Compressive Sensing: a crash introduction

Silvia Gazzola

BASP Group
Heriot-Watt University, Edinburgh, UK

ISSS society monthly seminar
November 26, 2015

A Classic in Signal Processing

The Nyquist-Shannon Theorem

"to exactly reconstruct an arbitrary band-limited signal
from its samples, the sampling rate needs to be
at least twice the bandwidth"

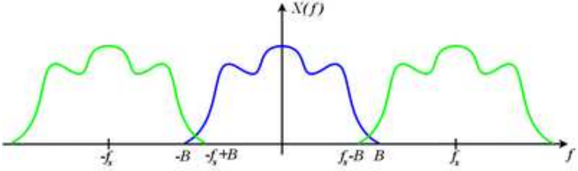
A Classic in Signal Processing

The Nyquist-Shannon Theorem

"to exactly reconstruct an arbitrary band-limited signal from its samples, the sampling rate needs to be at least twice the bandwidth"

Otherwise...aliasing effects!

[Image courtesy: Wikipedia]



A look at the what really happens

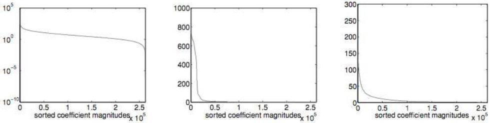
A look at the what really happens

[Image courtesy: Han et al. *Compressive Sensing for Wireless Networks*, 2013]

(a) DCT coefficients

(b) Haar wavelet coefficients

(c) Local variation



(d) DCT coeff's decay

(e) Haar wavelet coeff's decay

(f) Local variation decay

Sparsity

Sparsity

Some definitions:

Sparsity

Some definitions:

- $x \in \mathbb{C}^N$

Sparsity

Some definitions:

- $x \in \mathbb{C}^N$
- $\text{supp}(x) = \{ j : x_j \neq 0 \}$

Sparsity

Some definitions:

- $x \in \mathbb{C}^N$
- $\mathrm{supp}(x) = \{ j \, : \, x_j \neq 0 \}$
- $\|x\|_0 = |\mathrm{supp}(x)|$

Sparsity

Some definitions:

- $x \in \mathbb{C}^N$
- $\text{supp}(x) = \{ j : x_j \neq 0 \}$
- $\|x\|_0 = |\text{supp}(x)|$
- $\|x\|_1 = \sum_{i=1}^{N} |x_i|$

Sparsity

Some definitions:

- $x \in \mathbb{C}^N$
- $\text{supp}(x) = \{ j \,:\, x_j \neq 0 \}$
- $\|x\|_0 = |\text{supp}(x)|$
- $\|x\|_1 = \sum_{i=1}^{N} |x_i|$
- x sparse if $\|x\|_0 \ll N$

Sparsity

Some definitions:

- $x \in \mathbb{C}^N$
- $\mathrm{supp}(x) = \{ j \,:\, x_j \neq 0 \}$
- $\|x\|_0 = |\mathrm{supp}(x)|$
- $\|x\|_1 = \sum_{i=1}^{N} |x_i|$
- x sparse if $\|x\|_0 \ll N$
- x k-sparse if $\|x\|_0 \leq k$

Sparsity

Some definitions:

- $x \in \mathbb{C}^N$
- $\text{supp}(x) = \{\, j \,:\, x_j \neq 0 \,\}$
- $\|x\|_0 = |\text{supp}(x)|$
- $\|x\|_1 = \sum_{i=1}^N |x_i|$
- x sparse if $\|x\|_0 \ll N$
- x k-sparse if $\|x\|_0 \leq k$
- $\Sigma_k = \{\, x \in \mathbb{C}^N \,:\, \|x\|_0 \leq k \,\}$

Sparsity

Some definitions:

