JOINTLY SPARSE VECTOR RECOVERY VIA REWEIGHTED ℓ_1 MINIMIZATION

Mu-Hsin Wei, Waymond R. Scott, Jr., and James H. McClellan

Georgia Institute of Technology School of Electrical and Computer Engineering 777 Atlantic Drive NW, Atlanta, GA 30332-0250

ABSTRACT

An iterative reweighted algorithm is proposed for the recovery of jointly sparse vectors from multiple-measurement vectors (MMV). The proposed MMV algorithm is an extension of the iterative reweighted ℓ_1 algorithm for single measurement problems. The proposed algorithm (M-IRL1) is demonstrated to outperform non-reweighted MMV algorithms under noiseless measurements. A regularization of the M-IRL1 algorithm is also proposed to accommodate noise. The ability to robustly handle noise is demonstrated through an electromagnetic induction application.

Index Terms— Jointly sparse, multiple-measurement vector, basis pursuit, iterative reweighting.

1. INTRODUCTION

In many estimation problems, the signal of interest has a sparse representation in a specific basis. Furthermore, in some problems, the sparse representations of a set of signals are *jointly sparse*, having their nonzero elements occurring at the same entries. Such problems arise in neuromagnetic imaging [1], equalization of sparse communication channels [2], and electromagnetic induction (EMI) [3]. In these problems, a set of L measurements $b^{(l)} \in \mathbb{R}^m$ can be modeled as

$$b^{(l)} = Ax^{(l)} \quad l = 1, \dots, L,$$
 (1)

where $A \in \mathbb{R}^{m \times n}$ is the dictionary matrix and $x^{(l)} \in \mathbb{R}^n$ are jointly sparse vectors. The system of equations (1) is also known as the multiple-measurement-vector (MMV) problem and is often written as

$$B = AX, (2)$$

where the matrix $B \in \mathbb{R}^{m \times L}$ is the collection of $b^{(l)}$ and likewise for $X \in \mathbb{R}^{n \times L}$, which is row-sparse – having nonzero entries in only a few rows. When L = 1, we have the single-measurement vector problem (SMV), which has been widely studied.

Several MMV recovery methods were proposed in recent years. Most of the methods are extensions of existing SMV recovery methods. Extension of the Basis Pursuit (BP) to the MMV problem is considered in [1, 4], where the objective is to minimize the number of rows containing nonzero entries while satisfying (2). The problem can be formulated as

$$\min ||\mathcal{R}_{\ell_a}(X)||_0 \quad \text{s.t.} \quad B = AX, \tag{3}$$

where $\mathcal{R}_{\ell_q}(X)$ is a vector in \mathbb{R}^n whose i^{th} entry is the ℓ_q vector norm of the i^{th} row of X, and $|| \cdot ||_0$ is number of nonzero entries in a given vector. As in the SMV problem, (3) is NP-hard but can be convexified as an ℓ_1 minimization problem:

$$\min ||\mathcal{R}_{\ell_a}(X)||_1 \quad \text{s.t.} \quad B = AX. \tag{4}$$

When the nonzero rows in X are sparse enough, (4) recovers the same solution as (3). The condition for which (3) and (4) are equivalent can be found in [1, 4]. It was also shown that an exact recovery does not depend on the ℓ_q norm chosen for \mathcal{R}_{ℓ_q} . On the other hand, greedy algorithms have also been extended to accommodate MMV problems [5, 6, 7]. Various MMV methods based on Matching Pursuit (MP) were proposed, such as the MMV orthogonal matching pursuit (M-OMP). The condition for exact recovery was also established [1, 5].

From a slightly different approach, it is proposed in [8] the framework, ReMBo, that solves a MMV problem by recasting it into a series of SMV problems. The framework can incorporate both convex relaxation and greedy algorithms, and is shown to be robust.

