Overview of the Mathematics of Compressive Sensing

Simon Foucart

Reading Seminar on
"Compressive Sensing, Extensions, and Applications"
Texas A&M University
8 October 2015

Coherence-based Recovery Guarantees

Coherence

Coherence

For a matrix with ℓ_2-normalized columns $\mathbf{a}_1, \ldots, \mathbf{a}_N$, define

$$\mu := \max_{i \neq j} \left| \langle \mathbf{a}_i, \mathbf{a}_j \rangle \right| .$$

Coherence

For a matrix with ℓ_2-normalized columns $\mathbf{a}_1, \ldots, \mathbf{a}_N$, define

$$\mu := \max_{i \neq j} \left| \langle \mathbf{a}_i, \mathbf{a}_j \rangle \right| .$$

As a rule, the smaller the coherence, the better.

Coherence

For a matrix with ℓ_2-normalized columns $\mathbf{a}_1, \ldots, \mathbf{a}_N$, define

$$\mu := \max_{i \neq j} \left| \langle \mathbf{a}_i, \mathbf{a}_j \rangle \right| .$$

As a rule, the smaller the coherence, the better.

However, the Welch bound reads

$$\mu \geq \sqrt{\frac{N - m}{m(N-1)}}.$$

Coherence

For a matrix with ℓ_2-normalized columns $\mathbf{a}_1, \ldots, \mathbf{a}_N$, define

$$\mu := \max_{i \neq j} |\langle \mathbf{a}_i, \mathbf{a}_j \rangle| \, .$$

As a rule, the smaller the coherence, the better.

However, the Welch bound reads

$$\mu \geq \sqrt{\frac{N-m}{m(N-1)}}.$$

- Welch bound achieved at and only at equiangular tight frames.

Coherence

For a matrix with ℓ_2-normalized columns $\mathbf{a}_1, \ldots, \mathbf{a}_N$, define

$$\mu := \max_{i \neq j} |\langle \mathbf{a}_i, \mathbf{a}_j \rangle| .$$

As a rule, the smaller the coherence, the better.

However, the Welch bound reads

$$\mu \geq \sqrt{\frac{N-m}{m(N-1)}}.$$

- Welch bound achieved at and only at equiangular tight frames.
- Deterministic matrices with coherence $\mu \leq c/\sqrt{m}$ exist.

Recovery Conditions using Coherence

Recovery Conditions using Coherence

- Every s-sparse $\mathbf{x} \in \mathbb{C}^N$ is recovered from $\mathbf{y} = A\mathbf{x} \in \mathbb{C}^m$ via at most s iterations of OMP provided

$$\mu < \frac{1}{2s-1}.$$

Recovery Conditions using Coherence

- Every s-sparse $\mathbf{x} \in \mathbb{C}^N$ is recovered from $\mathbf{y} = A\mathbf{x} \in \mathbb{C}^m$ via at most s iterations of OMP provided

$$\mu < \frac{1}{2s-1}.$$

- Every s-sparse $\mathbf{x} \in \mathbb{C}^N$ is recovered from $\mathbf{y} = A\mathbf{x} \in \mathbb{C}^m$ via Basis Pursuit provided

$$\mu < \frac{1}{2s-1}.$$

Recovery Conditions using Coherence

- Every s-sparse $\mathbf{x} \in \mathbb{C}^N$ is recovered from $\mathbf{y} = A\mathbf{x} \in \mathbb{C}^m$ via at most s iterations of OMP provided

$$\mu < \frac{1}{2s-1}.$$

- Every s-sparse $\mathbf{x} \in \mathbb{C}^N$ is recovered from $\mathbf{y} = A\mathbf{x} \in \mathbb{C}^m$ via Basis Pursuit provided

$$\mu < \frac{1}{2s-1}.$$

- In fact, the Exact Recovery Condition can be rephrased as

$$\|A_S^{\dagger} A_{\overline{S}}\|_{1\to 1} < 1,$$

and this implies the Null Space Property.

Situation circa 2004

Situation circa 2004

The coherence conditions for s-sparse recovery are of the type

$$\mu \leq \frac{c}{s},$$

Situation circa 2004

The coherence conditions for s-sparse recovery are of the type

$$\mu \leq \frac{c}{s},$$

while the Welch bound (for $N \geq 2m$, say) reads

$$\mu \geq \frac{c}{\sqrt{m}}.$$

Situation circa 2004

The coherence conditions for s-sparse recovery are of the type

$$\mu \leq \frac{c}{s},$$

while the Welch bound (for $N \geq 2m$, say) reads

$$\mu \geq \frac{c}{\sqrt{m}}.$$

Therefore, arguments based on coherence necessitate

$$m \geq c\, s^2.$$

Situation circa 2004

The coherence conditions for s-sparse recovery are of the type

$$\mu \leq \frac{c}{s},$$

while the Welch bound (for $N \geq 2m$, say) reads

$$\mu \geq \frac{c}{\sqrt{m}}.$$

Therefore, arguments based on coherence necessitate

$$m \geq c\, s^2.$$

This is far from the ideal linear scaling of m in s...

Situation circa 2004

The coherence conditions for s-sparse recovery are of the type

$$\mu \leq \frac{c}{s},$$

while the Welch bound (for $N \geq 2m$, say) reads

$$\mu \geq \frac{c}{\sqrt{m}}.$$

Therefore, arguments based on coherence necessitate

$$m \geq c\, s^2.$$

This is far from the ideal linear scaling of m in s...

We now introduce new tools to break this *quadratic barrier*.

