Compressive Sensing (Sampling)

Mahdy Nabaee

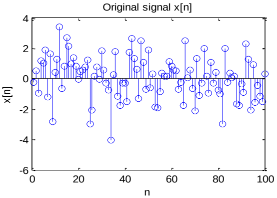
December 4, 2011

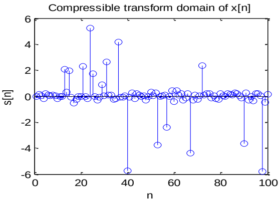
Consider an n-dim (signal) vector, $\underline{x} = [x_1, \cdots, x_i, \cdots, x_n]^T \in \mathbf{R}^n$:

$$\underline{x}_{n\times 1} = \Phi_{n\times n} \cdot \underline{s}_{n\times 1},$$

where:

- $\Phi = [\underline{\Phi}_1, \cdots, \underline{\Phi}_n]$,
- $\{\underline{\Phi}_i\}$ is an orthonormal *representation basis* for $\mathbf{R}^n$.



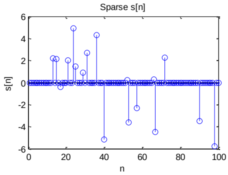


Sparse/Compressible Signal

Sparsity: $\underline{x}$ is K-sparse in some transform domain, if:

$$\|\underline{s}\|_{\ell_0} \leq K,$$

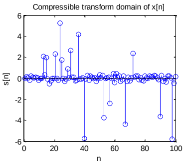
for all realizations.



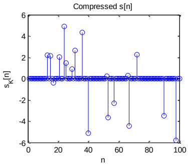
Compressibility: $\underline{x}$ is compressible, if:

$$\|\underline{x} - \underline{x}_K\|_{\ell_2} = \|\underline{s} - \underline{s}_K\|_{\ell_2},$$

is small, for all realizations.



$\Longrightarrow$



Transform Coding

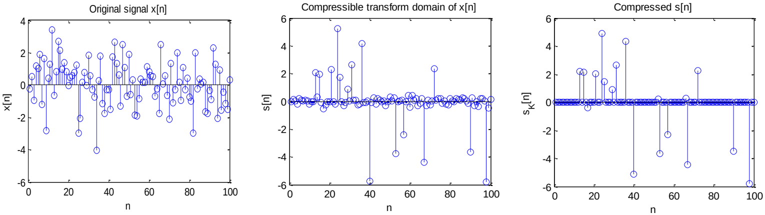
Encoding:

- Upon sensing all x_i's, the sparse/compressible transform domain, $\underline{s}$, is calculated:

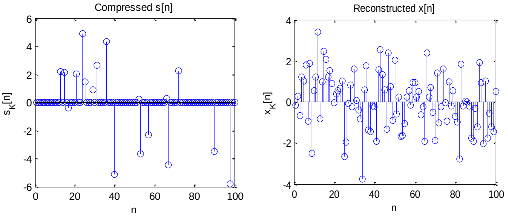
$$\underline{s} = \Phi^T \cdot \underline{x},$$

- Then, K largest coefficients, s_i, are transmitted, along their locations (corresponding i's):

$$\underline{s}_K \ .$$



Transform Coding



Decoding:

- $\underline{s}_K$ is reconstructed from the location of non-zero coefficients and their values,
- Then, the original signal can be reconstructed without/with loss, according to:

$$\underline{x}_K = \Phi \cdot \underline{s}_K \ .$$

Sensing + Compression

- Obviously, using only Shannon sampling is not efficient to store/transmit sparse/compressible signals:
  - since the signal may have large components in high frequencies,
  - while it has few non-zero frequency (transform domain) components.
- Shannon sampling + transform coding:
  - need to sense and store all the original n samples (large n),
  - need to calculate all the coefficients, s_i's, which may be computationally complex,
  - need to store/transmit the location of non-zero/large coefficients, s_i's, for sparse/compressible signals.

