Optimization Methods and Software
Vol. 00, No. 00, Month 200x, 1–9

RESEARCH ARTICLE

Approximate optimality conditions and stopping criteria in canonical DC programming

Giancarlo Bigi, Antonio Frangioni* and Qinghua Zhang

Dipartimento di Informatica, Università di Pisa,
Largo B.Pontecorvo, 3, 56127 Pisa, Italia

(*received December 2008*)

In this paper we study approximate optimality conditions for the Canonical DC (CDC) optimization problem and their relationships with stopping criteria for a large class of solution algorithms for the problem. In fact, global optimality conditions for CDC are very often restated in terms of a nonconvex optimization problem, that has to be solved each time the optimality of a given tentative solution has to be checked. Since this is in principle a costly task, it makes sense to only solve the problem approximately, leading to an inexact stopping criteria and therefore to approximate optimality conditions. In this framework, it is important to study the relationships between the approximation in the stopping criteria and the quality of the solutions that the corresponding approximated optimality conditions may eventually accept as optimal, in order to ensure that a small tolerance in the stopping criteria does not lead to a disproportionally large approximation of the optimal value of the CDC problem. We develop conditions ensuring that this is the case; these turn out to be closely related with the well-known concept of *regularity* of a CDC problem, actually coinciding with the latter if the reverse-constraint set is a polyhedron.

Keywords: DC problems; approximate optimality conditions; value function; Lipschitz property.

AMS Subject Classification: 90C26, 90C46.

1. Introduction

It is well-known that any DC optimization problem, that is a nonconvex program where the objective function is the difference of two convex functions and the constraint can be expressed as the set difference of two convex sets, can be reduced via standard transformations [13] to the *Canonical DC* problem

$$(CDC) \qquad \min\{\ dx \ \mid\ x \in \Omega \setminus \operatorname{int} C\ \}$$

where $\Omega \subseteq \mathbb{R}^n$ and $C \subseteq \mathbb{R}^n$ are full-dimensional closed convex sets, $d \in \mathbb{R}^n$ and dx denotes the scalar product between d and the vector of variables $x \in \mathbb{R}^n$. Notice that full-dimensionality here can be assumed without loss of generality: if Ω is not full-dimensional then the problem can be reformulated in the (affine) space generated by Ω, and if C is not full-dimensional then $\operatorname{int} C = \emptyset$ and (CDC) actually reduces

The work has been partially supported by the grant PRIN 2007 9PLLN7 “Nonlinear Optimization, Variational Inequalities and Equilibrium Problems”

*Corresponding author. Email: frangio@di.unipi.it

ISSN: 1055-6788 print/ISSN 1029-4937 online

DOI: 10.1080/03081080xxxxxxxxx
http://www.informaworld.com

to a convex program. For the same reason, we assume that the set C provides an essential constraint, i.e.,

$$\min\{\ dx \mid x \in \Omega\ \} < \min\{\ dx \mid x \in \Omega \setminus \operatorname{int} C\ \}.$$

Relying on the appropriate translation, this assumption can be equivalently stated through the following two conditions

$$0 \in \Omega \cap \operatorname{int} C, \tag{1}$$

$$dx > 0 \quad \forall x \in \Omega \setminus \operatorname{int} C. \tag{2}$$

Due to its status of canonical form for a very large class of difficult optimization problems, several algorithms have been proposed to solve (CDC) [3, 5, 7, 8, 12–19], though many of them are actually variants of the cutting plane algorithm proposed by Tuy in [13]. All these algorithms are based on (approximately) checking the well-known *necessary* global optimality condition

$$D(\gamma^*) \subseteq C \tag{3}$$

where

$$D(\gamma) := \{\ x \in \Omega \mid dx \leq \gamma\ \}$$

and γ^* is the optimal value of (CDC). To simplify the treatment, in the following we assume that optimal solutions do indeed exist; this can be easily obtained by requiring boundedness of $D(\gamma)$ for the feasible values γ, i.e., those values $\gamma = dx$ for some $x \in \Omega \setminus \operatorname{int} C$, which is the case in particular if Ω itself is bounded. It is also worth mentioning at this point that (3) is also a *sufficient* global optimality condition when problem (CDC) is *regular* [17, Proposition 10], i.e.,

$$\min\{\ dx \mid x \in \Omega \setminus \operatorname{int} C\ \} = \inf\{\ dx \mid x \in \Omega \setminus C\ \}. \tag{4}$$

Therefore, (4) is a crucial condition for all the above mentioned algorithm to converge to a global optimal solution.

