Reconstruction Algorithms for Compressive Sensing

Dheeraj Mekala
dmekala@ucsd.edu

Udayan Joshi
ujoshi@ucsd.edu

Abstract

Reconstructing a signal is a classical problem in the field of Signal Processing with wide ranged applications in medical imaging, electronics. This is possible by some assumptions about the signal and using some samples from signal. In this work, we introduce mathematical background, formulate the problem with noisy measurements and investigate the robustness of existing reconstruction techniques on this noisy formulation. We impose additional linear constraints and apply these reconstruction algorithms to reconstruct some famous images used in the image processing community.

1 Contributions

All the members of the team contributed equally to all parts of the project including literature survey, implementation, experimentation, analysis and the final report.

2 Introduction

The Compressive Sensing(CS) is a signal processing technique for efficiently acquiring signals in a compressive format and then reconstructing them(Wikipedia contributors, 2020). CS has a wide variety of applications in photography (Veeraraghavan et al., 2010), holography (Rivenson et al., 2010), facial recognition (Zhang et al., 2012) and also varied fields like Medical Image Processing (Lustig et al., 2007). Many have thought this task is impossible but with prior knowledge and some assumptions on signal, the reconstruction is possible from a series of measurements. These measurements could be noisy and ideally, the reconstruction techniques has to be robust for this noise. In this work, we restrict to reconstructing discrete signals and investigate the variation of reconstruction error with the number of measurements and the relative intensity of noise. We also try to improve it by imposing additional linear constraints. We will introduce mathematical background, relevant reconstruction algorithms without noise, formal problem description considering noise and its potential improvement for this problem.

Reproducibility: We made our code and signal data publicly available in our GitHub repository[1].

[1] https://github.com/dheeraj7596/compressiveSensing

3 Mathematical Background

A simple and most popular mechanism to reconstruct a signal is by linear functionals recording the values. This can be formulated as follows: Given a signal $x \in \mathbb{R}^n$, we take $m << n$ different measurements of the signal which can be written as a vector $y \in \mathbb{R}^m$. Each measurement y_j is of the form $y_j = \phi_j x$. We can thus express this in a matrix form as shown below:

$$y = \Phi x$$

Since, the signal is reconstructed using linear measurements, this problem is called Linear Compressive Sensing. Note that in the above equation $\Phi \in \mathbb{R}^{mxn}$ with $m < n$. Thus, this is an underdetermined system of equations in the variable x and thus has infinitely many solutions. In order to have any chance of reconstructing the original signal, we must impose additional constraints. There are two underlying constraints that makes this efficient reconstruction possible: 1.) Sparsity of signal of interest 2.) Incoherence of sensing modality.

The first condition that we assume is the original signal is utmost s-sparse in some basis Ψ, where $s << n$. This means that there exists some basis Ψ such that $x = \Psi\theta$ and θ is s-sparse i.e. θ has s non-zero entries. This condition ensures that the actual content of signal isn't as high as any arbitrary n-dimensional signal.

The second condition is imposed on the rows of matrix Φ. This condition states that the rows of the sensing matrix must be incoherent with the basis in which the signal is sparse. This condition ensures that our measurements actually contain enough information about the original signal. This aspect is captured by properties like Mutual Coherence and the Restricted Isometry Property (RIP) (Candès and Wakin, 2008).

With these two conditions, we can pose our reconstruction as an optimization problem.

$$min\ ||\theta||_0 \ \ subject\ to\ y = \Phi\Psi\theta \tag{1}$$

The above problem is NP-hard as it is currently stated.

4 Reconstruction Algorithms

All the methods for solving (1) can be broadly classified into two categories. The first category solves the problem using a convex relaxation as follows:

$$min\ ||\theta||_1 \ \ subject\ to\ y = \Phi\Psi\theta \tag{2}$$

The above problem can be easily posed as a linear program and can be thus solved efficiently. There are well known error bounds on the reconstruction error whenever our matrix $A = \Phi\Psi$ obeys certain properties like RIP or mutual incoherence. In some cases, the solution to (2) can exactly match the original signal or be very close (Candès, 2008).

