

AN EFFICIENT DYKSTRA-LIKE PROXIMAL ALGORITHM FOR SENSE MODEL SIGNAL RECONSTRUCTION

XU Wei-zhi ¹, YIN Hong ², JIANG Ling-yun ¹

(*1.School of Statistics, Hubei University of Economics, Wuhan 430205, China*)

(*2.School of Information, Renmin University of China, Beijing 100872, China*)

Abstract: We consider the problem of the efficient reconstruction from partial Fourier data in Magnetic Resonance (MR) images. Based on Dykstra-like proximal method and Bregman method, we propose an accelerated Dykstra-like proximal algorithm (ADPA-BI) for SENSE model signal reconstruction, and obtain the proof of its convergence property. Our numerical simulations on recovering MR images indicate that the proposed method is more efficient than the classic Split Bregman Iteration (SBI) method.

Keywords: MRI reconstruction; compressed sensing; Bregman method; Dykstra-like proximal algorithm; SENSE model

2010 MR Subject Classification: 90C30

Document code: A

Article ID: 0255-7797(2015)04-0881-08

1 Introduction

Magnetic Resonance Imaging (MRI) plays a very important role today in medical diagnosis because of its non-invasive manner and excellent depiction of soft tissue changes. Its use of nonionizing radio frequency emission for image acquisition is considered safe for repeated use in a clinical setting. Both fast acquisition and efficient reconstruction in MR image are very crucial to apply in clinical settings of practical MR image.

Compressed sensing theory shows that it is possible to accurately reconstruct the Magnetic Resonance images from highly undersampled k -space data and therefore significantly reduce the scanning duration. Lustig et al. [1] proposed their pioneering work for the MR images reconstruction. Their method can effectively reconstruct MR images with only 20% sampling. In order to improve the construction result, it considers the SENSE model that the constrained objective involves both a discrete gradient and a wavelet transform, which is formulated as follows:

$$\min_u \alpha \int_{\Omega} |\nabla u| dx + \beta \int_{\Omega} |\Psi u| dx \text{ such that } \|R\mathcal{F}u - f\|_2^2 < \sigma^2, \quad (1.1)$$

* **Received date:** 2014-09-10

Accepted date: 2014-12-24

Biography: Xu Weizhi(1977-), male, Xiantao, Hubei, lecturer, major in data analysis, scientific computing and approximation theory.

Corresponding author: Yin Hong.

where $\alpha, \beta > 0$, Ψ is a wavelet transform, \mathcal{F} represents the Fourier transform matrix, f represents the observed “k-space” data, and σ represents the variance of the signal noise. The matrix R represents a “row selector” matrix, which comprises a subset of the rows of an identity matrix. The related unconstrained problem is

$$\min_u \{J(u) + \frac{1}{2} \int_{\Omega} (\mathcal{A}u - f)^2 dx\}, \quad (1.2)$$

where $J(u) = \alpha \|\nabla u\|_1 + \beta \|\Psi u\|_1$ and $\mathcal{A} = R\mathcal{F}$. It is based on the fact that the piecewise smooth MR images of organs can be sparsely represented by the wavelet basis and framelet and should have small total variations. Since MR images commonly possess a blocky structure, the use of TV norm in regularization exploits image sparsity and preserves edges. In [8], a kernel-based online quantile regression algorithm associated with a sequence of insensitive pinball loss functions is considered. A fast reconstruction algorithm can make CS-MRI more clinically applicable.

However, the computational challenge for solving the TVL1-L2 Model (1.2) comes from the combination of two issues: one is the nondifferentiability of the TV and l_1 terms, and the other one is the large size and severe ill-conditioning of the matrix \mathcal{A} . The conjugate gradient (CG) and PDE methods were used to attack it [1]. However, they are very slow and impractical for real MR images. The computation became the bottleneck that prevented this model (1.2) from being used in practical MR image reconstruction. Therefore, the key problem in compressed MR image reconstruction is thus to develop efficient algorithms to solve problem (1.2) with nearly optimal reconstruction accuracy. The main purpose of the paper is to develop a fast numerical algorithm for solving the model (1.2).

