CONVERGENCE OF THE LINEARIZED BREGMAN ITERATION FOR ℓ_1-NORM MINIMIZATION

JIAN-FENG CAI, STANLEY OSHER, AND ZUOWEI SHEN

ABSTRACT. One of the key steps in compressed sensing is to solve the basis pursuit problem $\min_{u \in \mathbb{R}^n}\{\|u\|_1 : Au = f\}$. Bregman iteration was very successfully used to solve this problem in [40]. Also, a simple and fast iterative algorithm based on linearized Bregman iteration was proposed in [40], which is described in detail with numerical simulations in [35]. A convergence analysis of the smoothed version of this algorithm was given in [11]. The purpose of this paper is to prove that the linearized Bregman iteration proposed in [40] for the basis pursuit problem indeed converges.

1. INTRODUCTION

Let $A \in \mathbb{R}^{m \times n}$ with $n > m$ and $f \in \mathbb{R}^m$ be given. Assume that A is a surjective map, i.e., AA^T is invertible. Then there are infinitely many solutions for the system of linear equations $Au = f$, e.g., $u = A^T(AA^T)^{-1}f$ is the solution minimizing the ℓ_2-norm among all solutions. For applications in compressed sensing, it amounts to finding a minimal ℓ_1-norm solution, i.e., the solution should satisfy the following minimization problem:

$$(1.1) \qquad \min_{u \in \mathbb{R}^n}\{\|u\|_1 : Au = f\}.$$

The set of all solutions of $Au = f$ is convex. Since $\|\cdot\|_1$ is coercive, the set of all solutions of (1.1) is a nonempty convex set. The interested readers should consult [14–17, 27, 38] and references therein on theory of compressed sensing for more details.

One can transform (1.1) into a linear programming problem, and then solve it by a conventional linear programming solver in many cases. However, such solvers do not use, for example, the facts that matrices A are normally formed by rows of some orthonormal matrices corresponding to fast transforms where both Au and A^Tu can be computed by fast transforms, and that the solution to seek is sparse. These facts are indeed true in the applications of compressed sensing. More importantly, the algorithm should be robust to noise and should take care of the difficulties that the matrix A is huge and dense. Therefore, there is a need to find a more efficient algorithm that adapts to the above challenges.

For these purposes, a simple and fast algorithm based on linearized Bregman iteration was proposed in [40]. The linearized Bregman iteration for (1.1) is

(1.2) $$\begin{cases} v^{k+1} = v^k + A^T(f - Au^k), \\ u^{k+1} = \delta T_\mu(v^{k+1}), \end{cases}$$

where $u^0 = v^0 = 0$, and

(1.3) $$T_\mu(w) := [t_\mu(w(1)), t_\mu(w(2)), \ldots, t_\mu(w(n))]^T$$

is the soft thresholding operator given in [26] with

(1.4) $$t_\mu(\xi) = \begin{cases} 0, & \text{if } |\xi| \le \mu, \\ \text{sgn}(\xi)(|\xi| - \mu), & \text{if } |\xi| > \mu. \end{cases}$$

In the first step of (1.2), we add the transformed error $f - Au^k$ by A^T into v^k to obtain v^{k+1}. This can be understood as an updating of v^k by an approximation of a solution of the error equation $Au = f - Au^k$ which may not be sparse. In fact, when $AA^T = I$, $A^T(f - Au^k)$ is the solution of the error equation minimizing the ℓ_2-norm. In the second step of (1.2), we threshold v^{k+1} by T_μ. This step produces a sparse vector u^{k+1} and removes the noise. In fact, if we choose a large μ (as we will see later, this is the case in both theory and practice), only large components in v^{k+1} are nonzeroes in u^{k+1}. This implies that u^{k+1} is a sparse vector, and the noise contained can be efficiently removed. When the observed data f contains noise, one can stop iteration (1.2) whenever, e.g.,

$$\|Au^k - f\|^2 \le \sigma^2,$$

where σ^2 is the variance of the noise.

The linearized Bregman iteration (1.2) is very efficient and robust to noise in solving the problems in which the underlying solution is very sparse. This fact is shown by the numerical experiments in [35] for the applications arising from compressed sensing. It can be made even faster by introducing a simple numerical device called "kicking", which resembles line search (see [35] for details). Furthermore, the linearized Bregman iteration has led to a very fast frame based deblurring algorithm as shown in [12].

Finally, we remark that the idea of applying a thresholding operator to each iterate in an iterative algorithm to obtain a sparse and noise free solution has been used successfully in image and signal processing in many occasions. The interested readers should consult, e.g., [6–10, 13, 20–22, 24, 25], for details.