The second category of methods are greedy approximate algorithms that directly tackle the NP-hard problem (1). The hard part about this problem is identifying the support set of the solution. If we know the support set, then minimizing the cost function $||y - \Phi\Psi\theta||_2^2$ on the support set is easy. We will implement two important methods called Orthogonal Matching Pursuit(OMP) (Cai and Wang, 2011) and Gradient Support Pursuit(GraSP) Algorithm (Bahmani et al., 2013).

4.1 Orthogonal Matching Pursuit Algorithm

In the OMP algorithm, we build the support set of our solution iteratively. The size of the support set grows by one per iteration. We greedily pick the index with the largest magnitude of the gradient of the cost function and add it to the support set. The pseudocode is mentioned in Algorithm-1.

Algorithm 1: OMP algorithm

Input : $A\ (columns\ are\ normalized), y, \epsilon$
Output : θ_{T_i}, T_i
Initialize: $r_0 \leftarrow y$, $\theta \leftarrow 0$, $T_0 \leftarrow \Phi$, $i \leftarrow 0$
while $||r_i||^2 > \epsilon$ **do**
  $j \leftarrow argmax_j |r_i^T a_j|$
  $T_{i+1} \leftarrow T_i \cup j$
  $i \leftarrow i + 1$
  $\theta_{T_i} \leftarrow argmin_w ||y - A_{T_i} W||$
  $r_i \leftarrow y - A_{T_i}\theta_{T_i}$
end

4.2 Gradient Support Pursuit Algorithm

This algorithm computes the gradient of the cost function at the current estimate in each iteration. It then pick 2s co-ordinates of the gradient having the most absolute value. These co-ordinates, along with the support of the current estimate give us the 3s co-ordinates over which we minimize. At the end, we prune the estimate to have sparsity s. The pseudo code is mentioned in Algorithm-2.

Algorithm 2: GraSP algorithm

Input : $f(.), s$
Output : $\hat{\mathbf{x}}$
Initialize: $\hat{\mathbf{x}} \leftarrow 0$
repeat
  compute local gradients: $\mathbf{z} = \nabla f(\hat{\mathbf{x}})$
  identify directions: $\mathcal{Z} = supp(\mathbf{z}_{2s})$
  merge supports: $\mathcal{T} = \mathcal{Z} \cup supp(\hat{\mathbf{x}})$
  minimize over support: $\mathbf{b} = \text{argmin } f(\mathbf{x})$
   s.t. $\mathbf{x}|_{\mathcal{T}^c = 0}$
  prune estimate: $\hat{\mathbf{x}} = \mathbf{b}_s$
until *halting condition holds*;

5 Problem Formulation

5.1 Primal Formulation

5.1.1 Reconstruction from Linear Measurements

The primal problem for reconstruction using linear measurements is as follows:

$$\begin{gathered} \min ||\theta||_1 \\ \text{subject to: } y = \Phi\Psi\theta \end{gathered}$$

5.1.2 Reconstruction from Noisy Linear Measurements

By considering additive Gaussian noise in our linear measurements of the form $y = \Phi\Psi\theta + \eta$, the convex relaxation takes the form of an SOCP as follows:

$$\begin{gathered} \min ||\theta||_1 \\ \text{subject to: } ||y - \Phi\Psi\theta||_2 < \epsilon \end{gathered}$$

Since, min $||\theta||_1$ = min $1^T t$ where $|\theta_i| <= t_i$, we can rewrite the above problem as follows:

$$\begin{gathered} \min 1^T t \\ \text{subject to: } (y - \Phi\Psi\theta)(y - \Phi\Psi\theta)^T < \epsilon^2 \\ I\theta - It <= 0 \\ -I\theta - It <= 0 \end{gathered} \tag{3}$$

5.2 Dual Formulation

5.2.1 Reconstruction using Linear Measurements

From the primal problem mentioned above, the lagrangian is as follows:

$$L(\theta, \mu) = ||\theta||_1 + \mu^T(y - \Phi\Psi\theta)$$

The dual function is:

$$\begin{aligned} g(\mu) &= \inf_{\theta \in D}[||\theta||_1 - \mu^T\Phi\Psi\theta] + \mu^T y \\ &= -\sup_{\theta \in D}[-||\theta||_1 + \mu^T\Phi\Psi\theta)] + \mu^T y \end{aligned}$$

Since:

