Sharp minima and strong minima for robust recovery

TRAN T.A. NGHIA[1]

Variational Analysis and Optimisation Webinar

(Based on the joint work[2] with J. Fadili (NU) and T. Tran (OU))

[1]Department of Mathematics and Statistics, Oakland University, Rochester, MI 48309, USA. Email: nttran@oakland.edu

[2]J. Fadili, T. T. A. Nghia, T. T. T. Tran: Sharp, strong and unique minimizers for low complexity robust recovery, preprint, 2021

A particular situation in different areas of engineering and science is that one has the observation

$$y_0 = \Phi x_0 \tag{1}$$

via a known (or random) linear process $\Phi \in \mathbb{R}^{m\times n}$ and an unknown vector $x_0 \in \mathbb{R}^n$.

Solving this linear equation to recover x_0 is a challenging task especially for the case $m << n$.

With prior information on x_0, an optimization is considered to recover x_0:

$$\min_{x\in\mathbb{R}^n} \quad J(x) := J_0(D^* x) \quad \text{subject to} \quad \Phi x = y_0, \tag{2}$$

where $J_0 : \mathbb{R}^p \to \mathbb{R}_+$ is non-negative regularizer and D is an $n \times p$ matrix.

- Sparsity: $J_0(u) = \|u\|_1 = \sum_{k=1}^{p} |u_k|$ (ℓ_1 norm) for $u \in \mathbb{R}^p$.
- Group-sparsity: $J_0(u) = \|u\|_{1,2} = \sum_{k=1}^{q} \|u_{g_k}\|$ (ℓ_1/ℓ_2 norm) for $u = (u_{g_1}, \ldots, u_{g_q}) \in \mathbb{R}^p$.
- Low-rank: $J_0(U) = \|U\|_* = \sum_{k=1}^{r} \sigma_k(U)$ (nuclear norm) for $U \in \mathbb{R}^{t\times s} = \mathbb{R}^p$ and $r = \text{rank}\,(U)$.

When the observation is disrupted by noise, the system (1) is modified by

$$y = \Phi x_0 + \omega \tag{3}$$

with a small noise ω in $\mathbb{R}^m$ with $\|\omega\| \leq \delta$.

A typical way to recover x_0 via optimization is to solve the following problem

$$\min_{x\in\mathbb{R}^n} \quad J(x) \quad \text{subject to} \quad \|\Phi x - y\| \leq \delta \tag{4}$$

or its Lagrange form

$$\min_{x\in\mathbb{R}^n} \quad \frac{1}{2}\|\Phi x - y\|^2 + \mu J(x) \tag{5}$$

with parameter $\mu > 0$.

A stage of Robust Recovery occurs when

- Any solution x_δ to (4) converges to x_0 as $\delta \to 0$.
- Any solution x_μ to (5) converges to x_0 as $\delta \to 0$ and $\mu = c\delta$.

Some well-known results on robust recovery

Theorem 1 (Solution uniqueness for robust recovery)

If $J(x) = \|x\|_1$, x_0 is a unique minimizer to problem (2) if and only if $\|x_\delta - x_0\| = O(\delta)$ and $\|x_\mu - x_0\| = O(\delta)$ as $\mu = c\delta$.

Grasmair, M., Haltmeier, M., Scherzer, O.: Necessary and sufficient conditions for linear convergence of ℓ_1-regularization, *Comm. Pure Applied Math.* **64** (2011), 161–182.

[3] Fuchs, J. J.: Recovery of exact sparse representations in the presence of bounded noise, *IEEE Trans. Inf. Theory*, **51** (2005), 3601–3608.

[4] Bruckstein, A., M., Donoho, D. L., Elad, M.: From sparse solutions of systems of equations to sparse modeling of signals and images, *SIAM Review* **51** (2009), 34–81.

Some well-known results on robust recovery

Theorem 1 (Solution uniqueness for robust recovery)

If $J(x) = \|x\|_1$, x_0 is a unique minimizer to problem (2) if and only if $\|x_\delta - x_0\| = O(\delta)$ and $\|x_\mu - x_0\| = O(\delta)$ as $\mu = c\delta$.

