Mittagsseminar

Dimension reduction with the Johnson-Lindenstrauss Lemma

Sebastian Stich

ETH Zürich

September 29, 2011

Abstract

The Johnson-Lindenstrauss Lemma asserts that a set of n points in any Euclidean space can be mapped to an Euclidean space of dimension $k = O(\epsilon^{-2} \log n)$ so that all distances are preserved up to a multiplicative factor between $(1 - \epsilon)$ and $(1 + \epsilon)$. Known proofs obtain such a mapping as a linear map $\mathbb{R}^n \to \mathbb{R}^k$ with a suitable random matrix U. The structure of U can be surprisingly simple, e.g. independent Gaussian entries (Indyk and Motwani, 1998), independent -1,1 entries (Achlioptas 2001) and even sparse (Ailon and Chazelle 2006, Matoušek 2008).

We give an overview of these results and present an elementary proof of the result obtained by Indyk and Motwani.

Table of contents

Mittagsseminar

Dimension reduction with the Johnson-Lindenstrauss Lemma

Sebastian Stich

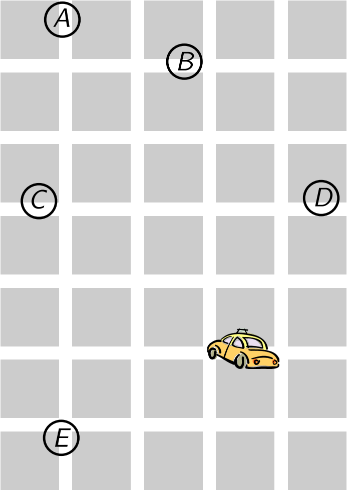
ETH Zürich

September 29, 2011

This talk

Main result	W.B. Johnson and J. Lindenstrauss (1984)
This talk	General overview & proof of the result of Indyk and Motwani (1998)
Motivated by	A. Adoni (2011), Madalgo & CTIC Summer School 2011.

The CAB-Problem



Problem in CAB-space

Diameter of set S of n points in L_1^d.

$$\|x\|_1 = \sum_{i=1}^{d} |x_i|$$

- Trivial algorithm: $O(dn^2)$
- Solution in time $O(2^d n)$ by embedding $f: L_1^d \to L_\infty^{2^d}$

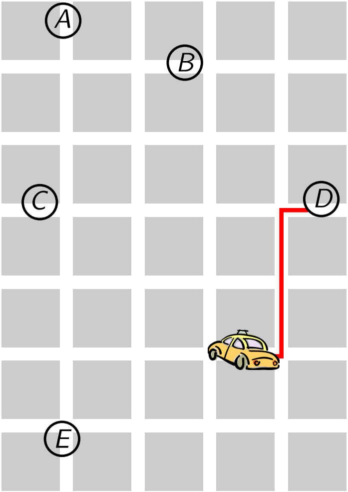
$$\|x-y\|_1 = \|f(x) - f(y)\|_\infty \quad \forall x, y \in S$$

Problem in 'Hyper-CAB'-space

Diameter of set S, of n points in $L_\infty^{2^d}$.

$$\|x\|_\infty = \max_i |x_i|$$

The CAB-Problem



Problem in CAB-space

Diameter of set S of n points in L_1^d.

$$\|x\|_1 = \sum_{i=1}^{d} |x_i|$$

- Trivial algorithm: $O(dn^2)$
- Solution in time $O(2^d n)$ by embedding $f : L_1^d \to L_\infty^{2^d}$

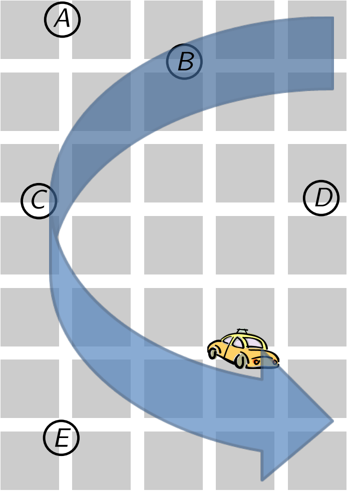
$$\|x - y\|_1 = \|f(x) - f(y)\|_\infty \quad \forall x, y \in S$$

