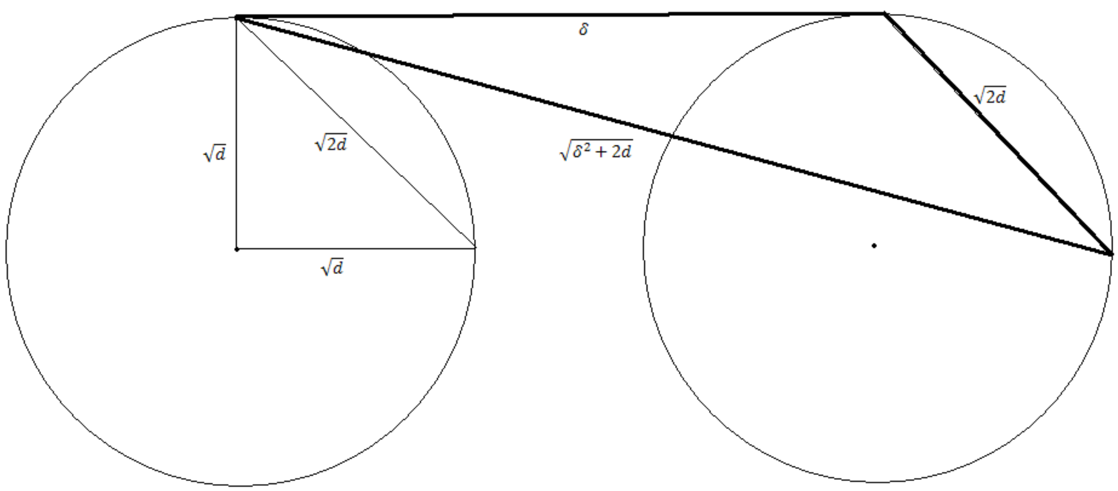
Lecture 3 – Math 4850

Notes by Eric Frackleton (epf2) and Dan Goldfein (dg254)

Gaussian Processes: Process generates a coordinate in d dimensions corrupted by random noise with Gaussian distribution

The density of the distribution is found by integrating over spheres around the center coordinate. This remains 0 until we hit a radius of $\sqrt{d}$.



If have 2 points, get either a distance of $\sqrt{2d}$ if on same sphere or $\sqrt{\delta^2 + 2d}$ if on different sphere. Here we assume that the bolded triangle is a right triangle in d-dimensional space.

If we have $\sqrt{\delta^2 + 2d} > \sqrt{2d} + \alpha$ for appropriate approximation term α, we can differentiate between the two spheres.

We want to choose δ such that $\sqrt{2d}\left(1 + \frac{\delta^2}{2d}\right)^{1/2} = \sqrt{2d}\left(1 + \frac{\delta^2}{4d} + \cdots\right) \geq \sqrt{2d} + \alpha$. Therefore take $\sqrt{2d}\left(\frac{\delta^2}{4d}\right) \geq \alpha$ so $\left(\frac{\delta^2}{2\sqrt{2d}}\right) \geq \alpha$. Hence choose δ such that $\delta^2 \geq 2\sqrt{2d}\alpha$, a constant based on α.

Suppose now we have k Gaussian processes instead of just 2. Also suppose we know the centers of the Gaussians. Then we can put a k-dimensional plane through the centers and project every point onto it. All other dimensions are noise, so projection throws away noise. Since centers are on the plane, they stay the same distance apart. Hence the signals are as strong with less noise and reduce to k dimensions.

Is it safe to assume the centers are known?

Single Value Decomposition (To be covered later) – Basically, reduce a matrix given by the points to a rank k matrix defining the k-dimensional space through the centers of the points.

Random Graphs:

Consider the worldwide web. Each webpage is a node on a graph, with a directed edge between node A and node B if page A links to page B. There are billions of nodes in this graph, so any property of the graph is probably the same if a few nodes or a few edges are deleted.

