Analysis of The Ratio of ℓ_1 and ℓ_2 Norms in Compressed Sensing

Yiming Xu[1], Akil Narayan[1,2], Hoang Tran[3] and Clayton Webster[4]

[1]Department of Mathematics, University of Utah, Salt Lake City, UT, US
[2]Scientific Computing and Imaging (SCI) Institute, University of Utah, Salt Lake City, UT, US
[3]Computer Science and Mathematics Division, Oak Ridge National Laboratory, Oak Ridge, TN, US
[4]Department of Mathematics, University of Tennessee, Knoxville, TN, US

Abstract

In [4], we study the theories behind using the non-convex objective ℓ_1/ℓ_2 in compressed sensing. Our work comprises the following aspects:

- Propose a novel criterion for local optimality guarantee;
- Provide the first uniform exact recovery condition and demonstrate its practicality;
- Analyze the stability of the method when noise pollutes data;
- Give a new initialization approach which empirically improves the current algorithm.

Problem set-up

Consider the compressed sensing problem [1]:

$$\min ||\mathbf{x}||_0 \quad \text{s.t. } \mathbf{Ax} = \mathbf{b}, \tag{1}$$

where $\mathbf{A} \in \mathbb{R}^{m\times n}$, $\mathbf{x} \in \mathbb{R}^n$ and $\mathbf{b} \in \mathbb{R}^m$. Here $m < n$ so the linear system is under-determined, making (1) NP-hard to solve. We study an alternative minimization problem whose objective is given by a norm ratio:

$$\min \frac{||\mathbf{x}||_1}{||\mathbf{x}||_2} \quad \text{s.t. } \mathbf{Ax} = \mathbf{b}. \tag{2}$$

Motivations

- Norm ratios are valid measures of complexity. For example, ℓ_1/ℓ_∞ of the singular value vector of a matrix is known as its numerical intrinsic rank.
- Norm ratios are scale-invariant.
- Better empirical performance over many other existing methods.
- Quick implementation based on manifold optimization [2] or the alternating direction method of multipliers (ADMM) [3].

The Main Results (Friendly Versions)

Theorem (A local optimality criterion)

Assume $s > 6$. If $\mathbf{A}$ satisfies that for every $T \subset [n]$ with $|T| \le s$,

$$||\mathbf{h}_T||_1 < \frac{1}{\sqrt{s+2}}||\mathbf{h}_{T^c}||_1 \quad \forall \mathbf{h} \in \ker(\mathbf{A}).$$

Then $\mathbf{x}_0$ is a local minimizer of ℓ_1/ℓ_2 for $\mathbf{Ax} = \mathbf{Ax}_0$.

Theorem (Uniform recoverability)

If, for some $s \in \mathbb{N}$, the matrix $\mathbf{A}$ satisfies

$$\inf_{\mathbf{h}\in\ker(\mathbf{A})\setminus\{\mathbf{0}\}} \frac{||\mathbf{h}||_1}{||\mathbf{h}||_2} > 3\sqrt{s} \tag{3}$$

then every s-sparse $\mathbf{x}_0$ ($||\mathbf{x}_0||_0 \le s$) is the unique solution to (2) with $\mathbf{b} = \mathbf{Ax}_0$. Particularly, if the rows of $\mathbf{A}$ are independent, isotropic and sub-gaussian, then (3) holds with probability at least $1 - 2e^{-u}$ for $m \gtrsim K^4 us\log n$, where K is the maximum sub-gaussian norm of $\mathbf{A}$.

Theorem (Stability)

Consider the noisy version of (2):

$$\min \frac{||\mathbf{x}||_1}{||\mathbf{x}||_2} \quad \text{s.t. } ||\mathbf{Ax} - \mathbf{Ax}_0||_2 \le \varepsilon, \tag{4}$$

where $\mathbf{x}_0$ is a sparse vector. Let $\mathbf{x}^*$ be a solution to (4). Under a similar condition as (3), if $\mathbf{x}^*$ and $\mathbf{x}_0$ are not 'close' in magnitude or coherence, then $||\mathbf{x}^* - \mathbf{x}_0||_1 \lesssim \mathcal{O}(\sqrt{n}\varepsilon)$.

