Journal of Computational Mathematics, Vol.24, No.4, 2006, 475–480.

ON THE CONVERGENCE OF AN APPROXIMATE PROXIMAL METHOD FOR DC FUNCTIONS *

A. Moudafi P-E. Maingé
(*Université Antilles Guyane, GRIMAAG, Département Scientifique interfacultaires, 97200 Schelcher, Martinique, France*)

Abstract

In this paper we prove the convergence of the approximate proximal method for DC functions proposed by Sun et al [6]. Our analysis also permits to treat the exact method. We then propose an interesting result in the case where the second component of the DC function is differentiable and provide some computational experiences which proved the efficiency of our method.

Mathematics subject classification: 49J53, 65K10, 49M37, 90C25.
Key words: DC minimization, Critical points, Subdifferentials, Proximal mappings.

1. Introduction

The proximal point algorithm was introduced by Martinet [2] for solving proper lower semi-continuous convex minimization problems and extensively studied by Rockafellar [4] in the context of monotone variational inequalities. It is well-known that if we drop the convexity assumption on the objective function several problems arise. The proximal mapping is not well-defined and in general it is not nonexpansive anymore even in arbitrary small neighbourhoods of minima. Only few research has been proposed concerning the construction of solutions in this nonconvex case, see for instance [3]. Here we focus our attention on the method recently proposed by Sun et al. [6]. To find a critical point of $f := g - h$, this method consists to increasing the function h along the direction of the subgradient and then decreasing the function f thanks to a proximal step. They proved that if the sequence generated by their algorithm is bounded, then every cluster point is a critical point of f. The aim of the paper is to provide a correct proof for the main result in the article [6]. Indeed, we propose a right and elementary proof of the convergence result for the approximate form and we provide conditions ensuring the boundedness of the generated sequences. Aftewards, by means of an epi-convergence argument, we propose an interesting result in the case where the second component of the DC function is differentiable. We then give some numerical experiments which proved the convergence of the algorithm PMDC to local solutions and showed at the same time its robustness and efficiency with respect to the algorithm DCA introduced by Pham Dinh Tao [3].

Let f be a DC function, i.e. $f = g - h$ where f and g are two convex lower semi-continuous and proper functions defined on $\mathbb{R}^n$ satisfying $dom g \cap dom h \neq \emptyset$. We consider the problem:

$$\min_{x\in\mathbb{R}^n} (g(x) - h(x)) \tag{1.1}$$

and the associated dual

$$\min_{y\in\mathbb{R}^n} (h^*(y) - g^*(y)), \tag{1.2}$$

* Received January 4, 2005; Final revised September 19, 2005.

$g^*(y) = \sup_{x \in \mathbb{R}^n}(\langle y, x\rangle - g(x)), h^*$ stand for the conjugate functions of g and h.
It is well known that

$$\inf_{x \in \mathbb{R}^n}(g(x) - h(x)) = \inf_{y \in \mathbb{R}^n}(h^*(y) - g^*(y)),$$

and that a necessary condition for $x \in domf$ to be a local minimizer of f is $\partial h(x) \subset \partial g(x)$. As in general this necessary condition is hard to reach, we will focus our attention on finding critical points of f, namely points satisfying the relaxed condition $\partial h(x) \cap \partial g(x) \neq \emptyset$.
Throughout the paper $f := g - h : \mathbb{R}^n \to \overline{\mathbb{R}}$ is a real DC function. We recall that a vector w is called an ε-subgradient (with $\varepsilon \geq 0$) of g at $x \in domg$, if

$$g(u) \geq g(x) + \langle w, u - x\rangle - \varepsilon \quad \forall u \in \mathbb{R}^n. \tag{1.3}$$

The set of all ε-subgradients of g at x, denoted by $\partial_\varepsilon g(x)$, is called the approximate subdifferential of g at x which reduces to exact subdifferential when $\varepsilon = 0$. We also recall that the Moreau-Yosida approximate and the approximate proximal mapping of g are defined for $c > 0$ by

$$g_c(x) := \inf_{u \in \mathbb{R}^n}\{g(u) + \frac{1}{2c}\|u - x\|^2\} \quad \text{and} \quad prox^{\varepsilon}_{cg}(x) := (I + c\partial_\varepsilon g)^{-1}(x).$$

It is worth mentioning the richness of the class of DC functions which contains the class of lower-$\mathcal{C}^2$ functions and constitutes a minimal realistic extension of the class of convex functions. It has been successfully used in many nonconvex applications such as finance, molecular biology, mutlicommodity network, image restoration processing and seems particularly well suited to model several nonconvex industrial problems (Robotic: computer's vision, fuel mixture ...).

2. Approximate Proximal Point Algorithms

2.1 ε-Proximal Method for DC functions

The method we will study is an approximate form of the scheme by Sun et al which is based on the following equivalence:

$$x \text{ is a critical point of } g - h \Leftrightarrow x = prox_{cg}(x + cw), \forall c > 0 \text{ and } w \in \partial h(x).$$

Thanks to this fixed-point formulation, Sun et al. [6] proposed an algorithm for finding a critical point of a DC function. This method combines the proximal point algorithm with the subgradient method. Here we consider the approximate version obtained by replacing the exact subdifferential by the approximate one, since the function h (respectively g) is assumed to be convex, proper and lower semicontinuous, $\partial_0 h(x) = \partial h(x)$, for any x. Furthermore, directly from the definition it follows that $0 \leq \varepsilon_1 \leq \varepsilon_2 \Rightarrow \partial_{\varepsilon_1} h(x) \subseteq \partial_{\varepsilon_2} h(x)$. Thus $\partial_\varepsilon h(x)$ is an enlargement of $\partial h(x)$. The use of elements in $\partial_\varepsilon h$ instead of ∂h allows an extra degree of freedom which is very useful in various applications. On the other hand, setting $\varepsilon = 0$ one retrieves the exact subdifferential, so that the exact method can be also treated. For all these reasons, we consider the following scheme:

Algorithm. Proximal Method for DC Functions (PMDC)
Setp 1: Given $x_0, c_0 \geq c$. Set $k = 0$.
Step 2: Compute $w_k \in \partial_{\varepsilon_k} h(x_k)$ and set $y_k = x_k + c_k w_k$.
Step 3: Compute $x_{k+1} \in prox^{\varepsilon_k}_{c_k g}(y_k)$ (Proximal step).
If $x_{k+1} = x_k$ stop. Otherwise increase k by 1 and loop to step 2.

The following proposition contains the convergence results of PMDC.

Theorem 2.1. *Assume that $f := g - h$ is bounded from below, $c_k \geq c > 0$ for any $k \in \mathbb{N}$ and suppose that $\sum_{k=0}^{+\infty} \varepsilon_k < +\infty$, then the sequence $(f(x_k))_{k \in \mathbb{N}}$ is convergent and the sequence (x_k) is asymptotically regular in the following sense: $\lim_{k \to +\infty} c_k^{-1}\|x_k - x_{k+1}\| = 0$. Moreover, if the sequences (x_k) and (w_k) are bounded, then every cluster-point x_∞ and w_∞ of the sequences (x_k) and (w_k) are critical points of the functions $g - h$ and $h^* - g^*$, respectively.*