Compressed sensing

Deterministic algorithm of sensing matrix generation

Agnieszka Lisowska[1] Jolanta Socala[1]

[1]University of Silesia
Institute of Computer Science
ul. Bedzinska 39, 41-200 Sosnowiec, Poland

Brenna, 2018

Outline

1. Preliminaries
   - History
   - Basic definitions
   - Motivation
2. Compressed sensing at glance
   - Sensing matrices
   - Signal recovery algorithms
3. The proposed algorithm
   - The algorithm of sensing matrix generation
   - Experimental results
4. Conclusions

History of signal representations

- Fourier Transform (localization), 1D and 2D signals [Fourier, 1822]
- Wavelet Transform, basis (localization, scale), 1D and 2D signals [Haar, 1910]
- Curvelet Transform and others, frames (localization, scale, orientation), only 2D signals [Candes, Donoho, 1999]
- Wedgelet<Smoothlet> Transform, dictionaries (localization, scale, orientation, <smoothness>), only 2D signals [Donoho, 1999]

Shannon-Nyquist sampling theorem

- The theorem fixes the limit for convertion of analog to digital signal when one wants to reconstruct the exact signal ;
- Theorem : Suppose the highest frequency component (in hertz) for a given analog signal is f_{max}. Then the sampling rate must be at least $2f_{max}$ in order to reconstruct the signal without loss.

Sensing sparse signals

- Let
$$y = Ax,$$
where $x \in \mathbb{R}^n$ is an original signal, A is an $m \times n$ matrix and $y \in \mathbb{R}^m$. We assume m is much smaller than n. The matrix A is called the *sensing matrix*.
- Vector x of length n is k-*sparse* if only k elements are non-zero ($k = ||x||_0$), $k < n$.
- Vector x is *compressible* if only a small number of elements are significantly non-zero.

Classical viewpoint

- Full signal acquisition, then its compression - large fraction of coefficients is discarded :
  measure all coefficients (pixels), transform them to obtain the sparse signal, keep k largest coefficients ;

Fundamental questions

- Why go to so much effort to acquire all the data when most of what we get will be thrown away ? [Donoho]
- Fundamental question : Can we directly acquire just the useful part of the signal ?
- Answer : Yes, we can achive this by random sensing mechanism.

Compressed sensing viewpoint

- Simultaneous signal acquisition and compression - achived by random sensing mechanism :
  take m random measurements $y_i = \langle x, a_i \rangle$, reconstruct the signal by linear programming.
- Important ! : The sparsity of a signal can be exploited to recover it from far fewer samples than required by the Shannon-Nyquist sampling theorem.

Computational problems

- There are two main problems in compressed sensing which have to be addressed :
  - How can we construct a proper sensing matrix ?
  - How can we recover the original signal ?
- In this work we address the first problem and show the method of sensing matrix generation. Even though the matrix is generated in deterministic way, it has properties typical for random matrices.

Compressed sensing problem

- We deal with the undetermined problem $y = Ax$.
- To have a unique solution we need to add a prior : the solution has to be sparse.

Spark of the matrix

- The *spark* of matrix A is the smallest number of linearly dependent columns.
- Spark is used for definition of the existence of the sparsest solution of the minimization problem

$$\min ||x||_0 \quad subject\ \ to \quad y = Ax.$$

- Theorem : If $||x||_0 = k$ then if $k < 1/2\, spark\{A\}$ the solution is unique.
- However, the computation of $spark\{A\}$ is the NP-hard problem.

Coherence of the matrix

- The *coherence* of matrix A is defined as

$$\mu(A) = \max_{1 \le i < j \le n} \frac{|\langle a_i, a_j \rangle|}{||a_i||_2 ||a_j||_2},$$

where a_i and a_j are two columns of A. This is the degree of linear dependence of columns (e.g. for an orthogonal matrix coherence equals to 0).

Coherence versus spark

- By definition :

$$spark\{A\} = 1 + \frac{1}{\mu(A)}$$

- We are looking for matrices with the smallest cohenence possible in order to make the matrix A as close as possible to the unitary matrix.