Sparsity could be further enhanced through iteratively reweighting. In particular, it was shown in [9, 10] that sparse solutions for a SMV problem can be found via iterative reweighted least-squares (IRLS), with which the FOCUSS algorithm [11] is identified. A M-FOCUSS algorithm that extends FOCUSS to the MMV problem was introduced in [10]. In addition, it was also shown in [12] that sparse solutions can be obtained via an iterative reweighted ℓ_1 algorithm (IRL1), as a result of minimizing a log-sum objective function.

Adopting the reweighting scheme, we propose a MMV method (M-IRL1) that iteratively reweights the M-BP algorithm. The proposed method can be seen as a natural MMV

This work is supported in part by the US Army REDCOM CERDEC Night Vision and Electronic Sensors Directorate, Science and Technology Division, Countermine Branch and in part by the U.S. Army Research Office.

extension of the IRL1 algorithm, just as M-FOCUSS is a natural extension of the FOCUSS (IRLS) algorithm. We also propose a regularized version of M-IRL1.

The paper is organized as follows: In Section 2, the proposed algorithm is presented, and the relationship of M-IRL1 to M-FOCUSS is discussed. In Section 3, the behavior of the proposed algorithm is examined using noiseless and noisy data. It is observed that through reweighting, the proposed algorithm outperforms M-BP, as well as other MMV methods. It is also demonstrated that the proposed method performs well, and better than M-FOCUSS, in the application of EMI spectrum estimation.

Notation: The remaining sections use the following notation:

$$\mathcal{R}_{\ell_q}(X) = [v_1, \dots, v_n]^T, \tag{5}$$

where
$$v_i = \left(\sum_{l=1}^L |x_{i,l}|^q\right)^{1/q}$$
 (6)

and $x_{i,l}$ are the entries of the matrix X.

2. METHOD

To obtain row-sparse solutions, we choose an objective function that penalizes non-row-sparse answers. The log-sum function is a closer approximation to the ℓ_0 quasi-norm than the ℓ_1 norm, so it better promotes sparsity. As shown in Fig. 1, the log-sum function has a slope that vanishes near the axes like $|| \cdot ||_0$ does. On the other hand, the ℓ_1 norm is less similar to the ℓ_0 quasi-norm. Other functions have also been suggested to better approximate the ℓ_0 quasi-norm [9, 12, 13].

We propose to approximate (3) by minimizing the logsum objective function on the norm of the rows:

$$\underset{X}{\operatorname{arg\,min}} \sum_{i=1}^{n} \log \left(v_i + \epsilon \right) \quad \text{s.t.} \quad B = AX, \tag{7}$$

where $\epsilon > 0$ is a small positive real number introduced for stability. Recall that $v_i \ge 0$ are the entries of $\mathcal{R}_{\ell_q}(X)$, the ℓ_q -norm of the rows of X defined in (6). For simplicity, we consider the case q = 1.

Following the reweighting scheme in [12], we can solve (7) with an iterative algorithm that we call M-IRL1:

Step 1) Initialize count k = 0 and $w_i^{(0)} = 1, i = 1, ..., n$. Step 2) Solve the weighted ℓ_1 minimization problem

$$X^{(k)} = \underset{X}{\operatorname{arg\,min}} ||W^{(k)} \mathcal{R}_{\ell_1}(X)||_1 \quad \text{s.t.} \quad B = AX,$$

where $W = \operatorname{diag}[w_1, w_2, \dots, w_n].$ (8)

Step 3) Update the weights:

$$w_i^{(k+1)} = \frac{1}{v_i^{(k)} + \epsilon}, \quad i = 1, \dots, n.$$
 (9)

Step 4) Terminate on convergence or when k reaches a specified maximum number of iterations k_{max} . Otherwise, iterate from Step 2.

While (7) better promotes sparsity, it is nonconvex and a unique solution is not guaranteed. The proposed M-IRL1 method can be trapped in local minima. However, when an initial point is properly chosen, the algorithm does converge to the global minimum, as shown empirically in Section 3. The proposed method converges as argued in [13].

2.1. Algorithm Justification

We show a justification of the proposed method. A similar derivation is found in [12], which is the following with L = 1.

