A Splitting-based Iterative Method for Sparse Reconstruction

Liquan Kang, Ying Chen, Zefeng Yu, Heng Wu, Zijun Zheng, Shanzhou Niu*

School of Mathematics and Computer Sciences, Gannan Normal University, Ganzhou, 341000, China.

Abstract   In this paper, we study a ℓ_1-norm regularized minimization method for sparse solution recovery in compressed sensing and X-ray CT image reconstruction. In the proposed method, an alternating minimization algorithm is employed to solve the involved ℓ_1-norm regularized minimization problem. Under some suitable conditions, the proposed algorithm is shown to be globally convergent. Numerical results indicate that the presented method is effective and promising.

Keywords   Compressed sensing, Sparse reconstruction, ℓ_1-norm regularized minimization, Variable splitting, Convergence

1. Introduction

In this paper, we mainly focus on the following ℓ_1-norm regularized minimization problem

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

where $A \in R^{m \times n}(m \ll n)$ is a linear operator, and $b \in R^m$ is an observation, parameter $\lambda > 0$ is used to trade off ℓ_2 term and ℓ_1 term for minimization. The above problem (1) is well-known as the unconstrained basis pursuit denoising problem.

   Up to now, numerous approaches have been proposed and analyzed for solving the aforementioned problem (1) [7, 12, 23, 25]. Figueiredo et.al presented a GPSR [12] method, in which the ℓ_1 problem (1) was reformulated as a box-constrained quadratic program and was minimization by the Barzilai-Borwein gradient projection method [1] with an efficient nonmonotone line search. Iterative shrinkage-thresholding algorithm (ISTA) is one of the most widely studied methods for solving problem (1), which was first proposed for wavelet-based image deconvolution [13, 18], and then was applied to many other filed [10, 11, 21, 22]. Using operator splitting, Hale, Yin and Zhang derived the iterative shrinkage-thresholding fixed point continuation algorithm (FPC)[16, 17]. Another similar method is the sparse reconstruction algorithm SpaRSA, which was proposed by Wright, Nowak and Figueiredo in [24]. Recently, Beck and Teboulle proposed a FISTA algorithm with $O(1/k^2)$ convergence rate [4]. In [25], Yin, Osher, Goldfarb and Darbon proposed a Bregman iterative regularization method to solving the problem (1). Furthermore, a linearized Bregman method was proposed and analyzed subsequently in [7, 8, 26]. Using the spectral gradient projection method with an efficient Euclidean projection on ℓ_1-norm ball, Friedlander and Van den Berg [27] proposed a SPGL1 method for problem (1). Becker, Bobin and Candès [3] proposed a NESTA algorithm based on the Nesterov's smoothing technique to solve the problem (1). Moreover, an alternating direction

---

*Correspondence to: Email: maszniu@163.com

algorithms for ℓ_1-norm regularized minimization problem were proposed, which was derived from either the primal or the dual form of ℓ_1-norm problem [28].

In this paper, based on the variable splitting scheme [14, 20], we present the following ℓ_1-norm regularized minimization model:

$$\min_{x,y\in R^n} \frac{1}{2}\|Ax - b\|_2^2 + \lambda\|u\|_1, \tag{2}$$

$$\text{s.t. } \|x - u\|_2^2 \leq \varepsilon.$$

It is clear that the constrained problem (2) will reduce to (1) when $\varepsilon = 0$. However, in general, the constrained problem (2) is a strict inequality constrained problem for the effect of noise. The parameter $\varepsilon > 0$ is the standard deviation of additive noise. The above constrained problem can be transformed into an unconstrained problem as follows:

$$\min_{x,u\in R^n} \mathcal{J}(x,u) = \min_{x,u\in R^n} \frac{1}{2\lambda_1}\|Ax - b\|_2^2 + \frac{1}{2\lambda_2}\|u - x\|_2^2 + \|u\|_1, \tag{3}$$

where λ_1 and λ_2 are positive regularization parameters. In fact, this variable splitting method has been proposed in image restoration [15], where the authors proposed a fast TV minimization method for image restoration.

For the problem (3), as we can see that, if we fixed u, it is a smooth quadratic optimization problem and if we fixed x, the optimization problem on u has a closed form that it can be explicitly calculated by Iterative Shrinkage/Thresholding Algorithm.

The rest of this paper is organized as follows. In Section 2, the alternating minimization algorithm is developed to solve the proposed ℓ_1-norm minimization problem (3) and its global convergence of the iterative algorithm is established under some suitable conditions. In Section 3, numerical examples are presented to demonstrate the effectiveness of the proposed model. Finally, we have a conclusion section.

