Computational Complexity of Certifying Restricted Isometry Property

Abhiram Natarajan, Yi Wu

Department of Computer Science, Purdue University

November 14, 2017

Introduction

Problem Relevance:

- Tons of data is generated by sensing systems
- Sampling at required rates (Nyquist rate) is impractical
- Construct compressible representations of signals

The sparse vector recovery problem

Sparse Vector Recovery Problem: Given a matrix $A \in \mathbb{R}^{n \times N}$, with $n \ll N$, and a vector $\mathrm{y} \in \mathbb{R}^n$, find a k-sparse vector $\mathrm{x} \in \mathbb{R}^N$ such that

$$\mathrm{y} = A\mathrm{x}$$

There exists efficient algorithm recovering x if A exhibits the **Restricted Isometry Property** *(RIP)*.

Restricted Isometry Property (RIP)

Definition

Given $k < n$ and $0 < \delta < 1$, a matrix $\mathrm{A} \in \mathbb{R}^{n \times N}$ is (k, δ)-RIP if, for any k-sparse vector $\mathrm{x} \in \mathbb{R}^n$,

$$(1 - \delta)\|\mathrm{x}\|_2 \leq \|\mathrm{Ax}\|_2 \leq (1 + \delta)\|\mathrm{x}\|_2$$

Ideally, a matrix that exhibits strong RIP has

- large k (called order)
- small δ (called *RIC*)

RIP and Sparse Recovery

Theorem (Candes, Romberg and Tao, 2005, 2006, 2008)

If A is $(2k, \delta)$-RIP for some $\delta < \sqrt{2} - 1$, we can find an k-sparse x efficiently by solving

$$\min_{\mathrm{a} \in \mathbb{R}^n} \|\mathrm{a}\|_1 \qquad \textit{subject to } A\mathrm{a} = \mathrm{y}$$

The above result just says that N dimensional k-sparse signals can be compressed into n dimensional signals if we use a matrix A that exhibits good RIP.

On Constructing RIP Matrices

For $N = poly(n)$,

- Best deterministic constructions can achieve $k \leq n^{0.50001}$ by Bourgain et al. (2011)
- Can be shown that randomized constructions give $k \in \Omega(n/polylog(n))$ by sampling a random symmetric ± 1 Bernoulli matrix or a random Gaussian matrix, *w.h.p.*.

Randomized constructions are much better than deterministic constructions!

RIP Certification

Definition

(RIP Certification Problem) Given a matrix M

- (Exact Version) Decide whether the matrix satisfies (k, δ)-RIP.
- (Approximate Version) Decide whether a matrix satisfies (k_1, δ_1)-RIP or does not satisfy (k_2, δ_2)-RIP.
  - We only need to have $\delta \leq \sqrt{2} - 1$ for most applications

"...an alternate approach, and one of interest in its own right, is to work on improving the time it takes to verify that a given matrix (possibly one of a special form) obeys the RIP.." – Terry Tao

Types of Exact Problems

In the exact optimization world, there are three kinds of problems

- Decision (e.g., "YES" if $\geq k$ clauses of a SAT instance are satisfiable, "NO" otherwise)
- Computation (e.g., find the max k such that k clauses can be satisfied)
- Search (e.g., find an assignment that satisfies maximum number of clauses)

$$\text{Decision} \equiv_P \text{Computation} \leq_P \text{Search}$$

Bringing in the 'g'

Parallelly, in the approximation world, there are three kinds of problems

- Verification Gap problems (e.g.,"YES" if we can satisfy $\geq k$ clauses, "NO" if we cannot more than $\frac{k}{g}$ clauses)
- Approximate Computation (e.g., find k' such that $k \geq k'$ and $k \leq gk'$ clauses can be satisfied)
- Approximate Search (e.g., find an assignment that satisfies at least $\frac{opt}{g}$ clauses)

Relation Between Approximation Versions

- Verification Gap $\leq_P$ Approx. Computation
  Get k'^* from Approx. Computation; We know that real $k^* \leq gk'^*$, so if $gk'^* \geq k$, say "YES"
- Approx. Computation $\leq_P$ Verification Gap
  Find largest k such that Verification Gap says "YES", return $\frac{k}{g}$ as answer

Verification Gap $\equiv_P$ Approx. Computation $\leq_P$ Approx. Search

Making Strong Hardness Statements

- Decision problems are easiest to work with, that is why we work with Verification Gap problems
- Proving NP hardness of Gap version is a strong statement
- Two kinds of reductions for inapproximability results:
  - Gap-Producing Reduction - No *gap* in original problem
  - Gap-Preserving Reduction - Reduction from one *gap* problem to other

