International Journal of Electrical and Electronics Engineers
Vol. No.7 Issue 02, July-December 2015
www.arresearchpublication.com
IJEEE ISSN 2321 - 2055

SPARSE APPROXIMATION MODELLING USING COMPRESSIVE SENSING

Saruti Gupta

School of Electronics, CDAC, Noida (India)

ABSTRACT

Compressed sensing (also known as compressive sensing, compressive sampling, or sparse sampling) is a signal processing technique for efficiently acquiring and reconstructing a signal, by finding solutions to underdetermined linear systems. Given signals of objects with sparse contents relative to its dimension, compressed sensing seeks to reconstruct the signals from as few non-adaptive linear measurements as possible. This paper deals with the CS paradigm exploring sparsity, incoherence, restricted isometric properties of the mathematical model. Using Sparse recovery Algorithm, Null Space Conditions and Thresholds for Rank Minimization a compressive sensing scheme with linear decoding complexity, deterministic performance guarantees of linear sparsity recovery, and explicitly constructible measurement matrices is analysed.

Keywords: Sparse Reconstruction, Compressed Sensing, Basis Pursuit (BP), Greedy Algorithms, Sensing Matrix, Incoherence, Signal Reconstruction

I. INTRODUCTION

The theory of compressive sensing was developed by Candes et al and Donoho in 2004. This method is different from traditional method as it sampled the signal below the Nyquist rate and it permits to exploit the sparse property at the signal acquisition stage of compression. Contrary to traditional Nyquist paradigm, the CS paradigm, banking on finding sparse solutions to underdetermined linear systems, can reconstruct the signals from far fewer samples than is possible using Nyquist sampling rate. In this method the signal is first transformed into a sparse domain and then the signal is reconstructed using numerical optimization technique using small number of linear measurements. In essence, CS combines the sampling and compression into one step by measuring minimum samples that contain maximum in- formation about the signal. This eliminates the need to acquire and store large number of samples only to drop most of them because of their minimal value. CS operates very differently, and performs as "if it were possible to directly acquire just the important information about the object of interest." By taking about O(Slog(n/S)) random projections as in "Random Sensing," one has enough information to reconstruct the signal with accuracy at least as good as that provided by f_S, the best S-term approximation—the best compressed representation—of the object.

This paper starts with the insight of the compressive sensing. Section II focuses on various parameters like sparsity, incoherence, analyzing the mathematical model, for the explicit constructions of sensing matrices and efficient decoding algorithms. Section III states the applications of compressive sensing and Section IV gives the conclusion for the article.

II. COMPRESSIVE SENSING

Compressive sensing theory asserts that we can recover certain signals from fewer samples than required in Nyquist paradigm. This recovery is exact if signal being sensed has a low information rate (means it is sparse in original or some transform domain). Number of samples needed for exact recovery depends on particular reconstruction algorithm being used. If signal is not sparse, then recovered signal is best reconstruction obtainable from s largest coefficients of signal. CS handles noise gracefully and reconstruction error is bounded for bounded perturbations in data. Various parameters related to the behavior of CS are analyzed as follows:

2.1 Sparsity

A signal which has only non-zero coefficients, is said to be s-Sparse. Vectors are often used to represent large amounts of data which can be difficult to store or transmit. CS theory relies first and foremost on the signal of interest, having a sparse representation in some basis,ψ=[ψ_1, $\psi_{2,\ldots\ldots}\psi_L$] such that f= ψx where is the coefficient vector f for under the basis ψ . For f to be sparse in ψ, the coefficients ,x_i , must be mostly zero or insignificant such that they can be discarded without any perceptual loss. If it has the most compact representation, then should be compressible if captured in some other basis. So sparseness also implies compressibility and vice versa. A familiar example of such a signal is a sine wave which requires many coefficients in time to represent, but requires only one non-zero coefficient in the Fourier domain[1].

By using a sparse approximation the amount of space needed to store the vector would be reduced to a fraction of what was originally needed. Sparse approximations can also be used to analyze data by showing how column vectors in a given basis come together to produce the data. There are many areas of science and technology which have greatly benefited from advances involving sparse approximations.