In order to deal with the nonconvex constraint $x \notin \operatorname{int}\ C$, the set C has often been represented as the zero sublevel set of a real-valued convex function, i.e.,

$$C = \{\ x \in \mathbb{R}^n \mid h(x) \leq 0\ \}$$

where $h : \mathbb{R}^n \to \mathbb{R}$ is a convex function such that $h(0) < 0$. This allows to express the nonconvex constraint as the reverse convex inequality

$$x \notin \operatorname{int}\ C \iff h(x) \geq 0.$$

Furthermore, this choice provides an "optimization version" of the optimality conditions. In fact, it is easy to check that (3) is equivalent to the condition

$$\max\{\ h(x) \mid x \in D(\gamma^*)\ \} \leq 0. \tag{5}$$

In turn, (5) has been exploited to propose concepts of *approximate optimality*, for instance in [7, 13, 14, 20], and especially to allow solution algorithms to rely on *approximate stopping criteria*, generally of the form

$$\max\{\ h(x)\ \mid\ x \in D(\gamma)\ \} \leq \varepsilon, \tag{6}$$

or, equivalently,

$$D(\gamma) \subseteq \{\ x \in \mathbb{R}^n\ \mid\ h(x) \leq \varepsilon\ \},$$

where $\varepsilon > 0$ is some given "small" tolerance. In this context, it is very important to ensure that small values of ε lead to computing a solution whose value is "not too far" from the optimal one. While it can be easily shown (cf. Proposition 3.2) that the limit of (6) for $\varepsilon \to 0$ actually gives (5) as expected, the exact relationships between a feasible value γ satisfying (6), the optimal value γ^* and the tolerance ε have not been studied in details.

It is clear that condition (6) largely depends upon the choice of h. For instance, though h and αh provide the same set C for any $\alpha > 0$, they may obviously lead to extremely different feasible values satisfying (6) for the same fixed tolerance ε.

In this paper we study approximate optimality conditions for (CDC), relying on an alternative equivalent formulation which is based on a polar characterization of the nonconvex constraint. This formulation allows to express the optimality conditions in a geometric fashion such that their "optimization version", which has already been exploited in the algorithmic schemes of [3, 15, 19], does not involve any representation of the set C through a convex function h. In Section 2 we discuss the polar formulation and the corresponding optimality conditions and we introduce the related approximate version that can be exploited as a stopping criterion in algorithms. In Section 3 we provide alternative characterizations of the approximate optimality conditions and we show that the relationship between the approximation error of the optimal value and the tolerance is linear under some reasonable (geometric) assumptions on the data of the problem.

2. Polar Formulation and Optimality Conditions

The constraint $x \notin \text{int}\ C$ is the source of nonconvexity in problem (CDC) and it is given just as a set relation. However, relying on the polarity between convex sets, we can express this nonconvex constraint in a different fashion. Let us recall that

$$C^* = \{\ w \in \mathbb{R}^n\ \mid\ wx \leq 1 \quad \forall x \in C\ \}$$

is the polar set of C and it is a closed convex set. Exploiting bipolarity relations (see, for instance, [11]), it is easy to check that the assumption $0 \in \text{int}\ C$ ensures that $x \notin \text{int}\ C$ if and only if $wx \geq 1$ for some $w \in C^*$. Therefore, problem (CDC) can be equivalently formulated as

