Hindawi Publishing Corporation
Abstract and Applied Analysis
Volume 2014, Article ID 410104, 6 pages
http://dx.doi.org/10.1155/2014/410104

Research Article

Nonmonotone Adaptive Barzilai-Borwein Gradient Algorithm for Compressed Sensing

Yuanying Qiu,[1] Jianlei Yan,[2] and Fanyong Xu[3]

[1] *School of Foreign Languages, Gannan Normal University, Ganzhou 341000, China*
[2] *School of Mathematics and Computer Sciences, Gannan Normal University, Ganzhou 341000, China*
[3] *Department of Radiology, The First Affiliated Hospital of Gannan Medical University, Ganzhou 341000, China*

Correspondence should be addressed to Jianlei Yan; majlyan@163.com

Received 24 January 2014; Accepted 17 March 2014; Published 7 April 2014

Academic Editor: Gaohang Yu

We study a nonmonotone adaptive Barzilai-Borwein gradient algorithm for l_1-norm minimization problems arising from compressed sensing. At each iteration, the generated search direction enjoys descent property and can be easily derived by minimizing a local approximal quadratic model and simultaneously taking the favorable structure of the l_1-norm. Under some suitable conditions, its global convergence result could be established. Numerical results illustrate that the proposed method is promising and competitive with the existing algorithms NBBL1 and TwIST.

1. Introduction

In recent years, algorithms for finding sparse solutions to underdetermined linear systems of equations have been intensively investigated in signal processing and compressed sensing. The fundamental principle of compressed sensing (CS) is that a sparse signal $\bar{x} \in R^n$ can be recovered from the underdetermined linear system $b = A\bar{x}$, where $A \in R^{m\times n}$ (often $m \ll n$). By defining l_0 norm ($\|x\|_0$) of a vector as the number of nonzero components in x, one natural way to reconstruct $\bar{x}$ from the system is to solve the following problem:

$$\min_{x\in R^n} \|x\|_0 \quad \text{s.t. } b = Ax \tag{1}$$

via certain reconstruction technique. However, the l_0 norm problem is combinatorial and generally computationally intractable. A fundamental decoding model in CS is to replace l_0 norm by l_1 norm, which is defined as $\|x\|_1 = \sum_{i=1}^{n} |x(i)|$. The resulting adaptation of (1) is the so-called basis pursuit (BP) problem [1]

$$\min_{x\in R^n} \|x\|_1 \quad \text{s.t. } b = Ax. \tag{2}$$

It is shown that, under some reasonable conditions, problem (2) can produce the desired solutions with high probability [2]. When b contains some noise in most practical applications, the constraint in (2) should be relaxed to the penalized least squares problem:

$$\min_{x\in R^n} \frac{1}{2}\|b - Ax\|_2^2 + \mu\|x\|_1. \tag{3}$$

Here, $\mu > 0$ is related to the Lagrange multiplier of the constraint in (2).

It follows from some existing results that if a signal is sparse or approximately sparse in some orthogonal basis, then an accurate recovery can be obtained when A is a random matrix projection [3]. Quite a number of algorithms have been proposed and studied for solving the aforementioned l_1-problems arising in CS. Recently, some first-order methods are popular for solving (3), such as the projection steepest descent method [4] and the gradient projection algorithm (GPSR) proposed by Figueiredo et al. [5]. Moreover, based on a smoothing technique studied in [6], a fast and accurate first-order algorithm called NESTA was proposed in [7], and so on. By an operator splitting technique, Hale et al. derive the iterative shrinkage/thresholding fixed-point continuation algorithm (FPC) [8]. One most widely studied first-order method is the iterative shrinkage/thresholding (IST) method [9–11], which is designed for wavelet-based

image deconvolution. TwIST [12, 13] and FISTA [14] speed up the performance of IST and have virtually the same complexity but with better convergence properties. Another closely related method is the sparse reconstruction algorithm SpaRSA [15], which is to minimize nonsmooth convex problem with separable structures. In [16], the authors proposed nonsmooth equations-based method for l_1-norm problems. SPGL1 [17] solves a least squares problem with l_1-norm constraint by the spectral gradient projection method with an efficient Euclidean projection on l_1-norm ball. In [18], Yun and Toh studied a block coordinate gradient descent (CGD) method for solving (3). Recently, the alternating direction method (ADM) has received much attention for solving total variation regularization problems for image restoration and is also capable of solving the l_1-norm regularization problems in CS [19, 20].

