CS 761 : Randomized Algorithms , Spring 2023 , Waterloo

Lecture 5: Dimension Reduction and Compressed Sensing

An important application of Chernoff bound is for dimension reduction of a point set while preserving distances. Similar technique can be used to construct a good sensing matrix for sparse recovery.

Dimension Reduction

Given n points in the Euclidean space, we can always represent the vectors in n-dimensions.

In general, we cannot do better if no distortion is allowed (e.g. the standard basis).

Surprisingly, if we allow just a little distortion, then the number of dimensions can be significantly reduced.

Theorem (Johnson-Lindenstrauss Lemma) Given any $\varepsilon \in (0, \frac{1}{2})$ and any set of points $X = \{x_1, x_2, \ldots, x_n\}$, there exists a map $A : X \to \mathbb{R}^k$ for $k = O(\frac{\log n}{\varepsilon^2})$ such that

$$1 - \varepsilon \leq \frac{\|A(x_i) - A(x_j)\|_2^2}{\|x_i - x_j\|_2^2} \leq 1 + \varepsilon.$$

Algorithm : The construction is very simple. It just projects the points in a random k-dimensional subspace.

Let d be the dimension of the original points.

Let M be a $k \times d$ matrix, such that each entry of M is drawn from the normal $N(0,1)$ distribution (Gaussian random variable with mean 0 and variance 1, with density $f(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}$.)

Define $A(x) = \frac{1}{\sqrt{k}} Mx$. This is efficiently computable.

Since A is a linear transformation (e.g. $A(x) - A(y) = A(x-y)$), the theorem can be reduced to the following.

Lemma If A is constructed by the above algorithm with $k = \Theta\left(\frac{1}{\varepsilon^2} \log\left(\frac{1}{\delta}\right)\right)$, then $\Pr\left(1 - \varepsilon \leq \|A(x)\|_2^2 \leq 1 + \varepsilon\right) \geq 1 - \delta$ for any unit vector $x \in \mathbb{R}^d$ and any $\varepsilon \in (0, \frac{1}{2})$.

First, we see how the lemma implies the theorem.

We set $\delta = \frac{1}{n^2}$ and thus $k = \Theta(\frac{\log n}{\varepsilon^2})$.

For any pair $i \neq j$, the squared length of $x_i - x_j$ is maintained to within $1 \pm \varepsilon$ with probability $\geq 1 - \frac{1}{n^2}$.

By the union bound, the distances of all pairs are maintained to within $1 \pm \varepsilon$ with probability $\geq \frac{1}{2}$.

Henceforth, we focus on proving the lemma.

Proof idea : Consider the elementary unit vector $e_1 = (1, 0, \ldots, 0)$.

Then, Me_1 is just the first column of M, with independent and identical Gaussian values.

We are interested in the length of this column, which is the sum of squares of these Gaussians.

Note that $E[\sum_{i=1}^{k} M_{i,1}^2] = \sum_{i=1}^{k} E[M_{i,1}^2] = k$ as the variance of each M_{ij} is one, and so the expected length of Ae_1 is one as $E[\|Ae_1\|_2^2] = \frac{1}{k} E[\|Me_1\|_2^2] = 1$.

By setting k to be large enough, we expect that the length is highly concentrated around its expectation.

From our intuition of Chernoff bound, if we set $k = O(\frac{\log n}{\varepsilon^2})$, then the error probability is at most $2e^{-\mu\varepsilon^2/3} \le \frac{1}{n^2}$.

Proof The actual proof is similar to the above idea. There are two issues to handle.

1. We cannot assume $x = e_1$, and we need to deal with any x.
2. The standard Chernoff-Hoeffding bound in L03 cannot be directly applied, because the random variables are unbounded (although with a small tail).

The first issue can be taken care of by the nice properties of Gaussian random variables.

Consider an arbitrary entry y_j of the vector Mx for an arbitrary unit vector x.

Then $y_j = \sum_{i=1}^{d} M_{ji} x_i$ where M_{ji} is an $N(0,1)$ random variable.

So, y_j is a sum of Gaussian variable, and it is a well-known fact that y_j is an $N(0, \sum_{i=1}^{d} x_i^2)$ random variable. Since x is a unit vector, y_j is just an $N(0,1)$ random variable.

So, each of the k coordinates of Mx is just independent Gaussian.