It is proven in [11] that if $\{u^k\}_{k \in \mathbb{N}}$ generated by (1.2) converges, its limit is the unique solution of

(1.5) $$\min_{u \in \mathbb{R}^n} \{\mu\|u\|_1 + \frac{1}{2\delta}\|u\|^2 : Au = f\}.$$

It was also shown in [11] that the limit of (1.2) becomes a solution of the basis pursuit problem (1.1) as $\mu \to \infty$. Furthermore, it was shown in [11] that the corresponding linearized Bregman iteration converges when a smoothed ℓ_1-norm is used. However, there is no result on the convergence of the sequence $\{u^k\}_{k \in \mathbb{N}}$ of (1.2). In this paper, we prove that iteration (1.2) does converge.

2. Backgrounds and Main Results

In this section, we start with some backgrounds of the algorithms, and end up with our main results.

2.1. Bregman and linearized Bregman iterations.
Iterative algorithms involving Bregman distance were introduced to image and signal processing by [18,19] and by many other authors. See [34] for an overview. In [34], a Bregman iteration was proposed for the nondifferentiable TV energy for image restoration. Then, in [40], it was shown to be remarkably successful for ℓ_1-norm minimization problems in compressed sensing. To further improve the performance of the Bregman iteration, a linearized Bregman iteration was invented in [23]; see also [40]. More details and an improvement called "kicking" of the linearized Bregman iteration was described in [35], and a rigorous theory for a smoothed ℓ_1-norm was given in [11]. The linearized Bregman iteration was applied to tight frame-based image deblurring in [12]. Recently, a new type of iteration based on Bregman distance, called split Bregman iteration, was introduced in [29], which extended the utility of the Bregman iteration and the linearized Bregman iteration to minimizations of more general ℓ_1-based regularizations including total variation, Besov norms, and sums of such things.

The Bregman iteration and the linearized Bregman iteration are all based on Bregman distance [3], which is defined by

$$(2.1) \qquad D_J^p(u,v) = J(u) - J(v) - \langle u - v, p \rangle,$$

where J is a convex function, $p \in \partial J(v)$ is a subgradient in the subdifferential of J at the point v. Notice that $D_J^p(u,v)$ is not a distance in the usual sense, since $D_J^p(u,v) \neq D_J^p(v,u)$ in general. However, it measures the closeness between u and v in the sense that $D_J^p(u,v) \geq 0$, whenever J is convex, and $D_J^p(u,v) \geq D_J^p(w,v)$ for all points w on the line segment connecting u and v.

The Bregman iteration for the general problem

$$(2.2) \qquad \min_{u \in \mathbb{R}^n} \{ J(u) : Au = f \}$$

is, given $u^0 = p^0 = 0$,

$$(2.3) \qquad \begin{cases} u^{k+1} = \arg\min_{u \in \mathbb{R}^n} \{ \frac{1}{2} \|Au - f\|^2 + \mu D_J^{p^k}(u, u^k) \}, \\ p^{k+1} = p^k - \frac{1}{\mu} A^T (Au^k - f). \end{cases}$$

This iteration was first proposed in [34] for total variation denoising, and then applied to the ℓ_1-norm minimization problem (1.1) in [40]. It was proven in [34] that the error $\|Au^k - f\|^2$ decreases to 0 for any convex function J. It was further shown in [40] that (2.3) with $J(u) = \|u\|_1$ reaches a solution of (1.1) in finitely many steps. By letting $f^k = \sum_{i=0}^{k-1} (f - Au^i)$, it was shown in [34, 40] that the Bregman iteration has a beautiful formulation

$$(2.4) \qquad \begin{cases} f^{k+1} = f^k + (f - Au^k) \\ u^{k+1} = \arg\min_{u \in \mathbb{R}^n} \{ \frac{1}{2} \|Au - f^{k+1}\|^2 + \mu \|u\|_1 \}, \end{cases}$$

where $f^0 = 0$ and $u^0 = 0$. Since the parameter μ is arbitrary, one can choose an optimal μ such that the condition number in the second step is optimized. Therefore, the Bregman iteration (2.4) is efficient in solving (2.2). It was also

explained in [29] why (2.4) with $J(u) = \|u\|_1$ is particularly efficient in solving (1.1).

The convergence and error analysis of the Bregman iteration were studied in, for examples, [4, 34, 37, 40]. It was pointed out in [40] that the Bregman iteration (2.3) or (2.4) is equivalent to an augmented Lagrangian method in [1, 28, 30, 33, 36].

However, in the second step of (2.4), we need to solve a minimization problem. To improve the performance, the linearized Bregman iteration was proposed in [23, 40]. The idea is to approximate the term $\|Au - f\|^2$ in (2.3) by its Taylor expansion around u^k,

$$\|Au - f\|^2 \approx \|Au^k - f\|^2 + 2\langle u, A^T(Au^k - f)\rangle + \frac{1}{\delta}\|u - u^k\|^2,$$

where δ is a fixed parameter. With this, we obtain the linearized Bregman iteration as follows: Given $u^0 = p^0 = 0$, we iterate

$$(2.5) \quad \begin{cases} u^{k+1} = \arg\min_{u\in\mathbb{R}^n} \left\{ \frac{1}{2\delta}\|u - (u^k - \delta A^T(Au^k - f))\|^2 + \mu D_J^{p^k}(u, u^k) \right\}, \\ p^{k+1} = p^k - \frac{1}{\mu\delta}(u^{k+1} - u^k) - \frac{1}{\mu}A^T(Au^k - f). \end{cases}$$

If $J(u) = \|u\|_1$, the linearized Bregman iteration (2.5) can be reformulated into the compact form (1.2); see [11, 40].