- $x \in \mathbb{C}^N$
- $\text{supp}(x) = \{ j \,:\, x_j \neq 0 \}$
- $\|x\|_0 = |\text{supp}(x)|$
- $\|x\|_1 = \sum_{i=1}^N |x_i|$
- x sparse if $\|x\|_0 \ll N$
- x k-sparse if $\|x\|_0 \leq k$
- $\Sigma_k = \{ x \in \mathbb{C}^N \,:\, \|x\|_0 \leq k \}$
- best k-sparse approximation: $x_{[k]} = \arg\min_{z \in \Sigma_k} \|x - z\|_p$

Sparsity

Some definitions:

- $x \in \mathbb{C}^N$
- $\text{supp}(x) = \{ j : x_j \neq 0 \}$
- $\|x\|_0 = |\text{supp}(x)|$
- $\|x\|_1 = \sum_{i=1}^{N} |x_i|$
- x sparse if $\|x\|_0 \ll N$
- x k-sparse if $\|x\|_0 \leq k$
- $\Sigma_k = \{ x \in \mathbb{C}^N : \|x\|_0 \leq k \}$
- best k-sparse approximation: $x_{[k]} = \arg\min_{z \in \Sigma_k} \|x - z\|_p$
- best k-term approximation error $\sigma_k(x)_p = \inf_{z \in \Sigma_k} \|x - z\|_p$

Sparsity

Some definitions:

- $x \in \mathbb{C}^N$
- $\operatorname{supp}(x) = \{ j \, : \, x_j \neq 0 \}$
- $\|x\|_0 = |\operatorname{supp}(x)|$
- $\|x\|_1 = \sum_{i=1}^{N} |x_i|$
- x sparse if $\|x\|_0 \ll N$
- x k-sparse if $\|x\|_0 \leq k$
- $\Sigma_k = \{ x \in \mathbb{C}^N \, : \, \|x\|_0 \leq k \}$
- best k-sparse approximation: $x_{[k]} = \arg\min_{z \in \Sigma_k} \|x - z\|_p$
- best k-term approximation error $\sigma_k(x)_p = \inf_{z \in \Sigma_k} \|x - z\|_p$
- x compressible if $\sigma_k(x)_p$ quickly decays

Sparsity

Some definitions:

- $x \in \mathbb{C}^N$
- $\mathrm{supp}(x) = \{ j \,:\, x_j \neq 0 \}$
- $\|x\|_0 = |\mathrm{supp}(x)|$
- $\|x\|_1 = \sum_{i=1}^{N} |x_i|$
- x sparse if $\|x\|_0 \ll N$
- x k-sparse if $\|x\|_0 \leq k$
- $\Sigma_k = \{ x \in \mathbb{C}^N \,:\, \|x\|_0 \leq k \}$
- best k-sparse approximation: $x_{[k]} = \arg\min_{z \in \Sigma_k} \|x - z\|_p$
- best k-term approximation error $\sigma_k(x)_p = \inf_{z \in \Sigma_k} \|x - z\|_p$
- x compressible if $\sigma_k(x)_p$ quickly decays

Most often: sparsity in a basis (dictionary) or under a transform

$$x = \Psi u \quad \text{or, equivalently,} \quad x = \sum_{i=1}^{N} \psi_i u_i$$

Sparsity

Some definitions:

- $x \in \mathbb{C}^N$
- $\mathrm{supp}(x) = \{ j \,:\, x_j \neq 0 \}$
- $\|x\|_0 = |\mathrm{supp}(x)|$
- $\|x\|_1 = \sum_{i=1}^{N} |x_i|$
- x sparse if $\|x\|_0 \ll N$
- x k-sparse if $\|x\|_0 \leq k$
- $\Sigma_k = \{ x \in \mathbb{C}^N \,:\, \|x\|_0 \leq k \}$
- best k-sparse approximation: $x_{[k]} = \arg\min_{z \in \Sigma_k} \|x - z\|_p$
- best k-term approximation error $\sigma_k(x)_p = \inf_{z \in \Sigma_k} \|x - z\|_p$
- x compressible if $\sigma_k(x)_p$ quickly decays

Most often: sparsity in a basis (dictionary) or under a transform

$$x = \Psi u \quad \text{or, equivalently,} \quad x = \sum_{i=1}^{N} \psi_i u_i$$

