Journal of Computational and Applied Mathematics 256 (2014) 36–51

Contents lists available at ScienceDirect

Journal of Computational and Applied Mathematics

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

A proximal parallel splitting method for minimizing sum of convex functions with linear constraints

Deren Han [a,*], Hongjin He [b], Lingling Xu [c]

[a] School of Mathematical Sciences, Key Laboratory for NSLSCS of Jiangsu Province, Nanjing Normal University, Nanjing 210023, PR China
[b] School of Science, Hangzhou Dianzi University, Hangzhou 310018, PR China
[c] School of Mathematical Sciences, Key Laboratory for NSLSCS of Jiangsu Province, Nanjing Normal University, Nanjing 210023, PR China

ARTICLE INFO

Article history:
Received 8 February 2011
Received in revised form 17 March 2013

Keywords:
Parallel splitting method
Proximal point algorithm
Separable structure
Linearized algorithm
Global convergence

ABSTRACT

In this paper, we propose a proximal parallel decomposition algorithm for solving the optimization problems where the objective function is the sum of m separable functions (i.e., they have no crossed variables), and the constraint set is the intersection of Cartesian products of some simple sets and a linear manifold. The m subproblems are solved simultaneously per iterations, which are sum of the decomposed subproblems of the augmented Lagrange function and a quadratic term. Hence our algorithm is named as the 'proximal parallel splitting method'. We prove the global convergence of the proposed algorithm under some mild conditions that the underlying functions are convex and the solution set is nonempty. To make the subproblems easier, some linearized versions of the proposed algorithm are also presented, together with their global convergence analysis. Finally, some preliminary numerical results are reported to support the efficiency of the new algorithms.

1. Introduction

For the index $i = 1, 2, \ldots, m$, let $\mathcal{X}_i \subseteq \mathcal{R}^{n_i}$ be closed convex sets, $\theta_i : \mathcal{R}^{n_i} \to \mathcal{R}$ be closed proper convex functions (not necessarily smooth), and $A_i \in \mathcal{R}^{l\times n_i}$. In this paper, we consider the following optimization problem

$$\min \left\{ \sum_{i=1}^{m} \theta_i(x_i) \,\middle|\, \sum_{i=1}^{m} A_i x_i = b, x_i \in \mathcal{X}_i, i = 1, 2, \ldots, m \right\}, \tag{1.1}$$

where $b \in \mathcal{R}^l$ is a given vector. This type of problems are often encountered in realities, e.g., convex programming, variational analysis and PDE, etc. From the numerical point of view, this problem has very special structure: its objective function has a separable structure, i.e., the objective function is the sum of m functions and each function θ_i only depends on its own variable x_i. This special structure provides us the opportunity of splitting it into m smaller subproblems and solve it via solving these smaller subproblems.

Traditional studies of splitting algorithms focus on the case with $m = 2$. For this special case, Gabay and Mercier [1,2] proposed the following alternating direction method (ADM), which generates the iterative sequence via the following recursion

$$\begin{cases} x_1^{k+1} = \arg\min \left\{ \mathcal{L}_H(x_1, x_2^k, \lambda^k) \mid x_1 \in \mathcal{X}_1 \right\}, \\ x_2^{k+1} = \arg\min \left\{ \mathcal{L}_H(x_1^{k+1}, x_2, \lambda^k) \mid x_2 \in \mathcal{X}_2 \right\}, \\ \lambda^{k+1} = \lambda^k - H\left(A_1 x_1^{k+1} + A_2 x_2^{k+1} - b\right), \end{cases} \tag{1.2}$$

* Corresponding author. Tel.: +86 25 85891630; fax: +86 25 85898785.
E-mail addresses: handr00@hotmail.com, handeren@njnu.edu.cn (D. Han), hehj2003@163.com (H. He), xulingling@njnu.edu.cn (L. Xu).

0377-0427/$ – see front matter © 2013 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.cam.2013.07.010

where $\lambda \in \mathcal{R}^l$ is the Lagrangian multiplier associated to the linear constraint, $H \in \mathcal{R}^{l\times l}$ is a positive definite matrix and

$$\mathcal{L}_H(x_1,\ldots,x_m,\lambda) := \sum_{i=1}^m \theta_i(x_i) - \lambda^\top\left(\sum_{i=1}^m A_i x_i - b\right) + \frac{1}{2}\left\|\sum_{i=1}^m A_i x_i - b\right\|_H^2 \tag{1.3}$$

is the augmented Lagrange function of the optimization problem (1.1). In the classical ADM, the matrix H is specified as $H = \beta I$, where β is a positive constant and I is the identity matrix. Due to its significant efficiency in tracking problem (1.1), ADM has received much attention from various areas, and a lot of variants are developed in the past decades, see, e.g., [3–13]. Recently, ADM finds a great number of applications in matrix optimization [14–16], image restoration [17,18], compressive sensing [19] and statistical learning [20–22].