2.2. Convergence of linearized Bregman iteration for general problems.
The following result of [11] says that if (2.5) converges, the limit is a minimizer of the right cost functional.

Proposition 2.1 ([11]). *Suppose that the sequence $\{u^k\}_{k\in\mathbb{N}}$ generated by (2.5) converges, and $\{p^k\}_{k\in\mathbb{N}}$ is bounded. Then the limit of $\{u^k\}_{k\in\mathbb{N}}$ is the unique solution of*

$$(2.6) \qquad \min_{u\in\mathbb{R}^n}\{\mu J(u) + \frac{1}{2\delta}\|u\|^2 : Au = f\}.$$

Furthermore, [11] also gave the following convergence result for (2.5):

Proposition 2.2 ([11]). *Suppose that $J(u)$ is convex and differentiable, and its gradient satisfies*

$$(2.7) \qquad \|\nabla J(u) - \nabla J(v)\|^2 \le \beta\langle\nabla J(u) - \nabla J(v), u - v\rangle, \quad \forall u, v \in \mathbb{R}^n.$$

Then both the sequences $\{u^k\}_{k\in\mathbb{N}}$ and $\{p^k\}_{k\in\mathbb{N}}$ generated by (2.5) with $0 < \delta < \frac{2}{\|AA^T\|}$ converge with a rate $\eta < 1$.

2.3. Convergence of linearized Bregman iteration for basis pursuit problem.
By applying Proposition 2.1, one obtains the following result for $J = \|\cdot\|_1$.

Proposition 2.3 ([11]). *Suppose that the sequence $\{u^k\}_{k\in\mathbb{N}}$ generated by (1.2) converges. Then the limit of $\{u^k\}_{k\in\mathbb{N}}$ is the unique solution of*

$$(2.8) \qquad \min_{u\in\mathbb{R}^n}\{\mu\|u\|_1 + \frac{1}{2\delta}\|u\|^2 : Au = f\}.$$

Let $u_{\mu,0}^*$ be the solution of (2.8). Then

$$\lim_{\mu\to\infty}\|u_{\mu,0}^* - u_1\| = 0,$$

where u_1 is the solution of (1.1) that has the minimal ℓ_2-norm among all the solutions of (1.1).

Since $J(u) = \|u\|_1$ does not satisfy the condition (2.7), one cannot apply Proposition 2.2 to prove the convergence of (1.2). To overcome this, we used in [11] a convex differentiable function $J_\epsilon(u)$ to approximate $\|u\|_1$,

$$(2.9) \qquad J_\epsilon(u) \;=\; \sum_{i=1}^n F_\epsilon(u(i)), \quad \text{with} \quad F_\epsilon(\xi) = \begin{cases} \frac{\xi^2}{2\epsilon}, & \text{if } |\xi| \le \epsilon, \\ |\xi| - \frac{\epsilon}{2}, & \text{if } |\xi| > \epsilon. \end{cases}$$

We remark that J_ϵ is also known as the Huber norm [32] and the Moreau-Yosida C^1-regularization [31] of the ℓ_1-norm. It is easy to verify that $J_0(u) = \|u\|_1$, and J_ϵ with $\epsilon > 0$ satisfies (2.7). It was also derived in [11] that the linearized Bregman iteration (2.5) with $J = J_\epsilon$ can be reformulated into the following compact formula:

$$(2.10) \qquad \begin{cases} v^{k+1} = v^k + A^T(f - Au^k), \\ u^{k+1} = \delta T_{\mu,\epsilon}(v^{k+1}), \end{cases}$$

where $u^0 = v^0 = 0$, and

$$(2.11) \qquad T_{\mu,\epsilon}(w) := [t_{\mu,\epsilon}(w(1)), t_{\mu,\epsilon}(w(2)), \dots, t_{\mu,\epsilon}(w(n))]^T$$

is the thresholding operator with

$$(2.12) \qquad t_{\mu,\epsilon}(\xi) = \begin{cases} \frac{\epsilon}{\mu+\epsilon}\xi, & \text{if } |\xi| \le \mu + \epsilon, \\ \mathrm{sgn}(\xi)(|\xi| - \mu), & \text{if } |\xi| > \mu + \epsilon. \end{cases}$$