By the same argument as in the proof idea, the expected length of $\frac{1}{\sqrt{k}} Mx$ is one.

The second issue requires some work, to derive a Chernoff bound for square of Gaussian random variables.

By elementary calculus, we can compute the moment generating function of the sum of squares of independent Gaussians (i.e. $E[e^{tX^2}] = \frac{1}{\sqrt{1-2t}}$ for $t < \frac{1}{2}$ for $X \sim N(0,1)$) and use it to prove that $\Pr(\|Ax\|_2^2 \ge 1+\varepsilon) \le e^{-k\varepsilon^2/8}$. The details are left as a (harder) exercise.

Similarly, we can bound the lower tail and get a similar result.

So, by setting $k = O(\frac{1}{\varepsilon^2} \ln(\frac{1}{\delta}))$, we have $\Pr(\left| \|Ax\|_2^2 - 1 \right| > \varepsilon) \le \delta$, proving the lemma. □

Remarks: The same result is true even when M is a random ± 1 matrix [Achiloptas].

The proof is more difficult but the algorithm is much easier to implement.

Actually, if we use a random ± 1 matrix, this is similar to what is used in data streaming algorithms.

Applications: One immediate and important application is to do approximate near neighbor search.

A linear scan takes $O(n^2)$ time, but only $O(n \log n)$ time after dimension reduction.

Note that it works for Euclidean distances only (e.g. not for say L_1-distances).

Another application is approximate matrix multiplication.

Given two $n \times n$ matrices A and B, we do dimension reductions on the rows of A and the columns of B to $O(\log n)$ dimensions, so that the product can be done in $O(n^2 \log n)$ time, while each entry is approximately the same as the inner products are approximated with high prob.

More recently, dimension reduction is frequently used in designing fast algorithms for numerical problems as well as combinatorial problems such as graph sparsification.

Compressed Sensing: The Setup

The general setting is that there is an unknown signal $x \in \mathbb{R}^n$, and we would to make "few measurements" of the signal x so that we can "efficiently" and "exactly" recover x.

The measurements that we consider are of the form $\langle a, x \rangle$ where $a \in \mathbb{R}^n$, where we have the freedom to choose $a \in \mathbb{R}^n$ and observe the outcome $\langle a, x \rangle$.

So, if we choose $a_1, a_2, \ldots, a_k \in \mathbb{R}^n$ and make k measurements, this gives us a <u>sensing matrix</u> $A \in \mathbb{R}^{k \times n}$ where the i-th row is a_i and a k-dimensional vector b of outcomes.

Then, the question is to recover x given $Ax = b$.

Of course, we can't hope to exactly recover every x when $k < n$.

The additional assumption that we make is that x is <u>s-sparse</u>, meaning that x has at most s nonzeros (but of course the locations of the nonzeros are unknown to us).

This is a reasonable assumption as most signals are sparse in an appropriate basis.

Now, the hope is that we can set k to be close to s and we can recover any s-sparse signal x exactly and efficiently.

A natural idea is to construct a sensing matrix A so that there is only one s-sparse solution to $Ax = b$.

One sufficient condition to enforce this is to ensure that the sensing matrix A has large <u>kruskal rank</u>.

We say an $m \times n$ matrix has kruskal rank r if every subset of r columns are linearly independent.

<u>Claim</u> If A has kruskal rank at least $2s$, then for any b we have $Ax = b$ for at most one s-sparse x.

<u>Proof</u> Suppose, by contradiction, that $Ax = Ax'$ for two s-sparse vectors $x \neq x'$.

This implies that $A(x - x') = 0$. But note that $x - x'$ is $2s$-sparse, contradicting A has kruskal rank $2s$. □

With this claim, given a sensing matrix A with kruskal rank $2s$, the problem of sparse recovery can be formulated as an ℓ_0-minimization problem $\min \{ \|x\|_0 \mid Ax = b \}$, where $\|x\|_0$ denotes the number of nonzeros in x.

But the problem of this formulation is that we don't know how to solve ℓ_0-minimization efficiently.

In general, given an arbitrary A and b, the ℓ_0-minimization problem is NP-hard.

In our setting, we can design the sensing matrix A, but still it is not clear how to solve the problem directly.

In the following, we will define a stronger property of the sensing matrix called the restricted isometry property, and then show that the ℓ_0-minimization problem can be solved by the ℓ_1-minimization problem efficiently.

