A Short Introduction to Compressive Sensing

Simon Foucart
University of Georgia

Semester Program on "High-Dimensional Approximation"
ICERM
26 September 2014

Compressive Sensing

Compressive Sensing

Keywords:

Compressive Sensing

Keywords:

- Sparsity (Essential)

Compressive Sensing

Keywords:

- Sparsity
  (Essential)
- Randomness
  (Nothing better so far)

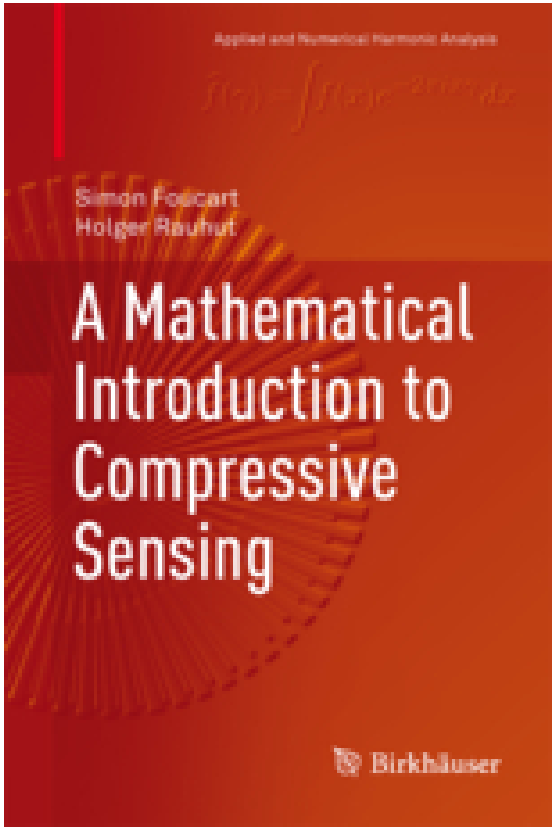
Compressive Sensing

Keywords:

- Sparsity
  (Essential)
- Randomness
  (Nothing better so far)
- Optimization
  (Merely preferred,
  but alternatives exist)

Compressive Sensing

Classical theory now found in:



Keywords:

- Sparsity
  (Essential)
- Randomness
  (Nothing better so far)
- Optimization
  (Merely preferred,
  but alternatives exist)

Part I:

The Standard Problem and First Algorithms

Initial Analysis

The Standard Compressive Sensing Problem

The Standard Compressive Sensing Problem

$\mathbf{x}$: unknown signal of interest in $\mathbb{K}^N$

The Standard Compressive Sensing Problem

$\mathbf{x}$: unknown signal of interest in $\mathbb{K}^N$

$\mathbf{y}$: measurement vector in $\mathbb{K}^m$

The Standard Compressive Sensing Problem

$\mathbf{x}$: unknown signal of interest in $\mathbb{K}^N$

$\mathbf{y}$: measurement vector in $\mathbb{K}^m$ with $m \ll N$,

The Standard Compressive Sensing Problem

$\mathbf{x}$	:	unknown signal of interest in $\mathbb{K}^N$
$\mathbf{y}$	:	measurement vector in $\mathbb{K}^m$ with $m \ll N$,
s	:	sparsity of $\mathbf{x}$

The Standard Compressive Sensing Problem

$\mathbf{x}$: unknown signal of interest in $\mathbb{K}^N$

$\mathbf{y}$: measurement vector in $\mathbb{K}^m$ with $m \ll N$,

s : sparsity of $\mathbf{x} = \mathrm{card}\big\{j \in \{1, \ldots, N\} : x_j \neq 0\big\}$.

The Standard Compressive Sensing Problem

$$\begin{aligned} \mathbf{x} &: \text{ unknown signal of interest in } \mathbb{K}^N \\ \mathbf{y} &: \text{ measurement vector in } \mathbb{K}^m \text{ with } m \ll N, \\ s &: \text{ sparsity of } \mathbf{x} = \mathrm{card}\big\{j \in \{1, \ldots, N\} : x_j \neq 0\big\}. \end{aligned}$$

Find concrete sensing/recovery protocols,

The Standard Compressive Sensing Problem

$$\begin{aligned} \mathbf{x} &: \text{ unknown signal of interest in } \mathbb{K}^N \\ \mathbf{y} &: \text{ measurement vector in } \mathbb{K}^m \text{ with } m \ll N, \\ s &: \text{ sparsity of } \mathbf{x} = \mathrm{card}\big\{j \in \{1, \ldots, N\} : x_j \neq 0\big\}. \end{aligned}$$

Find concrete sensing/recovery protocols, i.e., find

- measurement matrices $A : \mathbf{x} \in \mathbb{K}^N \mapsto \mathbf{y} \in \mathbb{K}^m$

The Standard Compressive Sensing Problem

$$\begin{aligned} \mathbf{x} &: \text{ unknown signal of interest in } \mathbb{K}^N \\ \mathbf{y} &: \text{ measurement vector in } \mathbb{K}^m \text{ with } m \ll N, \\ s &: \text{ sparsity of } \mathbf{x} = \mathrm{card}\big\{j \in \{1, \ldots, N\} : x_j \neq 0\big\}. \end{aligned}$$

Find concrete sensing/recovery protocols, i.e., find

- measurement matrices $A : \mathbf{x} \in \mathbb{K}^N \mapsto \mathbf{y} \in \mathbb{K}^m$
- reconstruction maps $\Delta : \mathbf{y} \in \mathbb{K}^m \mapsto \mathbf{x} \in \mathbb{K}^N$

The Standard Compressive Sensing Problem

$$\begin{aligned} \mathbf{x} &: \text{ unknown signal of interest in } \mathbb{K}^N \\ \mathbf{y} &: \text{ measurement vector in } \mathbb{K}^m \text{ with } m \ll N, \\ s &: \text{ sparsity of } \mathbf{x} = \mathrm{card}\big\{j \in \{1, \ldots, N\} : x_j \neq 0\big\}. \end{aligned}$$

Find concrete sensing/recovery protocols, i.e., find

- measurement matrices $A : \mathbf{x} \in \mathbb{K}^N \mapsto \mathbf{y} \in \mathbb{K}^m$
- reconstruction maps $\Delta : \mathbf{y} \in \mathbb{K}^m \mapsto \mathbf{x} \in \mathbb{K}^N$

such that

$$\Delta(A\mathbf{x}) = \mathbf{x} \qquad \text{for any } s\text{-sparse vector } \mathbf{x} \in \mathbb{K}^N.$$

The Standard Compressive Sensing Problem

$$\begin{aligned} \mathbf{x} &: \text{ unknown signal of interest in } \mathbb{K}^N \\ \mathbf{y} &: \text{ measurement vector in } \mathbb{K}^m \text{ with } m \ll N, \\ s &: \text{ sparsity of } \mathbf{x} = \mathrm{card}\big\{j \in \{1, \ldots, N\} : x_j \neq 0\big\}. \end{aligned}$$

Find concrete sensing/recovery protocols, i.e., find

- measurement matrices $A : \mathbf{x} \in \mathbb{K}^N \mapsto \mathbf{y} \in \mathbb{K}^m$
- reconstruction maps $\Delta : \mathbf{y} \in \mathbb{K}^m \mapsto \mathbf{x} \in \mathbb{K}^N$

such that

$$\Delta(A\mathbf{x}) = \mathbf{x} \qquad \text{for any } s\text{-sparse vector } \mathbf{x} \in \mathbb{K}^N.$$

In realistic situations, two issues to consider:

The Standard Compressive Sensing Problem

$$\begin{aligned} \mathbf{x} &: \text{ unknown signal of interest in } \mathbb{K}^N \\ \mathbf{y} &: \text{ measurement vector in } \mathbb{K}^m \text{ with } m \ll N, \\ s &: \text{ sparsity of } \mathbf{x} = \mathrm{card}\big\{j \in \{1, \ldots, N\} : x_j \neq 0\big\}. \end{aligned}$$

Find concrete sensing/recovery protocols, i.e., find

- measurement matrices $A : \mathbf{x} \in \mathbb{K}^N \mapsto \mathbf{y} \in \mathbb{K}^m$
- reconstruction maps $\Delta : \mathbf{y} \in \mathbb{K}^m \mapsto \mathbf{x} \in \mathbb{K}^N$

such that

$$\Delta(A\mathbf{x}) = \mathbf{x} \qquad \text{for any } s\text{-sparse vector } \mathbf{x} \in \mathbb{K}^N.$$

In realistic situations, two issues to consider:

Stability:	$\mathbf{x}$ not sparse but compressible,
Robustness:	measurement error in $\mathbf{y} = A\mathbf{x} + \mathbf{e}$.

A Selection of Applications*

A Selection of Applications*

- Magnetic resonance imaging

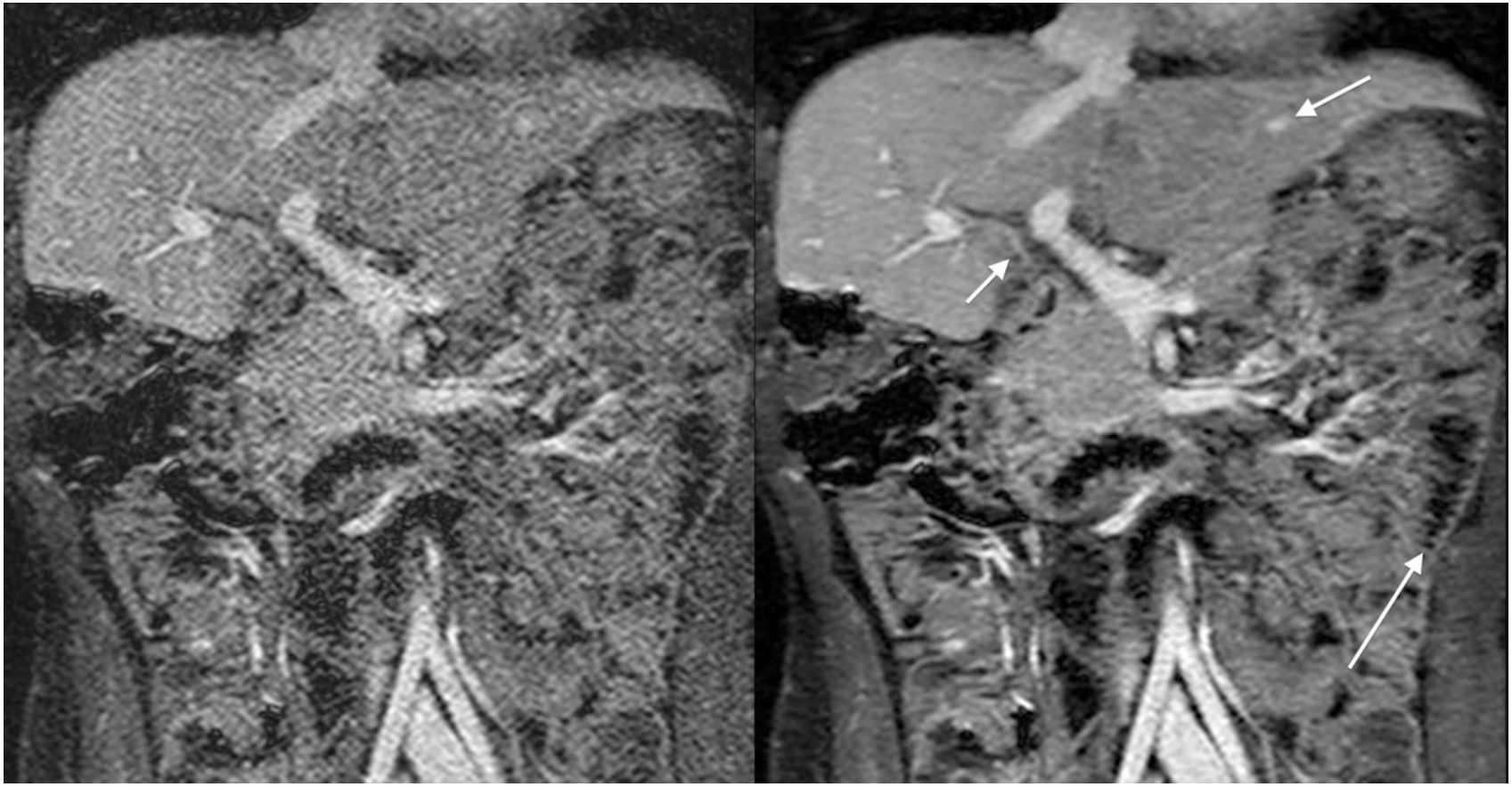


Figure : Left: traditional MRI reconstruction; Right: compressive sensing reconstruction (courtesy of M. Lustig and S. Vasanawala)

A Selection of Applications*

- Magnetic resonance imaging
- Sampling theory

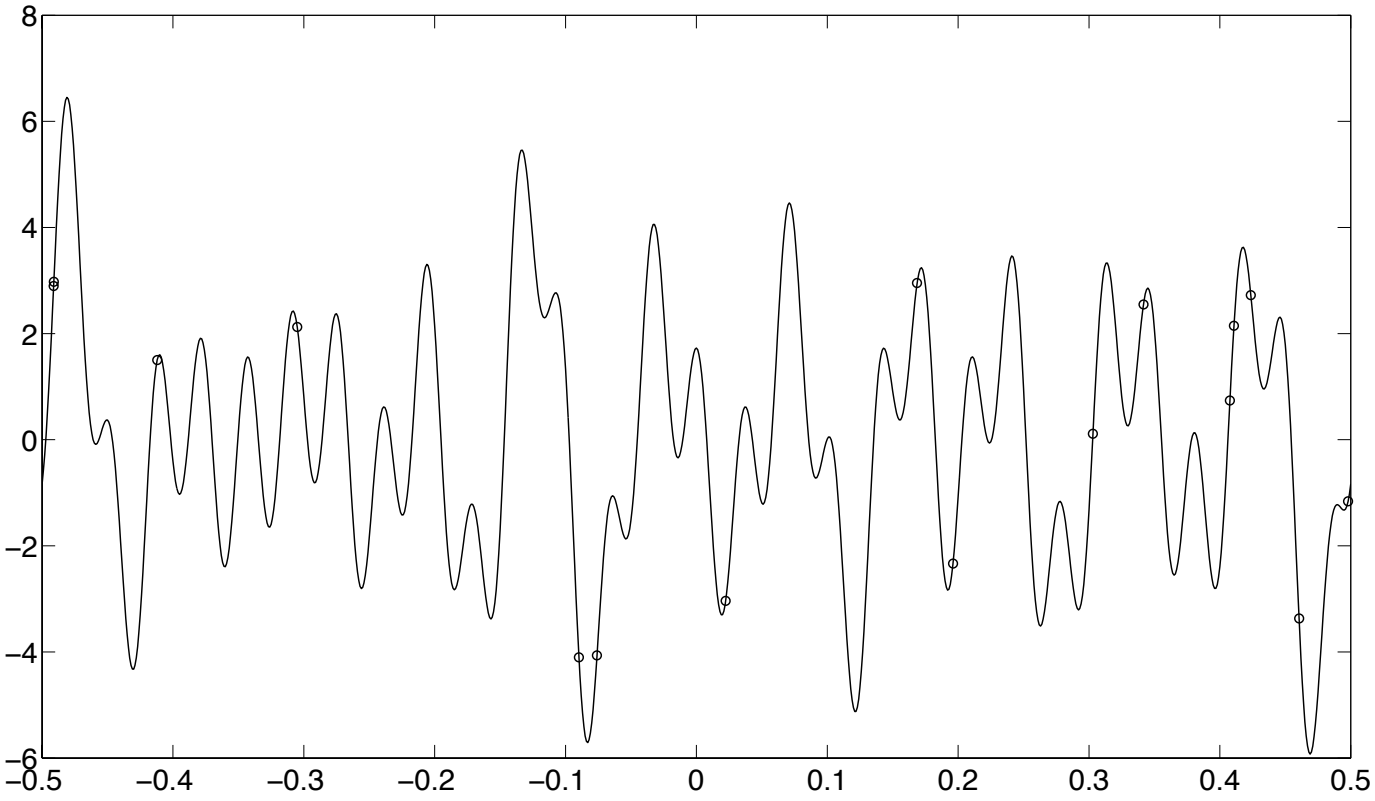


Figure : Time-domain signal with 16 samples.

A Selection of Applications*

- Magnetic resonance imaging
- Sampling theory
- Error correction

A Selection of Applications*

- Magnetic resonance imaging
- Sampling theory
- Error correction
- and many more...