$$\min\{\ dx\ \mid\ x \in \Omega\ ,\ w \in C^*\ ,\ wx \geq 1\ \} \tag{7}$$

where polar variables w have been introduced and the nonconvexity is given by the inequality constraint, which asks for some sort of reverse polar condition. Also, the assumption $0 \in \text{int}\ C$ ensures the compactness of C^* (see, for instance, [11, Corollary 14.5.1]). Again by bipolarity relations, the optimality condition (3) can be equivalently stated in a polar fashion as

$$D(\gamma) \times C^* \subseteq \{\ (x, w) \in \mathbb{R}^n \times \mathbb{R}^n\ \mid\ wx \leq 1\ \} \tag{8}$$

while the regularity condition (4) reads

$$\min\{\ dx \mid x \in \Omega,\ w \in C^*,\ wx \geq 1\ \} = \inf\{\ dx \mid x \in \Omega,\ w \in C^*,\ wx > 1\ \}. \quad (9)$$

As an immediate consequence of (8), any optimal solution (x^*, w^*) to (7) satisfies both $x^* \in \partial C$ and $w^* x^* = 1$, a fact that will be useful in the following.

Inclusion (8) leads to the following alternative "optimization version" of the optimality conditions:

$$\max\{\ vz - 1 \ \mid\ z \in D(\gamma)\ ,\ v \in C^*\ \}. \quad (10)$$

Obviously, (8) holds if and only if the optimal value of (10), which we will denote by $v(OC_\gamma)$, is nonpositive. While (10) provides a functional alternative to (5), the two problems are by no means equivalent. In particular, it is easy to see that (10) is completely independent of the choice of h. As a consequence, the objective function of (10), while clearly nonconvex, is "simple" and independent of the data of the instance at hand. On the other hand, (5) is defined on a smaller space. The different structure of the two problems may motivate different algorithms for their solution; for instance, a *lower bound* on $v(OC_\gamma)$ may be found by means of *alternating minimization* methods, whereby one of the two variables is kept fixed and a linear maximization problem is solved to optimize on the other, and then the role is reversed. Iterating this procedure provably leads to a local optima of the problem [6], and this approach has been experimentally proven to be remarkably effective in several fields, such as machine learning [4] and image processing [21]; the ability of efficiently computing good lower bounds may prove useful for some solution approaches to (CDC) [3]. However, for the purpose of testing optimality conditions *upper bounds* on $v(OC_\gamma)$ are required; these can be produced by solving suitable *relaxations* of (10), i.e., by replacing the non-concave objective function vz with a suitable concave upper approximation. In particular, one may use well-known results that characterize the *concave envelope* (lower concave approximation) of the function, which happens to be polyhedral in this particular case [9]. Upper and lower bounding techniques can then be combined in *exact algorithms* that can be used to obtain arbitrarily accurate upper bounds, albeit possibly at the cost of enumerative procedures [1, 2, 10]. Thus, while algorithms exist which can computationally prove or disprove that the *approximate stopping criterion*

$$v(OC_\gamma) \leq \varepsilon \quad (11)$$

holds for a given $\varepsilon > 0$, this may prove exceedingly expensive if ε is "very small". Thus, one could in principle want a "relatively large" ε; the drawback is that a feasible value γ needn't be optimal when (11) holds, and the convergence of γ (as a function of ε) to γ^* may be slow (see Example 3.3 below). The next section is therefore devoted to the study of the relationships between the "quality" of γ and the tolerance ε, with the aim to identifying conditions which ensure that the rate of convergence is at worst linear.

3. Approximate Optimality Conditions

The main tool for studying the impact of ε on the accuracy of the values γ for which (11) holds is the approximated problem

$$\min\{\ dx \ \mid \ x \in \Omega\ ,\ w \in C^*\ ,\ wx \geq 1+\varepsilon\ \}, \tag{12}$$

which is obtained by perturbing the right-hand side of the nonconvex constraint in (7). Our analysis does not require any regularity assumption on (12) and it is based on the following quantity

$$\phi(\varepsilon) := \inf\{\ dx \ \mid \ x \in \Omega\ ,\ w \in C^*\ ,\ wx > 1+\varepsilon\ \}.$$

Obviously, $\phi(\varepsilon)$ may be greater than the optimal value of (12); anyway, the value function ϕ provides the right tool to disclose the connections between γ, (11) and (12).