Problem in 'Hyper-CAB'-space

Diameter of set S, of n points in $L_\infty^{2^d}$.

$$\|x\|_\infty = \max_i |x_i|$$

The CAB-Problem



Problem in CAB-space

Diameter of set S of n points in L_1^d.

$$\|x\|_1 = \sum_{i=1}^{d} |x_i|$$

- Trivial algorithm: $O(dn^2)$
- Solution in time $O(2^d n)$ by embedding $f : L_1^d \to L_\infty^{2^d}$

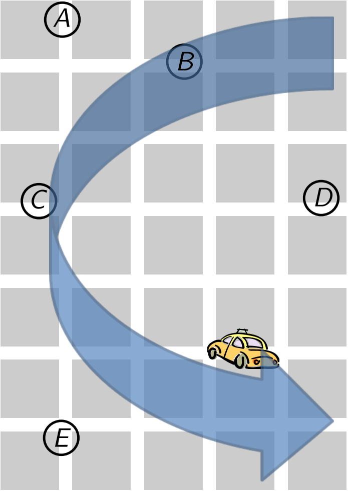
$$\|x - y\|_1 = \|f(x) - f(y)\|_\infty \quad \forall x, y \in S$$

Problem in 'Hyper-CAB'-space

Diameter of set S, of n points in $L_\infty^{2^d}$.

$$\|x\|_\infty = \max_i |x_i|$$

The CAB-Problem



Problem in CAB-space

Diameter of set S of n points in L_1^d.

$$\|x\|_1 = \sum_{i=1}^{d} |x_i|$$

- Trivial algorithm: $O(dn^2)$
- Solution in time $O(2^d n)$ by embedding $f : L_1^d \to L_\infty^{2^d}$

$$\|x-y\|_1 = \|f(x) - f(y)\|_\infty \quad \forall x, y \in S$$

Problem in 'Hyper-CAB'-space

Diameter of set S, of n points in $L_\infty^{2^d}$.

$$\|x\|_\infty = \max_i |x_i|$$

Solution in L_∞^k is easy

$$\begin{aligned}\operatorname{diam}(S) &= \max_{x,y\in S} \|x-y\|_\infty \\ &= \max_{x,y\in S} \max_i |x_i - y_i| \\ &= \max_i \max_{x,y\in S} |x_i - y_i|\end{aligned}$$

- Diameter can be found in time $O(dn)$
- Embedding f and the solution in L_∞ gives an $O(2^d n)$ algorithm for the problem in L_1^d.

The Embedding $f: L_1^d \to L_\infty^{2^d}$

$d = 2$; define $f(x)$ as follows:

$$f\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} := \begin{pmatrix} x_1 + x_2 \\ x_1 - x_2 \\ -x_1 + x_2 \\ -x_1 - x_2 \end{pmatrix}$$

And

$$\begin{aligned} \|x\|_1 &= |x_1| + |x_2| \\ &= \max\{x_1 + x_2, x_1 - x_2, -x_1 + x_2, -x_1 - x_2\} \\ &= \|f(x)\|_\infty \end{aligned}$$

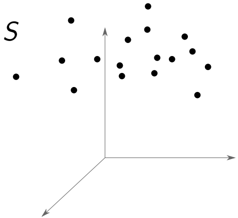
Embeddings in spaces of lower dimension

Want we want

Given set S of n points in L_2^d, $\epsilon > 0$,
a mapping $f\colon L_2^d \to L_2^k$, $k = k(\epsilon, n) < d$ with

$$(1-\epsilon)\,\|x-y\|_2 \leq \|f(x)-f(y)\|_2 \leq (1+\epsilon)\,\|x-y\|_2 \quad \forall x, y \in S\,.$$

(W.l.o.g. $d = n$.)