ℓ_0-Minimization

Since

$$\|\mathbf{x}\|_p^p := \sum_{j=1}^{N} |x_j|^p \xrightarrow[p\to 0]{} \sum_{j=1}^{N} \mathbf{1}_{\{x_j \neq 0\}},$$

ℓ_0-Minimization

Since

$$\|\mathbf{x}\|_p^p := \sum_{j=1}^{N} |x_j|^p \xrightarrow[p\to 0]{} \sum_{j=1}^{N} \mathbf{1}_{\{x_j \neq 0\}},$$

the notation $\|\mathbf{x}\|_0$ *[sic]* has become usual for

$$\|\mathbf{x}\|_0 := \mathrm{card}(\mathrm{supp}(\mathbf{x})), \quad \text{where } \mathrm{supp}(\mathbf{x}) := \{j \in [N] : x_j \neq 0\}.$$

ℓ_0-Minimization

Since

$$\|\mathbf{x}\|_p^p := \sum_{j=1}^N |x_j|^p \xrightarrow[p\to 0]{} \sum_{j=1}^N \mathbf{1}_{\{x_j \neq 0\}},$$

the notation $\|\mathbf{x}\|_0$ *[sic]* has become usual for

$$\|\mathbf{x}\|_0 := \mathrm{card}(\mathrm{supp}(\mathbf{x})), \quad \text{where } \mathrm{supp}(\mathbf{x}) := \{j \in [N] : x_j \neq 0\}.$$

For an s-sparse $\mathbf{x} \in \mathbb{K}^N$, observe the equivalence of

ℓ_0-Minimization

Since

$$\|\mathbf{x}\|_p^p := \sum_{j=1}^N |x_j|^p \xrightarrow[p\to 0]{} \sum_{j=1}^N \mathbf{1}_{\{x_j \neq 0\}},$$

the notation $\|\mathbf{x}\|_0$ *[sic]* has become usual for

$$\|\mathbf{x}\|_0 := \mathrm{card}(\mathrm{supp}(\mathbf{x})), \quad \text{where } \mathrm{supp}(\mathbf{x}) := \{j \in [N] : x_j \neq 0\}.$$

For an s-sparse $\mathbf{x} \in \mathbb{K}^N$, observe the equivalence of

- $\mathbf{x}$ is the unique s-sparse solution of $A\mathbf{z} = \mathbf{y}$ with $\mathbf{y} = A\mathbf{x}$,

ℓ_0-Minimization

Since

$$\|\mathbf{x}\|_p^p := \sum_{j=1}^N |x_j|^p \xrightarrow[p\to 0]{} \sum_{j=1}^N \mathbf{1}_{\{x_j \neq 0\}},$$

the notation $\|\mathbf{x}\|_0$ *[sic]* has become usual for

$$\|\mathbf{x}\|_0 := \mathrm{card}(\mathrm{supp}(\mathbf{x})), \quad \text{where } \mathrm{supp}(\mathbf{x}) := \{j \in [N] : x_j \neq 0\}.$$

For an s-sparse $\mathbf{x} \in \mathbb{K}^N$, observe the equivalence of

- $\mathbf{x}$ is the unique s-sparse solution of $A\mathbf{z} = \mathbf{y}$ with $\mathbf{y} = A\mathbf{x}$,
- $\mathbf{x}$ can be reconstructed as the unique solution of

$$(\mathrm{P}_0) \qquad \underset{\mathbf{z}\in\mathbb{K}^N}{\text{minimize}} \ \|\mathbf{z}\|_0 \quad \text{subject to } A\mathbf{z} = \mathbf{y}.$$

ℓ_0-Minimization

Since

$$\|\mathbf{x}\|_p^p := \sum_{j=1}^N |x_j|^p \xrightarrow[p\to 0]{} \sum_{j=1}^N \mathbf{1}_{\{x_j \neq 0\}},$$

the notation $\|\mathbf{x}\|_0$ *[sic]* has become usual for

$$\|\mathbf{x}\|_0 := \mathrm{card}(\mathrm{supp}(\mathbf{x})), \quad \text{where } \mathrm{supp}(\mathbf{x}) := \{j \in [N] : x_j \neq 0\}.$$

For an s-sparse $\mathbf{x} \in \mathbb{K}^N$, observe the equivalence of

- $\mathbf{x}$ is the unique s-sparse solution of $A\mathbf{z} = \mathbf{y}$ with $\mathbf{y} = A\mathbf{x}$,
- $\mathbf{x}$ can be reconstructed as the unique solution of

$$(\mathrm{P}_0) \qquad \underset{\mathbf{z}\in\mathbb{K}^N}{\text{minimize}} \ \|\mathbf{z}\|_0 \quad \text{subject to } A\mathbf{z} = \mathbf{y}.$$

This is a combinatorial problem, NP-hard in general.

Minimal Number of Measurements

Minimal Number of Measurements

Given $A \in \mathbb{K}^{m \times N}$, the following are equivalent:

Minimal Number of Measurements

Given $A \in \mathbb{K}^{m \times N}$, the following are equivalent:

1. Every s-sparse $\mathbf{x}$ is the unique s-sparse solution of $A\mathbf{z} = A\mathbf{x}$,

Minimal Number of Measurements

Given $A \in \mathbb{K}^{m \times N}$, the following are equivalent:

1. Every s-sparse $\mathbf{x}$ is the unique s-sparse solution of $A\mathbf{z} = A\mathbf{x}$,
2. $\ker A \cap \{\mathbf{z} \in \mathbb{K}^N : \|\mathbf{z}\|_0 \leq 2s\} = \{\mathbf{0}\}$,

Minimal Number of Measurements

Given $A \in \mathbb{K}^{m \times N}$, the following are equivalent:

1. Every s-sparse $\mathbf{x}$ is the unique s-sparse solution of $A\mathbf{z} = A\mathbf{x}$,
2. $\ker A \cap \{\mathbf{z} \in \mathbb{K}^N : \|\mathbf{z}\|_0 \leq 2s\} = \{\mathbf{0}\}$,
3. For any $S \subset [N]$ with $\mathrm{card}(S) \leq 2s$, the matrix A_S is injective,

Minimal Number of Measurements

Given $A \in \mathbb{K}^{m \times N}$, the following are equivalent:

1. Every s-sparse $\mathbf{x}$ is the unique s-sparse solution of $A\mathbf{z} = A\mathbf{x}$,
2. $\ker A \cap \{\mathbf{z} \in \mathbb{K}^N : \|\mathbf{z}\|_0 \leq 2s\} = \{\mathbf{0}\}$,
3. For any $S \subset [N]$ with $\mathrm{card}(S) \leq 2s$, the matrix A_S is injective,
4. Every set of $2s$ columns of A is linearly independent.

Minimal Number of Measurements

Given $A \in \mathbb{K}^{m \times N}$, the following are equivalent:

1. Every s-sparse $\mathbf{x}$ is the unique s-sparse solution of $A\mathbf{z} = A\mathbf{x}$,
2. $\ker A \cap \{\mathbf{z} \in \mathbb{K}^N : \|\mathbf{z}\|_0 \leq 2s\} = \{\mathbf{0}\}$,
3. For any $S \subset [N]$ with $\mathrm{card}(S) \leq 2s$, the matrix A_S is injective,
4. Every set of $2s$ columns of A is linearly independent.

As a consequence, exact recovery of every s-sparse vector forces

$$m \geq 2s.$$

Minimal Number of Measurements

Given $A \in \mathbb{K}^{m \times N}$, the following are equivalent:

1. Every s-sparse $\mathbf{x}$ is the unique s-sparse solution of $A\mathbf{z} = A\mathbf{x}$,
2. $\ker A \cap \{\mathbf{z} \in \mathbb{K}^N : \|\mathbf{z}\|_0 \leq 2s\} = \{\mathbf{0}\}$,
3. For any $S \subset [N]$ with $\mathrm{card}(S) \leq 2s$, the matrix A_S is injective,
4. Every set of $2s$ columns of A is linearly independent.

As a consequence, exact recovery of every s-sparse vector forces

$$m \geq 2s.$$

This can be achieved using partial Vandermonde matrices.

Exact s-Sparse Recovery from $2s$ Fourier Measurements*

Exact s-Sparse Recovery from $2s$ Fourier Measurements*

Identify an s-sparse $\mathbf{x} \in \mathbb{C}^N$ with a function x on $\{0, 1, \ldots, N-1\}$ with support S, $\mathrm{card}(S) = s$.

Exact s-Sparse Recovery from $2s$ Fourier Measurements*

Identify an s-sparse $\mathbf{x} \in \mathbb{C}^N$ with a function x on $\{0, 1, \ldots, N-1\}$ with support S, $\mathrm{card}(S) = s$. Consider the $2s$ Fourier coefficients

$$\hat{x}(j) = \sum_{k=0}^{N-1} x(k) e^{-i2\pi jk/N}, \qquad 0 \le j \le 2s-1.$$

Exact s-Sparse Recovery from $2s$ Fourier Measurements*

Identify an s-sparse $\mathbf{x} \in \mathbb{C}^N$ with a function x on $\{0, 1, \ldots, N-1\}$ with support S, $\mathrm{card}(S) = s$. Consider the $2s$ Fourier coefficients

$$\hat{x}(j) = \sum_{k=0}^{N-1} x(k) e^{-i2\pi jk/N}, \qquad 0 \le j \le 2s-1.$$

Consider a trigonometric polynomial vanishing exactly on S, i.e.,

$$p(t) := \prod_{k \in S} \left(1 - e^{-i2\pi k/N} e^{i2\pi t/N}\right).$$

Exact s-Sparse Recovery from $2s$ Fourier Measurements*

Identify an s-sparse $\mathbf{x} \in \mathbb{C}^N$ with a function x on $\{0, 1, \ldots, N-1\}$ with support S, $\mathrm{card}(S) = s$. Consider the $2s$ Fourier coefficients

$$\hat{x}(j) = \sum_{k=0}^{N-1} x(k) e^{-i2\pi jk/N}, \qquad 0 \le j \le 2s-1.$$

Consider a trigonometric polynomial vanishing exactly on S, i.e.,

$$p(t) := \prod_{k \in S} \left(1 - e^{-i2\pi k/N} e^{i2\pi t/N}\right).$$

Since $p \cdot x \equiv 0$, discrete convolution gives

$$0 = (\hat{p} * \hat{x})(j) = \sum_{k=0}^{N-1} \hat{p}(k) \hat{x}(j-k), \qquad 0 \le j \le N-1.$$

Exact s-Sparse Recovery from $2s$ Fourier Measurements*

Identify an s-sparse $\mathbf{x} \in \mathbb{C}^N$ with a function x on $\{0, 1, \ldots, N-1\}$ with support S, $\mathrm{card}(S) = s$. Consider the $2s$ Fourier coefficients

$$\hat{x}(j) = \sum_{k=0}^{N-1} x(k) e^{-i2\pi jk/N}, \qquad 0 \le j \le 2s-1.$$

Consider a trigonometric polynomial vanishing exactly on S, i.e.,

$$p(t) := \prod_{k \in S} \left(1 - e^{-i2\pi k/N} e^{i2\pi t/N}\right).$$

Since $p \cdot x \equiv 0$, discrete convolution gives

$$0 = (\hat{p} * \hat{x})(j) = \sum_{k=0}^{N-1} \hat{p}(k) \hat{x}(j-k), \qquad 0 \le j \le N-1.$$

Note that $\hat{p}(0) = 1$ and that $\hat{p}(k) = 0$ for $k > s$.

Exact s-Sparse Recovery from $2s$ Fourier Measurements*

Identify an s-sparse $\mathbf{x} \in \mathbb{C}^N$ with a function x on $\{0, 1, \ldots, N-1\}$ with support S, $\mathrm{card}(S) = s$. Consider the $2s$ Fourier coefficients

$$\hat{x}(j) = \sum_{k=0}^{N-1} x(k) e^{-i2\pi jk/N}, \qquad 0 \le j \le 2s-1.$$

Consider a trigonometric polynomial vanishing exactly on S, i.e.,

$$p(t) := \prod_{k \in S} \left(1 - e^{-i2\pi k/N} e^{i2\pi t/N}\right).$$

Since $p \cdot x \equiv 0$, discrete convolution gives

$$0 = (\hat{p} * \hat{x})(j) = \sum_{k=0}^{N-1} \hat{p}(k) \hat{x}(j-k), \qquad 0 \le j \le N-1.$$

Note that $\hat{p}(0) = 1$ and that $\hat{p}(k) = 0$ for $k > s$. The equations $s, \ldots, 2s-1$ translate into a Toeplitz system with unknowns $\hat{p}(1), \ldots, \hat{p}(s)$.

Exact s-Sparse Recovery from $2s$ Fourier Measurements*

Identify an s-sparse $\mathbf{x} \in \mathbb{C}^N$ with a function x on $\{0, 1, \ldots, N-1\}$ with support S, $\mathrm{card}(S) = s$. Consider the $2s$ Fourier coefficients

$$\hat{x}(j) = \sum_{k=0}^{N-1} x(k) e^{-i2\pi jk/N}, \qquad 0 \le j \le 2s-1.$$

Consider a trigonometric polynomial vanishing exactly on S, i.e.,

$$p(t) := \prod_{k \in S} \left(1 - e^{-i2\pi k/N} e^{i2\pi t/N}\right).$$

Since $p \cdot x \equiv 0$, discrete convolution gives

$$0 = (\hat{p} * \hat{x})(j) = \sum_{k=0}^{N-1} \hat{p}(k) \hat{x}(j-k), \qquad 0 \le j \le N-1.$$

Note that $\hat{p}(0) = 1$ and that $\hat{p}(k) = 0$ for $k > s$. The equations $s, \ldots, 2s-1$ translate into a Toeplitz system with unknowns $\hat{p}(1), \ldots, \hat{p}(s)$. This determines $\hat{p}$, hence p, then S, and finally $\mathbf{x}$.

Optimization and Greedy Strategies

ℓ_1-Minimization (Basis Pursuit)

ℓ_1-Minimization (Basis Pursuit)

Replace (P_0) by

$$(\mathrm{P}_1) \qquad \underset{\mathbf{z}\in\mathbb{K}^N}{\text{minimize}} \ \|\mathbf{z}\|_1 \quad \text{subject to } A\mathbf{z} = \mathbf{y}.$$

ℓ_1-Minimization (Basis Pursuit)

Replace (P_0) by

$$(\mathrm{P}_1) \qquad \underset{\mathbf{z}\in\mathbb{K}^N}{\text{minimize}} \ \|\mathbf{z}\|_1 \quad \text{subject to } A\mathbf{z} = \mathbf{y}.$$

- Geometric intuition

ℓ_1-Minimization (Basis Pursuit)

Replace (P_0) by

$$(\mathrm{P}_1) \qquad \underset{\mathbf{z}\in\mathbb{K}^N}{\text{minimize}} \ \|\mathbf{z}\|_1 \quad \text{subject to } A\mathbf{z} = \mathbf{y}.$$

- Geometric intuition
- Unique ℓ_1-minimizers are at most m-sparse (when $\mathbb{K} = \mathbb{R}$)

ℓ_1-Minimization (Basis Pursuit)

Replace (P_0) by

$$(\mathrm{P}_1) \qquad \underset{\mathbf{z}\in\mathbb{K}^N}{\text{minimize}} \ \|\mathbf{z}\|_1 \quad \text{subject to } A\mathbf{z} = \mathbf{y}.$$

- Geometric intuition
- Unique ℓ_1-minimizers are at most m-sparse (when $\mathbb{K} = \mathbb{R}$)
- Convex optimization program, hence solvable in practice

ℓ_1-Minimization (Basis Pursuit)

Replace (P_0) by

$$(\mathrm{P}_1) \qquad \underset{\mathbf{z}\in\mathbb{K}^N}{\text{minimize}} \ \|\mathbf{z}\|_1 \quad \text{subject to } A\mathbf{z} = \mathbf{y}.$$

- Geometric intuition
- Unique ℓ_1-minimizers are at most m-sparse (when $\mathbb{K} = \mathbb{R}$)
- Convex optimization program, hence solvable in practice
- In the real setting, recast as the linear optimization program

$$\underset{\mathbf{c},\mathbf{z}\in\mathbb{R}^N}{\text{minimize}} \ \sum_{j=1}^{N} c_j \quad \text{subject to } A\mathbf{z} = \mathbf{y} \text{ and } -c_j \leq z_j \leq c_j.$$

ℓ_1-Minimization (Basis Pursuit)

Replace (P_0) by

$$(\mathrm{P}_1) \qquad \underset{\mathbf{z}\in\mathbb{K}^N}{\text{minimize}} \ \|\mathbf{z}\|_1 \quad \text{subject to } A\mathbf{z} = \mathbf{y}.$$

- Geometric intuition
- Unique ℓ_1-minimizers are at most m-sparse (when $\mathbb{K} = \mathbb{R}$)
- Convex optimization program, hence solvable in practice
- In the real setting, recast as the linear optimization program

$$\underset{\mathbf{c},\mathbf{z}\in\mathbb{R}^N}{\text{minimize}} \ \sum_{j=1}^{N} c_j \quad \text{subject to } A\mathbf{z} = \mathbf{y} \text{ and } -c_j \leq z_j \leq c_j.$$

- In the complex setting, recast as a second order cone program

Basis Pursuit — Null Space Property

Basis Pursuit — Null Space Property

$\Delta_1(A\mathbf{x}) = \mathbf{x}$ for every vector $\mathbf{x}$ supported on S if and only if

Basis Pursuit — Null Space Property

$\Delta_1(A\mathbf{x}) = \mathbf{x}$ for every vector $\mathbf{x}$ supported on S if and only if

(NSP) $\qquad \|\mathbf{u}_S\|_1 < \|\mathbf{u}_{\overline{S}}\|_1, \qquad$ all $\mathbf{u} \in \ker A \setminus \{\mathbf{0}\}$.