On the other hand, recently the Bregman iteration has been successfully used to solve a wide variety of constrained l_1 -regularized optimization problems in compressed sensing, image processing and computer vision [2]. The basic idea is to transform a constrained optimization problem to a series of unconstrained problems. In each unconstrained problem, the object function is defined by Bregman distance of a convex functional.

The Bregman distance of a convex functional $J(u)$ is defined as the following nonnegative quantity

$$D_J^p(u, v) \equiv J(u) - J(v) - \langle p, u - v \rangle, \quad (1.3)$$

where $p \in \partial J(v)$.

Instead of solving (1.2) once, the Bregman iterative regularization procedure of Osher et al. [2] solves a sequence of convex problems

$$u^{k+1} \leftarrow \min_u D_J^{p^k}(u, u^k) + \frac{1}{2} \|\mathcal{A}u - f\|^2 \quad (1.4)$$

for $k = 0, 1, \dots$ starting with $u^0 = 0$ and $p^0 = 0$ (hence, for $k = 0$, one solves the original problem (1.2)). Since $J(u)$ is not differentiable everywhere, the subdifferential of $J(u)$ may contain more than one element. However, from the optimality of u^{k+1} in (1.4), it follows that $0 \in \partial J(u^{k+1}) - p^k + \mathcal{A}u^{k+1} - f$. The difference between (1.2) and (1.4) is in the use

of regularization. While (1.2) regularizes u by directly minimizing the energy $J(u)$, (1.4) regularizes u by minimizing the Bregman distance of u to a previous solution u^k based on the energy $J(u)$. In [2], it was proved that the problem (1.4) was equivalent to the solution of unconstrained problems of the form

$$u^{k+1} = \arg \min_u J(u) + \frac{1}{2} \|Au - f^k\|^2, \quad (1.5)$$

$$f^{k+1} = f^k + f - Au^{k+1} \quad (1.6)$$

for $k = 0, 1, \dots$ starting with $u^0 = 0, f^0 = f$.

We denote $\text{prox}_\rho(g)(x)$ as the proximal map of function g at the point $x \in \mathbb{R}^N$.

Definition 1 The proximal map: given a continuous convex function $g(x)$ and any scalar $\rho > 0$, the proximal map associated with function g is defined as follows

$$\text{prox}_\rho(g)(x) := \arg \min_u \{g(u) + \frac{1}{2\rho} \|u - x\|^2\}. \quad (1.7)$$

Combined the efficient acceleration scheme in FASTA method [5] and Dykstra-like proximal algorithm method [4], Xu [7] proposed an Accelerated Dykstra-like Proximal Algorithm (ADPA) for the image reconstruction subproblem (1.5). The algorithm ADPA can write as follows:

Algorithm 1 [7] ADPA(Accelerated Dykstra-like Proximal Algorithm)

1. Initial: $\rho = 1/L, \alpha, \beta > 0, \gamma_1, \gamma_2 > 0, t^1 = 1, z_1^1 = z_2^1 = \mathcal{F}^{-1}f$.
2. While it holds stop criterion;
3. $x_1 = \text{prox}_\rho(2\alpha\|x\|_{TV})(z_1^k)$;
4. $x_2 = \text{prox}_\rho(2\beta\|\Psi x\|_1)(z_2^k)$;
5. $x^k = \frac{\|x_2\|_{TV}}{\|\Psi x_1\|_1} x_1 + (1 - \frac{\|x_2\|_{TV}}{\|\Psi x_1\|_1}) x_2$;
6. $x^k = \text{project}(x^k, [l, u])$;
7. $t^{k+1} = (1 + \sqrt{1 + 4(t^k)^2})/2$;
8. $r^{k+1} = x^k + \frac{t^k - 1}{t^{k+1}}(x^k - x^{k-1})$;
9. $x^{k+1} = r^{k+1} - \rho \mathcal{A}^T(\mathcal{A}r^{k+1} - f)$;
10. $z_2^{k+1} = \gamma_1(z_1^k - x_1) + x^{k+1}$;
11. $z_1^{k+1} = \gamma_2(z_2^k - x_2) + x^{k+1}$;
12. end while.

