Expander Graphs and their Applications*
Lecture 9: Metric Embedding

Notes taken by Tamir Hazan

March 2002

Summary: We can embed any metric space into $\mathbb{R}^n$, but with some distortion of the distances. We show that the graph metric of expander graphs is the hardest metrics to embed, in the sense that of all finite metric spaces on a given number of points, expanders require the largest distortion.

1 Basic Definitions

(X, d) is a metric space if

- $d : X \times X \to \mathbb{R}^+$.
- $d(x, y) = 0$ if and only if $x = y$.
- $d(x, y) = d(y, x)$.
- $d(x, y) \leq d(x, z) + d(z, y)$.

In this lecture we will examine how to approximate a finite metric (X, d) by the metric space l_2. l_2 is the metric space $(\mathbb{R}^n, ||\cdot||)$ such that for every $y, z \in \mathbb{R}^n$, $||y - z||^2 = \sum_{i=1}^n (y_i - z_i)^2$.

Given the metric spaces (X, d) and $(\mathbb{R}^n, l_2)$ and a transformation $f : X \to \mathbb{R}^n$ we define:

- $expansion(f) = max_{x_1,x_2 \in X} \frac{||f(x_1)-f(x_2)||}{d(x_1,x_2)}$
- $contraction(f) = max_{x_1,x_2 \in X} \frac{d(x_1,x_2)}{||f(x_1)-f(x_2)||}$
- $distortion(f) = expansion(f) \cdot contraction(f)$

It is clear that there are metric spaces that need to be embedded with distortion. E.g the metric $(\{1, 2, 3, 4\}, d)$ with $d(1, 4) = d(2, 4) = d(3, 4) = 1$, and $d(i, j) = 2$ otherwise, since $\{1, 2, 4\}$, $\{1, 3, 4\}$, $\{2, 3, 4\}$ must be on the same line in $\mathbb{R}^n$.

2 Finding the Minimal Distortion

In this section we will present some of the properties of embeddings in l^2. Given a metric space (X, d) we denote its minimal distortion by $C_2(X, d)$.

Theorem 1. *Bourgain (1985) Any n-point metric space (X, d) can be embedded into $O(\log n)$ dimensional Euclidean space with $O(\log n)$ distortion.*

Theorem 2. *Linial, London, and Rabinovich (1995) There is a polynomial time algorithm which computes $C_2(X, d)$.*

* Lecture Notes for a course given by Nati Linial and Avi Wigderson at the Hebrew University, Israel.

Proof. The proof is based on semi definite programming. Let (X, d) be a metric space with $|X| = n$. Let $f : X \to l_2$. Since we can assume that without loss of generality that $contraction(f) = 1$, then $distortion(f) \leq \gamma$ if and only if for every $1 \leq i < j \leq n$:

$$(*) \quad d(x_i, x_j)^2 \leq \|f(x_i) - f(x_j)\|^2 \leq \gamma^2 d(x_i, x_j)^2$$

We say that a matrix Z is positive semi definite (denoted PSD) if Z is symmetric and $x^T Z x \geq 0$ for every $x \in \mathbb{R}^n$.

Let $u_i = f(x_i)$ be the rows of the matrix U. Let $Z = U \cdot U^T$. It is clear that $Z \in PSD$ since it is symmetric, and for every $x \in \mathbb{R}^n$, $x^T Z x = x^T U \cdot U^T x = (U^T x)^T \cdot (U^T x) = \|U^T x\|^2 \geq 0$.

But also the converse is true, if $Z \in PSD$ then $Z = U \cdot U^T$ for some matrix U. To see that, note that Z is symmetric, and therefore diagonizable. Thus $Z = A \Lambda A^T$ for some matrix A and a diagonal matrix Λ. Let $\sqrt{\Lambda}$ be the diagonal matrix which $(\sqrt{\Lambda})_{ii} = \sqrt{(\Lambda)_{ii}}$. Then $Z = A\sqrt{\Lambda}\sqrt{\Lambda}^T A^T = (A\sqrt{\Lambda}) \cdot (A\sqrt{\Lambda})^T$.

