Journal of Inequalities and Applications
a SpringerOpen Journal

RESEARCH

Open Access

The modified accelerated Bregman method for regularized basis pursuit problem

Yajun Xie[1,2], Yifen Ke[1] and Changfeng Ma[1*]

*Correspondence:
macf@fjnu.edu.cn
[1]School of Mathematics and
Computer Science, Fujian Normal
University, Fuzhou, 350007,
P.R. China
Full list of author information is
available at the end of the article

Abstract

In this paper, a modified accelerated Bregman method (MABM) for solving the regularized basis pursuit problem is considered and analyzed in detail. This idea is based on the fact that the linearized Bregman method (LBM) proposed by Osher *et al.* (Multiscale Model. Simul. 4(2):460-489, 2005) is equivalent to a gradient descent method applied to a certain dual formulation which converges to the solution of the regularized basis pursuit problem. The proposed method is based on an extrapolation technique which is used in accelerated proximal gradient methods presented by Nesterov (Dokl. Akad. Nauk SSSR 269:543-547, 1983). It is verified that the modified accelerated Bregman method (MABM) is equivalent to the corresponding accelerated augmented Lagrangian method (AALM). The theoretical results confirm that the method has a rapid convergence rate of $O(1/(k+1)^2)$.

Keywords: compressed sensing; linear constraint; Bregman method; modified accelerated method; convergence rate

1 Introduction

Compressed sensing, an interesting research field involving how to obtain information, acts as a crucial role in signal processing, image restoration, *etc.* The origin of its name is based on the idea of encoding a large sparse signal exploiting a relatively small number of linear measurements, then decoding the signal either through minimizing the 1-norm or utilizing a combinational algorithm, a greedy algorithm, *etc.* It resulted from elementary analysis and approximation theory by Kashin [1], but was brought into the forefront by the work of Candés *et al.* [2–4] and Donoho [5] who constructed specific methods and showed their application prospects. The concrete model considered in compressed sensing is the so-called basis pursuit problem [6, 7]

$$\min_{x \in \Re^n} \|x\|_1 \quad \text{subject to } Ax = b, \tag{1.1}$$

where $A \in \Re^{m \times n}$, $b \in \Re^m$. It can be deemed a classical transform in the compressed sensing field into the following NP-hard discrete optimization problem:

$$\min_{x \in \Re^n} \|x\|_0 \quad \text{subject to } Ax = b, \tag{1.2}$$

where $A \in \Re^{m \times n}$, $b \in \Re^m$, $\| \cdot \|_0$ denotes the number of nonzero elements.

In many practical circumstances, the data or signals are frequently represented by a few matrices, which, in fact, is convenient for data processing and analysis. However, the data

are often given with damage, loss or noise pollution, *etc.* In this case, how to restore the original data is among the practical difficulties faced when solving the matrix reconstruction problem. Being similar to handling the original signal (vector form) in the compressed sensing field, matrix remodeling refers to an original compressible or sparse representation matrix which can be exactly or approximately reconstructed with a proper model. The matrix reconstruction problem can be classified with matrix completion (MC) [8, 9] and matrix recovery (MR) [10, 11], which is met in various areas, *e.g.*, a famous application is the Netflix system [12] in matrix completion; while the matrix recovery problem originates from face image processing, background modeling and so forth, it can be applied to image alignment as by Peng *et al.* [13]. Generally speaking, the matrix reconstruction problem is written as

$$\min \text{rank}(X) \quad \text{subject to } X \in C, \tag{1.3}$$

where $X \in \Re^{m \times n}$ and C is a convex set. Normally, the following affine constrained minimization problem can be considered:

$$\min \text{rank}(X) \quad \text{subject to } A(X) = b, \tag{1.4}$$

where the linear map $A : \Re^{m \times n} \to \Re^d$, $b \in \Re^d$ is a given vector.

The model emerges in many fields, such as determining a low-order controller for a plant [14] and a minimum order linear system realization [15], and solving low-dimensional Euclidean embedding problems [16]. A natural generalization in the sense of compressed sensing, (1.4) can be formulated as the nuclear norm minimization problem

$$\min \|X\|_* \quad \text{subject to } A(X) = b, \tag{1.5}$$

where $\| \cdot \|_*$ denotes the sum of singular values of the corresponding matrix. The matrix completion problem

$$\min \text{rank}(X) \quad \text{subject to } X_{i,j} = M_{i,j}, (i,j) \in \Omega, \tag{1.6}$$

where $X \in \Re^{m \times n}$, $M \in \Re^{m \times n}$, Ω is a certain index set, is considered as a special case of (1.3). The idea is inspired by the method presented in [17, 18] *etc.* Now, we consider the so-called regularized basis pursuit problem

$$\min_{x \in \Re^n} \|x\|_1 + \frac{\mu}{2} \|x\|_2^2 \quad \text{subject to } Ax = b. \tag{1.7}$$

Adding the penalized term of the constraint condition to the objection function of the regularized basis pursuit problem, we obtain

$$\min_{x \in \Re^n} \|x\|_1 + \frac{\mu}{2} \|x\|_2^2 + \frac{\lambda}{2} \|Ax - b\|_2^2. \tag{1.8}$$

Particularly in [18], putting in the term in (1.7) yields the tractable object function, which is a strictly convex function. Thus, the linearly constrained basis pursuit problem has a sole solution and its dual problem is smooth. When μ is sufficiently small, the solution to (1.7)

is also the solution of (1.1). This exact regularization property of (1.7) was studied in [19, 20]. Problem (1.7) has a 2-norm in the regularizer term, thus it is considered less sensitive to noise than the basis pursuit problem (1.1).

To solve (1.1), a linearized Bregman iteration method was presented in [21], which was motivated by [22]. The idea of the linearized Bregman method is to combine a fixed point iteration with the Bregman method in [23, 24], but the linearized Bregman method is a bit slow in the convergence result. Hence, many accelerated schemes have been brought up in various theses. For example, in [21] the kicking technique was introduced. Yin [19] verified that the linearized Bregman method is equivalent to a gradient descent method applied to the Lagrangian dual problem of (1.7). They improved the linearized Bregman method, utilizing the Barzilai-Borwein line of searching [25], nonlinear conjugate gradient methods, and the method of limited memory BFGS [26]. Huang *et al.* [18] proposed an accelerated linearized Bregman method which is based on the fact that the linearized Bregman method is equivalent to the gradient descent method applied to a Lagrangian dual of problem (1.7) and the extrapolation technique, which is adopted in the accelerated proximal gradient methods [27] proposed by Nesterov *et al.* To solve problem (1.7), Goldstein *et al.* [28] used an alternating split technique and its Lagrange dual problem.

Based on these studies, we extend the accelerated Bregman method to solve (1.7) in which the object function might be not differentiable but have the 'good' performance (convex and continuous). We put forward a new improvement formula based on the accelerated Bregman method. It can be proved to have the property that the modified Bregman method is equivalent to the corresponding accelerated augmented Lagrangian method, and the latter has a rapid convergence rate which can be deemed to be an improvement of [17].