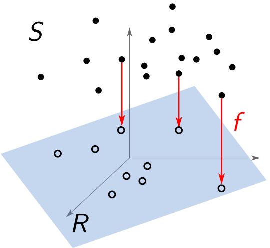
Embeddings in spaces of lower dimension

Want we want

Given set S of n points in L_2^d, $\epsilon > 0$,
a mapping $f\colon L_2^d \to L_2^k$, $k = k(\epsilon, n) < d$ with

$$(1-\epsilon)\left\|x-y\right\|_2 \leq \left\|f(x)-f(y)\right\|_2 \leq (1+\epsilon)\left\|x-y\right\|_2 \quad \forall x, y \in S\,.$$

(W.l.o.g. $d = n$.)



Johnson-Lindenstrauss Lemma

Theorem (Johnson-Lindenstrauss 1984)

Let S be a set of n points in $\mathbb{R}^n$, $\epsilon \in (0, \frac{1}{2}]$ and integer $k = O(\log n/\epsilon^2)$. Then there exists a (linear) mapping $T : \mathbb{R}^n \to \mathbb{R}^k$ such that

$$(1-\epsilon)\|x-y\|_2 \leq \|T(x)-T(y)\|_2 \leq (1+\epsilon)\|x-y\|_2$$

$\forall x, y \in S$.

Approximate nearest neighbour

ϵ-approximate nearest neighbour

Given set S of n points in $\mathbb{R}^d$, find the point $x \in S$ which is ϵ-closest to a query point q:

$$\|x - q\| \leq (1 + \epsilon)\|y - q\|, \quad \forall y \in S.$$

Solution 1: Store all points, for query q calculate distance to all points.

- Space: $O(dn)$
- Query-time: $O(dn)$

Solution 2: Use more space. (Indyk, Motwani 1998)

- Space: $O\left(n\epsilon^{-d}\right)$
- Query-time: $O(d)$

Approximate nearest neighbour

ϵ-approximate nearest neighbour

Given set S of n points in $\mathbb{R}^d$, find the point $x \in S$ which is ϵ-closest to a query point q:

$$\|x - q\| \leq (1 + \epsilon) \|y - q\| , \quad \forall y \in S .$$

Solution 1: Store all points, for query q calculate distance to all points.

- Space: $O(dn)$
- Query-time: $O(dn)$

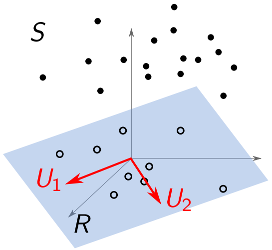
Solution 2: Use more space. (Indyk, Motwani 1998)

- Space: $O\left(n\epsilon^{-d}\right)$
- Query-time: $O(d)$

Solution 2 + Lemma:

- Space: $n^{O(\epsilon^{-2}\log \epsilon^{-1})}$
- Query-time: $O\left(d \log n/\epsilon^{-2}\right)$

Johnson Lindenstrauss 1984



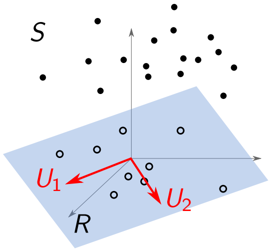
Linear Map T

- Choose random subspace R of dimension k uniformly.

$R = \text{span}\{U_1, \ldots, U_k\}$, $U_i \perp U_j$ $(i \neq j)$, $\alpha \in \mathbb{R}^n$

$$T(\alpha)_i = \sqrt{\frac{n}{k}} U_i^T \alpha$$

Johnson Lindenstrauss 1984



Linear Map T

- Choose random subspace R of dimension k uniformly.

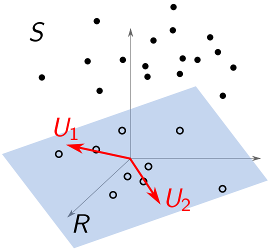
$R = \text{span}\{U_1, \dots, U_k\}$, $U_i \perp U_j$ $(i \neq j)$, $\alpha \in \mathbb{R}^n$

$$T(\alpha)_i = \sqrt{\frac{n}{k}} U_i^T \alpha$$

Proof idea:

- $E\left[\|T(\alpha)\|_2\right] = \|\alpha\|_2$ $\left(\text{or: } E\left[\|T(\alpha)\|_2^2\right] = \|\alpha\|_2^2\right)$
- $\|T(\alpha)\|_2$ tightly concentrated around mean $\left(\text{resp.: } \|T(\alpha)\|_2^2\right)$
- apply union bound
- Rotation invariance: fix $U_i = e_i$, $i = 1, \dots k$, choose α at random from sphere S^{n-1} (w.l.o.g. $\|\alpha\|_2 = 1$).