Restricted Isometry Property and ℓ_1-Minimization

The following property of the sensing matrix is stronger than the kruskal rank property above.

Restricted isometry property (RIP) A matrix A is (t,ε)-RIP if for any t-sparse vector x with $\|x\|_2 = 1$, it holds that $\|Ax\|_2^2 \in [1-\varepsilon, 1+\varepsilon]$.

This is a stronger property than kruskal rank because not only we require the columns to be linearly independent, but also that any subset of t columns are almost orthogonal.

Note that a (t,ε)-RIP matrix with $\varepsilon < 1$ has kruskal rank at least t, so if we have a $(2s,\varepsilon)$-RIP matrix A with $\varepsilon < 1$ then there is a unique s-sparse solution x to $Ax = b$.

Interestingly, the stronger RIP property allows us to solve the ℓ_0-minimization by solving the ℓ_1 minimization problem $\min_x \{ \|x\|_1 \mid Ax = b \}$.

Proposition If the sensing matrix A is $(3s,\varepsilon)$-RIP for some $\varepsilon \leq 1/9$, then the unique optimal solution to $\min_x \{ \|x\|_0 \mid Ax = b \}$ is also the unique optimal solution to $\min_x \{ \|x\|_1 \mid Ax = b \}$.

Note that the ℓ_1-minimization problem can be solved by linear programming, and so the above proposition gives us a polynomial time algorithm to recover the s-sparse signal from the sensing matrix A and b.

The remaining issue is how to find a good RIP matrix for sparse recovery.

Not surprisingly, we will generate such a matrix randomly, as we expect that random columns (with appropriate distributions) are almost orthogonal to each other.

Lemma There is a randomized algorithm to construct a matrix $A \in \mathbb{R}^{k \times n}$ with $k = \Theta(s \log n)$ so that A is $(3s,\varepsilon)$-RIP for some $\varepsilon \leq 1/9$ with high probability.

To summarize, for compressed sensing, we first generate a $(3s, \varepsilon)$-RIP matrix with $\varepsilon \leq 1/9$ using the lemma, and then we use linear programming to recover the s-sparse signal with the sensing matrix A and the measurement outcomes b whose correctness is by the proposition.

We will prove the proposition and the lemma in the following subsections.

Random Construction of RIP Matrices

Let's prove the lemma first, as it is easier and also it is randomized (while the other step is deterministic).

As discussed earlier, if we generate random columns from an appropriate distribution, then we expect them to be almost orthogonal to each other and thus satisfy the restricted isometry property.

Indeed, exactly the same matrix that we used to prove the Johnson-Lindenstrauss would work!

Construction: $A = \frac{1}{\sqrt{k}} M$ where M is a $k \times n$ matrix where each entry is an independent $N(0,1)$ variable.

Proof plan: Fix a subset of s coordinates, and we will use a union bound over all subsets of s coordinates.

For each subset of s coordinates, we consider an "ε-net" of its unit vectors, a finite subset that "covers" the whole sphere so that each unit vector is "close" to one of the vectors in the ε-net.

Then, we simply use the result proven for dimension reduction to show that the length of each vector in the ε-net is preserved with high probability, and then use it to conclude that the length of all unit vectors is preserved.

This proof technique of using ε-net is quite common and general, so it is good to learn it.

Epsilon-net: Given a metric space (X, d), an ε-net is a subset $N \subseteq X$ such that

(i) $d(x,y) \geq \varepsilon \ \forall x, y \in X$, and (ii) for each $x \in X$ there exists $y \in N$ with $d(x,y) \leq \varepsilon$.

Fact There is an ε-net of the sphere S^{s-1} (under Euclidean distances) with at most $\left(\frac{4}{\varepsilon}\right)^s$ vectors.

Proof We construct an ε-net by a simple greedy algorithm:

If some point $x \in X$ does not satisfy (ii), then we add x to the ε-net N.

We claim that at most $\left(\frac{4}{\varepsilon}\right)^s$ vectors are added to N.

To see this, for each point $x \in N$, define a ball of radius $\varepsilon/2$ around it.