Grasmair, M., Haltmeier, M., Scherzer, O.: Necessary and sufficient conditions for linear convergence of ℓ_1-regularization, *Comm. Pure Applied Math.* **64** (2011), 161–182.

Why is solution uniqueness?

- Naturally, x_δ may converge to a minimizer to problem (2) and we want that minimizer to be x_0 (recovering x_0).
- If x_0 is the unique solution to (2), it is also the unique solution to the ℓ_0-problem[3] [4]:

$$\min \quad \|x\|_0 \quad \text{subject to} \quad \Phi x = y_0.$$

[3] Fuchs, J. J.: Recovery of exact sparse representations in the presence of bounded noise, *IEEE Trans. Inf. Theory*, **51** (2005), 3601–3608.

[4] Bruckstein, A., M., Donoho, D. L., Elad, M.: From sparse solutions of systems of equations to sparse modeling of signals and images, *SIAM Review* **51** (2009), 34–81.

Theorem 2 (Robust recovery via descent cone)

Let J be a norm in $\mathbb{R}^n$. Suppose that there exists some $\alpha > 0$ such that $\|\Phi w\| \geq \alpha \|w\|$ for all $w \in \mathcal{T}_J(x_0)$, where $\mathcal{T}_J(x_0)$ is the descent cone to J at x_0 defined by

$$\mathcal{T}_J(x_0) := \text{cone}\,\{x - x_0 |\; J(x) \leq J(x_0)\}.$$

Then any solution x_δ to problem (4) satisfies

$$\|x_\delta - x_0\| \leq \frac{2\delta}{\alpha}.$$

Chandrasekaran, V., Recht, B., Parrilo, P.A, Willsky, A. S.: The convex geometry of linear inverse problems, *Found Comput Math*, **12** (2012), 805–849.

Theorem 2 (Robust recovery via descent cone)

Let J be a norm in $\mathbb{R}^n$. Suppose that there exists some $\alpha > 0$ such that $\|\Phi w\| \geq \alpha \|w\|$ for all $w \in \mathcal{T}_J(x_0)$, where $\mathcal{T}_J(x_0)$ is the descent cone to J at x_0 defined by

$$\mathcal{T}_J(x_0) := \mathrm{cone}\,\{x - x_0 |\ J(x) \leq J(x_0)\}.$$

Then any solution x_δ to problem (4) satisfies

$$\|x_\delta - x_0\| \leq \frac{2\delta}{\alpha}.$$

Chandrasekaran, V., Recht, B., Parrilo, P.A, Willsky, A. S.: The convex geometry of linear inverse problems, *Found Comput Math*, **12** (2012), 805–849.

- The closure of descent cone is indeed the critical cone to J at x_0:

$$\mathcal{C}_J(x_0) := \{w \in \mathbb{R}^n |\ dJ(x_0)(w) \leq 0\}.$$

- The descent cone $\mathcal{T}_J(x_0)$ is not closed. But it is closed in the case $J(x) = \|x\|_1$.

Theorem 2 (Robust recovery via descent cone)

Let J be a norm in $\mathbb{R}^n$. Suppose that there exists some $\alpha > 0$ such that $\|\Phi w\| \geq \alpha \|w\|$ for all $w \in \mathcal{T}_J(x_0)$, where $\mathcal{T}_J(x_0)$ is the descent cone to J at x_0 defined by

$$\mathcal{T}_J(x_0) := \text{cone}\,\{x - x_0 |\ J(x) \leq J(x_0)\}.$$

Then any solution x_δ to problem (4) satisfies

$$\|x_\delta - x_0\| \leq \frac{2\delta}{\alpha}.$$

Chandrasekaran, V., Recht, B., Parrilo, P.A, Willsky, A. S.: The convex geometry of linear inverse problems, *Found Comput Math*, **12** (2012), 805–849.

- The closure of descent cone is indeed the critical cone to J at x_0:

$$\mathcal{C}_J(x_0) := \{w \in \mathbb{R}^n |\ dJ(x_0)(w) \leq 0\}.$$

- The descent cone $\mathcal{T}_J(x_0)$ is not closed. But it is closed in the case $J(x) = \|x\|_1$.

