## 2.1 Embedding into Normed Spaces

We continue the study with the classic results of embedding into Euclidean normed space.

### 2.1.1 Embedding into $\ell_2$

As we have seen in the previous lecture there are metric spaces that do not admit isometric embedding into $\ell_2$. In addition, we have noted that the identity embedding of $C_4$ into the plane has distortion $\sqrt{2}$.

**Claim 2.1.** Any embedding of $C_4$ into $\ell_2$ has distortion at least $\sqrt{2}$.

Recall the following basic inequalities of $\ell_2$ space (the first is true in any inner product space).

**Parallelogram Law:** $\forall x, y \in \ell_2$ it holds that: 
\[ \|x + y\|_2^2 + \|x - y\|_2^2 = 2 (\|x\|_2^2 + \|y\|_2^2). \]

**Quadrilateral Inequality:** $\forall x, y, z, t \in \ell_2$ it holds that:
\[ \|x - z\|_2^2 + \|y - t\|_2^2 \leq \|x - y\|_2^2 + \|y - z\|_2^2 + \|z - t\|_2^2 + \|t - x\|_2^2. \]

To prove the inequality, it is enough to prove it for the points on the line.

**Proof of Claim 2.1.** Let $f: C_4 \to \ell_2$ be any embedding with distortion $\alpha$. Assume w.l.o.g. that $f$ is non-contractive (since distortion is invariant under scaling). Let $\{A, B, C, D\}$ denote the points of $C_4$, and let $f(A) = A', f(B) = B', f(C) = C', f(D) = D'$. By the quadrilateral inequality and by the definition of distortion we obtain:
\[ \|A' - C'\|_2^2 + \|B' - D'\|_2^2 \leq \|A' - B'\|_2^2 + \|B' - C'\|_2^2 + \|C' - D'\|_2^2 + \|D' - A'\|_2^2 \leq \alpha^2 (d(A, B)^2 + d(B, C)^2 + d(C, D)^2 + d(D, A)^2) = 4\alpha^2. \]

Also, $\|A' - C'\|_2^2 + \|B' - D'\|_2^2 \geq d(A, C)^2 + d(B, D)^2 = 8$. Thus, $4\alpha^2 \geq 8 \Rightarrow \alpha \geq \sqrt{2}$. 

As we mentioned in the previous lecture, Bourgain has shown that any metric space embeds into $\ell_2$ with distortion $O(\log n)$. Moreover, in [6], the bound on distortion was shown to be tight:

**Theorem 2.2 ([6]).** Every embedding of an $n$-vertex constant-degree expander into an $\ell_2$ space, of any dimension, has distortion $\Omega(\log n)$.

We will show a slightly weaker lower bound. Particularly, we will show that there exists an $n$-point metric space, any embedding of which into $\ell_2$ requires distortion of $\Omega(\sqrt{\log n})$. 

---

2-1
**Definition 2.1.** For an integer \( d \geq 1 \), the set \( \{0,1\}^d \) is called a hypercube of dimension \( d \). Each string in the set is called a vertex of the hypercube. The edges of a hypercube are the pairs of vertices that differ at exactly one coordinate.

![Hypercubes](image)

**Figure 2.1: Hypercubes.**

We can think of a hypercubes as unweighted graphs. Thus, we can consider metric spaces defined by this graphs. Let \( H_d \) denote the metric space defined over the vertices of a \( d \)-dimensional hypercube, imposed by its graph structure. Note that this metric is exactly the \( \ell_1 \)-distance, moreover, when the points are binary strings, the \( \ell_1 \)-distance is exactly the Hamming distance.

We have shown already that any embedding of \( H_2 (= C_4) \) into \( \ell_2 \) requires distortion of at least \( \sqrt{d} \).

**Theorem 2.3.** Any embedding of \( H_d \) (for \( d \geq 2 \)) into \( \ell_2 \), incurs distortion of at least \( \sqrt{d} \).

