Claremont Colleges
Scholarship @ Claremont

CMC Faculty Publications and Research   CMC Faculty Scholarship

12-11-2008

Greedy Signal Recovery Review

Deanna Needell
Claremont McKenna College

Joel A. Tropp
California Institute of Technology

Roman Vershynin
University of Michigan - Ann Arbor

Recommended Citation

GREEDY SIGNAL RECOVERY REVIEW

DEANNA NEEDELL, JOEL A. TROPP, ROMAN VERSHYNIN

ABSTRACT. The two major approaches to sparse recovery are L_1-minimization and greedy methods. Recently, Needell and Vershynin developed Regularized Orthogonal Matching Pursuit (ROMP) that has bridged the gap between these two approaches. ROMP is the first stable greedy algorithm providing uniform guarantees.

Even more recently, Needell and Tropp developed the stable greedy algorithm Compressive Sampling Matching Pursuit (CoSaMP). CoSaMP provides uniform guarantees and improves upon the stability bounds and RIC requirements of ROMP. CoSaMP offers rigorous bounds on computational cost and storage. In many cases, the running time is just $O(N \log N)$, where N is the ambient dimension of the signal. This review summarizes these major advances.

1. INTRODUCTION

Sparse signals are those that contain much less information than their ambient dimension suggests. The conventional signal compression scheme acquires the entire signal and then compresses it. This methodology has been questioned for decades, and new approaches in compressed sensing have been developed to overcome this seemingly wasteful approach.

Suppose $\boldsymbol{x}$ is a signal in $\mathbb{R}^N$, and define the ℓ_0 quasi-norm[1]

$$\|\boldsymbol{x}\|_0 = |\mathrm{supp}(\boldsymbol{x})| = |\{j : x_j \neq 0\}|.$$

When $\|\boldsymbol{x}\|_0 \leq s$, we say that the signal $\boldsymbol{x}$ is *s-sparse*. In practice, signals are not exactly sparse, but are rather close to sparse vectors. For example, compressible signals are those whose coefficients decay rapidly when sorted by magnitude. We say that a signal $\boldsymbol{x}$ is *p-compressible* with magnitude R if the sorted components of the signal decay at the rate

$$|x|_{(i)} \leq R \cdot i^{-1/p} \qquad \text{for } i = 1, 2, 3, \ldots. \tag{1.1}$$

The sparse recovery problem is the reconstruction of such signals from a set of nonadaptive linear measurements. The measurements are of the form $\boldsymbol{\Phi x}$ where $\boldsymbol{\Phi}$ is some $m \times N$ measurement matrix. Although in general this recovery is NP-Hard, work in compressed sensing has shown that for certain kinds of measurement matrices, recovery is possible when the number of measurements m is nearly linear in the sparsity s,

$$m = s \log^{O(1)}(N). \tag{1.2}$$

The survey [2] contains a discussion of these results.

The two major algorithmic approaches to sparse recovery are based on L_1-minimization and on greedy methods (Matching Pursuits). In this review we briefly describe these methods, as well as two new iterative methods that provide the advantages of both approaches. The first method, ROMP, is the first stable algorithm to provide uniform guarantees. The second method, CoSaMP, improves upon the results of ROMP, and provides rigorous computational bounds.

[1]We consider real numbers here for simplicity, but similar results hold for the complex case.

2. Major approaches

2.1. L_1-minimization. To recover the sparse signal $\boldsymbol{x}$ from its measurements $\boldsymbol{\Phi x}$, one needs to find the the solution to the highly non-convex problem

$$\min \|\boldsymbol{z}\|_0 \qquad \text{subject to} \qquad \Phi\boldsymbol{z} = \Phi\boldsymbol{x}. \qquad (L_0)$$

Donoho and his associates [6] suggested that for some measurement matrices Φ, the generally NP-Hard problem (L_0) should be equivalent to its convex relaxation:

$$\min \|\boldsymbol{z}\|_1 \qquad \text{subject to} \qquad \Phi\boldsymbol{z} = \Phi\boldsymbol{x}, \qquad (L_1)$$

where $\|\boldsymbol{z}\|_1 = \sum_i |z_i|$ denotes the ℓ_1-norm. The convex problem (L_1) can be solved using methods of linear programming.