The rest of this article is organized as follows. In Section 2, we sketch the original Bregman method and the linearized Bregman method which are useful for the subsequent analysis. In Section 3, we introduce the accelerated augmented Lagrangian method (AALM), and we present our modified accelerated Bregman method (MABM). Section 4 is devoted to the convergence of the regularized basis pursuit problem and here we analyze the error bound of the MABM in detail. In Section 5, we give some conclusions and discuss the research plans for our future work.

2 The original Bregman method and its linearized form

The Bregman method was introduced into image processing by Osher *et al.* in [23] for solving the total-variation (TV)-based image restoration problems. Let $\varphi(\cdot)$ be a Bregman function, the Bregman distance [29] of the point u to the point v is defined as

$$D_\varphi^p(u, v) := \varphi(u) - \varphi(v) - \langle p, u - v \rangle, \tag{2.1}$$

where $p \in \partial\varphi(v)$, the subdifferential of φ at v. Note that updating the formula for (1.7), this Bregman iterative regularization procedure recursively solves

$$x^{k+1} := \arg\min_x D_\varphi^{p^k}(x, x^k) + \frac{\lambda}{2}\|Ax - b\|_2^2 \tag{2.2}$$

for $k = 0, 1, \ldots$, starting with $x^0 = 0$, $p^0 = 0$. Since (2.2) is a convex programming problem, the optimality conditions can be given by

$$0 \in \partial\varphi(x^{k+1}) - p^k - \lambda A^T(b - Ax^{k+1}), \tag{2.3}$$

from which we obtain the updated formula

$$p^{k+1} := p^k + \lambda A^T(b - Ax^{k+1}). \tag{2.4}$$

Hence, the Bregman iterative scheme is given by

$$\begin{cases} x^{k+1} := \arg\min_x D_\varphi^{p^k}(x, x^k) + \frac{\lambda}{2}\|Ax - b\|_2^2, \\ p^{k+1} := p^k + \lambda A^T(b - Ax^{k+1}). \end{cases} \tag{2.5}$$

By the following lemma, we get the equivalent form of (2.5)

$$\begin{cases} x^{k+1} := \arg\min_x \varphi(x) + \frac{\lambda}{2}\|Ax - b^k\|_2^2, \\ b^{k+1} := b^k + (b - Ax^{k+1}), \end{cases} \tag{2.6}$$

starting with $x^0 = 0$, $b^0 = b$.

For the sake of the requirement for the whole theoretical analysis, in the following, we propose some significant equivalence results, and we give the detailed proofs.

Lemma 2.1 *The Bregman iterative scheme, which is given by* (2.5), *will obtain the same optimum point in the first term as* (2.6) *if*

$$p^k = \lambda A^T(b^k - b) \tag{2.7}$$

holds, where λ is a certain positive constant.

Proof From the first formula of (2.5) and the definition of Bregman distance (2.1), we get

$$\begin{aligned}
x^{k+1} &= \arg\min_x \varphi(x) - \varphi(x^k) - \langle p^k, x - x^k \rangle + \frac{\lambda}{2}\|Ax - b\|_2^2 \\
&= \arg\min_x \varphi(x) - \langle \lambda A^T(b^k - b), x \rangle + \frac{\lambda}{2}\|Ax - b\|_2^2 \\
&= \arg\min_x \varphi(x) - \lambda\langle b^k - b, Ax - b \rangle + \frac{\lambda}{2}\|Ax - b\|_2^2 + \frac{\lambda}{2}\|b^k - b\|_2^2 \\
&= \arg\min_x \varphi(x) - \frac{\lambda}{2}\|Ax - b^k\|_2^2,
\end{aligned}$$

therefore, we complete the proof. □

From the discussion above, we can give a crucial conclusion.

Theorem 2.2 *The original Bregman iterative scheme* (2.5) *is equivalent to its variant* (2.6).

Proof By induction, in fact, we only need to verify that (2.7) holds.

If $k = 0$, $p^0 = \lambda A^T(b^0 - b) = 0$, (2.5) holds by the initial conditions $p^0 = 0$ and $b^0 = b$.

Now, we suppose that (2.7) holds for $k - 1$. Then

$$\begin{aligned}
p^k &= p^{k-1} + \lambda A^T(b - Ax^k) \\
&= \lambda A^T(b^{k-1} - b) + \lambda A^T(b - Ax^k)
\end{aligned}$$

$$= \lambda A^T \left(b^{k-1} - b + b - Ax^k \right)$$
$$= \lambda A^T \left(b^k - b \right),$$

where the first equality is from the second term of (2.5), and the third equality is from the second term of (2.6). Therefore, the original Bregman iterative scheme (2.5) is equivalent to its variant (2.6). □

Noting that [24, 30, 31] and the references therein, it was argued that the Bregman iterative method is equivalent to the augmented Lagrangian method. The significance of this statement will be demonstrated in our later analysis in Section 4.

For the following analysis, we define the Lagrangian function of (1.7),

$$L(x, \sigma) := \varphi(x) + \langle \sigma, b - Ax \rangle,$$

where $\sigma \in \Re^m$ is the Lagrangian multiplier. The corresponding augmented Lagrangian function can be expressed as

$$L(x, \sigma, \lambda) := \varphi(x) + \langle \sigma, b - Ax \rangle + \frac{\lambda}{2} \|Ax - b\|_2^2, \tag{2.8}$$

where $\varphi(x) := \|x\|_1 + \frac{\mu}{2} \|x\|_2^2$ and $\frac{\lambda}{2} \|Ax - b\|_2^2$ is the penalty term.

The augmented Lagrangian iterative scheme is

$$\begin{cases} x^{k+1} := \arg\min_x \varphi(x) + \langle \sigma^k, b - Ax \rangle + \frac{\lambda}{2} \|Ax - b\|_2^2, \\ \sigma^{k+1} := \sigma^k + \lambda(b - Ax^{k+1}), \end{cases} \tag{2.9}$$

starting from $\sigma^0 = 0$.

Lemma 2.3 *The first item of the iterative sequence $\{x^k\}$ of* (2.9) *and that of* (2.6) *are equal if*

$$b^k = b + \frac{1}{\lambda}\sigma \tag{2.10}$$

holds.

Proof From the first formula of (2.6) and (2.10), we get

$$x^{k+1} = \arg\min_x \varphi(x) + \frac{\lambda}{2} \left\| Ax - b - \frac{\sigma^k}{\lambda} \right\|_2^2$$
$$= \arg\min_x \varphi(x) + \frac{\lambda}{2} \|Ax - b\|_2^2 + \langle \sigma^k, b - Ax \rangle.$$

The constant of the objective function does not affect the optimum point, so by comparing with the above equations, we get the conclusion. □

Theorem 2.4 *The augmented Lagrangian iterative scheme* (2.9) *is equivalent to the Bregman iterative method variant* (2.6).

