15-859(J): Metric Embeddings and Methods **Lecturer:** A. Gupta
Topic: Dimensionality Reduction and Random Projections **Date:** October 21, 2003
Scribe: Leonid Kontorovich

15.1 Introduction – adapted from Lecture 14

Given a set of points in some $\ell_p^{k'}$, can we reduce the number of dimensions required without distorting distances too much? Formally,

Problem 15.1.1 *Given $(X, \|\cdot\|_p) \subseteq \mathbf{R}^{k'}$, what is the least $k = k(n, D)$ such that*

$$(X, \|\cdot\|_p) \overset{D}{\hookrightarrow} \ell_p^k?$$

It turns out that the answer depends on p. E.g., for ℓ_∞, almost no dimension reduction is possible. For ℓ_1, a lower bound was recently shown by Brinkman and Charikar (2003), and much simpler exposition was given by Lee and Naor (2003):

Theorem 15.1.2 *There exists a set of n points in ℓ_1 which require $k = n^{\Omega(1/D^2)}$.*

However, the situation is very good for Euclidean space ℓ_2, as a seminal paper of Johnson and Lindenstrauss (1984) shows:

Theorem 15.1.3 *For any set of n points and $D = 1+\epsilon$, it is possible to get down to $k = O\left(\frac{\log n}{\epsilon^2}\right)$ dimensions.*

Furthermore, Alon (2001) showed that this result is almost tight:

Theorem 15.1.4 *With $D = (1+\epsilon)$, the uniform metric on n points requires at least $k = \Omega\left(\frac{\log n}{\epsilon^2 \log \frac{1}{\epsilon}}\right)$ dimensions.*

We will presently see a proof of Thm. 15.1.3, due to Indyk-Motwani (1998).

15.2 A proof of the Johnson-Lindenstrauss theorem

Theorem 15.2.1 *Let $(X, d) = \{\vec{v}^{(1)}, \ldots, \vec{v}^{(1)}\} \subset \mathbb{R}^n = \ell_2^n$. Then there exists an embedding $\phi : \mathbb{R}^n \to \mathbb{R}^k$, with $k = O\left(\frac{\log n}{\epsilon^2}\right)$ such that, for some ρ,*

$$(1-\epsilon) \le \frac{\|\phi(\vec{x}) - \phi(\vec{y})\|_2^2}{\rho \|\vec{x} - \vec{y}\|_2^2} \le (1+\epsilon) \qquad (15.2.1)$$

Proof: Choose $\vec{r}^{(1)}, \ldots, \vec{r}^{(k)} \in \mathbb{R}^n$ with $\vec{r}_i^{(t)} \sim N(0,1)$. We claim that $\phi(\vec{v}) = (\langle \vec{v}, \vec{r}^{(1)} \rangle, \ldots, \langle \vec{v}, \vec{r}^{(k)} \rangle)$ is an embedding satisfying (15.2.1).

Given a vector $\vec{v} = (v_1, \ldots, v_n) \in \mathbf{R}^n$, and a random vector $\vec{r} = (r_1, \ldots, r_n) \sim N(0, I)$, let $X = \langle \vec{v}, \vec{r} \rangle$. Since a linear combination of normal random variables is also normal, X is a normal random variable with variance ℓ^2, where $\ell = \|\vec{v}\|_2$:

$$\mathbf{E}[X] = \mathbf{E}\left[\sum_{i=1}^{n} v_i r_i\right] = \sum_{i=1}^{n} v_i \mathbf{E}[r_i] = 0, \text{ and} \tag{15.2.2}$$

$$\mathbf{Var}[X] = \sum_{i=1}^{n} \mathbf{Var}[v_i r_i] = \sum_{i=1}^{n} v_i^2 \mathbf{Var}[r_i] = \sum_{i=1}^{n} v_i^2 = \ell^2. \tag{15.2.3}$$

Observe that

$$\begin{aligned}
\mathbf{E}\left[\|\phi(\vec{v})\|_2^2\right] &= \mathbf{E}\left[\sum_{i=1}^{k} \left\langle \vec{v}, \vec{r}^{(k)} \right\rangle^2\right] && (15.2.4)\\
&= \sum_{i=1}^{k} \mathbf{E}\left[\left\langle \vec{v}, \vec{r}^{(k)} \right\rangle^2\right] && (15.2.5)\\
&= \sum_{i=1}^{k} \mathbf{E}\left[\sum_{i=1}^{n} v_i^2 r_i^2 + 2 \sum_{1 \le i < i' \le n} v_i v_{i'} r_i r_{i'}\right] && (15.2.6)\\
\text{(by (15.2.3) and independence of } r_i \text{ and } r_{i'}) \quad &= \sum_{i=1}^{k} \left(\ell^2 + \sum_{i \ne i'} v_i v_{i'} \mathbf{E}[r_i]\mathbf{E}[r_{i'}]\right) && (15.2.7)\\
\text{(since each } r_i \text{ has mean 0)} \quad &= k\ell^2 && (15.2.8)
\end{aligned}$$