Fundamental properties

- To recover the original signal some properties must be met.
- Null Space Property is a necessary condition, which guarantees that the algorithm can recover the k-sparse signal.
- NSP does not consider the noisy situation. If the signal is contaminated, we should consider stronger conditions.
- Restricted Isometry Property guarantees that the algorithm can recover the k-sparse signal even in the noisy situation.

Restricted Isometry Property

- Definition : A matrix A satisfies the RIP of order k if there exists $\delta_k \in (0,1)$ such that

$$(1-\delta_k)||x||_2^2 \leq ||Ax||_2^2 \leq (1+\delta_k)||x||_2^2$$

  holds for all k-sparse vectors x, where δ_k is called k-restricted isometry constant.
- The RIP condition provides the basic condition for the compressed sensing theory. It characterizes matrices which are nearly orthonormal, at least when operating on sparse vectors.

Recovery algorithms

- There are many recovery algorithms. They differ in speed of computation and recovered signal quality.
- There is no the best algorithm. Depending on applications one has to choose the appropriate one.

Looking for good matrices

- Compressed sensing matrices are usually generated in random way due to their good compression properties. They must have random properties.
- The generated matrix has to satisfy RIP. So, in order to get a good sensing matrice one may need to perform many random processes what takes too long time.

Why not to generate a deterministic matrix ?

- In the literature a few algorithms of deterministic generation of pseudorandom matrices were presented.
- In this work the deterministic generation of compressed sensing matrix is presented.
- As the experiments show, such obtained matrix has better properties than the random ones.

Sensing matrix details

- We are interested in generating sensing matrices with small coherence. The smaller the better.
- We generate matrix A of size $m \times n$ (with {1, -1} elements) such that $\mu(A) = 1/3$,

$$m = 3 \cdot 2^s, \qquad n = 2 \cdot 3^{s+1} - 3,$$

  where s is the number of steps in our algorithm.
- For example : number of steps = 5, n = 1455, m = 96.
- We can remove some columns of A and receive matrix A'. From the definition of μ we then have $\mu(A') \leq 1/3$.

Matrix generation algorithm

Input : number of steps
Initialize : $A = [[-1, 1, 1], [1, -1, 1], [1, 1, -1]]$,
Step : $A = [[m_0], [m_1], ..., [m_{3k+2}]]$, where $[m_0], ..., [m_{3k+2}]$ are columns of A; $k \in \mathbb{N}$

Algorithm

nextA $:=$ [ ] ; the empty matrix
for $i = 0$ to k do {
from the three columns $[m_{3i}], [m_{3i+1}], [m_{3i+2}]$ create
nine new columns of nextA as follows :
$[m_{3i}, m_{3i}],\ [m_{3i+1}, m_{3i+1}],\ [m_{3i+2}, m_{3i+2}],$
$[m_{3i}, m_{3i+1}],\ [m_{3i+1}, m_{3i+2}],\ [m_{3i+2}, m_{3i}],$
$[m_{3i}, m_{3i+2}],\ [m_{3i+1}, m_{3i}],\ [m_{3i+2}, m_{3i+1}]$; }
$A := nextA$;

Experimental results

size of matrix	the proposed matrix	20 random matrices		
		minimum	average	maximum
48×483	0.33	0.58	0.64	0.70
96×1455	0.33	0.46	0.50	0.54
192×4371	0.33	0.36	0.39	0.45

Table : Coherence values of the proposed matrix versus random matrices.

Conclusions

- For sparse images it is possible to recover them from far fewer samples than it is required by the Shannon-Nyquist sampling theorem.
- The sensing matrices must have random properties but can be generated in deterministic way. Such a way is faster and determined.
- The presented results are in the first stage. The image compression implementation is in progress...

References

Candes, Emmanuel J. ; Romberg, Justin K. ; Tao, Terence (2006). "Stable signal recovery from incomplete and inaccurate measurements". Communications on Pure and Applied Mathematics. 59(8) : 1207-1223.

Donoho, D.L. (2006). "Compressed sensing". IEEE Transactions on Information Theory. 52(4) : 1289-1306.