Convergence Analysis of DC Algorithm for DC programming with subanalytic data

Le Thi Hoai An[*], Huynh Van Ngai[†] and Pham Dinh Tao[‡]

Abstract

DC Programming and DCA have been introduced by Pham Dinh Tao in 1986 and extensively developed by Le Thi Hoai An and Pham Dinh Tao since 1993. These approaches have been successfully applied to solving real life problems in their large scale setting. In this paper, by using the Lojasiewicz inequality for nonsmooth subanalytic functions, we investigate the convergence of DC (difference of convex) algorithm (DCA) for solving DC program with subanalytic data. The convergent rate which depends on the so-called Lojiasiewicz exponent has been established.

Keywords: DC program, DC algorithm, subanalytic, subdifferential, Lojasiewicz exponent **Mathematics Subject Classification.** 90C30, 90C26, 47N10

1 Introduction

Let $\Gamma_0(\mathbb{R}^n)$ denote the convex cone of all lower semicontinuous proper convex functions on $\mathbb{R}^n$. The vector space of DC functions, $DC(\mathbb{R}^n) = \Gamma_0(\mathbb{R}^n) - \Gamma_0(\mathbb{R}^n)$, is quite large to contain almost real life objective functions and is closed under all the operations usually considered in Optimization.

Consider the standard DC program

$$\alpha := \inf\{f(x) := g(x) - h(x) \ : \ x \in \mathbb{R}^n\} \quad (P_{dc})$$

with $g, h \in \Gamma_0(\mathbb{R}^n)$. The function f is called a DC function on $\mathbb{R}^n$. Remark that the closed convex constraint set C is incorporated in the first convex DC component g with the help of its indicator function χ_C ($\chi_C\ (x) := 0$ if $x \in C, +\infty$ otherwise). Recall the natural convention $+\infty - (+\infty) = +\infty$ and that $\alpha \in \mathbb{R}$ implies *dom* $g := \{x \in \mathbb{R}^n : g(x) < +\infty\} \subset dom\ h$.

DC programming and DCA (DC Algorithms) have been introduced by Pham Dinh Tao in 1985 and extensively developed by Le Thi Hoai An and Pham Dinh Tao since 1994. DC programming plays a key role in nonconvex programming because almost nonconvex programs encountered in practice are DC programs. Based on local optimality conditions and DC duality, DCA is one of efficient algorithms for nonconvex programs, especially for nonsmooth ones. Due to its local character it cannot guarantee the globality of computed solutions for general DC programs. However, we observe that, with a suitable starting point, it converges quite often to a global one. In practice, DCA was successfully applied to a lot of different and various nonconvex programs to which it gave almost always global solutions and proved to be more robust and more efficient than related standard methods, especially in the large-scale setting. On the other hand, it is worth noting that, with appropriate DC decompositions, DCA permits to find again standard optimization algorithms for convex and nonconvex programming (see [7], [8], [14], [15] and references therein).

[*]Laboratoiry of Theoretical and Applied Computer Science, UFR MIM, Université de Metz, Ile du Saulcy 57045 Metz, France, lethi@univ-metz.fr

[†]Department of Mathematics, University of Quynhon, 170 An Duong Vuong, Qui Nhon, Vietnam, nghiakhiem@yahoo.com

[‡]Laboratory of Modelling, Optimization and Operations Research, LMI-National Institute for Applied Sciences-Rouen, BP 8, F 76 131 Mon Saint Aignan Cedex, France pham@insa-rouen.fr

In this paper, we consider DC programming with subanalytic data. The class of subanalytic functions which contains analytic functions, has been first introduced by Lojasiewicz ([10], [9], [11]) and, next has been extensively studied by many authors (see, e.g., [3], [5], [18] and references therein). This class of functions possesses many interesting properties; it is sufficiently rich to cover many practical optimizations problems. Due to the Stone-Weierstrass theorem, in principle, every continuous functions can be approximated by a subanalytic DC function with any desired precision. Recently, in [2], Attouch-Bolte have proved a result on the convergence of the proximal method for solving nonsmooth optimization problems by using the Lojiasiewicz inequality. This Lojasiewicz inequality is an ingredient tool in Subanalytic Geometry as well as in its applications to subanalytic continuous/descrete dynamical systems, which has been first established by Lojasiewicz for analytic functions, and recently generalized to nonsmooth subanalytic functions by Bolte-Daniilidis-Lewis in [3].

In this paper, by using the Lojiasiewicz inequality for nonsmooth subanalytic function as well as , we establish the convergence of the DC algorithm for DC programming with subanalytic data. As shown in [14], the proximal algorithm in Convex optimization can be consider as a special case of DCA, the convergent result in this paper therefore generalizes the result of Attouch-Bolte ([2]) in the context of Convex Optimization.

2 Subanalytic functions and Lojasiewicz inequality

We briefly recall the notion of subanalytic functions (see [10], [11]).