Iteration (2.10) can be understood as a smoothed version of iteration (1.2). By Propositions 2.1 and 2.2, the sequence $\{u^k\}_{k\in\mathbb{N}}$ generated by iteration (2.5) with $J = J_\epsilon$, $\epsilon > 0$, converges to $u^*_{\mu,\epsilon}$, the unique solution of

$$(2.13) \qquad \min_{u\in\mathbb{R}^n}\{\mu J_\epsilon(u) + \frac{1}{2\delta}\|u\|^2 : Au = f\}.$$

Moreover, we showed in [11] that, as $\mu\epsilon \to 0$, the limit $u^*_{\mu,\epsilon}$ converges to a solution (2.8). This, together with Proposition 2.3, implies that, for sufficiently small $\mu\epsilon$ and sufficiently large μ, the linearized Bregman iteration (2.10) gives a good approximation of the solution of the basis pursuit problem (1.1).

In short, what [11] has achieved is that the basis pursuit problem (1.1) can be solved as the limit of a smoothed version of iteration (1.2). However, there is no convergence result for iteration (1.2) itself. In Section 3, we will prove the following convergence theorem, the main result of the paper, for the linearized Bregman iteration (1.2) for (1.1). In this sense, this paper can be regarded as a continuation of [11].

Theorem 2.4 (Main Theorem). *Assume that $0 < \delta < \frac{1}{\|AA^T\|}$. Then the sequence $\{u^k\}_{k\in\mathbb{N}}$ generated by (1.2) converges to the unique solution of (1.5), i.e.,*

$$(2.14) \qquad \lim_{k\to\infty}\|u^k - u^*_{\mu,0}\| = 0,$$

where $u^*_{\mu,0}$ is the unique solution of (1.5). Furthermore,

$$\lim_{\mu\to\infty}\|u^*_{\mu,0} - u_1\| = 0,$$

where u_1 is the solution of (1.1) that has the minimal ℓ_2-norm among all the solutions of (1.1).

The second part of the theorem follows from the convergence of the sequence $\{u^k\}_{k\in\mathbb{N}}$ and Proposition 2.3. It only remains to prove that the sequence $\{u^k\}_{k\in\mathbb{N}}$ converges which is done in the next section.

The proof given here goes along with the approach of [11]. Since the area of the convex optimization is well developed, there is a rich literature on the convergence of iterative algorithms for constrained minimization. For example, one can show that the linearized Bregman iteration (2.5) is equivalent to gradient descent applied to the dual of problem (2.6), hence, it becomes the Uzawa's algorithm. The Uzawa's algorithm is a well studied subject and interested reader can find more details for example in [1,2]. This observation of connecting the linearized Bregman iteration to the Uzawa's algorithm is given in [5] where a singular value thresholding algorithm is developed for the matrix completion. A similar observation was also communicated to the second author of this paper by Wotao Yin in [39]. Once this linkage is established, the analysis of convergence can be go along with that of the Uzawa's algorithm. However, we keep our original proof here, since it helps us to develop an algorithm for applications beyond compressed sensing. For example, this proof motivates us to derive an efficient frame based deblurring algorithm and its convergence analysis in [12].

3. Proof of the Convergence

In this section, we prove the convergence of (1.2), i.e., Theorem 2.4. Our strategy is as follows. We first show that the energy $\|Au^k - f\|^2$ is decreasing by citing a lemma in [40]. Then, we show the boundedness of the sequence $\{u^k\}_{k\in\mathbb{N}}$, hence there exists at least one clustering point. Finally, we show that every clustering point is the unique solution of (1.5). Therefore, by the uniqueness of the solution of (1.5), we conclude that $\{u^k\}_{k\in\mathbb{N}}$ converges to the solution of (1.5).

We first cite a lemma which was shown in [40]. We include the proof to make the paper self contained.

Lemma 3.1 ([40]). *Assume that $\delta A^T A < I$. Then*

$$(3.1) \qquad \|Au^{k+1} - f\|^2 + \left(\frac{1}{\delta} - \|A^T A\|\right)\|u^{k+1} - u^k\|^2 \le \|Au^k - f\|^2.$$

Proof. By the first equation in (2.5), we have

(3.2)

$$\mu(\|u^{k+1}\|_1 - \|u^k\|_1 - \langle u^{k+1} - u^k, p^k\rangle) + \frac{1}{2\delta}\|u^{k+1} - (u^k - \delta A^T(Au^k - f))\|^2$$

$$\le \mu(\|u^k\|_1 - \|u^k\|_1 - \langle u^k - u^k, p^k\rangle) + \frac{1}{2\delta}\|u^k - (u^k - \delta A^T(Au^k - f))\|^2.$$

This, together with the nonnegativity of the Bregman distance, implies

$$\|u^{k+1} - u^k + \delta A^T(Au^k - f)\|^2 \le \|\delta A^T(Au^k - f)\|^2,$$

which is equivalent to

$$\|u^{k+1} - u^k\|^2 + 2\delta\langle u^{k+1} - u^k, A^T(Au^k - f)\rangle \le 0.$$