Very recently, Xiao et al. propose a Barzilai-Borwein gradient algorithm [21] for solving l_1 regularized nonsmooth minimization problems (NBBL1) [22], in which they approximate f locally by a convex quadratic model at each iteration, where the Hessian is replaced by the multiples of a spectral coefficient with an identity matrix. Motivated by them, we propose a nonmonotone adaptive Barzilai-Borwein gradient algorithm for l_1-norm minimization in compressed sensing, which is based on a new quasi-Newton equation [23] and a new adaptive spectral coefficient. Under reasonable assumptions, its convergence result could be established. Numerical experiments illustrate that the proposed method is efficient to recover a sparse signal arising in compressive sensing and outperforms NBBL1.

A full description of the proposed algorithm is presented in the next section. Meanwhile, we establish its global convergence under some suitable conditions. In Section 3, some numerical results were reported to illustrate the efficiency of the proposed method. Finally, we have a conclusion section.

2. Proposed Algorithm and Convergence Result

In this section, we construct an iterative algorithm to solve the l_1-norm regularization problems arising from the spare solution recovery in compressed sensing. Before stating the steps of our method, we first give a brief description of preliminary results for the following unconstrained optimization:

$$\min_{x\in\mathbb{R}^n} f(x), \tag{4}$$

where $f:\mathbb{R}^n \to \mathbb{R}$ is a continuously differentiable function.

In [23], Wei et al. proposed a new quasi-Newton equation and then derived a new conjugacy condition by using this new quasi-Newton equation. Using the Taylor formula for the objective function $f(x)$,

$$f(x) \approx f_k + g_k^T(x-x_k) + \frac{1}{2}(x-x_k)^T G_k (x-x_k), \tag{5}$$

where f_k (resp., G_k) denotes the function value (resp., Hessian matrix) and g_k denotes $\nabla f(x_k)$ at x_k. Hence, substituting $x = x_{k-1}$,

$$f_{k-1} \approx f_k - g_k^T s_{k-1} + \frac{1}{2} s_{k-1}^T G_k s_{k-1}. \tag{6}$$

Therefore,

$$\begin{aligned} s_{k-1}^T G_k s_{k-1} &\approx 2(f_{k-1}-f_k) + 2g_k^T s_{k-1} \\ &= 2(f_{k-1}-f_k) + (g_{k-1}+g_k)^T s_{k-1} + s_{k-1}^T y_{k-1}. \end{aligned} \tag{7}$$

Consider B_k as a new approximation of G_k such that

$$s_{k-1}^T B_k s_{k-1} = s_{k-1}^T y_{k-1} + \vartheta_{k-1}, \tag{8}$$

where $s_{k-1} = x_k - x_{k-1}$, $y_{k-1} = g_k - g_{k-1}$, $\bar{y}_{k-1} = y_{k-1} + (\vartheta_{k-1}/\|s_{k-1}\|^2)s_{k-1}$, and $\vartheta_{k-1} = 2(f_{k-1}-f_k)+(g_k+g_{k-1})^T s_{k-1}$. This suggests the following new quasi-Newton equation:

$$B_k s_{k-1} = \overline{y}_{k-1}. \tag{9}$$

In [24], Li et al. make a modification of the $\overline{y}_k$ in (9) as follows:

$$y_{k-1}^* = y_{k-1} + \frac{\max\{\vartheta_{k-1}, 0\}}{\|s_{k-1}\|^2} s_{k-1}, \tag{10}$$