Definition 2.1 *(i) A subset C of $\mathbb{R}^n$ is said to be semianalytic if each point of $\mathbb{R}^n$, there exists a neighborhood V such that $C \cap V$ is of the following form:*

$$C \cap V = \bigcup_{i=1}^{p} \bigcap_{j=1}^{q} \{x \in V : \; f_{ij}(x) = 0, \, g_{ij}(x) > 0\},$$

where $f_{ij}, \, g_{ij} : V \to \mathbb{R}$ $(1 \leq i \leq p, \, 1 \leq j \leq q$ $)$ are real-analytic functions.

(ii) A subset C of $\mathbb{R}^n$ is called subanalytic if each point of $\mathbb{R}^n$, there exists a neighborhood V such that

$$C \cap V = \{x \in \mathbb{R}^n : \; \exists y \in R^m, \, (x, y) \in D\},$$

where D is a bounded seminalytic subset of $\mathbb{R}^n \times \mathbb{R}^m$ with $m \geq 1$.

(iii) A function $f : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ is said to be subanalytic if its graph gph f *is a subanalytic subset of $R^n \times \mathbb{R}$.*

It is obvious that the class of subanalytic sets (resp. functions) contains all analytic sets (resp. functions). Let us list some of the elementary properties of subanalytic sets and subanalytic functions (see, e.g., [5], [10], [18]):

- Subanalytic sets are closed under locally finite union and intersection. The complement of a subanalytic set is subanalytic.

- The closure, the interior, the boundary of a subanalytic set are subanalytic.

-A closed $C \subseteq \mathbb{R}^n$ is subanalytic *iff* its indicator function χ_C, defined by $\chi_C(x) = 0$ if $x \in C$ and $+\infty$ otherwise, is subanalytic.

- Given a subanalytic set C, the distance function $d_C(x) := \inf_{z \in C} \|x - z\|$ is a subanalytic function.

- Let $f, g : X \to \mathbb{R}$ be continuous subanalytic functions, where $X \subseteq \mathbb{R}^n$ is a subanalytic set. Then the sum $f + g$ is subanalytic if f maps bounded sets on bounded sets, or if both two functions are bounded from below.

- Let $X \subseteq \mathbb{R}^n$, $T \subseteq \mathbb{R}^m$ be subanalytic sets, where T is compact. If $f : X \times T \to \mathbb{R}$ is a continuous subanalytic function, then $g(x) := \min_{t \in T} f(x, t)$ is continuous subanalytic.

-Let $X \subseteq \mathbb{R}^n$ be subanalytic set and let $f, g : X \to \mathbb{R}$ be subanalytic functions. Then, $f + g$ is subanalytic if f maps bounded sets on bounded sets, or both two functions are bounded from below.

The following proposition gives the subanalyticity of conjugate functions.

Proposition 2.2 *If $f : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ is a lower semicontinuous subanalytic strongly convex function then its conjugate f^* is a $C^{1,1}$ (the class of functions whose derivative is Lipschitz) subanalytic convex function.*

Proof. It is well-known that the conjugate of a lower semicontinuous strongly convex function is of the class $C^{1,1}$. Let us prove that f^* is subanalytic. Let $\rho > 0$ is the modulus of the strong convexity of f. Take $\rho_1 \in (0, \rho)$ and set $h(x) := f(x) - \frac{\rho_1}{2}\|x\|^2$, $x \in \mathbb{R}^n$. Then h is a strongly convex function. One has, by the definition of the conjugate

$$\begin{aligned} f^*(x) = -\inf_{y\in\mathbb{R}^n}\{f(y) - \langle x, y\rangle\} &= -\inf_{y\in\mathbb{R}^n}\{h(y) + \tfrac{\rho_1}{2}\|y\|^2 - \langle x, y\rangle\} \\ &= -\inf_{y\in\mathbb{R}^n}\{h(y) + \tfrac{\rho_1}{2}\|x/\rho_1 - y\|^2\} + \tfrac{1}{2}\|x\|^2 = -\varphi(x) + \frac{\rho_1^{-1/2}}{2}\|x\|^2, \end{aligned}$$

where,

$$\varphi(x) = \inf_{y\in\mathbb{R}^n}\{h(y) + \frac{\rho_1}{2}\|x/\rho_1 - y\|^2\},\ x \in \mathbb{R}^n.$$

According to Proposition 2.9 in ([3]), the function φ is subanalytic. Thus f^* is subanalytic. ■

Let us recall some notions from Convex Analysis and Nonsmooth Analysis, which will be needed thereafter (see [12], [16], [17]). In the sequel, the space $\mathbb{R}^n$ is equipped with the canonical inner product $\langle\cdot\rangle$. Its dual space is identified with $\mathbb{R}^n$ itself. $\mathcal{S}(\mathbb{R}^n)$ denotes the set of lower semicontinuous functions $f : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$. The open ball with the center $x \in \mathbb{R}^n$ and radius $\varepsilon > 0$ is denoted by $B(x, \varepsilon)$; while the unit ball (i.e., the ball with the center at the origin and unit radius) is denoted by B. A function $f : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ is called $\rho-$convex for some $\rho(f) \geq 0$, if for all $x, y \in \mathbb{R}^n$, $\lambda \in [0, 1]$ one has

$$f(\lambda x + (1-\lambda)y) \leq \lambda f(x) + (1-\lambda)f(y) - \frac{\rho}{2}\lambda(1-\lambda)\|x - y\|^2.$$

The supremum of all $\rho \geq 0$ such that the above inequality is verified is called the convex modulus of f, which is denoted by $\rho(f)$.