Indyk and Motwani 1998



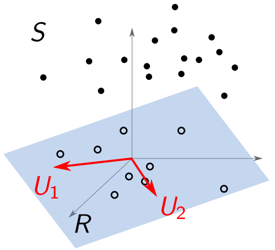
Linear Map T

- Choose set of k unit length vectors

$R = \text{span}\{U_1, \ldots, U_k\}$, ~~$U_i \perp U_j \ (i \neq j)$~~, $\alpha \in \mathbb{R}^n$

$$T(\alpha)_i = \sqrt{\frac{n}{k}} U_i^T \alpha$$

Indyk and Motwani 1998



Linear Map T

- Choose set of k Gaussian vectors

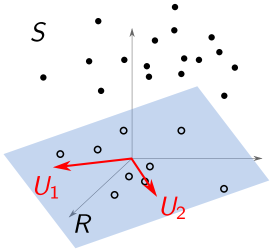
$R = \text{span}\{U_1, \ldots, U_k\}$, $U_{ij} \sim \mathcal{N}(0,1)$, $\alpha \in \mathbb{R}^n$

$$T(\alpha)_i = \frac{1}{\sqrt{k}} U_i^T \alpha$$

We have

$$\mathbb{E}[U_{ij}] = 0 \qquad \mathbb{E}\left[U_{ij}^2\right] = \text{Var}[U_{ij}] = 1$$

Indyk and Motwani 1998



Linear Map T

- Choose set of k Gaussian vectors

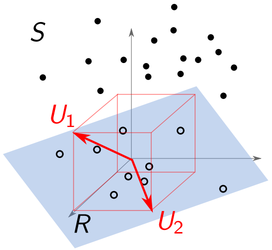
$R = \text{span}\{U_1, \ldots, U_k\}$, $U_{ij} \sim \mathcal{N}(0,1)$, $\alpha \in \mathbb{R}^n$

$$T(\alpha)_i = \frac{1}{\sqrt{k}} U_i^T \alpha$$

- Rotation and scale invariance: fix $\alpha = (1, 0, \ldots, 0)$

$$\mathbb{E}_U \left[\| T(\alpha) \|_2^2 \right] = \mathbb{E}_U \left[\frac{1}{k} \sum_i U_{i1}^2 \right] = 1 \qquad (\mathbb{E}\left[U_{i1}^2\right] = 1)$$

Achlioptas 2003



Linear Map T

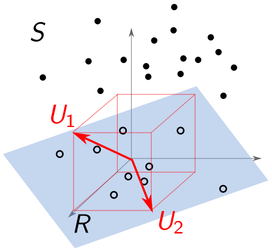
- Choose set of k $-1/1$ vectors

$R = \text{span}\{U_1, \ldots, U_k\}$, $\alpha \in \mathbb{R}^n$

$$\mathbb{P}[U_{ij} = 1] = \mathbb{P}[U_{ij} = -1] = \frac{1}{2}$$

$$T(\alpha)_i = \frac{1}{\sqrt{k}} U_i^T \alpha$$

Achlioptas 2003



Linear Map T

- Choose set of k $-1/1$ vectors

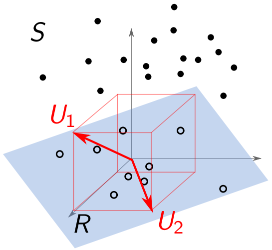
$R = \text{span}\{U_1, \ldots, U_k\}$, $\alpha \in \mathbb{R}^n$