RIP-based Recovery Guarantees

Restricted Isometry Constants

Restricted Isometry Constants

Restricted Isometry Constant δ_s = the smallest $\delta > 0$ such that

$$(1-\delta)\|\mathbf{z}\|_2^2 \leq \|A\mathbf{z}\|_2^2 \leq (1+\delta)\|\mathbf{z}\|_2^2 \quad \text{for all } s\text{-sparse } \mathbf{z} \in \mathbb{C}^N.$$

Restricted Isometry Constants

Restricted Isometry Constant δ_s = the smallest $\delta > 0$ such that

$$(1-\delta)\|\mathbf{z}\|_2^2 \le \|A\mathbf{z}\|_2^2 \le (1+\delta)\|\mathbf{z}\|_2^2 \quad \text{for all } s\text{-sparse } \mathbf{z} \in \mathbb{C}^N.$$

Alternative form:

$$\delta_s = \max_{\mathrm{card}(S) \le s} \|A_S^* A_S - \mathrm{Id}\|_{2\to 2}.$$

Restricted Isometry Constants

Restricted Isometry Constant δ_s = the smallest $\delta > 0$ such that

$$(1-\delta)\|\mathbf{z}\|_2^2 \le \|A\mathbf{z}\|_2^2 \le (1+\delta)\|\mathbf{z}\|_2^2 \quad \text{for all } s\text{-sparse } \mathbf{z} \in \mathbb{C}^N.$$

Alternative form:

$$\delta_s = \max_{\mathrm{card}(S) \le s} \|A_S^* A_S - \mathrm{Id}\|_{2\to 2}.$$

Suitable CS matrices have small restricted isometry constants.

Restricted Isometry Constants

Restricted Isometry Constant δ_s = the smallest $\delta > 0$ such that

$$(1-\delta)\|\mathbf{z}\|_2^2 \le \|A\mathbf{z}\|_2^2 \le (1+\delta)\|\mathbf{z}\|_2^2 \quad \text{for all } s\text{-sparse } \mathbf{z} \in \mathbb{C}^N.$$

Alternative form:

$$\delta_s = \max_{\mathrm{card}(S) \le s} \|A_S^* A_S - \mathrm{Id}\|_{2\to 2}.$$

Suitable CS matrices have small restricted isometry constants.

Typically, $\delta_s \le \delta_*$ holds for a number of random measurements

$$m \approx c(\delta_*)\, s \ln(eN/s).$$

Restricted Isometry Constants

Restricted Isometry Constant δ_s = the smallest $\delta > 0$ such that

$$(1-\delta)\|\mathbf{z}\|_2^2 \le \|A\mathbf{z}\|_2^2 \le (1+\delta)\|\mathbf{z}\|_2^2 \quad \text{for all } s\text{-sparse } \mathbf{z} \in \mathbb{C}^N.$$

Alternative form:

$$\delta_s = \max_{\mathrm{card}(S) \le s} \|A_S^* A_S - \mathrm{Id}\|_{2\to 2}.$$

Suitable CS matrices have small restricted isometry constants.

Typically, $\delta_s \le \delta_*$ holds for a number of random measurements

$$m \approx \frac{c}{\delta_*^2}\, s \ln(eN/s).$$

Restricted Isometry Constants

Restricted Isometry Constant δ_s = the smallest $\delta > 0$ such that

$$(1-\delta)\|\mathbf{z}\|_2^2 \le \|A\mathbf{z}\|_2^2 \le (1+\delta)\|\mathbf{z}\|_2^2 \quad \text{for all } s\text{-sparse } \mathbf{z} \in \mathbb{C}^N.$$

Alternative form:

$$\delta_s = \max_{\mathrm{card}(S) \le s} \|A_S^* A_S - \mathrm{Id}\|_{2\to 2}.$$

Suitable CS matrices have small restricted isometry constants.

Typically, $\delta_s \le \delta_*$ holds for a number of random measurements

$$m \approx \frac{c}{\delta_*^2} s \ln(eN/s).$$

In fact, $\delta_s \le \delta_*$ imposes $m \ge \frac{c}{\delta_*^2} s$.

Exact Sparse Recovery via ℓ_1-Minimization when $\delta_{2s} < 1/3$

Exact Sparse Recovery via ℓ_1-Minimization when $\delta_{2s} < 1/3$

Take $\mathbf{v} \in \ker A \setminus \{\mathbf{0}\}$.

Exact Sparse Recovery via ℓ_1-Minimization when $\delta_{2s} < 1/3$

Take $\mathbf{v} \in \ker A \setminus \{\mathbf{0}\}$. Consider sets $S_0, S_1, \ldots$ of s indices ordered by decreasing magnitude of entries of $\mathbf{v}$.

Exact Sparse Recovery via ℓ_1-Minimization when $\delta_{2s} < 1/3$

Take $\mathbf{v} \in \ker A \setminus \{\mathbf{0}\}$. Consider sets $S_0, S_1, \ldots$ of s indices ordered by decreasing magnitude of entries of $\mathbf{v}$.

$$\|\mathbf{v}_{S_0}\|_2^2 \leq$$

Exact Sparse Recovery via ℓ_1-Minimization when $\delta_{2s} < 1/3$

Take $\mathbf{v} \in \ker A \setminus \{\mathbf{0}\}$. Consider sets $S_0, S_1, \ldots$ of s indices ordered by decreasing magnitude of entries of $\mathbf{v}$.

$$\|\mathbf{v}_{S_0}\|_2^2 \leq \frac{1}{1-\delta_{2s}} \|A(\mathbf{v}_{S_0})\|_2^2$$

Exact Sparse Recovery via ℓ_1-Minimization when $\delta_{2s} < 1/3$

Take $\mathbf{v} \in \ker A \setminus \{\mathbf{0}\}$. Consider sets $S_0, S_1, \ldots$ of s indices ordered by decreasing magnitude of entries of $\mathbf{v}$.

$$\|\mathbf{v}_{S_0}\|_2^2 \quad \leq \quad \frac{1}{1-\delta_{2s}}\|A(\mathbf{v}_{S_0})\|_2^2 = \frac{1}{1-\delta_{2s}}\sum_{k\geq 1}\langle A(\mathbf{v}_{S_0}), A(-\mathbf{v}_{S_k})\rangle$$

Exact Sparse Recovery via ℓ_1-Minimization when $\delta_{2s} < 1/3$

Take $\mathbf{v} \in \ker A \setminus \{\mathbf{0}\}$. Consider sets $S_0, S_1, \dots$ of s indices ordered by decreasing magnitude of entries of $\mathbf{v}$.