However, in many cases, one often encounters the problem (1.1) with $m \geq 3$. For example, the robust principal component analysis model [23], the total-variation based image restoration problem [18], the superresolution image reconstruction problem [24,25], the multistage stochastic programming problem [26], and the deblurring Poissonian images problem [27]. For solving this general case, one naturally considers to extend the ADM directly to the following recursion:

$$\begin{cases} x_i^{k+1} = \arg\min\left\{\mathcal{L}_H(x_1^{k+1},\ldots,x_{i-1}^{k+1},x_i,x_{i+1}^k,\ldots,x_m^k,\lambda^k) \mid x_i \in \mathcal{X}_i\right\}, \quad i=1,\ldots,m,\\ \lambda^{k+1} = \lambda^k - H\left(\displaystyle\sum_{i=1}^m A_i x_i^{k+1} - b\right). \end{cases} \tag{1.4}$$

Unfortunately, till now there is no convergence proof for the iterative sequence generated by (1.4). Compared to the special case with $m = 2$, the study on the general case with $m \geq 3$ is on its infancy. He [28] proposed a parallel splitting algorithm for the case $m = 3$, where the main task per iteration is to solve the following subproblems (note that the method in [28] is presented under variational inequality framework; here we rewrite it for model (1.1) with $m = 3$):

$$\begin{cases} \tilde{x}_1^k = \arg\min\left\{\mathcal{L}_H(x_1,x_2^k,x_3^k,\lambda^k) \mid x_1 \in \mathcal{X}_1\right\},\\ \tilde{x}_2^k = \arg\min\left\{\mathcal{L}_H(x_1^k,x_2,x_3^k,\lambda^k) \mid x_2 \in \mathcal{X}_2\right\},\\ \tilde{x}_3^k = \arg\min\left\{\mathcal{L}_H(x_1^k,x_2^k,x_3,\lambda^k) \mid x_3 \in \mathcal{X}_3\right\}. \end{cases} \tag{1.5}$$

With the help of a simple correction step, it was proved that the generated sequence globally converges to a solution of (1.1). Most recently, Han et al. [29] proposed a new splitting method where the main subproblems are similar to those in (1.5), but with a different correction step in generating the new iterates.

The success of this type of splitting method relies on the efficient solvability of the subproblems. While recent applications of classical ADM type methods achieved great success due to the simplicity or closed-form solutions of the subproblems, in most cases the subproblems (1.5) may be time-consuming. The purpose of this paper is to design new parallel splitting algorithms with easier subproblems for solving convex programs with separable structure. We consider two strategies: first, we add a quadratic proximal term $\frac{r_i}{2}\|x_i - x_i^k\|^2$ with a constant $r_i \geq 0$ to the i-th subproblem in the splitting algorithms, then a strongly convex subproblem, which is easier to solve than the original subproblems, is obtained. Second, three linearized versions on the nonlinear terms in the subproblems are presented, i.e., we linearize the following two types of nonlinear terms:

$$\theta_i(x_i) \quad \text{or/and} \quad \left\|\sum_{j\neq i} A_j x_j^k + A_i x_i - b\right\|_H^2.$$

In many cases, the linearized subproblems possess closed-form solutions, which makes the iteration very simple, and the resulting algorithms may be very suitable to solve large-scale problems arising from real applications.

This paper is organized as follows. Some necessary preliminaries are provided in Section 2. By exploiting the separability of the models, in Section 3, we present a new proximal parallel splitting method. The global convergence of the present method is proved in Section 4. Based on the linearization techniques, Section 5 presents three different linearized schemes of the proposed proximal parallel splitting method. To investigate the numerical performance of our new algorithms, we implement the proposed algorithms for solving separable quadratic programs, and report some preliminary numerical results in Section 6. Finally, we complete this paper with some concluding remarks in Section 7.

2. Preliminaries

In this section, we summarize some basic concepts and their properties which will be useful for the subsequent sections.

Definition 2.1. An operator $f : \Omega \to \mathcal{R}^n$ is said to be

(a) monotone if

$$(u-v)^\top(f(u)-f(v)) \geq 0, \quad \forall u, v \in \Omega;$$