Denoting \( n = 2^d \), we get distortion at least \( \sqrt{d} = \sqrt{\log n} \). Note that the distortion of the identity embedding of \( H_d \) into \( \ell_2 \) is \( \sqrt{d} \).

**Proof.** (by Enflo [2]) Recall that for any \( u \neq v \in H^d \), \( (u, v) \) is an edge iff \( \|u - v\|_1 = 1 \). Let \( E \) denote the set of all edges of \( H_d \). Note that \( |E| = d2^{d-1} \), since for each vertex \( u \) there are exactly \( d \) edges that contain \( u \), and thus all the vertices together cover all the edges twice. Let \( F = \{(u, \bar{u})|u \in H_d\} \) denote the set of all the longest diagonals of \( H_d \), where \( \bar{u} \) denotes the negation of \( u \). Then \( |F| = 2^{d-1} \).

We prove the following generalization of the quadrilateral inequality.

**[Hyper-Quadrilateral Inequality]:** For any \( f : \{0,1\}^d \to \ell_2 \) it holds that

\[
\sum_{(u,\bar{u}) \in F} \|f(u) - f(\bar{u})\|_2^2 \leq \sum_{(u,v) \in E} \|f(u) - f(v)\|_2^2.
\]

Assume the inequality is correct, and let \( f: H_d \to \ell_2 \) be a (w.l.o.g.) non-contractive embedding with distortion \( \alpha \). Then, we have

\[
2^{d-1} \cdot d^2 \leq \sum_{(u,\bar{u}) \in F} \|f(u) - f(\bar{u})\|_2^2 \leq \sum_{(u,v) \in E} \|f(u) - f(v)\|_2^2 \leq \alpha^2 \cdot d \cdot 2^{d-1},
\]

implying \( \alpha \geq \sqrt{d} \) as required. Thus it remains to prove the inequality.
The proof is by induction on $d$. The claim is true for $d = 2$ (the quadrilateral inequality). Assume the claim is true for any $d' < d$. Denote by $X$ the set of $2^d$ vertices of the cube $\{0,1\}^d$. Let $X_0$ and $X_1$ be the sets of all points of $X$ with last coordinate being 0 and 1, respectively. Then $|X_0| = |X_1| = 2^{d-1}$. Let $E_0$ and $E_1$ be the sets of edges of $X_0$ and $X_1$, respectively, and let $F_0$ and $F_1$ be the sets of longest diagonals of $X_0$ and $X_1$, respectively. Note that $E_0, E_1$ are also edges of $X$, but $F_0, F_1$ are not the longest diagonals of $X$. By the induction hypothesis $X_0$ and $X_1$, we obtain that:

$$\sum_{(u_0,v_0) \in E_0} \|f(u_0) - f(v_0)\|_2 \leq \sum_{(u_0,v_0) \in E_0} \|f(u_0) - f(v_0)\|_2,$$

and

$$\sum_{(u_1,v_1) \in E_1} \|f(u_1) - f(v_1)\|_2 \leq \sum_{(u_1,v_1) \in E_1} \|f(u_1) - f(v_1)\|_2,$$

where for a string $u \in \{0,1\}^{d-1}$, and a bit $b \in \{0,1\}$, we denote by $ub$ the concatenation of $u$ with $b$ (i.e. $ub \in \{0,1\}^d$).

For each vertex $u_0 \in X_0$, consider the vertices $\bar{u}1 \in X_1$, $\bar{u}0 \in X_0$, $u_1 \in X_1$. Note that $(u_0, \bar{u}1)$ and $(\bar{u}0, u_1)$ form two longest diagonals of $X$. Moreover, by the quadrilateral inequality we have

$$\|f(u_0) - f(\bar{u}1)\|_2^2 + \|f(\bar{u}0) - f(u_1)\|_2^2 \leq \|f(u_0) - f(\bar{u}0)\|_2^2 + \|f(u_1) - f(\bar{u}1)\|_2^2 + \|f(\bar{u}0) - f(u_1)\|_2^2,$$

The first two arguments are the longest diagonals of $X_0$ and $X_1$ respectively, and the last two arguments are the edges of $X$, connecting the subcubes $X_0$ and $X_1$. Therefore, if we apply this argument to all the points of $X_0$ (essentially, to the half of the points, since we don’t want duplications) we will “cover” all the longest diagonals of $X$, by means of all the longest diagonals of $X_0$ and $X_1$, and all the edges of $X$ going between $X_0$ and $X_1$:

$$\sum_{(v,\bar{v}) \in F} \|f(v) - f(\bar{v})\|_2^2 \leq \sum_{(u_0,\bar{u}0) \in F_0} \|f(u_0) - f(\bar{u}0)\|_2^2 + \sum_{(u_1,\bar{u}1) \in F_1} \|f(u_1) - f(\bar{u}1)\|_2^2 + \sum_{(w,s) \in E_{0,1}} \|f(w) - f(s)\|_2^2,$$

where $E_{0,1} = \{(w,s) \mid w \in X_0, s \in X_1, \|w - s\|_1 = 1\}$. Applying the induction’s hypothesis, we conclude the lemma.

\[ \square \]

**Open Problem 2.1.** What is the lowest distortion $\alpha(n)$ such that all $n$-point metric spaces in $\ell_1$ are $\alpha(n)$ embeddable into $\ell_2$? The conjecture is that distortion is $O(\sqrt{\log n})$. In [1] the authors proved that $\alpha(n) = O(\sqrt{\log n \log(\log n)})$.

### 2.2 Embedding into $\ell_\infty$ With Low Distortion

In the last lecture we have seen the Frechet’s embedding [1]: Every $n$-point metric isometrically embeds into $\ell_{\infty}^{n-1}$. In fact, Frechet proved a stronger statement: Every separable metric space isometrically embeds into $\ell_\infty$ (of infinite dimension). The construction of the embedding is
similar to the finite case we have seen in class, where the “location” points are the points of the countable dense set of the metric space. Thus, we can use a finite $\epsilon$-net as the location points, loosing in precision for the sake of dimension.

**Theorem 2.4.** (Farago [3]) Let $(X, \|\cdot\|)$ be a normed space of dimension $d$. Let $S \subset X$ be a bounded subset. Then, for any $0 < \epsilon < \frac{1}{2}$, there is an embedding $f : S \to \ell_\infty^{O(\frac{1}{\epsilon^d})}$, with distortion $1 + \epsilon$.

Before the proof of the theorem, we shall define several concepts and prove some lemmas.

**Definition 2.2.** Let $(X, d)$ be a metric space and $N \subseteq X$.

$N$ is $\epsilon$-dense in $X$, if $\forall x \in X$ there exists $y \in N$ such that $d(x, y) \leq \epsilon$.

$N$ is $\epsilon$-separated, if $\forall x, y \in N$ it holds that $d(x, y) > \epsilon$.

$N$ is $\epsilon$-net of $X$, if $N$ is $\epsilon$-dense and is $\epsilon$-separated.

The idea is to show that there exists an $\epsilon$-net of $S$ of size $k = O\left(\frac{1}{\epsilon^d}\right)$, and that the Frechet embedding to the points of the net has distortion $1 + \epsilon$.

Namely, we will show that there is $N = \{z_1, z_2, \ldots, z_k\} \subseteq S$, such that $N$ is $\epsilon$-net of $S$, of size $k = O\left(\frac{1}{\epsilon^d}\right)$, and that for the embedding $f$ defined by $\forall x \in S$, $f(x) = (\|x - z_1\|, \|x - z_2\|, \ldots, \|x - z_k\|)$, it holds that

$$\forall x \neq y \in S, \quad \frac{1}{1 + \epsilon} \|x - y\| \leq \|f(x) - f(y)\|_{\infty} \leq \|x - y\|.$$

**Lemma 2.5** (A bounded, separated set in a normed space is not too large). Let $R > 0$, $0 < \epsilon \leq R/2$. Let $(X, \|\cdot\|)$ be a $d$ dimensional normed space, $B_R = \{x \in X \mid \|x\| \leq R\}$. If $A \subset B_R$ is $\epsilon$- separated, then $|A| = O\left(\frac{R}{\epsilon}\right)^d$.