$$\mathbb{P}[U_{ij} = 1] = \mathbb{P}[U_{ij} = -1] = \frac{1}{2}$$

$$T(\alpha)_i = \frac{1}{\sqrt{k}} U_i^T \alpha$$

Again $\mathbb{E}[U_{ij}] = 0$ $\text{Var}[U_{ij}] = 1$

Achlioptas 2003



Linear Map T

- Choose set of k $-1/1$ vectors

$R = \text{span}\{U_1, \dots, U_k\}$, $\alpha \in \mathbb{R}^n$

$$\mathbb{P}[U_{ij} = 1] = \mathbb{P}[U_{ij} = -1] = \frac{1}{2}$$

$$T(\alpha)_i = \frac{1}{\sqrt{k}} U_i^T \alpha$$

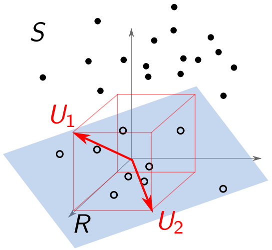
Again $\mathbb{E}[U_{ij}] = 0$ $\text{Var}[U_{ij}] = 1$

Also distribution

$$\mathbb{P}\left[U_{ij} = \sqrt{3}\right] = \mathbb{P}\left[U_{ij} = -\sqrt{3}\right] = \frac{1}{6}, \quad \mathbb{P}[U_{ij} = 0] = \frac{2}{3},$$

works: $\frac{2}{3}$ of the entries of U are zeros

Achlioptas 2003



Linear Map T

- Choose set of k $-1/1$ vectors

$R = \text{span}\{U_1, \ldots, U_k\}$, $\alpha \in \mathbb{R}^n$

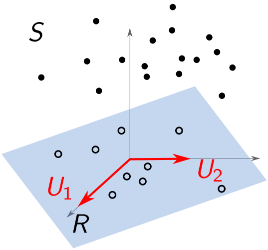
$$\mathbb{P}[U_{ij} = 1] = \mathbb{P}[U_{ij} = -1] = \frac{1}{2}$$

$$T(\alpha)_i = \frac{1}{\sqrt{k}} U_i^T \alpha$$

For $\|\alpha\|_2 = 1$ we calculate

$$\mathbb{E}\left[\|T(\alpha)\|_2^2\right] = \mathbb{E}\left[\frac{1}{k}\sum_i \left(\sum_j U_{ij}\alpha_j\right)^2\right] = \mathbb{E}\left[\frac{1}{k}\sum_{ij} U_{ij}^2 \alpha_j^2\right] = 1$$

Ailon and Chazelle 2006

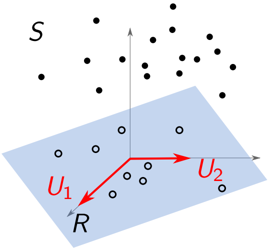


Linear Map T

- Sparse U is possible

$$T(\alpha)_i = \frac{1}{\sqrt{k}} U_i^T \alpha$$

Ailon and Chazelle 2006



Linear Map T

- Sparse U is possible

$$T(\alpha)_i = \frac{1}{\sqrt{k}} U_i^T \alpha$$

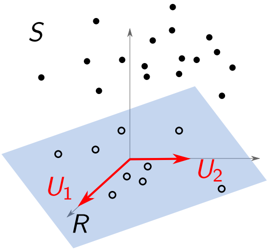
But... for 'sparse' vectors, like

$$\alpha = (1, 0, \ldots, 0) \text{ or } \alpha' = (0.99, 0.001.., \ldots, 0.001..)$$

concentration is not good enough.

Need to assume $\frac{1}{\sqrt{n}} \leq \|\alpha\|_\infty \leq \frac{1}{\sqrt{n}} + \delta$, ($\delta$ 'small'),
or change the statement $k = k(\epsilon, \delta)$.

Matoušek 2008



Linear Map T

- Sparse U is possible

$q \leq 1$ controlling the sparsity

$$\mathbb{P}\left[U_{ij} = \frac{1}{\sqrt{q}}\right] = \mathbb{P}\left[U_{ij} = -\frac{1}{\sqrt{q}}\right] = \frac{1}{2}q$$
$$\mathbb{P}\left[U_{ij} = 0\right] = 1 - q$$

$$T(\alpha)_i = \frac{1}{\sqrt{k}} U_i^T \alpha$$

Concentration bound

Consider distribution s.t. $\mathbb{E}\left[U_{ij}\right] = 0, \mathrm{Var}\left[U_{ij}\right] = 1$, satisfying