Proposition 3.1: *Let $\varepsilon \geq 0$. Then, the following statements are equivalent:*

(i) $v(OC_\gamma) \leq \varepsilon$;
(ii) $D(\gamma) \times C^* \subseteq \{\ (x,w) \in \mathbb{R}^n \times \mathbb{R}^n \ \mid \ wx \leq 1+\varepsilon\ \}$;
(iii) $\gamma \leq \phi(\varepsilon)$.

Proof: The equivalence between *(i)* and *(ii)* follows immediately from the definition of $v(OC_\gamma)$. Analogously, *(ii)* implies *(iii)* by the definition of $\phi(\varepsilon)$.

Suppose *(ii)* does not hold: there exist $x \in D(\gamma)$ and $w \in C^*$ such that $wx > 1+\varepsilon$. Take any $t \in (0,1)$ large enough to have $w(tx) > 1+\varepsilon$. Since $0 \in \Omega$, the convexity of Ω implies $tx \in \Omega$; obviously $d(tx) < dx \leq \gamma$. Therefore, (tx, w) guarantees $\phi(\varepsilon) < \gamma$ contradicting *(iii)*. □

Considering the optimal value of (12) as γ in Proposition 3.1, we get that *(ii)* is a necessary optimality condition for (12). Furthermore, if the problem is regular (i.e. $\phi(\varepsilon)$ is actually the optimal value), it is also sufficient. Choosing $\varepsilon = 0$, the known optimality conditions for (7) follow too. Therefore, inclusion *(ii)* can be considered as an approximate optimality condition for (7). It is easy to check that *(ii)* is equivalent to the inclusion $D(\gamma) \subseteq (1+\varepsilon)C$: perturbing the right-hand side of the nonconvex constraint in (7) corresponds to perturbing the reverse constraining set C in (CDC). As an immediate consequence of the proposition, we also have

$$\phi(\varepsilon) = \sup\{\ \gamma \ \mid \ D(\gamma) \times C^* \subseteq \{\ (x,w) \in \mathbb{R}^n \times \mathbb{R}^n \ \mid \ wx \leq 1+\varepsilon\ \}\ \}.$$

The stopping criterion *(i)* guarantees approximate optimality and condition *(iii)* provides the adequate tool to evaluate the quality of the approximation. In fact, supposing (7) to be regular, i.e., $\gamma^* = \phi(0)$, we have that

$$0 \leq \gamma - \gamma^* \leq \phi(\varepsilon) - \gamma^* = \phi(\varepsilon) - \phi(0)$$

holds for any feasible value γ which satisfies *(i)*. The following result guarantees that the approximation approaches the optimal value as ε goes to 0.

Proposition 3.2: *The value function ϕ is right-continuous at 0, i.e.*

$$\lim_{\varepsilon \downarrow 0} \phi(\varepsilon) = \phi(0).$$

Proof: Clearly ϕ is nondecreasing, that is $\phi(\varepsilon^1) \geq \phi(\varepsilon^2)$ whenever $\varepsilon^1 \geq \varepsilon^2 \geq 0$. As it is also bounded below by $\phi(0)$, there exist $\bar{\gamma} = \lim_{\varepsilon \downarrow 0} \phi(\varepsilon)$ and $\bar{\gamma} \geq \phi(0)$. Since $\bar{\gamma} \leq \phi(\varepsilon)$ for any $\varepsilon > 0$, Proposition 3.1 implies $v(OC^{\varepsilon}_{\bar{\gamma}}) \leq \varepsilon$ for any $\varepsilon > 0$. Since $v(OC_{\bar{\gamma}})$ does not depend upon ε, we get $v(OC_{\bar{\gamma}}) \leq 0$. Therefore, Proposition 3.1 guarantees $\bar{\gamma} \leq \phi(0)$. □

Although the approximation always converges to the optimal value, the rate of convergence may be less than linear as the following example shows.