Clearly if the measurement matrix $\boldsymbol{\Phi}$ is one-to-one on all $2s$-sparse vectors, then the s-sparse signal $\boldsymbol{x}$ will be recovered by solving (L_0). Candès and Tao [4] proved that if $\boldsymbol{\Phi}$ satisfies a stronger condition then recovery is possible by solving the convex problem (L_1).

Definition 2.1 (Restricted Isometry Condition). *A measurement matrix Φ satisfies the* Restricted Isometry Condition *(RIC) with parameters (n, δ) for $\delta \in (0,1)$ if we have*

$$(1 - \delta)\|\boldsymbol{v}\|_2 \le \|\Phi\boldsymbol{v}\|_2 \le (1 + \delta)\|\boldsymbol{v}\|_2 \qquad \text{for all } n\text{-sparse vectors.}$$

When δ is small, the restricted isometry condition says that every set of n columns of $\boldsymbol{\Phi}$ is approximately an orthonormal system. It has been shown (see [13] and [7]) that random Gaussian, Bernoulli and partial Fourier matrices satisfy the restricted isometry condition with number of measurements as in (1.2). In the more practical case when $\boldsymbol{x}$ is not exactly sparse and corrupted with noise, we consider the mathematical program:

$$\min \|\boldsymbol{y}\|_1 \qquad \text{subject to} \qquad \|\boldsymbol{\Phi y} - \boldsymbol{u}\|_2 \le \varepsilon. \qquad (2.1)$$

This program approximately recovers the signal $\boldsymbol{x}$ even in the presence of noise.

Theorem 2.2 (Recovery under RIC [3]). *Assume that the measurement matrix Φ satisfies the Restricted Isometry Condition with parameters $(3s, 0.2)$. Let $\boldsymbol{\Phi}$ be a measurement matrix and let $\boldsymbol{u} = \boldsymbol{\Phi x} + \boldsymbol{e}$ be a noisy measurement vector where $\boldsymbol{x} \in \mathbb{R}^N$ is an arbitrary signal and $\|\boldsymbol{e}\|_2 \le \varepsilon$. Then the program* (2.1) *produces an approximation $\boldsymbol{x}^\#$ that satisfies:*

$$\left\|\boldsymbol{x} - \boldsymbol{x}^\#\right\|_2 \le \mathrm{C}\left[\frac{1}{\sqrt{s}}\|\boldsymbol{x} - \boldsymbol{x}_s\|_1 + \varepsilon\right],$$

where $\boldsymbol{x}_s$ denotes the s-sparse vector consisting of the s largest coefficients in magnitude of $\boldsymbol{x}$.

In [5], Candès sharpened this theorem to work under the restricted isometry condition with parameters $(2s, \sqrt{2}-1)$. This theorem also demonstrates that in the noiseless case, the L_1 approach provides exact reconstruction. This was proved initially in [4].

2.2. Orthogonal Matching Pursuit (OMP). An alternative approach to sparse recovery is via greedy algorithms. These methods find the support of the signal $\boldsymbol{x}$ iteratively, and reconstruct the signal using the pseudoinverse.

Orthogonal Matching Pursuit (OMP) is such an algorithm, analyzed by Gilbert and Tropp in [14]. Since we expect the columns of the measurement matrix $\boldsymbol{\Phi}$ to be approximately orthonormal, $\boldsymbol{\Phi}^*\boldsymbol{\Phi x}$ is locally a good approximation to $\boldsymbol{x}$. OMP uses this idea to compute the support of a s-sparse signal $\boldsymbol{x}$. First, the residual $\boldsymbol{r}$ is set to the measurement vector $\boldsymbol{u}$. At each iteration, the observation vector is set, $\boldsymbol{y} = \Phi^*\boldsymbol{r}$, and the coordinate of its largest coefficient in magnitude is added to the index set I. Then by solving a least squares problem, the residual is updated to remove this coordinate's contribution,

$$\boldsymbol{y} = \arg\min_{\boldsymbol{z} \in \mathbb{R}^I} \|\boldsymbol{u} - \Phi\boldsymbol{z}\|_2; \qquad \boldsymbol{r} = \boldsymbol{u} - \Phi\boldsymbol{y}.$$

Repeating this s times yields an index set of s coordinates corresponding to the support of the signal $\boldsymbol{x}$. Tropp and Gilbert [14] showed that OMP recovers a sparse signal with high probability.