Compressive Sensing is based on the empirical observation that many types of real-world signals and images have a sparse expansion in terms of a suitable basis or frame, for instance a wavelet expansion. This means that the expansion has only a small number of significant terms, or in other words, that the coefficient vector can be well-approximated with one having only a small number of non vanishing entries. The support of a vector x is denoted $\mathrm{supp}(x) = \{j : x_j \neq 0\}$, and

$$\|x\|_0 := |\mathrm{supp}(x)|. \quad (1)$$

A vector x is called k-sparse if $\|x\|_0 \leq k$. For $k \in \{1,2,...,N\}$,

$$\Sigma_k := \{x \in C^N : \|x\|_0 \leq k\} \quad (2)$$

denotes the set of k-sparse vectors. Furthermore, the best k-term approximation error of a vector $x \in C^N$ in ℓp is defined as

$$\sigma_k(x)_p = \inf_{z\in\Sigma}\|x-z\|_p. \quad (3)$$

If $\sigma_k(x)$ decays quickly in k then x is called compressible. Indeed, in order to compress x one may simply store only the k largest entries. When reconstructing x from its compressed version the non stored entries are simply set to zero, and the reconstruction error is $\sigma_k(x)_p$. It is emphasized at this point that the procedure of obtaining the compressed version of x is adaptive and nonlinear since it requires the search of the largest entries of x in absolute value. In particular, the location of the non-zeros is a nonlinear type of information.

2.2 Incoherence

Coherence measures the maximum correlation between any two elements of two different matrices. These two matrices might represent two different basis/representation domains.

If Ψ is a n×n matrix with $\Psi_{1,\ldots}$ Ψ_N, as columns and Φ is an

m × n matrix with $\Phi_{1\ldots\ldots}$ Φ_m as rows[2]. Then, coherence μ is defined as

$\mu(\Phi,\Psi)=$ $\max|\Phi_k,\Psi_j|$ (4)

for 1<=j<j and 1<=k<=m. It follows from linear algebra that

$1<\mu(\Phi,\Psi)<=$

2.3 Mathematical Model

Let us consider a real-valued, finite-length, one dimensional, discrete-time signal X, which we view as an N × 1 column vector with elements x[n] , n=1, 2, . . . , N.

Any signal can be represented in terms of a basis of N × 1 vectors $\{\Psi_i\}$, i = 1: N.

We assume basis is orthonormal. Forming N X N basis matrix $\Psi:=\Psi_{1,\ldots}$ Ψ_N We can express signal X as

$X=\Sigma^{N}_{I=1}\, a_i\, \Psi_i$ or $X=a\,\Psi$ (5)

Where a is N × 1 column vector where its obtained by keeping only the terms corresponding to the k largest values of X(i).

We focus on signals that have sparse representation, where X is a linear combination of just k basis vectors, with K << N. That is only with K non zero and (N-K) are zero. We call K sparse such object with most K non zero entries[3]. Let A denote M × N measurement matrix which is obtained with randn function to which K and N are inputs, A with vectors $\Phi_1^*\ldots\ldots\ \Phi_M^*$ as rows. Then observation vector Y is obtained with help of equation

$Y=A\ X\ a_0$ (6)

Then the reconstruction algorithm OMP is used to reconstruct the audio signal using observation vector Y, sparsity level k. Initially in reconstruction method assign residual with observation vector $R_o=y$

Then calculate the inner product with equation

$G_n=A^T r_{n-1}$ (7)

where A^T is transpose of Measurement matrix A, r_{n-1} is previous residual.