Examples: wavelets, DCT, curvelets, shearlets, (generalized) Total Variation

Compressive Sensing: an innovative sensing paradigm

D.Donoho, *Compressed sensing*, IEEE Trans. on Inf.Theory, 2006
E.Candes, J.Romberg, and T.Tao, *Robust uncertainty principles*, IEEE Trans. on Information Theory, 2006

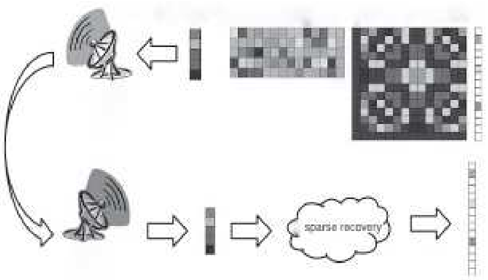
Compressive Sensing: an innovative sensing paradigm

D.Donoho, *Compressed sensing*, IEEE Trans. on Inf.Theory, 2006
E.Candes, J.Romberg, and T.Tao, *Robust uncertainty principles*, IEEE Trans. on Information Theory, 2006

The ingredients of Compressive Sensing (CS):

- signal sparse representation (sparsity basis);
- linear encoding and measurement collection (coherence and randomness);
- non-linear decoding (sparse recovery).

[Image courtesy: Han et al. *Compressive Sensing for Wireless Networks*, 2013]



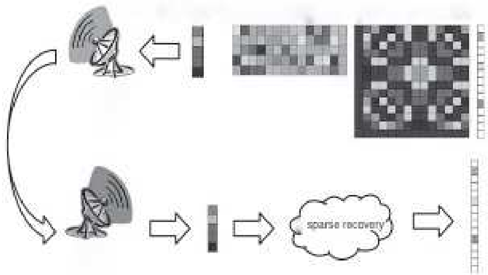
Compressive Sensing: an innovative sensing paradigm

D.Donoho, *Compressed sensing*, IEEE Trans. on Inf.Theory, 2006
E.Candes, J.Romberg, and T.Tao, *Robust uncertainty principles*, IEEE Trans. on Information Theory, 2006

The ingredients of Compressive Sensing (CS):

- signal sparse representation (sparsity basis);
- linear encoding and measurement collection (coherence and randomness);
- non-linear decoding (sparse recovery).

[Image courtesy: Han et al. *Compressive Sensing for Wireless Networks*, 2013]



CS senses less and computes more

Formulating the problem

Formulating the problem

Natural formulation:

$$x = \arg\min_{z \in \mathbb{C}^N} \|z\|_0 \quad \text{subject to} \quad Az = y$$

Formulating the problem

Natural formulation:

$$x = \arg\min_{z \in \mathbb{C}^N} \|z\|_0 \quad \text{subject to} \quad Az = y \quad \rightarrow \text{NP hard!}$$

Formulating the problem

Natural formulation:

$$x = \arg\min_{z \in \mathbb{C}^N} \|z\|_0 \quad \text{subject to} \quad Az = y \quad \rightarrow \text{NP hard!}$$

Convex relaxation:

$$x = \arg\min_{z \in \mathbb{C}^N} \|z\|_1 \quad \text{subject to} \quad Az = y$$

Formulating the problem

Natural formulation:

$$x = \arg\min_{z \in \mathbb{C}^N} \|z\|_0 \quad \text{subject to} \quad Az = y \quad \rightarrow \text{NP hard!}$$

Convex relaxation:

$$x = \arg\min_{z \in \mathbb{C}^N} \|z\|_1 \quad \text{subject to} \quad Az = y$$

Sensing matrix: $A \in \mathbb{C}^{m \times N}$, $m \ll N$.