and Yuan and Wei [25] make some further studies on it. Observe that this new quasi-Newton equation contains not only gradient value information but also function value information at the present and the previous step. In general, such B_{k+1} will be produced by updating B_k with some typical and popular formulae such as BFGS, DFP, and SR1. Furthermore, let the approximation Hessian B_k be a diagonal matrix with positive components; that is, $B_k = \lambda_k I$ with an identity matrix I and $\lambda_k > 0$. Then, the quasi-Newton condition (9) possesses the following form:

$$\lambda_k I s_{k-1} = y_{k-1}^*. \tag{11}$$

Multiplying both sides by s_{k-1}^T, it follows that

$$\lambda_k^{BB_1^*} = \frac{s_{k-1}^T y_{k-1}^*}{\|s_{k-1}\|_2^2}, \tag{12}$$

and, multiplying both sides by y_{k-1}^T, it gives

$$\lambda_k^{BB_2^*} = \frac{\|y_{k-1}^*\|_2^2}{s_{k-1}^T y_{k-1}^*}, \tag{13}$$

where y_{k-1}^* is defined as (10). If $s_{k-1}^T y_{k-1}^* > 0$ holds, then the matrix $\lambda_k I$ is positive definite which ensures that the search direction $-\lambda_k^{-1} g_k$ is descent at current point.

Now, we focus our attention on the l_1-norm minimization problem (3). The algorithm can be described as the iterative form

$$x_{k+1} = x_k + \alpha_k d_k, \tag{14}$$

Initialization: Give initial point $x_0 \in R^n$, set parameters $\mu > 0$, $\overline{\alpha} > 0$ and $\rho \in (0,1)$, $\delta \in (0,1)$, $h \in (0,1]$ and positive integer $\overline{m}$. Set $k = 0$.
Step 1. If $\|d_k\|_2 = 0$, then stop. Otherwise, continue.
Step 2. Compute d_k via (18).
Step 3. Compute α_k via following nonmonotone line search

$$F(x_k + \alpha_k d_k) \le \max_{0 \le j \le m(k)} F(x_{k-j}) + \delta\alpha_k \left(\nabla f(x_k)^T d_k + \frac{\mu\|x_k + hd_k\|_1 - \mu\|x_k\|_1}{h} \right),$$

where the smallest nonnegative integer j_k such as the stepsize $\alpha_k = \overline{\alpha}\rho^{j_k}$ satisfing above, $m(0) = 0$ and $0 \le m(k) \le \min\{m(k-1)+1, \overline{m}\}$.
Step 4. Let $x_{k+1} = x_k + \alpha_k d_k$.
Step 5. Let $k = k+1$. Go to Step 1.

Algorithm 1: NABBL1 for (3).

where α_k is a stepsize and d_k is a search direction defined by minimizing a quadratic approximated model of $F(x) = (1/2)\|Ax - b\|_2^2 + \mu\|x\|_1$. Since l_1-term is not differentiable, hence, at current x_k, objective function F is approximated by the quadratic approximation Q_k,

$$\begin{aligned} F(x_k + d) &= f(x_k + d) + \mu\|x_k + d\|_1 \\ &\approx f(x_k) + \nabla f(x_k)^T d + \frac{\lambda_k}{2}\|d\|_2^2 \\ &\quad + \mu\left[\|x_k\|_1 + \frac{\|x_k + hd\|_1 - \|x_k\|_1}{h}\right] \triangleq Q_k(d), \end{aligned} \tag{15}$$

where $f(x) = \|Ax - b\|_2^2$ and h is a small positive number. The term in $[\cdot]$ can be considered as an approximate Taylor expansion of $\|x_k + d\|_1$ with a small h, and the case $h = 1$ reduces the equivalent form $\|x_k + d\|_1$. Minimizing (15) yields