$$\begin{aligned}
\|\mathbf{v}_{S_0}\|_2^2 &\leq \frac{1}{1-\delta_{2s}} \|A(\mathbf{v}_{S_0})\|_2^2 = \frac{1}{1-\delta_{2s}} \sum_{k \geq 1} \langle A(\mathbf{v}_{S_0}), A(-\mathbf{v}_{S_k}) \rangle \\
&\leq \frac{1}{1-\delta_{2s}} \sum_{k \geq 1} \delta_{2s} \|\mathbf{v}_{S_0}\|_2 \|\mathbf{v}_{S_k}\|_2 .
\end{aligned}$$

Exact Sparse Recovery via ℓ_1-Minimization when $\delta_{2s} < 1/3$

Take $\mathbf{v} \in \ker A \setminus \{\mathbf{0}\}$. Consider sets $S_0, S_1, \ldots$ of s indices ordered by decreasing magnitude of entries of $\mathbf{v}$.

$$\begin{aligned} \|\mathbf{v}_{S_0}\|_2^2 &\leq \frac{1}{1-\delta_{2s}}\|A(\mathbf{v}_{S_0})\|_2^2 = \frac{1}{1-\delta_{2s}}\sum_{k\geq 1}\langle A(\mathbf{v}_{S_0}), A(-\mathbf{v}_{S_k})\rangle \\ &\leq \frac{1}{1-\delta_{2s}}\sum_{k\geq 1}\delta_{2s}\|\mathbf{v}_{S_0}\|_2\|\mathbf{v}_{S_k}\|_2. \end{aligned}$$

Simplify by $\|\mathbf{v}_{S_0}\|_2$

Exact Sparse Recovery via ℓ_1-Minimization when $\delta_{2s} < 1/3$

Take $\mathbf{v} \in \ker A \setminus \{\mathbf{0}\}$. Consider sets $S_0, S_1, \dots$ of s indices ordered by decreasing magnitude of entries of $\mathbf{v}$.

$$\begin{aligned}
\|\mathbf{v}_{S_0}\|_2^2 &\leq \frac{1}{1-\delta_{2s}}\|A(\mathbf{v}_{S_0})\|_2^2 = \frac{1}{1-\delta_{2s}}\sum_{k\geq 1}\langle A(\mathbf{v}_{S_0}), A(-\mathbf{v}_{S_k})\rangle \\
&\leq \frac{1}{1-\delta_{2s}}\sum_{k\geq 1}\delta_{2s}\|\mathbf{v}_{S_0}\|_2\|\mathbf{v}_{S_k}\|_2.
\end{aligned}$$

Simplify by $\|\mathbf{v}_{S_0}\|_2$ and observe that

$$\|\mathbf{v}_{S_k}\|_2 \leq \frac{1}{\sqrt{s}}\|\mathbf{v}_{S_{k-1}}\|_1$$

Exact Sparse Recovery via ℓ_1-Minimization when $\delta_{2s} < 1/3$

Take $\mathbf{v} \in \ker A \setminus \{\mathbf{0}\}$. Consider sets $S_0, S_1, \ldots$ of s indices ordered by decreasing magnitude of entries of $\mathbf{v}$.

$$\begin{aligned} \|\mathbf{v}_{S_0}\|_2^2 &\leq \frac{1}{1-\delta_{2s}} \|A(\mathbf{v}_{S_0})\|_2^2 = \frac{1}{1-\delta_{2s}} \sum_{k \geq 1} \langle A(\mathbf{v}_{S_0}), A(-\mathbf{v}_{S_k}) \rangle \\ &\leq \frac{1}{1-\delta_{2s}} \sum_{k \geq 1} \delta_{2s} \|\mathbf{v}_{S_0}\|_2 \|\mathbf{v}_{S_k}\|_2. \end{aligned}$$

Simplify by $\|\mathbf{v}_{S_0}\|_2$ and observe that

$$\|\mathbf{v}_{S_k}\|_2 \leq \frac{1}{\sqrt{s}} \|\mathbf{v}_{S_{k-1}}\|_1$$

to obtain

$$\|\mathbf{v}_{S_0}\|_2 \leq \frac{\delta_{2s}}{1-\delta_{2s}} \frac{1}{\sqrt{s}} \|\mathbf{v}\|_1,$$

Exact Sparse Recovery via ℓ_1-Minimization when $\delta_{2s} < 1/3$

Take $\mathbf{v} \in \ker A \setminus \{\mathbf{0}\}$. Consider sets $S_0, S_1, \ldots$ of s indices ordered by decreasing magnitude of entries of $\mathbf{v}$.

$$\begin{aligned}
\|\mathbf{v}_{S_0}\|_2^2 &\leq \frac{1}{1-\delta_{2s}} \|A(\mathbf{v}_{S_0})\|_2^2 = \frac{1}{1-\delta_{2s}} \sum_{k \geq 1} \langle A(\mathbf{v}_{S_0}), A(-\mathbf{v}_{S_k}) \rangle \\
&\leq \frac{1}{1-\delta_{2s}} \sum_{k \geq 1} \delta_{2s} \|\mathbf{v}_{S_0}\|_2 \|\mathbf{v}_{S_k}\|_2.
\end{aligned}$$

Simplify by $\|\mathbf{v}_{S_0}\|_2$ and observe that

$$\|\mathbf{v}_{S_k}\|_2 \leq \frac{1}{\sqrt{s}} \|\mathbf{v}_{S_{k-1}}\|_1$$

to obtain

$$\|\mathbf{v}_{S_0}\|_2 \leq \frac{\delta_{2s}}{1-\delta_{2s}} \frac{1}{\sqrt{s}} \|\mathbf{v}\|_1, \quad \text{hence } \|\mathbf{v}_{S_0}\|_1 \leq \frac{\delta_{2s}}{1-\delta_{2s}} \|\mathbf{v}\|_1.$$

Exact Sparse Recovery via ℓ_1-Minimization when $\delta_{2s} < 1/3$

Take $\mathbf{v} \in \ker A \setminus \{\mathbf{0}\}$. Consider sets $S_0, S_1, \ldots$ of s indices ordered by decreasing magnitude of entries of $\mathbf{v}$.

$$\begin{aligned}\|\mathbf{v}_{S_0}\|_2^2 &\leq \frac{1}{1-\delta_{2s}}\|A(\mathbf{v}_{S_0})\|_2^2 = \frac{1}{1-\delta_{2s}}\sum_{k\geq 1}\langle A(\mathbf{v}_{S_0}), A(-\mathbf{v}_{S_k})\rangle \\ &\leq \frac{1}{1-\delta_{2s}}\sum_{k\geq 1}\delta_{2s}\|\mathbf{v}_{S_0}\|_2\|\mathbf{v}_{S_k}\|_2.\end{aligned}$$

Simplify by $\|\mathbf{v}_{S_0}\|_2$ and observe that

$$\|\mathbf{v}_{S_k}\|_2 \leq \frac{1}{\sqrt{s}}\|\mathbf{v}_{S_{k-1}}\|_1$$

to obtain

$$\|\mathbf{v}_{S_0}\|_2 \leq \frac{\delta_{2s}}{1-\delta_{2s}}\frac{1}{\sqrt{s}}\|\mathbf{v}\|_1, \quad \text{hence } \|\mathbf{v}_{S_0}\|_1 \leq \frac{\delta_{2s}}{1-\delta_{2s}}\|\mathbf{v}\|_1.$$

Note that $\rho := \delta_{2s}/(1-\delta_{2s}) < 1/2$ whenever $\delta_{2s} < 1/3$.

Stable and Robust Sparse Recovery via ℓ_1-Minimization

Stable and Robust Sparse Recovery via ℓ_1-Minimization

Objective: for $p \in [1,2]$, for all $\mathbf{x} \in \mathbb{C}^N$ and $\mathbf{e} \in \mathbb{C}^m$ with $\|\mathbf{e}\|_2 \leq \eta$:

$$\|\mathbf{x} - \Delta(A\mathbf{x} + \mathbf{e})\|_p \leq \overbrace{\frac{C}{s^{1-1/p}} \min_{\mathbf{x}_s \; s-\text{sparse}} \|\mathbf{x} - \mathbf{x}_s\|_1}^{\text{stability}} + \overbrace{D \, s^{1/p-1/2} \, \eta}^{\text{robustness}},$$

where $\Delta(\mathbf{y}) = \Delta_{1,\eta}(\mathbf{y}) := \operatorname{argmin} \|\mathbf{z}\|_1 \quad$ subject to $\|A\mathbf{z} - \mathbf{y}\|_2 \leq \eta$.

Stable and Robust Sparse Recovery via ℓ_1-Minimization

Objective: for $p \in [1,2]$, for all $\mathbf{x} \in \mathbb{C}^N$ and $\mathbf{e} \in \mathbb{C}^m$ with $\|\mathbf{e}\|_2 \le \eta$:

$$\|\mathbf{x} - \Delta(A\mathbf{x} + \mathbf{e})\|_p \le \overbrace{\frac{C}{s^{1-1/p}} \min_{\mathbf{x}_s \; s-\text{sparse}} \|\mathbf{x} - \mathbf{x}_s\|_1}^{\text{stability}} + \overbrace{D\, s^{1/p-1/2}\, \eta}^{\text{robustness}},$$

where $\Delta(\mathbf{y}) = \Delta_{1,\eta}(\mathbf{y}) := \operatorname{argmin} \|\mathbf{z}\|_1 \quad$ subject to $\|A\mathbf{z} - \mathbf{y}\|_2 \le \eta$.

Taking $\mathbf{x} = \mathbf{v} \in \mathbb{C}^N$, $\mathbf{e} = -A\mathbf{v} \in \mathbb{C}^m$, and $\eta = \|A\mathbf{v}\|_2$ gives

$$\|\mathbf{v}\|_p \le \frac{C}{s^{1-1/p}} \|\mathbf{v}_{\overline{S}}\|_1 + D\, s^{1/p-1/2} \|A\mathbf{v}\|_2$$

for all $S \subseteq [N]$ with $\operatorname{card}(S) = s$.

Robust Null Space Property

Robust Null Space Property

For $q \in [1,2]$, $A \in \mathbb{C}^{m \times N}$ has the ℓ_q-robust null space property of order s (wrto $\|\cdot\|$) with constants $0 < \rho < 1$ and $\tau > 0$ if, for any set $S \subset [N]$ with $\mathrm{card}(S) \leq s$,

$$\|\mathbf{v}_S\|_q \leq \frac{\rho}{s^{1-1/q}} \|\mathbf{v}_{\overline{S}}\|_1 + \tau \|A\mathbf{v}\| \qquad \text{for all } \mathbf{v} \in \mathbb{C}^N.$$

Robust Null Space Property

For $q \in [1,2]$, $A \in \mathbb{C}^{m \times N}$ has the ℓ_q-robust null space property of order s (wrto $\|\cdot\|$) with constants $0 < \rho < 1$ and $\tau > 0$ if, for any set $S \subset [N]$ with $\mathrm{card}(S) \leq s$,

$$\|\mathbf{v}_S\|_q \leq \frac{\rho}{s^{1-1/q}} \|\mathbf{v}_{\overline{S}}\|_1 + \tau \|A\mathbf{v}\| \qquad \text{for all } \mathbf{v} \in \mathbb{C}^N.$$

- ℓ_q-RNSP (wrto $\|\cdot\|$) $\Rightarrow$ ℓ_p-RSNP (wrto $s^{1/p-1/q}\|\cdot\|$) whenever $1 \leq p \leq q \leq 2$.

Robust Null Space Property

For $q \in [1,2]$, $A \in \mathbb{C}^{m \times N}$ has the ℓ_q-robust null space property of order s (wrto $\|\cdot\|$) with constants $0 < \rho < 1$ and $\tau > 0$ if, for any set $S \subset [N]$ with $\mathrm{card}(S) \leq s$,

$$\|\mathbf{v}_S\|_q \leq \frac{\rho}{s^{1-1/q}} \|\mathbf{v}_{\overline{S}}\|_1 + \tau \|A\mathbf{v}\| \qquad \text{for all } \mathbf{v} \in \mathbb{C}^N.$$

- ℓ_q-RNSP (wrto $\|\cdot\|$) $\Rightarrow$ ℓ_p-RSNP (wrto $s^{1/p-1/q}\|\cdot\|$) whenever $1 \leq p \leq q \leq 2$.
- ℓ_q-RNSP (wrto $\|\cdot\|$) implies that, for $1 \leq p \leq q$,

$$\|\mathbf{z}-\mathbf{x}\|_p \leq \frac{C}{s^{1-1/p}} \big(\|\mathbf{z}\|_1 - \|\mathbf{x}\|_1 + 2\sigma_s(\mathbf{x})_1\big) + D\, s^{1/p-1/q} \|A(\mathbf{z}-\mathbf{x})\|$$

for all $\mathbf{x}, \mathbf{z} \in \mathbb{C}^N$.