We will show that if $k = O\left(\frac{\log n}{\epsilon^2}\right)$, then $\|\phi(\vec{v})\|_2^2 \simeq (1 \pm \epsilon)k\ell^2$ with probability $\ge 1 - \frac{1}{n^2}$ (in other words, $\|\phi(\vec{v})\|_2^2$ will be tightly concentrated around its mean). Then, since ϕ is linear (being a projection onto a k-dimensional subspace spanned by the $\vec{r}^{(t)}$'s), we'll have

$$\frac{\|\phi(\vec{x} - \vec{y})\|_2^2}{k\,\|\vec{x} - \vec{y}\|_2^2} = \frac{\|\phi(\vec{x}) - \phi(\vec{y})\|_2^2}{k\,\|\vec{x} - \vec{y}\|_2^2} \tag{15.2.9}$$

and the desired (15.2.1) will follow.

For convenience, let us rescale the lengths of our original vectors $\vec{v}^{(t)}$, dividing each by ℓ^2. Then, fixing a $\vec{v}$, and letting $X_i = \left\langle \vec{v}, \vec{r}^{(i)} \right\rangle$, and $X = \sum_{i=1}^{k} X_i^2$, we have $X_i \sim N(0, 1)$ and $\mathbf{E}[X] = k$. Let $\alpha = (1 + \epsilon)k$; we will compute the probability that X substantially exceeds its mean:

$$\begin{aligned}
\mathbf{Pr}[X \ge \alpha] &= \mathbf{Pr}\left[e^{sX} \ge e^{s\alpha}\right] \quad \text{(holds for all } s \ge 0) && (15.2.10)\\
&= \mathbf{Pr}\left[e^{sX} e^{-s\alpha} \ge 1\right] && (15.2.11)\\
\text{(Markov's inequality)} \quad &\le \mathbf{E}\left[e^{sX} e^{-s\alpha}\right] && (15.2.12)\\
&= e^{-s\alpha} \mathbf{E}\left[e^{s \sum_i X_i^2}\right] && (15.2.13)\\
&= e^{-s\alpha} \mathbf{E}\left[\prod_i e^{sX_i^2}\right] && (15.2.14)\\
\text{(by mutual independence of the } X_i^2\text{'s)} \quad &= e^{-s\alpha} \prod_i \mathbf{E}\left[e^{sX_i^2}\right] && (15.2.15)
\end{aligned}$$

It remains to compute $\mathbf{E}\left[e^{sX_i^2}\right]$. We know that $X_i \sim N(0,1)$. For $x \geq 0$,

$$\mathbf{Pr}[|X_i| \geq x] = 2\int_x^\infty \frac{1}{\sqrt{2\pi}} \exp(-t^2/2)dt$$

so $|X_i| \sim \frac{2}{\sqrt{2\pi}} \exp(-x^2/2)$ on $[0,\infty)$. Let $Y = |X_i|^2 = h(|X_i|)$. By the change of variable formula (see, for example, Spivak (1965), pp 67-72), we have

$$Y \sim \frac{1}{\sqrt{2\pi}} \exp(-y/2)y^{-\frac{1}{2}}.$$

Then the moment generating function of Y, $\mu_Y(s)$, is computed by

$$\begin{aligned}
\mu_Y(s) &= \mathbf{E}\left[e^{sY}\right] && (15.2.16)\\
&= \int_0^\infty e^{sy}\frac{1}{\sqrt{2\pi}} \exp(-y/2)y^{-\frac{1}{2}}dy && (15.2.17)\\
\text{(integrating by parts twice)} &= \frac{1}{\sqrt{1-2s}} \qquad \text{(defined for } -\infty < s < \tfrac{1}{2}). && (15.2.18)
\end{aligned}$$