For a function $f : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ be a lower semicontinuous real extended valued function. The *Fréchet subdifferential* of f at $x \in \text{Dom } f$ is defined by

$$\partial^F f(x) = \left\{x^* \in \mathbb{R}^n : \ \liminf_{h\to 0} \frac{f(x+h) - f(x) - \langle x^*, h\rangle}{\|h\|} \geq 0\right\}.$$

For $x \notin \text{Dom } f$, we set $\partial f(x) = \emptyset$. A point $x_0 \in \mathbb{R}^n$ is called a (Fréchet) *critical* point for the function f, if $0 \in \partial^F f(x_0)$.

When f is a convex function , then ∂^F coincides with the subdifferential in the sense of Convex Analysis. Moreover, if f is a d.c. function, i.e., $f := g - h$, where g, h is convex functions, then

$$\partial^F f(x) \subseteq \partial g(x) - \partial h(x)$$

wherever h is continuous at x. Especially, if h is differentiable at x, then one has the equality:

$$\partial^F f(x) = \partial g(x) - \nabla h(x).$$

Let us recall the nonsmooth version of the Lojasiewicz inequality established by Bolte-Daniliidis-Lewis ([3]), which is needed in the convergence analysis of DCA in the sequel.

Theorem 2.3 *(Theorem 3.1, [3]) Let $f : \mathbb{R}^n \to \mathbb{R}\cup\{+\infty\}$ be a subanalytic function such that its domain* Dom f *is closed and $f|_{\text{Dom } f}$ is continuous and let x_0 is a Fréchet critical point of f. Then there exist $\theta \in [0, 1)$, $L > 0$ and a neighborhood V of x_0 such that the following inequality holds.*

$$|f(x) - f(x_0)|^\theta \leq L\|x^*\| \quad \textit{for all}\, x \in V,\ x^* \in \partial^F f(x),$$

where a convention $0^0 = 1$ is used.

The number θ in the theorem is called a *Lojasiewicz exponent* of the critical point x_0.

3 Convergence Analysis of DCA

We consider now a standard d.c. program, that is, a optimization problem of the form:

$$(\mathcal{P}) \quad \alpha = \inf\{f(x) := g(x) - h(x) : \quad x \in \mathbb{R}^n\}$$

where g, h belong to $\Gamma_0(\mathbb{R}^n)$, the class of lower semicontinuous proper convex functions on $\mathbb{R}^n$. The dual problem of ($\mathcal{P}$ is defined by

$$(\mathcal{D}) \quad \alpha = \inf\{h^*(y) - g^*(y) : \quad y \in \mathbb{R}^n\}.$$

where g^*, h^* are the conjugate functions of g, h, respectively, i.e.,

$$g^*(y) := \sup\{\langle x, y\rangle - g(x) : x \in \mathbb{R}^n\}.$$

A point $x^* \in \mathbb{R}^n$ (resp. $y^* \in \mathbb{R}^n$) is called a *critical* point of d.c problem ($\mathcal{P}$) if $0 \in \partial g(x^*) \cap \partial h(x^*)$ (resp. $0 \in \partial g^*(y^*) \cap \partial h^*(y^*)$.

For a d.c. optimization problem with a set constraint:

$$\inf\{f(x) : \quad x \in C\},$$

where f is a d.c. function and $C \subseteq \mathbb{R}^n$ is a nonempty convex set, we can equivalently transform it into a standard d.c. program by using the indicator function of C as follow.

$$\inf\{f(x) + \chi_C(x) : \quad x \in \mathbb{R}^n\}.$$

Note that if f is a subanalytic function and C is a subanalytic set then so is the function $f + \chi_C$.

In the convex approach to DC programming, the DCA, based on local optimality and DC duality. The DCA consists in the construction of the two sequences $\{x^k\}$ and $\{y^k\}$ (candidates for being primal and dual solutions, respectively) that we improve at each iteration (thus, the sequences $\{g(x^k) - h(x^k)\}$ and $\{h^*(y^k) - g^*(y^k)\}$ are decreasing) in an appropriate way such that their corresponding limits x^∞ and y^∞ satisfy local optimality conditions.

DC Algorithm (DCA) (simplified form) consists of constructing the two sequences $\{x^k\}$ and $\{y^k\}$, starting a given $x^0 \in \text{Dom}\ g$ by setting

$$(DCA) \qquad y^k \in \partial h(x^k); \quad x^{k+1} \in \partial g^*(y^k).$$

The proofs of the following theorems are based on the nonsmooth version of the Lojasiewicz inequality (Theorem 3.1, [3]), inspired by Lojasiewicz [10] and Attouch- Bolte [2].