Robust Null Space Property

For $q \in [1,2]$, $A \in \mathbb{C}^{m \times N}$ has the ℓ_q-robust null space property of order s (wrto $\|\cdot\|$) with constants $0 < \rho < 1$ and $\tau > 0$ if, for any set $S \subset [N]$ with $\mathrm{card}(S) \leq s$,

$$\|\mathbf{v}_S\|_q \leq \frac{\rho}{s^{1-1/q}} \|\mathbf{v}_{\overline{S}}\|_1 + \tau \|A\mathbf{v}\| \qquad \text{for all } \mathbf{v} \in \mathbb{C}^N.$$

- ℓ_q-RNSP (wrto $\|\cdot\|$) $\Rightarrow$ ℓ_p-RSNP (wrto $s^{1/p-1/q}\|\cdot\|$) whenever $1 \leq p \leq q \leq 2$.
- ℓ_q-RNSP (wrto $\|\cdot\|$) implies that, for $1 \leq p \leq q$,

$$\|\mathbf{z}-\mathbf{x}\|_p \leq \frac{C}{s^{1-1/p}} \left(\|\mathbf{z}\|_1 - \|\mathbf{x}\|_1 + 2\sigma_s(\mathbf{x})_1\right) + D\, s^{1/p-1/q} \|A(\mathbf{z}-\mathbf{x})\|$$

for all $\mathbf{x}, \mathbf{z} \in \mathbb{C}^N$. The converse holds when $q = 1$.

Robust Null Space Property

For $q \in [1,2]$, $A \in \mathbb{C}^{m \times N}$ has the ℓ_q-robust null space property of order s (wrto $\|\cdot\|$) with constants $0 < \rho < 1$ and $\tau > 0$ if, for any set $S \subset [N]$ with $\mathrm{card}(S) \leq s$,

$$\|\mathbf{v}_S\|_q \leq \frac{\rho}{s^{1-1/q}} \|\mathbf{v}_{\overline{S}}\|_1 + \tau \|A\mathbf{v}\| \qquad \text{for all } \mathbf{v} \in \mathbb{C}^N.$$

- ℓ_q-RNSP (wrto $\|\cdot\|$) $\Rightarrow$ ℓ_p-RSNP (wrto $s^{1/p-1/q}\|\cdot\|$) whenever $1 \leq p \leq q \leq 2$.
- ℓ_q-RNSP (wrto $\|\cdot\|$) implies that, for $1 \leq p \leq q$,

$$\|\mathbf{z}-\mathbf{x}\|_p \leq \frac{C}{s^{1-1/p}} \big(\|\mathbf{z}\|_1 - \|\mathbf{x}\|_1 + 2\sigma_s(\mathbf{x})_1\big) + D\, s^{1/p-1/q} \|A(\mathbf{z}-\mathbf{x})\|$$

  for all $\mathbf{x}, \mathbf{z} \in \mathbb{C}^N$. The converse holds when $q = 1$.
- ℓ_2-RNSP (wrto $\|\cdot\|_2$) holds when $\delta_{2s} < 1/\sqrt{2}$.

Iterative Hard Thresholding and Hard Thresholding Pursuit

Iterative Hard Thresholding and Hard Thresholding Pursuit

- solving the rectangular system $A\mathbf{x} = \mathbf{y}$ amounts to solving the square system $A^*A\mathbf{x} = A^*\mathbf{y}$,

Iterative Hard Thresholding and Hard Thresholding Pursuit

- solving the rectangular system $A\mathbf{x} = \mathbf{y}$ amounts to solving the square system $A^* A\mathbf{x} = A^*\mathbf{y}$,
- classical iterative methods suggest the iteration $\mathbf{x}^{n+1} = \mathbf{x}^n + A^*(\mathbf{y} - A\mathbf{x}^n)$,

Iterative Hard Thresholding and Hard Thresholding Pursuit

- solving the rectangular system $A\mathbf{x} = \mathbf{y}$ amounts to solving the square system $A^* A\mathbf{x} = A^*\mathbf{y}$,
- classical iterative methods suggest the iteration $\mathbf{x}^{n+1} = \mathbf{x}^n + A^*(\mathbf{y} - A\mathbf{x}^n)$,
- at each iteration, keep s largest absolute entries and set the other ones to zero.

Iterative Hard Thresholding and Hard Thresholding Pursuit

- solving the rectangular system $A\mathbf{x} = \mathbf{y}$ amounts to solving the square system $A^*A\mathbf{x} = A^*\mathbf{y}$,
- classical iterative methods suggest the iteration $\mathbf{x}^{n+1} = \mathbf{x}^n + A^*(\mathbf{y} - A\mathbf{x}^n)$,
- at each iteration, keep s largest absolute entries and set the other ones to zero.

IHT: Start with an s-sparse $\mathbf{x}^0 \in \mathbb{C}^N$ and iterate:

$$\text{(IHT)} \qquad \mathbf{x}^{n+1} = H_s(\mathbf{x}^n + A^*(\mathbf{y} - A\mathbf{x}^n))$$

until a stopping criterion is met.

Iterative Hard Thresholding and Hard Thresholding Pursuit

- solving the rectangular system $A\mathbf{x} = \mathbf{y}$ amounts to solving the square system $A^* A\mathbf{x} = A^*\mathbf{y}$,
- classical iterative methods suggest the iteration $\mathbf{x}^{n+1} = \mathbf{x}^n + A^*(\mathbf{y} - A\mathbf{x}^n)$,
- at each iteration, keep s largest absolute entries and set the other ones to zero.

IHT: Start with an s-sparse $\mathbf{x}^0 \in \mathbb{C}^N$ and iterate:

$$\text{(IHT)} \qquad \mathbf{x}^{n+1} = H_s(\mathbf{x}^n + A^*(\mathbf{y} - A\mathbf{x}^n))$$

until a stopping criterion is met.