$$\begin{aligned} &\min_{d \in R^n} Q_k(d) \\ &\iff \min_{d \in R^n} \nabla f(x_k)^T d + \frac{\lambda_k}{2}\|d\|_2^2 + \frac{\mu\|x_k + hd\|_1}{h} \\ &\iff \min_{d \in R^n} \frac{h^2}{\lambda_k}\left(\nabla f(x_k)^T d + \frac{\lambda_k}{2}\|d\|_2^2 + \frac{\mu\|x_k + hd\|_1}{h}\right) \\ &\iff \min_{d \in R^n} \frac{1}{2}\left\|x_k + hd - \left(x_k - \frac{h}{\lambda_k}\nabla f(x_k)\right)\right\|_2^2 + \frac{\mu h}{\lambda_k}\|x_k + hd\|_1 \\ &\iff \min_{d \in R^n} \sum_{i=1}^{n} \left\{\frac{1}{2}\left(x_k^i + hd^i - \left(x_k^i - \frac{h}{\lambda_k}\nabla f^i(x_k)\right)\right)^2 \right. \\ &\qquad \left. + \frac{\mu h}{\lambda_k}\|x_k^i + hd^i\|_1\right\}, \end{aligned} \tag{16}$$

where x_k^i, d^i, and $\nabla f^i(x_k)$ denote the ith component of x_k, d, and $\nabla f(x_k)$, respectively. The favorable structure of (16) admits the explicit solution

$$\begin{aligned} x_k^i + hd_k^i &= \max\left\{\left|x_k^i - \frac{h}{\lambda_k}\nabla f^i(x_k^i)\right| - \frac{\mu h}{\lambda_k}, 0\right\} \\ &\quad \times \frac{x_k^i - (h/\lambda_k)\nabla f^i(x_k)}{|x_k^i - (h/\lambda_k)\nabla f^i(x_k)|}. \end{aligned} \tag{17}$$

Hence, the search direction at current point is

$$\begin{aligned} d_k = -\frac{1}{h}\Bigg[& x_k - \max\left\{\left|x_k^i - \frac{h}{\lambda_k}\nabla f^i(x_k^i)\right| - \frac{\mu h}{\lambda_k}, 0\right\} \\ & \times \frac{x_k^i - (h/\lambda_k)\nabla f^i(x_k)}{|x_k^i - (h/\lambda_k)\nabla f^i(x_k)|}\Bigg]. \end{aligned} \tag{18}$$

In this paper, we adopt the following adaptive Barzilai-Borwein step in (18):

$$\lambda_k^{ABB} = \begin{cases} \lambda_k^{BB_1^*}, & \text{if } \sqrt{\lambda_k^{BB_2^*}/\lambda_k^{BB_1^*}} > 0.9 \\ \lambda_k^{BB_2^*}, & \text{otherwise,} \end{cases} \tag{19}$$

where $\lambda_k^{BB_1^*}$ and $\lambda_k^{BB_2^*}$ are defined in (12) and (13), respectively.

In the light of all derivations above, we now describe the nonmonotone adaptive Barzilai-Borwein gradient algorithm (abbreviated as NABBL1) (see Algorithm 1).

Remark 1. We have shown that if $\lambda_k > 0$, then the generated direction is descent. However, in this case, the condition $\lambda_k > 0$ may fail to be fulfilled and the hereditary descent property is not guaranteed any more. To cope with this defect, we should keep the sequence $\{\lambda_k\}$ uniformly bounded; that is, for sufficiently small $\lambda_{\min} > 0$ and sufficiently large $\lambda_{\max} > 0$, the λ_k is forced as

$$\lambda_k = \min\{\lambda_{\max}, \max\{\lambda_k, \lambda_{\min}\}\}. \tag{20}$$