Proof It is not difficult to see that the proof is the same as that of Theorem 2.2. Similarly, by the mathematical induction, we simply show that (2.7) holds.

If $k = 0$, $b^0 = b + \frac{\sigma^0}{\lambda}$, (2.7) holds by the initial conditions $\sigma^0 = 0$ and $b^0 = b$. Now, we suppose that (2.7) holds for $n \le k - 1$. Then, when $n = k$,

$$b^k = b^{k-1} + \left(b - Ax^k\right) = b + \frac{1}{\lambda}\sigma^{k-1} + \left(b - Ax^k\right)$$

$$= b + \frac{1}{\lambda}\sigma^{k-1} + \frac{1}{\lambda}\left(\sigma^k - \sigma^{k-1}\right) = b + \frac{1}{\lambda}\sigma^k,$$

where the first equality is from the second term of (2.6), the second equality has its roots in induction, and the third equality is derived from the second term of (2.9). Therefore, the augmented Lagrangian iterative scheme (2.9) is equivalent to the Bregman iterative method variant (2.6). Moreover, we can get the equivalence of (2.5) and (2.9) from the two theorems above. □

For solving the subproblem of Bregman iterative method (2.2), different algorithms were conceived in [27, 32, 33], however, these methods have to perform significant iterations for solving the subproblem. In order to overcome the difficulties, the linearized Bregman method is obtained by linearizing the last term in (2.2) into $\langle A^T(Ax^k - b), x\rangle$ (where $\lambda = 1$) and adding the l_2-proximity term $\frac{1}{2\gamma}\|x - x^k\|^2$. The concrete scheme is

$$\begin{cases} x^{k+1} := \arg\min_x D_\varphi^{p^k}(x, x^k) + \langle A^T(Ax^k - b), x\rangle + \frac{1}{2\gamma}\|x - x^k\|_2^2, \\ p^{k+1} := p^k + A^T(Ax^{k+1} - b) - \frac{1}{\gamma}(x - x^k). \end{cases} \tag{2.11}$$

By the optimality conditions, we obtain the following formula:

$$0 \in \partial\varphi\left(x^{k+1}\right) - p^k - A^T\left(b - Ax^k\right) + \frac{1}{\gamma}\left(x^{k+1} - x^k\right).$$

That is, for the second term of (2.11) $p^{k+1} \in \partial\varphi(x^{k+1})$. In [7, 19, 21], the linearized Bregman method was analyzed; as $\gamma < \frac{2}{\|A\|^2}$ (where $\|\cdot\|$ denotes the spectral norm of the corresponding matrix) the iterates of the linearized Bregman method converge to the solution of the regularized basis pursuit problem

$$\min_{x\in\Re^n}\|x\|_1 + \frac{1}{2\gamma}\|x\|_2^2 \quad \text{subject to } Ax = b.$$

It has appeared in the accelerated Bregman algorithm in recent years, such as in [18] *etc.* But we shall argue that the accelerated Bregman method has large room to advance. We give some valid improvements on these accelerated Bregman methods. It can be verified that the theoretical result as regards the proposed method has a rapid convergence rate of $O(1/(k + 1)^2)$.

3 The modified accelerated Bregman method

In [24], the linearized Bregman algorithm could be written as follows.

Algorithm 3.1 (Linearized Bregman method (LBM))

Step 0. Input: $J(\cdot), H(\cdot), \gamma > 0$; initial point: x^0, p^0.

Step 1. Initialize: $k = 0$, let $x^0 = 0$ and $p^0 = 0$.

Step 2. Compute $x^{k+1} := \arg\min_x D_\varphi^{p^k}(x, x^k) + \langle \nabla H(x^k), x \rangle + \frac{1}{2\gamma} \|x - x^k\|_2^2$.

Step 3. Set $p^{k+1} := p^k - \nabla H(x^k) - \frac{1}{\gamma}(x^{k+1} - x^k)$.

Step 4. Set $k := k + 1$, go to step 2.

For the iterative scheme above, $H(x) = \frac{1}{2}\|Ax - b\|_2^2$. Next, we give its equivalence form from the following lemma.

Lemma 3.2 *The linearized Bregman method* (LBM) *in Algorithm* 3.1 *is equivalent to the iterative scheme*

$$\begin{cases} x^{k+1} := \arg\min_x \varphi(x) + \frac{1}{2\gamma}\|x - \gamma u^k\|_2^2, \\ u^{k+1} := u^k + A^T(b - Ax^{k+1}), \end{cases} \tag{3.1}$$

starting from $u^0 = A^T b$, *where* $u^k = p^k + A^T(b - Ax^k) + \frac{x^k}{\gamma}$.

Proof From step 3 of Algorithm 3.1, we have

$$p^{k+1} = p^k + A^T(b - Ax^k) - \frac{1}{\gamma}(x^{k+1} - x^k) = \cdots = \sum_{i=0}^{k} A^T(b - Ax^i) - \frac{1}{\gamma}x^{k+1}. \tag{3.2}$$

Thus, we denote

$$u^k := p^{k+1} + \frac{1}{\gamma}x^{k+1} = p^k + A^T(b - Ax^k) + \frac{1}{\gamma}x^k = \sum_{i=0}^{k} A^T(b - Ax^i);$$

then we simplify step 2 of Algorithm 3.1,

$$x^{k+1} = \arg\min_x \varphi(x) + \langle p^k, x \rangle - \langle A^T(b - Ax^k), x \rangle + \frac{1}{2\gamma}\|x - x^k\|_2^2$$

$$= \arg\min_x \varphi(x) + \frac{1}{2\gamma}\left\|x - \gamma\left[p^k + A^T(b - Ax^k) + \frac{x^k}{\gamma}\right]\right\|_2^2$$

$$= \arg\min_x \varphi(x) + \frac{1}{2\gamma}\|x - \gamma u^k\|_2^2.$$

On the other hand,

$$u^{k+1} = p^{k+1} + A^T(b - Ax^{k+1}) + \frac{1}{\gamma}x^{k+1}$$

$$= p^k + A^T(b - Ax^k) - \frac{1}{\gamma}(x^{k+1} - x^k) + A^T(b - Ax^{k+1}) + \frac{x^{k+1}}{\gamma}$$

$$= p^k + A^T(b - Ax^k) + \frac{1}{\gamma}x^k + A^T(b - Ax^{k+1})$$

$$= u^k + A^T(b - Ax^{k+1}),$$

where the second equality is from (3.2). So, step 2 and step 3 in Algorithm 3.1 can be rewritten as (3.1). □

Yin *et al.* presented all kinds of techniques, for example, line search, and L-BFGS and BB steps, to accelerate the linearized Bregman method. It is interesting that for the latter in [18] the accelerated linearized Bregman algorithm is argued for as follows.

Algorithm 3.3 (Accelerated linearized Bregman method (ALBM))

Step 0. Initialize: $x^0 = \widetilde{x}^0 = p^0 = \widetilde{p}^0 = 0$, $\gamma > 0$, $\lambda > 0$, $k = 0$.

