PRINCETON U. SP'02 COS 598B: ALGORITHMS AND COMPLEXITY

Lecture 1: Introductory Lecture

Lecturer: *Sanjeev Arora* Scribe:*Michael Dinitz*

1 Course Outline

This course will cover three main areas:

1. Geometric embeddings of metric spaces
2. Nonconstructive methods in combinatorics and theoretical computer science
3. Efficient algorithms for some geometric problems, especially convex optimization

In this lecture we begin our study of geometric embeddings of metric spaces.

2 Basic Definitions

We start with the definitions of metric spaces and normed spaces.

DEFINITION 1 *Let X be a set, and let $d : X \times X \to \mathbf{R}^+ \cup \{0\}$. The pair (X, d) is a* metric space *if for all $x, y, z \in X$,*

1. $d(x, y) = d(y, x)$
2. $d(x, y) = 0 \Leftrightarrow x = y$
3. $d(x, y) + d(y, z) \geq d(x, z)$ *(triangle inequality)*

DEFINITION 2 *A* normed space *is $\mathbf{R}^k$ for some finite k together with an associated mapping $a \mapsto \|a\|$ from $\mathbf{R}^k$ to $\mathbf{R}^+ \cup \{0\}$ such that:*

1. *For all $\lambda \in \mathbf{R}$, $\|\lambda a\| = |\lambda| \|a\|$*
2. $\|u + v\| \leq \|u\| + \|v\|$
3. $\|u\| = 0$ *if and only if $a = 0$ (the zero vector)*

Note that $(\mathbf{R}^k, d)$ where $d(u, v) = \|u - v\|$ is a metric space. The following are some examples of norms:

1. l_2 is the normal Euclidean norm. If $u = (u_1, u_2, \ldots, u_k)$ then $\|u\|_2 = (\sum |u_i|^2)^{1/2}$.
2. In general, l_p for $p \geq 1$ is the norm $\|u\|_p = (\sum |u_i|^p)^{1/p}$.
3. l_1 is sometimes referred to as the *Manhattan norm* since it is simply the sum of absolute values of the coordinates.

4. As p approaches infinity the value is dominated by the largest coordinate, so l_∞ is the *max norm* and is the largest coordinate (in absolute value).

Metric and normed spaces arise naturally in many practical applications (e.g., finding similar images in a database) but we will not talk about practical applications here. Rather, we will be interested in theoretical issues (which of course have a bearing on design of practical algorithms) such as how well metrics can be embedded into each other. We need a notion of distortion that measures how good the embedding is.

DEFINITION 3 *Let (X_1, d_1) and (X_2, d_2) be metric spaces. An embedding $f : X_1 \to X_2$ has* distortion C *if there is an $r > 0$ such that $\forall x, y \in X_1$,*

$$r \cdot d_1(x, y) \leq d_2(f(x), f(y)) \leq Cr \cdot d_1(x, y).$$

3 Three types of problems

3.1 Types

There are three main flavors of problems:

1. (Upperbound) Can metrics of class I embed into class II with distortion at most C?
2. (Lowerbound) Show that there is a metric in class I that does not embed into any metric of class II with distortion at most C.
3. (Algorithmic version) Given particular metrics (X_1, d_1) and (X_2, d_2), determine the minimum distortion required to embed X_1 into X_2.

Typically solutions to problems of type 1 also give rise to an algorithm for embedding all metrics of class I into a metric of class II.

3.2 Examples and Results

The following three theorems are examples of the first type of problem:

THEOREM 1 (BOURGAIN '85)
For $p > 1$, every n-point metric embeds into l_p with distortion $O(\log n)$.

We will prove the above theorem later in the course.

THEOREM 2 (FRECHÉT)
Every n-point metric embeds into l_∞ isometrically (distortion 1).

PROOF: Let (X, d) be a n-point metric, and let $f : X \to \mathbf{R}^{|X|}$ map each element $x \in X$ to $(d(x, z))_{z \in X}$ (i.e., vector in $\mathbf{R}^{|X|}$ whose ith coordinate is the distance between x and the ith element of X). We need to verify that $\|f(x) - f(y)\|_\infty = d(x, y)$ for all $x, y \in X$. Note that $f(x) - f(y) = (d(x, z) - d(y, z))_{z \in X}$. By the triangle inequality, $|d(x, z) - d(y, z)| \leq d(x, y)$ for all z, and thus $\|f(x) - f(y)\|_\infty \leq d(x, y)$. On the other hand, if $z = x$ then $|d(x, z) - d(y, z)| = |d(y, x)|$, so $\|f(x) - f(y)\|_\infty \geq d(x, y)$. Thus $\|f(x) - f(y)\|_\infty = d(x, y)$. $\square$

The following theorem will also be proved later in the course.