Basis Pursuit — Null Space Property

$\Delta_1(A\mathbf{x}) = \mathbf{x}$ for every vector $\mathbf{x}$ supported on S if and only if

(NSP) $\qquad \|\mathbf{u}_S\|_1 < \|\mathbf{u}_{\overline{S}}\|_1, \qquad$ all $\mathbf{u} \in \ker A \setminus \{\mathbf{0}\}$.

For real measurement matrices,

Basis Pursuit — Null Space Property

$\Delta_1(A\mathbf{x}) = \mathbf{x}$ for every vector $\mathbf{x}$ supported on S if and only if

(NSP) $$\|\mathbf{u}_S\|_1 < \|\mathbf{u}_{\overline{S}}\|_1, \qquad \text{all } \mathbf{u} \in \ker A \setminus \{\mathbf{0}\}.$$

For real measurement matrices, real and complex NSPs read

Basis Pursuit — Null Space Property

$\Delta_1(A\mathbf{x}) = \mathbf{x}$ for every vector $\mathbf{x}$ supported on S if and only if

(NSP) $\qquad \|\mathbf{u}_S\|_1 < \|\mathbf{u}_{\overline{S}}\|_1, \qquad$ all $\mathbf{u} \in \ker A \setminus \{\mathbf{0}\}$.

For real measurement matrices, real and complex NSPs read

$$\sum_{j \in S} |u_j| < \sum_{\ell \in \overline{S}} |u_\ell|, \qquad \text{all } \mathbf{u} \in \ker_{\mathbb{R}} A \setminus \{0\},$$

Basis Pursuit — Null Space Property

$\Delta_1(A\mathbf{x}) = \mathbf{x}$ for every vector $\mathbf{x}$ supported on S if and only if

(NSP) $\qquad \|\mathbf{u}_S\|_1 < \|\mathbf{u}_{\overline{S}}\|_1, \qquad$ all $\mathbf{u} \in \ker A \setminus \{\mathbf{0}\}$.

For real measurement matrices, real and complex NSPs read

$$\sum_{j \in S} |u_j| < \sum_{\ell \in \overline{S}} |u_\ell|, \qquad \text{all } \mathbf{u} \in \ker_{\mathbb{R}} A \setminus \{0\},$$

$$\sum_{j \in S} \sqrt{v_j^2 + w_j^2} < \sum_{\ell \in \overline{S}} \sqrt{v_\ell^2 + w_\ell^2}, \qquad \text{all } (\mathbf{v}, \mathbf{w}) \in (\ker_{\mathbb{R}} A)^2 \setminus \{0\}.$$

Basis Pursuit — Null Space Property

$\Delta_1(A\mathbf{x}) = \mathbf{x}$ for every vector $\mathbf{x}$ supported on S if and only if

(NSP) $\qquad \|\mathbf{u}_S\|_1 < \|\mathbf{u}_{\overline{S}}\|_1, \qquad$ all $\mathbf{u} \in \ker A \setminus \{\mathbf{0}\}$.

For real measurement matrices, real and complex NSPs read

$$\sum_{j \in S} |u_j| < \sum_{\ell \in \overline{S}} |u_\ell|, \qquad \text{all } \mathbf{u} \in \ker_{\mathbb{R}} A \setminus \{0\},$$

$$\sum_{j \in S} \sqrt{v_j^2 + w_j^2} < \sum_{\ell \in \overline{S}} \sqrt{v_\ell^2 + w_\ell^2}, \qquad \text{all } (\mathbf{v}, \mathbf{w}) \in (\ker_{\mathbb{R}} A)^2 \setminus \{0\}.$$

Real and complex NSPs are in fact equivalent.

Orthogonal Matching Pursuit

Orthogonal Matching Pursuit

Starting with $S^0 = \emptyset$ and $\mathbf{x}^0 = \mathbf{0}$, iterate

$$(\mathrm{OMP}_1) \qquad S^{n+1} = S^n \cup \left\{ j^{n+1} := \operatorname*{argmax}_{j \in [N]} \left\{ |(A^*(\mathbf{y} - A\mathbf{x}^n))_j| \right\} \right\},$$

$$(\mathrm{OMP}_2) \qquad \mathbf{x}^{n+1} = \operatorname*{argmin}_{\mathbf{z} \in \mathbb{C}^N} \left\{ \|\mathbf{y} - A\mathbf{z}\|_2, \operatorname{supp}(\mathbf{z}) \subseteq S^{n+1} \right\}.$$

Orthogonal Matching Pursuit

Starting with $S^0 = \emptyset$ and $\mathbf{x}^0 = \mathbf{0}$, iterate

$$(\mathrm{OMP}_1) \qquad S^{n+1} = S^n \cup \left\{ j^{n+1} := \operatorname*{argmax}_{j \in [N]} \left\{ |(A^*(\mathbf{y} - A\mathbf{x}^n))_j| \right\} \right\},$$

$$(\mathrm{OMP}_2) \qquad \mathbf{x}^{n+1} = \operatorname*{argmin}_{\mathbf{z} \in \mathbb{C}^N} \left\{ \|\mathbf{y} - A\mathbf{z}\|_2, \operatorname{supp}(\mathbf{z}) \subseteq S^{n+1} \right\}.$$

- The norm of the residual decreases according to

$$\|\mathbf{y} - A\mathbf{x}^{n+1}\|_2^2 \leq \|\mathbf{y} - A\mathbf{x}^n\|_2^2 - \left| (A^*(\mathbf{y} - A\mathbf{x}^n))_{j^{n+1}} \right|^2.$$

Orthogonal Matching Pursuit

Starting with $S^0 = \emptyset$ and $\mathbf{x}^0 = \mathbf{0}$, iterate

$$(\mathrm{OMP}_1) \qquad S^{n+1} = S^n \cup \left\{ j^{n+1} := \operatorname*{argmax}_{j \in [N]} \left\{ |(A^*(\mathbf{y} - A\mathbf{x}^n))_j| \right\} \right\},$$

$$(\mathrm{OMP}_2) \qquad \mathbf{x}^{n+1} = \operatorname*{argmin}_{\mathbf{z} \in \mathbb{C}^N} \left\{ \|\mathbf{y} - A\mathbf{z}\|_2, \operatorname{supp}(\mathbf{z}) \subseteq S^{n+1} \right\}.$$

- The norm of the residual decreases according to

$$\|\mathbf{y} - A\mathbf{x}^{n+1}\|_2^2 \leq \|\mathbf{y} - A\mathbf{x}^n\|_2^2 - \left| (A^*(\mathbf{y} - A\mathbf{x}^n))_{j^{n+1}} \right|^2.$$

- Every vector $\mathbf{x}_{\neq \mathbf{0}}$ supported on S, $\operatorname{card}(S) = s$, is recovered from $\mathbf{y} = A\mathbf{x}$ after at most s iterations of OMP if and only if A_S is injective and

$$(\mathrm{ERC}) \qquad \max_{j \in S} |(A^*\mathbf{r})_j| > \max_{\ell \in \overline{S}} |(A^*\mathbf{r})_\ell|$$

for all $\mathbf{r}_{\neq \mathbf{0}} \in \left\{ A\mathbf{z}, \operatorname{supp}(\mathbf{z}) \subseteq S \right\}$.

First Recovery Guarantees

Coherence

Coherence

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

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

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.

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

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

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.

Summary of Part I

Summary of Part I

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

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

Summary of Part I

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

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

Summary of Part I

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

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

Summary of Part I

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...

Summary of Part I

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...

Next, we will introduce new tools to break this *quadratic barrier* (with random matrices).

Part II:

The Restricted Isometry Property

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

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

Alternative form:

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

Restricted Isometry Constants

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

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

Alternative form:

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

Suitable CS matrices have small restricted isometry constants.

Restricted Isometry Constants

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

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

Alternative form:

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

Suitable CS matrices have small restricted isometry constants.

Typically, $\delta_s \leq \delta_*$ 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

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

Alternative form:

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

Suitable CS matrices have small restricted isometry constants.

Typically, $\delta_s \leq \delta_*$ 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

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

Alternative form:

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

Suitable CS matrices have small restricted isometry constants.

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

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

In fact, $\delta_s \leq \delta_*$ imposes $m \geq \dfrac{c}{\delta_*^2} s$.

RIP-based Recovery Guarantees

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 $\mathbf{v}$.

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

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 $\mathbf{v}$.

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

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 $S_0, S_1, \ldots$ of s indices ordered by decreasing magnitude of entries of $\mathbf{v}$.

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

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 $S_0, S_1, \ldots$ 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

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

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

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 $\|\mathbf{e}\|_2 \leq \eta$:

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

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

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

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

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 $1 \leq p \leq q$,

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

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

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 $1 \leq p \leq q$,

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

  for all $\mathbf{x}, \mathbf{z} \in \mathbb{C}^N$. The converse holds when $q = 1$.
- ℓ_2-RNSP (wrto $\|\cdot\|_2$) holds when $\delta_{2s} < 0.62$.

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:

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

until a stopping criterion is met.

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

(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:

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

$$\text{(HTP}_2\text{)}$$

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:

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

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

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:

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

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

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:

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

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

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

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

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

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.

Recovery by Orthogonal Matching Pursuit

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

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

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?

RIP for Random Matrices

Concentration Inequality

Concentration Inequality

- Let $A \in \mathbb{R}^{m \times N}$ be a random matrix with entries

$$a_{i,j} = \frac{g_{i,j}}{\sqrt{m}} \qquad \text{where the } g_{i,j} \text{ are independent } \mathcal{N}(0,1).$$

Concentration Inequality

- Let $A \in \mathbb{R}^{m \times N}$ be a random matrix with entries

$$a_{i,j} = \frac{g_{i,j}}{\sqrt{m}} \qquad \text{where the } g_{i,j} \text{ are independent } \mathcal{N}(0,1).$$

- For a fixed $\mathbf{x} \in \mathbb{R}^N$, note that $(A\mathbf{x})_i = \sum_{j=1}^N a_{i,j} x_j$, hence

$$\mathbb{E}\big((A\mathbf{x})_i^2\big) = \mathbb{V}\big(\sum a_{i,j} x_j\big) = \sum x_j^2 \mathbb{V}(a_{i,j}) = \frac{\|\mathbf{x}\|_2^2}{m},$$
$$\mathbb{E}\big(\|A\mathbf{x}\|_2^2\big) = \|\mathbf{x}\|_2^2.$$

Concentration Inequality

- Let $A \in \mathbb{R}^{m \times N}$ be a random matrix with entries

$$a_{i,j} = \frac{g_{i,j}}{\sqrt{m}} \qquad \text{where the } g_{i,j} \text{ are independent } \mathcal{N}(0,1).$$

- For a fixed $\mathbf{x} \in \mathbb{R}^N$, note that $(A\mathbf{x})_i = \sum_{j=1}^N a_{i,j} x_j$, hence

$$\mathbb{E}\big((A\mathbf{x})_i^2\big) = \mathbb{V}\big(\sum a_{i,j} x_j\big) = \sum x_j^2 \mathbb{V}(a_{i,j}) = \frac{\|\mathbf{x}\|_2^2}{m},$$
$$\mathbb{E}\big(\|A\mathbf{x}\|_2^2\big) = \|\mathbf{x}\|_2^2.$$

- In fact, $\|A\mathbf{x}\|_2^2$ concentrates around its mean: for $t \in (0,1)$,

$$\text{(CI)} \qquad \mathbb{P}\big(\big|\|A\mathbf{x}\|_2^2 - \|\mathbf{x}\|_2^2\big| > t\|\mathbf{x}\|_2^2\big) \leq 2\exp(-ct^2 m).$$

Covering Arguments

Covering Arguments

Suppose that the random matrix $A \in \mathbb{R}^{m \times N}$ satisfies (CI).

Covering Arguments

Suppose that the random matrix $A \in \mathbb{R}^{m \times N}$ satisfies (CI).
Let $S \subseteq [N]$ with $\mathrm{card}(S) = s$.

Covering Arguments

Suppose that the random matrix $A \in \mathbb{R}^{m \times N}$ satisfies (CI). Let $S \subseteq [N]$ with $\mathrm{card}(S) = s$. Then

$$\mathbb{P}\big(\|A_S^* A_S - \mathrm{Id}\|_{2\to 2} > \delta\big) \leq 2\exp(-c\delta^2 m)$$

Covering Arguments

Suppose that the random matrix $A \in \mathbb{R}^{m \times N}$ satisfies (CI). Let $S \subseteq [N]$ with $\mathrm{card}(S) = s$. Then

$$\mathbb{P}\big(\|A_S^* A_S - \mathrm{Id}\|_{2\to 2} > \delta\big) \leq 2\exp(-c\delta^2 m)$$

provided

$$m \geq \frac{c'}{\delta^2} s.$$

Covering Arguments

Suppose that the random matrix $A \in \mathbb{R}^{m \times N}$ satisfies (CI). Let $S \subseteq [N]$ with $\mathrm{card}(S) = s$. Then

$$\mathbb{P}\big(\|A_S^* A_S - \mathrm{Id}\|_{2\to 2} > \delta\big) \leq 2\exp(-c\delta^2 m)$$

provided

$$m \geq \frac{c'}{\delta^2} s.$$

The argument relies on the following fact:

A subset U of the unit ball of $\mathbb{R}^k$ relative to a norm $\|\cdot\|$ has covering and packing numbers satisfying

$$\mathcal{N}(U, \|\cdot\|, t) \leq \mathcal{P}(U, \|\cdot\|, t) \leq \left(1 + \frac{2}{t}\right)^k.$$

Restricted Isometry Property

Restricted Isometry Property

- Suppose that the random matrix $A \in \mathbb{R}^{m \times N}$ satisfies (CI). Then
$$\mathbb{P}(\delta_s > \delta) \leq 2\exp(-c\delta^2 m)$$
provided
$$m \geq \frac{c'}{\delta^2} s \ln(eN/s).$$

Restricted Isometry Property

- Suppose that the random matrix $A \in \mathbb{R}^{m \times N}$ satisfies (CI). Then
$$\mathbb{P}(\delta_s > \delta) \leq 2\exp(-c\delta^2 m)$$
provided
$$m \geq \frac{c'}{\delta^2} s \ln(eN/s).$$
- The arguments are also valid for subgaussian matrices (e.g. Bernoulli matrices), since these satisfy (CI), too.

Restricted Isometry Property

- Suppose that the random matrix $A \in \mathbb{R}^{m \times N}$ satisfies (CI). Then
$$\mathbb{P}(\delta_s > \delta) \leq 2\exp(-c\delta^2 m)$$
provided
$$m \geq \frac{c'}{\delta^2} s \ln(eN/s).$$
- The arguments are also valid for subgaussian matrices (e.g. Bernoulli matrices), since these satisfy (CI), too.
- For Gaussian matrices, more powerful techniques can provide an explicit value for c'.