With some manipulations, this leads to

$$\|u^{k+1} - u^k\|^2 + \delta\|Au^{k+1} - f\|^2 - \delta\langle u^{k+1} - u^k, A^T A(u^{k+1} - u^k)\rangle \le \delta\|Au^k - f\|^2.$$

It, in turn, gives

$$\langle (I - \delta A^T A)(u^{k+1} - u^k), u^{k+1} - u^k \rangle + \delta \|Au^{k+1} - f\|^2 \le \delta \|Au^k - f\|^2.$$

This leads to (3.1). □

The following lemma shows the boundedness of $\{u^k\}_{k \in \mathbb{N}}$. It is shown by orthogonally decomposing u^k into the sum of two components: one is in the range of A^T and the other is in the kernel of A.

Lemma 3.2. *Assume that $\delta A^T A < I$. Then the sequences $\{u^k\}_{k \in \mathbb{N}}$ and $\{v^k\}_{k \in \mathbb{N}}$ are bounded.*

Proof. Decompose u^k into the orthogonal sum of $u^k = x^k + y^k$, where x^k is in the range of A^T, and y^k is in the kernel of A. Then, we have

$$(3.3) \qquad\qquad x^k = A^\dagger A u^k,$$

where $A^\dagger$ is the pseudo-inverse of A, and $A^\dagger = A^T(AA^T)^{-1}$ when A is rectangular and surjective. Since $\{\|Au^k - f\|^2\}_{k \in \mathbb{N}}$ is a decreasing sequence by Lemma 3.1, the sequence $\{\|Au^k - f\|^2\}_{k \in \mathbb{N}}$ converges. Hence the sequence $\{Au^k\}_{k \in \mathbb{N}}$ is bounded. Since $A^\dagger$ is a bounded operator, the sequence $\{A^\dagger Au^k\}_{k \in \mathbb{N}}$ is bounded. By (3.3), $\{x^k\}_{k \in \mathbb{N}}$, the component of u^k in the range of A^T, is bounded.

It remains to show that $\{y^k\}_{k \in \mathbb{N}}$ is bounded. By the definition of $\{u^k\}_{k \in \mathbb{N}}$ in (1.2), $u^k = \delta T_\mu(v^k)$, which can be written as

$$(3.4) \qquad\qquad u^k = \delta v^k + \delta(T_\mu(v^k) - v^k).$$

Notice that by induction $v^k = A^T \sum_{j=0}^{k-1}(f - Au^j)$, which implies that v^k is in the range of A^T, so is δv^k. From this and (3.4), we deduce that y^k must be the orthogonal projection of $\delta(T_\mu(v^k) - v^k)$ onto the kernel of A, i.e., $y^k = \delta P_{\mathrm{Ker}(A)}(T_\mu(v^k) - v^k)$, where $P_{\mathrm{Ker}(A)}$ is the orthogonal projection onto the kernel of A. Since $(T_\mu(v^k) - v^k) \in [-\mu, \mu]^n$ by the definition of the soft thresholding in (1.3) and (1.4), $P_{\mathrm{Ker}(A)}(T_\mu(v^k) - v^k)$ is bounded. This concludes that $\{y^k\}_{k \in \mathbb{N}}$ is bounded.

Hence $\{u^k\}_{k \in \mathbb{N}}$ is bounded. The boundedness of $\{v^k\}_{k \in \mathbb{N}}$ is an immediate consequence of (3.4). □

Lemma 3.3. *Assume that $\delta A^T A < I$. Then*

$$(3.5) \qquad\qquad \lim_{k \to \infty} \|A^T(Au^k - f)\| = 0.$$

In particular, when A is rectangular and surjective, Au^k converges to f, i.e.,

$$(3.6) \qquad\qquad \lim_{k \to \infty} \|Au^k - f\| = 0.$$

Proof. Since $\{u^k\}_{k \in \mathbb{N}}$ hence $\{A^T Au^k\}_{k \in \mathbb{N}}$ are bounded by Lemma 3.2, there exist convergent subsequences of $\{A^T Au^k\}_{k \in \mathbb{N}}$. In order to prove (3.5), we show that each convergent subsequence of $\{A^T Au^k\}_{k \in \mathbb{N}}$ converges to $A^T f$. For this, let $\{A^T Au^{k_i}\}_{i \in \mathbb{N}}$ be an arbitrary given convergent subsequence, and denote $\lim_{i \to \infty} A^T(f - Au^{k_i}) = d$. We prove that $d = 0$ by contradiction.

