CS 880: Advanced Complexity Theory	04/03/06
Lecture 24: Metric Embeddings	
Instructor: Dieter van Melkebeek	Scribe: Mark Cowlishaw, Michael Kowalczyk

In this lecture we will discuss lower bounds on the distortion of metric embeddings, using expanders as a main component in the proof. In later lectures we will discuss some lower bounds on proof complexity.

1 Background on Metric Embeddings

We begin by defining metric spaces and embeddings:

Definition 1. *[Metric Space] A* metric space (X, δ) *consists of a universe of points X and a measure of distance $\delta : X \times X \to [0, \infty)$, where the distance function δ is symmetric, satisfies the triangle inequality, and is* 0 *only if $x = y$:*

$$\forall_{x,y \in X} \delta(x, y) = \delta(y, x)$$
$$\forall_{x,y,z \in X} \delta(x, y) \leq \delta(x, z) + \delta(z, y)$$
$$\forall_{x,y \in X} \delta(x, y) = 0 \Leftrightarrow x = y$$

A simple example of such a metric space is a weighted graph, with X the vertex set and δ the shortest weighted path between two vertices. A *metric embedding* is the transformation of one metric space into another:

Definition 2. *[Metric Embedding] A metric embedding from (X, δ) onto (X', δ') is a mapping $\phi : X \to X'$.*

In this class, we will only discuss embedding finite metric spaces onto euclidean space (with the familiar distance function as δ'). Given an embedding ϕ, define the following properties:

Definition 3. *[Contraction] Given two metric spaces $S_1 = (X, \delta)$ and $S_2 = (X', \delta')$, the* contraction *of an embedding $\phi : X \to X'$ is the maximum factor by which the distances in S_1 shrink under ϕ:*

$$\text{contraction of } \phi = \sup_{x,y \in X} \frac{\delta(x, y), x \neq y}{\delta'(\phi(x), \phi(y))}$$

Definition 4. *[Expansion] Given two metric spaces $S_1 = (X, \delta)$ and $S_2 = (X', \delta')$, the* expansion *of an embedding $\phi : X \to X'$ is the maximum factor by which the distances in S_1 stretch under ϕ:*

$$\text{expansion of } \phi = \sup_{x,y \in X} \frac{\delta'(\phi(x), \phi(y))}{\delta(x, y), x \neq y}$$

Definition 5. *[Distortion] Given two metric spaces $S_1 = (X, \delta)$ and $S_2 = (X', \delta')$, the* distortion *of an embedding $\phi : X \to X'$ is the product of the expansion and Contraction of ϕ.*

We say that an embedding is *isometric* if the distortion is 1.

The main motivation behind metric embeddings in computer science is in the development of approximation algorithms. Embedding a finite metric space (such as a graph with some distance measurement) into Euclidean space, then solving the problem in Euclidean space exactly, produces a solution within a predictable factor (i.e. the distortion) of an optimal solution. This method has been used to create approximation algorithms for NP-complete optimization problems, such as sparsest cut and balanced cut.

2 Proof of Lower Bounds for Distortion of Metric Embeddings

An upper bound on the distortion of metric embeddings is given by Bourgain:

Theorem 1 (Bourgain's Theorem). *Any N point metric space can be embedded into a Euclidean space of dimension $O(\log N)$, with distortion $O(\log N)$.*

We will show that this distortion result is the best you can do for any embedding into a Euclidean space of any dimension.

Theorem 2 (Lower Bounds for Metric Embedding). *Any embedding ϕ of (X, δ) where X represents the vertices of an expander $G = (V, E)$ with $|V| = N$ onto Euclidean space will have distortion $D(X, \delta) = \Omega(\log N)$.*

2.1 Outline of Proof

The proof of lower bounds proceeds in three parts:

1. Cast an optimal embedding (an embedding with optimal distortion) as a positive, semi-definite program.
2. Use duality to find a lower bound for this problem
3. Use expander properties to show that the lower bound is $\Omega(\log N)$.

2.2 Casting the Optimal Embedding as A Semi-definite Program

We represent the square of the distortion $(D(X, \delta)$ using the following problem:

$$\begin{aligned} &\min_{w,\phi} w \text{ s.t.} \\ &(\forall x, y \in X)\, \delta^2(x,y) \le \underbrace{\|\phi(x) - \phi(y)\|_2^2}_{(*)} \le w \cdot \delta^2(x,y) \end{aligned}$$

We will express this objective and constraints as a semi-definite program. We will need the following fact about positive, semi-definite matrices.

Claim 1. *Let Z be a symmetric matrix of dimension N. Z is positive, semi-definite if and only if there exists a matrix $V = [\phi_1, \phi_2, \ldots, \phi_N]$ of vectors ϕ_i such that $Z = V^T V$. In other words, Z is the Gram matrix of a set of vectors.*

Note that we can rewrite $(*)$ as $\|\phi(x)\|_2^2 + \|\phi(y)\|_2^2 - 2\phi(x)^T\phi(y)$, yielding the following semi-definite program:

$$\begin{array}{l} \min_{w,\mathrm{PSD}\,Z} w \text{ s.t.} \\ (\forall x, y \in X)\, \delta^2(x,y) \leq Z_{xx} + Z_{yy} - 2Z_{xy} \leq w \cdot \delta^2(x,y) \end{array}$$

2.3 Using Duality to find a Lower Bound

To lower bound this problem, we will use the principle of weak duality - any feasible solution to the dual of a minimizing program provides a lower bound on the objective of the primal program. Recall that the dual constraints are linear combinations of the primal constraints, bounded by the coefficients of the primal objective.

We will need to following key fact to form the dual constraints. The proof is left as an exercise.