In the step $x^k = \text{project}(x^k, [l, u])$, the function $y = \text{project}(x, [l, u])$ is defined as

$$y = \text{project}(x, [l, u]) = \begin{cases} l, & \text{if } x < l, \\ x, & \text{if } l \leq x \leq u, \\ u, & \text{if } x > u, \end{cases}$$

where $[l, u]$ is the range of x . Balance parameter τ depends on the original image and its noise. We investigate specially the adaptive weight for τ , then at the s -th iteration $\tau(s) = \frac{\|\nabla v^{(s)}\|_1}{\|\Psi u^{(s)}\|_1}$, where $\|\nabla v^{(s)}\|_1$ is the total variation semi-norm of the s -th iteration $v^{(s)}$ and $\|\Psi u^{(s)}\|_1$ is the

l_1 -norm of the wavelet transform for the s -th iteration $u^{(s)}$. γ_1 and γ_2 can be chosen as two monotone sequences as limited to 0 (e.g. $\gamma_1 = \gamma_2 = 1/k, e^{-\frac{1}{\sqrt[2]{k}}}$ or $e^{-\frac{1}{\sqrt[4]{k}}}$ (k is the iteration counter)). The total cost of each iteration in the algorithm is $\mathcal{O}(N \log(N))$. The algorithm ADPA has the convergence property as follows:

Theorem 1 [7] When $0 < \rho < \frac{1}{\|\mathcal{A}^T \mathcal{A}\|}$, both sequences $(z_1^n)_{n \in \mathbb{N}}$ and $(z_2^n)_{n \in \mathbb{N}}$ generated by Algorithm 1 converge to the solution to problem (1.2).

In this paper, we apply the Bregman iteration to the algorithm ADPA, and accelerate further to reconstruct magnetic resonance images from the model (1.2).

2 Main Results

For total variation regularity $J(u) = \mu \|\nabla u\|_1$ and wavelet regularity $J(u) = \mu \|\Psi u\|_1$, the FISTA can very efficiently solve the subproblems (1.5). But it cannot efficiently solve the composite L1 and TV regularization problem $J(u) = \alpha \|\nabla u\|_1 + \beta \|\Psi u\|_1$ in the SENSE model (1.2), since no efficient algorithm exists to solve the step

$$u^k = \text{prox}_\rho(\alpha \|u\|_{TV} + \beta \|\Psi u\|_1)(u_g). \quad (2.1)$$

However, the algorithm ADPA can efficiently deal the problem (2.1) by the Dykstra-like proximal algorithm and the FISTA technique [7]. By the Bregman iteration, problem (1.1) could be reduced to a sequence of unconstrained problems of the form

$$u^{k+1} = \arg \min_u J(u) + \frac{1}{2} \|R\mathcal{F}u - f^k\|^2, \quad (2.2)$$

$$f^{k+1} = f^k + f - R\mathcal{F}u^{k+1}. \quad (2.3)$$

Because both total variation and l_1 -norm regularization terms are nonsmooth, subproblem (2.2) becomes the bottleneck that prevent from being used in practical MR image reconstruction. Using the algorithm ADPA (1) to solve the subproblem (2.2), we obtain Algorithm 2 to solve the original problem (1.1). Our algorithm can write as follows:

Algorithm 2 ADPA-BI (ADPA via Bregman Iteration for Compressed Sensing)

1. Initial: $f^1 = f, \sigma, T, \rho = 1/L, \alpha, \beta > 0, \gamma_1, \gamma_2 > 0, t^1 = 1$.
2. While $\|R\mathcal{F}z_1^k - f\|_2^2 > \sigma^2$;
3. $z_1^k = z_2^k = \mathcal{F}^{-1}f^k$;
4. using the algorithm ADPA for $\mathcal{F}^{-1}f^k$;
5. $f^{k+1} = f^k + f - R\mathcal{F}z_1^{k+1}$;
6. end While.

Because the Step 3 and 5 only involve adding vectors or scalars, the costs are only $\mathcal{O}(\log(N))$ or $\mathcal{O}(1)$. In the step 4 the cost of Algorithm 1 is $\mathcal{O}(N \log(N))$. Thus, the total cost of each iteration in Algorithm 2 is $\mathcal{O}(N \log(N))$. When parameter values are properly chosen, it has been found that the outer “while” loop of this algorithm only needs to be executed a small number of times. As a result, for imaging applications this algorithm is very fast when properly chosen parameter values are used.