Theorem 2.3 (OMP Recovery [14]). *Let $\boldsymbol{\Phi}$ be a $m \times N$ subgaussian matrix, and fix a s-sparse signal $\boldsymbol{x} \in \mathbb{R}^N$. Then OMP recovers (the support of) $\boldsymbol{x}$ from the measurements $\boldsymbol{u} = \boldsymbol{\Phi}\boldsymbol{x}$ correctly with high probability, provided the number of measurements is $m \sim s \log N$.*

2.3. **Advantages and challenges of both approaches.** The L_1-minimization method provides *uniform guarantees* for sparse recovery. Once the measurement matrix Φ satisfies the restricted isometry condition, this method works correctly for all sparse signals $\boldsymbol{x}$. The method is also *stable*, so it works for non-sparse signals such as those which are compressible, as well as noisy signals. However, the method is based on linear programming, and there is no strongly polynomial time algorithm in linear programming yet.

OMP on the other hand, is quite *fast* both provably and empirically. The speed of OMP is a great advantage, but it lacks the strong guarantees that L_1 provides. Indeed, OMP works correctly for a *fixed* signal and measurement matrix with high probability, and so it must fail for some sparse signals and matrices [12]. It is also unknown whether OMP succeeds for compressible signals or on noisy measurements.

There has thus existed a gap between the approaches. The development of Regularized Orthogonal Matching Pursuit (ROMP) bridges this gap by providing a greedy algorithm with the same advantages as the L_1 method. Compressive Sampling Matching Pursuit (CoSaMP) improves upon these results and provides rigorous runtime guarantees. We now discuss these new algorithms.

3. Regularized OMP

Regularized Orthogonal Matching Pursuit (ROMP) is a new algorithm developed by Needell and Vershynin in [11, 10] for sparse recovery that performs correctly for all measurement matrices Φ satisfying the restricted isometry condition, and for all sparse signals. Again, since the restricted isometry condition gaurantees every set of s columns forms approximately an orthonormal system, every s coordinates of the observation vector $\boldsymbol{y} = \boldsymbol{\Phi}^*\boldsymbol{u}$ are in a loose sense good estimators of the corresponding s coefficients of $\boldsymbol{x}$. This notion suggests to use the s largest coefficients of the observation vector $\boldsymbol{y}$ rather than only the largest, as in OMP. We also include a regularization step to ensure that each coordinate carries close to an even share of the information. The ROMP algorithm is described in Algorithm 3.1.

For measurement matrices that satisfy the RIC, ROMP stably recovers all s-sparse signals. This is summarized in the following theorem from [10].

Theorem 3.1 (Recovery by ROMP [10]). *Assume a measurement matrix Φ satisfies the restricted isometry condition with parameters $(8s, \epsilon)$ for $\epsilon = 0.01/\sqrt{\log s}$. Consider an arbitrary vector $\boldsymbol{x}$ in $\mathbb{R}^N$. Suppose that the measurement vector $\Phi\boldsymbol{x}$ becomes corrupted, so we consider $\boldsymbol{u} = \Phi\boldsymbol{x} + \boldsymbol{e}$ where $\boldsymbol{e}$ is some error vector. Then ROMP produces a good approximation $\hat{\boldsymbol{x}}$ to $\boldsymbol{x}$:*

$$\|\hat{\boldsymbol{x}} - \boldsymbol{x}\|_2 \leq C\sqrt{\log s}\Big(\|\boldsymbol{e}\|_2 + \frac{\|\boldsymbol{x} - \boldsymbol{x}_s\|_1}{\sqrt{s}}\Big). \tag{3.1}$$

In the case where the signal $\boldsymbol{x}$ is exactly sparse without noise, this theorem guarantees exact reconstruction. Note also that in the noisy case, ROMP needs no knowledge about the error vector $\boldsymbol{e}$ to approximate the signal. In the special case where $\boldsymbol{x}$ is a compressible signal as in (1.1), the theorem provides the bound

$$\|\boldsymbol{x} - \hat{\boldsymbol{x}}\|_2 \leq R'\frac{\sqrt{\log s}}{s^{p-1/2}} + C\sqrt{\log s}\|\boldsymbol{e}\|_2. \tag{3.2}$$

ROMP thus provides the first greedy approach with these uniform and stable guarantees. See [11, 10] for runtime analysis and empirical results. The results are optimal up to the logarithmic factors,