In (7), we substitute for v_i using (5):

$$\underset{X}{\arg\min} \sum_{i=1}^{n} \log \left(\sum_{l=1}^{L} |x_{i,l}| + \epsilon \right) \quad \text{s.t.} \quad B = AX.$$
(10)

The minimization in (10) is equivalent to

$$\underset{s}{\operatorname{arg\,min}} \sum_{i=1}^{n} \log \left(\sum_{l=1}^{L} u_{i,l} + \epsilon \right) \quad \text{s.t.} \quad s \in \mathcal{C}, \qquad (11)$$

where s = (X, U) and C is the convex set $\{(x_{i,l}, u_{i,l}) | B = AX$ and $|x_{i,l}| \le u_{i,l}\}$.

Recognizing the objective function in (11) is concave, which is below its tangent, a guess $s^{(k)} \in C$ can be improved by minimizing a linearized objective function around $s^{(k)}$:

$$s^{(k+1)} = \arg\min_{s} g(s^{(k)}) + \nabla g(s^{(k)})(s - s^{(k)}) \text{ s.t. } s \in \mathcal{C},$$

where $g(s) = \sum_{i=1}^{n} \log\left(\sum_{l=1}^{L} u_{i,l} + \epsilon\right).$ (12)

It can be readily shown that

$$\frac{\partial g}{\partial u_{i,l}} = \left(\sum_{l=1}^{L} u_{i,l} + \epsilon\right)^{-1}.$$
 (13)

From (12) and (13), we obtain

$$(X^{(k+1)}, U^{(k+1)}) = \arg\min\sum_{i=1}^{n} \frac{\sum_{l=1}^{L} u_{i,l}}{\sum_{l=1}^{L} u_{i,l}^{(k)} + \epsilon}$$

s.t. $B = AX$ and $|x_{i,l}| \le u_{i,l}$, (14)

which is equivalent to

$$X^{(k+1)} = \arg\min\sum_{i=1}^{n} \frac{\sum_{l=1}^{L} |x_{i,l}|}{\sum_{l=1}^{L} |x_{i,l}^{(k)}| + \epsilon} \quad \text{s.t.} \quad B = AX.$$

Using (5), we obtain

$$X^{(k+1)} = \arg\min\sum_{i=1}^{n} \frac{v_i}{v_i^{(k)} + \epsilon}$$
 s.t. $B = AX$. (15)

The iteration and weights in (15) define the proposed algorithm.



Fig. 1. Surface plots of various (quasi-)norms in \mathbb{R}^2 . (a) ℓ_0 quasi-norm. (b) ℓ_1 norm. (c) $\sum \log(|x_i| + \epsilon)$. (d) $\ell_{1/2}$ quasi-norm [9].

2.2. M-IRL1 in relation to M-FOCUSS

Both the proposed method and M-FOCUSS promote sparsity by iterative reweighting. In this section, we examine the relation between the two.

The M-FOCUSS algorithm [5] derives its weights from minimizing the following diversity measure (e.g., Fig. 1(d)):

$$J^{(p)}(X) = ||\mathcal{R}_{\ell_2}(X)||_p^p = \sum_{i=1}^n (v_i)^p, \quad 0 \le p \le 1.$$
 (16)

By solving the Euler-Lagrange equation derived from (16), the iterative reweighted *least-squares* is obtained:

$$X^{(k)} = \tilde{W}^{(k)} A^T (A \tilde{W}^{(k)} A^T)^{-1} B, \qquad (17)$$

where $\tilde{W} = \text{diag}[\tilde{w}_1, \dots, \tilde{w}_n]$, and the weights are updated by

$$\tilde{w}_i^{(k+1)} = \left(v_i^{(k)}\right)^{p-2}.$$
(18)

The weights are updated similarly to that of the proposed method (9).

While similar, the two methods differ in that M-FOCUSS minimizes the diversity measure (16) and results in a reweighted *least-squares* method. While M-IRL1 minimizes the log-sum function (7) and results in a reweighted ℓ_1 minimization problem. Nonetheless, both methods are instances of a general iterative sparsity enhancing algorithm proposed in [13]. Both deliver robust performance, and can be regularized to accommodate noisy measurements, as shown in the following section.