Substituting (15.2.18) into (15.2.15) we get

$$\mathbf{Pr}[X \geq \alpha] \leq e^{-s\alpha}(1-2s)^{-k/2}. \tag{15.2.19}$$

Minimizing the r.h.s. of (15.2.19) as a function of s (to get the tightest bound possible), we get

$$s = \frac{1}{2}\left(1-\frac{k}{\alpha}\right) = \frac{1}{2}\left(1-\frac{k}{(1+\epsilon)k}\right). \tag{15.2.20}$$

Plugging (15.2.20) into (15.2.19), we have

$$\begin{aligned}
\mathbf{Pr}[X \geq \alpha] &\leq e^{-s\alpha}(1-2s)^{-k/2} && (15.2.21)\\
&= e^{-\frac{1}{2}(\alpha-k)}\left(\frac{\alpha}{k}\right)^{k/2} && (15.2.22)\\
&= e^{-\epsilon k/2}(1+\epsilon)^{k/2} = \left(\frac{e^\epsilon}{1+\epsilon}\right)^{-k/2} && (15.2.23)\\
&= \exp\left(-\frac{k}{2}\left(\epsilon - \log(1+\epsilon)\right)\right) && (15.2.24)\\
\text{(Taylor expanding)} &\leq \exp\left(-\frac{k}{2}\left(\frac{\epsilon^2}{2}-\frac{\epsilon^3}{3}\right)\right). && (15.2.25)\\
& && (15.2.26)
\end{aligned}$$

Hence, if

$$k = \frac{4\log n}{\left(\frac{\epsilon^2}{2}-\frac{\epsilon^3}{3}\right)},$$

we have

$$\mathbf{Pr}[X \geq (1+\epsilon)k] \leq \frac{1}{n^2}. \tag{15.2.27}$$

An entirely analogous calculation yields

$$\mathbf{Pr}[X \leq (1-\epsilon)k] \leq \frac{1}{n^2},$$

finishing off the proof.

■

15.3 Embedding ℓ_2 in ℓ_1

Theorem 15.3.1 *Any $(X, d) \subset \mathbb{R}^n = \ell_2^n$ can be embedded into ℓ_1 with no distortion.*

Proof: Pick a $\vec{v} \in X$, and let $\vec{r}$ be a random unit vector drawn uniformly from S^{n-1}. Then

$$\begin{aligned}
\mathbf{E}[|\langle \vec{v}, \vec{r} \rangle|] &= \int_{S^{n-1}} |\langle \vec{v}, \vec{x} \rangle| d\mu(\vec{x}) && (15.3.28)\\
&= \int_{S^{n-1}} \|\vec{v}\|_2 \|\vec{x}\|_2 |\cos\theta| d\mu(\vec{x}) && (15.3.29)\\
&= \|\vec{v}\|_2 \int_{S^{n-1}} |\cos\theta(x)| d\mu(\vec{x}) && (15.3.30)\\
\text{(by symmetry, } \theta \text{ is independent of } \vec{v}\text{)} \quad &= \alpha \|\vec{v}\|_2 && (15.3.31)
\end{aligned}$$

This yields an embedding ϕ of ℓ_2^n into $\ell_1^{S^{n-1}}$, where

$$\phi : \vec{v} \mapsto f_{\vec{v}} \tag{15.3.32}$$

such that $f_{\vec{v}} : S^{n-1} \to \mathbb{R}$, with

$$f_{\vec{v}} : \vec{r} \mapsto \langle \vec{v}, \vec{r} \rangle .$$

Note that $\ell_1^{S^{n-1}}$ is actually the space of all Lebesgue integrable functions on S^{n-1}, and we should be properly denoting this by $\mathscr{L}^1(S^{n-1})$.

This embedding is distortion-free because

$$\begin{aligned}
\|\phi(\vec{x}) - \phi(\vec{y})\|_1 &= \left\|f_{\vec{x}} - f_{\vec{y}}\right\|_1 && (15.3.33)\\
\text{(by linearity of } \phi \text{ \& letting } \vec{z} = \vec{x} - \vec{y}\text{)} \quad &= \left\|f_{\vec{x}-\vec{y}}\right\|_1 = \|f_{\vec{z}}\|_1 && (15.3.34)\\
\text{(definition of } \ell_1 \text{ norm)} \quad &= \int_{S^{n-1}} |f_{\vec{z}}(\vec{r})| d\mu(\vec{r}) && (15.3.35)\\
&= \int_{S^{n-1}} |\langle \vec{z}, \vec{r} \rangle| d\mu(\vec{r}) && (15.3.36)\\
&= \alpha \|\vec{z}\|_2 = \alpha \|\vec{x} - \vec{y}\|_2 . && (15.3.37)
\end{aligned}$$