Example 3.3 Consider (12) with $n = 2$, $d = (-1, 2)$, the convex constraint $\Omega = \operatorname{conv}\{(-2, 0), (1, -1), (-2, 3)\}$ and $C = \{ x \in \mathbb{R}^2 \mid x_1^2 + x_2^2 \leq 4 \}$. It is easy to check that (12) is regular for any $\varepsilon \geq 0$ and that

$$(x^*, w^*) = ((-2, 2\sqrt{(1+\varepsilon)^2 - 1}), (-1/2(1+\varepsilon), -2\sqrt{(1+\varepsilon)^2 - 1}/(1+\varepsilon)))$$

is an optimal solution to (12) for $\varepsilon \leq 1/4$. Therefore, we have

$$\phi(\varepsilon) = 4\sqrt{(1+\varepsilon)^2 - 1} + 2$$

and

$$\lim_{\varepsilon \downarrow 0} [\phi(\varepsilon) - \phi(0)]/\varepsilon = \lim_{\varepsilon \downarrow 0} 4\sqrt{1 + 2/\varepsilon} = +\infty.$$

Thus, regularity is not enough to achieve a linear rate of convergence. Additional assumptions on the problem are needed: the existence of an optimal solution with some particular properties guarantees the Lipschitz behaviour of ϕ.

Proposition 3.4: *If there exists an optimal solution (x^*, w^*) to (7) such that*

$$\{ x^* + \lambda u \mid \lambda > 0 \} \cap \Omega \neq \emptyset \quad \text{and} \quad w^* u > 0 \tag{13}$$

for some direction $u \in \mathbb{R}^n$, then the value function ϕ is locally Lipschitz at 0, i.e. there exist $L > 0$ and $\bar{\varepsilon} > 0$ such that

$$\phi(\varepsilon) - \phi(0) \leq L\varepsilon \qquad \forall\ \varepsilon \in [0, \bar{\varepsilon}].$$

Proof: Let $\bar{\lambda} > 0$ be such that $x^* + \bar{\lambda} u \in \Omega$; the convexity of Ω implies $x(\lambda) := x^* + \lambda u \in \Omega$ for any $\lambda \in [0, \bar{\lambda}]$; furthermore, $w^*(x^* + \lambda u) = 1 + \lambda w^* u > 1$ if $\lambda > 0$. Thus, the sequence $(x(\lambda), w^*)$ shows that the regularity condition (9) holds. Therefore, we have $\phi(0) = dx^*$.

Chosen $\bar{\varepsilon} := (w^* u/2)\bar{\lambda}$, let us consider $y(\varepsilon) := x^* + (2\varepsilon/w^* u)u$ for any $\varepsilon \in (0, \bar{\varepsilon}]$: we have $y(\varepsilon) \in \Omega$ and

$$w^* y(\varepsilon) = w^* x^* + (2\varepsilon/w^* u) w^* u = 1 + 2\varepsilon > 1 + \varepsilon,$$

where the last equality holds since optimality implies $w^* x^* = 1$. Therefore, $(y(\varepsilon), w^*)$ provides an upper bound for $\phi(\varepsilon)$, i.e. $\phi(\varepsilon) \leq dy(\varepsilon)$. Finally, we get

$$\phi(\varepsilon) - \phi(0) \leq dy(\varepsilon) - dx^* = (2du/w^* u)\varepsilon. \qquad □$$

Though regularity has not been explicitly required for (7), the assumption on the optimal solution implies it. A geometric view of this assumption can be achieved relying on the (Bouligand) tangent cone of C at x^*, namely the set

$$T(C,x) := \{\ u \in \mathbb{R}^n \ \mid\ \exists t_n \downarrow 0,\ u_n \to u \ \text{ s.t. } x + t_n u_n \in C\ \},$$

and its following characterization.