2. Optimization algorithm

In this section, we present a variable splitting method (VSM) to solve the problem (3). Starting from an initial guess u_0, this method yields a sequence

$$u_0 \rightarrow x_1 \rightarrow u_1 \rightarrow x_2 \rightarrow u_2 \rightarrow \cdots \rightarrow x_k \rightarrow u_k \cdots$$

such that

$$\begin{cases} \mathcal{S}_A(u_{k-1}) := x_k = \text{argmin}_x f(x) = \text{argmin}_x \frac{1}{2\lambda_1}\|Ax - b\|_2^2 + \frac{1}{2\lambda_2}\|u_{k-1} - x\|_2^2, & \text{(4a)} \\[2mm] \mathcal{S}_{\ell_1}(x_k) := u_k = \text{argmin}_u g(u) = \text{argmin}_u \frac{1}{2\lambda_2}\|u - x_k\|_2^2 + \|u\|_1. & \text{(4b)} \end{cases}$$

Then, we can obtain that

$$u_k = \mathcal{S}_{\ell_1}(\mathcal{S}_A(u_{k-1})).$$

Let $\mathcal{T}(\cdot) = \mathcal{S}_{\ell_1}(\mathcal{S}_A(\cdot))$, then

$$u_k = \mathcal{T}(u_{k-1}). \tag{5}$$

2.1. Minimization with respect to x

The subproblem (4a) is a quadratic convex optimization problem, which is equivalent to solving a linear system:

$$(A^T A + \frac{\lambda_1}{\lambda_2}I)x = A^T b + \frac{\lambda_1}{\lambda_2}u, \tag{6}$$

where $(A^T A + \frac{\lambda_1}{\lambda_2}I)$ is a symmetric positive definite matrix.

In this paper, the conjugate gradient method was utilized to solve (4a). The iterative scheme is as follows:

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

where $\alpha_k = \mathrm{argmin}_\alpha f(x_k + \alpha d_k)$, d_k is the conjugate gradient direction,

$$d_k = \begin{cases} -\nabla f(x_k) & \text{if } k = 0, \\ -\nabla f(x_k) + \beta_k d_{k-1} & \text{otherwise.} \end{cases}$$

where $\beta_k = \frac{\nabla f(x_k)^T y_{k-1}}{d_{k-1}^T y_{k-1}}$, $y_{k-1} = \nabla f(x_k) - \nabla f(x_{k-1})$.

Remark: The optimal step size can be determined by

$$\alpha_k = \frac{\lambda_2 (Ad_k)^T b + \lambda_1 d_k^T u_k - \lambda_2 (Ad_k)^T Ax_k - \lambda_1 d_k^T x_k}{\lambda_2 (Ad_k)^T Ad_k + \lambda_1 d_k^T d_k}.$$

2.2. Minimization with respect to u

Since u is separable in the subproblem (4b), each of its components can be independently obtained by the shrinkage operator in [9], i.e.,

$$u_k^i = \max\{x_k^i - \lambda_2, 0\}\mathrm{sgn}(x_k^i), i = 1, 2, \cdots, n \tag{7}$$

where u_k^i is the ith component of the u_k.

2.3. Convergence analysis

This section is devoted to analysis the convergence of the poposd algorithm for (3). First, let us give some definitions.

Definition 1
An operator $\mathcal{P}$ is called nonexpansive if, for any $x_1, x_2 \in R^N$, we have

$$\|\mathcal{P}(x_1) - \mathcal{P}(x_2)\| \leq \|x_1 - x_2\|.$$

If there exists some nonexpansive operator $\mathcal{A}$ and $\kappa \in (0,1)$ such that $\mathcal{P} = (1-\kappa)I + \kappa\mathcal{A}$, then $\mathcal{P}$ is called κ-averaged nonexpansive.

In the following, we give the Lemma 2.4 in [6].

Lemma 1
Let φ be convex and semicontinuous and $\beta > 0$. Suppose $\widehat{u}$ is defined as follows:

$$\widehat{u} = \mathrm{argmin}_u \|u - x\|_2^2 + \beta\varphi(u).$$

Define $\mathcal{S}$ such that $\widehat{u} = \mathcal{S}(x)$ for each x, Then $\mathcal{S}$ is $\frac{1}{2}$-averaged nonexpansive.

It is known that $\|u\|_1$ is convex and semicontinuous function. By Lemma 1, we can derive that $\mathcal{S}_{\ell_1}$ is $\frac{1}{2}$-averaged nonexpansive.

Lemma 2
The operator $\mathcal{T}$ defined in (5) is nonexpansive.

Proof: Since $A^T A + \frac{\lambda_1}{\lambda_2}I$ is symmetric positive definite and its eigenvalues are larger than $\frac{\lambda_1}{\lambda_2}$, then for any u, v

$$\|\frac{\lambda_2}{\lambda_1}(A^T A + \frac{\lambda_1}{\lambda_2}I)^{-1}(u-v)\|_2 \leq \|u-v\|_2.$$