Note that these balls are disjoint by property (i) of ε-net, and they are all contained in a ball of radius $1+\varepsilon$

Since the volume of a ball of radius r scales as r^s, it follows that $|N| \leq \left(\frac{1+\varepsilon}{\varepsilon/2}\right)^s \leq \left(\frac{4}{\varepsilon}\right)^s$. □

If we preserve the length of all vectors in an ε-net, then we preserve the length of all vectors.

Claim Let N be a δ-net of S^{s-1} for $\delta \le \frac{1}{3}$. If $1-\delta \le \|Ax\|_2 \le 1+\delta$ for all vectors in N,

then $1-3\delta \le \|Ax\|_2 \le 1+3\delta$ for all vectors in S^{s-1}.

Proof Let M be the maximum stretch $\max\{\|Ax\|_2 \mid x \in S^{s-1}\}$ and x be a point achieveing the maximum.

Let y be a point in N with $\|x-y\|_2 \le \delta$.

Then, $M = \|Ax\|_2 \le \|Ay\|_2 + \|A(x-y)\|_2 \le 1+\delta + M\delta$, which implies that $M \le \frac{1+\delta}{1-\delta} \le 1+3\delta \le 1+\varepsilon$ for $\delta \le \frac{1}{3}$.

Also, for any $x \in S^{s-1}$, $\|Ax\| \ge \|Ay\| - \|A(x-y)\| \ge 1-\delta - M\delta \ge 1-\delta-(1+3\delta)\delta \ge 1-3\delta$ for $\delta \le \frac{1}{3}$. □

Now we are ready to complete the proof of the lemma about RIP matrices.

Proof Let $A = \frac{1}{\sqrt{k}} M$ where M is a $k \times n$ matrix where each entry is an independent $N(0,1)$ variable.

We need to show that $1-\varepsilon \le \|Ax\|_2 \le 1+\varepsilon$ for any s-sparse unit vector x.

Fix a subset of s coordinates, and consider the s-sparse unit vectors supported in these coordinates.

To do so, let N be an δ-net of the sphere S^{s-1} defined by these coordinates. where we set $\delta := \frac{\varepsilon}{3}$.

For each unit vector $x \in N$, by the lemma in dimension reduction, $1-\delta \le \|Ax\|_2 \le 1+\delta$ holds with

probability at least $1 - e^{-c\delta^2 k}$ for some constant c.

By applying the union bound and the fact that $|N| \le \left(\frac{4}{\delta}\right)^s$, $1-\delta \le \|Ax\|_2 \le 1+\delta$ holds for

all unit vectors in N with probability at least $1 - \left(\frac{4}{\delta}\right)^s e^{-c\delta^2 k} = 1 - e^{s \ln(\frac{4}{\delta}) - c\delta^2 k}$.

If this holds, by the claim, $1-\varepsilon = 1-3\delta \le \|Ax\|_2 \le 1+3\delta = 1+\varepsilon$ holds for all vectors in S^{s-1}.

Therefore, by a union bound over all subsets of s coordinates, $1-\varepsilon \le \|Ax\|_2 \le 1+\varepsilon$ holds

for all s-sparse unit vectors with probability at least $1 - \binom{n}{s}\left(\frac{4}{\delta}\right)^s e^{-c\delta^2 k} = 1 - e^{s\ln n + s\ln(\frac{4}{\delta}) - c\delta^2 k}$.

For a constant $\varepsilon \le \frac{1}{9}$ and thus $\delta = \frac{\varepsilon}{3} \le \frac{1}{27}$, we just need to set $k = \theta(s \cdot \log n)$. □

ℓ_0-Minimization via ℓ_1-Minimization

Here we prove the proposition that using an $(3s, \varepsilon)$-RIP matrix with $\varepsilon \le \frac{1}{9}$, then an optimal solution to the ℓ_1-minimization problem is an optimal solution to the ℓ_0-minimization problem.

Let x^* be the optimal solution to the ℓ_0-minimization problem, and x be an optimal solution to the ℓ_1-minimization problem.

Our goal is to prove that $\Delta := x - x^* = 0$. Note that $A\Delta = A(x-x^*) = 0$.

(If Δ is $3s$-sparse, then we can conclude $0 = \|A\Delta\|_2 \ge (1-\varepsilon)\|\Delta\|_2$ and thus $\Delta = 0$. But we can't assume that Δ is $3s$-sparse.)

Let $S := \text{supp}(x^*)$ be the support of x^*, and $\bar{S}$ be the set of remaining coordinates.