Theorem 3.1 *Let us consider DC problem $\mathcal{P}$ with $\alpha \in \mathbb{R}$. Suppose that the sequences $\{x^k\}$, $\{y^k\}$ are defined by the DC Algorithm .*

(i) Suppose that the d.c. function $f := g - h$ is subanalytic such that Dom f *is closed; $f\mid_{\text{Dom}\ f}$ is continuous and that g or h is differentiable on* Dom h *or* Dom g, *respectively with locally Lipschitz derivative. Assume that $\rho(g) + \rho(h) > 0$, where $\rho(g)$, $\rho(h)$ are modulus of the strong convexity of g, h, respectively. If either the sequence $\{x^k\}$ or $\{y^k\}$ is bounded then $\{x^k\}$ and $\{y^k\}$ are convergent to critical points of ($\mathcal{P}$) and ($\mathcal{D}$), respectively.*

(ii) Similarly, for the dual problem, suppose that $h^* - g^*$ is subanalytic such that Dom $(h^* - g^*)$ *is closed; $(h^* - g^*)\mid_{\text{Dom}\,(h^*-g^*)}$ is continuous and that g^* or h^* is differentiable on* Dom g^* *or* Dom h^*, *respectively with locally Lipschitz derivative. If $\rho(g^*) + \rho(h^*) > 0$ and either the sequence $\{x^k\}$ or $\{y^k\}$ is bounded then $\{x^k\}$ and $\{y^k\}$ are convergent to critical points of ($\mathcal{P}$) and ($\mathcal{D}$), respectively.*

Proof. In view of the DC duality (see, e.g., [15]), it suffices to prove statement (i). Set $\rho = \rho(g) + \rho(h) > 0$. Due to Proposition 2.2, either g^* or h^* is of the class $C^{1,1}$. Since $x^{k+1} \in \partial g^*(y^k)$ and $x^k \in \partial h^*(y^k)$, then $\{x^k\}$ is bounded *iff* $\{y^k\}$ is bounded. Moreover, by the differentiability of either g or h, if $\{x^k\}$ is convergent then $\{y^k\}$ is also convergent. Thus, we can assume that $\{x^k\}$ is bounded.

If $x^{k+1} = x^k$ for some k, then x^k is a critical point of $(\mathcal{P})$ as shown in (Theorem 3.7, [15]). Assume now that $x^{k+1} \neq x^k$ for all k. According to Theorem 3.7 in [15], one has

$$f(x^k) - f(x^{k+1}) \geq \frac{\rho}{2}\|x^k - x^{k+1}\|^2, \quad \text{for all } k = 0, 1, ... \tag{1}$$

and any limit point of $\{x^k\}$ is a critical point of $(\mathcal{P})$. This follows that $\{f(x^k)\}$ is decreasing. By the boundedness, then $\{f(x^k)\}$ is a convergent sequence and therefore, $\lim_{k\to\infty} \|x^k - x^{k+1}\| = 0$. Without loss of generality, we can assume that $\lim_{k\to\infty} f(x^k) = 0$. Denote by $\mathcal{C}$ the set of all limit points of $\{x^k\}$. By assumption, $\{x^k\}$ is bounded, then obviously $\mathcal{C}$ is a compact set. Let us consider the following cases.

Case 1. g is differentiable and its derivative is locally Lipschitz. For each $x \in \mathcal{C}$, there exist $\kappa(x), \varepsilon(x) > 0$ such that

$$\|\nabla g(u) - \nabla g(v)\| \leq \kappa(x)\|u - v\| \quad \forall u, v \in B(x, \varepsilon(x)). \tag{2}$$

On the other hand, according to the Lojasiewicz inequality, applied for the subanalytic function $-f$, by shrinking $\varepsilon(x)$ if necessary, for each $x \in \mathcal{C}$, we can find $L(x) > 0$; $\theta(x) \in [0, 1)$ such that

$$|f(u)|^{\theta(x)} \leq L(x)\|u^*\| \quad \text{for all } u \in B(x, \varepsilon(x));\ u^* \in \partial^F(-f)(u). \tag{3}$$

By the compactness of $\mathcal{C}$, there exist $z^1, z^2, ..., z^p \in \mathcal{C}$ such that $\mathcal{C} \subseteq \bigcup_{i=1}^p B(z^i, \varepsilon(z^i)/2)$. By relabeling if necessary, we can assume that

$$x^k \in \bigcup_{i=1}^p B(z^i, \varepsilon(z^i)/2) \quad \text{and} \quad \|x^k - x^{k+1}\| < \varepsilon/2,\ k = 0, 1, ...,$$

where $\varepsilon = \min\{\varepsilon(z^i)/2 :\ i = 1, ..., p\}$. Hence, by relations and, one obtaines

$$\|\nabla g(x^k) - \nabla g(x^{k+1})\| \leq \kappa\|x^k - x^{k+1}\|; \ \ |f(x^k)|^\theta \leq L\|x^{k*}\| \quad \text{for all } k = 0, 1, ...;\ x^{k*} \in \partial^F(-f)(x^k), \tag{4}$$

where, $\theta = \max\{\theta(z^i) :\ i = 1, ..., p\}$; $\kappa = \min\{\kappa(z^i) :\ i = 1, ..., p\}$; $L = \min\{L(z^i) :\ i = 1, ..., p\}$. Since g is differentiable, and by the definition of $\{x^k\}$ by (DCA), then

$$\nabla g(x^{k+1}) - \nabla g(x^k) = y^k - \nabla g(x^k) \in \partial h(x^k) - \nabla g(x^k) = \partial^F(-f)(x^k).$$