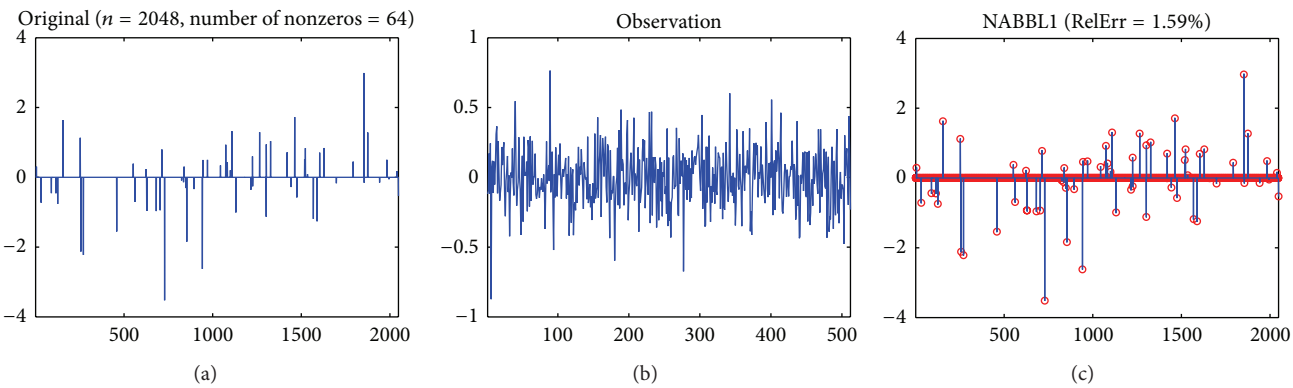

Figure 1: (a) Original signal with length of 2048 and 64 nonzero elements. (b) Noisy measurement with length of 512. (c) Recovered signal by NABBL1 (red circle) versus original signal (blue peaks).

This approach ensures that λ_k is bounded from zero and subsequently ensures that d_k is descent at per-iteration.

We prepare to show our main global convergence result of algorithm NABBL1. The desirable convergence is directly from Theorem 3.3 of [22]; we state it for completeness here.

Theorem 2. *Let the sequences $\{x_k\}$ and $\{d_k\}$ be generated by Algorithm 1. Then, there exists a subsequence K such that*

$$\lim_{k\to\infty, k\in K} \|d_k\|_2 = 0. \tag{21}$$

3. Experimental Results

In this section, we describe some experiments to illustrate the good performance of the algorithm NABBL1 for reconstructing sparse signals. These experiments are all tested in Matlab R2010a. The relative error is used to measure the quality of the reconstructive signals which is defined as

$$\text{RelErr} = \frac{\|\hat{x} - \bar{x}\|_2}{\|\bar{x}\|_2}, \tag{22}$$

where $\hat{x}$ denotes the reconstructive signal and $\bar{x}$ denotes the original signal.

In our experiments, we consider a typical compressive sensing scenario, where the goal is to reconstruct an n length sparse signal from m observations. The random A is the Gaussian matrix whose elements are generated from shape i.i.d. normal distributions $\mathcal{N}(0,1)$ (randn (m,n) in Matlab). In real applications, the measurement b is usually contaminated by noise; that is, $b = Ax + \eta$, where η is the Gaussian noise distributed as $\mathcal{N}(0,\sigma^2 I)$.

We test a small size signal with $n = 2^{11}$, $m = 2^9$; the original contains randomly $k = 2^6$ nonzero elements. The proposed algorithm starts at a zero point and terminates when the relative change of two successive points is sufficiently small; that is,

$$\frac{\|x_k - x_{k-1}\|_2}{\|x_{k-1}\|_2} < \tau. \tag{23}$$

In this experiment, we take $\tau = 10^{-4}$, $h = 0.8$, $\lambda_{\min} = 10^{-30}$, and $\lambda_{\max} = 10^{30}$. In the line search, we choose $\overline{\alpha} = 10^{-2}$, $\rho = 0.35$, $\delta = 10^{-4}$, and $\overline{m} = 5$. The original signal, the limited measurement, and the reconstructed signal when the noise level $\sigma^2 = 10^{-3}$ are given in Figure 1.

Comparing (a) to (c) in Figure 1, we clearly see that the original sparse signal is restored almost exactly. We see that all the blue peaks are circled by the red circles, which illustrates that the original signal has been found almost exactly. Altogether, this simple experiment shows that our algorithm performs quite well and provides an efficient approach to recover large sparse nonnegative signal.

We choose four different signals with noise level of $\sigma^2 = 10^{-3}$ compared with algorithms NBBL1 [22] and TwIST [13] in our next experiment. In order to test the speed of the algorithms more fairly, we list the average of the five results in Table 1. Numerical results are listed in Table 1, in which we report the CPU time in seconds (time) required for the whole reconstructing process and the relative error (RelErr). From Table 1, we can see that algorithm NABBL1 is faster than algorithms NBBL1 and TwIST, and the number of iterations of algorithm NABBL1 is less than that of the algorithms NBBL1 and TwIST with different signals.

From Figure 2, NABBL1 usually decreases relative errors faster than NBBL1 and TwIST throughout the entire iteration process. We conclude that NABBL1 provides an efficient approach for solving l_1 regularized nonsmooth problem from compressed sensing and is competitive with or performs better than NBBL1 and TwIST.

4. Conclusion

In this paper, we proposed a nonmonotone adaptive Barzilai-Borwein algorithm (NABBL1) for solving a l_1 regularized least squares problem arising from spare solution recovery in compressed sensing. At each iteration, the generated search direction enjoys descent property and can be easily derived by minimizing a local approximal quadratic model and simultaneously taking the favorable structure of the l_1-norm. Numerical results illustrate that the proposed method is promising and competitive with the existing algorithms

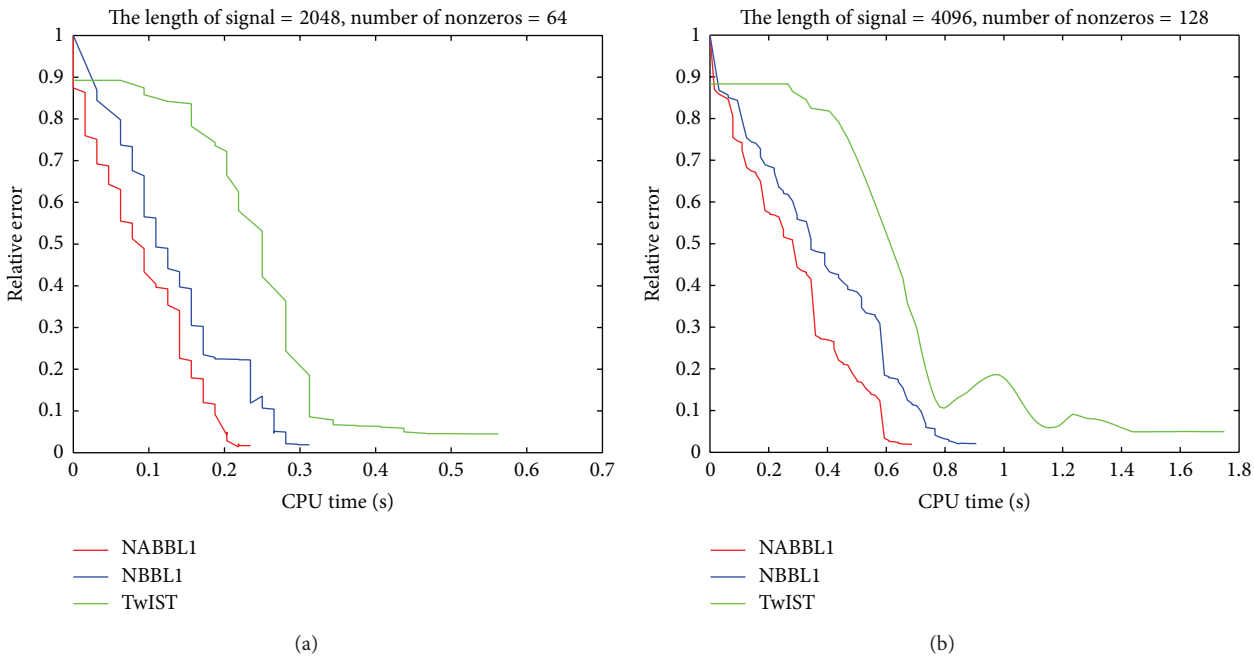

Figure 2: Comparison result of NABBL1, NBBL1, and TwIST with $\sigma^2 = 10^{-3}$. (a) Original signal with length of 2048 and 64 nonzero elements. (b) Original signal with length of 4096 and 128 nonzero elements. The x-axes represent the CPU time in seconds. The y-axes represent the relative error.

Table 1: Test results of NABBL1 and NBBL1 with different combinations of (n, m, k).

n	m	k	NABBL1		NBBL1		TwIST	
			Time	RelErr	Time	RelErr	Time	RelErr
8192	2048	256	2.9844	$1.73e-02$	3.5313	$1.97e-02$	4.0156	$4.83e-02$
4096	1024	128	0.6875	$1.93e-02$	0.9063	$2.09e-02$	1.7500	$4.96e-02$
2048	512	64	0.2344	$1.66e-02$	0.3125	$1.90e-02$	0.5625	$4.50e-02$
1024	256	32	0.0625	$1.55e-02$	0.0938	$2.23e-02$	0.1718	$4.54e-02$

NBBL1 and the two-step IST (TwIST). Our future topic is to extend NABBL1 method for solving matrix trace norm minimization problems in compressed sensing or some minimization problems in computed tomography reconstruction [26, 27].

Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

References

[1] S. S. Chen, D. L. Donoho, and M. A. Saunders, "Atomic decomposition by basis pursuit," *SIAM Journal on Scientific Computing*, vol. 20, no. 1, pp. 33–61, 1998.

[2] D. L. Donoho, "For most large underdetermined systems of linear equations the minimal l_1-norm solution is also the sparsest solution," *Communications on Pure and Applied Mathematics*, vol. 59, no. 6, pp. 797–829, 2006.

[3] M. F. Duarte and Y. C. Eldar, "Structured compressed sensing: from theory to applications," *IEEE Transactions on Signal Processing*, vol. 59, no. 9, pp. 4053–4085, 2011.

[4] I. Daubechies, M. Fornasier, and I. Loris, "Accelerated projected gradient method for linear inverse problems with sparsity constraints," *Journal of Fourier Analysis and Applications*, vol. 14, no. 5-6, pp. 764–792, 2008.

[5] M. A. T. Figueiredo, R. D. Nowak, and S. J. Wright, "Gradient projection for sparse reconstruction: application to compressed sensing and other inverse problems," *IEEE Journal on Selected Topics in Signal Processing*, vol. 1, no. 4, pp. 586–597, 2007.

[6] Y. Nesterov, "Smooth minimization of non-smooth functions," *Mathematical Programming*, vol. 103, no. 1, pp. 127–152, 2005.

[7] S. Becker, J. Bobin, and E. J. Candés, "NESTA: a fast and accurate first-order method for sparse recovery," *SIAM Journal on Imaging Sciences*, vol. 4, no. 1, pp. 1–39, 2011.

[8] E. T. Hale, W. Yin, and Y. Zhang, "Fixed-point continuation for l_1-minimization: methodology and convergence," *SIAM Journal on Optimization*, vol. 19, no. 3, pp. 1107–1130, 2008.

[9] M. A. T. Figueiredo and R. D. Nowak, "An EM algorthim for wavelet-besed image restoration," *IEEE Transactions on Image Processing*, vol. 12, no. 8, pp. 906–916, 2003.

[10] M. A. T. Figueiredo and R. D. Nowak, "A bound optimization approach to wavelet-based image deconvolution," in *Proceedings of the IEEE International Conference on Image Processing (ICIP '05)*, vol. 2, pp. 782–785, Genoa, Italy, September 2005.

[11] R. D. Nowak and M. A. T. Figueiredo, "Fast wavelet-based image deconvolution using the EM algorithm," in *Proceedings of the 35th Asilomar Conference on Signals, Systems and Computers*, vol. 1, pp. 371–375, Pacific Grove, Calif, USA, November 2001.