We consider the optimal solution of problem (1.2) with $\alpha, \beta > 0$. Note that the objective function for problem (1.2) is proper, lower semicontinuous and level-bounded [6]. Therefore, by Theorem 1.9 in [6], the set of minimizers is nonempty and compact. In addition, the objective function for (1.2) with $\alpha, \beta > 0$ is strictly convex. Hence the optimal solution for problem (1.2) with $\alpha, \beta > 0$ is unique. So the following Lemma 1 holds.

Lemma 1 There exists an optimal solution for problem (1.2). In addition, the optimal solution of problem (1.2) with $\alpha, \beta > 0$ is unique.

Now, we consider the convergence property of our proposed Algorithm 2. and give a concise proof.

Theorem 2 Let the sequence $(z_1^k, z_2^k, f^k)_{k \in \mathbb{N}}$ be generated by Algorithm 2. If $0 < \rho < \frac{1}{\|\mathcal{A}^T \mathcal{A}\|}$, then both sequences $(z_1^k)_{k \in \mathbb{N}}$ and $(z_2^k)_{k \in \mathbb{N}}$ converge to the unique solution to problem (1.2).

Proof We consider a Lagrangian formulation of the original constrained problem (1.2)

$$L(u, p) = J(u) - \langle \mathcal{A}u - f, p - f \rangle \quad \text{and} \quad \mathcal{A}u = f, \quad (2.4)$$

where a change of variable $p - f$ instead of p as Lagrangian multiplier.

We let u^* be an optimal solution of problem (1.2) and p^* be a Lagrangian multiplier, respectively. We denote $s^* = -\mathcal{A}^T(f - p^*)$, then we can see that s^* is a subgradient of J at u^* by the Lagrangian function. Therefore, the overall optimality conditions are as follows:

$$s^* + \mathcal{A}^T(f - p^*) = 0, \quad (2.5)$$

$$\mathcal{A}u^* - f = 0. \quad (2.6)$$

We let (z_1^k, z_2^k, f^k) be the sequence generated by Algorithm , then $f^{k+1} = f^k + f - \mathcal{A}z_1^{k+1}$. Because the convergence of Algorithm 1 for subproblem (2.2), it holds that $\mathcal{A}z_1^{k+1} \rightarrow f$ and $\mathcal{A}z_2^{k+1} \rightarrow f$ as $k \rightarrow \infty$. Then

$$\lim_{k \rightarrow \infty} \|\mathcal{A}z_1^{k+1} - f\| = \lim_{k \rightarrow \infty} \|\mathcal{A}z_2^{k+1} - f\| = \lim_{k \rightarrow \infty} \|f^{k+1} - f^k\| = 0.$$

So $\mathcal{A}z_1^{k+1} \rightarrow f = \mathcal{A}u^*$ and $\mathcal{A}z_2^{k+1} \rightarrow f = \mathcal{A}u^*$ as $k \rightarrow \infty$. By the nonnegative property of the Bregman distance, it holds that

$$\begin{aligned} \lim_{k \rightarrow \infty} D_J^{s^*}(z_1^{k+1}, u^*) &= \lim_{k \rightarrow \infty} (J(z_1^{k+1}) - J(u^*) - \langle s^*, z_1^{k+1} - u^* \rangle) = 0, \\ \lim_{k \rightarrow \infty} D_J^{s^*}(z_2^{k+1}, u^*) &= \lim_{k \rightarrow \infty} (J(z_2^{k+1}) - J(u^*) - \langle s^*, z_2^{k+1} - u^* \rangle) = 0. \end{aligned}$$