Summary of Part II

Summary of Part II

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

$$\delta_{\kappa s} < \delta_*.$$

Summary of Part II

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

$$\delta_{\kappa s} < \delta_*.$$

They guarantee stable and robust reconstructions in the form, say,

$$(1) \quad \|\mathbf{x} - \Delta(A\mathbf{x} + \mathbf{e})\|_2 \leq \frac{C}{\sqrt{s}} \sigma_s(\mathbf{x})_1 + D\|\mathbf{e}\|_2 \qquad \text{for all } \mathbf{x} \text{ and all } \mathbf{e}.$$

Summary of Part II

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

$$\delta_{\kappa s} < \delta_*.$$

They guarantee stable and robust reconstructions in the form, say,

$$(1) \quad \|\mathbf{x} - \Delta(A\mathbf{x} + \mathbf{e})\|_2 \leq \frac{C}{\sqrt{s}} \sigma_s(\mathbf{x})_1 + D\|\mathbf{e}\|_2 \qquad \text{for all } \mathbf{x} \text{ and all } \mathbf{e}.$$

Random matrices fulfill the RI conditions with high probability as soon as

$$(2) \qquad m \geq c\, s \ln(N/s).$$

Summary of Part II

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

$$\delta_{\kappa s} < \delta_*.$$

They guarantee stable and robust reconstructions in the form, say,

$$(1) \quad \|\mathbf{x} - \Delta(A\mathbf{x} + \mathbf{e})\|_2 \leq \frac{C}{\sqrt{s}} \sigma_s(\mathbf{x})_1 + D\|\mathbf{e}\|_2 \qquad \text{for all } \mathbf{x} \text{ and all } \mathbf{e}.$$

Random matrices fulfill the RI conditions with high probability as soon as

$$(2) \qquad m \geq c\, s \ln(N/s).$$

Next, we will see that this number of measurement is optimal, in the sense that estimates of type (1) require (2) to hold.

Summary of Part II

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

$$\delta_{\kappa s} < \delta_*.$$

They guarantee stable and robust reconstructions in the form, say,

$$(1) \quad \|\mathbf{x} - \Delta(A\mathbf{x} + \mathbf{e})\|_2 \leq \frac{C}{\sqrt{s}} \sigma_s(\mathbf{x})_1 + D\|\mathbf{e}\|_2 \qquad \text{for all } \mathbf{x} \text{ and all } \mathbf{e}.$$

Random matrices fulfill the RI conditions with high probability as soon as

$$(2) \qquad m \geq c\, s \ln(N/s).$$

Next, we will see that this number of measurement is optimal, in the sense that estimates of type (1) require (2) to hold. We will also examine the gain in replacing *for all* $\mathbf{x}$ in (1) by *for a fixed* $\mathbf{x}$.

Part III:

Optimality, Uniform vs Nonuniform Recovery

Optimality of Uniform Guarantees

Gelfand Widths

Gelfand Widths

For a subset K of a normed space X, define

$$E^m(K,X) := \inf\left\{\sup_{\mathbf{x}\in K}\|\mathbf{x}-\Delta(A\mathbf{x})\|,\ A : X \overset{\text{linear}}{\to} \mathbb{R}^m, \Delta : \mathbb{R}^m \to X\right\}$$

Gelfand Widths

For a subset K of a normed space X, define

$$E^m(K,X) := \inf\left\{\sup_{\mathbf{x}\in K}\|\mathbf{x}-\Delta(A\mathbf{x})\|,\ A : X \overset{\text{linear}}{\to} \mathbb{R}^m, \Delta : \mathbb{R}^m \to X\right\}$$

The Gelfand m-width of K in X is

$$d^m(K,X) := \inf\left\{\sup_{\mathbf{x}\in K\cap L^m}\|\mathbf{x}\|,\ \ L^m \text{ subspace of } X,\ \mathrm{codim}(L^m)\le m\right\}$$

Gelfand Widths

For a subset K of a normed space X, define

$$E^m(K,X) := \inf\left\{\sup_{\mathbf{x}\in K}\|\mathbf{x}-\Delta(A\mathbf{x})\|,\ A: X \overset{\text{linear}}{\to} \mathbb{R}^m, \Delta: \mathbb{R}^m \to X\right\}$$

The Gelfand m-width of K in X is

$$d^m(K,X) := \inf\left\{\sup_{\mathbf{x}\in K\cap L^m}\|\mathbf{x}\|,\ \ L^m \text{ subspace of } X,\ \mathrm{codim}(L^m)\le m\right\}$$

If $-K = K$, then

$$d^m(K,X) \le E^m(K,X),$$

Gelfand Widths

For a subset K of a normed space X, define

$$E^m(K,X) := \inf\left\{\sup_{\mathbf{x}\in K}\|\mathbf{x}-\Delta(A\mathbf{x})\|,\ A : X \overset{\text{linear}}{\to} \mathbb{R}^m, \Delta : \mathbb{R}^m \to X\right\}$$

The Gelfand m-width of K in X is

$$d^m(K,X) := \inf\left\{\sup_{\mathbf{x}\in K\cap L^m}\|\mathbf{x}\|,\ \ L^m \text{ subspace of } X,\ \mathrm{codim}(L^m)\le m\right\}$$

If $-K = K$, then

$$d^m(K,X) \le E^m(K,X),$$

and if in addition $K + K \subseteq a\,K$, then

$$E^m(K,X) \le a\,d^m(K,X).$$

Gelfand Widths of ℓ_1-Balls: Upper Bound

Gelfand Widths of ℓ_1-Balls: Upper Bound

Let $A \in \mathbb{R}^{m \times N}$ with $m \approx c\, s \ln(eN/m)$ be such that $\delta_{2s} < 1/2$.

Gelfand Widths of ℓ_1-Balls: Upper Bound

Let $A \in \mathbb{R}^{m \times N}$ with $m \approx c\, s \ln(eN/m)$ be such that $\delta_{2s} < 1/2$.
Let $\Delta_1 : \mathbb{R}^m \to \mathbb{R}^N$ be the ℓ_1-minimization map.

Gelfand Widths of ℓ_1-Balls: Upper Bound

Let $A \in \mathbb{R}^{m \times N}$ with $m \approx c\, s \ln(eN/m)$ be such that $\delta_{2s} < 1/2$. Let $\Delta_1 : \mathbb{R}^m \to \mathbb{R}^N$ be the ℓ_1-minimization map. Given $1 < p \leq 2$, for any $\mathbf{x} \in B_1^N$,

$$\|\mathbf{x} - \Delta_1(A\mathbf{x})\|_p \leq \frac{C}{s^{1-1/p}} \sigma_s(\mathbf{x})_1 \leq \frac{C}{s^{1-1/p}} \approx \frac{C'}{(m/\ln(eN/m))^{1-1/p}}.$$

Gelfand Widths of ℓ_1-Balls: Upper Bound

Let $A \in \mathbb{R}^{m \times N}$ with $m \approx c\, s \ln(eN/m)$ be such that $\delta_{2s} < 1/2$. Let $\Delta_1 : \mathbb{R}^m \to \mathbb{R}^N$ be the ℓ_1-minimization map. Given $1 < p \leq 2$, for any $\mathbf{x} \in B_1^N$,

$$\|\mathbf{x} - \Delta_1(A\mathbf{x})\|_p \leq \frac{C}{s^{1-1/p}} \sigma_s(\mathbf{x})_1 \leq \frac{C}{s^{1-1/p}} \approx \frac{C'}{(m/\ln(eN/m))^{1-1/p}}.$$

This gives an upper bound for $E^m(B_1^N, \ell_p^N)$,

Gelfand Widths of ℓ_1-Balls: Upper Bound

Let $A \in \mathbb{R}^{m \times N}$ with $m \approx c\, s \ln(eN/m)$ be such that $\delta_{2s} < 1/2$. Let $\Delta_1 : \mathbb{R}^m \to \mathbb{R}^N$ be the ℓ_1-minimization map. Given $1 < p \leq 2$, for any $\mathbf{x} \in B_1^N$,

$$\|\mathbf{x} - \Delta_1(A\mathbf{x})\|_p \leq \frac{C}{s^{1-1/p}} \sigma_s(\mathbf{x})_1 \leq \frac{C}{s^{1-1/p}} \approx \frac{C'}{(m/\ln(eN/m))^{1-1/p}}.$$

This gives an upper bound for $E^m(B_1^N, \ell_p^N)$, and in turn

$$d^m(B_1^N, \ell_p^N) \leq C \min\left\{1, \frac{\ln(eN/m)}{m}\right\}^{1-1/p}.$$

Gelfand Widths of ℓ_1-Balls: Lower Bound

Gelfand Widths of ℓ_1-Balls: Lower Bound

The Gelfand width of B_1^N in ℓ_p^N, $p > 1$, also satisfies

$$d^m(B_1^N, \ell_p^N) \geq c \min\left\{1, \frac{\ln(eN/m)}{m}\right\}^{1-1/p}.$$

Gelfand Widths of ℓ_1-Balls: Lower Bound

The Gelfand width of B_1^N in ℓ_p^N, $p > 1$, also satisfies

$$d^m(B_1^N, \ell_p^N) \geq c \min \left\{ 1, \frac{\ln(eN/m)}{m} \right\}^{1-1/p}.$$

Once established, this will show the optimality of the CS results.

Gelfand Widths of ℓ_1-Balls: Lower Bound

The Gelfand width of B_1^N in ℓ_p^N, $p > 1$, also satisfies

$$d^m(B_1^N, \ell_p^N) \geq c \min \left\{ 1, \frac{\ln(eN/m)}{m} \right\}^{1-1/p}.$$

Once established, this will show the optimality of the CS results. Indeed, suppose the existence of (A, Δ) such that

$$\|\mathbf{x} - \Delta(A\mathbf{x})\|_p \leq \frac{C}{s^{1-1/p}} \sigma_s(\mathbf{x})_1 \qquad \text{for all } \mathbf{x} \in \mathbb{R}^N.$$

Gelfand Widths of ℓ_1-Balls: Lower Bound

The Gelfand width of B_1^N in ℓ_p^N, $p > 1$, also satisfies

$$d^m(B_1^N, \ell_p^N) \geq c \min \left\{ 1, \frac{\ln(eN/m)}{m} \right\}^{1-1/p}.$$

Once established, this will show the optimality of the CS results. Indeed, suppose the existence of (A, Δ) such that

$$\|\mathbf{x} - \Delta(A\mathbf{x})\|_p \leq \frac{C}{s^{1-1/p}} \, \sigma_s(\mathbf{x})_1 \leq \frac{C}{s^{1-1/p}} \, \|\mathbf{x}\|_1 \qquad \text{for all } \mathbf{x} \in \mathbb{R}^N.$$

Gelfand Widths of ℓ_1-Balls: Lower Bound

The Gelfand width of B_1^N in ℓ_p^N, $p > 1$, also satisfies

$$d^m(B_1^N, \ell_p^N) \geq c \min\left\{1, \frac{\ln(eN/m)}{m}\right\}^{1-1/p}.$$

Once established, this will show the optimality of the CS results. Indeed, suppose the existence of (A, Δ) such that

$$\|\mathbf{x} - \Delta(A\mathbf{x})\|_p \leq \frac{C}{s^{1-1/p}}\, \sigma_s(\mathbf{x})_1 \leq \frac{C}{s^{1-1/p}}\, \|\mathbf{x}\|_1 \qquad \text{for all } \mathbf{x} \in \mathbb{R}^N.$$

Then,

$$c \min\left\{1, \frac{\ln(eN/m)}{m}\right\}^{1-1/p} \leq d^m(B_1^N, \ell_p^N) \leq E^m(B_1^N, \ell_p^N) \leq \frac{C}{s^{1-1/p}},$$

Gelfand Widths of ℓ_1-Balls: Lower Bound

The Gelfand width of B_1^N in ℓ_p^N, $p > 1$, also satisfies

$$d^m(B_1^N, \ell_p^N) \geq c \min\left\{1, \frac{\ln(eN/m)}{m}\right\}^{1-1/p}.$$

Once established, this will show the optimality of the CS results. Indeed, suppose the existence of (A, Δ) such that

$$\|\mathbf{x} - \Delta(A\mathbf{x})\|_p \leq \frac{C}{s^{1-1/p}}\, \sigma_s(\mathbf{x})_1 \leq \frac{C}{s^{1-1/p}}\, \|\mathbf{x}\|_1 \qquad \text{for all } \mathbf{x} \in \mathbb{R}^N.$$

Then,

$$c \min\left\{1, \frac{\ln(eN/m)}{m}\right\}^{1-1/p} \leq d^m(B_1^N, \ell_p^N) \leq E^m(B_1^N, \ell_p^N) \leq \frac{C}{s^{1-1/p}},$$

in other words

$$c' \min\left\{1, \frac{\ln(eN/m)}{m}\right\} \leq \frac{1}{s}.$$

Gelfand Widths of ℓ_1-Balls: Lower Bound

The Gelfand width of B_1^N in ℓ_p^N, $p > 1$, also satisfies

$$d^m(B_1^N, \ell_p^N) \geq c \min\left\{1, \frac{\ln(eN/m)}{m}\right\}^{1-1/p}.$$

Once established, this will show the optimality of the CS results. Indeed, suppose the existence of (A, Δ) such that

$$\|\mathbf{x} - \Delta(A\mathbf{x})\|_p \leq \frac{C}{s^{1-1/p}}\, \sigma_s(\mathbf{x})_1 \leq \frac{C}{s^{1-1/p}}\, \|\mathbf{x}\|_1 \qquad \text{for all } \mathbf{x} \in \mathbb{R}^N.$$

Then,

$$c \min\left\{1, \frac{\ln(eN/m)}{m}\right\}^{1-1/p} \leq d^m(B_1^N, \ell_p^N) \leq E^m(B_1^N, \ell_p^N) \leq \frac{C}{s^{1-1/p}},$$

in other words

$$c' \min\left\{1, \frac{\ln(eN/m)}{m}\right\} \leq \frac{1}{s}.$$

We derive either $s \leq 1/c'$ or $m \geq c'\, s\, \ln(eN/m)$,

Gelfand Widths of ℓ_1-Balls: Lower Bound

The Gelfand width of B_1^N in ℓ_p^N, $p > 1$, also satisfies

$$d^m(B_1^N, \ell_p^N) \geq c \min\left\{1, \frac{\ln(eN/m)}{m}\right\}^{1-1/p}.$$

Once established, this will show the optimality of the CS results. Indeed, suppose the existence of (A, Δ) such that

$$\|\mathbf{x} - \Delta(A\mathbf{x})\|_p \leq \frac{C}{s^{1-1/p}}\, \sigma_s(\mathbf{x})_1 \leq \frac{C}{s^{1-1/p}}\, \|\mathbf{x}\|_1 \qquad \text{for all } \mathbf{x} \in \mathbb{R}^N.$$

Then,

$$c \min\left\{1, \frac{\ln(eN/m)}{m}\right\}^{1-1/p} \leq d^m(B_1^N, \ell_p^N) \leq E^m(B_1^N, \ell_p^N) \leq \frac{C}{s^{1-1/p}},$$

in other words

$$c' \min\left\{1, \frac{\ln(eN/m)}{m}\right\} \leq \frac{1}{s}.$$

We derive either $s \leq 1/c'$ or $m \geq c'\, s\, \ln(eN/m)$, in which case

Gelfand Widths of ℓ_1-Balls: Lower Bound

The Gelfand width of B_1^N in ℓ_p^N, $p > 1$, also satisfies

$$d^m(B_1^N, \ell_p^N) \geq c \min\left\{1, \frac{\ln(eN/m)}{m}\right\}^{1-1/p}.$$

Once established, this will show the optimality of the CS results. Indeed, suppose the existence of (A, Δ) such that

$$\|\mathbf{x} - \Delta(A\mathbf{x})\|_p \leq \frac{C}{s^{1-1/p}}\, \sigma_s(\mathbf{x})_1 \leq \frac{C}{s^{1-1/p}}\, \|\mathbf{x}\|_1 \qquad \text{for all } \mathbf{x} \in \mathbb{R}^N.$$

Then,

$$c \min\left\{1, \frac{\ln(eN/m)}{m}\right\}^{1-1/p} \leq d^m(B_1^N, \ell_p^N) \leq E^m(B_1^N, \ell_p^N) \leq \frac{C}{s^{1-1/p}},$$

in other words

$$c' \min\left\{1, \frac{\ln(eN/m)}{m}\right\} \leq \frac{1}{s}.$$

We derive either $s \leq 1/c'$ or $m \geq c'\, s\, \ln(eN/m)$, in which case

$$m \geq c'\, s\, \ln\left(\frac{eN}{s}\right) + c'\, s\, \ln\left(\frac{s}{m}\right)$$

Gelfand Widths of ℓ_1-Balls: Lower Bound

The Gelfand width of B_1^N in ℓ_p^N, $p > 1$, also satisfies

$$d^m(B_1^N, \ell_p^N) \geq c \min\left\{1, \frac{\ln(eN/m)}{m}\right\}^{1-1/p}.$$

Once established, this will show the optimality of the CS results. Indeed, suppose the existence of (A, Δ) such that

$$\|\mathbf{x} - \Delta(A\mathbf{x})\|_p \leq \frac{C}{s^{1-1/p}}\, \sigma_s(\mathbf{x})_1 \leq \frac{C}{s^{1-1/p}}\, \|\mathbf{x}\|_1 \qquad \text{for all } \mathbf{x} \in \mathbb{R}^N.$$

