Applied Mathematics and Computation 336 (2018) 315–325

Contents lists available at ScienceDirect

Applied Mathematics and Computation

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

On the pointwise iteration-complexity of a dynamic regularized ADMM with over-relaxation stepsize

M.L.N. Gonçalves[1]

IME, Universidade Federal de Goiás, Goiânia, GO 74001-970, Brazil

ARTICLE INFO

MSC:
47H05
49M27
90C25
90C30
90C60
65K10

Keywords:
Alternating direction method of multipliers
Hybrid proximal extragradient framework
Pointwise iteration-complexity
Convex programming

ABSTRACT

In this paper, we extend the improved pointwise iteration-complexity estimation of a dynamic regularized alternating direction method of multipliers (ADMM) for a new stepsize domain. In this complexity analysis, the stepsize parameter can be chosen in the interval (0,2) instead of interval $(0, (1+\sqrt{5})/2)$. We illustrate, by means of a numerical experiment, that the enlargement of this stepsize domain can lead to better performance of the method in some applications. Our complexity study is established by interpreting this ADMM variant as an instance of a hybrid proximal extragradient framework applied to a specific monotone inclusion problem.

1. Introduction

We are interested in the following linearly constrained convex problem

$$\min\{f(x)+g(y) : Ax+By=b,\ x\in\mathbb{R}^n, y\in\mathbb{R}^p\}, \tag{1}$$

where $f:\mathbb{R}^n\to\mathbb{R}$ and $g:\mathbb{R}^p\to\mathbb{R}$ are convex functions, $A\in\mathbb{R}^{m\times n}$, $B\in\mathbb{R}^{m\times p}$ and $b\in\mathbb{R}^m$. We assume that the solution set of (1) is nonempty. Convex optimization problems with a separable structure such as (1) appear in many applications areas such as machine learning, compressive sensing and image processing. The augmented Lagrangian method (see, e.g., [3]) attempts to solve (1) directly without taking into account its particular structure. To overcome this drawback, a variant of the augmented Lagrangian method, namely, the alternating direction method of multipliers (ADMM), was proposed and studied in [10,12]. The ADMM takes full advantage of the special structure of the problem by considering each variable separably in an alternating form and coupling them into the Lagrange multiplier updating; for detailed reviews, see [5,11].

Lately, several versions of the ADMM for solving (1) have been proposed in the literature; see, for example, [1,4,6–8,14–20,25,26]. A dynamic regularized ADMM (DR-ADMM) with stepsize parameter $\theta\in(0,(1+\sqrt{5})/2)$ was proposed by Gonçalves et al. [14], whose pointwise iteration-complexity bound considerably improved the one of the ADMM. More specifically, for a given tolerance $\rho>0$, it follows from [14, Theorem 4.1] that the DR-ADMM terminates in at most

E-mail address: maxlng@ufg.br

[1] The work of this author was supported in part by CNPq Grants 406975/2016-7 and 302666/2017-6.

https://doi.org/10.1016/j.amc.2018.04.066
0096-3003/

$\mathcal{O}\big(\rho^{-1}\log(\rho^{-1})\big)$ iterations with triples (x, y, γ) and (v^x, v^y, v^γ) satisfying

$$v^x \in \partial f(x) - A^*\gamma, \quad v^y \in \partial g(y) - B^*\gamma, \quad v^\gamma = Ax + By - b, \qquad \|(v^x, v^y, v^\gamma)\| \le \rho, \tag{2}$$

where the asterisk superscript denotes transpose of a matrix, and ∂ denotes subdifferential operator. Although different criteria are used, in general the ADMM and its variants need $\mathcal{O}\big(\rho^{-2}\big)$ iterations to find triples satisfying (2) (see, e.g., [1,6–8,15–20,22,25]). It is worth pointing out that the best ergodic iteration-complexity bounds known of the aforementioned ADMMs are $\mathcal{O}\big(\rho^{-1}\big)$, which are better than the pointwise one of the DR-ADMM. However, the triples constructed in the ergodic case satisfy (2) with ρ-subdifferential operators instead of subdifferential operators.

It is well-known that some ADMMs with stepsize parameter close to $(1+\sqrt{5})/2$ (i.e. close to upper limit of its domain) have better performance in many applications (see, e.g., [9,11] for more details). Hence, it would be reasonable to analyze the numerical performance and/or complexity results for these methods with a larger stepsize domain; for example, stepsize belonging to the interval (0, 2) as in the augmented Lagrangian method. Recently, paper [13] established complexity results for the ADMM with stepsize $\theta \in (0, 2)$ for solving non-convex linearly constrained problems and, subsequently, paper [17] studied convergence and complexity results for the ADMM with the same stepsize domain of the present paper for the convex case.