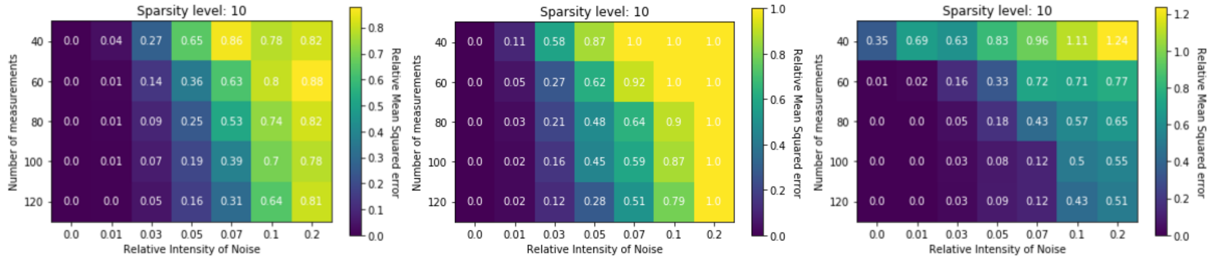
$$\sup_y z^T y - \lambda||y||_1 = \begin{cases} 0, & \text{if } ||z||_\infty \le \lambda \\ -\infty, & \text{otherwise} \end{cases}$$

The dual function can be written as follows:

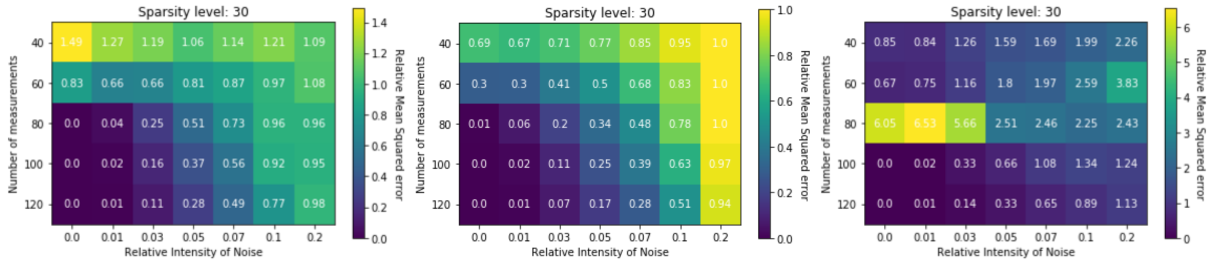
$$g(\mu) = \begin{cases} \mu^T y, & \text{if } ||\Psi^T\Phi^T\mu||_\infty \le 1 \\ -\infty, & \text{otherwise} \end{cases}$$

Thus, the dual problem is:

$$\begin{gathered} \max \mu^T y \\ \text{subject to: } ||\Psi^T\Phi^T\mu||_\infty \le 1 \end{gathered}$$



(a) Reconstruction error of a 10-sparse signal using OMP, L_1 Min, GraSP algorithms (left to right)



(b) Reconstruction error of a 30-sparse signal using OMP, L_1 Min, GraSP algorithms (left to right)

Figure 1: Reconstruction of Sparse signals

5.2.2 Reconstruction from Noisy Linear Measurements

From the primal problem mentioned in the previous section, the lagrangian is as follows:

$$L(\theta, \lambda_1, \lambda_2, \lambda_3) = 1^T t +$$
$$\lambda_1((y - \Phi\Psi\theta)(y - \Phi\Psi\theta)^T - \epsilon^2)$$
$$+{\lambda_2}^T(I\theta - It) + {\lambda_3}^T(-I\theta - It)$$

By computing the gradient with respect to θ and equating to zero, the dual problem is as follows:

$$\max 1^T t + \lambda_1((y - \Phi\Psi\hat{\theta})(y - \Phi\Psi\hat{\theta})^T - \epsilon^2)$$
$$+{\lambda_2}^T(I\hat{\theta} - It) + {\lambda_3}^T(-I\hat{\theta} - It)$$
$$\text{subject to: } \hat{\theta} =$$
$$(2\lambda_1\Psi^T\Phi^T\Phi\Psi)^{-1})(2\lambda_1\Psi^T\Phi^T y - I\lambda_2 + I\lambda_3)$$
$$\lambda_1 \geq 0$$
$$\lambda_2 \geq 0$$
$$\lambda_3 \geq 0$$

6 Experiments

We divide our experiments broadly into two parts: 1.) Reconstruction of sparse signals, 2.) Reconstruction of Images. In the first part, we reconstruct the sparse signals. This is comparatively easier than second part where we reconstruct an image, which is an nxn dense matrix. For all the experiments, we compare the performance between OMP, GraSP and solving the convex formulation mentioned in equation - 3 which we call it L_1-min. We wrote routines for OMP and GraSP by ourselves and used CVX to implement L_1-min. We didn't use any existing code, everything is written by us.