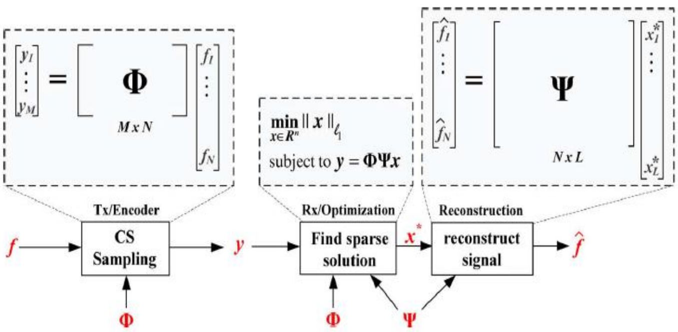

Figure 1.Compressive Sensing Block Diagram

Then find the index *K* that solves the optimization problem: K=arg max |gn[i]| i={1,2…n} Obtain a new signal estimate and calculate the new approximation of the new residual:

$xn[k]=x_{n-1}[k]+g_n[k]$ (8)

$r_n=r_{n-1} - g_n[k]A_K$ (9)

With this the reconstructed audio signal is obtained with same number of samples as that of input audio signal.

2.4 The Restricted Isometry Property

In linear algebra, an orthogonal matrix is a square matrix with real entries whose columns and rows are orthogonal unit vectors (*i.e.*, orthonormal vectors), *i.e.* $A^T A = A A^T = I$, where I is the identity matrix.

This leads to the equivalent characterization: a matrix A is orthogonal if its transpose is equal to its inverse; *i.e.*, $A^T = A^{-1}$. The determinant of any orthogonal matrix is either +1 or −1. As a linear transformation, an orthogonal matrix preserves the dot product of vectors, and therefore acts as an isometry of Euclidean space, such as a rotation or reflection[4].

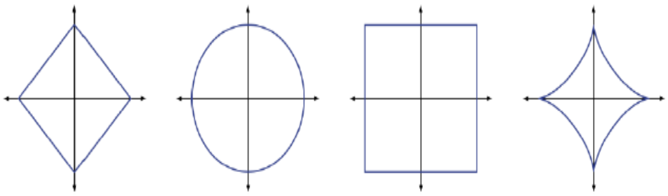

Figure 2. Unit spheres in R^2 for the lp norms with p =1;2;∞ , and for the l p quasi-norm with p = 0.5 . (a) Unit sphere for l1 norm; (b) Unit sphere for l2 norm; (c) Unit sphere for l∞ norm; (d) Unit sphere for l p quasi norm

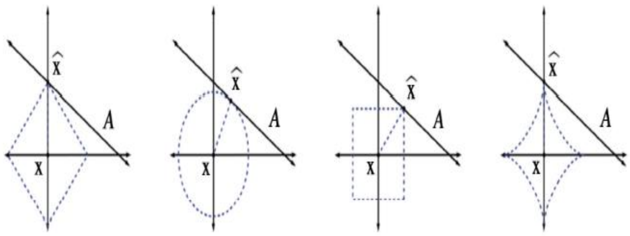

Figure 3. Best approximation of a point in R^2 by a one-dimensional subspace using l p norms with lp =1;2;∞ , and the l p quasi-norm with p = 0.5 (a) Approximation in l1 norm; (b) Approximation in l2 norm; (c) Approximation in l∞ norm; (d) Approximation in l p quasi-norm.

2.5 Explicit Constructions of Sensing Matrices

In technical literature, more attention has been paid to random sensing matrices generated by identical and independent distributions (i.i.d.) such as Gaussian, Bernoulli, and random Fourier ensembles, to name a few [6]. Their applications have been shown in medical images processing and other various signal processing problems . Even though random sensing matrices ensure high probability in reconstruction, they also have many drawbacks such as excessive complexity in reconstruction, significant space requirement for storage, and no efficient algorithm to verify whether a sensing matrix satisfies RIP property with small RIC value. Table 1 illustrates the comparison between random sensing and deterministic sensing .

Table 1. Comparison Between Random Sensing and Deterministic Sensing

Random Sensing	Deterministic Sensing
Outside the mainstream of signal processing: worst case signal processing	Aligned with the mainstream of signal processing: average case signal processing
Less efficient recovery time	More efficient recovery time
No explicit constructions	explicit constructions
Larger storage	Efficient Storage
Looser recovery bounds	Tight recovery bounds

Hence, exploiting specific structures of deterministic sensing matrices is required to solve these problems of the random sensing matrices. Recently, several deterministic sensing matrices have been proposed [4]. We can classify them into two categories. First are those matrices which are based on coherence . Second are those matrices which are based on RIP or some weaker RIPs . More recently in some highlighted results such as deterministic construction of sensing matrices via algebraic curves over finite fields in term of coherence and chirp sensing matrices have been introduced.