Graph Expansion

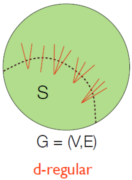
Definition

Given a graph d-regular graph $G(V, E)$, we define

$$\phi_G(S) = \frac{\text{No. of edges going out of } S}{\text{No. of edges incident on vertices of } S}$$

$$= \frac{|E(S, V - S)|}{d \cdot \min(|S|, |V - S|)}$$

$$\phi_G(\delta) = \min_{S:|S| \leq \delta |V|} \Phi_G(S) \qquad (\delta \leq \frac{1}{2})$$



- Expansion of S measures the probability of a random edge cross a set S
- Expansion is a very useful notion

Small-Set-Expansion Conjecture

Definition

SSE(ϵ, δ) problem: Given a graph $G = (V, E)$ of n vertices, and $\epsilon, \delta \leq \frac{1}{2}$, distinguish between the following cases

- *(non-expanding)* $\exists S \subset V$ *with* $|S| = \delta n$ *such that* $\Phi_G(S) \leq \epsilon$
- *(highly expanding)* $\forall S \subset V$ *with* $|S| = \delta n$, $\Phi_G(S) \geq 1 - \epsilon$

Conjecture (Raghavendra and Steurer 2010)

For every $\epsilon > 0$, $\exists 0 \leq \delta \leq \frac{1}{2}$, *such that it is NP-hard to solve SSE(ϵ, δ):*

Importance of SSE Conjecture

- Unique Games Conjecture (UGC) is big open question (Inapproximability results ...); For more, read Khot's survey (2010)
- Consequences of refutation of UGC was poorly understood until SSE
- SSE is more natural and easy to state
- RS (2010) gave a reduction from SSE to UG; Also gave other inapproximability results

Previous Work (1): Hardness of Exact RIP Certification

Exact Decision: Given δ, k, and a matrix M as input, decide if M satisfy (k, δ)-RIP.

- Bandeira et al. (2013) proved that it is NP-hard
- Tillmann and Pfetsch (2014) proved that it is co-NP-hard
- Both results work when $\delta = 1 - o_n(1)$

Previous Work (2): Inapproximability of RIP Certification

Koiran and Zouzias (2011, 2012) show inapproximability results by assuming hardness of hidden clique problem and densest k-subgraph problem

- Most results state that it is hard to distinguish (k, δ_1)-RIP from (k, δ_2)-RIP for some $\delta_1 < \delta_2 \in o_n(1)$
- Exception:
  - No polynomial time algorithm can distinguish matrices that satisfy the $(k, \frac{\kappa}{2})$-RIP from matrices that do not satisfy the (k, κ)-RIP

  where $\kappa \left(\leq \frac{\sqrt{5}}{3}\right)$ is an unknown constant depending on the correctness of certain hardness assumptions of densest k-subgraph.

Previous Work (3): Inapproximability of RIP Certification

- In practice, it is known that an RIP matrix is useful for many applications as long as $\delta \leq \sqrt{2} - 1$
- Their work does not rule out the existense of an algorithm for deciding whether the RIC of a matrix is $\leq \sqrt{2} - 1$. This is because there is no guarantee that $\kappa \in (\sqrt{2} - 1, 2\sqrt{2} - 2)$.

Our Results

Theorem

For any $0 \leq \delta \leq 1$ *and arbitrary large constant* C*, there exists* k *such that, given a matrix* M *it is* Small-Set-Expansion-hard *to distinguish between:*

- *(Highly RIP)* M *is* (k, δ)*-RIP.*
- *(Far away from RIP)* M *is not* $(\frac{k}{C}, 1-\delta)$*-RIP.*

This is the first hardness result that applies for any $0 < \delta < 1$ (including $\sqrt{2} - 1$).