HTP: Start with an s-sparse $\mathbf{x}^0 \in \mathbb{C}^N$ and iterate:

Iterative Hard Thresholding and Hard Thresholding Pursuit

- solving the rectangular system $A\mathbf{x} = \mathbf{y}$ amounts to solving the square system $A^*A\mathbf{x} = A^*\mathbf{y}$,
- classical iterative methods suggest the iteration $\mathbf{x}^{n+1} = \mathbf{x}^n + A^*(\mathbf{y} - A\mathbf{x}^n)$,
- at each iteration, keep s largest absolute entries and set the other ones to zero.

IHT: Start with an s-sparse $\mathbf{x}^0 \in \mathbb{C}^N$ and iterate:

(IHT) $$\mathbf{x}^{n+1} = H_s(\mathbf{x}^n + A^*(\mathbf{y} - A\mathbf{x}^n))$$

until a stopping criterion is met.

HTP: Start with an s-sparse $\mathbf{x}^0 \in \mathbb{C}^N$ and iterate:

(HTP_1)

(HTP_2)

Iterative Hard Thresholding and Hard Thresholding Pursuit

- solving the rectangular system $A\mathbf{x} = \mathbf{y}$ amounts to solving the square system $A^*A\mathbf{x} = A^*\mathbf{y}$,
- classical iterative methods suggest the iteration $\mathbf{x}^{n+1} = \mathbf{x}^n + A^*(\mathbf{y} - A\mathbf{x}^n)$,
- at each iteration, keep s largest absolute entries and set the other ones to zero.

IHT: Start with an s-sparse $\mathbf{x}^0 \in \mathbb{C}^N$ and iterate:

$$\text{(IHT)} \qquad \mathbf{x}^{n+1} = H_s(\mathbf{x}^n + A^*(\mathbf{y} - A\mathbf{x}^n))$$

until a stopping criterion is met.

HTP: Start with an s-sparse $\mathbf{x}^0 \in \mathbb{C}^N$ and iterate:

$$(\text{HTP}_1) \quad S^{n+1} = \left\{ s \text{ largest abs. entries of } \mathbf{x}^n + A^*(\mathbf{y} - A\mathbf{x}^n) \right\},$$

(HTP_2)

Iterative Hard Thresholding and Hard Thresholding Pursuit

- solving the rectangular system $A\mathbf{x} = \mathbf{y}$ amounts to solving the square system $A^*A\mathbf{x} = A^*\mathbf{y}$,
- classical iterative methods suggest the iteration $\mathbf{x}^{n+1} = \mathbf{x}^n + A^*(\mathbf{y} - A\mathbf{x}^n)$,
- at each iteration, keep s largest absolute entries and set the other ones to zero.

IHT: Start with an s-sparse $\mathbf{x}^0 \in \mathbb{C}^N$ and iterate:

$$\text{(IHT)} \qquad \mathbf{x}^{n+1} = H_s(\mathbf{x}^n + A^*(\mathbf{y} - A\mathbf{x}^n))$$

until a stopping criterion is met.

HTP: Start with an s-sparse $\mathbf{x}^0 \in \mathbb{C}^N$ and iterate:

$$\text{(HTP}_1\text{)} \quad S^{n+1} = \{s \text{ largest abs. entries of } \mathbf{x}^n + A^*(\mathbf{y} - A\mathbf{x}^n)\},$$
$$\text{(HTP}_2\text{)} \quad \mathbf{x}^{n+1} = \operatorname{argmin}\{\|\mathbf{y} - A\mathbf{z}\|_2, \operatorname{supp}(\mathbf{z}) \subseteq S^{n+1}\},$$

Iterative Hard Thresholding and Hard Thresholding Pursuit

- solving the rectangular system $A\mathbf{x} = \mathbf{y}$ amounts to solving the square system $A^*A\mathbf{x} = A^*\mathbf{y}$,
- classical iterative methods suggest the iteration $\mathbf{x}^{n+1} = \mathbf{x}^n + A^*(\mathbf{y} - A\mathbf{x}^n)$,
- at each iteration, keep s largest absolute entries and set the other ones to zero.

IHT: Start with an s-sparse $\mathbf{x}^0 \in \mathbb{C}^N$ and iterate:

$$\text{(IHT)} \qquad \mathbf{x}^{n+1} = H_s(\mathbf{x}^n + A^*(\mathbf{y} - A\mathbf{x}^n))$$

until a stopping criterion is met.

HTP: Start with an s-sparse $\mathbf{x}^0 \in \mathbb{C}^N$ and iterate:

$$(\text{HTP}_1) \quad S^{n+1} = \big\{ s \text{ largest abs. entries of } \mathbf{x}^n + A^*(\mathbf{y} - A\mathbf{x}^n) \big\},$$

$$(\text{HTP}_2) \quad \mathbf{x}^{n+1} = \operatorname{argmin}\big\{ \|\mathbf{y} - A\mathbf{z}\|_2, \operatorname{supp}(\mathbf{z}) \subseteq S^{n+1} \big\},$$

until a stopping criterion is met ($S^{n+1} = S^n$ is natural here).

Stable and Robust Sparse Recovery via IHT and HTP

Stable and Robust Sparse Recovery via IHT and HTP

Suppose that $\delta_{3s} < 1/\sqrt{3}$.