2.6 Reconstruction Model

In CS, we are concerned with the incoherence of matrix used to sample/sense signal of interest (hereafter referred as measurement matrix Φ) and the matrix representing a basis, in which signal of interest is sparse (hereafter referred as representation matrix Ψ). Within the CS framework, low coherence between Φ and Ψ translates to fewer samples required for reconstruction of signal.

Taking m linear measurements of a signal $x \in CN$ corresponds to applying a matrix $A \in C^{m\times N}$ — the measurement matrix

$$y = Ax. \qquad (10)$$

The vector $y \in C^m$ is called the measurement vector. The main interest is in the vastly undersampled case $m \ll N$. Without further information, it is, of course, impossible to recover x from y since the linear system Eq (10) is highly underdetermined, and has therefore infinitely many solutions. However, if the additional assumption that the vector x is k-sparse is imposed, then the situation dramatically changes as will be outlined. The approach for a recovery procedure that probably comes first to mind is to search for the sparsest vector x which is consistent with the measurement vector y = Ax. This leads to solving the ℓ0-miminization problem

$$\text{Min}||z||0 \text{ subject to } Az = y \qquad (11)$$

Unfortunately, this combinatorial minimization problem is NP–hard in general [9]. In other words, an algorithm that solves Eq (11) for any matrix A and any right hand side y is necessarily computationally intractable. Therefore, essentially two practical and tractable alternatives to Eq (11) have been proposed in the literature: convex relaxation leading to ℓ1-minimization — also called basis pursuit [12] and greedy algorithms, such as various matching pursuits [11]. Quite surprisingly for both types of approaches various recovery results are available, which provide conditions on the matrix A and on the sparsity $||x||_0$ such that the recovered solution coincides with the original x, and consequently also with the solution of Eq (11). This is no contradiction to the NP–hardness of Eq (11) since these results apply only to a subclass of matrices A and right-hand sides y. The ℓ1-minimization approach considers the solution of

$\text{Min}\|z\|_1$ subject to $Az = y$, (12)

which is a convex optimization problem and can be seen as a convex relaxation of Eq (11).

It is another important feature of compressive sensing that practical reconstruction can be performed by using efficient algorithms. Since the interest is in the vastly under sampled case, the linear system describing the measurements is underdetermined and therefore has infinitely many solution. The key idea is that the sparsity helps in isolating the original vector. The first naive approach to a reconstruction algorithm consists in searching for the sparsest vector that is consistent with the linear measurements. This leads to the combinatorial ℓ0-problem, see Eq (11) below, which unfortunately is NP-hard in general. There are essentially two approaches for tractable alternative algorithms. The first is convex relaxation leading to ℓ1-minimization also known as basis pursuit, see Eq (12) while the second constructs greedy algorithms. This overview focuses on ℓ1-minimization. By now basic properties of the measurement matrix which ensure sparse recovery by ℓ1-minimization are known: the null space property (NSP) and the restricted isometry property (RIP). The latter requires that all column sub- matrices of a certain size of the measurement matrix are well-conditioned. This is where probabilistic methods come into play because it is quite hard to analyze these properties for deterministic matrices with minimal amount of measurements. Among the provably good measurement matrices are Gaussian, Bernoulli random matrices, and partial random Fourier matrices.

