Applied Mathematics and Computation 270 (2015) 777–784

Contents lists available at ScienceDirect

Applied Mathematics and Computation

journal homepage: www.elsevier.com/locate/amc

A fast proximal point algorithm for ℓ_1-minimization problem in compressed sensing

Yun Zhu [a], Jian Wu [b], Gaohang Yu [c,*]

[a] School of Physics and Electronic Information, Gannan Normal University, China
[b] School of Computer Science and Technology, Harbin Institute of Technology Shenzhen Graduate School, China
[c] School of Mathematics and Computer Science, Gannan Normal University, Ganzhou 341000, China

ARTICLE INFO

Keywords:
Proximal point algorithm
ℓ_1-regularized least square
Compressed sensing

ABSTRACT

In this paper, a fast proximal point algorithm (PPA) is proposed for solving ℓ_1-minimization problem arising from compressed sensing. The proposed algorithm can be regarded as a new adaptive version of customized proximal point algorithm, which is based on a novel decomposition for the given nonsymmetric proximal matrix M. Since the proposed method is also a special case of the PPA-based contraction method, its global convergence can be established using the framework of a contraction method. Numerical results illustrate that the proposed algorithm outperforms some existing proximal point algorithms for sparse signal reconstruction.

1. Introduction

In this paper, we study proximal point algorithms for ℓ_1-minimization problem arising from the area of compressed sensing. In compressed sensing theory, one is interested in finding a solution $x^* \in \mathbb{R}^n$ to the following problem:

$$\min_{x\in\mathbb{R}^n} \|x\|_1 \quad \text{s.t. } Ax = b, \tag{1}$$

which is widely known as basis pursuit (BP) problem. Usually, the equality constraint is relaxed to an inequality constraint in the case of data being corrupted by outlying noises such as additive Gaussian noise. Hence the BP problem (1) could be reformulated in a form of ℓ_1-regularized least square problem as follows:

$$\min_{x\in\mathbb{R}^n} \lambda\|x\|_1 + \frac{1}{2}\|Ax-b\|^2, \lambda > 0. \tag{2}$$

The problem above is called as basis pursuit denoising (BPDN) problem, which is very close to the LASSO problem in statistics [14]. Actually the original LASSO problem can be transformed into an unconstrained form according to the Lagrangian multiplier theorem. In fact the parameter λ in (2) is just the Lagrangian multiplier corresponding to the sparsity constraint in the LASSO. On the other hand, the problems (1) and (2) can be unified in a general fashion though they are quiet different in structure. From a viewpoint of homotopy, we know that the solutions of BPDN problem converges to those of BP problem as $\lambda \to 0$ continuously. Hence a numerical solver for (2) also can be used to solve (1) if a procedure of continuation is casted on the parameter λ.

* Corresponding author. Tel.: +86 15270609677
E-mail addresses: zhuwwb@163.com (Y. Zhu), wujianhitsz@163.com (J. Wu), maghyu@163.com (G. Yu).

http://dx.doi.org/10.1016/j.amc.2015.08.082
0096-3003/

Fast numerical algorithms for problems with separable structure (for example, see (2)) were extensively studied in the last decade. Most of these algorithms take advantage of the separable structure based on operator splitting. Some augmented Lagrangian/Bregman distance based methods splits the involved problem via a block Gauss–Seidel approximation [5,11,15,16], while other algorithms decouple variables by formulating the original problem into a primal-dual formulation [1,3,8,19]. Moreover, some theoretical connections between these two types of methods have been analyzed in [3,15,16].

A class of so-called proximal point algorithms is of our main interests in this paper for its computational efficiency and simplicity in convergence analysis. Proximal point algorithm was originally proposed in [10] and then was naturally extended to metric form [4,7–9]. The proximal matrix is usually sophisticatedly chosen so that subproblems in the metric proximal point algorithm can be easily solved, for which reason these metric proximal point algorithms are called customized proximal point algorithms by authors in [7]. Since it fully takes advantage of special structure of problems, the customized proximal point algorithm usually shows remarkable efficiency in practice. Recently, some nonlinear proximal point algorithms for monotonic variational inequalities were given and analyzed in [13].

