Johnson –Lindenstrauss $Cont^d$

$$(1-\varepsilon)\left\|f(u)-f(v))\right\| \leq \|u-v\| \leq (1+\varepsilon)\left\|f(u)-f(v))\right\| \qquad (\star)$$

$$k \sim \frac{1}{\varepsilon^2} log(n)$$

$(\star)$ holds with probability if we choose

$$f(u) = \frac{1}{\sqrt{k}} Au$$

where A is a k x d Gaussian

1. Fast – JL Transforms (FJLT)

When the map f is realized via a Gaussian random matrix, the cost to evaluate $u \mapsto f(u)$ is $O(k \times d)$

Ailon & Chazelle proposed the projection

$$f(u) = PHDu$$

where

$D_{d\times d}$ is Diagonal
$H_{d\times d}$ is a Hessenberg Transform
$P_{k\times d}$ is a "sort of" subsampling

$f(u)$ can be evaluated in $O(dlog(d))$ time.

Ailon & Chazelle proved that the map can approximately preserve distance with high probability.

Later, the subsampled FFT was proposed

$$f((u) = SFD$$

where

F is the discrete Fourier Transform
D is Diagonal and $D(i,j) = e^{i\Theta j}$ with $\Theta j \epsilon \cup [0, 2\Pi]$

The cost to evaluate f is $O(mnlogk) \rightarrow$ reduced

Recall: We previously used SRFT to accelerate the randomized SVD from $O(mnk)$ to $O(mnlogk)$ for rank-k approximation of an $m \times n$ matrix

Recent work includes looking for sparse maps or matrices with integer entries.

Question: Can we generalize the J–L Theorem to metric spaces?

2. Bourgain Embedding

The idea of embedding a set Q of n points in Eculidean space $\mathbb{R}^d$ to $\mathbb{R}^k$ (for "small" k) while approximately preserving distances can be generalized to METRIC spaces.