Therefore,

$$
\begin{aligned}
\|\mathcal{T}(u) - \mathcal{T}(v)\|_2 &= \|\mathcal{S}_{\ell_1}(\mathcal{S}_A(u)) - \mathcal{S}_{\ell_1}(\mathcal{S}_A(v))\| \\
&\leq \|\mathcal{S}_A(u) - \mathcal{S}_A(v)\|_2 \\
&= \|(A^T A + \frac{\lambda_1}{\lambda_2}I)^{-1}(A^T b + \frac{\lambda_1}{\lambda_2}u) - (A^T A + \frac{\lambda_1}{\lambda_2}I)^{-1}(A^T b + \frac{\lambda_1}{\lambda_2}v)\|_2 \\
&= \|\frac{\lambda_2}{\lambda_1}(A^T A + \frac{\lambda_1}{\lambda_2}I)^{-1}(u - v)\| \\
&\leq \|u - v\|_2.
\end{aligned}
\tag{8}
$$

The result follows. $\square$

Definition 2
Let C be a convex closed set in Banach space X, a mapping $F : C \to C$ is asymptotically regular if, for any $x \in C$,

$$
\lim_{n \to \infty} \|F^{n+1}(x) - F^n(x)\|_2 = 0.
$$

Lemma 3
Let $\{u_k\}$ be generated by (5), then we have

$$
\lim_{k \to \infty} \|u_{k+1} - u_k\|_2 = 0.
$$

Proof: We consider the Taylor series expansion of $\mathcal{J}(x, u_k)$ in the first variable, and we have

$$
\begin{aligned}
\mathcal{J}(x_k, u_k) &= \mathcal{J}(x_{k+1}, u_k) + (x_k - x_{k+1})^T \frac{\partial \mathcal{J}}{\partial x}(x_{k+1}, u_k) \\
&\quad + \frac{1}{2}(x_k - x_{k+1})^T \frac{\partial^2 \mathcal{J}}{\partial x^2}(x_k - x_{k+1}).
\end{aligned}
\tag{9}
$$

Since $x_{k+1} = \operatorname{argmin}\mathcal{J}(x, u_k)$, it holds that

$$
\frac{\partial \mathcal{J}}{\partial x}(x_{k+1}, u_k) = 0,
$$

and

$$
\frac{\partial^2 \mathcal{J}}{\partial x^2}(x_{k+1}, u_k) = A^T A + \frac{\lambda_1}{\lambda_2}I.
$$

As $(A^T A + \frac{\lambda_1}{\lambda_2}I)$ is symmetric positive definite and its eigenvalues are larger than $\frac{\lambda_1}{\lambda_2}$, then we can obtain that

$$
\begin{aligned}
\mathcal{J}(x_k, u_k) - \mathcal{J}(x_{k+1}, u_k) &= \frac{1}{2}(x_k - x_{k+1})^T(A^T A + \frac{\lambda_1}{\lambda_2}I)(x_k - x_{k+1}) \\
&\geq \frac{\lambda_1}{\lambda_2}\|x_k - x_{k+1}\|_2^2.
\end{aligned}
\tag{10}
$$

Since $u_{k+1} = \operatorname{argmin}\mathcal{J}(x_{k+1}, u)$, then $\mathcal{J}(x_{k+1}, u_{k+1}) \leq \mathcal{J}(x_{k+1}, u_k)$, therefore,

$$
\begin{aligned}
\mathcal{J}(x_k, u_k) - \mathcal{J}(x_{k+1}, u_{k+1}) &\geq \mathcal{J}(x_k, u_k) - \mathcal{J}(x_{k+1}, u_k) \\
&\geq \frac{\lambda_1}{\lambda_2}\|x_k - x_{k+1}\|_2^2.
\end{aligned}
\tag{11}
$$

Since $\mathcal{S}_{\ell_1}$ is nonexpansive, by Lemma 1, we have that

$$
\|x_k - x_{k+1}\|_2^2 \geq \|\mathcal{S}_{\ell_1}(x_k) - \mathcal{S}_{\ell_1}(x_{k+1})\|_2 = \|u_k - u_{k+1}\|_2^2.
$$

Hence, we obtain that

$$\mathcal{J}(x_k, u_k) - \mathcal{J}(x_{k+1}, u_{k+1}) \geq \frac{\lambda_1}{\lambda_2} \|u_k - u_{k+1}\|_2^2.$$

It follows that $\sum_{k=0}^{\infty} \|u_k - u_{k+1}\|_2^2$ is bounded, and follows that $\lim_{k\to\infty} \|u_{k+1} - u_k\|_2 = 0$. □

Taking the Lemma 3 into account, we can prove that the operator $\mathcal{T}$ is asymptotically regular.

Lemma 4

For any initial guess u_0, suppose u_k is generated by (5); then $\mathcal{T}$ is asymptotically regular, i.e.,

$$
\begin{aligned}
\lim_{k\to\infty} \|\mathcal{T}^{k+1}(u_0) - \mathcal{T}^k(u_0)\|_2 &= \lim_{k\to\infty} \|\mathcal{T}^k(u_1) - \mathcal{T}^{k-1}(u_1)\|_2 \\
&\vdots \\
&= \lim_{k\to\infty} \|\mathcal{T}(u_k) - \mathcal{T}(u_{k-1})\|_2 \\
&= \lim_{k\to\infty} \|u_{k+1} - u_k\|_2 \\
&= 0.
\end{aligned}
\tag{12}
$$

To proof the set of the minimum of $\mathcal{J}(x, u)$ is nonempty, we shall show the coerciveness of $\mathcal{J}$.