Lemma 3.5: *Let $x^* \in \partial C$. Then, the following statements are equivalent:*

(*i*) $u \in T(C,x^*)$;
(*ii*) $wu \leq 0$ *for all* $w \in C^*$ *such that* $wx^* = 1$.

Proof: Take $u \in T(C,x^*)$: there exist $t_n \downarrow 0$ and $u_n \to u$ such that $x^* + t_n u_n \in C$. Therefore, we have $w(x^* + t_n u_n) \leq 1$ for any $w \in C^*$. If $wx^* = 1$, we get $wu_n \leq 0$ and taking the limit $wu \leq 0$.

Vice versa, suppose u satisfies (*ii*) but $u \notin T(C,x^*)$. Since the tangent cone is a closed set, there exists $\varepsilon > 0$ such that $\hat{u} = u - \varepsilon x^* \notin T(C,x^*)$. Consider any $t_n \downarrow 0$ and $u_n \to \hat{u}$ such that $x^* + t_n u_n \notin C$. Therefore, there exist $w_n \in C^*$ such that $w_n(x^* + t_n u_n) > 1$. Assumption (1) implies that C^* is compact. Thus, we can suppose $w_n \to \bar{w}$ for some $\bar{w} \in C^*$. Taking the limit in the above inequality, we get $\bar{w}x^* \geq 1$ and therefore $\bar{w}x^* = 1$. Since $t_n w_n u_n > 1 - w_n x^* \geq 0$, we also get $\bar{w}\hat{u} \geq 0$. The assumption on u guarantees also $\bar{w}u \leq 0$. Therefore, we get the contradiction $0 \leq \bar{w}\hat{u} = \bar{w}(u - \varepsilon x^*) \leq -\varepsilon$. □

The following characterization allows to formulate the assumption of Proposition 3.4 in a geometric fashion.

Proposition 3.6: *Let $x^* \in \partial C$. Then, the following statements are equivalent:*

(*i*) *there exist* $w^* \in C^*$ *and* $u \in \mathbb{R}^n$ *such that* $w^*x^* = 1$ *and (13) holds;*
(*ii*) $T(\Omega,x^*) \nsubseteq T(C,x^*)$.

Proof: Suppose (*ii*) does not hold and take any $w^* \in C^*$ and $u \in \mathbb{R}^n$ such that $w^*x^* = 1$ and $x^* + \bar{\lambda}u \in \Omega$ for some $\bar{\lambda} > 0$. The convexity of Ω implies $\Omega \subseteq x^* + T(\Omega,x^*)$ and therefore $\bar{\lambda}u \in T(\Omega,x^*) \subseteq T(C,x^*)$. By Lemma 3.5 we get $w^*u \leq 0$: hence (*i*) does not hold.

Vice versa, take any $u \in T(\Omega,x^*) \setminus T(C,x^*)$. Lemma 3.5 implies that there exists $w^* \in C^*$ such that $w^*x^* = 1$ and $w^*u > 0$. As $u \in T(\Omega,x^*)$, there exist $t_n \downarrow 0$ and $u_n \to u$ such that $x^* + t_n u_n \in \Omega$; if n is large enough, we also have $w^*u_n > 0$. Thus, w^* and u_n satisfy (13). □

It is worth to note that (*ii*) depends upon x^* only. Indeed, the original formulation of the canonical DC problem does not have polar variables. Anyway, x^* is an optimal solution to (CDC) if and only if (x^*, w^*) is an optimal solution to (7) for any $w^* \in C^*$ such that $w^*x^* = 1$. As a consequence, Propositions 3.4 and 3.6 lead to the main result of the section.

