Plan

Logistics

Review

Sketched Regression

Problem

due 5pm Friday

Project

Main focus

Spectral Graph Theory

$G = (V, E) \qquad |V| = n, \ |E| = m$

Adjacency $\quad A \in \mathbb{R}^{n \times n}$

Degree $\quad D \in \mathbb{R}^{n \times n}$

Laplacian $\quad L = D - A$

Edge-incidence $\quad B \in \mathbb{R}^{m \times n}$

$$b_{(i,j)} = \begin{array}{cccccc} & & j & & & i \\ \lbrack\, 0 & 0 & -1 & 0 & 0 & 1 \,\rbrack \end{array}$$

$$\underset{n \times n}{L} = \underset{n \times m}{B^T} \underset{m \times n}{B}$$

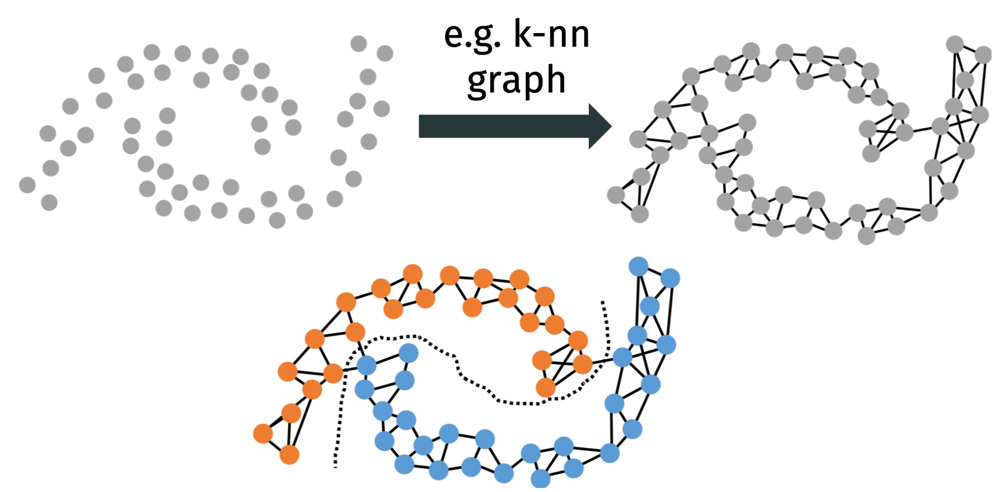
$$x^T L x = x^T B^T B x$$

$$= \|Bx\|_2^2 = \sum_{(i,j) \in E} (x_i - x_j)^2 \geq 0$$

$$v_n = \underset{\|v\|_2 = 1}{\operatorname{argmin}} \ v^T L v$$

$$v_{n-1} = \underset{\|v\|_2 = 1, \ \langle v_n, v \rangle = 0}{\operatorname{argmin}} \ v^T L v$$

Clustering



S, S^c

$c \in \{-1, 1\}^n$

$c_i = 1$ iff $i \in S$

$$c^T L c = \sum_{(i,j) \in E} (c_i - c_j)^2$$

$$= 4\,\mathrm{cut}(S, S^c)$$

Balanced Cut

$$\min c^T L c \quad \text{s.t. } c^T \mathbb{1} = 0$$

$c \in \{-1/\sqrt{n}, 1/\sqrt{n}\}^n$

Relaxed Balanced Cut

$$\min_{\|c\|_2 = 1} c^T L c \quad \text{s.t. } c^T \mathbb{1} = 0$$

$\iff$

$$v_{n-1} = \underset{\|v\|_2 = 1,\ \langle v_n, v \rangle = 0}{\operatorname{argmin}} v^T L v$$

1. Compute v_{n-1}
2. Round v_{n-1} to get partition S, S^c

Regression

$A \in \mathbb{R}^{n \times d}$ $\quad b \in \mathbb{R}^{n}$

$$x^* = \operatorname*{argmin}_{x \in \mathbb{R}^d} \|Ax - b\|_2^2$$

Problem:

↳ Show $x^* = (A^TA)^{-1}A^Tb$

↳ Computing takes $O(nd^2)$

Faster?