Using $s^* = -\mathcal{A}^T(f - p^*)$, $\mathcal{A}z_1^{k+1} \rightarrow f = \mathcal{A}u^*$ and $\mathcal{A}z_2^{k+1} \rightarrow f = \mathcal{A}u^*$ as $k \rightarrow \infty$, We have

$$\begin{aligned} 0 &= \lim_{k \rightarrow \infty} (J(z_1^{k+1}) - J(u^*) + \langle f - p^*, \mathcal{A}(z_1^{k+1} - u^*) \rangle) = \lim_{k \rightarrow \infty} (J(z_1^{k+1}) - J(u^*)), \\ 0 &= \lim_{k \rightarrow \infty} (J(z_2^{k+1}) - J(u^*) + \langle f - p^*, \mathcal{A}(z_2^{k+1} - u^*) \rangle) = \lim_{k \rightarrow \infty} (J(z_2^{k+1}) - J(u^*)). \end{aligned}$$

It means that $J(z_1^{k+1}) \rightarrow J(u^*)$ and $J(z_2^{k+1}) \rightarrow J(u^*)$ as $k \rightarrow \infty$. So Algorithm 2 is convergent.

So we apply the ADPA method for subproblem (2.2) and propose the algorithm ADPA-BI that can efficiently solve the original problem (1.2). It is one of the most contributions in this paper.

3 Numerical Experiments

Suppose a Magnetic Resonance image x has n pixels, the partial Fourier transform in problem (1.2) consists of m rows of a $n \times n$ matrix corresponding to the full 2D discrete Fourier transform. The m selected rows correspond to the acquired b (the observation measurement). The sampling ratio is defined as m/n . The scanning duration is shorter if the sampling ratio is smaller. In MR imaging, we have certain freedom to select rows, which correspond to certain frequencies. In the following experiments, we select the corresponding frequencies according to the following manner. In the k -space, we randomly obtain more samples in lower frequencies and less samples in higher frequencies. This sampling scheme has been widely used for compressed MR image reconstruction [1]. Practically, the sampling scheme and speed in MR imaging also depend on the physical and physiological limitations [1].

We implement our ADPA-BI algorithm for problem (1.2) and apply them on 2D real MR images. All experiments are conducted on a 1.8 GHz PC in Matlab environment. We compare the ADPA-BI algorithm with the classic Splitting Bregman Iteration (SBI) algorithm [2] for the reconstruction results and computation complexity of MR images. For example, the observation measurement b is synthesized as $b = Rx + n$, where n is the Gaussian white noise with standard deviation $\sigma = 0.001$. The regularization parameters α and β are set as 0.001 and 0.035. The parameters γ_1 and γ_2 are both set as $e^{-\sqrt[4]{k}}$, where k is the inner iteration counter. R and b are given as inputs, and x is the unknown target.

Let x_{true} be the original true image and x a reconstruction image. For quantitative evaluation, the Signal-to-Noise Ratio (SNR) is computed for each reconstruction result. The signal-to-noise ratio (SNR) defined by

$$SNR = 10 \log_{10} \frac{\|x_{true}\|_2^2}{\|x - x_{true}\|_2^2}. \quad (3.1)$$

The relative dynamic error between two consecutive iterations:

$$\frac{\|z_1^k - z_1^{k-1}\|_2}{\|z_1^k\|_2} < \varepsilon, \quad (3.2)$$

and $\|Az_1^k - b\|_2^2 > \delta^2$ for the prescribed tolerances δ, ε as the stopping rules.

We apply our proposed ADPA-BI method and Split Bregman Iteration (SBI) method from 21% K-space data of six 2D MR images: phantom, brain1, brain2, brain3, brain4 and brain5 respectively. Fig. 1 shows these images for simulations.

Table 1 shows that the SNR of the results of reconstruction by different methods. From the results of reconstruction, the proposed algorithm ADPA-BI is superior to the algorithm SBI.