Concentration bound

For $\epsilon \in (0, \frac{1}{4}]$ and $k = (3 \log n + \log 2)/\epsilon^2$, $\alpha \in \mathbb{R}^n$

$$\mathbb{P}\left[(1-\epsilon)\|\alpha\|_2 \leq \|T(\alpha)\|_2 \leq (1+\epsilon)\|\alpha\|_2\right] \geq 1 - \frac{1}{n^2}$$

Applying the union bound on all $\binom{n}{2}$ pairs of points $x, y \in S$, yields

$$\mathbb{P}\left[\text{distortion violated for any } x, y \in S\right] \leq \binom{n}{2}\frac{1}{n^2} < \frac{1}{2}$$

Gaussian variables

Markov ($a > 0$)

$$\mathbb{P}\left[|X| \geq a\right] \leq \frac{\mathbb{E}\left[|X|\right]}{a}$$

Lemma 1 ($\lambda > 0$)

$$\mathbb{E}\left[e^{\lambda U_{ij}^2}\right] = \frac{1}{\sqrt{1-2\lambda}}$$

$U_{ij} \sim \mathcal{N}(0,1)$; Rotation invariance: $\alpha = (1, 0, \ldots, 0)$, $T(\alpha) = \frac{1}{\sqrt{k}} \sum_i U_{i1}$

Gaussian variables

Markov ($a > 0$)

$$\mathbb{P}\left[|X| \geq a\right] \leq \frac{\mathbb{E}\left[|X|\right]}{a}$$

Lemma 1 ($\lambda > 0$)

$$\mathbb{E}\left[e^{\lambda U_{ij}^2}\right] = \frac{1}{\sqrt{1-2\lambda}}$$

$U_{ij} \sim \mathcal{N}(0,1)$; Rotation invariance: $\alpha = (1, 0, \ldots, 0)$, $T(\alpha) = \frac{1}{\sqrt{k}} \sum_i U_{i1}$

$$\mathbb{P}\left[\|T(\alpha)\|_2 \geq 1+\epsilon\right] \leq \mathbb{P}\left[\|T(\alpha)\|_2^2 \geq 1+2\epsilon\right] = \mathbb{P}\left[\sum_i U_{i1}^2 \geq (1+2\epsilon)k\right]$$

Gaussian variables

Markov ($a > 0$)

$$\mathbb{P}\left[|X| \geq a\right] \leq \frac{\mathbb{E}\left[|X|\right]}{a}$$

Lemma 1 ($\lambda > 0$)

$$\mathbb{E}\left[e^{\lambda U_{ij}^2}\right] = \frac{1}{\sqrt{1-2\lambda}}$$

$U_{ij} \sim \mathcal{N}(0,1)$; Rotation invariance: $\alpha = (1,0,\dots,0)$, $T(\alpha) = \frac{1}{\sqrt{k}} \sum_i U_{i1}$

$$\mathbb{P}\left[\|T(\alpha)\|_2 \geq 1+\epsilon\right] \leq \mathbb{P}\left[\|T(\alpha)\|_2^2 \geq 1+2\epsilon\right] = \mathbb{P}\left[\sum_i U_{i1}^2 \geq (1+2\epsilon)k\right]$$

$$= \mathbb{P}\left[e^{\lambda \sum_i U_{i1}^2} \geq e^{\lambda(1+2\epsilon)k}\right]$$

Gaussian variables

Markov ($a > 0$)

$$\mathbb{P}[|X| \geq a] \leq \frac{\mathbb{E}[|X|]}{a}$$

Lemma 1 ($\lambda > 0$)

$$\mathbb{E}\left[e^{\lambda U_{ij}^2}\right] = \frac{1}{\sqrt{1-2\lambda}}$$

$U_{ij} \sim \mathcal{N}(0,1)$; Rotation invariance: $\alpha = (1,0,\dots,0)$, $T(\alpha) = \frac{1}{\sqrt{k}} \sum_i U_{i1}$

$$\mathbb{P}\left[\|T(\alpha)\|_2 \geq 1+\epsilon\right] \leq \mathbb{P}\left[\|T(\alpha)\|_2^2 \geq 1+2\epsilon\right] = \mathbb{P}\left[\sum_i U_{i1}^2 \geq (1+2\epsilon)k\right]$$

$$= \mathbb{P}\left[e^{\lambda \sum_i U_{i1}^2} \geq e^{\lambda(1+2\epsilon)k}\right]$$

$$(\text{Markov}) \leq e^{-\lambda(1+2\epsilon)k}\,\mathbb{E}\left[e^{\lambda \sum_i U_{i1}^2}\right]$$