ALGORITHM 3.1: ROMP Recovery Algorithm [11]

ROMP($\boldsymbol{\Phi}$, $\boldsymbol{u}$, s)
Input: Measurement matrix $\boldsymbol{\Phi}$, measurement vector $\boldsymbol{u}$, sparsity level s
Output: Index set $I \subset \{1, \ldots, d\}$

Initialize: Let the index set $I = \emptyset$ and the residual $r = u$.
    Repeat the following steps s times or until $|I| \geq 2s$:
Identify: Choose a set J of the s biggest coordinates in magnitude of the observation vector $\boldsymbol{y} = \Phi^* \boldsymbol{r}$, or all of its nonzero coordinates, whichever set is smaller.
Regularize: Among all subsets $J_0 \subset J$ with comparable coordinates:
$$|\boldsymbol{y}(i)| \leq 2|\boldsymbol{y}(j)| \quad \text{for all } i, j \in J_0,$$
choose J_0 with the maximal energy $\|\boldsymbol{y}|_{J_0}\|_2$.
Update: Add the set J_0 to the index set: $I \leftarrow I \cup J_0$, and update the residual:
$$\boldsymbol{w} = \arg\min_{\boldsymbol{z} \in \mathbb{R}^I} \|\boldsymbol{u} - \Phi\boldsymbol{z}\|_2; \qquad \boldsymbol{r} = \boldsymbol{u} - \Phi\boldsymbol{w}.$$

and the runtime is similar to that of OMP. There is, however, room for improvement which leads us to the CoSaMP algorithm, developed by Needell and Tropp [9]. CoSaMP provides optimal guarantees as well as an important implementation analysis.

4. CoSaMP

CoSaMP is an iterative recovery algorithm that provides the same guarantees as even the best optimization approaches. As in the case of ROMP and the L_1 approach, CoSaMP recovers signals using measurement matrices that satisfy the RIC. Thus as before, the observation vector $\boldsymbol{y} = \boldsymbol{\Phi}^* \boldsymbol{u}$ serves as a good proxy for the signal $\boldsymbol{x}$. Using the largest coordinates, an approximation to the signal is formed at each iteration. After each new residual is formed, reflecting the missing portion of the signal, the measurements are updated. This is repeated until all the recoverable portion of the signal is found. (See [9] for halting criteria.) The algorithm is described in Algorithm 4.2.

For a recovery algorithm to be used efficiently in practice, the least squares step must be analyzed carefully. By the restricted isometry condition, the matrix $\boldsymbol{\Phi}_T$ in the estimation step is very well-conditioned. This suggests the use of an iterative method such as Richardson's iteration [1, Sec. 7.4] to apply the psuedoinverse $\boldsymbol{\Phi}_T^\dagger = (\boldsymbol{\Phi}_T^* \boldsymbol{\Phi}_T)^{-1} \boldsymbol{\Phi}_T^*$. This method is analyzed in the context of CoSaMP and shown to provide an efficient means of acquiring the estimation [9]. The following theorem from [9] summarizes the fundamentally optimal recovery guarantees and rigorous computational costs of CoSaMP.

Theorem 4.1 (Recovery by CoSaMP [9]). *Suppose that $\boldsymbol{\Phi}$ is an $m \times N$ measurement matrix satisfying the restricted isometry condition with parameters $(2s, \mathrm{c})$. Let $\boldsymbol{u} = \boldsymbol{\Phi}\boldsymbol{x} + \boldsymbol{e}$ be a vector of samples of an arbitrary signal, contaminated with arbitrary noise. For a given precision parameter η, the algorithm* CoSaMP *produces an s-sparse approximation $\boldsymbol{a}$ that satisfies*

$$\|\boldsymbol{x} - \boldsymbol{a}\|_2 \leq \mathrm{C} \cdot \max\left\{\eta, \frac{1}{\sqrt{s}}\left\|\boldsymbol{x} - \boldsymbol{x}_{s/2}\right\|_1 + \|\boldsymbol{e}\|_2.\right\}$$

The running time is $\mathrm{O}(\mathscr{L} \cdot \log(\|\boldsymbol{x}\|_2 / \eta))$, *where $\mathscr{L}$ bounds the cost of a matrix–vector multiply with $\boldsymbol{\Phi}$ or $\boldsymbol{\Phi}^*$. Working storage is* $\mathrm{O}(N)$.