Assume that $d \neq 0$. By the first equation in (1.2), for any finite integer ℓ, we have

$$(3.7) \quad v^{k_i + \ell} - v^{k_i} = \sum_{j=0}^{\ell-1} A^T(f - Au^{k_i + j}) = \sum_{j=0}^{\ell-1} A^T(f - Au^{k_i}) + \sum_{j=0}^{\ell-1} A^T A(u^{k_i} - u^{k_i + j}).$$

On the other hand, by (3.1) in Lemma 3.1, one has that $\sum_{k=1}^{\infty} \|u^{k+1} - u^k\|^2 < \infty$, which implies that

$$\text{(3.8)} \qquad\qquad \lim_{k \to \infty} \|u^{k+1} - u^k\| = 0.$$

By letting $i \to \infty$ in (3.7), (3.8) implies that, for any finite integer ℓ, $\lim_{i\to\infty} v^{k_i+\ell} - v^{k_i} = \ell d$. Therefore, there exist an i_0 such that for all $i > i_0$, $\|v^{k_i+\ell} - v^{k_i} - \ell d\| \leq 1$. Hence

$$\text{(3.9)} \qquad\qquad \|v^{k_i+\ell}\| \geq \|v^{k_i} + \ell d\| - 1 \geq \ell\|d\| - \|v^{k_i}\| - 1.$$

By Lemma 3.2, the sequence $\{v^k\}_{k\in\mathbb{N}}$ is bounded, i.e.,

$$\text{(3.10)} \qquad\qquad \|v^k\| \leq B, \qquad \forall k.$$

However, if we choose $\ell = \lceil (2B+2)/\|d\| \rceil$ that is finite because $d \neq 0$, then by (3.9) $\|v^{k_i+\ell}\| \geq B + 1$. It contradicts (3.10). $\qquad\square$

Remark 3.4. *Lemma 3.2 and Lemma 3.3 still hold even when A is not surjective, since the proofs do not depend on the surjectivity of A. In particular, the conclusion of Lemma 3.3 is that the limit of u^k is a solution of $\min_u \|Au - f\|^2$, i.e., a least square solution.*

With the above lemmas, we are ready to prove the main result of the paper, Theorem 2.4.

Proof of Theorem 2.4. In order to prove Theorem 2.4, it only remains to show that the sequence $\{u^k\}_{k\in\mathbb{N}}$ converges. Since $\{u^k\}_{k\in\mathbb{N}}$ is bounded by Lemma 3.2, there exist convergent subsequences. We show that the limit of an arbitrary convergent subsequence is the unique solution $u^*_{\mu,0}$ of (2.8). By the uniqueness of $u^*_{\mu,0}$, we obtain (2.14).

Let $\{u^{k_i}\}_{i\in\mathbb{N}}$ be an arbitrary given convergent subsequence and $\widetilde{u}$ its limit. Next, we prove that $\widetilde{u} = u^*_{\mu,0}$. Since $p^0 = u^0 = 0$, by the second equation in (2.5), we have $\mu p^k + \frac{1}{\delta} u^k = A^T \sum_{j=0}^{k-1}(f - Au^j)$. Define $w^k = \sum_{j=0}^{k-1}(f - Au^j)$. Then, since A is surjective and both $\{p^k \in [-1,1]^n\}_{k\in\mathbb{N}}$ and $\{u^k\}_{k\in\mathbb{N}}$ are bounded, we have that $\{w^k\}_{k\in\mathbb{N}}$ is bounded, i.e., $\|w^k\| \leq C$ for all k.

Let $H(u) = \mu\|u\|_1 + \frac{1}{2\delta}\|u\|^2$. It is obviously that $p^0 \in \partial\|u^0\|_1$ since $p^0 = u^0 = 0$. By the definition of p^k and u^k in (2.5), we have $p^k \in \partial\|u^k\|_1$ by induction and differentiating the energy in the first equation of (2.5). Hence, $p^{k_i} \in \partial\|u^{k_i}\|_1$. Therefore, $\mu p^{k_i} + \frac{1}{\delta} u^{k_i} \in \partial H(u^{k_i})$. By the nonnegativity of the Bregman distance $D_H^{\mu p^{k_i}+\frac{1}{\delta}u^{k_i}}(u^*_{\mu,0}, u^{k_i})$ for $H(u)$ as in (2.1), we obtain

$$
\begin{aligned}
\text{(3.11)} \qquad H(u^{k_i}) &\leq H(u^*_{\mu,0}) - \langle u^*_{\mu,0} - u^{k_i}, \mu p^{k_i} + \frac{1}{\delta} u^{k_i} \rangle \\
&= H(u^*_{\mu,0}) - \langle u^*_{\mu,0} - u^{k_i}, A^T w^{k_i} \rangle \\
&= H(u^*_{\mu,0}) - \langle A(u^*_{\mu,0} - u^{k_i}), w^{k_i} \rangle.
\end{aligned}
$$