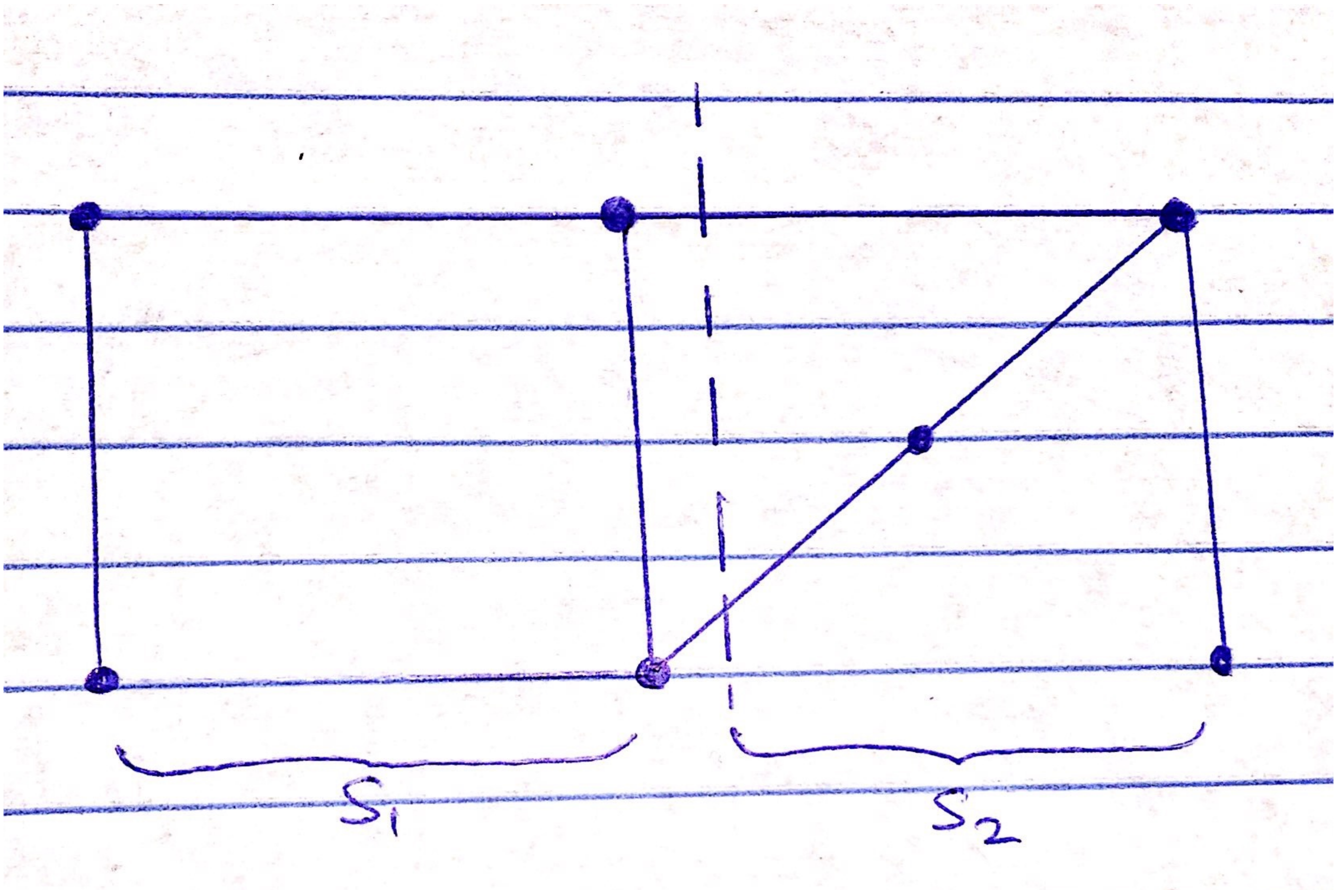
Theorem (Bourgain Embedding): Let Q be a set of n points and let $d\ Q \times Q \rightarrow [0,\infty)$ be a metric. Then, $\exists$ a map $f : Q \rightarrow \mathbb{R}^k$ for some $k = O((logn)^2)$ such that

$$\|f(u)-f(v)\|_{l^1} < d(u,v) \leqslant \|f(u)-f(v)\|_{l^1} .c.log(n)$$

where c is a universal constant.

A map f that realizes the bound can "with high probability" be built in polynomial time n.

NOTE: The "sparsest–cut" problem for a graph V, E is to find a partition $V = S1 \cup S2$ such that $\frac{|E(S_1,S_2)|}{|S_1||S_2|}$ is minimized.



The sparsest–cut minimization problem can be expressed in terms of certain matrices on V. The Bourgain embedding techniques are useful in solving the optimization problem efficiently ("In probability" since the sparsest–cut problems are NP hard).

3. Connection to Central Limit Theorem

Theorem(Central Limit: Let $\{\underline{X}_i\}_{i=1}^n$ be a set of i.i.d random variables with mean μ and variance σ^2

$$S_k = \frac{1}{k}\sum_{j=1}^n \underline{X}_j \text{ (average)}$$

As K increases, the distribution of S_k will approach a normal distribution with mean μ and variance $\frac{1}{k}\sigma^2$