Formulating the problem

Natural formulation:

$$x = \arg\min_{z \in \mathbb{C}^N} \|z\|_0 \quad \text{subject to} \quad Az = y \quad \rightarrow \text{NP hard!}$$

Convex relaxation:

$$x = \arg\min_{z \in \mathbb{C}^N} \|z\|_1 \quad \text{subject to} \quad Az = y$$

Sensing matrix: $A \in \mathbb{C}^{m \times N}$, $m \ll N$.
When using a dictionary: $A = \Phi\Psi$, $\Phi \in \mathbb{C}^{m \times N}$, $\Psi \in \mathbb{C}^{N \times N}$

Formulating the problem

Natural formulation:

$$x = \arg\min_{z \in \mathbb{C}^N} \|z\|_0 \quad \text{subject to} \quad Az = y \quad \to \text{NP hard!}$$

Convex relaxation:

$$x = \arg\min_{z \in \mathbb{C}^N} \|z\|_1 \quad \text{subject to} \quad Az = y$$

Sensing matrix: $A \in \mathbb{C}^{m \times N}$, $m \ll N$.
When using a dictionary: $A = \Phi\Psi$, $\Phi \in \mathbb{C}^{m \times N}$, $\Psi \in \mathbb{C}^{N \times N}$

$$\min_{u \in \mathbb{C}^N} \|u\|_1 \qquad \text{subject to} \quad Au = y$$

Formulating the problem

Natural formulation:

$$x = \arg\min_{z \in \mathbb{C}^N} \|z\|_0 \quad \text{subject to} \quad Az = y \quad \rightarrow \text{NP hard!}$$

Convex relaxation:

$$x = \arg\min_{z \in \mathbb{C}^N} \|z\|_1 \quad \text{subject to} \quad Az = y$$

Sensing matrix: $A \in \mathbb{C}^{m \times N}$, $m \ll N$.
When using a dictionary: $A = \Phi\Psi$, $\Phi \in \mathbb{C}^{m \times N}$, $\Psi \in \mathbb{C}^{N \times N}$

$$\min_{u \in \mathbb{C}^N} \|u\|_1 \qquad \text{subject to} \quad Au = y \quad \text{SYNTHESIS}$$

Formulating the problem

Natural formulation:

$$x = \arg\min_{z \in \mathbb{C}^N} \|z\|_0 \quad \text{subject to} \quad Az = y \quad \to \text{NP hard!}$$

Convex relaxation:

$$x = \arg\min_{z \in \mathbb{C}^N} \|z\|_1 \quad \text{subject to} \quad Az = y$$

Sensing matrix: $A \in \mathbb{C}^{m \times N}$, $m \ll N$.
When using a dictionary: $A = \Phi\Psi$, $\Phi \in \mathbb{C}^{m \times N}$, $\Psi \in \mathbb{C}^{N \times N}$

$$\min_{u \in \mathbb{C}^N} \|u\|_1 \quad \text{subject to} \quad Au = y \quad \text{SYNTHESIS}$$
$$\min_{x \in \mathbb{C}^N} \|\Psi^T x\|_1 \quad \text{subject to} \quad \Phi x = y$$

Natural formulation:

$$x = \arg\min_{z \in \mathbb{C}^N} \|z\|_0 \quad \text{subject to} \quad Az = y \quad \to \text{NP hard!}$$

Convex relaxation:

$$x = \arg\min_{z \in \mathbb{C}^N} \|z\|_1 \quad \text{subject to} \quad Az = y$$

Sensing matrix: $A \in \mathbb{C}^{m \times N}$, $m \ll N$.
When using a dictionary: $A = \Phi\Psi$, $\Phi \in \mathbb{C}^{m \times N}$, $\Psi \in \mathbb{C}^{N \times N}$

$$\min_{u \in \mathbb{C}^N} \|u\|_1 \quad \text{subject to} \quad Au = y \quad \text{SYNTHESIS}$$
$$\min_{x \in \mathbb{C}^N} \|\Psi^T x\|_1 \quad \text{subject to} \quad \Phi x = y \quad \text{ANALYSIS}$$

Formulating the problem

Natural formulation:

$$x = \arg\min_{z \in \mathbb{C}^N} \|z\|_0 \quad \text{subject to} \quad Az = y \quad \rightarrow \text{NP hard!}$$

Convex relaxation:

$$x = \arg\min_{z \in \mathbb{C}^N} \|z\|_1 \quad \text{subject to} \quad Az = y$$