**Proof.** We use the standard volume argument to bound the size of $A$. Denote $k = |A|$. The closed balls of radius $\frac{\epsilon}{2}$ around the points in $A$ are disjoint, and their union is contained in a closed ball of radius $(R + \frac{\epsilon}{2})$ around 0.

There are various definitions for the volume of the set in some normed space. We can use a common one – the Jordan measure, which can be defined for all normed spaces over $\mathbb{R}$. Recall that not any set is Jordan measurable, but it is known that a bounded set is Jordan measurable if and only if its boundary has Jordan measure zero. In addition, it can be shown that the boundary of a convex set has Jordan measure zero. Therefore, since the closed ball of radius $r$ of any normed space in $\mathbb{R}$ is a convex set, it is Jordan measurable.

We can assume w.l.o.g. that $V = \mathbb{R}^d$. In addition, denote by $B_{2^}\frac{\epsilon}{2}$ and by $B_{(R + \frac{\epsilon}{2})}$ the closed balls of the appropriate radii under the $\|\cdot\|_X$. Note that $B_{(R + \frac{\epsilon}{2})} = \left(\frac{R + \frac{\epsilon}{2}}{\frac{\epsilon}{2}}\right) \cdot B_{2^}\frac{\epsilon}{2}$. Therefore,

$$Vol\left(B_{(R + \frac{\epsilon}{2})}\right) = \left(\frac{R + \frac{\epsilon}{2}}{\frac{\epsilon}{2}}\right)^d Vol\left(B_{2^}\frac{\epsilon}{2}\right).$$

Therefore,

$$k \leq \left(\frac{R + \epsilon/2}{\epsilon/2}\right)^d = O\left(\frac{R}{\epsilon}\right)^d.$$

\[\square\]
Claim 2.6. Let $R > 0$, $0 < \epsilon \leq R/2$. In every normed space $X$ of dimension $d$, there is an $\epsilon$-net $N \subset B_R$, of size $O\left(\frac{R}{\epsilon}\right)^d$.

Proof. We build $N$ greedily. At the beginning $N = \emptyset$. Choose an arbitrary point in $B_R$ and add it to $N$. Let $B$ be the union of closed balls of radius $\epsilon$ around the points in $N$. As long as there is a point in $B_R \setminus B$, add it to $N$. When the algorithm stops, the resulting set $N$ is clearly $\epsilon$-dense and $\epsilon$-separated. Therefore, $N$ is an $\epsilon$-net and by lemma 2.5, $|N| = O\left(\frac{R}{\epsilon}\right)^d$.

In fact, one can prove that for any $\epsilon > 0$, for any metric space $X$ there exists an $\epsilon$-net of $X$.

It is not clear how to find a net efficiently in a general normed space. However, if the norm is $\ell_p$-norm then we can take the grid (that covers the set $S$) with side length of $\epsilon/d^p$, and sieve the points that are too close to the points that have been taken to the net. This can be done in running time of $O(kk'd)$ where $k$ is the number of points in the grid and $k'$ is the number of points that survived in the net.

Note that we can assume w.l.o.g. that $S \subset B_1$, so it is enough to prove the following theorem.

Theorem 2.7. Let $0 < \epsilon \leq 1/2$, and let $X$ be a normed space of dimension $d$. Then $B_1$ can be embedded into $l_\infty$ with distortion $1 + \epsilon$ and $k = O\left(\frac{1}{\epsilon}\right)^d$.