The following claim is the only place that we use that x is an optimizer for the ℓ_1-minimization problem.

Claim $\|\Delta_S\|_1 \ge \|\Delta_{\bar{S}}\|_1$.

Proof Informally, $\Delta_{\bar{S}}$ increases the 1-norm of $x_{\bar{S}}$, so Δ_S must be as large to decrease the 1-norm of x_S, so that x is a solution of minimum 1-norm.

Formally, $\|x^*\|_1 \geq \|x^*+\Delta\|_1 = \|x_S^*+\Delta_S\|_1 + \|\Delta_{\bar{S}}\|_1 \geq \|x_S^*\|_1 - \|\Delta_S\|_1 + \|\Delta_{\bar{S}}\|_1 = \|x^*\|_1 - \|\Delta_S\|_1 + \|\Delta_{\bar{S}}\|_1$, where the last inequality is by triangle inequality. The claim follows by rearranging. $\square$

Next, we sort the coordinates of $\bar{S}$ in Δ in decreasing order of their absolute value, and group them into buckets of $2s$ coordinates, and call them $B_1, B_2, \ldots$, and so on.

We will prove the following claim bounding the "total length of the tail".

<u>Claim</u> $\sum_{j\geq 2} \|\Delta_{B_j}\|_2 \leq \|\Delta_S\|_2 / \sqrt{2}$.

Now, we use the claim to prove that Δ_S must be zero, as otherwise the error cannot be canceled out, because the error in $S \cup B_1$ is preserved by RIP.

Formally,
$$
\begin{aligned}
0 = \|A\Delta\|_2 &\geq \|A\Delta_{S\cup B_1}\|_2 - \sum_{j\geq 2} \|A\Delta_{B_j}\|_2 && \text{// triangle inequality} \\
&\geq (1-\varepsilon)\|\Delta_{S\cup B_1}\|_2 - (1+\varepsilon)\sum_{j\geq 2}\|\Delta_{B_j}\|_2 && \text{// RIP for } 3s\text{-sparse vectors; note that } |B\cup S_1| \leq 3s. \\
&\geq (1-\varepsilon)\|\Delta_S\|_2 - \frac{1+\varepsilon}{\sqrt{2}}\|\Delta_S\|_2 . && \text{// by the claim right above.}
\end{aligned}
$$

This implies that $\Delta_S = 0$ as $1-\varepsilon > \frac{1+\varepsilon}{\sqrt{2}}$ when $\varepsilon \leq 1/9$.

Then, by the previous claim that $\|\Delta_S\|_1 \geq \|\Delta_{\bar{S}}\|_1$, we conclude that $\Delta_{\bar{S}} = 0$ and thus $\Delta = 0$, as desired.

It remains to prove the claim to complete the proof.

<u>Proof of claim</u> For any bucket B_j with $j \geq 2$, each entry of Δ in this bucket is smaller that the smallest entry in B_{j-1}, and hence smaller than the average entry of B_{j-1}.

Since there are $2s$ entries in the bucket B_j, $\|\Delta_{B_j}\|_2^2 \leq 2s \cdot \left(\frac{\|\Delta_{B_{j-1}}\|_1}{2s}\right)^2 \Rightarrow \|\Delta_{B_j}\|_2 \leq \frac{\|\Delta_{B_{j-1}}\|_1}{\sqrt{2s}}$.

Therefore, $\sum_{j\geq 2} \|\Delta_{B_j}\|_2 \leq \sum_{j\geq 2} \frac{\|\Delta_{B_{j-1}}\|_1}{\sqrt{2s}} = \frac{\|\Delta_{\bar{S}}\|_1}{\sqrt{2s}} \leq \frac{\|\Delta_S\|_1}{\sqrt{2s}} \leq \frac{\|\Delta_S\|_2}{\sqrt{2}}$,

where the second last inequality is by the previous claim and the last inequality is by Cauchy-Schwarz. $\square$

This completes the proof of the proposition and thus the result of compressed sensing.

<u>**Reference**</u>: This lecture is **heavily** based on the notes by Anupam Gupta, Lecture 11 of 15-850 Spring 2023 in CMU.

The original work on compressed sensing was done by many people including Donoho, Candes, Tao, Rudelson, Vershynin.

The idea of using ℓ_1-minimization for ℓ_0-minimization was extended in the line of work for matrix completion.