Sketched Regression

$$\tilde{x} = \operatorname*{argmin}_{x \in \mathbb{R}^d} \|\Pi A x - \Pi b\|_2^2$$

$\Pi \in \mathbb{R}^{m \times n}$ $\quad m \ll n$

$$\underbrace{\begin{bmatrix} \\ \\ \\ \end{bmatrix}}_{A}^{\,n \times d} \underbrace{\begin{bmatrix} \\ \end{bmatrix}}_{x}^{\,d} = \underbrace{\begin{bmatrix} \\ \\ \\ \end{bmatrix}}_{b}^{\,n}$$

$O(nd^2)$

vs

$$\underbrace{\begin{bmatrix} \\ \end{bmatrix}}_{\substack{\Pi A \\ m \times d}} \underbrace{\begin{bmatrix} \\ \end{bmatrix}}_{\substack{x \\ d \times 1}} = \underbrace{\begin{bmatrix} \\ \end{bmatrix}}_{\substack{\Pi b \\ m \times 1}}$$

$O(md^2)$

$m = O(d/\epsilon^2)$

Theorem 1 : $\|A\tilde{x} - b\|_2^2 \leq (1+\epsilon)\|Ax^* - b\|_2^2$ wp $9/10$

Claim: $(1-\epsilon)\|Ax - b\|_2^2 \overset{1}{\leq} \|\Pi Ax - \Pi b\|_2^2 \overset{2}{\leq} (1+\epsilon)\|Ax - b\|_2^2$

$\forall x$ wp $9/10$

Proof of theorem 1:

$$\|A\tilde{x} - b\|_2^2 \overset{1}{\leq} \frac{1}{1-\epsilon}\|\Pi A\tilde{x} - \Pi b\|_2^2$$

$$\leq \frac{1}{1-\epsilon}\|\Pi Ax^* - \Pi b\|_2^2 \text{ by optimality of } \tilde{x}$$

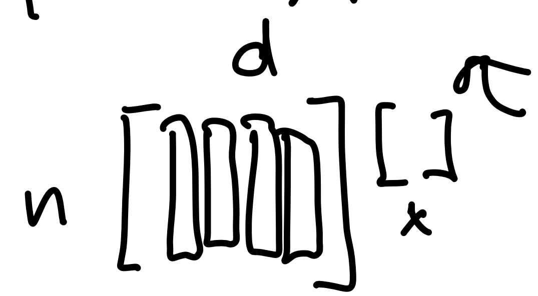
$$\overset{2}{\leq} \frac{(1+\epsilon)}{(1-\epsilon)}\|Ax^* - b\|_2^2$$

Distributional JL

$m = O\left(\frac{\log 1/\delta}{\epsilon^2}\right)$

$(1-\epsilon)\|y\|_2^2 \le \|\Pi y\|_2^2 \le (1+\epsilon)\|y\|_2^2$ wp $1-\delta$ for fixed y

We want it to hold for $Ax-b$ for all $x \in \mathbb{R}^d$



$\leftarrow$ d dimensional

Goal

Subspace Embedding Theorem

$m = O\left(\frac{d\log(1/\epsilon) + \log(1/\delta)}{\epsilon^2}\right)$ $\qquad U \subset \mathbb{R}^n$ is d-dimensionsal

$(1-\epsilon)\|y\|_2^2 \le \|\Pi y\|_2^2 \le (1+\epsilon)\|y\|_2^2$ for all $y \in U$ wp $1-\delta$