6.1 Reconstruction of Sparse signals

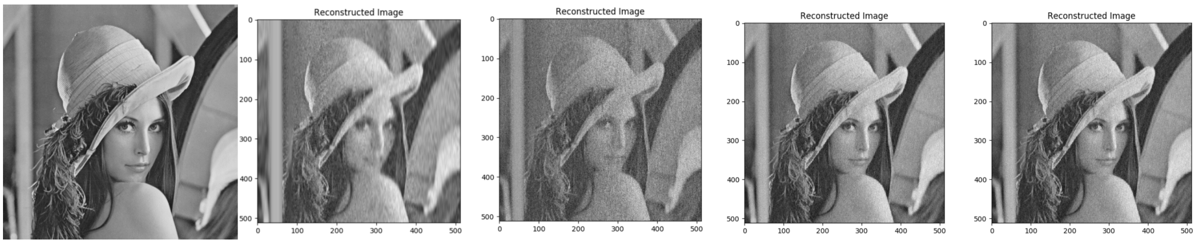
We synthetically generate s-sparse signals of dimension $n = 200$. Then, we take m noisy measurements, where the noise is additive Gaussian in nature with relative intensity σ. The performance of all three algorithms is analysed with varying the values of s, m and σ. The sparsity s could take values in $10, 30$. The number of measurements m could take values in $40, 60, 80, 100, 120$, while the relative noise intensity could take values in $0.00, 0.01, 0.03, 0.05, 0.07, 0.1, 0.2$. The variation of reconstruction error with m, σ is plotted as colormap for 10-sparse signals in Figure-1(a) and for a 30-sparse signals in Figure-1(b).

An intuitive pattern that can be easily observed from the colormaps is that decreasing the amount of noise and increasing the number of measurements improve our reconstruction. We see that all the algorithms perform equally well when the sparsity level is 10. However, in the case of s = 30, we see that L_1 minimization performs marginally better than both the greedy algorithms, especially in the high noise regions. On the other hand, greedy algorithms take significantly less time than the l_1 minimization formulation.

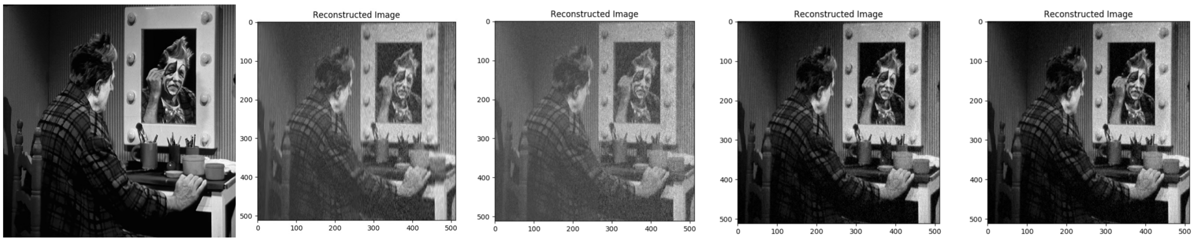
6.2 Reconstruction of Images

Images are usually dense nxn matrices. But the discrete cosine transform (DCT) represents an image as a sum of sinusoids of varying magnitudes and frequencies. In the space with DCT basis, image can be represented as a sparse matrix (Pennebaker and Mitchell, 1992).