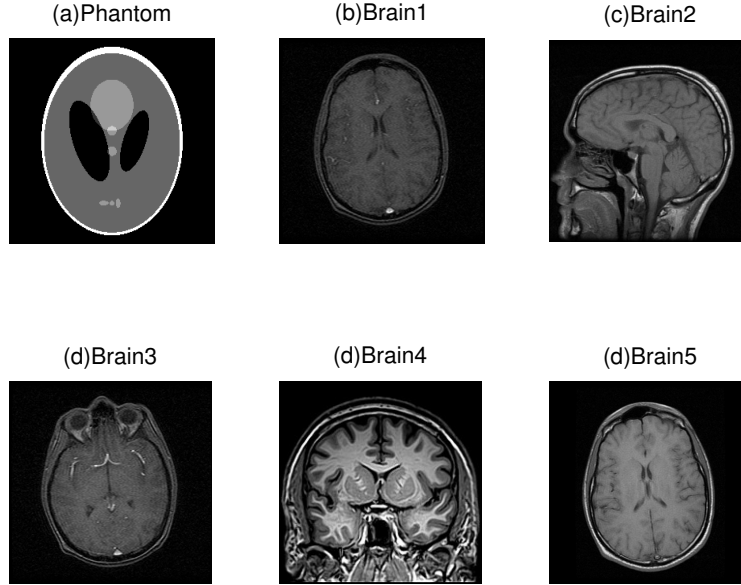


Figure 1: MR images: (a)Phantom; (b)Brain1; (c)Brain2; (d)Brain3; (e)Brain4 and (f)Brain5

Table 1: The SNR (db) of results of reconstruction from 21% K-space data by SBI and ADPA-BI.

	Phantom	Brain1	Brain2	Brain3	Brain4	Brain5
SBI	21.72	17.33	14.94	20.73	10.58	20.73
ADPA-BI	35.53	18.64	18.55	27.76	16.65	35.07

4 Conclusion

We have proposed an efficient and fast algorithm for the compressed MR image reconstruction from partial Fourier data. Our work has the following contributions. First, the proposed algorithm ADPA-BI can efficiently solve a composite regularization problem including both TV term and L1 norm term, which can be easily extended to other medical image applications. Second, the computation complexity of the ADPA-BI algorithm is analysed. Third, the convergence properties of the proposed algorithm ADPA-BI are analysed. These properties make the real compressed MR image reconstruction much more feasible than before. Finally, we conduct numerical simulations on recovering magnetic resonance images from partial Fourier data. Numerical simulations show that the proposed algorithm is very efficient.

References

- [1] Lustig M, Donoho D, Pauly J M. Sparse MRI: The application of compressed sensing for rapid mr imaging[J]. Magnetic Resonance in Medicine, 2007, 58(6), 1182–1195.
- [2] Goldstein T, Osher S. The split Bregman method for L_1 regularized problems[J]. SIAM J. Imag. Sci., 2009, 2(2): 323–343.
- [3] Combettes P L, Pesquet J C. A Douglas-Rachford splitting approach to nonsmooth convex variational signal recovery[J]. IEEE J. Selected Topics Sig. Proc., 2007, 1(4): 564–574.
- [4] Bauschke H H, Combettes P L. A Dykstra-like algorithm for two monotone operators[J]. Pacific J. Optim., 2008, 4(1): 383–391.
- [5] Beck A, Teboulle M. A fast iterative shrinkage-thresholding algorithm for linear inverse problems[J]. SIAM J. Imag. Sci., 2009, 2(1): 183–202.
- [6] Rockafellar R T, Wets R J. Variational analysis[M]. New York: Springer-Verlag, 2009.
- [7] Xu W Z, Zhang L Q. Fast and efficient signal reconstruction from structurally sampling partial Fourier data by chaotic dynamical system[J]. Fifth Intern. Confer. Machine Vision, Proc. SPIE 8784, 2013, doi: 10.1117/12.2014152.
- [8] Wang B B, Yin H. Varying quantile regression with online scheme and unbounded sampling[J]. J. Mathematics, 2014, 34(2): 281–286.

一种SENSE模型下信号重建的类-Dykstra近点有效算法

许伟志¹, 殷 弘², 蒋凌云¹

(1.湖北经济学院统计学院, 湖北 武汉 430205)

(2.中国人民大学信息学院, 北京 100872)

摘要: 本文研究了SENSE模型下从部分傅里叶数据中信号的重建问题. 利用类Dykstra近点方法和Bregman迭代方法, 我们获得了一种SENSE模型下信号重建的加速类-Dykstra近点有效算法, 并证明了该算法的收敛性. 实验仿真显示, 该方法比经典的分裂Bregman方法有效.

关键词: 核磁共振图像重建; 压缩感知; Bregman方法; 类Dykstra近点算法; SENSE模型

MR(2010)主题分类号: 90C30 中图分类号: O221.2