Compressive Sensing Problem

As a combination of sampling and transform coding, compressive sensing problem is defined as in below:

Definition: Compressive Sampling (Sensing) problem

For the measurement equation:

$$\underline{y}_{m\times 1} = \Psi_{m\times n} \cdot \underline{x}_{n\times 1} = \Psi\Phi \cdot \underline{s} = \Theta_{m\times n} \cdot \underline{s},$$

where $m < n$ (usually $m \ll n$), find:

- a stable measurement matrix, Ψ, and,
- a reconstruction algorithm to find $\hat{\underline{s}}$ (or $\hat{\underline{x}}$),

for sparse/compressible signals.

Under-determined Set of Linear Measurements

- Generally,

$$\underline{y}_{m\times 1} = \Theta_{m\times n} \cdot \underline{s}_{n\times 1}, \quad m < n,$$

  is ill-conditioned and there are infinite number of solutions:

$$\hat{\underline{s}} = \underline{s} + \underline{s}_0,$$

  where $\underline{s}_0$ is in the null-space of $\Theta_{m\times n}$.

- However, if $\underline{s}$ is K-sparse, and location of non-zero elements is known, there exists only a single unique solution if:

$$m \geq K.$$

Under-determined Set of Linear Measurements

- For this simplified version (known locations), the necessary and sufficient condition is:

$$1 - \delta_K \leq \frac{\|\Theta \cdot \underline{v}\|_{\ell_2}}{\|\underline{v}\|_{\ell_2}} \leq 1 + \delta_K,$$

  for any $\underline{v} \in \mathbf{R}^n$ which share the same K non-zero elements and small δ_K.

Under-determined Set of Linear Measurements

- For this simplified version (known locations), the necessary and sufficient condition is:

$$1 - \delta_K \leq \frac{\|\Theta \cdot \underline{v}\|_{\ell_2}}{\|\underline{v}\|_{\ell_2}} \leq 1 + \delta_K,$$

for any $\underline{v} \in \mathbf{R}^n$ which share the same K non-zero elements and small δ_K.

Definition: Restricted Isometry Property (RIP)

A matrix, $\Theta_{m\times n}$, is called to satisfy RIP of order K, if there exists a small $0 \leq \delta_K < 1$ such that:

$$1 - \delta_K \leq \frac{\|\Theta \cdot \underline{v}\|_{\ell_2}}{\|\underline{v}\|_{\ell_2}} \leq 1 + \delta_K,$$

for any K-sparse vector $\underline{v} \in \mathbf{R}^n$.

More on RIP

- For both K-sparse and compressible signals, to have stable Θ (and correspondingly Ψ), it is sufficient that Θ satisfy RIP of order $3K$.
- Moreover, Ψ is universal in the sense that RIP of Θ does not depend on the choice of Φ.

Choices of Matrices which satisfy RIP with high probability:

- random matrix with iid zero-mean, $\frac{1}{n}$ variance Gaussian entries, if $m \geq cK \log(\frac{n}{K})$, for small constant c.
- random matrix with iid Bernouli half entries.

Sparse Signal Reconstruction: ℓ_0 Minimization

We need to find $\underline{\hat{s}} \in \mathbf{R}^n$ such that:

- $\Theta \underline{\hat{s}} = \underline{y}$,
- $\underline{\hat{s}}$ is the sparsest solution.

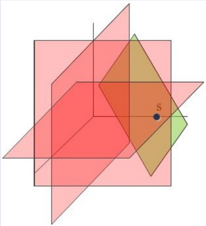
Sparse Signal Reconstruction: ℓ_0 Minimization

We need to find $\hat{\underline{s}} \in \mathbf{R}^n$ such that:

- $\Theta\hat{\underline{s}} = \underline{y}$,
- $\hat{\underline{s}}$ is the sparsest solution.