■

Now using a large-deviation bound similar to that used for Theorem 15.1.3 implies that if we pick $k = O\left(\frac{\log n}{\epsilon^2}\right)$ random unit vectors $\vec{r}^{(1)}, \ldots, \vec{r}^{(k)} \in \mathbb{R}^n$ and let $\phi(\vec{v}) = (\langle \vec{v}, \vec{r}^{(1)} \rangle, \ldots, \langle \vec{v}, \vec{r}^{(k)} \rangle)$, then

$$(1-\epsilon) \leq \frac{\|\phi(\vec{x}) - \phi(\vec{y})\|_1}{\|\vec{x}-\vec{y}\|_2} \leq (1+\epsilon) \tag{15.3.38}$$

with high probability.

15.4 Embedding via random walks

Let (X, d) be a finite subset of the ℓ_2^n metric space. We will exhibit an embedding of (X, d) into ℓ_1^k with distortion $\sqrt{3}$.

Consider a given fixed $\vec{v} = (v_1, \ldots, v_n) \in \mathbb{R}^n$ and let $\vec{\omega}$ be a random variable with range $\{-1, 1\}^n$, consisting of n independent Bernoulli events with equiprobable outcome ± 1.

If we start at the origin of the real line and at time i take a step in the direction of $\pm v_i$ with equal probability, how far do we expect to be from the origin at time n? In other words, if $X_i = \omega_i v_i$ and $X = X(\vec{\omega}) = \sum_{i=1}^n X_i$, what is $\mathbf{E}[|X|]$? We will show that

$$\frac{\|\vec{v}\|_2}{\sqrt{3}} \leq \mathbf{E}[|X|] \leq \|\vec{v}\|_2 \,. \tag{15.4.39}$$

Then let $Y = Y(\vec{v})$ be a 2^n-dimensional vector, its dimensions indexed by the 2^n values that $\vec{\omega}$ can take on, with the $\vec{\omega}$'th component of Y equal to $\frac{1}{2^n} X(\vec{\omega})$:

$$Y = \frac{1}{2^n}\left(X(-1, -1, \ldots, -1), \ldots, X(1, 1, \ldots, 1)\right). \tag{15.4.40}$$

It is trivial to see that $\|Y\|_1 = \mathbf{E}[|X|]$, and that this implies that the embedding $\phi : \vec{v} \mapsto Y(v)$ is a $\sqrt{3}$-distortion embedding.

It remains to verify (15.4.39). Let us compute

$$\begin{aligned}
\mathbf{E}\left[X^2\right] &= \mathbf{E}\left[\left(\sum X_i^2\right)^2\right] && (15.4.41)\\
&= \mathbf{E}\left[\sum_i X_i^2 + 2\sum_{1\leq i<j\leq n} X_i X_j\right] && (15.4.42)\\
\text{(X_i and X_j are independent, zero-mean r.v.'s)} \quad &= \sum \mathbf{E}\left[\omega_i^2 v_i^2\right] = \sum v_i^2 = \|v\|_2^2\,. && (15.4.43)
\end{aligned}$$

Now we claim that

$$\begin{aligned}
\frac{\sqrt{\mathbf{E}\left[X^2\right]}}{\sqrt{3}} &\leq \mathbf{E}[|X|] && (15.4.44)\\
&\leq \sqrt{\mathbf{E}\left[X^2\right]}. && (15.4.45)
\end{aligned}$$

We obtain (15.4.45) immediately from Jensen's inequality, since $f(x) = x^2$ is a convex function and so $f(\mathbf{E}[|X|]) \leq \mathbf{E}[f(X)]$. We will obtain (15.4.44) from Hölder's inequality: $\langle \vec{x}, \vec{y} \rangle \leq \|\vec{x}\|_p \|\vec{y}\|_q$ for $\frac{1}{p} + \frac{1}{q} = 1$. Hölder yields

$$\mathbf{E}[YZ] \leq (\mathbf{E}[Y^p])^{\frac{1}{p}} (\mathbf{E}[Z^q])^{\frac{1}{q}}.$$

Taking $Y = |X|^{\frac{2}{3}}$ and $Z = |X|^{\frac{4}{3}}$, with $p = \frac{3}{2}$ and $q = 3$, we have

$$\mathbf{E}\left[X^2\right] \leq (\mathbf{E}[|X|])^{\frac{2}{3}} \left(\mathbf{E}\left[X^4\right]\right)^{\frac{1}{3}} \tag{15.4.46}$$

so

$$\mathbf{E}[|X|] \geq \frac{\left(\mathbf{E}\left[X^2\right]\right)^{\frac{3}{2}}}{\left(\mathbf{E}\left[X^4\right]\right)^{\frac{1}{2}}}. \tag{15.4.47}$$

We computed $\mathbf{E}\left[X^2\right]$ in (15.4.41) to be $\|v\|_2^2$. Now let us calculate $\mathbf{E}\left[X^4\right]$:

$$\begin{aligned}
\mathbf{E}\left[X^4\right] &= \mathbf{E}\left[\left(\sum X_i\right)^4\right] && (15.4.48) \\
&= \mathbf{E}\Bigg[\sum_i X_i^4 + \sum_{i \neq j \neq k} X_i^2 X_j X_k + 6 \sum_{i<j} X_i^2 X_j^2 \\
&\qquad + \sum_{i \neq j} X_i^3 X_j + \sum_{i \neq j \neq k \neq \ell} X_i X_j X_k X_\ell\Bigg] && (15.4.49) \\
\text{(all the terms with an odd power multiple vanish)} &= \sum_i \mathbf{E}\left[X_i^4\right] + 6 \sum_{i<j} \mathbf{E}\left[X_i^2\right] \mathbf{E}\left[X_j^2\right] && (15.4.50) \\
&= \sum_i v_i^4 + 6 \sum_{i<j} v_i^2 v_j^2 && (15.4.51) \\
&\leq 3 \left(\sum_i v_i^2\right)^2 && (15.4.52) \\
&= 3 \|v\|_2^4 && (15.4.53)
\end{aligned}$$

Plugging (15.4.53) into (15.4.47) we prove (15.4.44) and obtain the desired result:

Theorem 15.4.1

$$\ell_2^n \overset{\sqrt{3}}{\hookrightarrow} \ell_1^k$$

(with $k = 2^n$).

15.5 Lower bounds on ℓ_2 embeddings

The Johnson-Lindenstrauss theorem is a rather strong positive result, which makes one wonder if it is possible to do better yet. Alon (2001) showed that one cannot do much better; in this section we will build up to his result.

We start with a general result about the rank of a "perturbed" identity matrix:

Theorem 15.5.1 *Let A be an $n \times n$ symmetric matrix, with $A_{ii} = 1$ and $|A_{ij}| \le \epsilon = \frac{1}{\sqrt{n}}$ for $i \neq j$. Then* $\mathrm{rank}(A) > \frac{n}{2}$.

Proof: Let $\lambda_1, \ldots, \lambda_d$ be the non-zero eigenvalues of A. Then

$$\sum_{i=1}^{d} \lambda_i = \mathrm{tr}(A) = \sum_{i=1}^{n} A_{ii} = n. \tag{15.5.54}$$

Since A is symmetric, we have $A^2 = A^T A$, and the latter's non-zero eigenvalues are $\lambda_1^2, \ldots, \lambda_d^2$. Now

$$\sum_{i=1}^{d} \lambda_i^2 = \mathrm{tr}(A^T A) = \sum_{i,j}^{n} a_{ij}^2 \le n + n(n-1)\epsilon^2 < 2n \tag{15.5.55}$$

where the two inequalities follows from the definition of A. Thus we have

$$2n > \sum_{i=1}^{d} \lambda_i^2 \tag{15.5.56}$$

$$\ge \frac{1}{d}\left(\sum_{i=1}^{d} \lambda_i\right)^2 = \frac{n^2}{d} \tag{15.5.57}$$

where (15.5.57) follows by straightforward application of Jensen's inequality (taking $f(x) = x^2$ as the convex function). But the latter implies $d > \frac{n}{2}$, meaning that there are at least this many non-zero eigenvalues; this proves the theorem. ■

We will need two more lemmas:

Lemma 15.5.2 *If $B = (b_{ij})$ is an $n \times n$ matrix of rank d and $P(\cdot)$ is a polynomial of degree k, then the rank of the matrix $C = (P(b_{ij}))$ is at most $\binom{k+d}{d}$.*

Proof: Let $\vec{v}_1 = (v_{1,j})_{j=1}^n, \ldots, \vec{v}_d = (v_{d,j})_{j=1}^n$ be a basis for the row space of B. Then the vectors of the form $v_j = \prod_i v_{i,j}^{k_i}$, where $k_i \ge 0$ and $\sum_i k_i \le k$ span the rows of C. It is well-known (see, for example, Brualdi, ch. 3) that the number of d-tuples of non-negative k_i's whose sum is k is $\binom{k+d-1}{d-1}$. ■

Lemma 15.5.3 *Let B be an $n \times n$ matrix, with $B_{ii} = 1$ and $|B_{ij}| \le \epsilon$ for $i \neq j$. Then* $d = \mathrm{rank}(B) \ge \Omega\left(\frac{1}{\epsilon^2 \log(1/\epsilon)} \log n\right)$.

Proof: Assume without loss of generality that B is symmetric and that $\epsilon \geq 1/n^\delta$ for some $\delta > 0$ (B can be symmetrized via $\frac{1}{2}(B + B^T)$ with at most a twofold gain in the rank; if $\epsilon < 1/n^\delta$ we can apply the Thm. 15.5.1 to a $\frac{1}{\epsilon} \times \frac{1}{\epsilon}$ submatrix of B to obtain the result). Put $k = \frac{\log n}{2\log(1/\epsilon)}$ and $C = (b_{ij}^k)$ and note that by Lem. 15.5.2, $\mathrm{rank}(C) \leq \binom{d+k}{k} \leq \left(\frac{e(k+d)}{k}\right)^k$. But by Thm. 15.5.1, $\mathrm{rank}(C) \geq n/2$. Thus

$$d \geq (1 + o(\delta)) \frac{\log n}{2e \log(1/\epsilon)} \left(\frac{1}{\epsilon^2} - e\right). \tag{15.5.58}$$

■

Finally, we have

Theorem 15.5.4 *Suppose $\vec{v}^{(0)}, \ldots, \vec{v}^{(n)} \in \mathbb{R}^k$ such that*

$$1 \leq \left\| v^{(i)} - v^{(j)} \right\| \leq 1 + \epsilon. \tag{15.5.59}$$

Then $k \geq O\left(\frac{\log n}{\epsilon^2 \log \frac{1}{\epsilon}}\right)$.

Proof: Put $\vec{v}^{(0)}$ at the origin and shift all other points by at most ϵ to ensure that $\left\| v^{(i)} - v^{(0)} \right\| = 1$ for all $i > 0$. By the triangle inequality, after the shift the pairwise distances are still $1 + O(\epsilon)$. Therefore, if $C = (c_{ij}) = \left\langle v^{(i)}, v^{(j)} \right\rangle$, then C is $n \times n$, $c_{ii} = 1$, and for $i \neq j$, $c_{ij} = \frac{1}{2} + O(\epsilon)$, and $\mathrm{rank}(C) \leq k$. So the rank of $B = 2C - J$ is at most $k + 1$, where J is the all-1's matrix. By Lemma 15.5.3, $\mathrm{rank}(B) \geq \Omega\left(\frac{1}{\epsilon^2 \log(1/\epsilon)} \log n\right)$, supplying the requisite lower bound for k. ■

References

Achlioptas, D. (2000). Database friendly random projections. In *Proceedings of the Twentieth ACM Symposium on Principles of Database Systems*.

Alon, N. (2001). Problems and results in extremal combinatorics, I, Discrete Math., in press.

Brinkman, B. and Charikar, M. S. (2003). On the impossibility of dimension reduction in ℓ_1. In *Proceedings of the 44th Annual IEEE Symposium on Foundations of Computer Science*, pages 514–523.

Brualdi R. A. (1999) Introductory Combinatorics. Prentice Hall.

Engebretsen, L., Indyk, P., and O'Donnell, R. (2002). Derandomized dimensionality reduction with applications. In *Proceedings of the 13th Annual ACM-SIAM Symposium on Discrete Algorithms*, pages 715–712.

Indyk, P. and Motwani, R. (1998). Approximate nearest neighbors: Towards removing the curse of dimensionality. In *Proceedings of the 30th Annual ACM Symposium on Theory of Computing*, pages 604–613.

Johnson, W. B. and Lindenstrauss, J. (1984). Extensions of Lipschitz maps into a Hilbert space. *Contemporary Mathematics*, 26:189–206.

Lee, J. R. and Naor, A. (to appear). Embedding the diamond graph in ℓ_p and dimension reduction in ℓ_1. In *Geometric and Functional Analysis (GAFA)*.

Sivakumar, D. (2002). Algorithmic derandomization via complexity theory. In *Proceedings of the thiry-fourth annual ACM symposium on Theory of computing*, pages 619–626. ACM Press.

Spivak, M. (1965). Calculus on Manifolds. W.A. Benjamin, Inc.