It can be shown that when p = 0, M-FOCUSS effectively minimizes the log-sum objective [9], as with M-IRL1. In this case, iterations on (17) and (18) lead to minimizing a logsum with summand entries of \mathcal{R}_{ℓ_2} . However, when regularized, M-FOCUSS (p = 0) and M-IRL1 have different objectives because of the regularization introduced at each iteration. Simulations show differences in the performance of the two algorithms, see Fig. 3.



Fig. 2. Performance of MMV methods under noiseless data, including M-IRL1 and two M-FOCUSS cases.

3. NUMERICAL EXPERIMENTS

The proposed method is tested on noiseless and noisy data. M-IRL1 is found to provide a robust exact recovery rate for noise-free data. Also, a regularized M-IRL1 is shown to accommodate noisy data in the context of an EMI application. We choose $k_{max} = 20$ and $\epsilon = 10^{-6}$.

3.1. Noiseless Data

We proceed with a numerical experiment investigating the behavior of the proposed M-IRL1 algorithm. Other methods considered for examination include M-BP [1], M-OMP [1], M-FOCUSS [5], and ReMBo (with BP) [8].

We choose, m = 20, n = 30, and L = 5. The entries of the real-valued matrix A are *i.i.d.* Gaussian with zero mean and variance one. The multiple-measurement vectors are constructed by $B = AX_0$ where X_0 has N rows with nonzero entries. The locations of the N rows are selected uniformly at random, and the nonzero entries of X_0 are drawn as in A. The above simulation is repeated 500 times per MMV method.

The empirical rate of exact recovery for different N is shown in Fig. 2. If we concentrate on very high recovery rates above 0.95, then M-IRL1 and both M-FOCUSS cases achieve the same recovery rate for $N \leq 13$. For N > 13, M-IRL1 exhibits performance comparable to M-FOCUSS (p = 0), the IRLS counterpart of the proposed method.



Fig. 3. Performance under noisy EMI measurements.

3.2. Noisy Data

We consider recovery from measurements corrupted by additive white Gaussian noise:

B = AX + i.i.d. Gaussian noise.

To handle the noise, we regularize (8):

$$X^{(k+1)} = \arg\min_{X} ||W\mathcal{R}_{\ell_1}(X)||_1 + \lambda ||AX - B||_F, \quad (19)$$

where λ is a regularization parameter, and $|| \cdot ||_F$ is the Frobenius norm. The selection of λ can be found via a crossvalidation-like experiment for each specific applications, i.e., for each matrix A [14].

We demonstrate the performance of the regularized M-IRL1 in the application of spectrum estimation in EMI [3]. In this case, the matrix A has the structure

$$A = \begin{bmatrix} 1 & \frac{1-j\omega_1/\zeta_1}{1+j\omega_1/\zeta_1} & \frac{1-j\omega_1/\zeta_1}{1+j\omega_1/\zeta_2} & \dots \\ 1 & \frac{1-j\omega_2/\zeta_1}{1+j\omega_2/\zeta_1} & \frac{1-j\omega_2/\zeta_2}{1+j\omega_2/\zeta_2} & \dots \\ \vdots & \vdots & \vdots & \ddots \\ 1 & \frac{1-j\omega_m/\zeta_1}{1+j\omega_m/\zeta_1} & \frac{1-j\omega_m/\zeta_2}{1+j\omega_m/\zeta_2} & \dots \end{bmatrix},$$

where ω is the angular frequency and ζ the relaxation frequencies. We choose m = 20, n = 50, L = 5, and N = 3. The matrix B and X_0 are generated in the same way as the noiseless case. Because of the noise, it is difficult to perfectly recover X_0 . In the noisy case, we declare a "perfect recovery" when all the N selected columns of A are correctly identified.