Where is solution uniqueness?

- It is hidden in the red condition, which indeed means $\text{Ker}\,\Phi \cap C_J(x_0) = \{0\}$.
- Solution uniqueness is characterized by $\text{Ker}\,\Phi \cap \mathcal{T}_J(x_0) = \{0\}$.

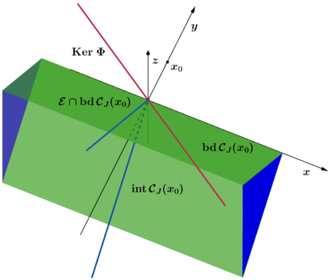
Example 3 (Solution uniqueness for group-sparsity problems)

Consider the following ℓ_1/ℓ_2 optimization problem:

$$\min_{x\in\mathbb{R}^3} \quad J(x) = \sqrt{x_1^2 + x_2^2} + |x_3| \quad \text{subject to} \quad \Phi x = \Phi x_0$$

with $\Phi = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 0 & -1 \end{bmatrix}$, $x_0 = (0,1,0)^T$, and $y_0 = \Phi x_0 = (1,0)^T$. We have

$$\operatorname{Ker}\Phi \cap \mathcal{C}_J(x_0) \neq \{0\} \quad \text{but} \quad \operatorname{Ker}\Phi \cap \mathcal{T}_J(x_0) = \{0\}.$$



Some questions to answer

Q1: What does the condition $\operatorname{Ker}\Phi \cap \mathcal{C}_J(x_0) = \{0\}$ mean?

Q2: Does J have to be a norm?

Q3: Can solution uniqueness sufficiently guarantee robust recovery with linear rate?

Some questions to answer

Q1: What does the condition $\operatorname{Ker}\Phi \cap \mathcal{C}_J(x_0) = \{0\}$ mean?

Q2: Does J have to be a norm?

Q3: Can solution uniqueness sufficiently guarantee robust recovery with linear rate?

We observe that the above condition is equivalent to the so-called sharp minima at x_0: there exists $c > 0$ such that

$$J(x) - J(x_0) \geq c\|x - x_0\| \quad \text{for} \quad \Phi x = y_0.$$

This is also equivalent to solution uniqueness in the case of ℓ_1 optimization problem.

Sharp minima

Definition 4 (Sharp minima)

We say $\bar{x}$ to be a sharp solution/minimizer to the (not necessarily convex) function $\varphi : \mathbb{R}^n \to \overline{\mathbb{R}} := \mathbb{R} \cup \{\infty\}$ with a constant $c > 0$ if there exists $\varepsilon > 0$ such that

$$\varphi(x) - \varphi(\bar{x}) \geq c\|x - \bar{x}\| \quad \text{for all} \quad x \in \mathbb{B}_\varepsilon(\bar{x}).$$

Polyak, B. T.: Sharp minima, Institute of Control Sciences Lecture Notes, Moscow, USSR, 1979.

- Sharp minima is a global property when φ is a convex function.
- Sharp minima plays significant roles in algorithms as it usually guarantees finite termination.
- Sharp minima at $\bar{x}$ can be characterized by:

$$d\varphi(\bar{x})(w) \geq c\|w\| \quad \text{for all} \quad w \in \mathbb{R}^n.$$

Recall problem (2):

$$\min_{x\in\mathbb{R}^n} \quad J(x) \quad \text{subject to} \quad \Phi x = y_0,$$

Suppose that $J:\mathbb{R}^n\to\mathbb{R}$ is a continuous convex function (not necessary, but for simplification in this talk).

Proposition 1 (Solution uniqueness for robust recovery)

x_0 is a unique solution to problem (2) if and only if:

1. *Any solution x_δ to problem (4):*

$$\min_{x\in\mathbb{R}^n} \quad J(x) \quad \textit{subject to} \quad \|\Phi x - y\| \le \delta$$

   converges to x_0 as $\delta\to 0$.

2. *For any constant $c_1>0$, any solution x_μ to problem (5):*

$$\min_{x\in\mathbb{R}^n} \quad \frac{1}{2}\|\Phi x - y\|^2 + \mu J(x)$$

   with $\mu = c_1\delta$ converges to x_0 as $\delta\to 0$.

