Journal of Computational and Applied Mathematics 330 (2018) 475–487

Contents lists available at ScienceDirect

Journal of Computational and Applied Mathematics

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

A novel method for a class of structured low-rank minimizations with equality constraint[☆]

Jianchao Bai [a], Jicheng Li [a,*], Fengmin Xu [b], Pingfan Dai [a,c]

[a] School of Mathematics and Statistics, Xi'an Jiaotong University, Xi'an 710049, PR China
[b] School of Economics and Finance, Xi'an Jiaotong University, Xi'an 710049, PR China
[c] Department of Information Engineering, Sanming University, Sanming 365004, PR China

ARTICLE INFO

Article history:
Received 14 April 2016
Received in revised form 26 June 2017

MSC:
15B48
41A29
65F30
68W25

Keywords:
Low-rank
Equality constraint
Positive semidefinite constraint
Accelerated augmented Lagrangian method

ABSTRACT

The positive semidefinite constraint and equality constraint arise widely in matrix optimization problems of different areas including signal/image processing, finance and risk management. In this paper, an inexact accelerated Augmented Lagrangian Method (ALM) relying on a parameter m is designed to solve the structured low-rank minimization with equality constraint, which is more general and flexible than the existing ALM and its variants. We prove a worst-case $\mathcal{O}(1/k^2)$ convergence rate of the new method in terms of the residual of the Lagrangian function, and we analyze that when $m \in [0, 1)$ the residual of our method is smaller than that of the traditional accelerated ALM. Compared with several state-of-the-art methods, preliminary numerical experiments on solving the Q-weighted low-rank correlation matrix problem from finance validate the efficiency of the proposed method.

1. Introduction

Consider the following structured low-rank matrix minimization with equality constraint

$$\begin{array}{ll} \min & f(Y) \\ \text{s.t.} & g(Y) = 0, \\ & Y \in \mathcal{S}_+^n, \operatorname{rank}(Y) \le r, \end{array} \tag{1.1}$$

where $f(Y) : \mathbb{R}^{n\times n} \to \mathbb{R}$, $g(Y) : \mathbb{R}^{n\times n} \to \mathbb{R}$ are respectively convex and concave functions (but not necessarily smooth); $\mathbb{R}^{n\times n}$, $\mathbb{R}^n$, $\mathbb{R}$ denote the set of $n \times n$ dimensional real matrices, the set of n dimensional real column vectors and the set of real numbers, respectively; $Y \in \mathcal{S}_+^n$ means a symmetric positive semidefinite matrix in $\mathbb{R}^{n\times n}$, and r $(1 \le r \le n)$ is a given positive integer. Throughout the discussions, we assume that the solution set of the problem (1.1) is nonempty.

In 2010, He et al. [1] introduced an accelerated Augmented Lagrangian Method (ALM) to tackle the following linearly constrained optimization problem

$$\min\{f(x) | \, Ax = b, x \in \Omega\}, \tag{1.2}$$

☆ The work was supported by the National Natural Science Foundation of China (11671318; 11571271; 11101325) and the Natural Science Foundation of Fujian province (2016J01028).

* Corresponding author.
E-mail addresses: baijianchaok@126.com (J.-C. Bai), jcli@mail.xjtu.edu.cn (J.-C. Li).

http://dx.doi.org/10.1016/j.cam.2017.09.021
0377-0427/

where $f(x)$ is a differentiable convex function and Ω is a closed convex set. Later, this method was applied to deal with the complex matrix case of (1.2) with $tr(Y) = 1$ and Ω being the set of positive semidefinite Hermitian matrices, whose effectiveness was verified by testing a density matrix optimization problem [2]. Clearly, the problem (1.1) is more complex and is an extension of (1.2) from vector to matrix and linear to nonlinear constraints. Moreover, the functions $f(Y)$ and $g(Y)$ can be sub-differentiable in the subsequent analysis of our proposed General Inexact Accelerated ALM (GIALM), while both the objective function and equality constraint of (1.2) are differentiable with respect to the variable x. This paper is strongly motivated by the work [1] and recent investigations on the popular Q-weighted low-rank correlation matrix problem [3,4] arising in finance:

$$\begin{array}{ll} \min & \|C - Y\|_Q^2 \\ \text{s.t.} & Y_{ii} = 1, i = 1, 2, \ldots, n, \\ & Y \in \mathcal{S}_+^n, \operatorname{rank}(Y) \le r, \end{array} \tag{1.3}$$

where $C \in \mathbb{R}^{n\times n}$ is a given symmetric matrix; $Q \in \mathcal{S}_+^{n^2}$ is a weighted matrix defined by $Q = Q_1 \otimes Q_2$ with $Q_1, Q_2 \in \mathcal{S}_+^n$; the symbol $\otimes$ denotes the Kronecker product defined by $A \otimes B = (a_{ij}B)$ for any matrices $A, B \in \mathbb{R}^{m\times n}$ and a_{ij} is the ijth entry of matrix A; the notation $\|\cdot\|_Q$ stands for the Q-weighted norm which is defined by

$$\|C\|_Q = \sqrt{\operatorname{vec}(C)^{\mathsf T} Q \operatorname{vec}(C)}. \tag{1.4}$$

