European Journal of Operational Research 240 (2015) 315–327

Contents lists available at ScienceDirect

European Journal of Operational Research

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

Continuous Optimization

Inexact subgradient methods for quasi-convex optimization problems

Yaohua Hu [a], Xiaoqi Yang [b,*], Chee-Khian Sim [c]

[a] Department of Mathematics, Zhejiang University, Hangzhou 310027, PR China
[b] Department of Applied Mathematics, The Hong Kong Polytechnic University, Kowloon, Hong Kong
[c] Department of Decision Sciences, National University of Singapore, 15 Kent Ridge Drive, Singapore 119245, Singapore

ARTICLE INFO

Article history:
Received 8 April 2013
Accepted 12 May 2014
Available online 20 May 2014

Keywords:
Subgradient method
Quasi-convex optimization
Noise
Weak sharp minima

ABSTRACT

In this paper, we consider a generic inexact subgradient algorithm to solve a nondifferentiable quasi-convex constrained optimization problem. The inexactness stems from computation errors and noise, which come from practical considerations and applications. Assuming that the computational errors and noise are deterministic and bounded, we study the effect of the inexactness on the subgradient method when the constraint set is compact or the objective function has a set of generalized weak sharp minima. In both cases, using the constant and diminishing stepsize rules, we describe convergence results in both objective values and iterates, and finite convergence to approximate optimality. We also investigate efficiency estimates of iterates and apply the inexact subgradient algorithm to solve the Cobb–Douglas production efficiency problem. The numerical results verify our theoretical analysis and show the high efficiency of our proposed algorithm, especially for the large-scale problems.

1. Introduction

Subgradient methods are popular and practical techniques used to minimize a nondifferentiable convex function. Subgradient methods originated with the works of Polyak (1967) and Ermoliev (1966) and were further developed by Shor, Kiwiel, and Ruszczyński (1985). In the last 40 years, many properties of subgradient methods have been discovered, generalizations and extensions have been proposed, and various applications have been found (see Auslender & Teboulle, 2004; Bertsekas, Nedić, & Ozdaglar, 2003; Hiriart-Urruty & Lemaréchal, 1996; Larsson, Patriksson, & Strömberg, 1996; Nedić & Bertsekas, 2001; Nesterov, 2009; Patriksson, 2008; Shor et al., 1985 and references therein). Nowadays, the subgradient method still remains an important tool for nonsmooth and stochastic optimization problems, special for large-scale problems, due to its simple formulation and low storage requirement.

Motivated by practical reasons, approximate subgradient methods (also called ϵ-subgradient methods) are widely studied in Auslender and Teboulle (2004), D'Antonio and Frangioni (2009), Kiwiel (2004), Larsson, Patriksson, and Strömberg (2003), Shor et al. (1985). Kiwiel (2004) proposed a unified convergence framework for approximate subgradient methods. The author presented convergence in objective values and convergence to a neighborhood of the optimal solution set, using both the diminishing and nonvanishing stepsize rules. Larsson et al. (2003) proposed and analyzed conditional ϵ-subgradient methods to solve convex optimization problems and convex–concave saddle-point problems. Improving conditional subgradient methods, D'Antonio and Frangioni (2009) combined the deflection and the conditional subgradient technique into one iterative process, and investigated the unified convergence analysis for the deflected conditional approximate subgradient methods, using both the Polyak-type and diminishing stepsize rules. Furthermore, Auslender and Teboulle (2004) proposed and developed an interior ϵ-subgradient method for convex constrained optimization problems over polyhedral sets, in particular $\mathbb{R}^n_+$, via replacing the Euclidean distance function by a logarithmic-quadratic distance-like function.

Recently, Nedić and Bertsekas (2010) investigated the effect of noise on subgradient methods for convex optimization problems. Their work was motivated by the distributed optimization in networks where the data is quantized before being transmitted between nodes (see Kashyap, Basar, & Srikant, 2007; Rabbat & Nowak, 2005 and references therein). When the constraint set is compact or the objective function has a set of weak sharp minima, the authors established convergence properties to the optimal value within some tolerance, which is expressed in terms of errors and noise, under the bounded subgradient assumption.