Gaussian variables

Markov ($a > 0$)

$$\mathbb{P}\left[|X| \geq a\right] \leq \frac{\mathbb{E}\left[|X|\right]}{a}$$

Lemma 1 ($\lambda > 0$)

$$\mathbb{E}\left[e^{\lambda U_{ij}^2}\right] = \frac{1}{\sqrt{1-2\lambda}}$$

$U_{ij} \sim \mathcal{N}(0,1)$; Rotation invariance: $\alpha = (1,0,\ldots,0)$, $T(\alpha) = \frac{1}{\sqrt{k}} \sum_i U_{i1}$

$$\mathbb{P}\left[\|T(\alpha)\|_2 \geq 1+\epsilon\right] \leq \mathbb{P}\left[\|T(\alpha)\|_2^2 \geq 1+2\epsilon\right] = \mathbb{P}\left[\sum_i U_{i1}^2 \geq (1+2\epsilon)k\right]$$

$$= \mathbb{P}\left[e^{\lambda \sum_i U_{i1}^2} \geq e^{\lambda(1+2\epsilon)k}\right]$$

$$(\text{Markov}) \leq e^{-\lambda(1+2\epsilon)k}\, \mathbb{E}\left[e^{\lambda \sum_i U_{i1}^2}\right]$$

$$(\text{independence}) = e^{-\lambda(1+2\epsilon)k} \left(\mathbb{E}\left[e^{\lambda U_{11}}\right]\right)^k$$

Gaussian variables

Markov ($a > 0$)

$$\mathbb{P}\left[|X| \geq a\right] \leq \frac{\mathbb{E}\left[|X|\right]}{a}$$

Lemma 1 ($\lambda > 0$)

$$\mathbb{E}\left[e^{\lambda U_{ij}^2}\right] = \frac{1}{\sqrt{1-2\lambda}}$$

$U_{ij} \sim \mathcal{N}(0,1)$; Rotation invariance: $\alpha = (1, 0, \ldots, 0)$, $T(\alpha) = \frac{1}{\sqrt{k}} \sum_i U_{i1}$

$$\mathbb{P}\left[\|T(\alpha)\|_2 \geq 1+\epsilon\right] \leq \mathbb{P}\left[\|T(\alpha)\|_2^2 \geq 1+2\epsilon\right] = \mathbb{P}\left[\sum_i U_{i1}^2 \geq (1+2\epsilon)k\right]$$

$$= \mathbb{P}\left[e^{\lambda \sum_i U_{i1}^2} \geq e^{\lambda(1+2\epsilon)k}\right]$$

$$(\text{Markov}) \leq e^{-\lambda(1+2\epsilon)k}\, \mathbb{E}\left[e^{\lambda \sum_i U_{i1}^2}\right]$$

$$(\text{independence}) = e^{-\lambda(1+2\epsilon)k} \left(\mathbb{E}\left[e^{\lambda U_{11}}\right]\right)^k$$

$$(\text{Lemma 1}) = \left(\frac{e^{-\lambda(1+2\epsilon)}}{\sqrt{1-2\lambda}}\right)^k$$