Methodology

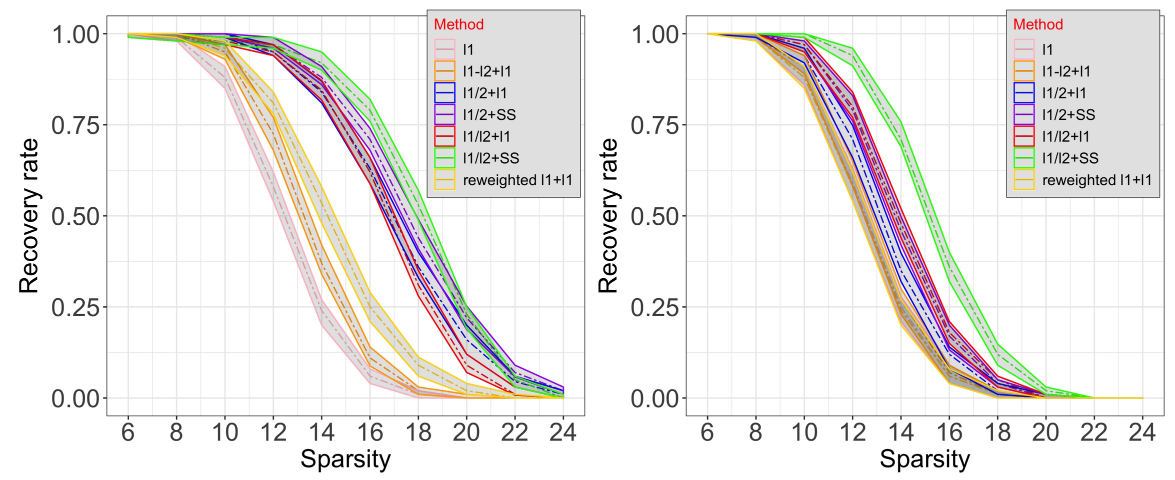
Non-convex problems are hard to tackle in general. However, the norm ratio objective in our case has close connections to the area of *high-dimensional geometry*, which provides us with the original idea. Other tools used in this work include fundamental results in linear algebra, random matrices and concentration inequalities. We remark that none of the results are proved tight; they only serve as the initial insight into certain aspects of the method that have been observed in practice.

Remarks

1. Our local optimality criterion improves a similar result in [3]. Moreover, the worst-case bound $\frac{1}{\sqrt{s+2}}$ can be raised to $\mathcal{O}(\frac{1}{\sqrt{\log s}})$ under mild assumptions.
2. Our exact recovery condition for sparse vectors is asymptotically comparable to the one given by the state-of-the-art ℓ_1, but not better. A promising direction to look for improvement is to consider recovery of compressible signals, which requires further definition.

Numerical simulations

The average performance of ℓ_1/ℓ_2 is compared to some well-established algorithms. A new initialization called *support selection* (SS) is tested against the commonly adopted ℓ_1 initial. $\mathbf{A}$ is chosen as a Gaussian random matrix of size 50×250. The nonzero components of the ground truth are sampled from two different distributions with varying dynamic range. Left: large dynamic range. Right: small dynamic range.



Conclusions

We have theoretically and numerically investigated the ℓ_1/ℓ_2 minimization problem in the context of recovery of sparse signals from a small number of measurements. Two recoverability conditions as well as a stability analysis are provided to back up the use of ℓ_1/ℓ_2 as a sparsity-promoting objective in practice.

References

[1] D. Donoho, IEEE T. Inform. Theory, volume 52, 1289-1306, 2006
[2] A. Petrosyana, H.Tran, C. Webster, Applied Numerical Mathematics, volume 144, 140-150, 2019.
[3] Y. Rahimi, C. Wang, H. Dong, Y. Lou, SIAM J. on Sci. Comp., 41 (6), 3649-3672, 2019.
[4] Y. Xu, A. Narayan, H.Tran, C. Webster, arXiv:2004.05873 [math.NA], 2020.

Acknowledgements

Yiming Xu would like to thank Tom Alberts, You-Cheng Chou and Dong Wang for useful discussions.