From relations in (4), we obtain

$$|f(x^k)|^\theta \leq \frac{L}{\kappa}\|x^k - x^{k+1}\| \quad \text{for all } k = 0, 1, ... \tag{5}$$

By using the concavity of the function $t \in \mathbb{R} \mapsto t^{1-\theta}$ on $(0, +\infty)$, one has

$$f(x^k)^{1-\theta} - f(x^{k+1})^{1-\theta} \geq (1-\theta)f(x^k)^{-\theta}(f(x^k) - f(x^{k+1})) \geq \frac{(1-\theta)L\rho}{\kappa}\|x^k - x^{k+1}\|. \tag{6}$$

It follows that

$$\sum_{k=1}^{\infty} \|x^k - x^{k+1}\| \leq \frac{\kappa}{L\rho(1-\theta)} f(x^0)^{1-\theta}.$$

Thus, the sequence $\{x^k\}$ converges to an unique limit point.

Case 2. h is differentiable and its derivative is locally Lipschitz. Similar to Case 1, but instead of using the Lojasiewicz inequality for the function $-f$, we apply this inequality to the function f, we can find $\kappa > 0$; $L > 0$ such that

$$\|\nabla h(x^k) - \nabla h(x^{k-1})\| \leq \kappa \|x^k - x^{k-1}\|; \ \ |f(x^k)|^\theta \leq L\|x^{k*}\| \quad \text{for all } k = 1, 2, ...; \ x^{k*} \in \partial^F f(x^k). \tag{7}$$

Since h is differentiable, and by the definition of $\{x^k\}$ by (DCA), then

$$\nabla h(x^{k-1}) - \nabla h(x^k) = y^{k-1} - \nabla h(x^k) \in \partial g(x^k) - \nabla h(x^k) = \partial^F f(x^k).$$

By (7), one has

$$|f(x^k)|^\theta \leq \frac{L}{\kappa}\|x^k - x^{k-1}\| \quad \text{for all } k = 1, 2, ... \tag{8}$$

Therefore,

$$f(x^k)^{1-\theta} - f(x^{k+1})^{1-\theta} \geq (1-\theta) f(x^k)^{-\theta}(f(x^k) - f(x^{k+1})) \geq \frac{(1-\theta)L\rho\|x^k - x^{k+1}\|^2}{\kappa\|x^k - x^{k-1}\|}.$$

Hence, by using the inequality $a \leq \frac{a^2}{b} + b/4$ for any $a, b > 0$, we derive that

$$\|x^k - x^{k+1}\| \leq \frac{\|x^k - x^{k+1}\|^2}{\|x^k - x^{k-1}\|} + \|x^k - x^{k-1}\|/4 \leq \|x^k - x^{k-1}\|/4 + \frac{\kappa}{L\rho(1-\theta)}\big(f(x^k)^{1-\theta} - f(x^{k+1})^{1-\theta}\big), \ k = 1, 2, ... \tag{9}$$

Consequently,

$$\sum_{k=1}^{\infty} \|x^k - x^{k+1}\| \leq \frac{\|x^1 - x^0\|}{3} + \frac{4\kappa}{3L\rho(1-\theta)} f(x^0)^{1-\theta}.$$

This implies that the sequence $\{x^k\}$ is convergent. Therefore, it follows obviously the convergence of $\{y^k\}$. The proof is completed. ■

Corollary 3.2 *Suppose that $g - h$ and $h^* - g^*$ are subanalytic functions with closed domain such that $(g - h)\,|_{\text{Dom}\,(g-h)}$ and $(h^* - g^*)\,|_{\text{Dom}\,(h^*-g^*)}$ are continuous. Assume that $\rho(g) + \rho(h) > 0$ as well as $\rho(g^*) + \rho(h^*) > 0$. If either the sequence $\{x^k\}$ or $\{y^k\}$ is bounded then these sequences converge to critical points of $(\mathcal{P})$ and $(\mathcal{D})$, respectively.*

Proof. It follows directly from Theorem 3.1 by the fact that the conjugate of a strongly convex function is differentiable with Lipschitz derivative. ■

The following theorem gives convergent rates of the sequence $\{x^k\}$.