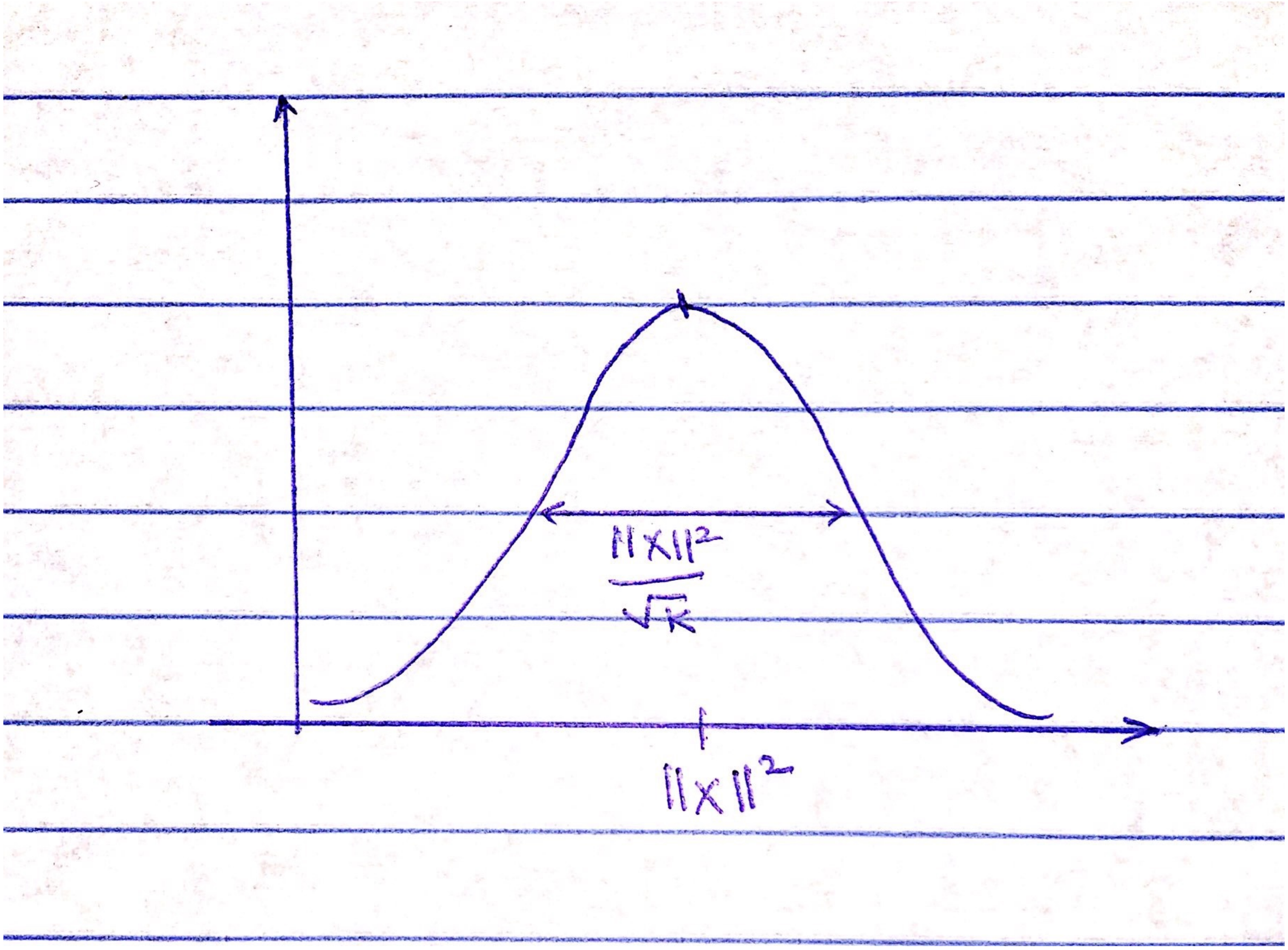
$k \times d$ Gaussian matrix and $y = \frac{1}{\sqrt{k}}Ax$

Then

$$\|y\|^2 = \frac{1}{k}\|g\|^2 \|x\|^2$$

where $\|g\|^2$ has a χ_k^2 distribution.

So, the distribution of $\|y\|^2$ looks like.



As k grows, the variance in $\|y\|^2$ shrinks, but pretty slowly.

This result is similar to classical *MONTE–CARLO*, where the expected errors shrink as $\frac{1}{\sqrt{k}}$ where $k = \#$ samples