Therefore instead of finding $u_i = f(x_i)$ which satisfies $(*)$ we can find a $Z \in PSD$ that satisfies

$$(**) \quad d(x_i, x_j)^2 \leq z_{ii} + z_{jj} - 2z_{ij} \leq \gamma^2 d(x_i, x_j)^2,$$

since $\|u_i - u_j\|^2 = z_{ii} + z_{jj} - 2z_{ij}$ for $Z = U \cdot U^T$.

Thus we conclude that $C_2(X, d) \leq \gamma$ if and only if there is a positive semi definite matrix Z such that for all i, j (**) holds. This is a linear programming problem (more precisely a convex programming problem) which can be solved in polynomial time by the ellipsoid algorithm. $\square$

The algorithm above constructs a primal problem and solves it by the ellipsoid algorithm. Looking at the dual problem gives us an interesting characterization of $C_2(X, d)$. When we transform a primal problem to its dual we take a non negative combination of its constraints. But how do we look at the constraint $Z \in PSD$?

Lemma 3. *$Z \in PSD$ if and only if for all $Q \in PSD$, $\sum_{i,j} q_{ij} z_{ij} \geq 0$.*

Proof. Let Q be a matrix such that $Q_{ij} = q_i \cdot q_j$. Previously we showed that such a matrix is PSD. Therefore $q^T Z q = \sum_{i,j} q_{ij} z_{ij} \geq 0$ implying that $Z \in PSD$.

It can be easily seen that any $Q \in PSD$ of rank 1 must have the form $Q_{ij} = q_i \cdot q_j$ for some values $q_1, ..., q_n$. Thus, if Z is PSD then for every symmetric matrix $Q \in PSD$ of rank 1, $\sum_{i,j} q_{ij} z_{ij} \geq 0$. The lemma follows from the fact that any PSD matrix is a non negative linear combination of rank 1 PSD matrices. To see this, note that any $P \in PSD$ can be written as $A \Lambda A^T$, where Λ is a diagonal matrix of the same rank as P. Therefore $P = \sum_{i=1}^{rank(P)} A \Lambda_{ii} A^T$. $\square$

Our primal problem is:

- $\sum_{ij} q_{ij} z_{ij} \geq 0$ for all $Q \in PSD$.
- $d(x_i, x_j)^2 \leq z_{ii} + z_{jj} - 2z_{ij}$ for all i, j.
- $z_{ii} + z_{jj} - 2z_{ij} \leq \gamma^2 d(x_i, x_j)^2$ for all i, j.

We proceed by deriving an explicit formula for $C_2(X, d)$ from the dual problem:

Theorem 4.

$$C_2(X, d) = \max_{P \in PSD,\, P \cdot \vec{1} = \vec{0}} \sqrt{\frac{\sum_{p_{ij} > 0} p_{ij} d(x_i, x_j)^2}{-\sum_{p_{ij} < 0} p_{ij} d(x_i, x_j)^2}}$$

Proof. In the following proof we shall assume that $\gamma < C_2(X, d)$. Therefore when we inspect a non negative combination of the constraints of the primal problem (the dual problem) we must get a contradiction.

Let us look at the constraints of the first type (for all $Q \in PSD, < q, z >= \sum_{ij} q_{ij} z_{ij} \geq 0$). Since the collection of PSD matrices is convex, a non negative combination $\sum_k a_k < q, z >$ is equal to $< p, z >$ for some matrix $p \in PSD$. Thus the combination of the first type constraints gives us $\sum_{ij} p_{ij} z_{ij} \geq 0$ for some $P \in PSD$.

A contradiction can be reached if a combination of all the constraints result in $0 > 0$. Unfortunately so far we have $\sum_{ij} p_{ij} z_{ij} \geq 0$ for some $P \in PSD$. So to eliminate the z_{ij} for $i \neq j$, we take the following linear combination of the rest of the constraints:

- If $p_{kl} = 0$ then we multiply the constraints involving z_{kl} by zero.
- If $p_{kl} > 0$ then we multiply by $p_{kl}/2$ the constraint $d(x_k, x_l)^2 \leq z_{kk} + z_{ll} - 2z_{kl}$
- If $p_{kl} < 0$ then we multiply by $p_{kl}/2$ the constraint $z_{kk} + z_{ll} - 2z_{kl} \leq \gamma^2 d(x_k, x_l)^2$

To eliminate z_{ii}, we have to choose P such that

$$p_{ii} + \sum_{j \neq i} p_{ij} = 0.$$

Therefore the combination of all the constraints gives

$$(*) \quad \sum_{p_{ij}>0} p_{ij} d(x_i, x_j)^2 + \gamma^2 \sum_{p_{ij}<0} p_{ij} d(x_i, x_j)^2 \leq 0.$$

We get our contradiction if $(*)$ is violated. Thus we conclude the theorem.