Theorem 3.3 *Suppose that the assumptions of Theorem 3.1 (i) are satisfied. Let x^∞ be the limit point of $\{x^k\}$ with a Lojasiewicz exponent $\theta \in [0, 1)$. Then there exists constants $\tau_1, \tau_2 > 0$ such that*

$$\|x^k - x^\infty\| \leq \sum_{j=k}^{\infty} \|x^j - x^{j+1}\| \leq \tau_1 \|x^k - x^{k-1}\| + \tau_2 \|x^k - x^{k-1}\|^{\frac{1-\theta}{\theta}}, \ \ k = 1, 2, ... \tag{10}$$

As a result, one has

- *If $\theta \in (1/2, 1)$ then $\|x^k - x^\infty\| \leq ck^{\frac{1-\theta}{1-2\theta}}$ for some $c > 0$.*
- *If $\theta \in (0, 1/2]$ then $\|x^k - x^\infty\| \leq cq^k$ for some $c > 0; q \in (0, 1)$.*
- *If $\theta = 0$ then $\{x^k\}$ is convergent in a finite number of steps.*

Proof. The first inequality in (10) is obvious. Set $r_k = \sum_{j=k}^{\infty} \|x^j - x^{j+1}\|$, $k = 0, 1, ...$

If g is differentiable with Lipschitz derivative, then by (6) and (5), one has

$$r_k = \sum_{j=k}^{\infty} \|x^j - x^{j+1}\| \leq \frac{\kappa}{L\rho} f(x^k)^{1-\theta} \leq \frac{\kappa}{L\rho(1-\theta)} f(x^{k-1})^{1-\theta} \leq \frac{L^{\frac{1-2\theta}{\theta}}}{(1-\theta)\rho\kappa^{\frac{1-2\theta}{\theta}}} \|x^k - x^{k-1}\|^{\frac{1-\theta}{\theta}}.$$

Suppose now h is differentiable with Lipschitz derivative. By relations (8) and (9), one has

$$r_k \leq \frac{r_k + \|x^k - x^{k-1}\|}{4} + \frac{\kappa}{L\rho}\big(f(x^k)^{1-\theta} \leq \frac{r_k + \|x^k - x^{k-1}\|}{4} + \frac{L^{\frac{1-2\theta}{\theta}}}{(1-\theta)\rho\kappa^{\frac{1-2\theta}{\theta}}} \|x^k - x^{k-1}\|^{\frac{1-\theta}{\theta}}.$$

It follows that

$$r_k \leq \frac{\|x^k - x^{k-1}\|}{3} + \frac{4L^{\frac{1-2\theta}{\theta}}}{3(1-\theta)\rho\kappa^{\frac{1-2\theta}{\theta}}} \|x^k - x^{k-1}\|^{\frac{1-\theta}{\theta}},$$

and (10) is proved.

Since $\lim_{k\to\infty} \|x^k - x^{k+1}\| = 0$, then without loss of generality, we can assume that $\|x^k - x^{k+1}\| < 1$ for all $k = 0, 1, ...$ By setting $\tau = \tau_1 + \tau_2$ in (10), one has

- If $\theta \in (1/2, 1]$, then $r_k \leq \tau(r_{k-1} - r_k)^{\frac{1-\theta}{\theta}}$. Hence, by the concavity of the function $t \mapsto t^{\frac{1-2\theta}{1-\theta}}$ on $(0, +\infty)$,

  $$r_k^{\frac{1-2\theta}{1-\theta}} - r_{k-1}^{\frac{1-2\theta}{1-\theta}} \geq \frac{1-2\theta}{1-\theta} r_k^{\frac{-\theta}{1-\theta}} (r_k - r_{k-1}) \geq \frac{2\theta - 1}{1-\theta)\tau}.$$

  Thus,

  $$r_k^{\frac{1-2\theta}{1-\theta}} = r_0^{\frac{1-2\theta}{1-\theta}} + \sum_{j=1}^{k} \Big(r_k^{\frac{1-2\theta}{1-\theta}} - r_{k-1}^{\frac{1-2\theta}{1-\theta}} \Big) \geq r_0^{\frac{1-2\theta}{1-\theta}} + \frac{2\theta - 1}{1-\theta)\tau} k.$$

  Consequently, $r_k \leq ck^{\frac{1-\theta}{1-2\theta}}$, for some $c > 0$.

- If $\theta \in (0, 1/2]$, then $r_k \leq \tau(r_{k-1} - r_k)$, $k = 1.2, ...$ Therefore, $r_k \leq \frac{\tau}{\tau+1} r_{k-1}$. It follows that $r_k \leq cq^k$, where, $q = \tau/(\tau + 1)$; $c = r_0$.

- If $\theta = 0$ then $\|x_k - x_{k+1}\| \geq 1/L$ for $x^k \neq x^{k+1}$ with k sufficiently large. From the inequality

  $$f(x^k) - f(x^{k+1}) \geq \frac{\rho}{2} \|x^k - x^{k+1}\|^2 \geq \frac{\rho}{2L^2}$$

  for $x^k \neq x^{k+1}$, we derive that (DCA) terminates after a finite number of iteration steps.