Sensing matrix: $A \in \mathbb{C}^{m \times N}$, $m \ll N$.
When using a dictionary: $A = \Phi\Psi$, $\Phi \in \mathbb{C}^{m \times N}$, $\Psi \in \mathbb{C}^{N \times N}$

$$\min_{u \in \mathbb{C}^N} \|u\|_1 \quad \text{subject to} \quad Au = y \quad \text{SYNTHESIS}$$
$$\min_{x \in \mathbb{C}^N} \|\Psi^T x\|_1 \quad \text{subject to} \quad \Phi x = y \quad \text{ANALYSIS}$$

Moreover, noise: $y = Ax + w$, $\|w\|_2 \leq \varepsilon$.

A geometrical glimpse into the problem

A geometrical glimpse into the problem

$A \in \mathbb{R}^{1\times 2}$

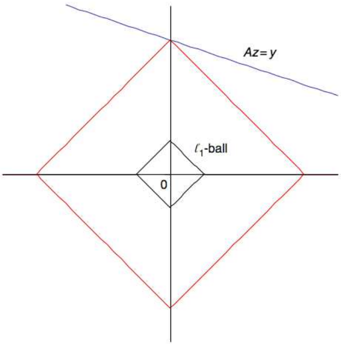
$\|z\|_1 = |z_1| + |z_2|$

A geometrical glimpse into the problem

$A \in \mathbb{R}^{1\times 2}$

$\|z\|_1 = |z_1| + |z_2|$

[Image courtesy: Fornasier and Rauhut. *Compressive Sensing*, 2011]

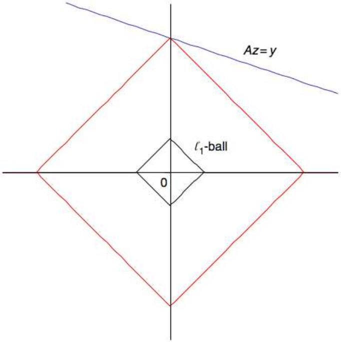


A geometrical glimpse into the problem

$A \in \mathbb{R}^{1\times 2}$

$\|z\|_1 = |z_1| + |z_2|$

[Image courtesy: Fornasier and Rauhut. *Compressive Sensing*, 2011]



the sparsest solution is recovered!

Key question:
What A allows the recovery of x from $y = Az$ given that x is sparse?

The mathematical theory

Key question:
What A allows the recovery of x from $y = Az$ given that x is sparse?

Some more definitions:

The mathematical theory

Key question:

What A allows the recovery of x from $y = Az$ given that x is sparse?

Some more definitions:

- Null space property (NSP)

Key question:

What A allows the recovery of x from $y = Az$ given that x is sparse?

Some more definitions:

- Null space property (NSP)

  $A \in \mathbb{C}^{m \times N}$ satisfies the NSP of order k with constant $\gamma_k \in (0, 1)$ if

$$\|\eta_T\|_1 \leq \gamma_k \|\eta_{T^c}\|_1$$

  for all $\eta \in \ker(A)$, for all $|T| \leq k$.

Key question:
What A allows the recovery of x from $y = Az$ given that x is sparse?

Some more definitions:

- Null space property (NSP)
  $A \in \mathbb{C}^{m \times N}$ satisfies the NSP of order k with constant $\gamma_k \in (0, 1)$ if

$$\|\eta_T\|_1 \leq \gamma_k \|\eta_{T^c}\|_1$$

  for all $\eta \in \ker(A)$, for all $|T| \leq k$.

- Restricted Isometry property (RIP)

The mathematical theory

Key question:
What A allows the recovery of x from $y = Az$ given that x is sparse?

Some more definitions:

- Null space property (NSP)
  $A \in \mathbb{C}^{m \times N}$ satisfies the NSP of order k with constant $\gamma_k \in (0, 1)$ if

$$\|\eta_T\|_1 \leq \gamma_k \|\eta_{T^c}\|_1$$

  for all $\eta \in \ker(A)$, for all $|T| \leq k$.