Prove Subspace Embedding

w on unit sphere in $\mathcal{U}$

$$(1-\epsilon) \|w\|_2 \leq \|\Pi w\|_2 \leq (1+\epsilon) \|w\|_2$$

$y = c \cdot w \qquad y \in \mathcal{U}$

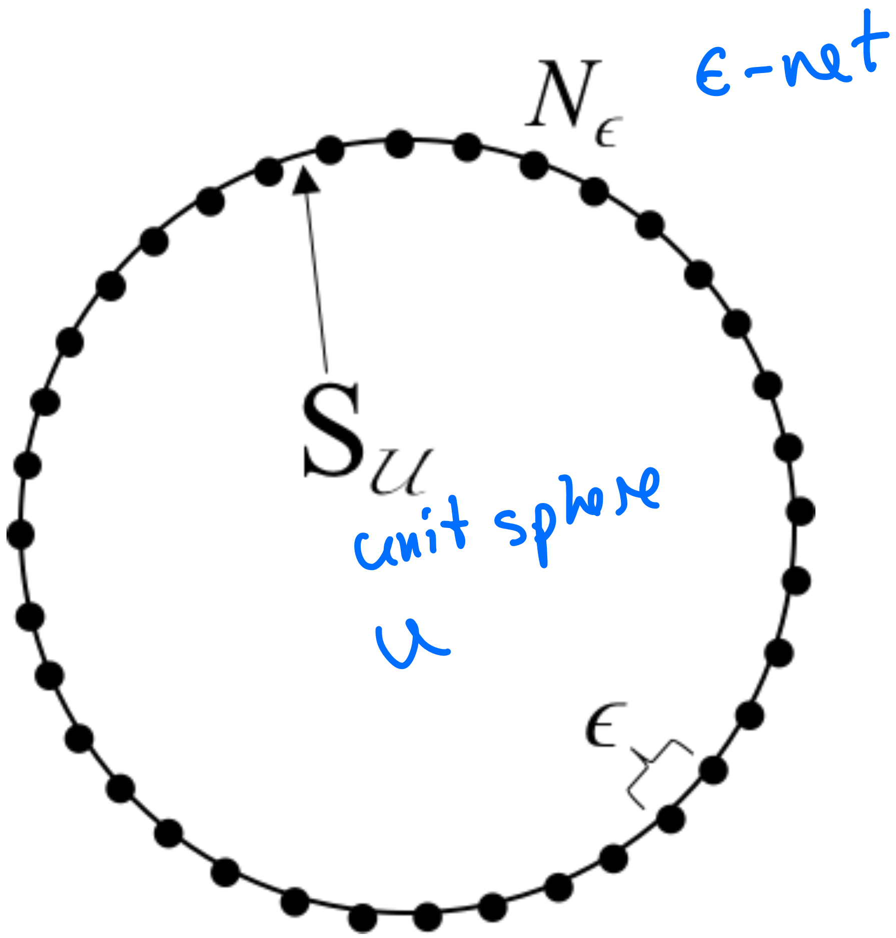
$$(1-\epsilon)\, c \|w\|_2 \leq c \|\Pi w\|_2 \leq (1+\epsilon)\, c \|w\|_2$$

$$(1-\epsilon) \|c w\|_2 \leq \|c \Pi w\|_2 \leq (1+\epsilon) \|c w\|_2$$

$$(1-\epsilon) \|y\|_2 \leq \|\Pi y\|_2 \leq (1+\epsilon) \|y\|_2$$

$$c\|w\|_2 = c \cdot \sqrt{\sum_{i=1}^{d} w_i^2} = \sqrt{c^2 \sum_{i=1}^{d} w_i^2} = \sqrt{\sum_{i=1}^{d} (c w_i)^2} = \sqrt{\|c w\|_2^2} = \|c w\|_2$$

Not "too" many different points on sphere



Goal:

1. $(1-\epsilon)\|w\|_2 \le \|\Pi w\|_2 \le (1+\epsilon)\|w\|_2$

   for all $w \in N_\epsilon$

2. for all $v \in S_U$

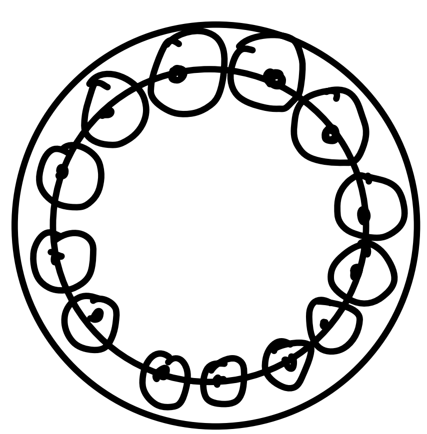
   $$\min_{w \in N_\epsilon} \|v - w\|_2 \le \epsilon$$

3. $|N_\epsilon| \le \left(\frac{3}{\epsilon}\right)^d$

Construct $\mathcal{N}_\epsilon$

$\mathcal{N}_\epsilon = \{\}$

while point in S_u that is more than ϵ away from everything in $\mathcal{N}_\epsilon$:
add point $\mathcal{N}_\epsilon$