Theorem 3.7: *If there exists an optimal solution (x^*, w^*) to (7) such that $T(\Omega,x^*) \nsubseteq T(C,x^*)$, then ϕ is locally Lipschitz at 0.*

An illustration of Theorem 3.7 is provided in Figure 3.3, which depicts the problem of Example 3.3: ϕ is not locally Lipschitz, in fact $T(\Omega,x^*) \subseteq T(C,x^*)$.

The assumption on the tangent cones can be considered as a *strong regularity* condition. In fact, it implies regularity but they are not equivalent, as the problem of Example 3.3 shows for $\varepsilon = 0$. Anyway, when C is a polyhedron, strong regularity collapses to regularity.

Theorem 3.8: *If C is a polyhedron, then (7) is regular if and only if there exists an optimal solution (x^*, w^*) to (7) such that $T(\Omega,x^*) \nsubseteq T(C,x^*)$.*

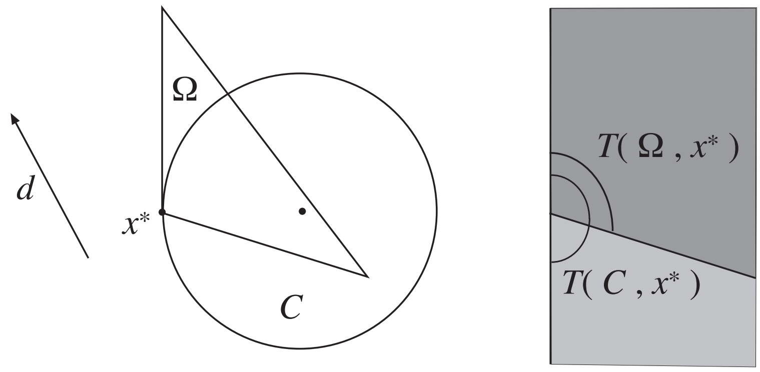

Figure 1. Optimal solution and corresponding tangent cones in Example 3.3

Proof: Suppose (7) is regular: there exists a sequence $\{(x^k, w^k)\} \subseteq \Omega \times C^*$ such that $w^k x^k > 1$ and $dx^k \downarrow \gamma^*$. Since $D(dx^1)$ is compact by assumption and (1) implies the compactness of C^*, the sequence $\{(x^k, w^k)\}$ admits at least one cluster point (x^*, w^*). Since $w^k x^k > 1$, we have $x^k \notin C$ and therefore $x^* \in \mathrm{cl}\,(\Omega \setminus C)$. Since Ω and C^* are closed, we have $x^* \in \Omega$ and $w^* \in C^*$. We also have $w^* x^* \geq 1$: this means that (x^*, w^*) is feasible for (7) and hence optimal as $dx^* = \gamma^*$. Optimality guarantees $x^* \in \partial C$ and therefore $x^* \notin \Omega \setminus C$. As a consequence, we have $x^* \in \mathrm{bd}\,(\Omega \setminus C)$. Now, suppose also $T(\Omega, x^*) \subseteq T(C, x^*)$. Since C is a polyhedron, there exists $\varepsilon > 0$ such that

$$[x^* + T(C, x^*)] \cap B(x^*, \varepsilon) = C \cap B(x^*, \varepsilon).$$

Since the convexity of Ω implies $\Omega \subseteq x^* + T(\Omega, x^*)$, we have

$$\Omega \cap B(x^*, \varepsilon) \subseteq C \cap B(x^*, \varepsilon)$$

in contradiction with $x^* \in \mathrm{bd}\,(\Omega \setminus C)$.

The if part follows from Proposition 3.6 and the proof of Proposition 3.4. □

Corollary 3.9: *Suppose C is a polyhedron. If (7) is regular, then ϕ is locally Lipschitz at 0.*

The corollary shows that for the large class of (CDC) problems where C is a polyhedron, the regularity condition (4), which is needed for algorithms to converge to a global optimal solution, is also sufficient to ensure that ϕ is locally Lipschitz.

