Journal of Computational Mathematics, Vol.21, No.4, 2003, 451–462.

PROXIMAL POINT ALGORITHM FOR MINIMIZATION OF DC FUNCTION [*1)]

Wen-yu Sun
(*School of Mathematics and Computer Science, Nanjing Normal University, Nanjing 210097, China*)

Raimundo.J.B. Sampaio
(*Programa de Pos-Graduacao em Informatica Aplicada, Pontificia Universidade Catolica do Parana (PUCPR), CEP: 80215-901, Curitiba, PR, Brazil*)

M.A.B. Candido
(*Programa de Pos-Graduacao em Informatica Aplicada, Pontificia Universidade Catolica do Parana (PUCPR), CEP: 80215-901, Curitiba, PR, Brazil*)

Abstract

In this paper we present some algorithms for minimization of DC function (difference of two convex functions). They are descent methods of the proximal-type which use the convex properties of the two convex functions separately. We also consider an approximate proximal point algorithm. Some properties of the ϵ-subdifferential and the ϵ-directional derivative are discussed. The convergence properties of the algorithms are established in both exact and approximate forms. Finally, we give some applications to the concave programming and maximum eigenvalue problems.

Key words: Nonconvex optimization, Nonsmooth optimization, DC function, Proximal point algorithm, ϵ-subgradient.

1. Introduction

In this paper we consider solving a special class of nonconvex optimization problems:

$$\min_{x\in R^n} \quad f(x), \tag{1.1}$$

where $f : R^n \to R$ is a nonconvex function. In many cases, for example, in optimal control and engineering design, the nonconvex function f can be dealt with as a difference of two convex functions

$$f(x) = g(x) - h(x), \quad \forall x \in R^n, \tag{1.2}$$

where $g : R^n \to R$ and $h : R^n \to R$ are proper, convex, and lower semi-continuous (l.s.c.). In this case, the function f is called DC function.

The interest for studying DC function (i.e. difference of two convex functions) is motivated by the possibility of using twice the underlying convex structure of such representation when dealing with nonconvex problems. This is especially attractive when one of these convex functions or both is nonsmooth. Although there is a lot of papers devoted to the theory of DC functions in the literature (see for example, [6] [7] [8]), only a few have proposed some specific algorithms and reported some numerical experiments. Here we quote some methods which use the regularization approach [3] [19], the dual approach [1] and the subgradient method [13], respectively.

[*] Received February 13, 2001; final revised July 16, 2001.

[1)] This work was supported by the National Natural Science Foundation of China, the Oversea Exchange Fund of Nanjing Normal University, and CNPq of Brazil.

It is well-known that proximal point algorithm (PPA) is an effective method for solving nonsmooth convex optimization problems. Its remarkable feature is that a nonsmooth convex optimization problem can be converted to a continuously differentiable convex optimization problem. Consequently, we can use some methods for smooth optimization to deal with it. This paper aims to study using proximal point algorithm to minimize a DC function.

Let $\langle \cdot, \cdot \rangle$ denote the inner product in R^n, Γ_0 the set of convex proper and l.s.c functions on R^n. Let $f: R^n \to R$ be a DC function on R^n, i.e. there exist g and h both in Γ_0 such that

$$f(x) = g(x) - h(x), \ \forall x \in R^n. \tag{1.3}$$

Moreover, suppose that $\mathrm{Dom}(g) \cap \mathrm{Dom}(h) \neq \phi$, where $\mathrm{Dom}(g)$ denotes the domain of g

$$\mathrm{Dom}(g) := \{x \in R^n \mid g(x) < \infty\}.$$

The functions g and h can be chosen as strongly convex since one can always add a strongly convex function to each function, for example,

$$f(x) = [g(x) + \omega(x)] - [h(x) + \omega(x)],$$

where $\omega : R^n \to R$ is a strongly convex function. The corresponding conjugate function of g and h are denoted by g^* and h^*, and their respective subdifferentials by ∂g, $\partial h, \partial g^*$ and ∂h^*.

Proposition 1.1. (see [23] [8])

1.

$$\inf_{x \in R^n} \{g(x) - h(x)\} = \inf_{y \in R^n} \{h^*(y) - g^*(y)\}. \tag{1.4}$$

2. A necessary condition for $x \in Dom(f)$ to be a local minimizer of f is

$$\partial h(x) \subset \partial g(x). \tag{1.5}$$

In general, the condition 2 above is hard to be reached and one may relax it to

$$\partial g(x) \cap \partial h(x) \neq \phi. \tag{1.6}$$

We say that x^* is a critical point of f if it satisfies (1.6).

The method presented in this paper is closely related to the proximal point algorithm (see e.g. [17]). This class of algorithms finds a zero of a maximal monotone operator T by means of the following iteration:

$$x_{k+1} = (I + c_k T)^{-1} x_k, \tag{1.7}$$

where $c_k > c > 0$, where c is a suitably small positive number such that $I + cT$ is nonsingular. The operator $P_k = (I + c_k T)^{-1}$ which is the resolvent of T, is nonexpansive, single-valued on the whole space, and Lipschitz continuous. When T is the subdifferential of a convex l.s.c. function g, i.e., $T = \partial g$, the iteration (1.7) becomes

$$x_{k+1} = \arg\min\{g(x) + \frac{1}{2c_k} \parallel x - x_k \parallel^2\}. \tag{1.8}$$

Rockafellar [17] [18] has developed a detailed study of the convergence on proximal point algorithm. In particular, the algorithm converges linearly at least. If $c_k \to \infty$, the convergence is superlinear. In addition, the attractive approximate versions of proximal point algorithm are established by [17] [10].

With this strategy, we propose a new descent algorithm for finding a critical point of a DC function which satisfies necessary optimality conditions. Each iteration combines an ascent subgradient step on the second function with a proximal step on the first function. In addition, the approximate version of our algorithm is also discussed.