High-Dimensional Data Visualisation

By: Inon Sinn - 209632009

1. Introduction

This is an Incomplete Project, But i will still go through my work and what i found currently

The Goal of this Project is test a few Theories that i had about high Dimensional Data Projection wheter it is correct or not, I intend to check in an empircal way.

1. The Volume of High Dimensional Geometric Shapes gather at the corners creating spikes
2. If we where to use an Image dataset and take them as Hypercube data and project to 2D, we could see cluster for images with same theme/animal/person.

2. Code Overview

The Code is in jupyter notebook, and is divided the following way:

I. **Hyperdimensional Shapes:** A collection of function for sampling random data of high dimensional shapes such as Balls, Spheres, Cubes and Simplexes.

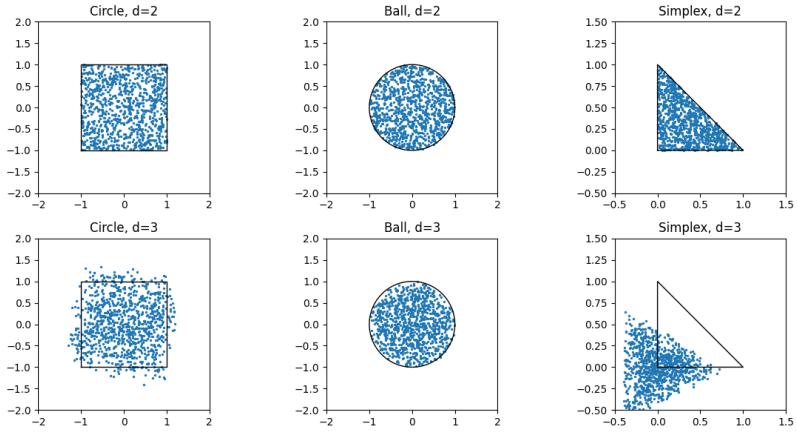
II. **Projection Methods:** All the projection Methods used in the Code, currently only Random Projection and PCA. Both come from the scikit-learn Library

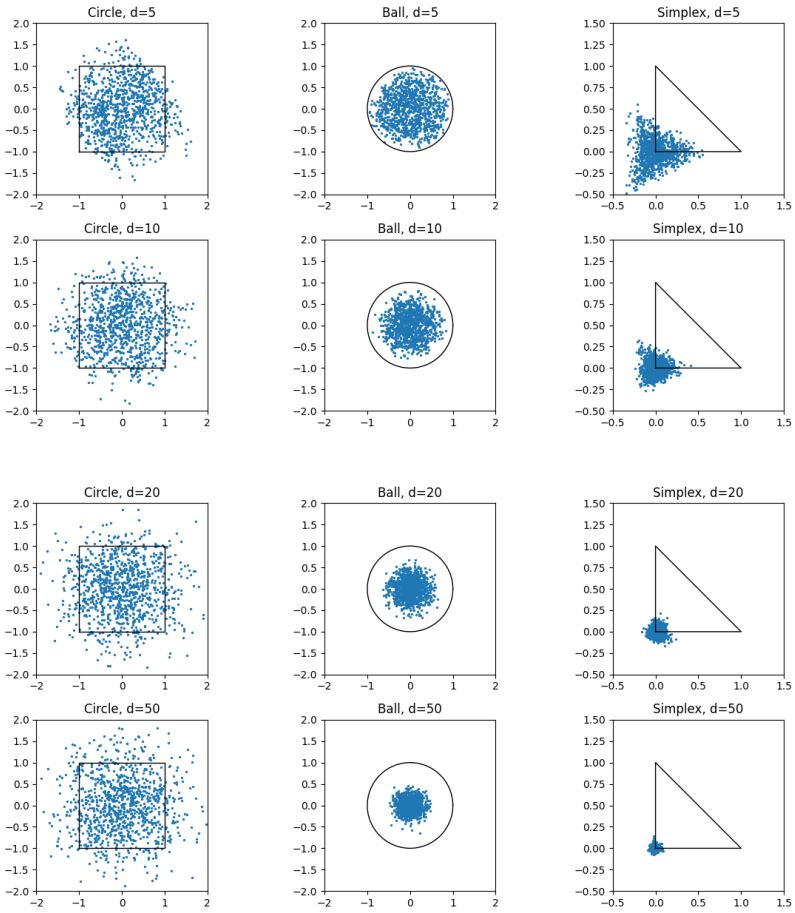
III. **Plotting Functions:** Methods to plot 2D and 3D points, 2D and 3D outlines for all hyperdimensional Shapes and lastly a function to create subplots for comparison of different projections