Step 1. Compute $x^{k+1} := \arg\min_x D_\varphi^{\widetilde{p}^k}(x, \widetilde{x}^k) + \lambda\langle A^T(A\widetilde{x}^k - b), x\rangle + \frac{1}{2\gamma}\|x - \widetilde{x}^k\|_2^2$.

Step 2. Set $p^{k+1} := \widetilde{p}^k - \lambda A^T(A\widetilde{x}^k - b) - \frac{1}{\gamma}(x^{k+1} - \widetilde{x}^k)$.

Step 3. Set $\widetilde{x}^{k+1} := \alpha_k x^{k+1} + (1 - \alpha_k)x^k$.

Step 4. Set $\widetilde{p}^{k+1} := \alpha_k p^{k+1} + (1 - \alpha_k)\widetilde{x}^k$.

Step 5. Set $k := k + 1$, go to step 1.

These methods motivated us to consider the following algorithm, and its acceleration idea is based on the extrapolation technique proposed by Nesterov in [34, 35]; see also the references therein.

Algorithm 3.4 (Modified accelerated Bregman method (MABM))

Step 0. Initialize: $x^0 = p^0 = \widetilde{p}^0 = 0$, $t_0 = 1$, $\lambda > 0$, $k = 0$.

Step 1. Compute $x^{k+1} := \arg\min_x D_\varphi^{p^k}(x, x^k) + \frac{\lambda}{2}\|Ax - b\|_2^2$.

Step 2. Set $\widetilde{p}^{k+1} := p^k + \lambda A^T(b - Ax^k)$.

Step 3. Set $p^{k+1} := (1 - \frac{1-2t_k}{t_{k+1}})\widetilde{p}^{k+1} + \frac{1-2t_k}{t_{k+1}}p^k$.

Step 4. Set $t_{k+1} = \frac{1}{3}(\sqrt{1 + 9t_k^2} + 1)$.

Step 5. Set $k := k + 1$, go to step 2.

The basic idea of the equivalence between MABA and the corresponding AALM can be traced back to [31]. Especially, we can see that our updated iterative for p^k is obviously better than the pre-iterative $\widetilde{p}^k$ [17], since we consider a sufficient amount of information about the former iterative. In this way, most of the better iterative efficiency could be expected, which is just our purpose in improving the method. Then we will be dependent on a series of transformations in preparation for the convergence proof in Section 4.

Lemma 3.5 *The MABM in Algorithm* 3.4 *is equal to the following iterative scheme*:

$$\begin{cases} x^{k+1} := \arg\min_x \varphi(x) + \frac{\lambda}{2}\|Ax - b^k\|_2^2, \\ \widetilde{b}^{k+1} := b^k + (b - Ax^{k+1}), \\ b^{k+1} := (1 - \frac{1-2t_k}{t_{k+1}})\widetilde{b}^{k+1} + \frac{1-2t_k}{t_{k+1}}b^k, \end{cases} \tag{3.3}$$

starting with $b^0 = b$.

Proof Recalling (2.7) and noting Theorem 2.2, we can prove that (2.7) holds for all k by induction.

If $k = 0$, $p^0 = \lambda A^T(b^0 - b) = 0$, (2.7) holds by the initial conditions $p^0 = 0$ and $b^0 = b$. Now, we assume that (2.7) holds for $n \leq k - 1$, then

$$\widetilde{p}^n = p^{n-1} + \lambda A^T(b - Ax^n) = \lambda A^T(b^{n-1} - b) + \lambda A^T(b - Ax^n)$$

$$= \lambda A^T(b^{n-1} - b + b - Ax^n) = \lambda A^T(\widetilde{b}^n - b), \tag{3.4}$$

where the first equality is directly derived from step 2 of Algorithm 3.4, the first equality is derived from the induction hypothesis, the fourth equality utilizes the second step of (3.3).

Moreover, when $n = k$, we have

$$
\begin{aligned}
p^k &= \left(1 - \frac{1 - 2t_k}{t_{k+1}}\right)\widetilde{p}^k + \frac{1 - 2t_k}{t_{k+1}}p^{k-1} \\
&= \left(1 - \frac{1 - 2t_k}{t_{k+1}}\right)\lambda A^T\left(\widetilde{b}^k - b\right) + \frac{1 - 2t_k}{t_{k+1}}\lambda A^T\left(b^{k-1} - b\right) \\
&= \lambda A^T\left[\left(1 - \frac{1 - 2t_k}{t_{k+1}}\right)\widetilde{b}^k - \left(1 - \frac{1 - 2t_k}{t_{k+1}}\right)b + \frac{1 - 2t_k}{t_{k+1}}b^{k-1} - \frac{1 - 2t_k}{t_{k+1}}b\right] \\
&= \lambda A^T\left(b^k - b\right),
\end{aligned}
$$

where the first equality is from step 3 of Algorithm 3.4, the second term is from (3.4) and the induction hypothesis, the fourth equality is from the third term of (3.3), so (2.7) holds for all k. Namely, the MABM in Algorithm 3.4 is equal to (3.3). □

Lemma 3.6 *Iterative scheme* (3.3) *is equal to the following AALM iterative scheme*:

$$
\begin{cases}
x^{k+1} := \arg\min_x \varphi(x) + \langle \sigma^k, b - Ax \rangle + \frac{\lambda}{2}\|Ax - b^k\|_2^2, \\
\widetilde{\sigma}^{k+1} := \sigma^k + (b - Ax^{k+1}), \\
\sigma^{k+1} := (1 - \frac{1 - 2t_k}{t_{k+1}})\widetilde{\sigma}^{k+1} + \frac{1 - 2t_k}{t_{k+1}}\sigma^k,
\end{cases}
\tag{3.5}
$$

starting from $\sigma^0 = 0$, *where* $\sigma^k \in \Re^m$ *is the Lagrangian multiplier*.

Proof It is not difficult to see that the idea has likeness to Theorem 2.4, we are just required to verify that (2.7) holds. To this end, we proceed by mathematical induction.

If $k = 0$, $b^0 = b + \frac{\sigma^0}{\lambda}$, (2.7) holds by the initial conditions $\sigma^0 = 0$ and $b^0 = b$. Suppose that (2.7) holds for $n \leq k - 1$, then, when $n = k$, we have

$$
\begin{aligned}
\widetilde{b}^k &= b^{k-1} + \left(b - Ax^k\right) = b + \frac{1}{\lambda}\sigma^{k-1} + \left(b - Ax^k\right) \\
&= b + \frac{1}{\lambda}\sigma^{k-1} + \frac{1}{\lambda}\left(\widetilde{\sigma}^k - \sigma^{k-1}\right) = b + \frac{1}{\lambda}\widetilde{\sigma}^k,
\end{aligned}
$$

where the first equality is from the second term of (3.3), the second equality stems from induction, and the third equality is derived from the second term of (3.5).