Gaussian variables

Markov ($a > 0$)

$$\mathbb{P}\left[|X| \geq a\right] \leq \frac{\mathbb{E}\left[|X|\right]}{a}$$

Lemma 1 ($\lambda > 0$)

$$\mathbb{E}\left[e^{\lambda U_{ij}^2}\right] = \frac{1}{\sqrt{1-2\lambda}}$$

$U_{ij} \sim \mathcal{N}(0,1)$; Rotation invariance: $\alpha = (1, 0, \ldots, 0)$, $T(\alpha) = \frac{1}{\sqrt{k}} \sum_i U_{i1}$

$$\mathbb{P}\left[\|T(\alpha)\|_2 \geq 1+\epsilon\right] \leq \mathbb{P}\left[\|T(\alpha)\|_2^2 \geq 1+2\epsilon\right] = \mathbb{P}\left[\sum_i U_{i1}^2 \geq (1+2\epsilon)k\right]$$

$$= \mathbb{P}\left[e^{\lambda \sum_i U_{i1}^2} \geq e^{\lambda(1+2\epsilon)k}\right]$$

$$(\text{Markov}) \leq e^{-\lambda(1+2\epsilon)k}\, \mathbb{E}\left[e^{\lambda \sum_i U_{i1}^2}\right]$$

$$(\text{independence}) = e^{-\lambda(1+2\epsilon)k} \left(\mathbb{E}\left[e^{\lambda U_{11}}\right]\right)^k$$

$$(\text{Lemma 1}) = \left(\frac{e^{-\lambda(1+2\epsilon)}}{\sqrt{1-2\lambda}}\right)^k$$

$$\left(\lambda = \frac{\epsilon}{1+2\epsilon}\right) \leq e^{-\frac{2}{3}k\epsilon^{-2}} \leq \frac{1}{2n^2}.$$

Remarks and open questions

- The conditions $\mathbb{E}[U_{ij}] = 0$, $\text{Var}[U_{ij}] = 1$ and *subgaussian tail* $\mathbb{P}[\pm U_{ij} > \lambda] \leq e^{-a\lambda^2}$ are sufficient. (Matoušek 2008)
- Dependence of k on (ϵ, n) is optimal up constant factors. (Alon 2003)
- Structural parameter of S instead of $(\log n)$ (Klartag, Mendelson 2005)

- No such result for L_1^n, $k = n^{\Omega\left(\frac{1}{(1+\epsilon)^2}\right)}$ (Brinkman, Charikar 2003).
- Embedding of L_1^n in space with different norm?

References

- W.B. Johnson and J. Lindenstrauss. Extensions of Lipschitz mappings into a Hilbert space, 1984.
- P. Indyk and R. Motwani, Approximate nearest neighbors: Towards removing the curse of dimensionality, 1998.
- D. Achlioptas, Database-friendly random projections: Johnson-Lindenstrauss with binary coins, 2003.
- N. Ailon and B. Chazelle, Approximate nearest neighbors and the fast Johnson-Lindenstrauss transform, 2006.
- J. Matoušek, On Variants of the Johnson-Lindenstrauss Lemma, 2008.

$-1, 1$ random Variables (I)

$$\mathbb{P}[U_{ij} = 1] = \mathbb{P}[U_{ij} = -1] = \frac{1}{2}$$

$$\text{Define } Y_i := \sum_j U_{ij}\alpha_j$$

Lemma 2 ($\lambda > 0$)

$$\mathbb{E}\left[e^{\lambda Y_i^2}\right] \leq e^{C\lambda^2}$$

See for instance (Matoušek 2008).

Now the proof is the same as for the Gaussian case.

$-1, 1$ random Variables (II)

For Lemma 2 you need:

- Chernoff bound
- (And Lemma 3)

Lemma 3

$$\mathbb{E}\left[Y_i^4\right] \leq 3$$

Proof of Lemma 3

$$\mathbb{E}\left[Y_i^4\right] = \mathbb{E}\left[\sum_{|\iota|=4} \binom{4}{\iota} Y_i^{\iota}\right] = \mathbb{E}\left[\sum_j \alpha_j^4 Uij^4 + 3\sum_{k \neq l} \alpha_k^2 \alpha_l^2 U_{ik}^2 U_{il}^2\right]$$

$$\leq \mathbb{E}\left[3\sum_{k,l} \alpha_k^2 \alpha_l^2 U_{ik}^2 U_{il}^2\right] = 3\sum_k \alpha_k^2 \sum_l \alpha_l^2 = 3 \qquad \|\alpha\| = 1$$