Then,

$$c \min\left\{1, \frac{\ln(eN/m)}{m}\right\}^{1-1/p} \leq d^m(B_1^N, \ell_p^N) \leq E^m(B_1^N, \ell_p^N) \leq \frac{C}{s^{1-1/p}},$$

in other words

$$c' \min\left\{1, \frac{\ln(eN/m)}{m}\right\} \leq \frac{1}{s}.$$

We derive either $s \leq 1/c'$ or $m \geq c'\, s\, \ln(eN/m)$, in which case

$$m \geq c'\, s\, \ln\left(\frac{eN}{s}\right) + c'\, m \frac{s}{m}\, \ln\left(\frac{s}{m}\right)$$

Gelfand Widths of ℓ_1-Balls: Lower Bound

The Gelfand width of B_1^N in ℓ_p^N, $p > 1$, also satisfies

$$d^m(B_1^N, \ell_p^N) \geq c \min\left\{1, \frac{\ln(eN/m)}{m}\right\}^{1-1/p}.$$

Once established, this will show the optimality of the CS results. Indeed, suppose the existence of (A, Δ) such that

$$\|\mathbf{x} - \Delta(A\mathbf{x})\|_p \leq \frac{C}{s^{1-1/p}}\, \sigma_s(\mathbf{x})_1 \leq \frac{C}{s^{1-1/p}}\, \|\mathbf{x}\|_1 \qquad \text{for all } \mathbf{x} \in \mathbb{R}^N.$$

Then,

$$c \min\left\{1, \frac{\ln(eN/m)}{m}\right\}^{1-1/p} \leq d^m(B_1^N, \ell_p^N) \leq E^m(B_1^N, \ell_p^N) \leq \frac{C}{s^{1-1/p}},$$

in other words

$$c' \min\left\{1, \frac{\ln(eN/m)}{m}\right\} \leq \frac{1}{s}.$$

We derive either $s \leq 1/c'$ or $m \geq c'\, s \ln(eN/m)$, in which case

$$m \geq c'\, s \ln\left(\frac{eN}{s}\right) + \underbrace{c'\, m \frac{s}{m} \ln\left(\frac{s}{m}\right)}_{\geq -1/e}$$

Gelfand Widths of ℓ_1-Balls: Lower Bound

The Gelfand width of B_1^N in ℓ_p^N, $p > 1$, also satisfies

$$d^m(B_1^N, \ell_p^N) \geq c \min\left\{1, \frac{\ln(eN/m)}{m}\right\}^{1-1/p}.$$

Once established, this will show the optimality of the CS results. Indeed, suppose the existence of (A, Δ) such that

$$\|\mathbf{x} - \Delta(A\mathbf{x})\|_p \leq \frac{C}{s^{1-1/p}}\, \sigma_s(\mathbf{x})_1 \leq \frac{C}{s^{1-1/p}}\, \|\mathbf{x}\|_1 \qquad \text{for all } \mathbf{x} \in \mathbb{R}^N.$$

Then,

$$c \min\left\{1, \frac{\ln(eN/m)}{m}\right\}^{1-1/p} \leq d^m(B_1^N, \ell_p^N) \leq E^m(B_1^N, \ell_p^N) \leq \frac{C}{s^{1-1/p}},$$

in other words

$$c' \min\left\{1, \frac{\ln(eN/m)}{m}\right\} \leq \frac{1}{s}.$$

We derive either $s \leq 1/c'$ or $m \geq c'\, s \ln(eN/m)$, in which case

$$m \geq c'\, s \ln\left(\frac{eN}{s}\right) + \; c'\, m \underbrace{\frac{s}{m} \ln\left(\frac{s}{m}\right)}_{\geq -1/e} \geq \cdots$$

Gelfand Widths of ℓ_1-Balls: Lower Bound

The Gelfand width of B_1^N in ℓ_p^N, $p > 1$, also satisfies

$$d^m(B_1^N, \ell_p^N) \geq c \min\left\{1, \frac{\ln(eN/m)}{m}\right\}^{1-1/p}.$$

Once established, this will show the optimality of the CS results. Indeed, suppose the existence of (A, Δ) such that

$$\|\mathbf{x} - \Delta(A\mathbf{x})\|_p \leq \frac{C}{s^{1-1/p}}\, \sigma_s(\mathbf{x})_1 \leq \frac{C}{s^{1-1/p}}\, \|\mathbf{x}\|_1 \qquad \text{for all } \mathbf{x} \in \mathbb{R}^N.$$

Then,

$$c \min\left\{1, \frac{\ln(eN/m)}{m}\right\}^{1-1/p} \leq d^m(B_1^N, \ell_p^N) \leq E^m(B_1^N, \ell_p^N) \leq \frac{C}{s^{1-1/p}},$$

in other words

$$c' \min\left\{1, \frac{\ln(eN/m)}{m}\right\} \leq \frac{1}{s}.$$

We derive either $s \leq 1/c'$ or $m \geq c'\, s \ln(eN/m)$, in which case

$$m \geq c'\, s \ln\left(\frac{eN}{s}\right) + \; c'\, m \underbrace{\frac{s}{m} \ln\left(\frac{s}{m}\right)}_{\geq -1/e} \geq \cdots \geq c''\, s \ln\left(\frac{eN}{s}\right).$$

The lower estimate for $d^m(B_1^N, \ell_p^N)$: two key insights

The lower estimate for $d^m(B_1^N, \ell_p^N)$: two key insights

- small width implies ℓ_1-recovery of s-sparse vectors for large s.

The lower estimate for $d^m(B_1^N, \ell_p^N)$: two key insights

- small width implies ℓ_1-recovery of s-sparse vectors for large s.

- ℓ_1-recovery of s-sparse vectors only possible for moderate s.

The lower estimate for $d^m(B_1^N, \ell_p^N)$: two key insights

- small width implies ℓ_1-recovery of s-sparse vectors for large s.

  There is a matrix $A \in \mathbb{R}^{m \times N}$ such that every s-sparse $\mathbf{x} \in \mathbb{R}^N$ is a minimizer of $\|\mathbf{z}\|_1$ subject to $A\mathbf{z} = A\mathbf{x}$ for

$$s \approx \left(\frac{1}{2d^m(B_1^N, \ell_p^N)} \right)^{\frac{p}{p-1}}.$$

- ℓ_1-recovery of s-sparse vectors only possible for moderate s.

The lower estimate for $d^m(B_1^N, \ell_p^N)$: two key insights

- small width implies ℓ_1-recovery of s-sparse vectors for large s.

  There is a matrix $A \in \mathbb{R}^{m \times N}$ such that every s-sparse $\mathbf{x} \in \mathbb{R}^N$ is a minimizer of $\|\mathbf{z}\|_1$ subject to $A\mathbf{z} = A\mathbf{x}$ for

$$s \approx \left(\frac{1}{2d^m(B_1^N, \ell_p^N)} \right)^{\frac{p}{p-1}}.$$

- ℓ_1-recovery of s-sparse vectors only possible for moderate s.

  For $s \geq 2$, if $A \in \mathbb{R}^{m \times N}$ is a matrix such that every s-sparse vector $\mathbf{x}$ is a minimizer of $\|\mathbf{z}\|_1$ subject to $A\mathbf{z} = A\mathbf{x}$, then

$$m \geq c_1 \, s \, \ln\left(\frac{N}{c_2 s}\right),$$

The lower estimate for $d^m(B_1^N, \ell_p^N)$: two key insights

- small width implies ℓ_1-recovery of s-sparse vectors for large s.

  There is a matrix $A \in \mathbb{R}^{m \times N}$ such that every s-sparse $\mathbf{x} \in \mathbb{R}^N$ is a minimizer of $\|\mathbf{z}\|_1$ subject to $A\mathbf{z} = A\mathbf{x}$ for

$$s \approx \left(\frac{1}{2d^m(B_1^N, \ell_p^N)} \right)^{\frac{p}{p-1}}.$$

- ℓ_1-recovery of s-sparse vectors only possible for moderate s.

  For $s \geq 2$, if $A \in \mathbb{R}^{m \times N}$ is a matrix such that every s-sparse vector $\mathbf{x}$ is a minimizer of $\|\mathbf{z}\|_1$ subject to $A\mathbf{z} = A\mathbf{x}$, then

$$m \geq c_1 \, s \, \ln\left(\frac{N}{c_2 s}\right), \qquad c_1 \geq 0.45, \; c_2 = 4.$$

Deriving the lower estimate*

Deriving the lower estimate*

Suppose that $d^m(B_1^N, \ell_p^N) < (c\mu)^{1-1/p}/2$, where

$$\mu := \min\left\{1, \frac{\ln(eN/m)}{m}\right\} \leq 1.$$

Deriving the lower estimate*

Suppose that $d^m(B_1^N, \ell_p^N) < (c\mu)^{1-1/p}/2$, where

$$\mu := \min\left\{1, \frac{\ln(eN/m)}{m}\right\} \leq 1.$$

Setting $s \approx 1/(c\mu) \geq 2$,

Deriving the lower estimate*

Suppose that $d^m(B_1^N, \ell_p^N) < (c\mu)^{1-1/p}/2$, where

$$\mu := \min\left\{1, \frac{\ln(eN/m)}{m}\right\} \leq 1.$$

Setting $s \approx 1/(c\mu) \geq 2$, there exists $A \in \mathbb{R}^{m\times N}$ allowing ℓ_1-recovery of all s-sparse vectors.

Deriving the lower estimate*

Suppose that $d^m(B_1^N, \ell_p^N) < (c\mu)^{1-1/p}/2$, where

$$\mu := \min\left\{1, \frac{\ln(eN/m)}{m}\right\} \leq 1.$$

Setting $s \approx 1/(c\mu) \geq 2$, there exists $A \in \mathbb{R}^{m\times N}$ allowing ℓ_1-recovery of all s-sparse vectors. Therefore

$$m \geq c_1\, s\, \ln\left(\frac{N}{c_2 s}\right)$$

Deriving the lower estimate*

Suppose that $d^m(B_1^N, \ell_p^N) < (c\mu)^{1-1/p}/2$, where

$$\mu := \min\left\{1, \frac{\ln(eN/m)}{m}\right\} \leq 1.$$

Setting $s \approx 1/(c\mu) \geq 2$, there exists $A \in \mathbb{R}^{m\times N}$ allowing ℓ_1-recovery of all s-sparse vectors. Therefore

$$m \geq c_1\, s\, \ln\left(\frac{N}{c_2 s}\right) \geq c_1\, s\, \ln\left(\frac{N}{c_2' m}\right)$$

Deriving the lower estimate*

Suppose that $d^m(B_1^N, \ell_p^N) < (c\mu)^{1-1/p}/2$, where

$$\mu := \min\left\{1, \frac{\ln(eN/m)}{m}\right\} \leq 1.$$

Setting $s \approx 1/(c\mu) \geq 2$, there exists $A \in \mathbb{R}^{m\times N}$ allowing ℓ_1-recovery of all s-sparse vectors. Therefore

$$m \geq c_1\, s\, \ln\left(\frac{N}{c_2 s}\right) \geq c_1\, s\, \ln\left(\frac{N}{c_2' m}\right) \geq c_1'\, s\, \ln\left(\frac{eN}{m}\right)$$

Deriving the lower estimate*

Suppose that $d^m(B_1^N, \ell_p^N) < (c\mu)^{1-1/p}/2$, where

$$\mu := \min\left\{1, \frac{\ln(eN/m)}{m}\right\} \leq 1.$$

Setting $s \approx 1/(c\mu) \geq 2$, there exists $A \in \mathbb{R}^{m\times N}$ allowing ℓ_1-recovery of all s-sparse vectors. Therefore

$$m \geq c_1\, s\, \ln\left(\frac{N}{c_2 s}\right) \geq c_1\, s\, \ln\left(\frac{N}{c_2' m}\right) \geq c_1'\, s\, \ln\left(\frac{eN}{m}\right)$$
$$\geq \frac{c_1'}{c}\frac{\ln(eN/m)}{\min\left\{1, \frac{\ln(eN/m)}{m}\right\}}$$

Deriving the lower estimate*

Suppose that $d^m(B_1^N, \ell_p^N) < (c\mu)^{1-1/p}/2$, where

$$\mu := \min\left\{1, \frac{\ln(eN/m)}{m}\right\} \leq 1.$$

Setting $s \approx 1/(c\mu) \geq 2$, there exists $A \in \mathbb{R}^{m\times N}$ allowing ℓ_1-recovery of all s-sparse vectors. Therefore

$$m \geq c_1\, s\, \ln\left(\frac{N}{c_2 s}\right) \geq c_1\, s\, \ln\left(\frac{N}{c_2' m}\right) \geq c_1'\, s\, \ln\left(\frac{eN}{m}\right)$$
$$\geq \frac{c_1'}{c}\frac{\ln(eN/m)}{\min\left\{1, \frac{\ln(eN/m)}{m}\right\}} \geq \frac{c_1'}{c} m$$

Deriving the lower estimate*

Suppose that $d^m(B_1^N, \ell_p^N) < (c\mu)^{1-1/p}/2$, where

$$\mu := \min\left\{1, \frac{\ln(eN/m)}{m}\right\} \leq 1.$$

Setting $s \approx 1/(c\mu) \geq 2$, there exists $A \in \mathbb{R}^{m\times N}$ allowing ℓ_1-recovery of all s-sparse vectors. Therefore

$$\begin{aligned} m &\geq c_1\, s\, \ln\left(\frac{N}{c_2 s}\right) \geq c_1\, s\, \ln\left(\frac{N}{c_2' m}\right) \geq c_1'\, s\, \ln\left(\frac{eN}{m}\right) \\ &\geq \frac{c_1'}{c}\frac{\ln(eN/m)}{\min\left\{1, \frac{\ln(eN/m)}{m}\right\}} \geq \frac{c_1'}{c} m > m, \end{aligned}$$

provided c is chosen small enough.

Deriving the lower estimate*

Suppose that $d^m(B_1^N, \ell_p^N) < (c\mu)^{1-1/p}/2$, where

$$\mu := \min\left\{1, \frac{\ln(eN/m)}{m}\right\} \leq 1.$$

Setting $s \approx 1/(c\mu) \geq 2$, there exists $A \in \mathbb{R}^{m\times N}$ allowing ℓ_1-recovery of all s-sparse vectors. Therefore

$$\begin{aligned} m &\geq c_1\, s\, \ln\left(\frac{N}{c_2 s}\right) \geq c_1\, s\, \ln\left(\frac{N}{c_2' m}\right) \geq c_1'\, s\, \ln\left(\frac{eN}{m}\right) \\ &\geq \frac{c_1'}{c}\frac{\ln(eN/m)}{\min\left\{1, \frac{\ln(eN/m)}{m}\right\}} \geq \frac{c_1'}{c} m > m, \end{aligned}$$

provided c is chosen small enough. This is a contradiction.

Insight 1: width and ℓ_1-recovery*

Insight 1: width and ℓ_1-recovery*

Every s-sparse $\mathbf{x} \in \mathbb{R}^N$ is a minimizer of $\|\mathbf{z}\|_1$ subject to $A\mathbf{z} = A\mathbf{x}$ if and only if the null space property of order s holds,

Insight 1: width and ℓ_1-recovery*

Every s-sparse $\mathbf{x} \in \mathbb{R}^N$ is a minimizer of $\|\mathbf{z}\|_1$ subject to $A\mathbf{z} = A\mathbf{x}$ if and only if the null space property of order s holds, i.e.,

$$\|\mathbf{v}_S\|_1 \leq \|\mathbf{v}_{\overline{S}}\|_1 \qquad \text{for all } \mathbf{v} \in \ker A \text{ and all } S \in [N] \text{ with } |S| \leq s.$$

Insight 1: width and ℓ_1-recovery*

Every s-sparse $\mathbf{x} \in \mathbb{R}^N$ is a minimizer of $\|\mathbf{z}\|_1$ subject to $A\mathbf{z} = A\mathbf{x}$ if and only if the null space property of order s holds, i.e.,

$$\|\mathbf{v}_S\|_1 \leq \|\mathbf{v}\|_1/2 \quad \text{for all } \mathbf{v} \in \ker A \text{ and all } S \in [N] \text{ with } |S| \leq s.$$

Insight 1: width and ℓ_1-recovery*

Every s-sparse $\mathbf{x} \in \mathbb{R}^N$ is a minimizer of $\|\mathbf{z}\|_1$ subject to $A\mathbf{z} = A\mathbf{x}$ if and only if the null space property of order s holds, i.e.,

$$\|\mathbf{v}_S\|_1 \leq \|\mathbf{v}\|_1/2 \quad \text{for all } \mathbf{v} \in \ker A \text{ and all } S \in [N] \text{ with } |S| \leq s.$$

Setting $d := d^m(B_1^N, \ell_p^N)$, there exists $A \in \mathbb{R}^{m \times N}$ such that

$$\|\mathbf{v}\|_p \leq d\|\mathbf{v}\|_1 \quad \text{for all } \mathbf{v} \in \ker A.$$

Insight 1: width and ℓ_1-recovery*

Every s-sparse $\mathbf{x} \in \mathbb{R}^N$ is a minimizer of $\|\mathbf{z}\|_1$ subject to $A\mathbf{z} = A\mathbf{x}$ if and only if the null space property of order s holds, i.e.,

$$\|\mathbf{v}_S\|_1 \leq \|\mathbf{v}\|_1/2 \quad \text{for all } \mathbf{v} \in \ker A \text{ and all } S \in [N] \text{ with } |S| \leq s.$$

Setting $d := d^m(B_1^N, \ell_p^N)$, there exists $A \in \mathbb{R}^{m \times N}$ such that

$$\|\mathbf{v}\|_p \leq d\|\mathbf{v}\|_1 \quad \text{for all } \mathbf{v} \in \ker A.$$

Then, for $\mathbf{v} \in \ker A$ and $S \in [N]$ with $|S| \leq s$,

Insight 1: width and ℓ_1-recovery*

Every s-sparse $\mathbf{x} \in \mathbb{R}^N$ is a minimizer of $\|\mathbf{z}\|_1$ subject to $A\mathbf{z} = A\mathbf{x}$ if and only if the null space property of order s holds, i.e.,

$$\|\mathbf{v}_S\|_1 \leq \|\mathbf{v}\|_1/2 \quad \text{for all } \mathbf{v} \in \ker A \text{ and all } S \in [N] \text{ with } |S| \leq s.$$

Setting $d := d^m(B_1^N, \ell_p^N)$, there exists $A \in \mathbb{R}^{m \times N}$ such that

$$\|\mathbf{v}\|_p \leq d\|\mathbf{v}\|_1 \quad \text{for all } \mathbf{v} \in \ker A.$$

Then, for $\mathbf{v} \in \ker A$ and $S \in [N]$ with $|S| \leq s$,

$$\|\mathbf{v}_S\|_1 \leq s^{1-1/p}\|\mathbf{v}_S\|_p$$

Insight 1: width and ℓ_1-recovery*

Every s-sparse $\mathbf{x} \in \mathbb{R}^N$ is a minimizer of $\|\mathbf{z}\|_1$ subject to $A\mathbf{z} = A\mathbf{x}$ if and only if the null space property of order s holds, i.e.,

$$\|\mathbf{v}_S\|_1 \leq \|\mathbf{v}\|_1/2 \quad \text{for all } \mathbf{v} \in \ker A \text{ and all } S \in [N] \text{ with } |S| \leq s.$$

Setting $d := d^m(B_1^N, \ell_p^N)$, there exists $A \in \mathbb{R}^{m \times N}$ such that

$$\|\mathbf{v}\|_p \leq d\|\mathbf{v}\|_1 \quad \text{for all } \mathbf{v} \in \ker A.$$

Then, for $\mathbf{v} \in \ker A$ and $S \in [N]$ with $|S| \leq s$,

$$\|\mathbf{v}_S\|_1 \leq s^{1-1/p}\|\mathbf{v}_S\|_p \leq s^{1-1/p}\|\mathbf{v}\|_p$$

Insight 1: width and ℓ_1-recovery*

Every s-sparse $\mathbf{x} \in \mathbb{R}^N$ is a minimizer of $\|\mathbf{z}\|_1$ subject to $A\mathbf{z} = A\mathbf{x}$ if and only if the null space property of order s holds, i.e.,

$$\|\mathbf{v}_S\|_1 \leq \|\mathbf{v}\|_1/2 \quad \text{for all } \mathbf{v} \in \ker A \text{ and all } S \in [N] \text{ with } |S| \leq s.$$

Setting $d := d^m(B_1^N, \ell_p^N)$, there exists $A \in \mathbb{R}^{m \times N}$ such that

$$\|\mathbf{v}\|_p \leq d\|\mathbf{v}\|_1 \quad \text{for all } \mathbf{v} \in \ker A.$$

Then, for $\mathbf{v} \in \ker A$ and $S \in [N]$ with $|S| \leq s$,

$$\|\mathbf{v}_S\|_1 \leq s^{1-1/p}\|\mathbf{v}_S\|_p \leq s^{1-1/p}\|\mathbf{v}\|_p \leq ds^{1-1/p}\|\mathbf{v}\|_1.$$

Insight 1: width and ℓ_1-recovery*

Every s-sparse $\mathbf{x} \in \mathbb{R}^N$ is a minimizer of $\|\mathbf{z}\|_1$ subject to $A\mathbf{z} = A\mathbf{x}$ if and only if the null space property of order s holds, i.e.,

$$\|\mathbf{v}_S\|_1 \leq \|\mathbf{v}\|_1/2 \quad \text{for all } \mathbf{v} \in \ker A \text{ and all } S \in [N] \text{ with } |S| \leq s.$$

Setting $d := d^m(B_1^N, \ell_p^N)$, there exists $A \in \mathbb{R}^{m \times N}$ such that

$$\|\mathbf{v}\|_p \leq d\|\mathbf{v}\|_1 \quad \text{for all } \mathbf{v} \in \ker A.$$

Then, for $\mathbf{v} \in \ker A$ and $S \in [N]$ with $|S| \leq s$,

$$\|\mathbf{v}_S\|_1 \leq s^{1-1/p}\|\mathbf{v}_S\|_p \leq s^{1-1/p}\|\mathbf{v}\|_p \leq ds^{1-1/p}\|\mathbf{v}\|_1.$$

Choose $s \approx \left(\frac{1}{2d}\right)^{\frac{p}{p-1}}$

Insight 1: width and ℓ_1-recovery*

Every s-sparse $\mathbf{x} \in \mathbb{R}^N$ is a minimizer of $\|\mathbf{z}\|_1$ subject to $A\mathbf{z} = A\mathbf{x}$ if and only if the null space property of order s holds, i.e.,

$$\|\mathbf{v}_S\|_1 \leq \|\mathbf{v}\|_1/2 \quad \text{for all } \mathbf{v} \in \ker A \text{ and all } S \in [N] \text{ with } |S| \leq s.$$

Setting $d := d^m(B_1^N, \ell_p^N)$, there exists $A \in \mathbb{R}^{m \times N}$ such that

$$\|\mathbf{v}\|_p \leq d\|\mathbf{v}\|_1 \quad \text{for all } \mathbf{v} \in \ker A.$$

Then, for $\mathbf{v} \in \ker A$ and $S \in [N]$ with $|S| \leq s$,

$$\|\mathbf{v}_S\|_1 \leq s^{1-1/p}\|\mathbf{v}_S\|_p \leq s^{1-1/p}\|\mathbf{v}\|_p \leq ds^{1-1/p}\|\mathbf{v}\|_1.$$

Choose $s \approx \left(\frac{1}{2d}\right)^{\frac{p}{p-1}}$ to derive the null space property of order s.

A combinatorial lemma

A combinatorial lemma

Lemma

There exists $n \geq \left(\frac{N}{4s}\right)^{\frac{s}{2}}$ *subsets* $S^1, \ldots, S^n$ *of size* s *such that*

$$|S^j \cap S^k| < \frac{s}{2} \quad \text{for all } 1 \leq j \neq k \leq n.$$

A combinatorial lemma

Lemma

There exists $n \geq \left(\frac{N}{4s}\right)^{\frac{s}{2}}$ *subsets* $S^1, \ldots, S^n$ *of size* s *such that*

$$|S^j \cap S^k| < \frac{s}{2} \quad \textit{for all } 1 \leq j \neq k \leq n.$$

Define s-sparse vectors $\mathbf{x}^1, \ldots, \mathbf{x}^n$ by

$$(\mathbf{x}^j)_i = \begin{cases} 1/s & \text{if } i \in S^j, \\ 0 & \text{if } i \notin S^j. \end{cases}$$

A combinatorial lemma

Lemma

There exists $n \geq \left(\frac{N}{4s}\right)^{\frac{s}{2}}$ *subsets* $S^1, \ldots, S^n$ *of size* s *such that*

$$|S^j \cap S^k| < \frac{s}{2} \quad \textit{for all } 1 \leq j \neq k \leq n.$$

Define s-sparse vectors $\mathbf{x}^1, \ldots, \mathbf{x}^n$ by

$$(\mathbf{x}^j)_i = \begin{cases} 1/s & \text{if } i \in S^j, \\ 0 & \text{if } i \notin S^j. \end{cases}$$

Note that $\|\mathbf{x}^j\|_1 = 1$

A combinatorial lemma

Lemma

There exists $n \geq \left(\frac{N}{4s}\right)^{\frac{s}{2}}$ *subsets* $S^1, \ldots, S^n$ *of size* s *such that*

$$|S^j \cap S^k| < \frac{s}{2} \quad \textit{for all } 1 \leq j \neq k \leq n.$$

Define s-sparse vectors $\mathbf{x}^1, \ldots, \mathbf{x}^n$ by

$$(\mathbf{x}^j)_i = \begin{cases} 1/s & \text{if } i \in S^j, \\ 0 & \text{if } i \notin S^j. \end{cases}$$

Note that $\|\mathbf{x}^j\|_1 = 1$ and $\|\mathbf{x}^j - \mathbf{x}^k\|_1 > 1$ for all $1 \leq j \neq k \leq n$.

Insight 2: ℓ_1-recovery and number of measurements*

Insight 2: ℓ_1-recovery and number of measurements*

For $A \in \mathbb{R}^{m \times N}$, suppose that every $2s$-sparse vector $\mathbf{x} \in \mathbb{R}^N$ is a minimizer of $\|\mathbf{z}\|_1$ subject to $A\mathbf{z} = A\mathbf{x}$.

Insight 2: ℓ_1-recovery and number of measurements*

For $A \in \mathbb{R}^{m \times N}$, suppose that every $2s$-sparse vector $\mathbf{x} \in \mathbb{R}^N$ is a minimizer of $\|\mathbf{z}\|_1$ subject to $A\mathbf{z} = A\mathbf{x}$.

In the quotient space $\ell_1^N / \ker A$, this means

$$\|[\mathbf{x}]\| := \inf_{\mathbf{v} \in \ker A} \|\mathbf{x} - \mathbf{v}\|_1$$

Insight 2: ℓ_1-recovery and number of measurements*

For $A \in \mathbb{R}^{m \times N}$, suppose that every $2s$-sparse vector $\mathbf{x} \in \mathbb{R}^N$ is a minimizer of $\|\mathbf{z}\|_1$ subject to $A\mathbf{z} = A\mathbf{x}$.

In the quotient space $\ell_1^N / \ker A$, this means

$$\|[\mathbf{x}]\| := \inf_{\mathbf{v} \in \ker A} \|\mathbf{x} - \mathbf{v}\|_1 = \|\mathbf{x}\|_1 \qquad \text{for all } 2s\text{-sparse } \mathbf{x} \in \mathbb{R}^N.$$

Insight 2: ℓ_1-recovery and number of measurements*

For $A \in \mathbb{R}^{m \times N}$, suppose that every $2s$-sparse vector $\mathbf{x} \in \mathbb{R}^N$ is a minimizer of $\|\mathbf{z}\|_1$ subject to $A\mathbf{z} = A\mathbf{x}$.
In the quotient space $\ell_1^N / \ker A$, this means

$$\|[\mathbf{x}]\| := \inf_{\mathbf{v} \in \ker A} \|\mathbf{x} - \mathbf{v}\|_1 = \|\mathbf{x}\|_1 \qquad \text{for all } 2s\text{-sparse } \mathbf{x} \in \mathbb{R}^N.$$

In particular,

$$\|[\mathbf{x}^j]\| = 1$$

Insight 2: ℓ_1-recovery and number of measurements*

For $A \in \mathbb{R}^{m \times N}$, suppose that every $2s$-sparse vector $\mathbf{x} \in \mathbb{R}^N$ is a minimizer of $\|\mathbf{z}\|_1$ subject to $A\mathbf{z} = A\mathbf{x}$.
In the quotient space $\ell_1^N / \ker A$, this means

$$\|[\mathbf{x}]\| := \inf_{\mathbf{v} \in \ker A} \|\mathbf{x} - \mathbf{v}\|_1 = \|\mathbf{x}\|_1 \qquad \text{for all } 2s\text{-sparse } \mathbf{x} \in \mathbb{R}^N.$$

In particular,

$$\|[\mathbf{x}^j]\| = 1 \quad \text{and} \quad \|[\mathbf{x}^j] - [\mathbf{x}^k]\| > 1, \qquad \text{all } 1 \le j \ne k \le n.$$

Insight 2: ℓ_1-recovery and number of measurements*

For $A \in \mathbb{R}^{m \times N}$, suppose that every $2s$-sparse vector $\mathbf{x} \in \mathbb{R}^N$ is a minimizer of $\|\mathbf{z}\|_1$ subject to $A\mathbf{z} = A\mathbf{x}$.
In the quotient space $\ell_1^N / \ker A$, this means

$$\|[\mathbf{x}]\| := \inf_{\mathbf{v} \in \ker A} \|\mathbf{x} - \mathbf{v}\|_1 = \|\mathbf{x}\|_1 \qquad \text{for all } 2s\text{-sparse } \mathbf{x} \in \mathbb{R}^N.$$

In particular,

$$\|[\mathbf{x}^j]\| = 1 \quad \text{and} \quad \|[\mathbf{x}^j] - [\mathbf{x}^k]\| > 1, \qquad \text{all } 1 \le j \ne k \le n.$$

The size of this 1-separating set of the unit sphere satisfies

$$n \le \left(1 + \frac{2}{1}\right)^m$$

Insight 2: ℓ_1-recovery and number of measurements*

For $A \in \mathbb{R}^{m \times N}$, suppose that every $2s$-sparse vector $\mathbf{x} \in \mathbb{R}^N$ is a minimizer of $\|\mathbf{z}\|_1$ subject to $A\mathbf{z} = A\mathbf{x}$.
In the quotient space $\ell_1^N / \ker A$, this means

$$\|[\mathbf{x}]\| := \inf_{\mathbf{v} \in \ker A} \|\mathbf{x} - \mathbf{v}\|_1 = \|\mathbf{x}\|_1 \qquad \text{for all } 2s\text{-sparse } \mathbf{x} \in \mathbb{R}^N.$$

In particular,

$$\|[\mathbf{x}^j]\| = 1 \quad \text{and} \quad \|[\mathbf{x}^j] - [\mathbf{x}^k]\| > 1, \qquad \text{all } 1 \le j \ne k \le n.$$

The size of this 1-separating set of the unit sphere satisfies

$$n \le \left(1 + \frac{2}{1}\right)^m = 3^m.$$

Insight 2: ℓ_1-recovery and number of measurements*

For $A \in \mathbb{R}^{m \times N}$, suppose that every $2s$-sparse vector $\mathbf{x} \in \mathbb{R}^N$ is a minimizer of $\|\mathbf{z}\|_1$ subject to $A\mathbf{z} = A\mathbf{x}$.
In the quotient space $\ell_1^N / \ker A$, this means

$$\|[\mathbf{x}]\| := \inf_{\mathbf{v} \in \ker A} \|\mathbf{x} - \mathbf{v}\|_1 = \|\mathbf{x}\|_1 \qquad \text{for all } 2s\text{-sparse } \mathbf{x} \in \mathbb{R}^N.$$

In particular,

$$\|[\mathbf{x}^j]\| = 1 \quad \text{and} \quad \|[\mathbf{x}^j] - [\mathbf{x}^k]\| > 1, \qquad \text{all } 1 \le j \ne k \le n.$$

The size of this 1-separating set of the unit sphere satisfies

$$\left(\frac{N}{4s}\right)^{\frac{s}{2}} \le n \le \left(1 + \frac{2}{1}\right)^m = 3^m.$$

Insight 2: ℓ_1-recovery and number of measurements*

For $A \in \mathbb{R}^{m \times N}$, suppose that every $2s$-sparse vector $\mathbf{x} \in \mathbb{R}^N$ is a minimizer of $\|\mathbf{z}\|_1$ subject to $A\mathbf{z} = A\mathbf{x}$.
In the quotient space $\ell_1^N / \ker A$, this means

$$\|[\mathbf{x}]\| := \inf_{\mathbf{v} \in \ker A} \|\mathbf{x} - \mathbf{v}\|_1 = \|\mathbf{x}\|_1 \qquad \text{for all } 2s\text{-sparse } \mathbf{x} \in \mathbb{R}^N.$$