So each point in $\mathcal{N}_\epsilon$ has ball radius $\epsilon/2$ and no balls overlap

$$\mathrm{Vol}(d, r) = c_d r^d$$

volume of radius r ball in d-dimension

balls: vol(balls) $\leq$ vol(larger sphere)

$$|\mathcal{N}_\epsilon| \, c_d \left(\frac{\epsilon}{2}\right)^d \leq \left(1+\frac{\epsilon}{2}\right)^d \cdot c_d$$

$$|\mathcal{N}_\epsilon| \leq \frac{\left(1+\frac{\epsilon}{2}\right)^d}{(\epsilon/2)^d} = \left(\frac{1}{\epsilon/2} + \frac{\epsilon/2}{\epsilon/2}\right)^d$$

$$\leq \left(\frac{3}{\epsilon}\right)^2$$

Set $\delta' = \frac{1}{|N_\epsilon|} \cdot \delta \approx \left(\frac{\epsilon}{3}\right)^d \delta$

$$\log 1/\delta' = \log \left(\frac{3}{\epsilon}\right)^d + \log 1/\delta$$
$$= d \log (3/\epsilon) + \log (1/\delta)$$

$$\underbrace{m}_{\text{from distributional JL}} = O\left(\frac{\log (1/\delta')}{\epsilon^2}\right) = O\left(\frac{d \log (1/\epsilon) + \log(1/\delta)}{\epsilon^2}\right)$$

$(1-\epsilon)\|w\|_2 \leq \|\Pi w\|_2 \leq (1+\epsilon)\|w\|_2$ for all $w \in N_\epsilon$ w.p. $1-\delta$

What about $v \in S_u$ but not in N_ϵ?

$$v \in S_u \setminus N_\epsilon$$

$$w_0 = \underset{w \in N_\epsilon}{\operatorname{argmin}} \|v - w\|_2 \qquad r_0 = v - w_0 \qquad c_1 = \|r_0\|_2$$

$$w_1 = \underset{w \in N_\epsilon}{\operatorname{argmin}} \left\|\frac{r_0}{c_1} - w\right\|_2 \qquad r_1 = v - w_0 - c_1 w_1 \qquad c_2 = \|r_1\|_2$$

$$w_2 = \underset{w \in N_\epsilon}{\operatorname{argmin}} \left\|\frac{r_1}{c_2} - w\right\|_2 \qquad r_2 = v - w_0 - c_1 w_1 - c_2 w_2 \qquad c_3 = \|r_2\|_2$$

$$\vdots$$

Induction: $\|r_i\|_2 \leq \epsilon^i$, $\left\|\frac{r_{i-1}}{c_i} - w_i\right\|_2 \leq \epsilon$ by N_ϵ

$$\|r_i\|_2 = \|r_{i-1} - c_i w_i\|_2 \leq \epsilon \cdot c_i = \epsilon \|r_{i-1}\|_2 \leq \epsilon \cdot \epsilon^{i-1} = \epsilon^i$$

$$\|\Pi v\|_2 \overset{\text{want}}{\leq} (1+\epsilon)\|v\|_2 = 1+\epsilon$$

$$= \|\Pi (w_0 + c_1 w_1 + c_2 w_2 + \dots)\|_2$$

triangle inequality

$$\leq \|\Pi w_0\|_2 + \|\Pi c_1 w_1\|_2 + \|\Pi c_2 w_2\|_2 + \dots$$

$$\leq (1+\epsilon)\|w_0\|_2 + c_1(1+\epsilon)\|w_1\|_2 + c_2(1+\epsilon)\|w_2\|_2 + \dots$$

$$\leq (1+\epsilon)\cdot 1 + \epsilon(1+\epsilon) + \epsilon^2(1+\epsilon) + \dots$$

$$= 1 + 2\epsilon + 2\epsilon^2 + 2\epsilon^3 + \dots$$

$$= 1 + \frac{2\epsilon}{1-\epsilon} \leq 1+4\epsilon \quad \text{for } 0<\epsilon<1/2$$

triangle inequality

$$\|a+b\|_2 \leq \|a\|_2 + \|b\|_2$$