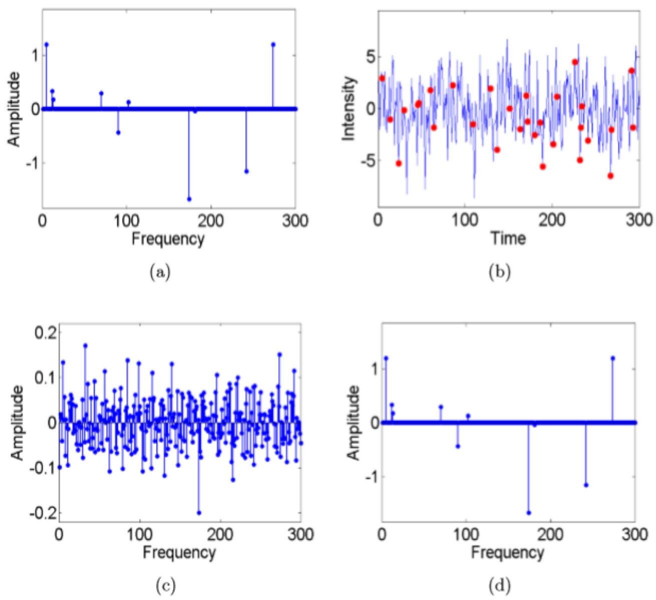

Figure 4 (a) 10-sparse Fourier spectrum, (b) time domain signal of length 300 with 30 samples, (c) reconstruction via ℓ2-minimization, (d) exact reconstruction via ℓ1-minimization

Currently the reconstruction algorithms are classified as three categories. One is the greedy algorithms, including the matching pursuit algorithms such as MP (Matching Pursuit), OMP(Orthogonal Matching Pursuit) , StOMP (Stagewise Orthogonal Matching Pursuit),CoSaMP (Compressive Sampling Matching Pursuit),and the gradient pursuit algorithms such as GP (Gradient Pursuit) and CGP (Conjugate Gradient Pursuit) .The second one is the convex relaxation, especially the projected gradient methods, iterative thresholding, Iterative hard thresholding, Bregman iterative algorithms, Basis Pursuit and Basis Pursuit De- Noising, the Least Absolute Shrinkage and Selection Operator (LASSO). The third is combinatorial algorithms. These methods acquire highly structured samples of the signal that support rapid reconstruction via group testing. Fourier sampling, chaining pursuit, and HH Pursuit.

The idea of Reconstruction In order to achieve an optimal recovery algorithm, there are several requirements that should be satisfied. The requirements are illustrated as below:

(1) Stability. The algorithm should be stable. That means when the signals or the measurements are perturbed slightly by noise, recovery should still be approximately accurate.

(2) Fast. The algorithm should be fast if we want to apply it into practice.

(3) Uniform guarantees. When acquiring linear measurements by using a specific method, these linear measurements can apply to all sparse signals.

(4) Efficiency. The algorithm should require as few measurements as possible.

2.6 Null Space Conditions and Thresholds for Rank Minimization

Minimizing the rank of a matrix subject to constraints is a challenging problem that arises in many applications in machine learning, control theory, and discrete geometry. This class of optimization problems, known as rank minimization, is NPHARD, and for most practical problems there are no efficient algorithms that yield exact solutions. A popular heuristic replaces the rank function with the nuclear norm—equal to the sum of the singular values—of the decision variable and has been shown to provide the optimal low rank solution in a variety of scenarios .It deals with the problem of recovering matrices of low ranks from compressed linear measurements through a heuristic of nuclear norm minimization. We assessed the practical performance of this heuristic for finding the minimum rank matrix subject to linear constraints. By characterizing a necessary and sufficient condition for the nuclear norm minimization to succeed, we provided the probabilistic performance bounds on the ranks as a function of the matrix dimensions and the number of constraints, for which the nuclear norm minimization heuristic succeeds with overwhelming probability[12]. The performance bounds we derived are tight in some regimes, especially the number of measurements and provided accurate predictions of the heuristic's performance in non-asymptotic scenarios. This suggests that a different parameterization of the null space of A could be the key to a better bound for small values of β. For large values of β, the bound is a rather good approximation of empirical results, and it might not be possible to further tighten this bound.

2.7 Applications

The fact that a compressible signal can be captured efficiently using a number of *incoherent* measurements that is proportional to its information level $S<<n$ has implications zthat are far reaching and concern a number of possible applications:

Data compression. In some situations, the sparse basis Ψ may be unknown at the encoder or impractical to implement for data compression. Φ can be considered a universal encoding strategy, as it need not be designed with regards to the structure of Ψ . (The knowledge and ability to implement. This universality may be particularly helpful for distributed source coding in multi-signal settings such as sensor networks [7].

Channel coding. CS principles (sparsity, randomness, and convex optimization) can be turned around and applied to design fast error correcting codes over the reals to protect from errors during transmission.