Theorem 5 (Sharp minima for robust recovery)

If x_0 is a sharp solution to (2), i.e., $\bar{x}$ is a sharp solution to the function $\varphi(x) := J(x) + \delta_{\Phi^{-1}(y_0)}(x)$, we have:

1. *Any solution x_δ to problem (4) satisfies*

$$\|x_\delta - x_0\| \leq O(\delta).$$

2. *For any $c_1 > 0$ and $\mu = c_1\delta$, every minimizer x_μ to (5) satisfies*

$$\|x_\mu - x_0\| \leq O(\delta).$$

- This theorem covers lots of well-known results.
- The precise calculation for $O(\delta)$ can be obtained.
- The cost function J is not necessarily convex.

Theorem 5 (Sharp minima for robust recovery)

If x_0 *is a sharp solution to* (2)*, i.e.,* $\bar{x}$ *is a sharp solution to the function* $\varphi(x) := J(x) + \delta_{\Phi^{-1}(y_0)}(x)$*, we have:*

1. *Any solution* x_δ *to problem* (4) *satisfies*

$$\|x_\delta - x_0\| \le O(\delta).$$

2. *For any* $c_1 > 0$ *and* $\mu = c_1\delta$*, every minimizer* x_μ *to* (5) *satisfies*

$$\|x_\mu - x_0\| \le O(\delta).$$

- This theorem covers lots of well-known results.
- The precise calculation for $O(\delta)$ can be obtained.
- The cost function J is not necessarily convex.

In Example 3, we have solution uniqueness, which is not sharp minima. However, this unique solution is indeed a strong solution.

Strong minima

Definition 6 (Strong minima)

We say $\bar{x}$ is said to be a strong solution/minimizer to the function $\varphi : \mathbb{R}^n \to \overline{\mathbb{R}}$ with a constant $\kappa > 0$ if there exists $\delta > 0$ such that

$$\varphi(x) - \varphi(\bar{x}) \geq \frac{\kappa}{2}\|x - \bar{x}\|^2 \quad \text{for all} \quad x \in \mathbb{B}_\delta(\bar{x}).$$

- Strong minima is desired in many nonlinear algorithms to guarantee fast convergences.
- Strong minima can be characterized by second-order analysis: $\bar{x}$ is a strong minima to φ if and only if $0 \in \partial\varphi(\bar{x})$ and

$$d^2\varphi(\bar{x}|0)(w) > 0 \quad \text{for all} \quad w \neq 0.$$

Theorem 7 (Strong minima for robust recovery)

If x_0 is a strong solution to problem (2), the following statements hold:

1. *Any solution x_δ to problem (4) satisfies:*
$$\|x_\delta - x_0\| \leq O(\delta^{\frac{1}{2}}).$$
2. *For any constant $c_1 > 0$, every minimizer x_μ to problem (5) with $\mu = c_1\delta$ satisfies*
$$\|x_\mu - x_0\| \leq O(\delta^{\frac{1}{2}}).$$

- The precise calculation for $O(\delta^{\frac{1}{2}})$ can be obtained.
- The cost function J is not necessarily convex.
- Robust recovery for the problem in Example 3 has rate $O(\delta^{\frac{1}{2}})$.

Does this rate remain for any group-sparsity problems?

Decomposable norms

In this section, consider the following problem (2)

$$\min_{x\in\mathbb{R}^n} \quad J(x) := \|D^* x\|_{\mathcal{A}} \quad \text{subject to} \quad \Phi x = y_0.$$

Definition 8 (Decomposable norms)

A norm $\|\cdot\|_{\mathcal{A}}$ is called to be decomposable at $\bar{u}$ if there is a subspace $V \subset \mathbb{R}^p$ and a vector $e \in V$ such that

$$\partial\|\bar{u}\|_{\mathcal{A}} = \{z \in \mathbb{R}^p |\ P_V z = e \quad \text{and} \quad \|P_{V^\perp} z\|_{\mathcal{A}}^* \le 1\},$$

where $\|\cdot\|_{\mathcal{A}}^*$ is the dual norm to $\|\cdot\|_{\mathcal{A}}$.