In particular,

$$\|[\mathbf{x}^j]\| = 1 \quad \text{and} \quad \|[\mathbf{x}^j] - [\mathbf{x}^k]\| > 1, \qquad \text{all } 1 \le j \ne k \le n.$$

The size of this 1-separating set of the unit sphere satisfies

$$\left(\frac{N}{4s}\right)^{\frac{s}{2}} \le n \le \left(1 + \frac{2}{1}\right)^m = 3^m.$$

Taking the logarithm yields

$$m \ge \frac{s}{\ln 9} \ln\left(\frac{N}{4s}\right).$$

Uniform vs. Nonuniform Guarantees*

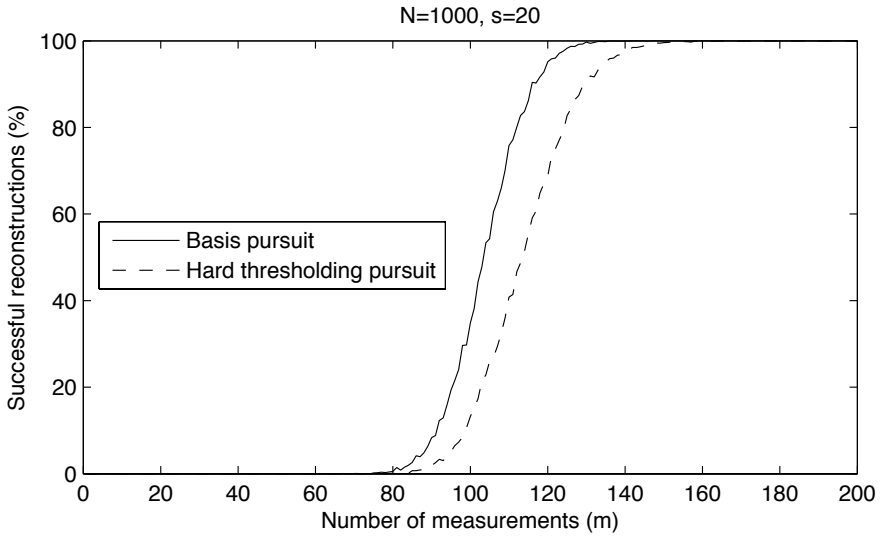
Empirical Performances: HTP and ℓ_1-Minimization

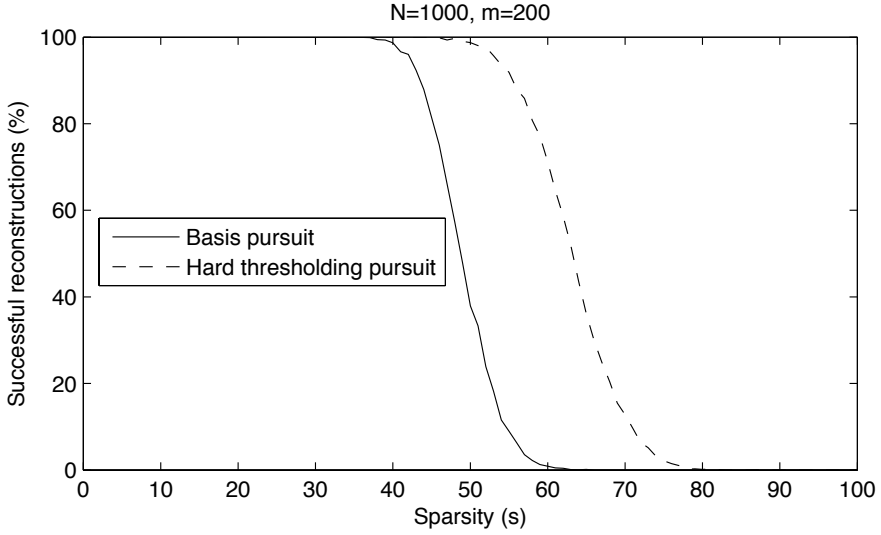
Empirical Performances: HTP and ℓ_1-Minimization

Gaussian matrices with Rademacher then with Gaussian vectors

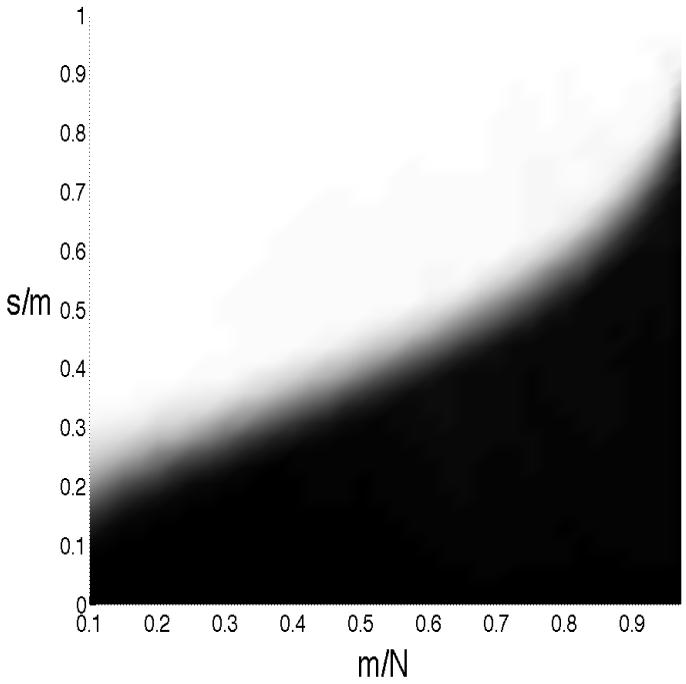
Empirical Performances: HTP and ℓ_1-Minimization

Gaussian matrices with Rademacher then with Gaussian vectors





Empirical Performances: Phase Transitions*



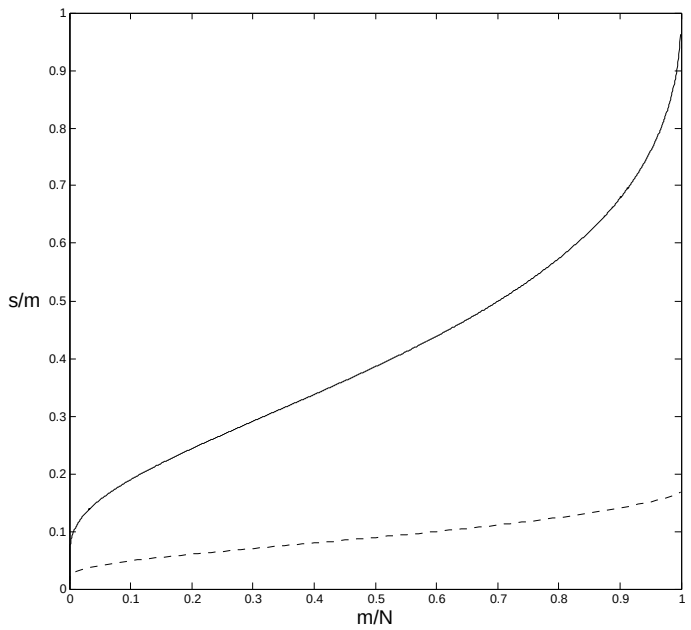


Figure : L: empirically observed weak threshold
R: strong (dashed) and weak (solid) thresholds (courtesy of J. Tanner)

Uniform vs. Nonuniform Sparse Recovery

Uniform vs. Nonuniform Sparse Recovery

When A is a random matrix, uniform results read

$$\mathbb{P}(\forall\, s\text{-sparse } \mathbf{x},\ \text{recovery of } \mathbf{x} \text{ is successful using } A) \geq 1 - \varepsilon,$$

Uniform vs. Nonuniform Sparse Recovery

When A is a random matrix, uniform results read

$$\mathbb{P}(\forall\, s\text{-sparse } \mathbf{x},\ \text{recovery of } \mathbf{x} \text{ is successful using } A) \geq 1 - \varepsilon,$$

while nonuniform results have the form

$$\forall\, s\text{-sparse } \mathbf{x},\ \mathbb{P}(\text{recovery of } \mathbf{x} \text{ is successful using } A) \geq 1 - \varepsilon.$$

Uniform vs. Nonuniform Sparse Recovery

When A is a random matrix, uniform results read

$$\mathbb{P}(\forall\, s\text{-sparse } \mathbf{x},\ \text{recovery of } \mathbf{x} \text{ is successful using } A) \geq 1 - \varepsilon,$$

while nonuniform results have the form

$$\forall\, s\text{-sparse } \mathbf{x},\ \mathbb{P}(\text{recovery of } \mathbf{x} \text{ is successful using } A) \geq 1 - \varepsilon.$$

Using Gaussian matrices, nonuniform s-sparse recovery results hold for exactly s iterations of OMP in the regime $m \geq cs\ln(N/s)$, but uniform results do not hold in this regime.

Nonuniform Recovery via BP when $m \gtrsim 2s \ln(N/s)$*

Nonuniform Recovery via BP when $m \gtrsim 2s \ln(N/s)$*

- A necessary and sufficient condition for the recovery of a fixed $\mathbf{x} \in \mathbb{C}^N$ supported on S via ℓ_1-minimization is

$$\left| \sum_{j \in S} \overline{\operatorname{sgn}(x_j)} v_j \right| < \|\mathbf{v}_{\overline{S}}\|_1 \qquad \text{for all } \mathbf{v}_{\neq \mathbf{0}} \in \ker A.$$

Nonuniform Recovery via BP when $m \gtrsim 2s\ln(N/s)$*

- A necessary and sufficient condition for the recovery of a fixed $\mathbf{x} \in \mathbb{C}^N$ supported on S via ℓ_1-minimization is

$$\left| \sum_{j \in S} \overline{\operatorname{sgn}(x_j)} v_j \right| < \|\mathbf{v}_{\overline{S}}\|_1 \qquad \text{for all } \mathbf{v}_{\neq \mathbf{0}} \in \ker A.$$

- This is implied by (in the real case, equivalent to) the injectivity of A_S and the existence of $\mathbf{h} \in \mathbb{C}^m$ such that

$$(A^*\mathbf{h})_j = \operatorname{sgn}(x_j),\ j \in S, \qquad \left|(A^*\mathbf{h})_\ell\right| < 1,\ \ell \in \overline{S}.$$

Nonuniform Recovery via BP when $m \gtrsim 2s\ln(N/s)$*

- A necessary and sufficient condition for the recovery of a fixed $\mathbf{x} \in \mathbb{C}^N$ supported on S via ℓ_1-minimization is

$$\left| \sum_{j \in S} \overline{\operatorname{sgn}(x_j)} v_j \right| < \|\mathbf{v}_{\overline{S}}\|_1 \qquad \text{for all } \mathbf{v}_{\neq \mathbf{0}} \in \ker A.$$

- This is implied by (in the real case, equivalent to) the injectivity of A_S and the existence of $\mathbf{h} \in \mathbb{C}^m$ such that

$$(A^*\mathbf{h})_j = \operatorname{sgn}(x_j),\ j \in S, \qquad \left|(A^*\mathbf{h})_\ell\right| < 1,\ \ell \in \overline{S}.$$

- The latter is in turn implied by

$$\left|\langle A_S^\dagger \mathbf{a}_\ell, \operatorname{sgn}(\mathbf{x}_S)\rangle\right| < 1 \qquad \text{for all } \ell \in \overline{S}.$$

Nonuniform Recovery via BP when $m \gtrsim 2s\ln(N/s)$*

- A necessary and sufficient condition for the recovery of a fixed $\mathbf{x} \in \mathbb{C}^N$ supported on S via ℓ_1-minimization is

$$\left| \sum_{j \in S} \overline{\operatorname{sgn}(x_j)} v_j \right| < \|\mathbf{v}_{\overline{S}}\|_1 \qquad \text{for all } \mathbf{v}_{\neq \mathbf{0}} \in \ker A.$$

- This is implied by (in the real case, equivalent to) the injectivity of A_S and the existence of $\mathbf{h} \in \mathbb{C}^m$ such that

$$(A^*\mathbf{h})_j = \operatorname{sgn}(x_j),\ j \in S, \qquad \left|(A^*\mathbf{h})_\ell\right| < 1,\ \ell \in \overline{S}.$$

- The latter is in turn implied by

$$\left|\langle A_S^\dagger \mathbf{a}_\ell, \operatorname{sgn}(\mathbf{x}_S)\rangle\right| < 1 \qquad \text{for all } \ell \in \overline{S}.$$

- For subgaussian matrices, this occurs with probability $\geq 1 - \varepsilon$ provided

$$m \gtrsim 2s\ln(2N/\varepsilon)$$

Nonuniform Recovery via BP when $m \gtrsim 2s\ln(N/s)$*

- A necessary and sufficient condition for the recovery of a fixed $\mathbf{x} \in \mathbb{C}^N$ supported on S via ℓ_1-minimization is

$$\left| \sum_{j \in S} \overline{\operatorname{sgn}(x_j)} v_j \right| < \|\mathbf{v}_{\overline{S}}\|_1 \qquad \text{for all } \mathbf{v}_{\neq \mathbf{0}} \in \ker A.$$

- This is implied by (in the real case, equivalent to) the injectivity of A_S and the existence of $\mathbf{h} \in \mathbb{C}^m$ such that

$$(A^*\mathbf{h})_j = \operatorname{sgn}(x_j),\ j \in S, \qquad \left|(A^*\mathbf{h})_\ell\right| < 1,\ \ell \in \overline{S}.$$

- The latter is in turn implied by

$$\left|\langle A_S^\dagger \mathbf{a}_\ell, \operatorname{sgn}(\mathbf{x}_S)\rangle\right| < 1 \qquad \text{for all } \ell \in \overline{S}.$$

- For subgaussian matrices, this occurs with probability $\geq 1 - \varepsilon$ provided

$$m \gtrsim 2s\ln(2N/\varepsilon)$$

(which is replaced by $m \gtrsim 2s\big(\sqrt{\ln(eN/s)} + \sqrt{\ln(1/\varepsilon)/s}\big)^2$ for Gaussian matrices using Gordon's *escape through the mesh*).

Bounded Orthonormal Systems

Bounded Orthonormal Systems

Let $\mathcal{D} \in \mathbb{R}^d$ be endowed with a probability measure ν. A bounded orthonormal system (BOS) with constant $K \geq 1$ is a system $(\phi_1, \ldots, \phi_N)$ of function of $\mathcal{D}$ satisfying

$$\int_{\mathcal{D}} \phi_j(t)\overline{\phi_k(t)} d\nu(t) = \delta_{j,k},$$

$$\sup_{1 \leq j \leq N} \sup_{t \in \mathcal{D}} |\phi_j(t)| \leq K.$$

Bounded Orthonormal Systems

Let $\mathcal{D} \in \mathbb{R}^d$ be endowed with a probability measure ν. A bounded orthonormal system (BOS) with constant $K \geq 1$ is a system $(\phi_1, \ldots, \phi_N)$ of function of $\mathcal{D}$ satisfying

$$\int_{\mathcal{D}} \phi_j(t)\overline{\phi_k(t)} d\nu(t) = \delta_{j,k},$$

$$\sup_{1 \leq j \leq N} \sup_{t \in \mathcal{D}} |\phi_j(t)| \leq K.$$

1. trigonometric polynomials:
   take $\phi_k(t) = e^{i2\pi kt}$ for $t \in \mathcal{D} = [0, 1]$, $d\nu(t) = dt$,

Bounded Orthonormal Systems

Let $\mathcal{D} \in \mathbb{R}^d$ be endowed with a probability measure ν. A bounded orthonormal system (BOS) with constant $K \geq 1$ is a system $(\phi_1, \ldots, \phi_N)$ of function of $\mathcal{D}$ satisfying

$$\int_{\mathcal{D}} \phi_j(t)\overline{\phi_k(t)} d\nu(t) = \delta_{j,k},$$

$$\sup_{1 \leq j \leq N} \sup_{t \in \mathcal{D}} |\phi_j(t)| \leq K.$$

1. trigonometric polynomials:
   take $\phi_k(t) = e^{i2\pi kt}$ for $t \in \mathcal{D} = [0,1]$, $d\nu(t) = dt$,
2. discrete orthonormal systems — let $U \in \mathbb{C}^{N \times N}$ be a unitary matrix with $\sqrt{N}|U_{k,j}| \leq K$ for all $k, j \in \{1, \ldots, N\}$ (e.g. Fourier or Hadamard matrix):
   take $\phi_k(t) = \sqrt{N} U_{k,t}$ for $t \in \{1, \ldots, N\}$, $\nu(B) = \dfrac{\mathrm{card}(B)}{N}$.

