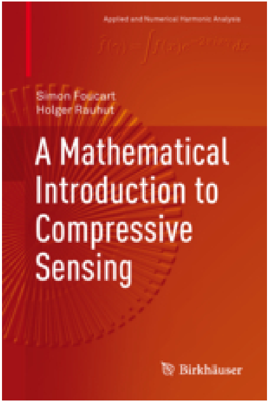
A Tutorial on Compressive Sensing

Simon Foucart
Drexel University / University of Georgia

CIMPA13
New Trends in Applied Harmonic Analysis
Mar del Plata, Argentina, 5-16 August 2013

This minicourse acts as an invitation to the elegant theory of Compressive Sensing. It aims at giving a solid overview of the fundamental mathematical aspects. Its content is partly based on a recent textbook coauthored with Holger Rauhut:



Part 1: The Standard Problem and First Algorithms

The lecture introduces the question of sparse recovery, establishes its theoretical limits, and presents an algorithm achieving these limits in an idealized situation. In order to treat more realistic situations, other algorithms are necessary. Basis Pursuit (that is, ℓ_1-minimization) and Orthogonal Matching Pursuit make their first appearance. Their success is proved using the concept of coherence of a matrix.

Keywords

Keywords

- Sparsity

  Essential

Keywords

- Sparsity

  Essential

- Randomness

  Nothing better so far
  (measurement process)

Keywords

- Sparsity

  Essential

- Randomness

  Nothing better so far
  (measurement process)

- Optimization

  Preferred, but competitive alternatives are available
  (reconstruction process)

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

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

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

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

Find concrete sensing/recovery protocols,

The Standard Compressive Sensing Problem

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

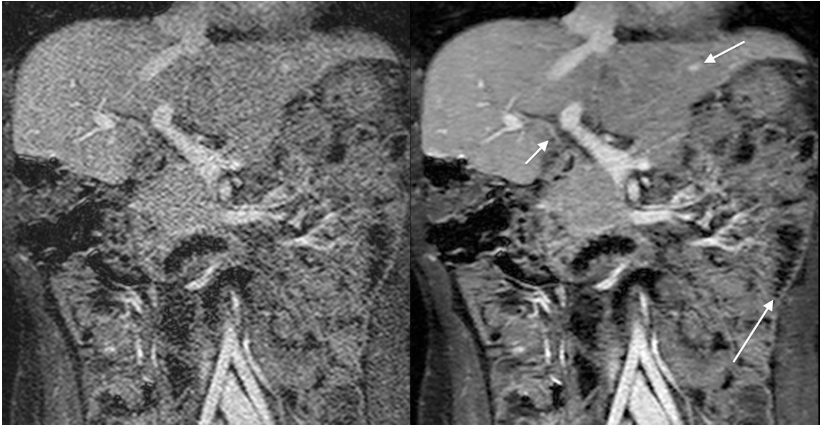


Figure: Left: traditional MRI reconstruction; Right: compressive sensing reconstruction (courtesy of M. Lustig and S. Vasanawala)

A Selection of Applications

- Magnetic resonance imaging

- Sampling theory

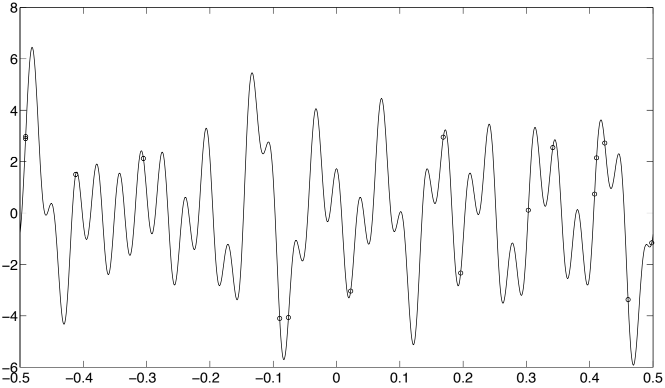


Figure: Time-domain signal with 16 samples.

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

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

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

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

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

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

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

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

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

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

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, $\text{card}(S) = s$. Consider the $2s$ Fourier coefficients

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

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

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

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

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

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

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

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

Identify an s-sparse $\mathbf{x} \in \mathbb{C}^N$ with a function x on $\{0, 1, \dots, N-1\}$ 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

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

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

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

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

$$(P_1) \qquad \underset{\mathbf{z}\in\mathbb{K}^N}{\text{minimize}} \; \|\mathbf{z}\|_1 \quad \text{subject to } A\mathbf{z} = \mathbf{y}.$$

- 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

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

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

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

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

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

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

For real measurement matrices, real and complex NSPs read

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

Basis Pursuit — Null Space Property

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

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

For real measurement matrices, real and complex NSPs read

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

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

Basis Pursuit — Null Space Property

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

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

For real measurement matrices, real and complex NSPs read

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

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

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

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

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

Orthogonal Matching Pursuit

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

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

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

- 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

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

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

- The norm of the residual decreases according to

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

- 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 $\mathbf{r}_{\neq \mathbf{0}} \in \big\{ A\mathbf{z}, \operatorname{supp}(\mathbf{z}) \subseteq S \big\}$.

First Recovery Guarantees

Coherence

Coherence

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

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

Coherence

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

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

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 $\mathbf{a}_1, \dots, \mathbf{a}_N$, 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

Summary

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

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

Summary

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

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

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

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).