Corollaries

As corollaries, we have that

Corollary

Given a matrix M *and* k*, it is* SMALL-SET-EXPANSION-HARD *to distinguish whether the matrix is* (k, δ)*-RIP or not* $(k, 1-\delta)$*-RIP.*

Corollary

Given a fixed δ *and matrix* M*, it is* SMALL-SET-EXPANSION-HARD *to get a constant approximation for the smallest* k *such that* M *exhibits* (k, δ)*-RIP.*

Proof Overview (1)

- If A is the adjacency matrix of a d regular graph, we consider the matrix M such that $\mathrm{M}^{\mathrm{T}}\mathrm{M} = \mathrm{I} - \frac{1}{d}\mathrm{A} = \mathrm{L}$ for RIP certification
- (Completenss of the Reduction) If there is a small set S with expansion less than ϵ, then $\phi_G(S) = \frac{\|\mathrm{M}\mathrm{x}_S\|_2^2}{\|\mathrm{x}_s\|_2^2} \leq \epsilon$, where $\mathrm{x}_S \in \{0,1\}^n$ is the indicator vector on S. This gives us $\|\mathrm{M}\mathrm{x}_S\|_2 \leq \sqrt{\epsilon}\|\mathrm{x}_S\|_2$, which suggests that M is far away from RIP.

Proof Overview (2)

- (Soundess of the Reduction) show that if $\exists$ a k-sparse $\mathrm{x} \in \mathbb{R}^n$ such that

$$\frac{\mathrm{x}^{\mathrm{T}}\mathrm{M}^{\mathrm{T}}\mathrm{M}\mathrm{x}}{\|\mathrm{x}\|_2^2} \leq (1 - \Omega(1))$$

  then we can find a small set S such that $\phi(S)$ is also bounded away from 1. This uses the Sparse Cheeger's Inequality

Sparse Cheeger's Inequality

We prove the following Cheeger's Inequality on sparse vectors.

Theorem

(Sparse Cheeger's Inequality) Let A *be the adjacency matrix of a d-regular graph G, and* $\mathrm{L} = \mathrm{I} - \frac{1}{d}\mathrm{A}$ *be its normalized Laplacian matrix. For any* $\delta \leq 1/2$*, we have that*

$$\lambda_\delta \leq \phi_G(\delta) \leq \sqrt{(2-\lambda_\delta)\lambda_\delta}$$

where $\lambda_\delta = \min_{\|\mathrm{x}\|_0 \leq \delta|V|} \frac{\mathrm{x}^T\mathrm{L}\mathrm{x}}{\mathrm{x}^T\mathrm{x}}$

Comparison with Cheeger's Inequality

Theorem

Let A be the adjacency matrix of a graph G, and $\mathrm{L} = \mathrm{I} - \frac{1}{d}\mathrm{A}$ be its normalized Laplacian matrix. We have that

$$\frac{\lambda_2}{2} \leq \phi(G) \leq \sqrt{2\lambda_2}$$

where

$$\lambda_2 = \min_{\substack{\mathrm{x} \in \mathbb{R}^n \\ \mathrm{x} \cdot \vec{1} = 0}} \frac{\|\mathrm{x}^T \mathrm{L} \mathrm{x}\|_2}{\|\mathrm{x}\|_2^2}$$

is the second smallest eigenvalue of L.

It must be noted that the relation between λ_δ and $\phi_\delta(G)$ is tighter in this case.

Proof of Sparse Cheeger's Inequality

- Lower bound of $\phi_\delta(G)$ is easy to prove

$$\phi_\delta(G) = \min_{\substack{S \subseteq V \\ S \leq \delta n}} \phi(S) = \min_{\substack{\mathrm{x} \in \{0,1\}^n \\ \|\mathrm{x}\|_0 \leq \delta}} \frac{\mathrm{x}^\mathrm{T} \mathrm{L} \mathrm{x}}{\mathrm{x}^\mathrm{T} \mathrm{x}} \geq \lambda_\delta$$

- Upper bound is called *hard direction*. Here, we assume we are given the vector x that gives us $\frac{\mathrm{x}^\mathrm{T}\mathrm{L}\mathrm{x}}{\mathrm{x}^\mathrm{T}\mathrm{x}} = \lambda_\delta$.
- The same randomized rounding as the proof of Cheeger's Inequality, we can create a cut set in the graph, and that the expansion of the cut is restricted.

Concluding Remarks

Summary:

- We have proved that RIP certification is hard to approximate in a strong sense assuming the SMALL-SET-EXPANSION HYPOTHESIS
- We developed a variant of Cheeger's inequality for sparse vectors

Future directions:

- It will be interesting to see if RIP certification is hard even when the matrix satisfies certain natural properties such as coherence
- It will also be interesting to prove NP/UG-hardness, because correctness of the SMALL-SET-EXPANSION HYPOTHESISis uncertain
- Subexponential algorithm for RIP certification