Nonuniform Recovery for BOS*

Nonuniform Recovery for BOS*

For $f = \sum_{j=1}^{N} x_j \phi_j$, the samples of f at $t_1, \ldots, t_m \in \mathcal{D}$ are

$$f(t_k) = \sum_{j=1}^{N} x_j \phi_j(t_k) = (A\mathbf{x})_k,$$

where the sampling matrix $A \in \mathbb{C}^{m \times N}$ have entries

$$A_{k,j} = \phi_j(t_k). \tag{3}$$

Nonuniform Recovery for BOS*

For $f = \sum_{j=1}^{N} x_j \phi_j$, the samples of f at $t_1, \ldots, t_m \in \mathcal{D}$ are

$$f(t_k) = \sum_{j=1}^{N} x_j \phi_j(t_k) = (A\mathbf{x})_k,$$

where the sampling matrix $A \in \mathbb{C}^{m \times N}$ have entries

$$A_{k,j} = \phi_j(t_k). \tag{3}$$

Let an s-sparse $\mathbf{x} \in \mathbb{C}^N$ be fixed.

Nonuniform Recovery for BOS*

For $f = \sum_{j=1}^{N} x_j \phi_j$, the samples of f at $t_1, \ldots, t_m \in \mathcal{D}$ are

$$f(t_k) = \sum_{j=1}^{N} x_j \phi_j(t_k) = (A\mathbf{x})_k,$$

where the sampling matrix $A \in \mathbb{C}^{m \times N}$ have entries

$$A_{k,j} = \phi_j(t_k). \tag{3}$$

Let an s-sparse $\mathbf{x} \in \mathbb{C}^N$ be fixed. Let sampling points $t_1, \ldots, t_m$ be randomly selected according to ν, and let $A \in \mathbb{C}^{m \times N}$ be the sampling matrix (3).

Nonuniform Recovery for BOS*

For $f = \sum_{j=1}^{N} x_j \phi_j$, the samples of f at $t_1, \ldots, t_m \in \mathcal{D}$ are

$$f(t_k) = \sum_{j=1}^{N} x_j \phi_j(t_k) = (A\mathbf{x})_k,$$

where the sampling matrix $A \in \mathbb{C}^{m \times N}$ have entries

$$A_{k,j} = \phi_j(t_k). \tag{3}$$

Let an s-sparse $\mathbf{x} \in \mathbb{C}^N$ be fixed. Let sampling points $t_1, \ldots, t_m$ be randomly selected according to ν, and let $A \in \mathbb{C}^{m \times N}$ be the sampling matrix (3). If

$$m \geq C\, K^2 s \ln(N) \ln(\varepsilon^{-1}),$$

then $\mathbf{x}$ is the unique minimizer of $\|\mathbf{z}\|_1$ subject to $A\mathbf{z} = A\mathbf{x}$ with probability at least $1 - \varepsilon$.

Nonuniform Recovery for BOS*

For $f = \sum_{j=1}^{N} x_j \phi_j$, the samples of f at $t_1, \ldots, t_m \in \mathcal{D}$ are

$$f(t_k) = \sum_{j=1}^{N} x_j \phi_j(t_k) = (A\mathbf{x})_k,$$

where the sampling matrix $A \in \mathbb{C}^{m \times N}$ have entries

$$A_{k,j} = \phi_j(t_k). \tag{3}$$

Let an s-sparse $\mathbf{x} \in \mathbb{C}^N$ be fixed. Let sampling points $t_1, \ldots, t_m$ be randomly selected according to ν, and let $A \in \mathbb{C}^{m \times N}$ be the sampling matrix (3). If

$$m \geq C\, K^2 s \ln(N) \ln(\varepsilon^{-1}),$$

then $\mathbf{x}$ is the unique minimizer of $\|\mathbf{z}\|_1$ subject to $A\mathbf{z} = A\mathbf{x}$ with probability at least $1 - \varepsilon$.

(Proof based on the *golfing scheme*).

Restricted Isometry Property for BOS*

Restricted Isometry Property for BOS*

Let $A \in \mathbb{C}^{m \times N}$ be the random sampling matrix associated to a BOS with constant $K \geq 1$.

Restricted Isometry Property for BOS*

Let $A \in \mathbb{C}^{m \times N}$ be the random sampling matrix associated to a BOS with constant $K \geq 1$. For $\delta \in (0,1)$, if

$$m \geq C\, K^2 \delta^{-2} s \ln^4(N),$$

then, with probability at least $1 - N^{-\ln^3(N)}$, the matrix $\frac{1}{\sqrt{m}} A$ has a restricted isometry constant satisfying $\delta_s \leq \delta$.

Restricted Isometry Property for BOS*

Let $A \in \mathbb{C}^{m \times N}$ be the random sampling matrix associated to a BOS with constant $K \geq 1$. For $\delta \in (0,1)$, if

$$m \geq C\, K^2 \delta^{-2} s \ln^4(N),$$

then, with probability at least $1 - N^{-\ln^3(N)}$, the matrix $\frac{1}{\sqrt{m}} A$ has a restricted isometry constant satisfying $\delta_s \leq \delta$.

(Proof uses *Dudley's inequality*, *empirical method of Maurey*, etc.)

Stability: Uniform Setting*

Stability: Uniform Setting*

- Given $q \geq p \geq 1$, a pair (A, Δ) is mixed (ℓ_q, ℓ_p)-instance optimal of order s with constant $C > 0$ if

$$\|\mathbf{x} - \Delta(A\mathbf{x})\|_q \leq \frac{C}{s^{1/p-1/q}} \sigma_s(\mathbf{x})_p \qquad \text{for all } \mathbf{x} \in \mathbb{C}^N.$$

Stability: Uniform Setting*

- Given $q \geq p \geq 1$, a pair (A, Δ) is mixed (ℓ_q, ℓ_p)-instance optimal of order s with constant $C > 0$ if

$$\|\mathbf{x} - \Delta(A\mathbf{x})\|_q \leq \frac{C}{s^{1/p-1/q}} \sigma_s(\mathbf{x})_p \qquad \text{for all } \mathbf{x} \in \mathbb{C}^N.$$

- Let $A \in \mathbb{R}^{m \times N}$.

Stability: Uniform Setting*

- Given $q \geq p \geq 1$, a pair (A, Δ) is mixed (ℓ_q, ℓ_p)-instance optimal of order s with constant $C > 0$ if

$$\|\mathbf{x} - \Delta(A\mathbf{x})\|_q \leq \frac{C}{s^{1/p-1/q}} \sigma_s(\mathbf{x})_p \qquad \text{for all } \mathbf{x} \in \mathbb{C}^N.$$

- Let $A \in \mathbb{R}^{m \times N}$. If there exists Δ making (A, Δ) mixed (ℓ_q, ℓ_p)-instance optimal of order s with constant C, then

$$\|\mathbf{v}\|_q \leq \frac{C}{s^{1/p-1/q}} \sigma_{2s}(\mathbf{v})_p \qquad \text{for all } \mathbf{v} \in \ker A. \tag{4}$$

Stability: Uniform Setting*

- Given $q \geq p \geq 1$, a pair (A, Δ) is mixed (ℓ_q, ℓ_p)-instance optimal of order s with constant $C > 0$ if

$$\|\mathbf{x} - \Delta(A\mathbf{x})\|_q \leq \frac{C}{s^{1/p-1/q}} \sigma_s(\mathbf{x})_p \qquad \text{for all } \mathbf{x} \in \mathbb{C}^N.$$

- Let $A \in \mathbb{R}^{m \times N}$. If there exists Δ making (A, Δ) mixed (ℓ_q, ℓ_p)-instance optimal of order s with constant C, then

$$\|\mathbf{v}\|_q \leq \frac{C}{s^{1/p-1/q}} \sigma_{2s}(\mathbf{v})_p \qquad \text{for all } \mathbf{v} \in \ker A. \tag{4}$$

Conversely, if (4) holds, then there exists Δ making (A, Δ) mixed (ℓ_q, ℓ_p)-instance optimal of order s with constant $2C$.

Stability: Uniform Setting*

- Given $q \geq p \geq 1$, a pair (A, Δ) is mixed (ℓ_q, ℓ_p)-instance optimal of order s with constant $C > 0$ if

$$\|\mathbf{x} - \Delta(A\mathbf{x})\|_q \leq \frac{C}{s^{1/p-1/q}} \sigma_s(\mathbf{x})_p \qquad \text{for all } \mathbf{x} \in \mathbb{C}^N.$$

- Let $A \in \mathbb{R}^{m \times N}$. If there exists Δ making (A, Δ) mixed (ℓ_q, ℓ_p)-instance optimal of order s with constant C, then

$$\|\mathbf{v}\|_q \leq \frac{C}{s^{1/p-1/q}} \sigma_{2s}(\mathbf{v})_p \qquad \text{for all } \mathbf{v} \in \ker A. \tag{4}$$

  Conversely, if (4) holds, then there exists Δ making (A, Δ) mixed (ℓ_q, ℓ_p)-instance optimal of order s with constant $2C$.

- If there is an ℓ_2-instance optimal pair of order $s \geq 1$ with constant C, then

$$m \geq c\, N$$

  for some constant c depending only on C.

Stability: Nonuniform Setting*

Stability: Nonuniform Setting*

Uniform ℓ_2-instance optimality may be irrelevant, but nonuniform ℓ_2-instance optimality is quite relevant...

Stability: Nonuniform Setting*

Uniform ℓ_2-instance optimality may be irrelevant, but nonuniform ℓ_2-instance optimality is quite relevant...

Let an s-sparse $\mathbf{x} \in \mathbb{R}^N$ be fixed.

Stability: Nonuniform Setting*

Uniform ℓ_2-instance optimality may be irrelevant, but nonuniform ℓ_2-instance optimality is quite relevant...

Let an s-sparse $\mathbf{x} \in \mathbb{R}^N$ be fixed. Let $A \in \mathbb{R}^{m \times N}$ be a matrix with ind $\mathcal{N}(0, m^{-1/2})$ entries.

Stability: Nonuniform Setting*

Uniform ℓ_2-instance optimality may be irrelevant, but nonuniform ℓ_2-instance optimality is quite relevant...

Let an s-sparse $\mathbf{x} \in \mathbb{R}^N$ be fixed. Let $A \in \mathbb{R}^{m \times N}$ be a matrix with ind $\mathcal{N}(0, m^{-1/2})$ entries. If $N \geq c_2 m$ and $m \geq c_3 s \ln(N/m)$, then

$$\|\mathbf{x} - \Delta_1(A\mathbf{x} + \mathbf{e})\|_2 \leq C\, \sigma_s(\mathbf{x})_2 + D\, \|\mathbf{e}\|_2$$

holds for all $\mathbf{e} \in \mathbb{R}^m$ with probability at least $1 - 5\exp(-c_1 m)$.

Stability: Nonuniform Setting*

Uniform ℓ_2-instance optimality may be irrelevant, but nonuniform ℓ_2-instance optimality is quite relevant...

Let an s-sparse $\mathbf{x} \in \mathbb{R}^N$ be fixed. Let $A \in \mathbb{R}^{m \times N}$ be a matrix with ind $\mathcal{N}(0, m^{-1/2})$ entries. If $N \geq c_2 m$ and $m \geq c_3 s \ln(N/m)$, then

$$\|\mathbf{x} - \Delta_1(A\mathbf{x} + \mathbf{e})\|_2 \leq C\, \sigma_s(\mathbf{x})_2 + D\, \|\mathbf{e}\|_2$$

holds for all $\mathbf{e} \in \mathbb{R}^m$ with probability at least $1 - 5\exp(-c_1 m)$.

This uses the ℓ_1-quotient property of A with constant d relative to the ℓ_2-norm $\|\cdot\|$ on $\mathbb{R}^m$, which is expressed in one of the forms:

Stability: Nonuniform Setting*

Uniform ℓ_2-instance optimality may be irrelevant, but nonuniform ℓ_2-instance optimality is quite relevant...

Let an s-sparse $\mathbf{x} \in \mathbb{R}^N$ be fixed. Let $A \in \mathbb{R}^{m \times N}$ be a matrix with ind $\mathcal{N}(0, m^{-1/2})$ entries. If $N \geq c_2 m$ and $m \geq c_3 s \ln(N/m)$, then

$$\|\mathbf{x} - \Delta_1(A\mathbf{x} + \mathbf{e})\|_2 \leq C\, \sigma_s(\mathbf{x})_2 + D\, \|\mathbf{e}\|_2$$

holds for all $\mathbf{e} \in \mathbb{R}^m$ with probability at least $1 - 5\exp(-c_1 m)$.

This uses the ℓ_1-quotient property of A with constant d relative to the ℓ_2-norm $\|\cdot\|$ on $\mathbb{R}^m$, which is expressed in one of the forms:

- for all $\mathbf{e} \in \mathbb{R}^m$, there exists $\mathbf{u} \in \mathbb{R}^N$ with

$$A\mathbf{u} = \mathbf{e} \qquad \text{and} \qquad \|\mathbf{u}\|_1 \leq d\sqrt{s_*}\|\mathbf{e}\|, \quad s_* := m/\ln(N/m),$$

Stability: Nonuniform Setting*

Uniform ℓ_2-instance optimality may be irrelevant, but nonuniform ℓ_2-instance optimality is quite relevant...

Let an s-sparse $\mathbf{x} \in \mathbb{R}^N$ be fixed. Let $A \in \mathbb{R}^{m \times N}$ be a matrix with ind $\mathcal{N}(0, m^{-1/2})$ entries. If $N \geq c_2 m$ and $m \geq c_3 s \ln(N/m)$, then

$$\|\mathbf{x} - \Delta_1(A\mathbf{x} + \mathbf{e})\|_2 \leq C\, \sigma_s(\mathbf{x})_2 + D\, \|\mathbf{e}\|_2$$

holds for all $\mathbf{e} \in \mathbb{R}^m$ with probability at least $1 - 5\exp(-c_1 m)$.

This uses the ℓ_1-quotient property of A with constant d relative to the ℓ_2-norm $\|\cdot\|$ on $\mathbb{R}^m$, which is expressed in one of the forms:

- for all $\mathbf{e} \in \mathbb{R}^m$, there exists $\mathbf{u} \in \mathbb{R}^N$ with

$$A\mathbf{u} = \mathbf{e} \qquad \text{and} \qquad \|\mathbf{u}\|_1 \leq d\sqrt{s_*}\|\mathbf{e}\|, \quad s_* := m/\ln(N/m),$$

- $\|[\mathbf{e}]\|_{\ell_1/\ker A} \leq d\sqrt{s_*}\|\mathbf{e}\|$ for all $\mathbf{e} \in \mathbb{R}^m$,

Stability: Nonuniform Setting*

Uniform ℓ_2-instance optimality may be irrelevant, but nonuniform ℓ_2-instance optimality is quite relevant...

Let an s-sparse $\mathbf{x} \in \mathbb{R}^N$ be fixed. Let $A \in \mathbb{R}^{m \times N}$ be a matrix with ind $\mathcal{N}(0, m^{-1/2})$ entries. If $N \geq c_2 m$ and $m \geq c_3 s \ln(N/m)$, then

$$\|\mathbf{x} - \Delta_1(A\mathbf{x} + \mathbf{e})\|_2 \leq C\, \sigma_s(\mathbf{x})_2 + D\, \|\mathbf{e}\|_2$$

holds for all $\mathbf{e} \in \mathbb{R}^m$ with probability at least $1 - 5\exp(-c_1 m)$.

This uses the ℓ_1-quotient property of A with constant d relative to the ℓ_2-norm $\|\cdot\|$ on $\mathbb{R}^m$, which is expressed in one of the forms:

- for all $\mathbf{e} \in \mathbb{R}^m$, there exists $\mathbf{u} \in \mathbb{R}^N$ with

$$A\mathbf{u} = \mathbf{e} \qquad \text{and} \qquad \|\mathbf{u}\|_1 \leq d\sqrt{s_*}\|\mathbf{e}\|, \quad s_* := m/\ln(N/m),$$

- $\|[\mathbf{e}]\|_{\ell_1/\ker A} \leq d\sqrt{s_*}\|\mathbf{e}\|$ for all $\mathbf{e} \in \mathbb{R}^m$,
- $\|\mathbf{e}\|_* \leq d\sqrt{s_*}\|A^*\mathbf{e}\|_\infty$ for all $\mathbf{e} \in \mathbb{R}^m$.