Thus, we get

$$
\begin{aligned}
b^k &= \left(1 - \frac{1 - 2t_k}{t_{k+1}}\right)\widetilde{b}^k + \frac{1 - 2t_k}{t_{k+1}}b^{k-1} \\
&= \left(1 - \frac{1 - 2t_k}{t_{k+1}}\right)\left(b + \frac{1}{\lambda}\widetilde{\sigma}^n\right) + \frac{1 - 2t_k}{t_{k+1}}\left(b + \frac{1}{\lambda}\sigma^{n-1}\right) \\
&= b + \frac{1}{\lambda}\left(\frac{1 - 2t_k}{t_{k+1}}\widetilde{\sigma}^n + \frac{1 - 2t_k}{t_{k+1}}\sigma^{n-1}\right) \\
&= b + \frac{1}{\lambda}\sigma^n.
\end{aligned}
$$

Therefore, AALM (3.5) is equivalent to iterative scheme (3.3). Moreover, we can get the equivalence of the MABM in Algorithm 3.4 to AALM from the above two lemmas. □

Theorem 3.7 *The MABM in Algorithm* 3.4 *is equivalent to the corresponding AALM* (3.5).

4 The convergence analysis

A practical challenge for the regularized basis pursuit problem is to offer an efficient method to solve the non-smooth optimization problems. Many algorithms have been proposed in recent years [36]. In these methods, some schemes of approximation that have to do with the non-smooth norm term are usually employed. However, a fast global convergence is difficult to guarantee. Due to the non-smooth nature of the 1-norm, a simple method to solve these problems is the subgradient approach [37], which converges only as $O(\frac{1}{\sqrt{k}})$, where k is the iteration counter.

In this paper, we present an efficient method with fast global convergence rate to solve the regularized basis pursuit problem. Particularly, we verify that this result is an extended gradient algorithm with the convergence rate of $O(\frac{1}{\sqrt{k}})$, like that for smooth problems. Following the Nesterov method for accelerating the gradient method [34, 35], we show that the MABM can be further accelerated to converge as $O(1/(k+1)^2)$.

A series of lemmas in the following are to ensure the convergence rate of the MABM.

Lemma 4.1 *Let* (x^n, σ^n) *be generated by the augmented Lagrangian iterative scheme* (2.9), *and let* (x^*, σ^*) *be a globally optimal solution of the problem*

$$\min_{x \in \Re^n} \varphi(x) \quad s.t. \ Ax = b, \tag{4.1}$$

then the inequality

$$\varphi(x^k) - \varphi(x) \geq \langle \sigma, A(x^k - x) \rangle \tag{4.2}$$

for any $(x, \sigma) = (x^n, \sigma^n)$ *and* (x^*, σ^*) *holds.*

Proof By the optimality conditions, we get

$$0 \in \partial\varphi(x^n) - A^T\sigma^{n-1} - \lambda A^T(b - Ax^n).$$

From the second term of (2.9), we obtain

$$A^T\sigma^n \in \partial\varphi(x^n). \tag{4.3}$$

By the definition of subdifferential, we have

$$\varphi(x^k) \geq \varphi(x^n) + \langle \sigma^n, A(x^k - x^n) \rangle.$$

So (4.1) holds for $(x, \sigma) = (x^n, \sigma^n)$. Owing to the fact that (x^*, σ^*) of (4.1) satisfies the KKT condition, it is easy to obtain

$$0 \in \partial\varphi(x^*) - A^T\sigma^* \quad \text{and} \quad Ax^* = b.$$

Compared with (4.3), inequality (4.2) holds for $(x, \sigma) = (x^*, \sigma^*)$, which completes the proof. □

Lemma 4.2 *Let (x^k, σ^k) be generated by the augmented Lagrangian iterative scheme* (2.9). *For any (x, σ), the inequality*

$$L(x^k, \sigma^k) - L(x, \sigma) \geq \frac{1}{\lambda} \left\| \sigma^k - \sigma^{k-1} \right\|_2^2 - \frac{1}{\lambda} \langle \sigma^{k-1} - \sigma, \sigma^{k-1} - \sigma^k \rangle \qquad (4.4)$$

holds.

Proof By the definition of the Lagrangian function and Lemma 4.1, we get the following result:

$$
\begin{aligned}
L(x^k, \sigma^k) - L(x, \sigma) &= \varphi(x^k) - \varphi(x) - \langle \sigma, b - Ax \rangle + \langle \sigma^k, b - Ax^k \rangle \\
&\geq \langle \sigma, A(x^k - x) \rangle - \langle \sigma, b - Ax \rangle + \langle \sigma^k, b - Ax^k \rangle \\
&= \langle \sigma^k - \sigma, b - Ax^k \rangle \\
&= \frac{1}{\lambda} \langle \sigma - \sigma^k, \sigma^k - \sigma^{k-1} \rangle \\
&= \frac{1}{\lambda} \left\| \sigma^k - \sigma^{k-1} \right\|_2^2 - \frac{1}{\lambda} \langle \sigma^{k-1} - \sigma, \sigma^{k-1} - \sigma^k \rangle.
\end{aligned}
$$

The second inequality is derived from (4.2), and the third equality is from (2.9). □

From the fact that $A^T \sigma^k \in \partial \varphi(x^k)$, $A^T \sigma^* \in \partial \varphi(x^*)$, and the definition of subdifferential, we have

$$\varphi(x^*) \geq \varphi(x^k) + \langle \sigma^n, A(x^* - x^n) \rangle,$$

then

$$\varphi(x^*) - \langle \sigma^*, Ax^* - b \rangle \geq \varphi(x^k) - \langle \sigma^n, Ax^n - b \rangle,$$

thus

$$L(x^*, \sigma^*) \geq L(x^k, \sigma^k), \qquad (4.5)$$

where we exploit the fact that $Ax^* = b$.

Lemma 4.3 *Let (x^k, σ^k) be generated by the augmented Lagrangian iterative scheme* (2.9), *then*

$$\left\| \sigma^k - \sigma^* \right\|_2^2 \geq \left\| \sigma^{k-1} - \sigma^* \right\|_2^2 - \left\| \sigma^k - \sigma^{k-1} \right\|_2^2 - 2\lambda \left(L(x^*, \sigma^*) - L(x^k, \sigma^k) \right).$$