□

3 Embeddings in l_2

3.1 Embedding the cube

Given a cube $\{0,1\}^r$ we can easily find an embedding in l_2 with distortion $\sqrt{r}$. Given the embedding $id : \{0,1\}^r \to \mathbb{R}^r$ such that $id(x) = x$, we get that $contraction(id) = \sqrt{r}$ and $expansion(id) = 1$. Using our main theorem we can show that this is the best embedding of the cube. Let define the $2^r \times 2^r$ matrix P:

- $P(i,j) = -1$ if $d(i,j) = 1$.
- $P(i,j) = r - 1$ if $i = j$.
- $P(i,j) = 1$ if $d(i,j) = r$.
- $P(i,j) = 0$ otherwise.

It is easy to check that $P\vec{1} = 0$, and that $P \in PSD$ (the later holds, since P has the same eigenvectors as the cube). Since $\sum_{p_{ij}>0} p_{ij} d(x_i, x_j)^2 = 2^r \cdot r^2$, and $-\sum_{p_{ij}<0} p_{ij} d(x_i, x_j)^2 = 2^r \cdot r$, we get that $C_2(X,d) \geq \sqrt{r}$.

3.2 Embedding expander graphs

Let $G = (V, E)$ be a k-regular graph, $|V| = n$, with $\lambda_2 \geq k - \epsilon$. As before it is simple to see that an expander can be embedded with distortion $O(\log n)$ in l_2. Indeed, take the expander and put it as a simplex in $\mathbb{R}^n$. Since every two nodes of the simplex have distance 1 we get that $expansion = 1$, and $contraction = diam(G)$. Since G is an expander then $diam(G) = O(\log n)$. As before this result is tight.

Lemma 5. *Let $H = (V, E')$ be a graph with the same vertex set as G. 2 vertices are adjacent in H if their distance in G is at least $\log_k n - 2$. Then H has a matching B of $n/2$ edges.*

Proof. G is k-regular graph thus every vertex has at most k^r vertices at distance $\leq r$ from it. If $r = \log_k n - 2$ then there are at most $n/2$ nodes at this distance. Since all the vertices of H have degree $\geq n/2$ then it has a matching of the desired size. This follows from Dirac's sufficient condition for a Hamiltonian circuit. (Modern Graph Theory B. Bollobas p. 106-107). □

Theorem 6. *Let $G = (V, E)$ be a k-regular graph, $|V| = n$, with $\lambda_2 \geq k - \epsilon$. Then $C_2(G) = \Omega(\log n)$.*

Proof. Let B be the adjacency matrix of the matching we found in H. For simplicity we assume that B is a complete matching in H. Let $P = kI - A_G + \frac{\epsilon}{2}(B - I)$ in H. It is easy to see that $P\vec{1} = 0$. $P \in PSD$ since for every $x \perp \vec{1}$,

$$x^T P x = x^T(kI - A_G)x + x^T(\epsilon/2)(B - I)x$$

$x^T(kI - A_G)x \geq \epsilon||x||^2$ since G is expander.

$$x^T(B - I)x = \sum_{(i,j)\ edge\ in\ B} (2x_i x_j - x_i^2 - x_j^2) \geq -2\sum(x_i^2 + x_j^2) = 2||x||^2$$

To get the lower bound on $C_2(G)$ we note that

$$\sum_{p_{ij}>0} d(i,j)^2 p_{ij} \geq \frac{\epsilon}{2} \cdot n(\log_k n - 2)^2$$

since the distances of the entries in B are at least $\log_k n - 2$.

$$-\sum_{p_{ij}<0} d(i,j)^2 p_{ij} = kn$$

Thus $C_2(G) = \Omega(\log n)$.

$\square$