Negahban, S., Ravikumar, P., Wainwright, M.J., Yu, B.: A unified framework for high-dimensional analysis of m-estimators with decomposable regularizers. In: Advances in Neural Information Processing Systems, conference proceeding (2009)

Candès, E. and Retch, B.: Simple bounds for recovering low-complexity models, *Math. Program.*, **141** (2013), 577–589.

For $\|\cdot\|_{\mathcal{A}} = \|\cdot\|_1$, we have

$$e = (\text{sign}\,\{\bar{u}_I\}, 0_K)^T \quad \text{and} \quad V = \mathbb{R}^I \times \{0_K\},$$

where $I = \text{supp}\,(\bar{u}) = \{i \in \{1, \ldots, p|\ \bar{u}_i \neq 0\}$ and $K = I^c$.

Decomposable norm includes the ℓ_1 norm, ℓ_1/ℓ_2 norm, and nuclear norm.

Theorem 9 (Characterizations for sharp solution)

The following statements are equivalent:

1. x_0 *is a sharp solution to problem* (2).
2. *The* Restricted Injectivity *holds at* x_0 *in the sense that*

$$\operatorname{Ker}\Phi \cap \operatorname{Ker}(D^*_{V^\perp}) = \{0\}$$

with $D_{V^\perp} = DP_{V^\perp}$ *and the* Source Identity $\rho(e)$ *is less than* 1, *where* $\rho(e)$ *is the optimal value to the following convex optimization problem*

$$\min \quad \|z\|^*_{\mathcal{A}} \quad \textit{subject to} \quad NDz = -NDe \quad \textit{and} \quad z \in V^\perp. \tag{6}$$

with N *being the matrix forming the basis to* $\operatorname{Ker}\Phi$.

Theorem 9 (Characterizations for sharp solution)

The following statements are equivalent:

1. x_0 *is a sharp solution to problem* (2).
2. *The* Restricted Injectivity *holds at* x_0 *in the sense that*

$$\operatorname{Ker}\Phi \cap \operatorname{Ker}(D^*_{V^\perp}) = \{0\}$$

with $D_{V^\perp} = DP_{V^\perp}$ *and the* Source Identity $\rho(e)$ *is less than* 1*, where* $\rho(e)$ *is the optimal value to the following convex optimization problem*

$$\min \quad \|z\|^*_{\mathcal{A}} \quad \textit{subject to} \quad NDz = -NDe \quad \textit{and} \quad z \in V^\perp. \tag{6}$$

with N *being the matrix forming the basis to* $\operatorname{Ker}\Phi$.

- Restricted Injectivity is a traditional condition necessary for solution uniqueness.
- The matrix N can be chosen from the singular value decomposition of $\Phi = U\Sigma V^*$.
- Particular cases of Source Condition are sufficient for solution uniqueness and robust recovery used in many papers.
- Problem (6) is a convex problem, we use `cvxopt` package to solve it.

Corollary 10 (Sufficient condition for sharp solution)

If the Restricted Injectivity holds at x_0 and the Analysis Exact Recovery Condition [a]

$$\tau(e) := \|(ND_{V^\perp})^\dagger ND_V e\|_{\mathcal{A}}^* < 1$$

is satisfied, then x_0 is a sharp solution to problem (2).

[a] appeared in "Nam, S., Davies, M. E., Elad, M., Gribonval, R.: The cosparse analysis model and algorithms, *Applied and Computational Harmonic Analysis*, **34** (2013), 30–56" for the case $J_0 = \|\cdot\|_1$.

Corollary 10 (Sufficient condition for sharp solution)

If the Restricted Injectivity holds at x_0 and the Analysis Exact Recovery Condition [a]

$$\tau(e) := \|(ND_{V^\perp})^\dagger ND_V e\|_{\mathcal{A}}^* < 1$$

is satisfied, then x_0 is a sharp solution to problem (2).

[a] appeared in "Nam, S., Davies, M. E., Elad, M., Gribonval, R.: The cosparse analysis model and algorithms, *Applied and Computational Harmonic Analysis*, **34** (2013), 30–56" for the case $J_0 = \|\cdot\|_1$.