3. Projections of Volume of High Dimensional Geometric Shapes

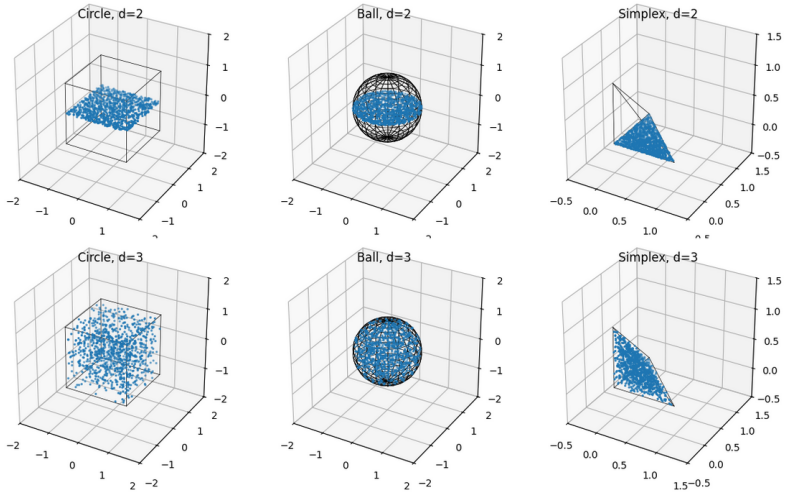
These are the results of testing the First theory, Here i used PCA as the projection methods, for 1000 sampled points.

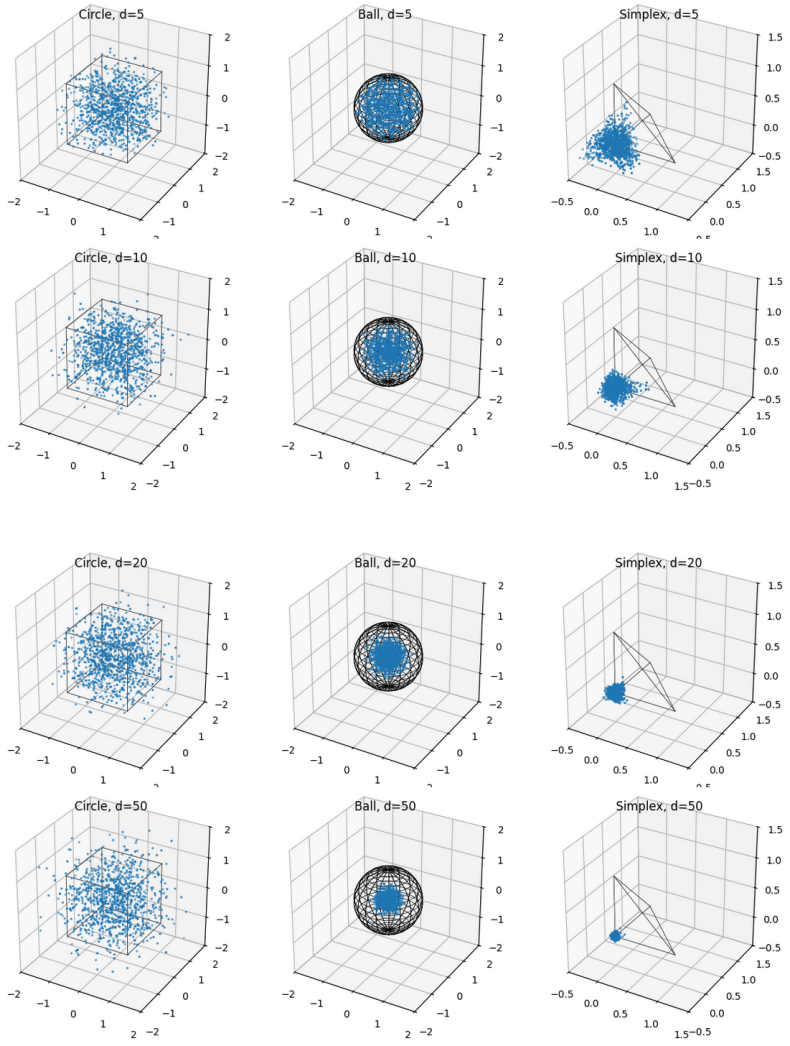
3.1. 2D Projection





3.2. 3D Projection





3.3. Results

The result while interesting is very different from what I expected to get, In additoin there are huge changes arcoding to which projection method is chosen. One the explanation that I have read is Because its a linear combination and all points are sampled iid, then because of the Central Limit Theorem we get the that projection would like a gaussian. Now I need to research the methods more before I can explain why exatly it looks like this.

3.4. Final Word

The Goal was to see the spikes in Hypercube and the Shell in the ball. But during my search on the web i have not found one real projection to do that. So either it is not possible with our projection techniques or its just much more complicated to do. Most of the things i found are graphs of the Volume that show the values of the Volume formula and why for example in a circle it goes to 0, but these are not what I wanted to do!

4. References

- Strohmer, T. (n.d.). Surprises in high dimensions (Lecture 1) [PDF]. Department of Mathematics, University of California, Davis. Retrieved from https://www.math.ucdavis.edu/~strohmer/courses/180BigData/180lecture1.pdf