In [8] it is shown that only a fixed number of iterations is required to reduce the error to an optimal amount. The report [8] also discusses variations on the CoSaMP algorithm that may improve results in practice.

<div style="text-align:center">

ALGORITHM 4.2: CoSaMP Recovery Algorithm [9]

</div>

CoSaMP($\boldsymbol{\Phi}$, $\boldsymbol{u}$, s)
Input: Sampling matrix $\boldsymbol{\Phi}$, noisy sample vector $\boldsymbol{u}$, sparsity level s
Output: An s-sparse approximation $\boldsymbol{a}$ of the target signal

$\boldsymbol{a}^0 \leftarrow \boldsymbol{0}$, $\boldsymbol{v} \leftarrow \boldsymbol{u}$, $k \leftarrow 0$ { Trivial initial approximation }

repeat
   $k \leftarrow k+1$
   $\boldsymbol{y} \leftarrow \boldsymbol{\Phi}^* \boldsymbol{v}$ { Form signal proxy }
   $\Omega \leftarrow \operatorname{supp}(\boldsymbol{y}_{2s})$ { Identify large components }
   $T \leftarrow \Omega \cup \operatorname{supp}(\boldsymbol{a}^{k-1})$ { Merge supports }

   $\boldsymbol{b}|_T \leftarrow \boldsymbol{\Phi}_T^{\dagger} \boldsymbol{u}$ { Signal estimation by least-squares }
   $\boldsymbol{b}|_{T^c} \leftarrow \boldsymbol{0}$

   $\boldsymbol{a}^k \leftarrow \boldsymbol{b}_s$ { Prune to obtain next approximation }
   $\boldsymbol{v} \leftarrow \boldsymbol{u} - \boldsymbol{\Phi} \boldsymbol{a}^k$ { Update current samples }
until halting criterion *true*

REFERENCES

[1] Å. Björck. *Numerical Methods for Least Squares Problems*. SIAM, Philadelphia, 1996.

[2] E. Candès. Compressive sampling. In *Proc. International Congress of Mathematics*, volume 3, pages 1433–1452, Madrid, Spain, 2006.

[3] E. Candès, J. Romberg, and T. Tao. Stable signal recovery from incomplete and inaccurate measurements. *Communications on Pure and Applied Mathematics*, 59(8):1207–1223, 2006.

[4] E. Candès and T. Tao. Decoding by linear programming. *IEEE Trans. Inform. Theory*, 51:4203–4215, 2005.

[5] E. J. Candès. The restricted isometry property and its implications for compressed sensing. *C. R. Math. Acad. Sci. Paris, Serie I*, 346:589–592, 2008.

[6] D. L. Donoho and P. B. Stark. Uncertainty principles and signal recovery. *SIAM J. Appl. Math.*, 49(3):906–931, June 1989.

[7] S. Mendelson, A. Pajor, and N. Tomczak-Jaegermann. Uniform uncertainty principle for Bernoulli and subgaussian ensembles. To appear, *Constr. Approx.*, 2009.

[8] D. Needell and J. A. Tropp. CoSaMP: Iterative signal recovery from incomplete and inaccurate samples. ACM Technical Report 2008-01, California Institute of Technology, Pasadena, July 2008.

[9] D. Needell and J. A. Tropp. CoSaMP: Iterative signal recovery from noisy samples. *Appl. Comput. Harmon. Anal.*, 2008. DOI: 10.1016/j.acha.2008.07.002.

[10] D. Needell and R. Vershynin. Signal recovery from incomplete and inaccurate measurements via regularized orthogonal matching pursuit. Submitted for publication, October 2007.

[11] D. Needell and R. Vershynin. Uniform uncertainty principle and signal recovery via regularized orthogonal matching pursuit. DOI: 10.1007/s10208-008-9031-3, 2007.

[12] H. Rauhut. On the impossibility of uniform sparse reconstruction using greedy methods. To appear, *Sampl. Theory Signal Image Process.*, 2008.

[13] M. Rudelson and R. Vershynin. On sparse reconstruction from Fourier and Gaussian measurements. *Comm. Pure Appl. Math.*, 61:1025–1045, 2008.

[14] J. A. Tropp and A. C. Gilbert. Signal recovery from random measurements via orthogonal matching pursuit. *IEEE Trans. Info. Theory*, 53(12):4655–4666, 2007.