Quasi-convex optimization problems can be found in important applications in various areas, such as economics, engineering,

* Corresponding author. Tel.: +852 27666954; fax: +852 23629045.
E-mail addresses: hyh19840428@163.com (Y. Hu), mayangxq@polyu.edu.hk (X. Yang), bizsimc@nus.edu.sg (C.-K. Sim).

http://dx.doi.org/10.1016/j.ejor.2014.05.017
0377-2217/

management science and various applied sciences (see Avriel, Diewert, Schaible, & Zang, 1988; Crouzeix, Martinez-Legaz, & Volle, 1998; Hadjisavvas, Komlósi, & Schaible, 2005 and references therein). The study of using subgradient methods to solve quasi-convex optimization problems has been limited. Using the diminishing stepsize rule, Kiwiel (2001) studied convergence properties and efficiency estimates of the exact subgradient method for solving a quasi-convex optimization problem under the assumption that the objective function is upper semi-continuous. On the other hand, modified dual subgradient algorithms were investigated in Gasimov (2002) and Burachik, Gasimov, Ismayilova, and Kaya (2006) for solving a general nonconvex optimization problem with equality constraints by virtue of a sharp augmented Lagrangian.

Motivated by practical and theoretical reasons, in this paper, we focus on an inexact subgradient algorithm for solving the following quasi-convex optimization problem:

$$\begin{aligned} \min \quad & f(x) \\ \text{s.t.} \quad & x \in X, \end{aligned} \tag{1.1}$$

where $f: \mathbb{R}^n \to \mathbb{R}$ is a quasi-convex function and the constraint set X is nonempty, closed and convex. We denote the optimal solution set and the optimal value respectively by X^* and f_*, and we assume that X^* is nonempty and compact.

Inspired by the idea in Nedić and Bertsekas (2010) and references therein, we investigate the influence of inexact terms, including both computation errors and noise, on the inexact subgradient algorithm. The computation errors, which give rise to the ϵ-subgradient, is inevitable in computing process. On the other hand, the noise may come from practical considerations and applications, and is manifested in inexact computation of subgradients. Considering a generic inexact subgradient algorithm for the quasi-convex optimization problem (1.1) and assuming that the computational errors and noise are deterministic and bounded, we establish convergence properties in both objective values and iterates within some tolerance given explicitly in terms of errors and noise. We also describe the finite convergence behavior to approximate optimality and efficiency estimates of iterates.

The quasi-convex function is more difficult to deal with, as the epigraph of a convex function is convex; while only the sublevel set of a quasi-convex function is convex. Lacking the convexity assumed in Nedić and Bertsekas (2010), the main technical challenges are defining a suitable subdifferential of a quasi-convex function, establishing the proper basic inequality, which is a key tool needed in this area of study, and applying the convexity of the sublevel set instead of that of the epigraph of a convex function, when analyzing the inexact subgradient method algorithm for the quasi-convex optimization problem. To meet these challenges, we adopt the closure of Greenberg–Pierskalla subdifferential as the quasi-convex subdifferential, introduce the Hölder condition to relate the quasi-convex subgradient with objective function values and establish the basic inequality, which is only a local property though, and then obtain the convergence property in objective values and finite convergence under the Hölder condition, instead of the upper semi-continuity of the objective function used in Kiwiel (2001). Another contribution is to describe the convergence property in iterates, which are absent in Nedić and Bertsekas (2010), by virtue of convexity of a sublevel set. When X is noncompact, we need to assume an additional generalized weak sharp minima condition. This condition extends the concept of weak sharp minima in Nedić and Bertsekas (2010) and is presented by using $\text{dist}(x, X^*)$, the distance of the decision variable x to X^*.

We also investigate the quantification of the influence of errors and noise by using both the constant and diminishing stepsize rules, while only the diminishing stepsize rule is considered in studying convergence properties and efficiency estimates of an exact subgradient method in Kiwiel (2001).

