 = 0,\ x_2^2 + x_3^2 + ... + x_d^2 = 1$$

Generating Random Points on a sphere:

To generate random points on the surface of a sphere, generate d points independently from a Gaussian and then normalize:

$$\frac{1}{\sqrt{x_1^2 + x_2^2 + ... + x_d^2}}\,(x_1, x_2, x_3, ...x_d)$$

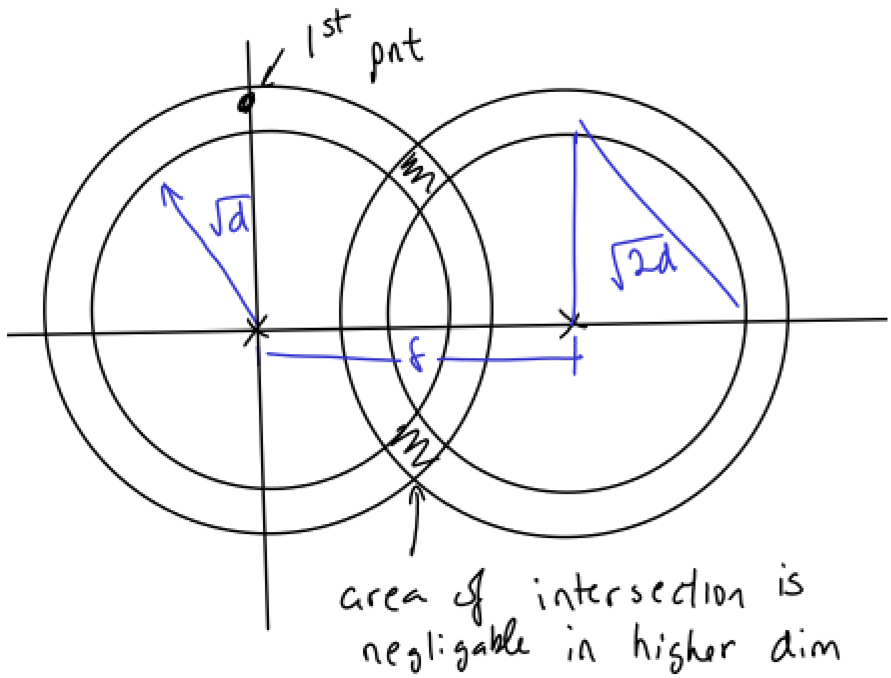
We can compute the square of the distance between two random points on the sphere by the formula below:

$$\begin{aligned}
& (1 - \frac{x_1}{\sqrt{x_1^2 + ... + x_d^2}})^2 + \frac{x_2^2}{x_1^2 + ... + x_d^2} + \frac{x_3^2}{x_1^2 + ... + x_d^2} + ... + \frac{x_d^2}{x_1^2 + ... + x_d^2} \\
= \quad & 2 - \frac{2x_1}{\sqrt{x_1^2 + ... + x_d^2}} \\
= \quad & 2 - \frac{1}{\sqrt{d}}
\end{aligned}$$

We see that as the dimensions increases, $\frac{1}{\sqrt{d}}$ becomes insignificant and the average distance between two random points on the sphere approaches $\sqrt{2}$.

An application

Suppose we have two random processes generating points. We want to know which process generated each point.



$\sqrt{d}$= radius

$\sqrt{2d}$=diagonal of right triangle

$\sqrt{2d}$ is the distance between two points on the same annulus

$\sqrt{d^2 + 2d}$ is the distance between points generated by different processes

and $\sqrt{d^2 + 2d} > \sqrt{2d} + \alpha$