Available online at www.sciencedirect.com
ScienceDirect
Nonlinear Analysis 66 (2007) 2016–2027
Nonlinear Analysis
www.elsevier.com/locate/na

A proximal bundle method with inexact data for convex nondifferentiable minimization[☆]

Jie Shen[a,b,*], Zun-Quan Xia[a], Li-Ping Pang[a,c]

[a] *Department of Applied Mathematics, Dalian University of Technology, Dalian 116024, China*
[b] *School of Mathematics, Liaoning Normal University, Dalian 116029, China*
[c] *Department of Physics, Dalian University of Technology, Dalian 116024, China*

Received 9 January 2006; accepted 24 February 2006

Abstract

A proximal bundle method with inexact data is presented for minimizing an unconstrained nonsmooth convex function f. At each iteration, only the approximate evaluations of f and its ε-subgradients are required and its search directions are determined via solving quadratic programmings. Compared with the pre-existing results, the polyhedral approximation model that we offer is more precise and a new term is added into the estimation term of the descent from the model. It is shown that every cluster of the sequence of iterates generated by the proposed algorithm is an exact solution of the unconstrained minimization problem.

MR Subject Classification: 49J52; 49M37; 90C20; 90C25; 90C30

Keywords: Approximate subgradient; Bundle method; Convex optimization; Nonlinear programming; Nonsmooth optimization; Proximal bundle method

1. Introduction

We consider the problem

$$\min\{f(x) \mid x \in R^n\}, \tag{1.1}$$

☆ Supported by the National Science Foundation of China under Grant 10471015.
* Corresponding author at: Department of Applied Mathematics, Dalian University of Technology, Dalian 116024, China.

E-mail addresses: shenjietongtong@sina.com (J. Shen), zqxia@dlut.edu.cn (Z.-Q. Xia), lppang@dlut.edu.cn (L.-P. Pang).

0362-546X/$ - see front matter © 2007 Published by Elsevier Ltd
doi:10.1016/j.na.2006.02.039

where $f : R^n \to R$ is a nonsmooth convex function. It is well known that many practical problems can be formulated as (1.1), for example the problems of catastrophe, ruin, vitality, data mining, and finance. At the moment, bundle methods are the most efficient and promising methods for solving nonsmooth optimization problems.

Traditional bundle methods use the exact values of f and its subgradients. But, generally, it is difficult to evaluate them. For example, consider the situation of Lagrangian relaxation; see [1–3, 7]. The primal problem is

$$\max\{q(\xi) \mid \xi \in P, h(\xi) = 0\}, \tag{1.2}$$

where P is a compact subset of R^m and $q : R^m \to R$, $h : R^m \to R^n$. Lagrangian relaxation of the equality constraints in this problem leads to problem (1.1), where

$$f(x) = \max_{\xi \in P}\{q(\xi) + \langle x, h(\xi)\rangle\} \tag{1.3}$$

is the dual function. Trying to solve problem (1.2) by means of solving its dual problem (1.1) makes sense in many situations. In this case, evaluating the function value $f(x)$ and a subgradient $z(x) \in \partial f(x)$ requires solving the optimization problem (1.3) exactly. Actually, in some cases, computing exact values of f is unnecessary and inefficient. For this reason, some modifications of bundle methods using inexact data were proposed; for instance, see [4,6] and [9]. In [4], the proposed algorithm only requires the ε-subgradient $\xi(x)$ of f at a given point $x \in R^n$ for fixed $\varepsilon \geq 0$, i.e., $\xi(x) \in \partial_\varepsilon f(x) = \{y \in R^n \mid f(\zeta) \geq f(x) + \langle y, \zeta - x\rangle - \varepsilon, \forall \zeta \in R^n\}$, the ε-subdifferential of f at x; see [8]. An ε-optimal solution of (1.1) can be found without assuming the knowledge of ε. But the exact function values of f are needed, so this setting is not suitable for some important applications of bundle methods, for example as discussed above, for the Lagrangian relaxation.

On the basis of the above observation, throughout this paper we make the following assumptions: at each given point $x \in R^n$, and for $\varepsilon \geq 0$, we can find some $\tilde{f} \in R$ and $z(x) \in R^n$ satisfying

$$\begin{aligned} &\tilde{f} \in [f(x) - \varepsilon, f(x)] \\ &f(\zeta) \geq f(x) + \langle z(x), \zeta - x\rangle - \bar{\varepsilon}, \quad \forall \zeta \in R^n, \end{aligned} \tag{1.4}$$

where $\bar{\varepsilon} = (f(x) - \tilde{f})$. At the same time, it can be ensured that

$$\tilde{f}_j \to \tilde{f}_k \quad \text{if } y_j \to x^k, \tag{$*$}$$

where $\tilde{f}_j \in [f(y^j) - \varepsilon^j, f(y^j)]$ and $\tilde{f}_k \in [f(x^k) - \varepsilon^k, f(x^k)]$ for given $\varepsilon^j > 0$ and $\varepsilon^k > 0$. This setting is realistic in many applications; see [5,9]. The condition (1.4) means that $z(x) \in \partial_\varepsilon f(x)$. A similar assumption was once used in [6], in which (1.4) is replaced by $\tilde{f} \in [f(x) - \varepsilon, f(x) + \varepsilon]$ and its ε-subgradients. But, when compared with [6], it will be seen that, under the assumption (1.4), the polyhedral approximation model of f constructed in Section 2 is above the one proposed in [6], such that the algorithm proposed in Section 5 has a much stronger convergence property.

This paper is organized as follows. In Sections 2 and 3 we introduce a polyhedral approximation of the objective function by employing the inexact data of the involved functions and find a possible descent direction by solving a quadratic programming. The inner iteration for determining the serious step and the null step is presented in Section 4. In the last section, we offer an algorithm for solving (1.1) and analyze its convergence as well.