Proof Considering (4.4) and replacing (x, σ) with (x^*, σ^*), we get

$$
\begin{aligned}
\left\| \sigma^k - \sigma^* \right\|_2^2 &= \left\| \sigma^k - \sigma^{k-1} + \sigma^{k-1} - \sigma^* \right\|_2^2 \\
&= \left\| \sigma^k - \sigma^{k-1} \right\|_2^2 + 2 \langle \sigma^k - \sigma^{k-1}, \sigma^{k-1} - \sigma^* \rangle + \left\| \sigma^{k-1} - \sigma^* \right\|_2^2
\end{aligned}
$$

$$\leq \left\| \sigma^k - \sigma^{k-1} \right\|_2^2 + \left\| \sigma^{k-1} - \sigma^* \right\|_2^2$$
$$- 2\left\| \sigma^k - \sigma^{k-1} \right\|_2^2 - 2\lambda\left(L\left(x^*,\sigma^*\right) - L\left(x^k,\sigma^k\right)\right)$$
$$= \left\| \sigma^{k-1} - \sigma^* \right\|_2^2 - \left\| \sigma^k - \sigma^{k-1} \right\|_2^2 - 2\lambda\left(L\left(x^*,\sigma^*\right) - L\left(x^k,\sigma^k\right)\right),$$

which completes the proof. □

Moreover, $\|\sigma^k - \sigma^*\|_2^2 \leq \|\sigma^{k-1} - \sigma^*\|_2^2 - \|\sigma^k - \sigma^{k-1}\|_2^2$ holds due to (4.5), this implies the global convergence of (2.9). By summing the above inequality over $k = 1, 2, \ldots, n$, we get

$$\sum_{k=1}^{n} \left\| \sigma^k - \sigma^{k-1} \right\|_2^2 \leq \left\| \sigma^1 - \sigma^* \right\|_2^2. \tag{4.6}$$

Hence,

$$\lim_{k \to \infty} \left\| \sigma^k - \sigma^{k-1} \right\|_2^2 = 0. \tag{4.7}$$

Theorem 4.4 *Let (x^k, σ^k) be generated by the augmented Lagrangian iterative scheme* (2.9), *then $L(x^*, \sigma^*) - L(x^k, \sigma^k) = O(\frac{1}{k})$.*

Proof Noting (4.4) of Lemma 4.2 and substituting (x^{n-1}, σ^{n-1}) for (x, σ), we get

$$L\left(x^n, \sigma^n\right) - L\left(x^{n-1}, \sigma^{n-1}\right) \geq \frac{1}{\lambda} \left\| \sigma^n - \sigma^{n-1} \right\|_2^2. \tag{4.8}$$

By multiplying with $n - 1$ and summing it over $n = 1, \ldots, k$, we obtain

$$\sum_{n=1}^{k} \left\{ (n-1)L\left(x^n, \sigma^n\right) - (n-1)L\left(x^{n-1}, \sigma^{n-1}\right) \right\}$$
$$= \sum_{n=1}^{k} \left\{ nL\left(x^n, \sigma^n\right) - (n-1)L\left(x^{n-1}, \sigma^{n-1}\right) - L\left(x^n, \sigma^n\right) \right\}$$
$$= kL\left(x^k, \sigma^k\right) - \sum_{k=1}^{n} L\left(x^n, \sigma^n\right)$$
$$\geq \sum_{n=1}^{k} \frac{n-1}{\lambda} \left\| \sigma^n - \sigma^{n-1} \right\|_2^2 \geq 0.$$

From the above inequality, we have

$$\sum_{n=1}^{k} L\left(x^n, \sigma^n\right) \leq kL\left(x^k, \sigma^k\right). \tag{4.9}$$

On the other hand, it follows from Lemma 4.3 that

$$L\left(x^*, \sigma^*\right) - L\left(x^n, \sigma^n\right) \leq \frac{1}{2\lambda}\left(\left\| \sigma^{n-1} - \sigma^* \right\|_2^2 - \left\| \sigma^n - \sigma^{n-1} \right\|_2^2 - \left\| \sigma^n - \sigma^* \right\|_2^2\right).$$

Summing the inequality over $n = 1, \ldots, k$, we obtain

$$\sum_{n=1}^{k} \left(L(x^*, \sigma^*) - L(x^n, \sigma^n) \right)$$

$$= -\sum_{n=1}^{k} L(x^n, \sigma^n) + nL(x^*, \sigma^*)$$

$$\leq \frac{1}{2\lambda} \sum_{n=1}^{k} \left(\left\| \sigma^{n-1} - \sigma^* \right\|_2^2 - \left\| \sigma^n - \sigma^* \right\|_2^2 - \left\| \sigma^n - \sigma^{n-1} \right\|_2^2 \right)$$

$$= \frac{1}{2\lambda} \left(\left\| \sigma^0 - \sigma^* \right\|_2^2 - \left\| \sigma^k - \sigma^* \right\|_2^2 - \sum_{n=1}^{k} \left\| \sigma^n - \sigma^{n-1} \right\|_2^2 \right). \tag{4.10}$$

Combining (4.9) with (4.10), we get

$$-kL(x^k, \sigma^k) + kL(x^*, \sigma^*) \leq -\sum_{n=1}^{k} L(x^n, \sigma^n) + kL(x^*, \sigma^*)$$

$$\leq \frac{1}{2\lambda} \left(\left\| \sigma^0 - \sigma^* \right\|_2^2 - \left\| \sigma^k - \sigma^* \right\|_2^2 - \sum_{n=1}^{k} \left\| \sigma^n - \sigma^{n-1} \right\|_2^2 \right)$$

$$\leq \frac{1}{2\lambda} \left\| \sigma^0 - \sigma^* \right\|_2^2.$$

Consequently, we obtain

$$L(x^*, \sigma^*) - L(x^k, \sigma^k) \leq \frac{1}{2k\lambda} \left\| \sigma^* - \sigma^0 \right\|_2^2 = O\left(\frac{1}{k} \right).$$

This completes the proof. □

Comparing (2.9) with (3.5), we have the following two lemmas by replacing (x^k, σ^k) with $(x^k, \widetilde{\sigma}^k)$.

Lemma 4.5 *Let $(x^k, \widetilde{\sigma}^k)$ be generated by the augmented Lagrangian iterative scheme* (3.5). *For any (x, σ), we have the inequality*

$$L(x^k, \widetilde{\sigma}^k) - L(x, \sigma) \geq \frac{1}{\lambda} \left\| \widetilde{\sigma}^k - \sigma^{k-1} \right\|_2^2 - \frac{1}{\lambda} \langle \sigma^{k-1} - \sigma, \sigma^{k-1} - \widetilde{\sigma}^k \rangle. \tag{4.11}$$

Lemma 4.6 *Let $(x^k, \widetilde{\sigma}^k)$ be generated by the augmented Lagrangian iterative scheme* (3.5), *then*

$$\left\| \widetilde{\sigma}^k - \sigma^* \right\|_2^2 \geq \left\| \sigma^{k-1} - \sigma^* \right\|_2^2 - \left\| \widetilde{\sigma}^k - \sigma^{k-1} \right\|_2^2 - 2\lambda \left(L(x^*, \sigma^*) - L(x^k, \widetilde{\sigma}^k) \right). \tag{4.12}$$