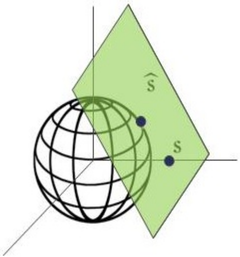
$$\hat{\underline{s}} = \arg\min_{\underline{s}'} \|\underline{s}'\|_{\ell_0}, \quad s.t.\ \Theta\underline{s}' = \underline{y}$$

suffers from computational complexity



Sparse Signal Reconstruction: Other Norms

ℓ_2 Minimization

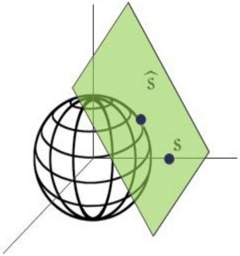


$$\underline{\hat{s}} = \arg\min_{\underline{s}'} \|\underline{s}'\|_{\ell_2}, \quad s.t.\ \Theta \cdot \underline{s}' = \underline{y}$$

- Closed Form: $\underline{\hat{s}} = \Theta^T (\Theta\Theta^T)^{-1} \cdot \underline{y}$,
- The resulting $\underline{\hat{s}}$ is almost always not sparse.

Sparse Signal Reconstruction: Other Norms

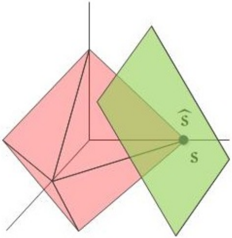
ℓ_2 Minimization



$$\underline{\hat{s}} = \arg\min_{\underline{s}'} \|\underline{s}'\|_{\ell_2}, \quad s.t.\ \Theta \cdot \underline{s}' = \underline{y}$$

- Closed Form: $\underline{\hat{s}} = \Theta^T(\Theta\Theta^T)^{-1} \cdot \underline{y}$,
- The resulting $\underline{\hat{s}}$ is almost always not sparse.

ℓ_1 Minimization



$$\underline{\hat{s}} = \arg\min_{\underline{s}'} \|\underline{s}'\|_{\ell_1}, \quad s.t.\ \Theta \cdot \underline{s}' = \underline{y}$$

- sparse solution with high probability,
- solved using linear programming.

Noisy Compressive Sensing

Consider the case of noisy measurements:

$$\underline{y}_{m\times 1} = \Theta_{m\times n} \cdot \underline{s}_{n\times 1} + \underline{w}_{m\times 1}, \quad \mathbf{E}[\underline{w}\ \underline{w}^T] = \sigma^2 \mathbf{I}_m$$

Noisy Compressive Sensing

Consider the case of noisy measurements:

$$\underline{y}_{m\times 1} = \Theta_{m\times n} \cdot \underline{s}_{n\times 1} + \underline{w}_{m\times 1}, \quad \mathbf{E}[\underline{w}\ \underline{w}^T] = \sigma^2 \mathbf{I}_m$$

Basic Pursuit

$$\arg\min_{\underline{s}'} \|\underline{s}'\|_{\ell_1}, \quad s.t.\ \|\underline{y} - \Theta \underline{s}'\|_{\ell_2} \leq \epsilon$$

where ϵ bounds the measurement noise, $\underline{w}$.

Noisy Compressive Sensing

Consider the case of noisy measurements:

$$\underline{y}_{m\times 1} = \Theta_{m\times n} \cdot \underline{s}_{n\times 1} + \underline{w}_{m\times 1}, \quad \mathbf{E}[\underline{w}\,\underline{w}^T] = \sigma^2 \mathbf{I}_m$$

Basic Pursuit

$$\arg\min_{\underline{s}'} \|\underline{s}'\|_{\ell_1}, \quad s.t.\ \|\underline{y} - \Theta \underline{s}'\|_{\ell_2} \leq \epsilon$$

where ϵ bounds the measurement noise, $\underline{w}$.

Lasso De-Noising

$$\arg\min_{\underline{s}'} \|\underline{y} - \Theta \underline{s}'\|_{\ell_2}, \quad s.t.\ \|\underline{s}'\|_{\ell_1} \leq q$$

where q ranges from zero to infinity.

Of the very first and major applications of CS: