

Expander Graphs in Algorithms Theory and in Data Applications

Lecturer: Michael Mahoney

Scribes: Erik Goldman and Richa Bhayani

**Unedited Notes*

1 Expanders

1.1 Introduction

Intuition: expanders reduce the need for randomness.

Definition: (S, d) is a metric space if $d: S \times S \rightarrow \mathbb{R}^+$, $d(x, y) \geq 0$, $d(x, y) = 0$ iff $x = y$, $d(x, y) = d(y, x)$ and $d(x, y) \leq d(x, z) + d(z, y)$.

Distances: can be a Gram matrix, kernel, allowing algorithms in an infinite-dimensional space.

Can we replace d with something "nicer" while preserving distances? If so, use the function d' that's faster but introduces a little error.

Given $(X, d) = (\mathbb{R}^n, L_2)$ and a mapping $f: X \rightarrow \mathbb{R}$:

$$\text{expansion}(f) = \max_{x_1, x_2 \in X} \frac{\|f(x_1) - f(x_2)\|_2}{d(x_1, x_2)}$$

$$\text{contraction}(f) = \max \frac{d(x_1, x_2)}{\|f(x_1) - f(x_2)\|}$$

$$\text{distortion}(f) = \text{expansion}(f) * \text{contraction}(f)$$

NB: Johnson-Lindenstrauss says that we can map $x_i \rightarrow f(x)$ such that distance is within $1 \pm \epsilon$ of the original.

Theorem: every n -point metric space can be embedded in L_2 with distortion $\leq O(\log n)$. Proof idea: given (X, d) , map each point $x \rightarrow \phi(x)$ in $O(\log^2 n)$ with coordinates equal to the distance to $S \subseteq V$ where S is chosen randomly.

Q1: Is this tight?

Q2: $C_2(X, d)$ is d'

Q3: Are there metrics that are better? (yes)

Theorem [LLR]: \exists a polynomial time algorithms that, given a metric space (X, d) , will compute (C_2, d) with C_2 as the minimum distortion of the space.

Proof: $|X| = n$, without loss of generality, scale f so $\text{contraction}(f) = 1$. $\text{distortion}(x) \leq \gamma$ iff $d(x_1, x_2) \leq \|f(x_1) - f(x_2)\|_2^2 = \gamma^2 d(x_1, x_2)^2$.

Let $u_i = f(x_i) = i^{\text{th}}$ row of embedding matrix. Let $Z = uu^T$, Z is positive definite.

$$\text{note: } \|f(x_i) - f(x_j)\|_2^2 = (u_i - u_j)^T (u_i - u_j) = u_i^T u_i + u_j^T u_j - 2u_i^T u_j = Z_{ii} + Z_{jj} - 2Z_{ij}$$

So I can look for a PSD matrix Z s.t. $d(x_i, x_j) \leq Z_{ii} + Z_{jj} - 2Z_{ij}$, so $C_2(X, d) \leq \gamma$ iff $\exists Z \in \text{PSD}$ s.t. the previous inequality holds.

This can be solved as an optimization problem!

1.2 Dual

Note: By looking at the dual, we get an idea about the lower bound. (This will come up again in graph partitioning). we've to look at the constraints again to go from the primal to the dual. Here, $z \in PSD$ is the constraint causing a problem.

1.3 Lemma

$z \in PSD < - >$ For all PSD , Q , $\sum q_{ij}z_{ij} \geq 0$
Primal:

- I. $\sum q_{ij}z_{ij} \geq 0$ for all $Q \in PSD$
- II. $z_{ii} + z_{jj} - 2z_{ij} \geq d(x_i, x_j)^2$
- III. $\gamma^2 d(x_i, x_j)^2 \geq z_{ii} + z_{jj} - 2z_{ij}$

1.4 Theorem:(LLR contd.)

$$C_2(x, d) = \max_{(P \in PSD, P.1=0)} \sqrt{(\sum_{P_{ij}>0} P_{ij}d(x_i, x_j)^2)/(-(\sum_{(P_{ij}<0)} P_{ij}d(x_i, x_j)^2))}$$

Proof: Assume $\gamma < C_2(x, d)$. Look for a linear combination of constraints such that, $Q, Z = \sum q_{ij}z_{ij} \geq 0$
Note the cone of PSD matrices is convex. $\sum k\alpha_k G.k.z = P.z$ for some P . So, modifying first constraint from the primal, you get

- I'. $\sum P_{ij}z_{ij} \geq 0$ for some $P \in PSD$
Construct P. Choose the elements such that you zero out z_{ij} . i.e. if $P_{ij} > 0$ multiply second constraint from primal by $P_{ij}/2$, if $P_{ij} < 0$ multiply second constraint from primal by $-P_{ij}/2$
Therefore , you can modify the other constraints from primal to be :
- II'. $\sum_{i,j, P_{ij}>0} P_{ij}/2(z_{ii} + z_{jj} - 2z_{ij}) \geq P_{ij}/2d(x_i, x_j)^2$
- III'. $\sum_{i,j, P_{ij}<0} P_{ij}/2(z_{ii} + z_{jj} - 2z_{ij}) \leq P_{ij}/2 \sum \gamma^2(d(x_i, x_j)^2)$

Adding the second and third constraints above,

$$0 \geq \sum_{i,j, P_{ij}>0} P_{ij}d(x_i, x_j)^2 + \sum_{i,j, P_{ij}<0} P_{ij} \sum \gamma^2(d(x_i, x_j)^2)$$

This will be false if you choose γ to be small . Will be false if $\gamma^2 \leq top/bottom$

2 Expanders: Some Q&A

2.1 Why is this bidirectional?

If its not, then in one of the assumption we make(not specified which one) you get a value greater than zero instead of lesser.

2.2 Why Expanders?

They are extremal. So, if your algorithm works for expanders, hypercubes and low dimensional graphs, you're good.

2.3 Can I draw a nice picture of an expander?

No. Because there are no good cuts for an expander.

3 Expanders: Some definitions

From now on, unless stated otherwise, a graph $G=(V,E)$ is undirected and d -regular. (all vertices have the same degree d).

For $S, T \subset V$, where V is the vertex set, you can denote the set of edges from S to T as:

$$E(S, T) = \{(u, v) | u \in S, v \in T, (u, v) \in E\}$$

The edge expansion ratio of G , denoted $h(G)$ is defined as:

$$h(G) = \min_{(S:|S|<n/2)} E(S, S')/|S|$$

Extensions for this definition: The numerator is known as edge boundary. You could try for a different definition of edge boundary. Another possibility is to look at expansion in terms of vertices instead of edges.

3.1 Definition

A graph G is (d, ϵ) expanded if

- I. d -regular
- II. $h(G) \geq \epsilon$ (Think of ϵ is a constant. eg $1/10$.)

3.2 FACT

(This will be expanded in the next lecture)

Cheeger's inequality from Spectral Graph Theory: Let G be a d -regular graph such that $\lambda_1 > \lambda_2 > \dots > \lambda_n$, then,

$$(d - \lambda_2)/2 < h(G) \leq \sqrt{2d(d - \lambda_2)}$$

If the graph is an expander, $h(G) = 1/n$