Theorem 4.7 *Let $(x^k, \widetilde{\sigma}^k)$ be generated by the augmented Lagrangian iterative scheme* (3.5), *then*

$$L(x^*, \sigma^*) - L(x^k, \widetilde{\sigma}^k) = O\left(1/(k+1)^2 \right).$$

Proof We work from step 4 of Algorithm 3.4, *i.e.*, $t_{k+1} = \frac{1}{3}(\sqrt{1 + 9t_k^2} + 1)$, where $t_0 = 1$. By a simple calculation, we have

$$t_{k+1}(3t_{k+1} - 2) = 3t_k^2 \tag{4.13}$$

and

$$t_k \geq \frac{k+2}{3}. \tag{4.14}$$

Based on Lemma 4.5, by replacing (x, σ) with $(x^{n-1}, \widetilde{\sigma}^{n-1})$ and (x^*, σ^*), setting $k = n$, we get the two inequalities

$$w_{n-1} - w_n \geq \frac{1}{\lambda} \|\widetilde{\sigma}^n - \sigma^{n-1}\|_2^2 - \frac{1}{\lambda}\langle\sigma^{n-1} - \widetilde{\sigma}^{n-1}, \sigma^{n-1} - \widetilde{\sigma}^n\rangle, \tag{4.15}$$

$$-w_n \geq \frac{1}{\lambda} \|\widetilde{\sigma}^n - \sigma^{n-1}\|_2^2 - \frac{1}{\lambda}\langle\sigma^{n-1} - \sigma^*, \sigma^{n-1} - \widetilde{\sigma}^n\rangle, \tag{4.16}$$

where

$$w_n := L(x^*, \sigma^*) - L(x^n, \widetilde{\sigma}^n).$$

By multiplying both sides of (4.15), (4.16) with $3t_n - 2$ and 2, respectively, and adding the two sides, we obtain

$$(3t_n - 2)w_n - 3t_n w_{n+1}$$
$$\geq \frac{3t_n}{\lambda} \|\widetilde{\sigma}^n - \sigma^{n-1}\|_2^2 + \frac{1}{\lambda}\langle 2\sigma^* + (3t_n - 2)\widetilde{\sigma}^{n-1} - 3t_n\sigma^{n-1}, \sigma^{n-1} - \widetilde{\sigma}^n\rangle. \tag{4.17}$$

For the sake of convenience, we denote

$$P_n := 2\sigma^* + (3t_n - 2)\widetilde{\sigma}^{n-1} - 4t_n\widetilde{\sigma}^n + t_n\sigma^{n-1}$$

and

$$Q_n := 2\sigma^* + (3t_n - 2)\widetilde{\sigma}^{n-1} - 2t_n\widetilde{\sigma}^n - t_n\sigma^{n-1}.$$

By multiplying both sides of (4.17) with t_n, we have

$$t_n(3t_n - 2)w_n - 3t_n^2 w_{n+1}$$
$$\geq \frac{3}{\lambda} \|t_n(\widetilde{\sigma}^n - \sigma^{n-1})\|_2^2$$
$$\quad + \frac{1}{\lambda}\langle 2\sigma^* + (3t_n - 2)\widetilde{\sigma}^{n-1} - 3t_n\sigma^{n-1}, t_n(\sigma^{n-1} - \widetilde{\sigma}^n)\rangle$$
$$= \frac{1}{\lambda}\langle 2\sigma^* + (3t_n - 2)\widetilde{\sigma}^{n-1} - 3t_n\widetilde{\sigma}^n, t_n(\sigma^{n-1} - \widetilde{\sigma}^n)\rangle$$
$$= \frac{1}{4\lambda} \|2\sigma^* + (3t_n - 2)\widetilde{\sigma}^{n-1} - 4t_n\widetilde{\sigma}^n + t_n\sigma^{n-1}\|_2^2$$

$$-\frac{1}{4\lambda}\left\|2\sigma^* + (3t_n - 2)\widetilde{\sigma}^{n-1} - 2t_n\widetilde{\sigma}^n - t_n\sigma^{n-1}\right\|_2^2$$

$$= \frac{1}{4\lambda}\|P_n\|_2^2 - \frac{1}{4\lambda}\|Q_n\|_2^2,$$

where the second equality is derived from the fact that $\|\alpha + \beta\|_2^2 - \|\alpha - \beta\|_2^2 = 4\langle\alpha, \beta\rangle$.

Then we have

$$
\begin{aligned}
P_n - Q_n &= \left[2\sigma^* + (3t_n - 2)\widetilde{\sigma}^{n-1} - 4t_n\widetilde{\sigma}^n + t_n\sigma^{n-1}\right] \\
&\quad - \left[2\sigma^* + (3t_n - 2)\widetilde{\sigma}^{n-1} - 2t_n\widetilde{\sigma}^n - t_n\sigma^{n-1}\right] \\
&= 2t_n\sigma^{n-1} - 2t_n\widetilde{\sigma}^n \\
&= 2t_n\sigma^{n-1} - 2t_n\frac{t_{n+1}\sigma^n - (1 - 2t_n)\sigma^{n-1}}{t_{n+1} + 2t_n - 1} \\
&=: K_1\left(\sigma^{n-1} - \sigma^n\right),
\end{aligned}
$$

where

$$K_1 := \frac{2t_n t_{n+1}}{t_{n+1} + 2t_n - 1},$$

and the third equality is from the third iterative scheme of (3.5) and we make a transposition to the third term. Thus,

$$
\begin{aligned}
3t_{n-1}w_n - 3t_n^2 w_{n+1} &\geq \frac{1}{4\lambda}\|P_n\|_2^2 - \frac{1}{4\lambda}\left\|P_n - (P_n - Q_n)\right\|_2^2 \\
&\geq \frac{1}{4\lambda}\|P_n\|_2^2 - \frac{1}{4\lambda}\left(\|P_n\|^2 + \|P_n - Q_n\|_2^2\right) \\
&\geq -\frac{1}{4\lambda}\|P_n - Q_n\|_2^2 \\
&:= -\frac{K_1}{4\lambda}\left\|\sigma^{n-1} - \sigma^n\right\|_2^2.
\end{aligned}
$$

Summing the inequality over $n = 1, \ldots, k$, we get

$$3t_0^2 w_1 - 3t_k^2 w_{k+1} \geq -\sum_{n=1}^{k}\frac{K_1}{4\lambda}\left\|\sigma^{n-1} - \sigma^n\right\|_2^2. \tag{4.18}$$

Then, combining (4.7), (4.14) with (4.18), we have

$$
\begin{aligned}
w_{k+1} &\leq \frac{3t_0^2 w_1 + \sum_{n=1}^{k}\frac{K_1}{4\lambda}\|\sigma^{n-1} - \sigma^n\|_2^2}{3t_k^2} \\
&\leq \frac{9t_0^2 w_1 + 3\varepsilon}{(k+2)^2},
\end{aligned}
$$

as k tends to infinity, where ε is an arbitrarily small positive number. Hence,

$$L\left(x^*, \sigma^*\right) - L\left(x^k, \widetilde{\sigma}^k\right) = O\left(1/(k+1)^2\right).$$

Thus, we complete the convergence proof. □

Remark The right formula of (4.18) exploits the fact that the choice of the penalty factor λ can be seen as a monotonically increasing sequence (such as λ_k) that depends on the selection of t_k in Algorithm 3.4. In this way, we are not only able to guarantee the convergence in the formula of K_1 divided by λ_k, but also to play a critical role of punishment to the constraint condition.