Definition 3

A function $\phi : R^N \to R$ is proper over a set $X \subset R^N$ if $\phi(x) < \infty$ at least one $x \in X$ and $\phi(x) > -\infty$ for all $x \in X$. A function $\phi : R^N \to R$ is coercive over a set $X \subset R^N$ if for every sequence $\{x_k\} \in X$ such that $\|x_k\| \to \infty$, we have $\lim_{k\to\infty} \phi(x) \to \infty$.

Lemma 5

The function $\mathcal{J}(x, u)$ defined in (3) is coercive.

Proof:

$$
\begin{aligned}
\mathcal{J}(x, u) &= \frac{1}{2\lambda_1} \|Ax - b\|_2^2 + \frac{1}{2\lambda_2} \|u - x\|_2^2 + \|u\|_1 \\
&\geq \frac{1}{2\lambda_1} \|Ax - b\|_2^2 + \frac{1}{2\lambda_2} \|u - x\|_2^2 + \|u\|_2 \\
&= \frac{1}{2} \left\| \begin{pmatrix} \frac{1}{\sqrt{\lambda_1}}A & \mathbf{O} \\ \frac{1}{\sqrt{\lambda_2}}I & -\frac{1}{\sqrt{\lambda_2}}I \end{pmatrix} \begin{pmatrix} x \\ u \end{pmatrix} - \begin{pmatrix} b \\ \mathbf{0} \end{pmatrix} \right\|_2^2 + \left\| \begin{pmatrix} \mathbf{0} & I \end{pmatrix} \begin{pmatrix} x \\ u \end{pmatrix} \right\|_2.
\end{aligned}
\tag{13}
$$

Using the matrix factorization, we note that

$$
\begin{pmatrix} \frac{1}{\sqrt{\lambda_1}}A & \mathbf{O} \\ \frac{1}{\sqrt{\lambda_2}}I & -\frac{1}{\sqrt{\lambda_2}}I \\ \mathbf{0} & I \end{pmatrix} = \begin{pmatrix} I & \mathbf{O} & \mathbf{O} \\ \mathbf{O} & I & \mathbf{O} \\ \mathbf{O} & -\sqrt{2\lambda_2}I & I \end{pmatrix} \begin{pmatrix} \frac{A}{\sqrt{\lambda_1}} & \mathbf{O} \\ \frac{1}{\sqrt{\lambda_2}}I & -\frac{1}{\sqrt{\lambda_2}}I \\ I & \mathbf{O} \end{pmatrix}.
$$

It is obvious that the above matrix is full rank as Null(A) ∩ Null(I)=∅. In fact, Null(A) ∩ Null(I)={**0**}, and **0** is not taken into consideration in practice. therefore, when $\left\| \begin{pmatrix} x \\ u \end{pmatrix} \right\|_2 \to \infty$, we have $\mathcal{J}(x, u) \to \infty$. The result follows

. □

Lemma 6

The the set of fixed points of $\mathcal{T}$ is nonempty.

Proof: Since the objective function $\mathcal{J}$ is coercive, the set of minimizers of $\mathcal{J}$ is nonempty. Assume that (x^*, u^*) is a minimizer of $\mathcal{J}(x, u)$, i.e.,

$$\frac{\partial \mathcal{J}}{\partial x}(x^*, u^*) = 0, \quad \frac{\partial \mathcal{J}}{\partial x}(x^*, u^*) = 0.$$

And

$$
\begin{cases}
x^* = \mathcal{S}_A(u^*) = \mathrm{argmin}\mathcal{J}(\cdot, u^*), \\
u^* = \mathcal{S}_{\ell_1}(x^*) = \mathrm{argmin}\mathcal{J}(x^*, \cdot).
\end{cases}
$$

Hence we get $u^* = \mathcal{S}_{\ell_1}(\mathcal{S}_A(u^*)) = \mathcal{T}(u^*)$, and u^* is a fixed point of $\mathcal{T}$. The result follows.   □

We remark that the objective function $\mathcal{J}$ in (3) is strictly convex as A is a matrix of full column rank, Therefore, a fixed point of $\mathcal{T}$ is also a global minimizer of $\mathcal{J}$. According to the Opial theorem [19], the sequence $\{u_k\}$ converges to a fixed point of $\mathcal{J}$, i.e., a minimizer of $\mathcal{J}$.

Theorem 1
For any initial guess $u^0 \in R^n$, suppose $\{u_k\}$ is generated by (5), and u_k converges to a stationary point of $\mathcal{J}$.

3. Numerical experiments

In this section, we test the performance of the variable splitting method for the ℓ_1-norm regularized minimization problem with application to computed tomography (CT) image reconstruction and compressed sensing.

