The University of Oxford

MSc (Mathematics and Foundations of Computer Science)

Finite dimensional normed spaces

Trinity Term 2015

The steps of (each) miniproject are for your guidance; if you wish to take an alternative route to the desired goal, you are free to do so. But, if you follow the suggested route and find yourself unable to carry out any particular step, you may simply assume it so that you can continue with the miniproject, but should make this assumption clear in your presentation.

Please write or print on one side of the paper only.

The aim of this project is to establish a variant of the Johnson-Lindenstrauss Theorem for metric spaces. All the normed spaces we consider will be over $\mathbb{R}$.

Given two metric spaces $X = (X, d_X)$ and $Y = (Y, d_Y)$ we say that X **embeds** in Y with **distortion** $D \geqslant 1$ if there is some $\phi : X \to Y$ and $C > 0$ such that

$$Cd_X(x, x') \leqslant d_Y(\phi(x), \phi(x')) \leqslant DCd_X(x, x') \text{ for all } x, x' \in X.$$

When this is the case we write

$$X \overset{D}{\hookrightarrow} Y.$$

We shall be interested in embedding *finite* metric spaces into metric spaces induced by normed space. Suppose that (X, d) is a metric space with N points.

1. By considering the map $\phi : X \to \ell_\infty^N; x \mapsto (d(x, x'))_{x' \in X}$ or otherwise show that (X, d) can be isometrically embedded into ℓ_∞^N. In the above notation that is

$$X \overset{1}{\hookrightarrow} \ell_\infty^N.$$

In the first instance we shall be interested in embeddings into lower dimensional spaces at the expense, perhaps, of the distortion. We shall make use of the Landau big-O notation[1] to capture the relationships between the quantities we are considering.

It is instructive to begin with an example.

[1]For functions $f, g : T \to \mathbb{C}$ we write

$$f = O(g) \text{ if and only if } \exists C > 0 \text{ such that } |f(t)| \leqslant C|g(t)| \text{ for all } t \in T,$$

and $f = O(g)$ if and only if $g = \Omega(f)$.

2. Suppose that $2 \geqslant d(x,x') \geqslant 2^{-1}$ for all $x, x' \in X$ with $x \neq x'$. Show that

$$X \overset{O(1)}{\hookrightarrow} \ell_\infty^{O(\log N)},$$

i.e. there is some $k = O(\log N)$ such that X embeds into ℓ_∞^k with distortion $O(1)$.

Returning to the general case when (X, d) is an N-point metric space not necessarily satisfying any additional hypotheses, we plan to construct an embedding for some suitable $k \in \mathbb{N}$ of the form

$$\Phi : (X, d) \to \ell_\infty^k; x \mapsto (d(x, S_i))_{i=1}^k. \tag{1.1}$$

The (non-empty) sets $(S_i)_i$ will be chosen *randomly* in such a way to make this an embedding.

3. Show that any map of the form (1.1) is a contraction *i.e.*

$$\|\Phi(x) - \Phi(x')\|_{\ell_\infty^k} \leqslant d(x, x') \text{ for all } x, x' \in X.$$

The question now is how to choose the S_is; for this we need a little notation. We write

$$B(x,r) := \{x' \in X : d(x,x') < r\} \text{ and } \overline{B}(x,r) := \{x' \in X : d(x,x') \leqslant r\},$$

with the usual warning that $\overline{B(x,r)}$, the closure of $B(x,r)$ in X, is *not* necessarily equal to $\overline{B}(x,r)$.

4. Suppose that $x, x' \in X$ and $r, r' > 0$ are such that $B(x,r)$ and $\overline{B}(x',r')$ are disjoint. Let $k, l \in \mathbb{N}$ be such that

$$k := \max\{|B(x,r)|, |\overline{B}(x',r')|\} \text{ and } l = \frac{N}{10k} + O(1),$$

and $X_1, \dots, X_l$ be independent uniform X-valued random variables; put $S := \{X_1, \dots, X_l\}$. Show that

$$\mathbb{P}(d(x,S) - d(x',S) > r - r') = \Omega(1).$$

There are N^2 pairs $(x,x') \in X^2$, and so it is unlikely the same set S will 'work' for all of them, but by considering many independently selected sets S we can ensure that at least one 'works'.

We can use averaging to find some suitable radii r and r' for use in Exercise 4. For an element $x \in X$ we define the quantity $r_i(x)$ to be the smallest radius such that

$$|\overline{B}(x, r_i(x))| \geqslant 2^i,$$

so that $r_0(x) = 0$ for all $x \in X$.

5. Show that for every $x, x' \in X$ there is some $j \in \mathbb{N}$ such that

$$\max\{r_j(x), r_j(x')\} - \max\{r_{j-1}(x), r_{j-1}(x')\} = \Omega\left(\frac{d(x,x')}{\log N}\right).$$

6. By combining the above exercises or otherwise show that

$$X \overset{O(\log N)}{\hookrightarrow} \ell_\infty^{O(\log^2 N)}$$

An immediate application of the triangle inequality to this result tells us that

$$X \overset{O(\log^3 N)}{\hookrightarrow} \ell_\infty^{O(\log^2 N)},$$

but we can do better.

To translate ℓ_∞^n to ℓ_1^n a little more efficiently we can use a Chernoff-type bound to show that if select many independent sets S then not just one of them 'works', but many 'work'.

7. Using the method above and a Chernoff-type bound to show that $X \overset{O(\log N)}{\hookrightarrow} \ell_1^{O(\log^2 N)}$.

8. Using Hölder's inequality or otherwise extend these results to ℓ_p for $1 < p < \infty$.