Therefore, from the theoretical viewpoint, the main goal of this work is to present complexity results for a variant of the DR-ADMM [14]. More specifically, we extend the improved pointwise iteration-complexity results for this ADMM for any stepsize $\theta \in (0, (1-\alpha+\sqrt{\alpha^2+6\alpha+5})/2)$, where α is a nonnegative proximal factor associated to the proximal term added to the second subproblem of the method (see the DR-ADMM in Section 3). Since the limit of $(\sqrt{\alpha^2+6\alpha+5}-\alpha)$ as α goes to infinity is 3, the latter stepsize domain becomes (0, 2) (resp. $(0, (1+\sqrt{5})/2)$) when α is sufficiently large (resp. $\alpha = 0$). It is worth pointing out, however, that stepsize θ close to 2 implies a large proximal factor α, which may turn the subproblem associated to α much slower to solve due the weight of the proximal term in the objective function of the subproblem. From the applicability viewpoint, we report a numerical experiment, which shows that this new variant of the DR-ADMM with suitable proximal term α and $\theta \ge (1+\sqrt{5})/2)$ can perform better in some applications. We mention that, as in [14], our complexity study is done by rewriting problem (1) as a monotone inclusion problem and by analyzing the DR-ADMM in the setting of a generalized hybrid proximal extragradient (HPE) framework.

This paper is organized as follows. In Section 2, we present a dynamic semi-regularized HPE framework as well as its pointwise iteration-complexity estimation. In Section 3, we describe the DR-ADMM and show that it can be viewed as an instance of the framework of Section 2 applied to a particular monotone inclusion problem. The pointwise iteration-complexity result for this ADMM variant is also obtained in this section. A numerical experiment is reported in Section 4.

Notation: The set of real numbers is denoted by $\mathbb{R}$. The set of non-negative real numbers and the set of positive real numbers are denoted by $\mathbb{R}_+$ and $\mathbb{R}_{++}$, respectively. For $t > 0$, we let $\log^+(t) := \max\{\log t, 0\}$. For a finite-dimensional real vector space $\mathcal{X}$ with inner product $\langle\cdot,\ \cdot\rangle$, its induced norm is denoted by $\|\cdot\|$. Denote by $\mathcal{M}_+^{\mathcal{X}}$ the space of selfadjoint positive semidefinite linear operators on $\mathcal{X}$. For each $H \in \mathcal{M}_+^{\mathcal{X}}$, the seminorm and (extended) dual seminorm induced by H on $\mathcal{X}$ are defined by $\|\cdot\|_H := \sqrt{\langle H(\cdot), \cdot\rangle}$ and $\|\cdot\|_H^* := \sup\{\langle\cdot, x'\rangle : \|x'\|_H \le 1\}$, respectively.

2. Preliminaries results

In this section, we present a dynamic regularized HPE framework and its pointwise iteration-complexity estimation. This framework is an instance of one studied in [14].

Consider the monotone inclusion problem (MIP)

$$0 \in T(z), \tag{3}$$

where $\mathcal{Z}$ is a finite-dimensional real vector space and $T : \mathcal{Z} \rightrightarrows \mathcal{Z}$ is a maximal monotone operator.[2] We assume that the solution set of (3), denoted by $T^{-1}(0)$, is nonempty.

The dynamic regularized HPE framework attempts to solve the inclusion (3) by solving approximately a sequence of regularized MIP of the following form

$$0 \in T(z) + \mu M(z - z_0), \tag{4}$$

where $z_0 \in \mathcal{Z}$, $\mu > 0$ and $M \in \mathcal{M}_+^{\mathcal{Z}}$ are fixed. We also assume that the solution set of (4)

$$\bar{Z}_\mu(M) := \{z \in \mathcal{Z} :\ 0 \in T(z) + \mu M(z - z_0)\} \tag{5}$$

is nonempty for every $\mu > 0$. It can be shown that if M is positive definite, then the operator $T(\cdot) + \mu M(\cdot - z_0)$ is maximal $(\mu\rho_0)$-strongly monotone for every $\mu > 0$, where ρ_0 is the smallest eigenvalue of M, and hence the latter assumption holds

[2] An operator $T : \mathcal{Z} \rightrightarrows \mathcal{Z}$ is said to be monotone if $\langle z - z', s - s'\rangle \ge 0$, for every $z, z' \in \mathcal{Z}, s \in T(z)$ and $s' \in T(z')$. Moreover, T is maximal monotone if it is monotone and, additionally, if S is a monotone operator such that $T(z) \subset S(z)$ for every $z \in \mathcal{Z}$ implies $T = S$.