All the programs were coded in MATLAB R2007a and run on a personal computer with an Intel Core 2 Duo CPU at 2.5 GHz and 2GB of memory.

3.1. Compressed sensing problem

In this subsection, the SVM algorithm was applied to solve the compressed sensing with comparison to FISTA method. In the numerical test, the original sparse signal xs and the matrix A are randomly generated, and the observation b is generated according to b=A*xs+ sigma*randn(m,1), where sigma is the standard deviation of additive Gaussian noise. Let k denote the number of nonzero components in xs. The stopping criterion was adopted as follows:

$$
\frac{\|x_{k+1} - x_k\|}{\|x_{k+1}\|} < 5 \times 10^{-3},
$$

where $\{x_k\}$ is the sequence generated by the test algorithms. And we measure the quality of recovered signal by its mean square error (MSE), i.e.,

$$
\mathrm{MSE} = \frac{\|\overline{x} - x^*\|}{n},
$$

where $\overline{x}$ is the signal restored by certain algorithm, x^* is the original signal, n is the length of the signal.

The value of the objective function (log10(fun)) and MSE are given in Table 1 and Table 2, respectively. Figure 1 and Figure 3 show the restored signal by FISTA algorithm and SVM algorithm, respectively. Figure 2 and Figure 4 show the objective function value and MSE vs the iteration numbers, respectively. The results indicate that the VSM algorithm outperforms FISTA algorithm.

Table 1. Comparison results in CS (sigma=0.01).

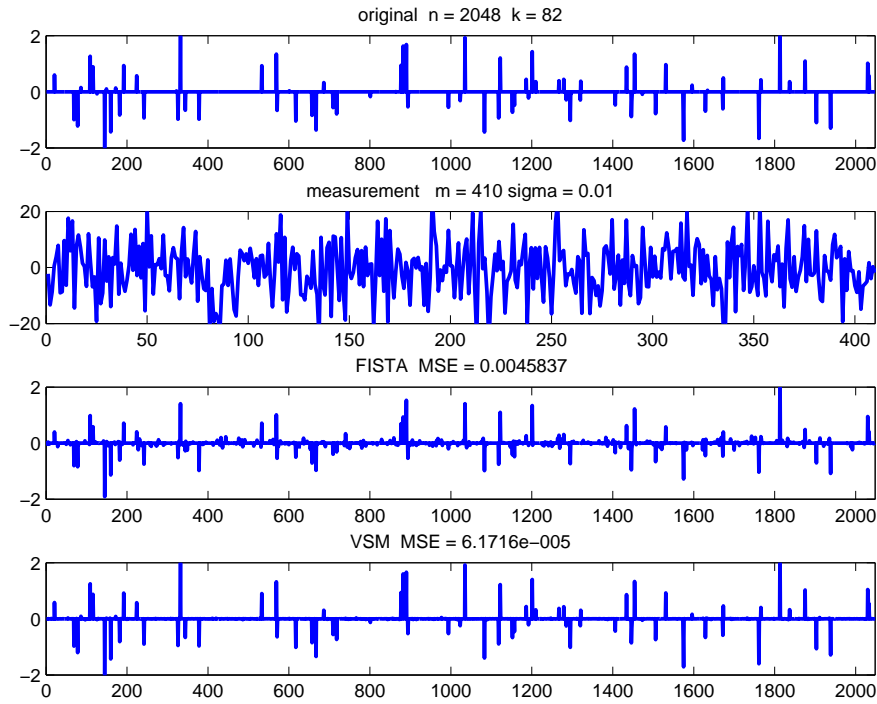| n=2048		FISTA			VSM		
i (m/n)	j (k/m)	iter	log10(fun)	MSE	iter	log10(fun)	MSE
0.3	0.1	232.000	2.515	2.312e-004	90.000	2.516	2.356e-005
0.3	0.2	195.000	2.963	1.315e-002	158.000	2.903	2.063e-006
0.2	0.1	260.000	2.225	1.104e-003	108.000	2.192	9.684e-006
0.2	0.2	203.000	2.703	1.016e-002	165.000	2.653	2.236e-004
0.1	0.1	278.000	1.785	4.486e-004	121.000	1.768	5.094e-006
0.1	0.2	207.000	2.197	9.346e-003	149.000	2.183	5.007e-003