Stable and Robust Sparse Recovery via IHT and HTP

Suppose that $\delta_{3s} < 1/\sqrt{3}$. Then, for some $\rho < 1$ and $\tau > 0$, the following facts hold for all $\mathbf{x} \in \mathbb{C}^N$ and $\mathbf{e} \in \mathbb{C}^m$, where S denotes an index set of s largest absolute entries of $\mathbf{x}$:

Stable and Robust Sparse Recovery via IHT and HTP

Suppose that $\delta_{3s} < 1/\sqrt{3}$. Then, for some $\rho < 1$ and $\tau > 0$, the following facts hold for all $\mathbf{x} \in \mathbb{C}^N$ and $\mathbf{e} \in \mathbb{C}^m$, where S denotes an index set of s largest absolute entries of $\mathbf{x}$:

- $\|\mathbf{x}_S - \mathbf{x}^{n+1}\|_2 \leq \rho \|\mathbf{x}_S - \mathbf{x}^n\|_2 + (1-\rho)\tau \|A\mathbf{x}_{\overline{S}} + \mathbf{e}\|_2,$

Stable and Robust Sparse Recovery via IHT and HTP

Suppose that $\delta_{3s} < 1/\sqrt{3}$. Then, for some $\rho < 1$ and $\tau > 0$, the following facts hold for all $\mathbf{x} \in \mathbb{C}^N$ and $\mathbf{e} \in \mathbb{C}^m$, where S denotes an index set of s largest absolute entries of $\mathbf{x}$:

- $\|\mathbf{x}_S - \mathbf{x}^{n+1}\|_2 \leq \rho \|\mathbf{x}_S - \mathbf{x}^n\|_2 + (1-\rho)\tau \|A\mathbf{x}_{\overline{S}} + \mathbf{e}\|_2$,
- $\|\mathbf{x}_S - \mathbf{x}^n\|_2 \leq \rho^n \|\mathbf{x}_S - \mathbf{x}^0\|_2 + \tau \|A\mathbf{x}_{\overline{S}} + \mathbf{e}\|_2$,

Stable and Robust Sparse Recovery via IHT and HTP

Suppose that $\delta_{3s} < 1/\sqrt{3}$. Then, for some $\rho < 1$ and $\tau > 0$, the following facts hold for all $\mathbf{x} \in \mathbb{C}^N$ and $\mathbf{e} \in \mathbb{C}^m$, where S denotes an index set of s largest absolute entries of $\mathbf{x}$:

- $\|\mathbf{x}_S - \mathbf{x}^{n+1}\|_2 \leq \rho \|\mathbf{x}_S - \mathbf{x}^n\|_2 + (1-\rho)\tau \|A\mathbf{x}_{\overline{S}} + \mathbf{e}\|_2,$
- $\|\mathbf{x}_S - \mathbf{x}^n\|_2 \leq \rho^n \|\mathbf{x}_S - \mathbf{x}^0\|_2 + \tau \|A\mathbf{x}_{\overline{S}} + \mathbf{e}\|_2,$
- The outpout $\mathbf{x}^\sharp$ of the algorithms satisfy

$$\|\mathbf{x}_S - \mathbf{x}^\sharp\|_2 \leq \tau \|A\mathbf{x}_{\overline{S}} + \mathbf{e}\|_2,$$

Stable and Robust Sparse Recovery via IHT and HTP

Suppose that $\delta_{3s} < 1/\sqrt{3}$. Then, for some $\rho < 1$ and $\tau > 0$, the following facts hold for all $\mathbf{x} \in \mathbb{C}^N$ and $\mathbf{e} \in \mathbb{C}^m$, where S denotes an index set of s largest absolute entries of $\mathbf{x}$:

- $\|\mathbf{x}_S - \mathbf{x}^{n+1}\|_2 \leq \rho \|\mathbf{x}_S - \mathbf{x}^n\|_2 + (1-\rho)\tau \|A\mathbf{x}_{\overline{S}} + \mathbf{e}\|_2$,
- $\|\mathbf{x}_S - \mathbf{x}^n\|_2 \leq \rho^n \|\mathbf{x}_S - \mathbf{x}^0\|_2 + \tau \|A\mathbf{x}_{\overline{S}} + \mathbf{e}\|_2$,
- The outpout $\mathbf{x}^\sharp$ of the algorithms satisfy

$$\|\mathbf{x}_S - \mathbf{x}^\sharp\|_2 \leq \tau \|A\mathbf{x}_{\overline{S}} + \mathbf{e}\|_2,$$

- With $2s$ in place of s in the algorithms, if $\delta_{6s} < 1/\sqrt{3}$, then

$$\|\mathbf{x} - \mathbf{x}^\sharp\|_p \leq \frac{C}{s^{1-1/p}} \sigma_s(\mathbf{x})_1 + D\, s^{1/p-1/2} \|\mathbf{e}\|_2, \quad 1 \leq p \leq 2,$$

Stable and Robust Sparse Recovery via IHT and HTP

Suppose that $\delta_{3s} < 1/\sqrt{3}$. Then, for some $\rho < 1$ and $\tau > 0$, the following facts hold for all $\mathbf{x} \in \mathbb{C}^N$ and $\mathbf{e} \in \mathbb{C}^m$, where S denotes an index set of s largest absolute entries of $\mathbf{x}$:

- $\|\mathbf{x}_S - \mathbf{x}^{n+1}\|_2 \leq \rho \|\mathbf{x}_S - \mathbf{x}^n\|_2 + (1-\rho)\tau \|A\mathbf{x}_{\overline{S}} + \mathbf{e}\|_2$,
- $\|\mathbf{x}_S - \mathbf{x}^n\|_2 \leq \rho^n \|\mathbf{x}_S - \mathbf{x}^0\|_2 + \tau \|A\mathbf{x}_{\overline{S}} + \mathbf{e}\|_2$,
- The outpout $\mathbf{x}^\sharp$ of the algorithms satisfy

$$\|\mathbf{x}_S - \mathbf{x}^\sharp\|_2 \leq \tau \|A\mathbf{x}_{\overline{S}} + \mathbf{e}\|_2,$$

- With $2s$ in place of s in the algorithms, if $\delta_{6s} < 1/\sqrt{3}$, then

$$\|\mathbf{x} - \mathbf{x}^\sharp\|_p \leq \frac{C}{s^{1-1/p}} \sigma_s(\mathbf{x})_1 + D\, s^{1/p-1/2} \|\mathbf{e}\|_2, \quad 1 \leq p \leq 2,$$

- For HTP, (pseudo)robustness is achieved in $\leq c\, s$ iterations.