Besides, several other popular problems also inspire us to develop a unified algorithm framework for solving (1.1), for instance, the low-rank semidefinite programming problem [5], the low-rank approximation of the positive semidefinite Hankel matrix [6] and the low-rank solution of structured matrix optimization problems [7–9]. In such cases, these problems can be naturally extended to the general structured low-rank minimization model (1.1).

There are two major contributions of our paper. One contribution is that the usual vector minimization problem with linear constraints is extended to a class of structured matrix minimization problems with nonlinear constraints. Since the involved functions are possibly non-differentiable, by using the properties of sub-differential operators we develop a novel GIALM depending on a parameter m to tackle the problem (1.1), where a worst-case $\mathcal{O}(1/k^2)$ convergence rate of the algorithm is established and here k denotes the iteration number. Moreover, we analyze that the residual of GIALM is smaller than that of the methods in [1,10] when m is restricted into [0, 1). The other contribution is that the proposed method is applied to handle the Q-weighted correlation matrix problem (1.3). By making use of the Gramian representation, the corresponding subproblem is equivalently transformed into an unconstrained optimization problem. Then, we use a quasi-Newton algorithm with Armijo line search to solve it. Numerical examples from finance are tested to illustrate that our method could outperform several others.

The remaining parts of this article are organized as follows. In Section 2, we introduce the concept of the sub-differential with its properties and review the classical ALM. Then, the new GIALM is constructed to deal with the original problem. Section 3 applies the proposed algorithm to the Q-weighted correlation matrix problem, whose subproblem is solved by the quasi-Newton algorithm. Section 4 investigates the performance of our proposed method and comparative experiments are also carried out. Finally, in Section 5 some conclusions are introduced.

2. General inexact accelerated augmented Lagrangian method

In this section we first prepare some preliminaries that will be of use in the subsequent sections. Then, the GIALM is developed to solve the problem (1.1), whose convergence is analyzed afterwards.

Any matrix $M \in \mathbb{R}^{n\times n}$ is called the sub-gradient (see [11], p. 214) of a convex function $f(X) : \mathbb{R}^{n\times n} \to \mathbb{R}$ at $Y \in \mathbb{R}^{n\times n}$, if it satisfies

$$f(X) - f(Y) \ge \langle M, X - Y\rangle, \tag{S}$$

where $\langle\cdot,\cdot\rangle$ denotes the standard inner product defined by $\langle A, B\rangle = tr(A^{\mathsf T}B)$ for any $A, B \in \mathbb{R}^{m\times n}$. The set of all sub-gradients of $f(X)$ is called its sub-differential and is usually denoted by $\partial f(Y)$.

Clearly, the set $\partial f(Y)$ is closed convex, since by definition $M \in \partial f(Y)$ if and only if M satisfies a certain infinite system of weak linear inequalities. If f is convex, then it is continuous in the domain and its sub-differential is nonempty and compact. Especially, if it is differentiable, then the element of its sub-differential is unique and amounts to its gradient. One significant function of the sub-differential is to verify whether the optimality condition of an unconstrained problem $\min_{Y\in\mathbb{R}^{n\times n}} f(Y)$ is satisfied, i.e., $\mathbf{0} \in \partial f(Y)$. In other words, it follows that

$$\widetilde{\nabla} f(Y^*) = \mathbf{0},$$

where $\widetilde{\nabla} f(Y^*)$ denotes the sub-gradient of $f(Y)$ at the minimum Y^* and $\mathbf{0}$ is the $n \times n$ zero matrix.

Next, we review the classical ALM for the problem (1.1). For convenience, let

$$\Omega = \left\{Y \in \mathbb{R}^{n\times n} |\; Y \in \mathcal{S}_+^n, \operatorname{rank}(Y) \le r\right\}.$$

For any $\mu > 0$, the augmented Lagrangian function of (1.1) is given by

$$\mathcal{L}_\mu(Y, \lambda) = \mathcal{L}(Y, \lambda) + \frac{\mu}{2}[g(Y)]^2,$$