Figure 1. The original signal (top), the measurement (second), and the reconstructed signals by FISTA (third) and VSM (bottom).
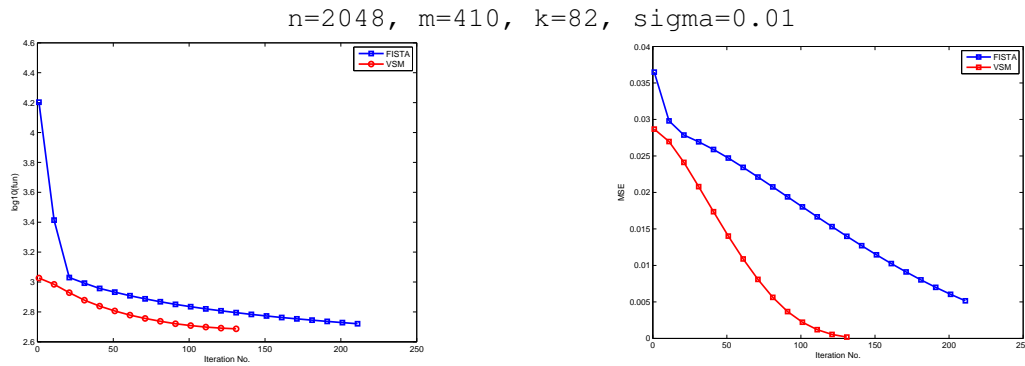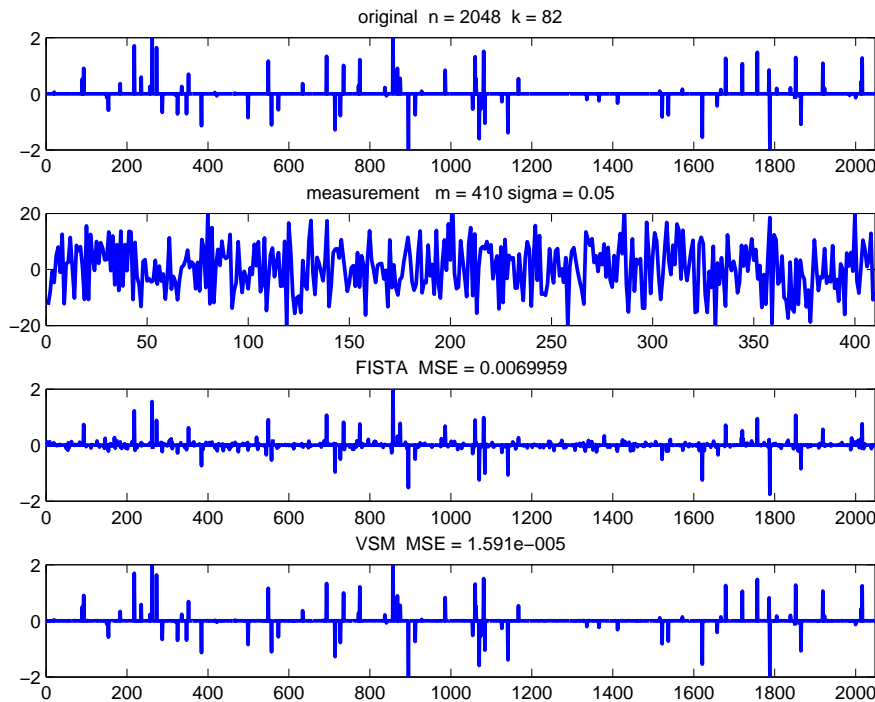

Figure 2. Function value vs. iteration (left) and MSE vs. iteration (right) for FISTA and VSM.

Table 2. Comparison results in CS (`sigma=0.05`).

n=2048		FISTA			VSM		
i (m/n)	j (k/m)	iter	log10(fun)	MSE	iter	log10(fun)	MSE
0.3	0.1	232.000	2.673	8.374e-005	94.000	2.692	4.676e-005
0.3	0.2	201.000	2.985	1.054e-002	147.000	2.933	4.523e-006
0.2	0.1	262.000	2.302	2.216e-004	105.000	2.307	1.405e-005
0.2	0.2	209.000	2.771	9.721e-003	177.000	2.731	1.105e-004
0.1	0.1	266.000	1.961	3.048e-004	124.000	1.958	6.486e-006
0.1	0.2	206.000	2.131	7.515e-003	145.000	2.119	4.442e-003

Figure 3. The original signal (top), the measurement (second), and the reconstructed signals by FISTA (third) and VSM (bottom).
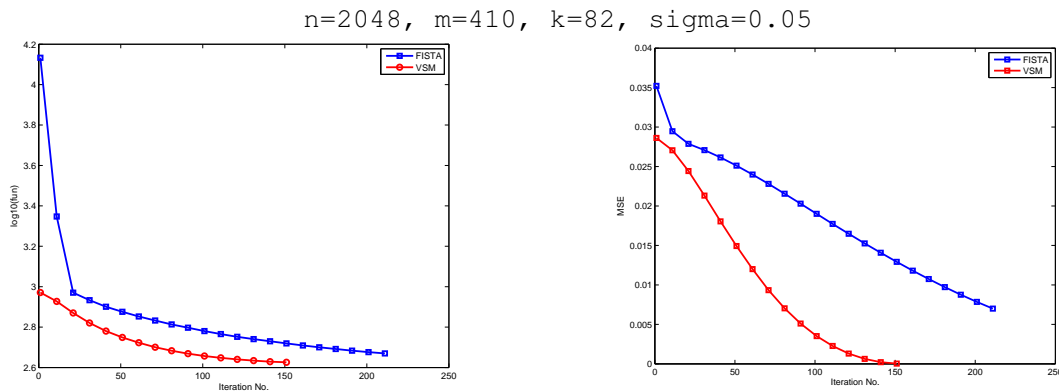

Figure 4. Function value vs. iteration (left) and MSE vs. iteration (right) for FISTA and VSM.