5 Conclusion

In this paper, we put forward the modified accelerated Bregman method (MABM) for solving the regularized basis pursuit problem. We give some beneficial improvement tasks on the basis of some recent literature on the accelerated Bregman method, and we perform the theoretical feasibility analysis in detail. It can be showed that the proposed MABM has a rapid convergence rate of $O(1/(k+1)^2)$. We will devote our future study to combining the advantages of LBM with our MABM as regards theory and numerical results.

Author details
[1]School of Mathematics and Computer Science, Fujian Normal University, Fuzhou, 350007, P.R. China. [2]Fujian Jiangxia University, Fuzhou, 350108, P.R. China.

References
1. Kashin, B: The widths of certain finite dimensional sets and classes of smooth functions. Izv. Akad. Nauk SSSR, Ser. Mat. **41**, 334-351 (1977)
2. Candés, E, Romberg, J, Tao, T: Robust uncertainty principles: exact signal reconstruction from highly incomplete frequency information. IEEE Trans. Inf. Theory **52**, 489-509 (2006)
3. Candés, E, Romberg, J, Tao, T: Stable signal recovery from incomplete and inaccurate measurements. Commun. Pure Appl. Math. **59**, 1207-1223 (2006)
4. Candés, E, Tao, T: Decoding by linear programming. IEEE Trans. Inf. Theory **51**, 4203-4215 (2005)
5. Donoho, D: Compressed sensing. IEEE Trans. Inf. Theory **52**, 1289-1306 (2006)
6. Candés, E, Wakin, M: An introduction to compressive sampling. IEEE Signal Process. Mag. **21**, 21-30 (2008)
7. Cai, J, Osher, S, Shen, Z: Linearized Bregman iterations for compressed sensing. Math. Comput. **78**, 1515-1536 (2009)
8. Cai, J, Candés, E, Shen, Z: A singular value thresholding algorithm for matrix completion. SIAM J. Optim. **20**, 1956-1982 (2010)
9. Candés, E, Recht, B: Exact matrix completion via convex optimization. Found. Comput. Math. **9**, 717-772 (2009)
10. Candés, E, Li, X, Ma, Y, Wright, J: Robust principal component analysis. Preprint (2009)
11. Figueiredo, MAT, Nowak, RD, Wright, SJ: Gradient projection for sparse reconstruction: application to compressed sensing and other inverse problems. IEEE J. Sel. Top. Signal Process. **1**(4), 586-597 (2007)
12. Beck, A, Teboulle, M: Fast iterative shrinkage-thresholding algorithm for linear inverse problems. SIAM J. Imaging Sci. **2**, 183-202 (2009)
13. Peng, Y, Ganesh, A, Wright, J, Xu, W, Ma, Y: RASL: robust alignment by sparse and low-rank decomposition for linearly correlated images. In: Proceedings of IEEE Conference on Computer Vision and Pattern Recognition (CVPR) (2010)
14. Ghaoui, LE, Gahinet, P: Rank minimization under LMI constraints: a framework for output feedback problems. In: Proceedings of the European Control Conference (1993)
15. Fazel, M, Hindi, H, Boyd, SP: A rank minimization heuristic with application to minimum order system approximation. In: Proceedings of the American Control Conference, vol. 6, pp. 4734-4739 (2001)
16. Linial, N, London, E, Rabinovich, Y: The geometry of graphs and some of its algorithmic applications. Combinatorica **15**, 215-245 (1995)
17. Kang, M, Yun, S, Woo, H, Kang, M: Accelerated Bregman method for linearly constrained $l_1 - l_2$ minimization. J. Sci. Comput. **56**, 515-534 (2012)
18. Huang, B, Ma, SQ, Goldfarb, D: Accelerated linearized Bregman method. J. Sci. Comput. **54**, 428-453 (2013)
19. Yin, W: Analysis and generalizations of the linearized Bregman method. SIAM J. Imaging Sci. **3**, 856-877 (2010)
20. Friedlander, M, Tseng, P: Exact regularization of convex programs. SIAM J. Optim. **18**, 1326-1350 (2007)

21. Osher, S, Mao, Y, Dong, B, Yin, W: Fast linearized Bregman iteration for compressive sensing and sparse denoising. Commun. Math. Sci. **8**, 93-111 (2010)
22. Darbon, J, Osher, S: Fast discrete optimization for sparse approximations and deconvolutions. Preprint (2007)
23. Osher, S, Burger, M, Goldfarb, D, Xu, J, Yin, W: An iterative regularization method for total variation-based image restoration. Multiscale Model. Simul. **4**(2), 460-489 (2005) (electronic). MR2162864 (2006c:49051)
24. Yin, W, Osher, S, Goldfarb, D, Darbon, J: Bregman iterative algorithms for ℓ_1-minimization with applications to compressed sensing. SIAM J. Imaging Sci. **1**(1), 143-168 (2008)
25. Barzilai, J, Borwein, J: Two point step size gradient methods. IMA J. Numer. Anal. **8**, 141-148 (1988)
26. Liu, D, Nocedal, J: On the limited memory method for large scale optimization. Math. Program., Ser. B **45**, 503-528 (1989)
27. Bennett, J, Lanning, S: The Netflix prize. In: Proceedings of KDD Cup and Workshop (2007)
28. Goldstein, T, Osher, S: The split Bregman method for L1-regularized problems. SIAM J. Imaging Sci. **2**(2), 323-343 (2009)
29. Bregman, L: The relaxation method of finding the common points of convex sets and its application to the solution of problems in convex programming. USSR Comput. Math. Math. Phys. **7**, 200-217 (1967)
30. Powell, MJD: A method for nonlinear constraints in minimization problems. In: Fletcher, R (ed.) Optimization, pp. 283-298 (1972)
31. Rockafellar, RT: Augmented Lagrangians and applications of the proximal point algorithm in convex programming. Math. Oper. Res. **1**, 97-116 (1976)
32. Hale, ET, Yin, W, Zhang, Y: Fixed-point continuation for l_1 minimization: methodology and convergence. SIAM J. Optim. **19**, 1107-1130 (2008)
33. Berg, EVD, Friedlander, MP: Probing the Pareto frontier for basis pursuit solutions. SIAM J. Sci. Comput. **31**, 890-912 (2008)
34. Nesterov, YE: A method for unconstrained convex minimization problem with the rate of convergence $O(1/k^2)$. Dokl. Akad. Nauk SSSR **269**, 543-547 (1983)
35. Nesterov, YE: Introductory Lectures on Convex Optimization, vol. 87, pp. 220-236 (2004)
36. Rennie, JDM, Srebro, N: Fast maximum margin matrix factorization for collaborative prediction. In: Proceedings of the International Conference on Machine Learning, pp. 713-719 (2005)
37. Bertsekas, D: Nonlinear Programming. Athena Scientific, Nashua (1999)