Inverse problems. In still other situations, the only way to acquire *f* may be to use a measurement system Φ of a certain modality. However, assuming a sparse basis Ψ exists for *f* that is also incoherent with Φ, then efficient sensing will be possible. One such application involves MR angiography and other types of MR setups [8], where Φ records a subset of the Fourier transform, and the desired image *f* is sparse in the time or wavelet domains.

Data acquisition. Finally, in some important situations the full collection of *n* discrete-time samples of an analog signal may be difficult to obtain (and possibly difficult to subsequently compress). Here, it could be helpful to design physical sampling devices that directly record discrete, low-rate incoherent measurements of the incident analog signal.

Analog to Information Conversion: Analog to-digital converters: (ADC) have been used in sensing and communications due to the advancement in digital signal processing. The process of ADC is based on the Nyquist sampling theorem which uniformly samples the signal with a rate of at least twice its bandwidth in order to reconstruct the signal perfectly. Emerging applications like radar detection and ultra-wideband communication are pushing the limit.

III. CONCLUSION

Compressive sensing (CS) is a novel sampling paradigm that samples signals in a much more efficient way than the established Nyquist sampling theorem. CS has recently gained a lot of attention due to its exploitation of signal sparsity .It gives a brief background on the origin, reviews the basic mathematical foundation. Numerous reconstruction algorithms aiming to achieve computational efficiency and high speeds are developed. Sparse recovery minimization, Null Space Conditions with Thresholds for Rank Minimization in compressive sensing scheme are analysed. With linear decoding complexity, deterministic performance guarantees of linear sparsity recovery, and explicitly constructible measurement matrices for the sparse signals.

REFERENCES

[1] M Fornasier, H Rauhut,"Compressive Sensing", Handbook of mathematical methods in imaging,2011, pp 187-228.

[2] Qaisar, S. ,Bilal, R.M. ,Iqbal, W., Naureen, M. ," Compressive Sensing: From Theory to Applications", Communications and Networks, Journal of Vol 15, pp 443-456,Oct 2013.

[3] Siddamal, K.V. Bhat, S.P. ; Saroja, V.S.,"A survey on compressive sensing" *IEEE Proceedings of International Conference* Electronics and Communication Systems (ICECS) pp 639-643,Feb 2015.

[4] Mohammed M. Abo-Zahhad, Aziza I. Hussein, Abdelfatah M. ohamed," Compressive Sensing Algorithms for Signal Processing Applications: A Survey", International Journal of Communications, Network and System Sciences, 2015,vol 8, pp197-216.

[6] Baraniuk, R. Compressive Sensing. *IEEE Signal Processing Magazine*, vol 24, 118-121,2007.

[5] EJ Candès, MB Wakin, "An introduction to compressive sampling" Signal Processing Magazine, vol 25, no 2, pp. 21 - 30, March 2008.

[6] S. S. Chen, D. L. Donoho, and M. A. Saunders. Atomic decomposition by Basis Pursuit. SIAM J. Sci. Comput., 20(1):33–61, 1999.

[7] D. Baron, M.B. Wakin, M.F. Duarte, S. Sarvotham, and R.G. Baraniuk, "Distributed compressed sensing," 2005, Preprint.

[8] M. Lustig, D.L. Donoho, and J.M. Pauly, "Rapid MR imaging with compressed sensing and randomly under-sampled 3DFT trajectories," in *Proc. 14th Ann. Meeting ISMRM*, Seattle, WA, May 2006.

[9] J. Tropp and D. Needell. CoSaMP: Iterative signal recovery from incomplete and inaccurate samples. Appl. Comput. Harmon. Anal., page 30, 2008.

[10] B. K. Natarajan. Sparse approximate solutions to linear systems. SIAM J. Comput., 24:227–234, 1995.

[11] E. J. Candes. The restricted isometry property and its implications for compressed sensing. C. R. Acad. Sci. I, 346:589–592, 2008.

[12] D.L. Donoho, "Compressed sensing," *IEEE Transactions on Information Theory*, vol. 52, no. 4, pp. 1289–1306, April 2006.