On the other hand, by Cauchy-Schwarz inequality, we have

$$|\langle A(u^*_{\mu,0} - u^{k_i}), w^{k_i} \rangle| \leq \|A(u^*_{\mu,0} - u^{k_i})\|\|w^{k_i}\| \leq C\|A(u^*_{\mu,0} - u^{k_i})\|.$$

Letting $i \to \infty$, and noticing that $\lim_{i\to\infty} Au^{k_i} = A\widetilde{u} = f = Au^*_{\mu,0}$, we obtain that $\lim_{i\to\infty} |\langle A(u^*_{\mu,0} - u^{k_i}), w^{k_i} \rangle| = 0$. By letting $i \to \infty$ in (3.11), we have

$H(\widetilde{u}) \leq H(u^*_{\mu,0})$. This, together with $A\widetilde{u} = f$ and the uniqueness of $u^*_{\mu,0}$, implies that $\widetilde{u} = u^*_{\mu,0}$. $\qquad\square$

We observe that, in the proof of Theorem 2.4, one needs only the convexity of $\|u\|_1$ and the boundedness of the subgradient $\partial\|u\|_1$. Therefore, one can extend Theorem 2.4 to the following theorem. We omit the proof here, since it follows from that of Theorem 2.4.

Theorem 3.5. *Assume that $0 < \delta < \frac{1}{\|AA^T\|}$. Suppose that $J(u)$ is convex and $\partial J(u)$ is bounded. Then the sequence $\{u^k\}_{k\in\mathbb{N}}$ generated by (2.5) converges to the unique solution of (2.6).*

References

1. D. P. Bertsekas, *Constrained optimization and Lagrange multiplier methods*, Computer Science and Applied Mathematics, Academic Press Inc. [Harcourt Brace Jovanovich Publishers], New York, 1982. MR MR690767 (84k:90068)
2. D. P. Bertsekas, *Nonlinear Programming*, Athena Scientific, Belmont, Massachussets, 1999.
3. L. M. Brègman, *A relaxation method of finding a common point of convex sets and its application to the solution of problems in convex programming*, Ž. Vyčisl. Mat. i Mat. Fiz. **7** (1967), 620–631. MR MR0215617 (35 #6457)
4. M. Burger, E. Resmerita, and L. He, *Error estimation for Bregman iterations and inverse scale space methods in image restoration*, Computing **81** (2007), no. 2-3, 109–135. MR MR2354192 (2008k:94002)
5. J.-F. Cai, E.J. Candès, and Z. Shen, *A singular value thresholding algorithm for matrix completion*, 2008, preprint.
6. J.-F. Cai, R. Chan, L. Shen, and Z. Shen, *Restoration of chopped and nodded images by framelets*, SIAM J. Sci. Comput. **30** (2008), no. 3, 1205–1227. MR MR2398862
7. J.-F. Cai, R. H. Chan, L. Shen, and Z. Shen, *Convergence analysis of tight framelet approach for missing data recovery*, Adv. Comput. Math., to appear.
8. J.-F. Cai, R. H. Chan, L. Shen, and Z. Shen, *Simultaneously inpainting in image and transformed domains*, 2008, preprint.
9. J.-F. Cai, R. H. Chan, and Z. Shen, *A framelet-based image inpainting algorithm*, Appl. Comput. Harmon. Anal. **24** (2008), no. 2, 131–149. MR MR2393979
10. J.-F. Cai, R. H. Chan, and Z. Shen, *Simultaneous cartoon and texture inpainting*, 2008, preprint.
11. J.-F. Cai, S. Osher, and Z. Shen, *Linearized Bregman iterations for compressed sensing*, 2008, Math. Comp., to appear; see also UCLA CAM Report (08-06).
12. J.-F. Cai, S. Osher, and Z. Shen, *Linearized Bregman iterations for frame-based image deblurring*, SIAM J. Imaging Sci., **2** (2009), no. 1, 226–252.
13. J.-F. Cai and Z. Shen, *Deconvolution: A wavelet frame approach, II*, 2008, preprint.
14. E. J. Candès, *Compressive sampling*, International Congress of Mathematicians. Vol. III, Eur. Math. Soc., Zürich, 2006, pp. 1433–1452. MR MR2275736
15. E. J. Candès and J. Romberg, *Quantitative robust uncertainty principles and optimally sparse decompositions*, Found. Comput. Math. **6** (2006), no. 2, 227–254. MR MR2228740 (2007a:94035)
16. E. J. Candès, J. Romberg, and T. Tao, *Robust uncertainty principles: exact signal reconstruction from highly incomplete frequency information*, IEEE Trans. Inform. Theory **52** (2006), no. 2, 489–509. MR MR2236170 (2007e:94020)
17. E. J. Candes and T. Tao, *Near-optimal signal recovery from random projections: universal encoding strategies?*, IEEE Trans. Inform. Theory **52** (2006), no. 12, 5406–5425. MR MR2300700
18. E. Cetin, *Reconstruction of signals from Fourier transform samples*, Signal Processing, **16** (1989), 129–148.
19. E. Cetin, *An iterative algorithm for signal reconstruction from bispectrum*, IEEE Trans. Signal Proc., **39** (1991), 2621–2628.
20. A. Chai and Z. Shen, *Deconvolution: A wavelet frame approach*, Numer. Math. **106** (2007), no. 4, 529–587.