3.2. Sparse-view X-ray CT image reconstruction

To validate and evaluate the performance of the SVM method, in this subsection, a modified digital Shepp-Logan phantom (Figure 5) was used in the experiments. In this subsection, the total variation regularization was used in (3) for sparse-view X-ray CT image reconstruction, the (4b) is solved by Chambolle's algorithm [5]. For the CT projection simulation, we chose a geometry that was representative of a monoenergetic fan-beam CT scanner setup. This geometry was modeled with 672 bins on a 1-D detector for 2-D image reconstruction, and projection data with 25 projection numbers of views was simulated at equal angular increment on 360 around the phantom. The imaging parameters of the CT scanner were as follows: (1) the distance from the detector arrays to the x-ray source was 1040 mm; (2) the distance from the rotation center to the x-ray source was 570 mm; (3) and the space
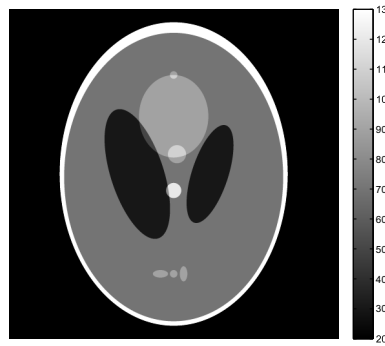
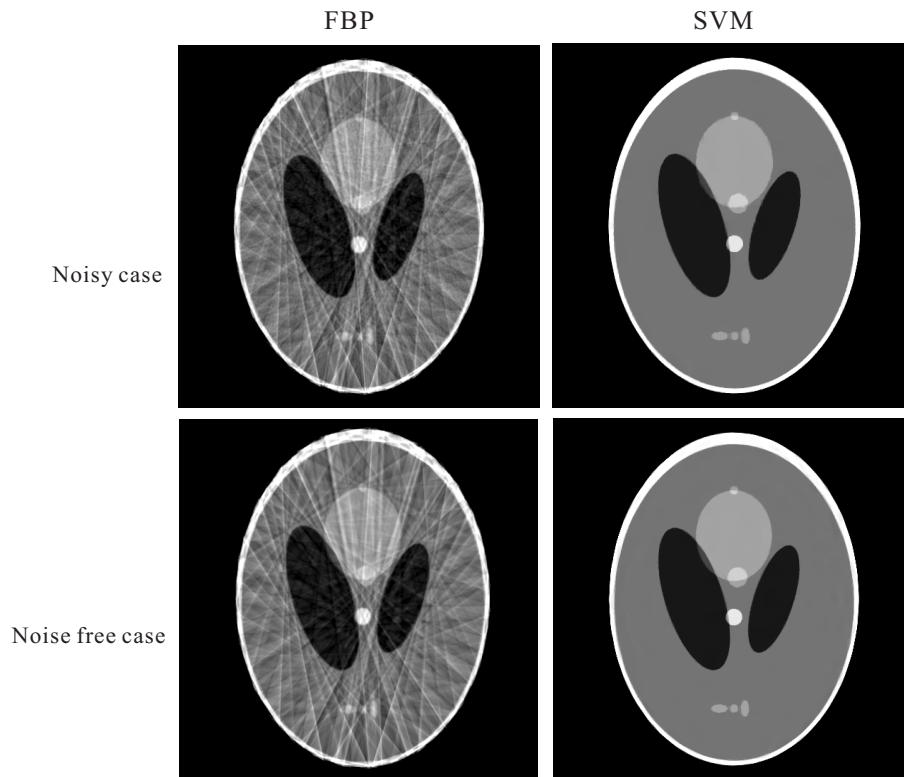Figure 5. A modified digital Shepp-Logan phantom.

Figure 6. Images reconstructed by FBP and SVM algorithms from 25-view projections.

of each detector bin was 1.407 mm. Each phantom is composed by 512×512 square pixels with each pixel size of 1.25 mm $\times$ 1.25 mm.

The results from the 25-view projection data in the noise-free and noisy cases are shown in Fig. 6. Serious noise-induced streak artifacts can be observed in the images reconstructed by the FBP method with ramp filter. The SVM algorithm can yield accurate reconstructions from spare-view projection data.

4. Conclusion

In this paper, we presented a splitting-based iterative method for sparse solution recovery in compressed sensing and X-ray CT image reconstruction. The numerical experiment results show that the proposed method is effective and promising.