Proof. By Claim 2.6 we can find an $(\epsilon/8)$-net $N$ of $B_2$ of size $k = O\left(\frac{1}{\epsilon}\right)^d$. Let $z_1 \ldots z_k$ be the points of $N$. We define the embedding $f : B_1 \rightarrow l_\infty$ by setting

$$f(x) = (\|x - z_1\|, \ldots, \|x - z_k\|), \quad \forall x \in B_1.$$ 

We show that for all $x \neq y \in B_1$, it holds that $\frac{1}{1 + \epsilon} \leq \frac{\|f(x) - f(y)\|_\infty}{\|x - y\|} \leq 1$.

From the triangle inequality, for all $1 \leq i \leq k$, $\|x - z_i\| - \|y - z_i\| \leq \|x - y\|$, so the right hand side of the inequality is correct. To prove the left hand side of the inequality we divide into two cases: $\|x - y\| \geq 1/2$ and $\|x - y\| < 1/2$. (Note that if $x = z_i$ or $y = z_j$ then trivially $\frac{\|f(x) - f(y)\|_\infty}{\|x - y\|} \geq 1$.)

We show that there exists $1 \leq i \leq k$ such that $\frac{\|y - z_i\| - \|x - z_i\|}{\|y - x\|} \geq \frac{1}{1 + \epsilon}$. (Namely, not all the coordinates are very close to each other. From this follows that the embedding is one to one.)

Case A: $\|x - y\| \geq 1/2$.

Let $z_i$ be the closest point of $N$ to $x$. (Note that $\|y - z_i\| \neq \|x - z_i\|$, as otherwise, by the triangle inequality $1/2 \leq \|x - y\| \leq \|x - z_i\| + \|y - z_i\| \leq \epsilon/4 \leq 1/8$). Therefore,

$$\frac{\|y - z_i\| - \|x - z_i\|}{\|y - x\|} \geq \frac{\|y - x\| - 2\|x - z_i\|}{\|y - x\|} = 1 - 2\frac{\|x - z_i\|}{\|y - x\|} \geq 1 - 2\frac{\epsilon/8}{1/2} = 1 - \epsilon/2 \geq \frac{1}{1 + \epsilon}.$$ 

Case B: $\|x - y\| < 1/2$.

We want to choose a point $x'$, such that $x'$ is on the line passing through $x, y$, and $\|x' - y\| = 1$. Denote $\lambda = \|x - y\|$. Let $x' = 1/\lambda(x - (1 - \lambda)y)$. Then:

- $x = \lambda x' + (1 - \lambda)y$. 

2-5
\[ \|x' - y\| = 1, \]
\[ x' \in B_2, \text{ since by the triangle inequality, } \|x'\| \leq \|x' - y\| + \|y\| \leq 2. \]

Therefore, by the case A, if \( z_i \) is the closest point of \( N \) to \( x' \), then (recall \( \|y - x'\| = 1 \))

\[ \|y - z_i\| - \|x' - z_i\| \geq \frac{1}{1 + \epsilon}. \]

By the triangle inequality (or convexity of the norm) we have

\[ \|x - z_i\| \leq \lambda \|x' - z_i\| + (1 - \lambda) \|y - z_i\|. \]

Therefore,

\[
\left| \frac{\|y - z_i\| - \|x - z_i\|}{\lambda} \right| \geq \left| \frac{\|y - z_i\| - (\lambda \|x' - z_i\| + (1 - \lambda) \|y - z_i\|)}{\lambda} \right| = \left| \|y - z_i\| - \|x' - z_i\| \right| \geq \frac{1}{1 + \epsilon},
\]

which completes the proof.

\[ \square \]

Remark 2.8. Essentially, making a more complicated arguments (from Banach space theory) can prove a stronger result: every \( d \)-dimensional normed space \( X \) can be embedded into \( \ell_\infty^{O(1/\epsilon^d)} \), with distortion \( 1 + \epsilon \). See [5] for an outline of such proof.

References