21. P. L. Combettes and J.-C. Pesquet, *Proximal thresholding algorithm for minimization over orthonormal bases*, SIAM J. Opt. **18** (2007), no. 4, 1351–1376 (electronic).
22. P. L. Combettes and V. R. Wajs, *Signal recovery by proximal forward-backward splitting*, Multiscale Model. Simul. **4** (2005), no. 4, 1168–1200 (electronic). MR MR2203849
23. J. Darbon and S. Osher, *Fast discrete optimization for sparse approximations and deconvolutions*, 2007, preprint.
24. I. Daubechies, M. Defrise, and C. De Mol, *An iterative thresholding algorithm for linear inverse problems with a sparsity constraint*, Comm. Pure Appl. Math. **57** (2004), no. 11, 1413–1457. MR MR2077704 (2005k:65106)
25. I. Daubechies, G. Teschke, and L. Vese, *Iteratively solving linear inverse problems under general convex constraints*, Inverse Probl. Imaging **1** (2007), no. 1, 29–46. MR MR2262744
26. D. L. Donoho, *De-noising by soft-thresholding*, IEEE Trans. Inform. Theory **41** (1995), no. 3, 613–627. MR MR1331258 (96b:94002)
27. D. L. Donoho, *Compressed sensing*, IEEE Trans. Inform. Theory **52** (2006), no. 4, 1289–1306. MR MR2241189 (2007e:94013)
28. M. Fortin and R. Glowinski, *Augmented Lagrangian methods*, Studies in Mathematics and its Applications, vol. 15, North-Holland Publishing Co., Amsterdam, 1983, Applications to the numerical solution of boundary value problems, Translated from the French by B. Hunt and D. C. Spicer. MR MR724072 (85a:49004)
29. T. Goldstein and S. Osher, *The split Bregman algorithm for l1 regularized problems*, 2008, UCLA CAM Report (08-29).
30. M. R. Hestenes, *Multiplier and gradient methods*, J. Optimization Theory Appl. **4** (1969), 303–320. MR MR0271809 (42 #6690)
31. J.-B. Hiriart-Urruty and C. Lemaréchal, *Convex analysis and minimization algorithms. I*, Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences], vol. 305, Springer-Verlag, Berlin, 1993, Fundamentals. MR MR1261420 (95m:90001)
32. P. J. Huber, *Robust regression: asymptotics, conjectures and Monte Carlo*, Ann. Statist. **1** (1973), 799–821. MR MR0356373 (50 #8843)
33. K. Ito and K. Kunisch, *The augmented Lagrangian method for equality and inequality constraints in Hilbert spaces*, Math. Programming **46** (1990), no. 3, (Ser. A), 341–360. MR MR1054143 (91i:90062)
34. S. Osher, M. Burger, D. Goldfarb, J. Xu, and W. Yin, *An iterative regularization method for total variation-based image restoration*, Multiscale Model. Simul. **4** (2005), no. 2, 460–489 (electronic). MR MR2162864 (2006c:49051)
35. S. Osher, Y. Mao, B. Dong, and W. Yin, *Fast linearized bregman iteration for compressed sensing and sparse denoising*, 2008, UCLA CAM Reprots (08-37).
36. M. J. D. Powell, *A method for nonlinear constraints in minimization problems*, Optimization (Sympos., Univ. Keele, Keele, 1968), Academic Press, London, 1969, pp. 283–298. MR MR0272403 (42 #7284)
37. F. Schöpfer, A. K. Louis, and T. Schuster, *Nonlinear iterative methods for linear ill-posed problems in Banach spaces*, Inverse Problems **22** (2006), no. 1, 311–329. MR MR2194197 (2006j:65151)
38. Y. Tsaig and D. L. Donono, *Extensions of compresssed sensing*, Signal Processing **86** (2005), 533–548.
39. W. Yin, *On linearized Bregman iterative algorithms*, private communication.
40. W. Yin, S. Osher, D. Goldfarb, and J. Darbon, *Bregman iterative algorithms for ℓ_1-minimization with applications to compressed sensing*, SIAM J. Imaging Sci. **1** (2008), no. 1, 143–168.

Temasek Laboratories, National University of Singapore
E-mail address: `tslcaij@nus.edu.sg`

Department of Mathematics, UCLA, 520 Portola Plaza, Los Angeles, CA 90095, USA
E-mail address: `sjo@math.ucla.edu`

Department of Mathematics, National University of Singapore, 2 Science Drive 2, Singapore 117543
E-mail address: `matzuows@nus.edu.sg`