We further consider the fractional programming as an application of the quasi-convex model, describe the Cobb–Douglas production efficiency problem as an example, and perform some numerical experiments on this problem via applying the inexact subgradient method. The numerical results verify our theoretical analysis and show that the quasi-subgradient type method is highly efficient for the production efficiency problem, even when the problem is large-scale.

This paper is organized as follows. In Section 2, we present the notations used in this paper, the quasi-subdifferential theory and the inexact subgradient algorithm. In Section 3, we establish convergence properties in both objective values and iterates, and finite convergence behavior of our algorithm when the constraint set X is compact. Section 4 presents the convergence behavior when f has a set of generalized weak sharp minima over noncompact X, and Section 5 gives the efficiency estimates. Finally in Section 6, we apply our algorithm to the Cobb–Douglas production efficiency problem, and demonstrate the numerical results.

2. Preliminaries

2.1. Notation and terminology

We consider the n-dimensional Euclidean space $\mathbb{R}^n$. We view vector as a column vector, and denote by $\langle x, y\rangle$ the inner product of two vectors $x, y \in \mathbb{R}^n$. We use $\|x\|$ to denote the standard Euclidean norm, $\|x\| = \sqrt{\langle x, x\rangle}$. For $x \in \mathbb{R}^n$ and $\delta \in \mathbb{R}_+$, $B(x, \delta)$ denotes the closed ball of radius δ centered at x and specially B denotes the unit closed ball at the origin. For a set $Z \subseteq \mathbb{R}^n$, we denote the closure of Z by cl$Z$. We also write $\text{dist}(x, Z)$ to denote the Euclidean distance of a vector x from the set Z, i.e.,

$$\text{dist}(x, Z) = \inf_{z \in Z} \|x - z\|.$$

A function $f: \mathbb{R}^n \to \mathbb{R}$ is said to be quasi-convex if for all $x, y \in \mathbb{R}^n$ and $\alpha \in [0, 1]$, the following inequality holds

$$f((1-\alpha)x + \alpha y) \leqslant \max\{f(x), f(y)\}.$$

f is said to be upper semi-continuous (usc) on $\mathbb{R}^n$ if $f(x) = \limsup_{y \to x} f(y)$ for all $x \in \mathbb{R}^n$. For each $\alpha \in \mathbb{R}$, we denote the (strict) sublevel sets of f by

$$S_{f,\alpha} = \{x \in \mathbb{R}^n : f(x) < \alpha\}, \quad S_f(x) = S_{f,f(x)},$$
$$\bar{S}_{f,\alpha} = \{x \in \mathbb{R}^n : f(x) \leqslant \alpha\}, \quad \bar{S}_f(x) = \bar{S}_{f,f(x)}.$$

It is well-known that f is quasi-convex if and only if $S_{f,\alpha}(\bar{S}_{f,\alpha})$ is convex for all $\alpha \in \mathbb{R}$, and that f is usc on $\mathbb{R}^n$ if and only if $S_{f,\alpha}$ is open for all $\alpha \in \mathbb{R}$.

2.2. Quasi-subdifferential theory

There are many different types of subdifferential, such as Clarke–Rockafellar subdifferential, Dini subdifferential, Fréchet subdifferential (see Aussel, Corvellec, & Lassonde, 1995 and references therein) and so on. They are the same for convex functions, but different for nonconvex functions. Here we introduce the Greenberg–Pierskalla subdifferential, defined by Greenberg and Pierskalla (1973), as follows.

Definition 2.1 (*see Greenberg & Pierskalla, 1973*). The z-quasi-conjugate of f is a function $f_z^*: \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$, defined by

$$f_z^*(x) = z - \inf\{f(y) : \langle x, y\rangle \geqslant z\}.$$

It is recalled in Greenberg and Pierskalla (1973, Theorem 1) that the z-quasi-conjugate function provides a lower bound for the corresponding convex conjugate function, and indeed, the