Theorem 3 (Arora, Lee, Naor '05)
Every n-point l_1 metric embeds into l_2 with distortion $O(\sqrt{\log n}\log\log n)$.

Examples of lowerbounds include the result that the Theorems 1 and 2 are tight up to a constant factor, and that Theorem 3 is tight up to a $\log\log n$ factor. A result of the third type is the following:

Theorem 4
For any metric space (X, d) we can determine in polynomial time the minimum C such that (X, d) embeds into l_2 with distortion C.

Proof: Note that the best C could be a transcendental number so "computing" it just means being able to compute it upto additive error ϵ in time poly$(\log(1/\epsilon))$. Such algorithms are called fully polynomial-time approximation scheme (FPAS). For any C we can determine in polynomial time if there exist vectors $u_1, u_2, \ldots, u_n$ such that $\forall i, j \in X,\ d(i, j)^2 \le \|u_i - u_j\|_2^2 \le C^2 d(i, j)^2$ by using semidefinite programming. The FPAS uses this algorithm and does a binary search on C. $\square$

3.3 Lower Bound Example

What is an example of a metric that does not embed isometrically into l_2? The three point metric with all distances equal to 1 embeds as a triangle, and the four point metric with all distances equal to 1 embeds as a simplex. What about the four point metric where the outside square has distances of 1 and the diagonals have distance 2? In other words, the shortest-path metric for the graph of Figure 1 whose edges AB, BC, CD, DA all have unit length. (Thus $d(A, C) = d(B, D) = 2$.) This obviously does not embed isometrically into l_2, since any three of the points determine a unique embedding (in fact a line) that makes the fourth point impossible to embed. This argument does not lead immediately to a lower bound on the distortion, but there is a different argument that does give a bound.

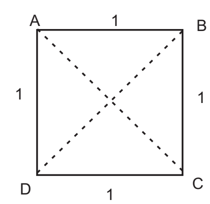

Figure 1: A metric that does not embed isometrically into l_2

Theorem 5
Every embedding of (X, d) into l_2 has distortion at least $\sqrt{2}$.

Proof: This is an example of a popular methodology for proving lower bounds. We first come up with two weightings $w_1 : X \times X \to \mathbf{R}^+ \cup \{0\}$ and $w_2 : X \times X \to \mathbf{R}^+ \cup \{0\}$. Then we show that for all embeddings f,

$$\frac{\sum w_1(x, y) d(x, y)^2}{\sum w_2(x, y) d(x, y)^2} \tag{1}$$

is different by a factor of k from

$$\frac{\sum w_1(x,y)\|f(x)-f(y)\|^2}{\sum w_2(x,y)\|f(x)-f(y)\|^2} \tag{2}$$

This obviously implies that the required distortion is at least $\sqrt{k}$, since we can just use the embedding with the smallest distortion for f. In our specific case, we will let w_2 be the weighting that puts equal weight on AB, BC, CD, DA and w_1 be the weighting that puts equal weight on AC, BD. Then equation (1) equals $\frac{|AC|^2+|BD|^2}{|AB|^2+|BC|^2+|CD|^2+|DA|^2} = \frac{4+4}{1+1+1+1} = 2$. Equation (2) becomes $\frac{|AC|_2^2+|BD|_2^2}{|AB|_2^2+|BC|_2^2+|CD|_2^2+|AD|_2^2}$ which, by the quadrilateral inequality from high school geometry, is less than or equal to 1 for any four points in $\mathbf{R}^k$. Thus the required distortion is at least $\sqrt{2}$. $\square$

The Boolean hypercube consists of the $n = 2^k$ elements of $\{-1,1\}^k$, where we define the distance between two vectors to be the number of coordinates that they differ in (i.e. the Hamming distance). Then if x, y are any two points, $d(x,y) = \frac{1}{2}\|x-y\|_1$. By an inductive argument from the previous theorem, every embedding of the Boolean hypercube into l_2 has distortion at least $\sqrt{k} = \sqrt{\log n}$. This was left as homework (and we went over it in the following lecture).

By the way, the above methodology involving two weight functions is "complete" for estimating distortion into l_2; in other words there is a choice of weight functions that reveals the true distortion. This can be shown using a duality argument (see Matousek's book).

Finally, we indicate a proof of the quadrilateral inequality mentioned above. (Note that "quadrilateral" is a misnomer since the four points need not be coplanar.) One just notices that the inequality holds iff it holds in one dimension ("only if" is trivial, and "if" follows from the observation that the squared distance is the sum of the squares of the contributions from the coordinates) and the one-dimensional case is checked by simple algebra.