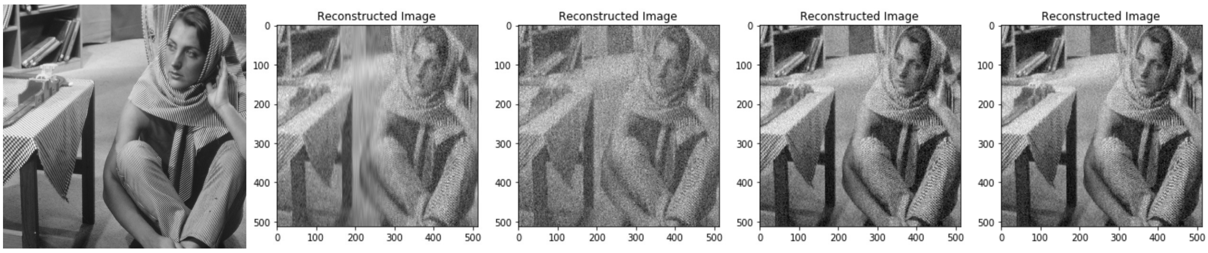
Novelty: We impose linear additional constraints during our L_1 minimization. Prior knowledge that the pixel values in an image lie in a range [0,255] is incorporated in our reconstruction algorithm. We add the



(a) Original image, Reconstructed image using OMP, GraSP, L_1-min with additional constraint, L_1-min without additional constraint (left to right)



(b) Original image, Reconstructed image using OMP, GraSP, L_1-min with additional constraint, L_1-min without additional constraint (left to right)



(c) Original image, Reconstructed image using OMP, GraSP, L_1-min with additional constraint, L_1-min without additional constraint (left to right)

Figure 2: Reconstruction of Famous images

constraint $0 \leq \Psi\theta \leq 255$ to our convex optimization problem.

Due to computation constraints, we operate on the images in a column-wise fashion, i.e., we reconstruct the image one column at a time. For each column, we take $m = 300$ measurements and add Gaussian noise with 1% the intensity of the original signal and reconstruct each column. The reconstruction of Lenna image is shown in Figure-2(a), Clown image in Figure-2(b) and Barbara image in Figure-2(c).

From the Figure-2, we can observe that the image reconstructed using L_1-min algorithm with additional constraint is of the highest quality compared to greedy algorithms and L_1-min without constraint. The greedy algorithms take significantly less time than L_1-min algorithm. L_1-min is relatively more robust to noise than greedy algorithms. Adding additional linear constraints gives a minute improvement.

7 Conclusion

We compared various algorithms on synthetic as well as natural(image) signals. We found that each type of algorithm has its pros and cons. Solving (1) using l_1 convex relaxation results in better reconstruction over greedy approximation but the running time is too high. Imposing additional constraints gives little improvements in the reconstruction error. This gives us the future direction to improve greedy algorithms and come up with an algorithm which is the best of both worlds i.e. low reconstruction error and low running time.

References

Sohail Bahmani, Bhiksha Raj, and Petros T Boufounos. 2013. Greedy sparsity-constrained optimization. *Journal of Machine Learning Research*, 14(Mar):807–841.

T Tony Cai and Lie Wang. 2011. Orthogonal matching pursuit for sparse signal recovery with noise. *IEEE Transactions on Information theory*, 57(7):4680–4688.

Emmanuel J Candès and Michael B Wakin. 2008. An introduction to compressive sampling. *IEEE signal processing magazine*, 25(2):21–30.

Emmanuel J. Candès. 2008. The restricted isometry property and its implications for compressed sensing.

Michael Lustig, David Donoho, and John M Pauly. 2007. Sparse mri: The application of compressed sensing for rapid mr imaging. *Magnetic Resonance in Medicine: An Official Journal of the International Society for Magnetic Resonance in Medicine*, 58(6):1182–1195.

William B Pennebaker and Joan L Mitchell. 1992. *JPEG: Still image data compression standard*. Springer Science & Business Media.

Yair Rivenson, Adrian Stern, and Bahram Javidi. 2010. Compressive fresnel holography. *Journal of Display Technology*, 6(10):506–509.

Ashok Veeraraghavan, Dikpal Reddy, and Ramesh Raskar. 2010. Coded strobing photography: Compressive sensing of high speed periodic videos. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 33(4):671–686.

Wikipedia contributors. 2020. Compressed sensing — Wikipedia, the free encyclopedia. [Online; accessed 18-March-2020].

Shiqing Zhang, Xiaoming Zhao, and Bicheng Lei. 2012. Robust facial expression recognition via compressive sensing. *Sensors*, 12(3):3747–3761.