■

In general, the Lojasiewicz exponent is unknown. But in several special cases, one can determine the Lojasiewicz exponent. Let us consider the case of trust-region subproblems, that is, the problems of minimizing a (nonconvex) quadratic function over an Euclidean ball in $\mathbb{R}^n$.

$$\min\{\frac{1}{2} x^T Q x + \langle b, x \rangle : \ \ x \in \mathbb{R}^n, \ \|x\| \leq R\},$$

or equivalently,

$$\min\{f(x) := \frac{1}{2} x^T Q x + \langle b, x \rangle + \chi_C(x) : \ \ x \in \mathbb{R}^n\}, \tag{11}$$

where, Q is a $n \times n$ real symmetric matrix, $b \in \mathbb{R}^n$, R is a positive scalar and $C := \{x \in \mathbb{R}^n : \ \|x\| \leq R\}$; $\chi_C(x)$ stands for the indicator function of the set C.

The following theorem gives a Lojasiewicz exponent of some critical point of the objective function f under a suitable assumption. First, we recall the following lemma in ([13], Lemma 4.1).

Lemma 3.4 *Let A be a $n \times n$ real symmetric matrix. Then there exists a constant $M > 0$ such that the following inequality holds*

$$|x^T Ax|^{1/2} \le M\|Ax\| \quad \text{for all } x \in \mathbb{R}^n.$$

Theorem 3.5 *Suppose $x^0 \in C$ is a critical point of problem (11). That is, there exists $\lambda_0 \ge 0$ such that*

$$Qx^0 + b + \lambda_0 x^0 = 0; \;\; \lambda_0(\|x^0\| - R) = 0. \tag{12}$$

In addition, assume that either $\|x^0\| < R$ or $\|x^0\| = R$ and

$$\langle v, Qv + \lambda_0 v\rangle > 0 \quad \text{for all } v \in \mathbb{R}^n \;\; \text{with } \langle x^0, v\rangle = 0, \; \|v\| = 1. \tag{13}$$

The Lojasiewicz exponent of the critical point x^0 is $1/2$.

Proof. Obviously, one has

$$\partial^F f(x) = \begin{cases} \{Qx + b + \lambda x : \; \lambda \ge 0\} & \text{if} \quad \|x\| = R \\ \{Qx + b\} & \text{if} \quad \|x\| < R. \end{cases} \tag{14}$$

Let $x^0 \in C$ be a critical point of the function f satisfying (12). If $\|x^0\| < R$ then $Qx^0 + b = 0$. According to Lemma 3.4, there exists $M > 0$ such that for all $x \in C$, one has

$$M\|Qx + b\| = M\|Q(x - x^0)\| \ge \frac{1}{2}|(x - x^0)^T Q(x - x^0)|^{1/2} = |f(x) - f(x^0)|^{1/2}.$$

Suppose now $\|x^0\| = R$, such that (13) is satisfied. Then, there exists $\lambda_0 \ge 0$ such that $Qx^0 + b + \lambda_0 x^0 = 0$. By again Lemma 3.4, we can find $M > 0$ such that for all $x \in C$, one has

$$M\|Qx + b + \lambda_0 x\| = M\|Q(x - x^0) + \lambda_0(x - x^0)\| \ge |(x - x^0)^T Q(x - x^0) + \lambda_0\|x - x^0\|^2|^{1/2}$$

On the other hand, by using the Taylor expansion,

$$f(x) - f(x^0) = \frac{1}{2}(x - x^0)^T Q(x - x^0) + \langle Qx^0 + b, x - x^0\rangle = \frac{1}{2}(x - x^0)^T Q(x - x^0) - \lambda_0\langle x^0, x - x^0\rangle.$$

Combining the later two relations, one obtains

$$M\|Qx + b + \lambda_0 x\| \ge |f(x) - f(x^0) + \lambda_0(\|x\|^2 - \|x^0\|^2)|^{1/2}. \tag{15}$$

Let $x \in C$. For $\|x\| < R$, if $\lambda^0 = 0$ then by (15),

$$M\|Qx + b\| \ge |f(x) - f(x^0)|^{1/2};$$

otherwise, $\lambda_0 > 0$, then for x sufficiently near x^0, one has

$$M\|Qx + b\| \ge M\lambda_0\|x\| - M\|Qx + b + \lambda_0 x\| \ge |f(x) - f(x^0)|^{1/2}.$$

Let us show that there exist $m, \varepsilon > 0$ such that for all $\lambda \in (\lambda_0 - \varepsilon, \lambda_0 + \varepsilon)$, $x \in B(x^0, \varepsilon)$ with $\|x\| = R$, one has

$$\|Qx + b + \lambda x\| \ge m\|Qx + b + \lambda_0 x\|. \tag{16}$$

Indeed, assume to contrary that there exist sequences $x^k \to x^0$ with $\|x^k\| = R$ and $\lambda_k \to \lambda_0$ such that

$$\lim_{k\to\infty} \frac{\|Qx^k + b + \lambda_k x^k\|}{\|Qx^k + b + \lambda_0 x^k\|} = 0.$$

