Limitations in Approximating RIP

Alok Puranik
Mentor: Adrian Vladu

Fifth Annual PRIMES Conference, 2015
Outline

1. Background
   - The Problem
   - Motivation
   - Construction
   - Certification

2. Planted model
   - Planting eigenvalues
   - Analysis
   - Distinguishing random from planted model

3. Hardness of RIP
   - Proof of hardness
   - Reductions

4. Conjectures
RIP Definition

Definition
A vector $x$ is $k$-sparse if it has at most $k$ nonzero components

Definition
An matrix $V$ satisfies the Restricted Isometry Property with order $k$ and Restricted Isometry Constant $\delta$ if for every $k$-sparse vector $x$,

$$(1 - \delta)\|x\|^2 \leq \|Vx\|^2 \leq (1 + \delta)\|x\|^2$$
Alternate Definition

Definition

A matrix $V$ is RIP-\(k, \delta\) if for every submatrix $A$ created by selecting $k$ columns from $V$ and every $k$-dimensional vector $x$,\

\[(1 - \delta)\|x\|^2 \leq \|Ax\|^2 \leq (1 + \delta)\|x\|^2\]
Given a compressible (sparse) signal vector and a few measurements with noise, can we reconstruct the original signal accurately? (Candès and Tao)
Given a compressible (sparse) signal vector and a few measurements with noise, can we reconstruct the original signal accurately? (Candès and Tao)

If the sensing matrix satisfies RIP with $\delta = \sqrt{2} - 1$, we can recover the original
Random Construction

- Draw elements from certain sufficiently concentrated distributions e.g. $\mathcal{N}(0, \frac{1}{\sqrt{n}})$
- This (almost) always works in theory, but is non-deterministic (Baraniuk et al.)
- Deterministic algorithms currently don’t achieve the same bounds
Random construction succeeds with very high probability, but is not guaranteed

A certification algorithm to verify generated matrices would be useful
Naive Algorithm

- To verify RIP-$k, \delta$ for a matrix $V$, check every $k$-column submatrix $A$ of $V$
- Inspect eigenvalues of $A^T A$
- Requires time exponential in $k$
Naive Algorithm

- To verify RIP-$k, \delta$ for a matrix $V$, check every $k$-column submatrix $A$ of $V$
- Inspect eigenvalues of $A^T A$
- Requires time exponential in $k$
- Certification is actually NP-Hard
Adversarial Matrices

- We can alter the generation process to produce matrices that "look" random
- We try to fool a decision algorithm: try to plant a large eigenvalue and break RIP
Breaking RIP with Singular Values

- Large eigenvalues in $V^T V$ correspond to large singular values of $V$
- We leave most of $V$ completely random, fix $k$ columns to have a large singular value

$$V = \begin{bmatrix} Q & QM \end{bmatrix}$$

$$V^T V = \begin{bmatrix} Q^T Q & Q^T QM \\ M^T Q^T Q & M^T Q^T QM \end{bmatrix}$$
We plant a large eigenvalue in $M^T Q^T Q M$

- $M$ must have a large singular value
- We can manipulate the singular value decomposition of $M$:
  1. Decompose random matrix as $U \Sigma V^T$ where $U$ and $V$ are unitary
  2. Construct $\Sigma'$ by setting first diagonal entry of $\Sigma$ to a planted singular value, setting the rest to something convenient
  3. Reconstruct $M$ as $U \Sigma' V^T$
Statistical Analysis

- Elements of the matrix $Q$ are independent, identically distributed Gaussian, $\mathcal{N}\left(0, \frac{1}{\sqrt{n}}\right)$
- Distribution of elements of $Q^TQ$ is highly concentrated: within $O\left(\frac{\log n}{n}\right)$
- Elements of $M^TQ^TQM$ follow the same bounds with high probability
Limitations in Approximating RIP
Planted model
Distinguishing random from planted model

Distinguishing Random from Planted

- Inspecting elements directly give no indication
- Inspecting eigenvalues of full matrix detects this implementation of planted model
An oracle that certifies RIP would be able to enable an efficient solution to Spark, and therefore subset sum.

**Theorem (Bandeira et al.)**

*Certifying RIP for arbitrary $k$ and $\delta$ is NP-Hard*
Limitations of proof

- Weak result: shows hardness only for arbitrary matrix
- Says nothing about approximability
Reductions

- Small set expansion: if approximating SSE is hard, then approximating RIP is hard (Natrajan and Wu)

- Densest $k$-subgraph: if detecting an $n^{\frac{1}{2}-\epsilon}$ clique in a random graph $G(n, \frac{1}{2})$ is hard, approximating RIP is hard (Koiran and Zouzias)
Sum Of Squares

- SOS: a framework for proving statements using the trivial inequality and basic rules of algebra
- A degree-2n SOS proof proves a statement using only intermediate inequalities of polynomials of degree at most 2n
- Unbounded degree SOS is a complete proof system, bounded is not
- Max clique with an $n^{1/3}$ clique embedded in a random graph $G(n, \frac{1}{2})$ is unsolvable by degree-4 SOS
- For this particular implementation of the planted model, degree 2 SOS proof is sufficient
Theorem (Koiran and Souzias)

Assume a matrix $\Phi$ has unit column vectors and satisfies RIP of order $k$ and parameter $\epsilon$. For $m \geq k$, $\Phi$ also satisfies RIP of order $m$ and parameter $\epsilon \left( \frac{m-1}{k-1} \right)$.

We can set $n = m$ and examine the matrix’s eigenvalues to get a very coarse approximation.

Theorem

Sum of squares of degree 2 can differentiate between a matrix that is RIP of order $k$ with parameter $\delta$ and one that is not RIP of order $k$ with parameter $\delta \left( \frac{n-1}{k-1} \right)$. 
Limitations in Approximating RIP

Conjectures

Future Research

Conjecture

Planted model is complete - planted framework can be improved in order to prove any hardness results.

Conjecture

Degree-4 SOS is insufficient to approximate RIP to within any constant factor.
Acknowledgements

I would like to thank the following:

- MIT PRIMES
- Adrian Vladu
- My parents