How can we check sharp minima?

- First, check the Restricted Injectivity.
- If the Restricted Injectivity holds and $\tau(e) << 1$, then x_0 is a sharp minima.
- If the Restricted Injectivity holds and $\tau(e) \approx 1$, check condition $\rho(e) < 1$.

Group-sparsity problems

Suppose that $\mathbb{R}^p$ is decomposed into q groups by

$$\mathbb{R}^p = \bigoplus_{g=1}^{q} V_g,$$

where each V_g is a subspace of $\mathbb{R}^p$ with the same dimension G. For any $u \in \mathbb{R}^p$ and $1 \leq g \leq q$, we write $u = \sum_{g=1}^{q} u_g$ with $u_g \in V_g$ being the vector in group V_g of u.

The ℓ_1/ℓ_2 norm in $\mathbb{R}^p$ is defined by

$$\|u\|_{\ell_1/\ell_2} = \sum_{g=1}^{q} \|u_g\|_2.$$

Its dual is the ℓ_∞/ℓ_2 norm:

$$\|u\|_{\ell_\infty/\ell_2} = \max_{1\leq g\leq q} \|u_g\|_2.$$

ℓ_1/ℓ_2 norm is decomposable at $\bar{u} = D^* x_0$ with $V = \bigoplus_{g\in I} V_g$,

$I := \{g \in \{1, \ldots, q\} |\ u_g \neq 0\}$, and

$$e = \sum_{g\in I} \frac{\bar{u}_g}{\|\bar{u}_g\|_2}.$$

Theorem 11 (Characterizations for unique/strong solutions to group-sparsity problems)

The following assertions are equivalent:

1. x_0 *is a unique solution to problem* (7).
2. x_0 *is a strong solution to problem* (7).
3. x_0 *is a solution to* (7), $\operatorname{Ker}\Phi \cap \mathcal{E} \cap \operatorname{Ker} D_S^* = \{0\}$, *and*

$$\zeta(e) := \min_{u \in \operatorname{Ker} MD_{V^\perp}} \|(MD_{V^\perp})^\dagger MD_V e - u\|_{\ell_\infty/\ell_2} < 1,$$

where $\mathcal{E}$ *is defined by*

$$\mathcal{E} := \{w \in \mathbb{R}^n \,|\, D_V^* w \in \operatorname{span}\{e_g |\, g \in I\}\}$$

and M^* *is a matrix forming a basis matrix to* $\operatorname{Ker}\Phi \cap \mathcal{E}$.

Several known sufficient conditions [5] [6] for solution uniqueness to group-sparsity problems are stronger than the about one.

[5] Grasmair, M.: Linear convergence rates for Tikhonov regularization with positively homogeneous functionals, *Inverse Problems*, **27**(2011) 075014.

[6] Roth, V. and Fischer, B.: The group-lasso for generalized linear models: uniqueness of solutions and efficient algorithms, Proceedings of the 25th ICML, 2008.

Corollary 12 (Robust convergence for group-sparsity under solution uniqueness)

If x_0 is a unique minimizer to problem (2) with $J_0 = \|\cdot\|_{1,2}$, the following statements hold :

1. *Any solution x_δ to problem (4) with noise $\|\omega\| \le \delta$ satisfies:* $\|x_\delta - x_0\| \le O(\delta^{\frac{1}{2}})$.
2. *For any constant $c_1 > 0$, every minimizer x_μ to problem (5) with noise $\|\omega\| \le \delta$ and $\mu = c_1\delta$ satisfies $\|x_\mu - x_0\| \le O(\delta^{\frac{1}{2}})$.*

Computing the Strong Source Condition $\zeta(e)$

- $\zeta^2(e)$ is the optimal value to the following smooth convex optimization problem (second-order cone programming):