5. Acknowledgement

REFERENCES

1. J. Barzilai and J. Borwein, Two point step size gradient methods, IMA J Numer. Anal., vol. 8, pp. 141-148, 1988.
2. E. G. Birgin, J. M. Martínez and M. Raydan, Nonmonotone spectral projected gradient methods on convex sets, SIAM J. Optim., vol. 10, pp. 1196-1121, 2000.
3. S. Becker, J. Bobin and E. Candès, NESTA: A fast and accurate first-order method for sparse recovery, SIAM J. Imag. Sci., vol. 4, pp. 1-39, 2011.
4. A. Beck and M. Teboulle, A fast iterative shrinkage-thresholding algorithm for linear inverse problems, SIAM J. Imag. Sci., vol. 2, no. 1, pp. 183-202, 2009.
5. A. Chambolle, An algorithm for total variation minimization and applications, Journal of Mathematical imaging and vision, vol. 20, pp. 89-97, 2004.
6. L. Combettes and R. Wajs, Signal recovery by proximal forward-backward splitting, Multiscale Model. Simul., vol. 4, pp. 1168-1200, 2005.
7. J. Cai, S. Osher and Z. Shen, Linearized Bregman iterations for compressive sensing, Mathematics of Computation, vol. 78, no. 267, pp. 1515-1536, 2009.
8. J. Cai, S. Osher and Z. Shen, Convergence of the linearized Bregman iteration for ℓ_1-norm minimization, Mathematics of Computation, vol. 78, no. 268, pp. 2127-2136, 2009.
9. L. Donoho, De-noising by soft-thresholding, Information Theory, IEEE Transactions on, vol. 41, no. 3, pp. 613-627, 1995..
10. I. Daubechies, M. Defriese and C. De Mol, An iterative thresholding algorithm for linear inverse problems with a sparsity constraint, Commun. Pure Appl. Math., vol. LVII, pp. 1413-1457, 2004.
11. M. Elad, Why simple shrinkage is still relevant for redundant representations? IEEE Trans. Inform. Theory, vol. 52, no.12, pp. 5559-5569, 2006.
12. M. Figueiredo, R. Nowak and S. J. Wright, Gradient projection for sparse reconstruction: application to compressed sensing and other inverse problems, IEEE J. Sel. Top. Signa., vol. 1, pp. 586-597, 2007.
13. M. Figueiredo and R. Nowak, An EM algorithm for wavelet-based image restoration, IEEE Trans. Imag. Process., vol. 12, no. 8, pp. 906-916, 2003.
14. B.S. He, M. Tao and X.M. Yuan, A splitting method for separable convex programming, IMA Journal of Numerical Analysis, vol. 35, no. 1, pp. 394-426, 2015.
15. Y. Huang, K. Ng and W.Wen, A fast total variation minimization method for image restoration, SIAM Multiscale Model. Simul. vol. 7, no. 2, pp. 774-795, 2008.
16. E. Hale, W. Yin and Y. Zhang, Fixed-point continuation for ℓ_1-minimization: methodology and convergence, SIAM J. Optim., vol. 19, no. 3, pp. 1107-1130, 2008.
17. E. Hale, W. Yin and Y. Zhang, Fixed-point continuation applied to compressed sensing: implementation and numerical experiments, J Comput. Math., vol. 28, no. 2, pp. 170-194, 2010.
18. R. Nowak and M. 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, 2001.
19. Z. Opial, Weak convergence of the sequence of successive approximations for nonexpansive mappings, Bull. Amer. Math. Soc.,vol. 73, pp. 591-597, 1967.
20. Z. Shen, Z. Geng and J. Yang, Image reconstruction from incomplete convolution data via total variation regularization, Stat., Optim. Inf. Comput., Vol. 3, pp. 1-14, 2015.
21. J. Starck, E. Candès and D. Donoho, Astronomical image representation by the curvelet transform, Astron. Astrophys., vol. 398, pp. 785-800, 2003.
22. J.-L. Starck, M. Nguyen and F. Murtagh, Wavelets and curvelets for image deconvolution: a combined approach, Signal Process., vol. 83, pp. 2279-2283, 2003.
23. W. Wang and Q. Wang, Approximated Function Based Spectral Gradient Algorithm for Sparse Signal Recovery, Stat., Optim. Inf. Comput., Vol. 2, pp 10-20, 2014.
24. S. Wright, R. Nowak and M. Figueiredo, Sparse reconstruction by separable approximation, in Proceedings of the International Conference on Acoustics, Speech, and Signal Processing, October 2008.

25. W. Yin, S. Osher, D. Goldfarb and J. Darbon, Bregman iterative algorithms for ℓ_1-minimization with applications to compressed sensing, SIAM J. Imag. Sci., vol. 1, no. 1, pp. 143-168, 2008.
26. W. Yin, The linearized Bregman method: reviews, analysis, and generalizations, CAAM TR09C02, Rice University.
27. M. Friedlander and E. Van den Berg, Probing the pareto frontier for basis pursuit solutions, SIAM J. Sci. Comput., vol. 31, no. 2, pp. 890-912, 2008.
28. J. Yang and Y. Zhang, Alternating direction algorithms for ℓ_1-problem in compressive sensing, SIAM J. Sci. Comput., vol. 33, pp. 250-278, 2011.