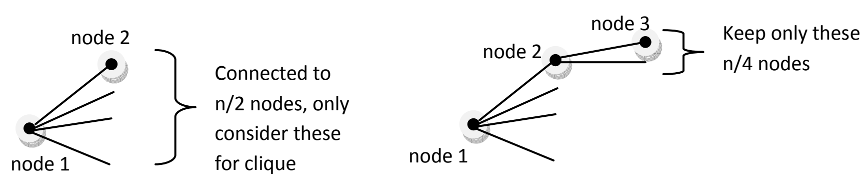
Definition:

G(n,p) is a graph with *n* nodes, then for each pair of nodes, there is an edge connecting the nodes with probability *p*.

To create *G(1000,1/2)*: create 1000 nodes, then for each pair of nodes, flip a coin and create or leave out an edge.

G(n,1/2): expected degree (number of outgoing edges for a node) is $\frac{n}{2}$.

We'll show that in this graph, there is a clique (a fully connected subgraph) of size *log n*.



Here we have in expectation n/2 nodes connected to node 1 and remove the rest from the set of nodes to consider. Then there are n/2 nodes in expectation connected to node 2, but half of these (n/4) are expected to have been removed already for not being connected to node 1. Repeating this process, at each step the node has half as many choices as the previous one in expectation to choose from.

We can perform *log n* of these steps, and clearly every node in the subgraph is connected to every other node. Thus we have a clique of size *log n*.

With high probability, for any $\varepsilon > 0$, there is a clique in *G(n,1/2)* of size $(2-\varepsilon)$*log n*. However, there is almost surely not a clique of size 2*log n*. However, there is no known polynomial time algorithm for finding such cliques.

Consider a less dense graph, $G(n,\frac{d}{n})$ where *d* is a constant. *d* ends up being the expected degree.

Problems with degree distribution:

Given *G(n,p)*, what is the probability that a given vertex has degree *k*?

prob(k) = $\binom{n-1}{k} p^k (1-p)^{n-1-k} \approx \binom{n}{k} p^k (1-p)^{n-k}$ for large n. This is the binomial distribution.

If $p = \frac{1}{2}$, then $prob(k) = \binom{n}{k}\left(\frac{1}{2}\right)^n$. $\sum_{k=0}^{n}\binom{n}{k} = 2^n$ so the $\left(\frac{1}{2}\right)^n$ is a normalizing constant.

Expected value: $E(k) = \frac{n}{2}$ variance: $\sigma^2 = \frac{n}{4}$

In general, $\sigma^2 = \int_{-\infty}^{\infty}\left(x - E(x)\right)^2 p(x)dx$ and standard deviation is square root of variance. In a Gaussian, almost all values are within 3 standard deviations of the mean.

Degrees in G(n, ½) are $\frac{n}{2} \pm const\sqrt{n}$

For *G(n,1/n)*:

$E(d) = 1$ however, highest degree you should find of a vertex is $\frac{\log n}{\log\log n}$

For large n with $k = o(n)$:

$$P(k) = \binom{n}{k}\left(\frac{1}{n}\right)^k\left(1-\frac{1}{n}\right)^{n-k} \approx \frac{n(n-1)\ldots(n-k+1)}{k!}\frac{1}{n^k}\left(1-\frac{1}{n}\right)^n \approx \frac{n^k}{k!}\frac{1}{n^k}e^{-1} = \frac{1}{ek!}$$

Note (1-c)^n for constant $0 < c < 1$ has $\lim_{n\to\infty}(1-c)^n = 0$, but $0 < 1/n < 1$ is not constant so doesn't need to have limit go to 0.

Now, if we let $k = \frac{\log n}{\log\log n}$, $\log k^k = k\log k = \frac{\log n}{\log\log n}(\log\log n - \log\log\log n) \approx \log n$

So, $k^k \approx n$ and $k! \leq k^k \approx n$