Ann. of Appl. Math.
34:2(2018), 138-152

A BLOCK-COORDINATE DESCENT METHOD FOR LINEARLY CONSTRAINED MINIMIZATION PROBLEM*†

Xuefang Liu, Zheng Peng‡

(*College of Math. and Computer Science, Fuzhou University, Fuzhou 350116, Fujian, PR China*)

Abstract

In this paper, a block coordinate descent method is developed to solve a linearly constrained separable convex optimization problem. The proposed method divides the decision variable into a few blocks based on certain rules. Then the candidate solution is iteratively obtained by updating one block at each iteration. The problem, whether or not there are overlapping regions between two immediately adjacent blocks, is investigated. The global convergence of the proposed method is established under some suitable assumptions. Numerical results show that the proposed method is effective compared with some "full-type" methods.

Keywords linearly constrained optimization; block coordinate descent; Gauss-Seidel fashion

2000 Mathematics Subject Classification 90C25; 90C30; 65K10

1 Introduction

We consider in this paper a linearly constrained convex optimization problem of the form:

$$\begin{cases} \min\limits_{x\in R^n} h(x), \\ \text{s.t. } Ax \geq b, \end{cases} \tag{1.1}$$

where $h(x)$ is a proper, convex and lower semi-continuous (lsc) function, $A \in R^{m\times n}$ and $b \in R^m$. In this paper, we assume that the objective function $h(x)$ has a (block)-separable form, that is, $h(x) = \sum\limits_{s=1}^{p} h_s(x_s)$, where x_s is the s-th block of the decision variable x and h_s is proper, convex and lower semi-continuous.

*This work was supported by the Natural Science Foundation of China (Grant No.11571074) and the Natural Science Foundation of Fujian Province (Grant No.2015J01010).

†Manuscript received September 30, 2017; Revised November 30, 2017

‡Corresponding author. E-mail: pzheng@fzu.edu.cn

Problem (1.1) is very general which generalizes the constrained minimum distance sum problem, where the Euclidean distance, l_1 norm distance or l_∞ norm distance can be selected. It has widely applications in business, industry areas, facility location, integrated circuits placement and so on.

The coordinate gradient descent method has been used in many considerable discussions, see for example Fukushima [4]. Tseng and Yun [13] gave a coordinate gradient descent method with an Armijo-type line search under an assumption that h is a separable and smooth function. Bai, Ng and Qi [8] showed that the gradient-based methods cannot be employed directly to solve problem (1.1) when $h(x)$ is nonsmooth. In addition, problem (1.1) has also raised in many applications, including signal denosing, image processing and data classification. In this case, those methods updating one coordinate variable at each iteration have been well suited due to their low computational cost and easy implementation [6]. For example, the (block-) successive over-relaxation (SOR) method is used to find sparse representation of signals, and some decomposition methods are used to support vector machine (SVM) training, etc.

The block coordinate descent (BCD) method divides the variable x into a few small blocks, then minimizes at each iteration the objective function with respect to one-block of the variable by fixing the other blocks. The BCD method can be found under numerous names, including linear or nonlinear Gauss-Seidel method, subspace correction method and alternating minimization approaches. Since the dimension of each block is often small, the computational cost of each iteration is relatively cheap which is suitable for large scale problem.

The blocks of the variable are often updated by using several types of strategies in the BCD method. The cyclic (Gauss-seidel) strategy updates blocks one by one in turn and the values of the newly updated blocks are used in the subsequent operations on the other blocks. The essentially cyclic rule selects each block at least once every successive T iterations, where T is an integer not less than the number of blocks. The Gauss-Southwell rule computes a positive value q_s for the s-th block according to some criteria, and then chooses the block with the largest value of q_s to work on next, or chooses the k-th block to work on, where $q_k \geq \beta \times \max_s q_s$ for $\beta \in (0,1]$. Recently, some randomized rules have been proposed by Zhao [16], Nesterov [7] and Richtrik [10], etc. The greedy coordinate block descent method [14] selected a few blocks of the variable and a few variables in each selected block, by greedy means, and updated the latter in parallel fashion at each iteration. The Jacobian-type iteration updated all the blocks simultaneously, hence it is suitable for parallel computation and distributed optimization [2,3].

In this paper, the BCD method for the unconstrained composite optimization pro-