Proximal point algorithm is quite simple in theoretic analysis. Its convergence can be obtained by directly extending the result developed in [10] to metric form when the proximal matrix is symmetric and positive definite [7]. It has been shown in [8] that the first-order primal-dual algorithm is equivalent to a proximal point algorithm, hence the convergence analysis is greatly simplified. In the case of asymmetric proximal matrix, an extra correction step is easily performed so that the convergence is guaranteed [6,8]. The correction ensures the algorithm having a contraction property, hence these methods are called contraction methods. A proximal point algorithm based contraction method using an adaptive proximal matrix was proposed for total variation image restoration in [2].

Steplength in customized proximal point algorithm is usually constant. Naturally an adaptive steplength could achieve better performance. The main contribution of this paper is to propose an adaptive customized proximal point algorithm and to show its improvement to the existing algorithms. This work is mainly motivated by [6,8], and the basic analytic tool is the contraction method extensively developed in [6,8].

The paper is organized as follows. First, we introduce a primal-dual formulation of the BPDN problem and briefly review some basic notations of proximal point algorithm and contraction method in Section 2. Then we present a new contraction method and show its connection to a customized proximal point algorithm in Section 3, and then discuss the global convergence of the proposed algorithm in Section 4. Finally, we did a sparse signal reconstruction experiment to show the efficiency of the proposed algorithm numerically in Section 5.

2. Preliminaries

In this section, we provide some basic notations which will be useful for subsequent analysis. In particular, we review some basic ideas of variational inequalities (VIs), the proximal point algorithms and contraction methods.

2.1. VI formulation of (2)

Let $y \in \mathbb{R}^m$ be Lagrangian multiplier. Then the optimization problem (2) is equivalent to finding a saddle point of the min–max problem

$$\min_x \max_y \Phi(x, y) := \|x\|_1 - y^T(Ax - b) - \frac{\lambda}{2}\|y\|^2, \tag{3}$$

and in the following analysis we mainly aim to solve the problem (3).

From the first-order optimality conditions of (3) and the convexity of the involved functions, the min–max problem (3) can be characterized by the following mixed variational inequality problem: find a solution pair $x^* \in \mathbb{R}^n$ and $y^* \in \mathbb{R}^m$ such that

$$\|x\|_1 - \|x^*\|_1 + \begin{pmatrix} x - x^* \\ y - y^* \end{pmatrix}^T \begin{pmatrix} -A^T y^* \\ Ax^* - b + \lambda y^* \end{pmatrix} \geq 0, \forall x \in \mathbb{R}^n, y \in \mathbb{R}^m.$$

Denote

$$u = \begin{pmatrix} x \\ y \end{pmatrix}, u^* = \begin{pmatrix} x^* \\ y^* \end{pmatrix}, F(u^*) = \begin{pmatrix} -A^T y^* \\ Ax^* - b + \lambda y^* \end{pmatrix}, \Omega = \mathbb{R}^n \times \mathbb{R}^m,$$

then the VI problem above can be rewritten in the following compact form: find $u^* \in \Omega$ such that

$$\|x\|_1 - \|x^*\|_1 + (u - u^*)^T F(u^*) \geq 0, \forall u \in \Omega. \tag{4}$$

It is easy to verify that the mapping F in (4) is monotone with respect to Ω in the sense of

$$(u - v)^T \{F(u) - F(v)\} \geq 0, \ \forall\, u, v \in \Omega.$$

Therefore, the VI (4) is monotonic. Denote the solution set as Ω^*. Throughtout this paper, we always assume that Ω^* is not empty.

It is worth noting that $\Phi(x, y)$ is the Lagrangian function of (1) in the case of $\lambda = 0$, which means that the min-max formulation (3) is an extension of Lagrangian duality, and that our algorithm, as we will see, is an extension of classical augmented Lagrangian method.

Finally before reviewing existing algorithms, two notions of norms about matrices are worthy of being showcased in advance.