4. Conclusions

The results of this paper show that a relatively simple geometric condition, strongly related with that of regularity of a canonical DC program and coinciding with the latter if the reverse convex constraint is polyhedral, guarantees a linear relationship between the approximation ε in the optimality conditions and that of the attained feasible value. This opens the interesting question about how to algorithmically compute an estimate of the Lipschitz constant for the value function ϕ, a fundamental step for being able to actually tune ε in numerical algorithm as a function of the desired final accuracy. We intend to pursue this subject in future research.

References

[1] F. Al-Khayyal and H. Sherali, *On finitely terminating branch-and-bound algorithms for some global optimization problems*, SIAM J. Optimization 10 (2000), pp. 1049–1057.

[2] I. Androulakis, C. Maranas, and C. Floudas, *α–BB: A global optimization method for general constrained nonconvex problems*, J. Global Optim. 7 (1995), pp. 337–363.

[3] G. Bigi, A. Frangioni and Q. Zhang, *Outer approximation algorithms for canonical DC problems*, J. Global Optim. DOI 10.1007/s10898-009-9415-1

[4] O. Chapelle, V. Sindhwani and S.S. Keerthi, *Optimization techniques for semi-supervised support vector machines*, J. Mach. Learn. Res. 9 (2008), pp. 203–233.

[5] J. Fulop, *A finite cutting plane method for solving linear programs with an additional reverse constraint*, European J. Oper. Res. 44 (1990), pp. 395–409.

[6] L. Grippo and M. Sciandrone, *On the convergence of the block nonlinear Gauss–Seidel method under convex constraints*, Oper. Res. Lett. 26 (2000), pp. 127–136.

[7] M.D. Nghia and N.D. Hieu, *A method for solving reverse convex programming problems*, Acta Math. Vietnam. 11 (1986), pp. 241–252.

[8] D.T. Pham and S. El Bernoussi, *Numerical methods for solving a class of global nonconvex optimization problems*, International Series of Numerical Mathematics 87 (1989), pp. 97–132.

[9] A.D. Rikun, *A convex envelope formula for multilinear functions*, J. Global Optim. 10 (1997), pp. 425–437.

[10] H. Ryoo and N. Sahinidis, *Global optimization of multiplicative programs*, J. Global Optim. 26 (2003), pp. 387–418.

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

[12] P.T. Thach, *Convex programs with several additional reverse convex constraints*, Acta Math. Vietnam. 10 (1985), pp. 35–57.

[13] H. Tuy, *A general deterministic approach to global optimization via d.c. programming*, in *FERMAT Days 85: Mathematics for Optimization*, J.B. Hiriart-Urruty, ed., North-Holland, Amsterdam, 1986, pp. 273–303.

[14] H. Tuy, *Convex programs with an additional reverse convex constraint*, J. Optim. Theory Appl. 52 (1987), pp. 463–486.

[15] H. Tuy, *On nonconvex optimization problems with separated nonconvex variables*, J. Global Optim. 2 (1992), pp. 133–144.

[16] H. Tuy, *Canonical DC programming problem: outer approximation methods revisited*, Oper. Res. Lett. 18 (1995), pp. 99–106.

[17] H. Tuy, *D.C. optimization: theory, methods and algorithms*, in *Handbook of Global Optimization*, R. Horst and P.M. Pardalos, eds., Kluwer Academic Publishers, Dordrecht, 1995, pp. 149–216.

[18] H. Tuy, *Convex analysis and global optimization*, Kluwer Academic Publishers, Dordrecht, 1998.

[19] H. Tuy and B.T. Tam, *Polyhedral annexation vs outer approximation for the decomposition of monotonic quasiconcave minimization problems*, Acta Math. Vietnam. 20 (1995), pp. 99–114.

[20] H. Tuy, *Robust solution of nonconvex global optimization problems*, J. Global Optim. 32 (2005), pp. 307–323.

[21] Y.-W. Wen, M.K. Ng and Y.-M. Huang, *Efficient total variation minimization methods for color image restoration*, IEEE Trans. Image Process. 17 (2008), pp. 2081–2088.