Recovery by Orthogonal Matching Pursuit

Recovery by Orthogonal Matching Pursuit

OMP: Start with $S^0 = \emptyset$ and $\mathbf{x}^0 = \mathbf{0}$, and iterate:

$$S^n = S^{n-1} \cup \{j^n := \mathrm{argmax}_j \big|\big(A^*(\mathbf{y} - A\mathbf{x}^{n-1})\big)_j\big|\},$$
$$\mathbf{x^n} = \mathrm{argmin}_{\mathbf{z}}\{\|\mathbf{y} - A\mathbf{z}\|_2, \mathrm{supp}(\mathbf{z}) \subseteq S^n\}.$$

Recovery by Orthogonal Matching Pursuit

OMP: Start with $S^0 = \emptyset$ and $\mathbf{x}^0 = \mathbf{0}$, and iterate:

$$S^n = S^{n-1} \cup \left\{ j^n := \mathrm{argmax}_j \left| \left(A^*(\mathbf{y} - A\mathbf{x}^{n-1}) \right)_j \right| \right\},$$
$$\mathbf{x^n} = \mathrm{argmin}_{\mathbf{z}} \left\{ \|\mathbf{y} - A\mathbf{z}\|_2, \mathrm{supp}(\mathbf{z}) \subseteq S^n \right\}.$$

If A has ℓ_2-normalized columns $\mathbf{a}_1, \ldots, \mathbf{a}_N$, then

$$\|\mathbf{y} - A\mathbf{x}^n\|_2^2 = \|\mathbf{y} - A\mathbf{x}^{n-1}\|_2^2 - \frac{\left| \left(A^*(\mathbf{y} - A\mathbf{x}^{n-1}) \right)_{j^n} \right|^2}{d(\mathbf{a}_{j^n}, \mathrm{span}[\mathbf{a}_j, j \in S^{n-1}])^2}$$

Recovery by Orthogonal Matching Pursuit

OMP: Start with $S^0 = \emptyset$ and $\mathbf{x}^0 = \mathbf{0}$, and iterate:

$$S^n = S^{n-1} \cup \left\{ j^n := \mathrm{argmax}_j \left| \left(A^*(\mathbf{y} - A\mathbf{x}^{n-1}) \right)_j \right| \right\},$$
$$\mathbf{x}^{\mathbf{n}} = \mathrm{argmin}_{\mathbf{z}} \left\{ \|\mathbf{y} - A\mathbf{z}\|_2, \mathrm{supp}(\mathbf{z}) \subseteq S^n \right\}.$$

If A has ℓ_2-normalized columns $\mathbf{a}_1, \ldots, \mathbf{a}_N$, then

$$\|\mathbf{y} - A\mathbf{x}^n\|_2^2 = \|\mathbf{y} - A\mathbf{x}^{n-1}\|_2^2 - \frac{\left| \left(A^*(\mathbf{y} - A\mathbf{x}^{n-1}) \right)_{j^n} \right|^2}{d(\mathbf{a}_{j^n}, \mathrm{span}[\mathbf{a}_j, j \in S^{n-1}])^2}$$
$$\leq \|\mathbf{y} - A\mathbf{x}^{n-1}\|_2^2 - \left| \left(A^*(\mathbf{y} - A\mathbf{x}^{n-1}) \right)_{j^n} \right|^2.$$

Recovery by Orthogonal Matching Pursuit

OMP: Start with $S^0 = \emptyset$ and $\mathbf{x}^0 = \mathbf{0}$, and iterate:

$$S^n = S^{n-1} \cup \left\{ j^n := \mathrm{argmax}_j \left| \left(A^*(\mathbf{y} - A\mathbf{x}^{n-1}) \right)_j \right| \right\},$$
$$\mathbf{x}^{\mathbf{n}} = \mathrm{argmin}_{\mathbf{z}} \left\{ \|\mathbf{y} - A\mathbf{z}\|_2, \mathrm{supp}(\mathbf{z}) \subseteq S^n \right\}.$$

If A has ℓ_2-normalized columns $\mathbf{a}_1, \ldots, \mathbf{a}_N$, then

$$\begin{aligned} \|\mathbf{y} - A\mathbf{x}^n\|_2^2 &= \|\mathbf{y} - A\mathbf{x}^{n-1}\|_2^2 - \frac{\left| \left(A^*(\mathbf{y} - A\mathbf{x}^{n-1}) \right)_{j^n} \right|^2}{d(\mathbf{a}_{j^n}, \mathrm{span}[\mathbf{a}_j, j \in S^{n-1}])^2} \\ &\leq \|\mathbf{y} - A\mathbf{x}^{n-1}\|_2^2 - \left| \left(A^*(\mathbf{y} - A\mathbf{x}^{n-1}) \right)_{j^n} \right|^2. \end{aligned}$$

- Suppose that $\delta_{13s} < 1/6$. Then s-sparse recovery via $12s$ iterations of OMP is stable and robust.

Recovery by Orthogonal Matching Pursuit

OMP: Start with $S^0 = \emptyset$ and $\mathbf{x}^0 = \mathbf{0}$, and iterate:

$$S^n = S^{n-1} \cup \left\{ j^n := \operatorname{argmax}_j \left| \left(A^*(\mathbf{y} - A\mathbf{x}^{n-1}) \right)_j \right| \right\},$$
$$\mathbf{x}^{\mathbf{n}} = \operatorname{argmin}_{\mathbf{z}} \left\{ \|\mathbf{y} - A\mathbf{z}\|_2, \operatorname{supp}(\mathbf{z}) \subseteq S^n \right\}.$$

If A has ℓ_2-normalized columns $\mathbf{a}_1, \ldots, \mathbf{a}_N$, then

$$\begin{aligned} \|\mathbf{y} - A\mathbf{x}^n\|_2^2 &= \|\mathbf{y} - A\mathbf{x}^{n-1}\|_2^2 - \frac{\left| \left(A^*(\mathbf{y} - A\mathbf{x}^{n-1}) \right)_{j^n} \right|^2}{d(\mathbf{a}_{j^n}, \operatorname{span}[\mathbf{a}_j, j \in S^{n-1}])^2} \\ &\leq \|\mathbf{y} - A\mathbf{x}^{n-1}\|_2^2 - \left| \left(A^*(\mathbf{y} - A\mathbf{x}^{n-1}) \right)_{j^n} \right|^2. \end{aligned}$$

- Suppose that $\delta_{13s} < 1/6$. Then s-sparse recovery via $12s$ iterations of OMP is stable and robust.
- Challenge: natural proof of OMP success in $c\, s$ iterations?