Claim 2. *Define* PSD_N *the set of positive, semi-definite matrices of dimension* N*. A matrix* Z *of dimension* N *is positive, semi-definite if and only if:*

$$\forall_{C \in \mathrm{PSD}_N} \underbrace{\sum_{x,y} C_{xy} Z_{xy}}_{Tr(C^T \cdot Z)} \geq 0$$

Since we would like all inequalities in the semi-definite program to be $\geq$ inequalities, we will multiply the $\geq$ inequality in our semi-definite program by the positive coefficients in C, and the $\leq$ inequalities by the negative coefficients in C:

$$\begin{array}{lll} C_{xy} \times & (Z_{xx} + Z_{yy} \geq \delta^2(x,y)) & C_{xy} > 0 \\ C_{xy} \times & (Z_{xx} + Z_{yy} \leq w \cdot \delta^2(x,y)) & C_{xy} < 0 \end{array}$$

Yielding the following constraints:

$$\sum_{x,y} C_{xy}(Z_{xx} + Z_{yy} - 2Z_{xy}) \geq \sum_{C_{xy} > 0} C_{xy}\delta^2(x,y) + w \cdot \sum_{C_{xy} < 0} C_{xy}\delta^2(x,y)$$

Now if we choose C to be a symmetric matrix with each row sum equal to zero (i.e. $C \cdot u = 0$), then we have

$$\sum_{x,y} C_{xy}(Z_{xx} + Z_{yy} - 2Z_{xy}) = -2\sum_{x,y} C_{xy} Z_{xy} \leq 0,$$

where the inequality holds since $C \in PSD_N$. Rearranging terms, we get

$$w \quad \geq \quad \frac{\sum_{C_{xy}>0} C_{xy}\delta^2(x,y)}{-\sum_{C_{xy}<0} C_{xy}\delta^2(x,y)} = \frac{\alpha}{\beta},$$

and the maximum of this expression over all possible C is exactly the minimum distortion (by semidefinite programming duality).

2.4 Using Expanders to get a lower bound of $\Omega(\log N)$

Now we want to find some symmetric matrix $C \in PSD_N$ such that $C \cdot u = 0$ and α/β in the previous expression is maximized. We will let (X, δ) be the shortest path metric on an (unweighted) expander, and we want to find a C such that the ratio is large. We will assume that N is even, as the argument generalizes to the the case when N is odd.

We could start by considering $I - A$, which is a natural choice (A is the normalized adjacency matrix, as usual). Clearly, every row sum is zero, so one condition is satisfied. Also, since the eigenvalues of $I - A$ have the form $1 - \lambda$ where λ is an eigenvalue of A, the eigenvalues are nonnegative and the matrix is positive semidefinite. The negative entries of $I - A$ correspond to precisely the edges that are not self-loops. But these entries sum up to at least -1 for each row, and by definition the associated vertices are all at distance 1 from each other, thus $\beta \leq N$. Also, $\alpha = 0$ since the positive entries are on the diagonal, and they all have distance zero. This is a good start, but it leaves something to be desired.

The main idea now is we need to give positive coefficients to pairs of vertices which are far apart. The following result will become useful.

Lemma 1. *For any graph G with maximum degree at most d and $|G| = N$ is even, there exists a perfect matching of vertices that are at distance at least $\Delta = \log_d N - 2$ from each other.*

Proof. Define a graph H on the same vertices as G, but with an edge between x and y if and only if they are at distance at least Δ in G. Then the degree of every vertex in H is at least $N/2$, and we apply the Lemma 2 (below) to complete the proof.

□

Let M be the adjacency matrix of such a perfect matching. Now consider $C = M - I$. It is clear that $Cu = 0$ is satisfied, since each vertex is adjacent to exactly one edge. In this case we get $\alpha \geq N\Delta^2$ and $\beta = 0$, and although C isn't positive semidefinite, it's pretty close. We compute:

$$\begin{aligned} Z^T(M-I)Z &= \sum_{\rho(i,j)\in E(M)} (2Z_iZ_j - Z_i^2 - Z_j^2) \\ &= \sum_{\rho(i,j)\in E(M)} ((Z_i + Z_j)^2 - 2Z_i^2 - 2Z_j^2) \\ &\geq -2Z^TZ. \end{aligned}$$

Now we combine the two ideas to get something that works. We use $C = (I-A)+(\frac{1-\lambda_2}{2})(M-I)$ and this is symmetric, positive semidefinite, and satisfies the $Cu = 0$. Now $\alpha \geq (\frac{1-\lambda_2}{2})N\Delta^2$ and $\beta \leq N$, thus $w \geq \frac{1-\lambda_2}{2}\Delta^2$. Taking square roots yields $\sqrt{w} \geq k\Delta$ where k is some constant. All that remains is to prove lemma 2.

Lemma 2. *Any simple graph H with at least 3 vertices such that each pair (x, y) of nonadjacent vertices satisfies $d_x + d_y \geq N$ is Hamiltonian. In particular, if N is even then H has a perfect matching.*

Proof. Consider a Hamiltonian cycle Y in the complete graph $K_N \supseteq H$ which is maximal in the sense that no other Hamiltonian cycle in K_N has more edges in H. If Y is not completely within H, then there is some edge $(x, y) \in E(Y) \setminus E(H)$. Without loss of generality, we can write

$V(Y) = \{x = x_1, x_1, \cdots, x_N = y\}$. But since the vertices x and y have high degree and H is simple, there must be some edge $(x_w, x_z) \in E(Y)$ such that $(x, x_z) \in H$ and $(x_w, y) \in H$. Finally (assuming $z > w$), the Hamiltonian cycle $(x_z, x_{z+1}, \cdots, x_N = y, x_w, x_{w-1}, \cdots, x_1 = x)$ has at least one more edge within H as does Y, contradicting the supposition that Y was maximal. This closes the proof. $\square$