

Homework assignment 2 due Friday February 12

To expedite grading please submit each problem on a separate sheet of paper.

1. We have claimed that a randomly generated point lies on the equator of the sphere independent of where we pick the North Pole. To test this claim randomly generate ten vectors in 128 dimensions whose coordinates have value ± 1 . Think of these ten vectors as ten choices for the North Pole. Then generate some additional random vectors with ± 1 coordinates. For each of the new vectors determine how many of the original vectors they are close to being perpendicular to. That is, they lie close to the equator.
2. List the five most important things you learned about high dimensions.
3. Look on the web for a data base that can be converted to an undirected graph. For example, in Science there is a data base of proteins and their interactions. Each protein can be represented by a vertex and two proteins that interact are connected with an edge. Find a data set that will yield a graph with at least 1000 vertices and a number of edges that is roughly in the range or one to two times the number of vertices. Thus, the graph will be quite sparse and have many connected components. Find all connected components and create a table with the number of components of each size. For example, 312 isolated vertices. 124 components consisting of two vertices connected by an edge, 29 components with three vertices, etc.
4. Generate a 1000 vertex graph adding edges randomly one at a time. How many edges are added before all isolated vertices disappear? Try the experiment enough times to determine how concentrated the number of edges is. To do the experiment, you do not need to actually generate the graph. Generate pairs of vertices at random and when a vertex is generated mark that vertex as not isolated.
5. Generate $G(1000, 1/2)$ and find the largest clique you can. A clique is a complete sub graph, that is, a set of vertices each pair of which is connected by an edge.