By setting $u^k = x^k - x^0$, and without loss of generality, assume that $u^k/\|u^k\|$ converges to v, then $\|v\| = 1$; $\langle x^0, v\rangle = 0$, and one has

$$\lim_{k\to\infty} \frac{\|Qx^k + b + \lambda_k x^k\|}{\|Qx^k + b + \lambda_0 x^k\|} = \lim_{k\to\infty} \frac{\|Qu^k + \lambda_k u^k + (\lambda_k - \lambda_0)x^0\|}{\|Qu^k + \lambda_0 u^k\|} = \lim_{k\to\infty} \frac{\|Qv + \lambda_0 v + (\lambda_k - \lambda_0)x^0/\|u^k\|\|}{\|Qv + \lambda_0 v\|} = 0.$$

This implies that there exists $t \in \mathbb{R}$ such that $Qv + \lambda_0 v = tx^0$. therefore, $\langle v, Qv + \lambda_0 v\rangle = 0$, which contradicts (13).

From relation (15) and (16), one obtains, for some $L, \varepsilon > 0$,

$$L\|Qx + b + \lambda x\| \geq |f(x) - f(x^0)|^{1/2},\ \forall\lambda \in (\lambda_0 - \varepsilon, \lambda_0 + \varepsilon);\ \forall x \in B(x^0, \varepsilon),\ \|x\| = R. \tag{17}$$

Finally, for any $\lambda \in R$ with $|\lambda - \lambda_0| \geq \varepsilon$, for all x near x^0, one has

$$M\|Qx + b + \lambda x\| \geq M|\lambda - \lambda_0|\|x\| - M\|Qx + b + \lambda_0 x\| \geq |f(x) - f(x^0)|^{1/2}.$$

The proof is completed. ■

Remark 3.6 *Note that if the critical point x^0 satisfies (13) then x^0 is either a local-nonglobal minimizer or a global minimizer of problem (11) (see, for instance, [15]).*

References

[1] P.-A. Absil, R. Mahony, B. Andrews, Convergence of the iterates of descent methods for analytic cost functions, *SIAM J. Optim.,* **16**, 531-547 (2005).

[2] H. Attouch, J. Bolte, The convergence of the proximal algorithm for nonsmooth functions involving analytic features, *Math. Program.* **116**, 5-16 (2009).

[3] J. Bolte, A. Daniliidis, A. Lewis, Lojasiewicz inequality for nonsmooth subanalytic functions with applications to subgradient dynamic systems, *SIAM Optim.*, **17** (4), 1205-1223 (2007).

[4] J. Bolte, A. Daniliidis, O. Ley, L. Mazet, Characterizations of Lojasiewicz inequalities: Subgradient flows, talweg, convexity, *Trans. Amer. Math. Soc.*, **362**, 3319-3363 (2010).

[5] E. Bierstone, P. Milman, Semianalytic and subanalytic sets, *IHES Publ. Math.*, **67**, 5-42 (1988).

[6] J.B. Hiriart-Urruty, C. Lemaréchal, *Convex Analysis and Minimization Algorithms*, Parts I&II, Springer Verlag (1991).

[7] Le Thi Hoai An, Pham Dinh Tao, Large scale global molecular optimization from distance matrices by a DC optimization appoach, *SIAM J. Optim,* **14**(1), 77-116, (2003).

[8] Le Thi Hoai An and Pham Dinh Tao (2005), The DC (difference of convex functions) Programming and DCA revisited with DC models of real world nonconvex optimization problems. *Annals of Operations Research*, **133**, pp. 23-48.

[9] S. Lojasiewicz, Sur le problème de la division, *Studia Mathematica,* **18**, 87-136(1959).

[10] S. Lojasiewicz, Une propriété topologique des sous-ensembles analytiques réels, *in les Equations aux dérivées Partielles, Editions du Centre National de la Recherche Scientifique, Paris,* 87-89 (1963).

[11] S. Lojasiewicz, Sur la géométrie semi- et sous-analytique, *Anal de l'institute Fourier,* **43**, 1575-1595 (1993).

[12] Mordukhovich B.S, *Variational analysis and generalized differentiation. I. Basic theory.* Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences], 330. Springer-Verlag, Berlin, 2006.

[13] Ngai H. V., Théra M., Error bounds for systems of lower semicontinuous functions in Asplund spaces, *Math. Program.* , **116**, 397-427 (2009).

[14] Pham Dinh Tao, Le Thi Hoai An, Convex analysis approach to D.C. Programming: Theory, algorithms and applications, *Acta Mathematica Vietnamica,* **22**, 289-355(1997).

[15] Pham Dinh Tao, Le Thi Hoai An, A DC Optimization algorithm for solving the trust region subproblem, *SIAM J. Optim.,* **8**, 2, 476-505(1998).

[16] R.T. Rockafellar, *Convex Analysis,* Princeton University Press, 1970.

[17] R.T. Rockafellar, R. J-B. Wets, *Variational Analysis*, Springer, New York (1998).

[18] M. Shiota, *Geometry of subanalytic and semialgebraic sets,* Progress in Math., 150. Birkhauser Boston, Inc., Boston, MA (1997).