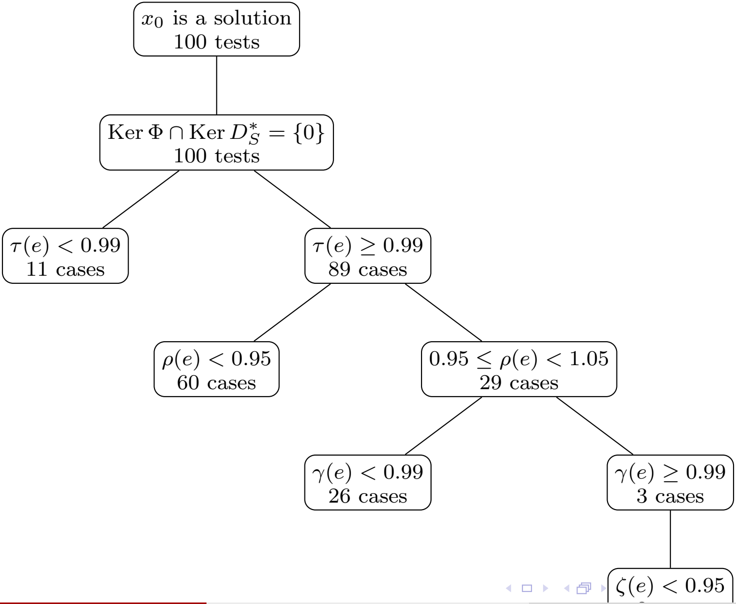
$$\min \quad t \quad \text{subject to} \quad MDz = -MDe, \quad \|z_g\|^2 - t \leq 0, \quad g \in I^c \quad \text{and} \quad z \in \bigoplus_{g \in I} V_g$$

  with $|I|G + 1$ variables, which can be solved by available packages such as `cvxopt`.

- An simple upper bound for $\zeta(e)$ is

$$\zeta(e) \leq \gamma(e) := \|(MD_{V^\perp})^\dagger MD_V e\|_{\ell_\infty/\ell_2}.$$

Unique solutions and sharp solutions in group-sparsity problems



	number of cases
Sharp solution	71
Strong solution (non-sharp)	29

Table: Number of cases with strong and sharp solutions

[7] Rao, N., Recht, B., Nowak,R.: Universal measurement bounds for structured sparse signal recovery, Proceedings of AISTATS (2012).

[8] Candès, E. and Retch, B.: Simple bounds for recovering low-complexity models, *Math. Program.*, **141** (2013), 577–589.

	number of cases
Sharp solution	71
Strong solution (non-sharp)	29

Table: Number of cases with strong and sharp solutions

How can all solutions in these 100 random problems be unique?

- This belongs to the area of Exact recovery with high probability. [7][8]

- Exact Recovery is strongly studied by involving sharp minima. Strong minima in Exact Recovery is not yet discovered.

[7] Rao, N., Recht, B., Nowak,R.: Universal measurement bounds for structured sparse signal recovery, Proceedings of AISTATS (2012).

[8] Candès, E. and Retch, B.: Simple bounds for recovering low-complexity models, *Math. Program.*, **141** (2013), 577–589.

Nuclear norm minimization problem

Example 13 (Difference between unique solution and strong solution to NNM)

Consider the following optimization problem

$$\min_{X\in\mathbb{R}^{2\times 2}} \|X\|_* \quad \text{subject to} \quad \Phi(X) := \begin{bmatrix} X_{11}+X_{22} \\ X_{12}-X_{21}+X_{22} \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}.$$

$\overline{X} = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$ is a unique solution, but it is neither strong nor sharp solution to NNM.

A study about unique/sharp/strong solutions for NNM is interesting and open.

In this work, we show that:

- Sharp, strong, and unique minimizers play significant roles in robust recovery.
- Sharp minima can be characterized numerically.
- Unique and strong solutions for group-sparsity problems are the same.
- Solution uniqueness for group-sparsity problems is equivalent to robust recover with $O(\delta^{\frac{1}{2}})$ rate.

In this work, we show that:

- Sharp, strong, and unique minimizers play significant roles in robust recovery.
- Sharp minima can be characterized numerically.
- Unique and strong solutions for group-sparsity problems are the same.
- Solution uniqueness for group-sparsity problems is equivalent to robust recover with $O(\delta^{\frac{1}{2}})$ rate.

We are working on several open questions:

- How does solution uniqueness affect on robust recover when dealing with nuclear norm minimization problems?
- What happens when solution uniqueness does not occur?
- How to use strong minima in Exact Recovery?