- Restricted Isometry property (RIP)
  $A \in \mathbb{C}^{m \times N}$ satisfies the RIP of order k with constant $\delta_k \in (0, 1)$ if

$$(1 - \delta_k)\|z\|_2^2 \leq \|Az\|_2^2 \leq (1 + \delta_k)\|z\|_2^2$$

  for all $z \in \Sigma_k$.

The mathematical theory

Key question:
What A allows the recovery of x from $y = Az$ given that x is sparse?

Some more definitions:

- Null space property (NSP)
  $A \in \mathbb{C}^{m \times N}$ satisfies the NSP of order k with constant $\gamma_k \in (0, 1)$ if

$$\|\eta_T\|_1 \leq \gamma_k \|\eta_{T^c}\|_1$$

  for all $\eta \in \ker(A)$, for all $|T| \leq k$.

- Restricted Isometry property (RIP)
  $A \in \mathbb{C}^{m \times N}$ satisfies the RIP of order k with constant $\delta_k \in (0, 1)$ if

$$(1 - \delta_k)\|z\|_2^2 \leq \|Az\|_2^2 \leq (1 + \delta_k)\|z\|_2^2$$

  for all $z \in \Sigma_k$.

Examples: Gaussian, Bernoulli, randomized orthonormal systems

Conditions for recovery

Conditions for recovery

- Theorem (exploits NSP)

Conditions for recovery

- Theorem (exploits NSP)
  Assumptions:
  - $A \in \mathbb{C}^{m \times N}$ satisfies the k-NSP with constant $\gamma_k \in (0, 1)$
  - $x^* = \arg\min_{z \in \mathbb{C}^N} \|z\|_1 \quad$ subject to $\quad Az = y$

  Then:

$$\|x - x^*\|_1 \leq \frac{2(1 + \gamma_k)}{1 - \gamma_k} \sigma_k(x)_1$$

Conditions for recovery

- Theorem (exploits NSP)

  Assumptions:

  - $A \in \mathbb{C}^{m \times N}$ satisfies the k-NSP with constant $\gamma_k \in (0, 1)$
  - $x^* = \arg\min_{z \in \mathbb{C}^N} \|z\|_1 \quad$ subject to $\quad Az = y$

  Then:

$$\|x - x^*\|_1 \leq \frac{2(1 + \gamma_k)}{1 - \gamma_k} \sigma_k(x)_1$$

- Theorem (exploits RIP)

Conditions for recovery

- Theorem (exploits NSP)

  Assumptions:

  - $A \in \mathbb{C}^{m \times N}$ satisfies the k-NSP with constant $\gamma_k \in (0, 1)$
  - $x^* = \arg\min_{z \in \mathbb{C}^N} \|z\|_1$ subject to $Az = y$

  Then:

$$\|x - x^*\|_1 \leq \frac{2(1+\gamma_k)}{1-\gamma_k} \sigma_k(x)_1$$

- Theorem (exploits RIP)

  Assumptions:

  - $A \in \mathbb{C}^{m \times N}$ satisfies the $2k$-RIP with constant $\delta_{2k} < 0.4931$
  - $x^* = \arg\min_{z \in \mathbb{C}^N} \|z\|_1$ subject to $\|Az - y\| \leq \varepsilon$

  Then:

$$\|x^* - x\|_2 \leq C_1 \varepsilon + C_2 \frac{\sigma_k(x)_1}{\sqrt{k}}$$

This is just the beginning of the story...

- design of sensing matrices
- applications
- optimization theory
- algorithmic developments
- high-performance computing

If you are thrilled about CS (or even not...)
please come on December 17, in the afternoon (tbc)!

- M. Fornasier and H. Rauhut
  Compressive Sensing in Handbook of Math. Meth. in Imaging
  pp. 187-228, Springer, 2011.
- Z. Han, H. Li, and W. Yin
  Compressive Sensing for Wireless Networks
  Cambridge, 2013.

If you are thrilled about CS (or even not...)
please come on December 17, in the afternoon (tbc)!