The empirical perfect recovery rate at different signal-tonoise ratios is shown in Fig. 3. The SNRs chosen are seemingly high, but fall in the normal range for the application of EMI sensors [3], where the estimation is sensitive to perturbation of the highly correlated columns of *A*. Both regularized M-IRL1 and M-FOCUSS are shown to robustly handle noise. The regularization parameter for M-FOCUSS is found via the same cross-validation-like technique described in [14]. The proposed M-IRL1 algorithm has higher exact recovery rates compared to M-FOCUSS, for this case. However, the M-FOCUSS algorithm takes less computation time. As a final comment, we note that regularized M-IRL1 has also been successfully applied to data taken in laboratory experiments and field tests.

In conclusion, we have shown that reweighting has advantages in recovering sparse signals using MMV problems. Though the noisy-data simulation presented is EMI specific, the proposed M-IRL1 method should be useful in other applications, such as the ones mentioned in Section 1.

4. REFERENCES

- J. Chen and X. Huo, "Theoretical results on sparse representations of multiple-measurement vectors," *IEEE Trans. Signal Process.*, vol. 54, no. 12, pp. 4634–4643, 2006.
- [2] S. F. Cotter and B. D. Rao, "Sparse channel estimation via matching pursuit with application to equalization," *IEEE Trans. Commun.*, vol. 50, no. 3, pp. 374–377, 2002.
- [3] M. H. Wei, W. R. Scott, Jr., and J. H. McClellan, "Robust estimation of the discrete spectrum of relaxations for electromagnetic induction responses," *IEEE Trans. Geosci. Remote Sens.*, vol. 48, no. 3, pp. 1169–1179, Mar. 2010.
- [4] J. A. Tropp, "Algorithms for simultaneous sparse approximation. Part II: Convex relaxation," *Signal Processing*, vol. 86, no. 3, pp. 589–602, 2006.
- [5] S. F. Cotter, B. D. Rao, K. Engan, and K. Kreutz-Delgado, "Sparse solutions to linear inverse problems with multiple measurement vectors," *IEEE Trans. Signal Process.*, vol. 53, no. 7, pp. 2477–2488, 2005.
- [6] J. A. Tropp, "Algorithms for simultaneous sparse approximation. Part I: Greedy Pursuit," *Signal Processing*, vol. 86, no. 3, pp. 572–588, 2006.
- [7] A. Lutoborski and V. N. Temlyakov, "Vector greedy algorithms," J. Complexity, vol. 19, no. 4, pp. 458–473, 2003.
- [8] M. Mishali and Y. C. Eldar, "Reduce and boost: Recovering arbitrary sets of jointly sparse vectors," *IEEE Trans. Signal Process.*, vol. 56, no. 10, pp. 4692–4702, 2008.
- [9] R. Chartrand and W. Yin, "Iteratively reweighted algorithms for compressive sensing," in *ICASSP*, Las Vegas, NV, Mar. 2008, pp. 3869–3872.
- [10] B. D. Rao and K. Kreutz-Delgado, "An affine scaling methodology for best basis selection," *IEEE Trans. Signal Process.*, vol. 47, no. 1, pp. 187–200, Jan. 1999.
- [11] I. F. Gorodnitsky and B. D. Rao, "Sparse signal reconstruction from limited data using focuss: A re-weighted minimum norm algorithm," *IEEE Trans. Signal Process.*, vol. 45, no. 3, pp. 600–616, Mar. 1997.
- [12] E. J. Candès, M. B. Wakin, and S. P. Boyd, "Enhancing sparsity by reweighted ℓ₁ minimization," *J. Fourier Anal. Appl.*, vol. 14, no. 5, pp. 877–905, 2008.
- [13] G. Harikumar and Y. Bresler, "A new algorithm for computing sparse solutions to linear inverse problems," in *ICASSP*. IEEE, 1996, vol. 3, pp. 1331–1334.
- [14] M. H. Wei, J. H. McClellan, and W. R. Scott, "Estimation of the discrete spectrum of relaxations for electromagnetic induction responses using ℓ_p -regularized least squares for $0 \le p \le 1$," *IEEE Trans. Geosci. Remote Sens.*, vol. 8, no. 2, pp. 233–237, Mar. 2011.