[12] J. M. Bioucas-Dias and M. A. T. Figueiredo, "A new TwIST: two-step iterative shrinkage/thresholding algorithms for image restoration," *IEEE Transactions on Image Processing*, vol. 16, no. 12, pp. 2992–3004, 2007.

[13] J. M. Bioucas-Dias and M. A. T. Figueiredo, "Two-step algorithms for linear inverse problems with non-quadratic regularization," in *Proceedings of the 14th IEEE International Conference on Image Processing (ICIP '07)*, vol. 1, pp. 105–108, San Antonio, Tex, USA, September 2007.

[14] A. Beck and M. Teboulle, "A fast iterative shrinkage-thresholding algorithm for linear inverse problems," *SIAM Journal on Imaging Science*, vol. 2, no. 1, pp. 183–202, 2009.

[15] S. J. Wright, R. D. Nowak, and M. A. T. Figueiredo, "Sparse reconstruction by separable approximation," *IEEE Transactions on Signal Processing*, vol. 57, no. 7, pp. 2479–2493, 2009.

[16] Y.-H. Xiao, Q.-Y. Wang, and Q.-J. Hu, "Non-smooth equations based method for l_1-norm problems with applications to compressed sensing," *Nonlinear Analysis: Theory, Methods & Applications*, vol. 74, no. 11, pp. 3570–3577, 2011.

[17] M. P. Friedlander and E. van den Berg, "Probing the pareto frontier for basis pursuit solutions," *SIAM Journal on Scientific Computing*, vol. 31, no. 2, pp. 890–912, 2008.

[18] S. Yun and K.-C. Toh, "A coordinate gradient descent method for l_1-regularized convex minimization," *Computational Optimization and Applications*, vol. 48, no. 2, pp. 273–307, 2011.

[19] J. Yang, Y. Zhang, and W. Yin, "A fast alternating direction method for TVL1-L2 signal reconstruction from partial Fourier data," *IEEE Journal of Selected Topics in Signal Processing*, vol. 4, no. 2, pp. 288–297, 2010.

[20] J. Yang and Y. Zhang, "Alternating direction algorithms for l_1-problems in compressive sensing," *SIAM Journal on Scientific Computing*, vol. 33, no. 1, pp. 250–278, 2011.

[21] J. Barzilai and J. M. Borwein, "Two-point step size gradient methods," *IMA Journal of Numerical Analysis*, vol. 8, no. 1, pp. 141–148, 1988.

[22] Y.-H. Xiao, S.-Y. Wu, and L.-Q. Qi, "Nonmonotone Barzilai-Borwein gradient algorithm for l_1-regularized nonsmooth minimization in compressive sensing," *Journal of Scientific Computing*, 2014.

[23] Z. Wei, G. Li, and L. Qi, "New quasi-Newton methods for unconstrained optimization problems," *Applied Mathematics and Computation*, vol. 175, no. 2, pp. 1156–1188, 2006.

[24] G. Li, C. Tang, and Z. Wei, "New conjugacy condition and related new conjugate gradient methods for unconstrained optimization," *Journal of Computational and Applied Mathematics*, vol. 202, no. 2, pp. 523–539, 2007.

[25] G. Yuan and Z. Wei, "Convergence analysis of a modified BFGS method on convex minimizations," *Computational Optimization and Applications*, vol. 47, no. 2, pp. 237–255, 2010.

[26] J. Ma, J. Huang, H. Zhang et al., "Low-dose computed tomography image restoration using previous normal-dose scan," *Medical Physics*, vol. 38, no. 10, pp. 5713–5731, 2011.

[27] J. Ma, H. Zhang, Y. Gao et al., "Iterative image reconstruction for cerebral perfusion CT using a pre-contrast scan induced edge-preserving prior," *Physics in Medicine and Biology*, vol. 57, no. 22, pp. 7519–7542, 2012.