OPTIMIZATION ERUDITORUM
Volume 1 (2024), No. 2, 121–127
https://doi.org/10.69829/oper-024-0102-ta04
Tulipa Opera Scholarum

ABOUT DC AND SPLIT COMPOSITE MINIMIZATION PROBLEMS

A. MOUDAFI*

Aix-Marseille Université, Laboratoire d'Informatique et Systèmes
(LIS UMR 7020 CNRS / AMU /UTLN) Marseille, France

Abstract. The aim of this paper is twofold. Firstly, to prove the linear convergence of a linearized proximal algorithm for solving DC composite optimization problems under suitable assumptions. In the DC optimization setting, this sharpens recent results. Secondly, to sketch out some ideas for solving split convex composite minimization problems by means of an alternative formulation based on a generalization of an infimal post-composition approach developed very recently. This will most probably rely on a conjugacy and subdifferential calculus for convex convex-composite functions in finite-dimensional space even in infinite dimensional spaces together very possibly with new notions. The goal is to move from an approach based on the composition of a convex function by a linear operator to a composition by a suitable map.

Keywords. Composite function, Proximal linearized algorithm, DC function, DCA linearized algorithm, critical point, infimal post-composition.

1. Introduction and Preliminaries

We are interested first in the following classe of composite minimization problems

$$\min_{x\in I\!R^d} \varphi(x) := f(x) - g(c(x)), \tag{1.1}$$

where $f : I\!R^d \to I\!R \cup \{+\infty\}$, $g : I\!R^m \to I\!R \cup \{+\infty\}$ are closed convex functions and $c : I\!R^d \to I\!R^m$ is a smooth map.

Regularized nonlinear least squares and exact penalty formulations of nonlinear programs are classical examples, while notable contemporary instances include robust phase retrieval and matrix factorization problems. The setting where c is the identity function, (1.1) reduces to

$$\min_{x\in I\!R^d} \varphi(x) := f(x) - g(x), \tag{1.2}$$

which is nothing else than DC minimization (minimization of a difference of two convex functions that use convex properties of the two convex functions separately) which is now common place in large-scale optimization. The DCA investigated by Tao [16], is a popular algorithm for DC minimization and a proximal point algorithm was proposed in [15]. Moudafi and Maingé [10] proposed an alternative proof of the main result in [15] with some extensions and a proximal linearized algorithm for minimizing DC functions was then proposed in [14]. DC optimization algorithms have been proved to be particularly successful for analyzing and solving a variety of highly structured and practical problems.

Our goal in [9] was centered around prox-linear methods and shares the same idea, namely, the linearization the component $g(\cdot)$ or $c(\cdot)$; or both, which we proposed to extend to the entire problem

*Corresponding author.
E-mail address: abdellatif.moudafi@univ-amu.fr
2020 Mathematics Subject Classification: 65K10, 90C26.
Accepted: August 22, 2024.

classes (1.1). We proved that every cluster point of the sequence generated by our method for minimizing composite DC functions is a critical point. In each iteration of (1.3), the prox-linear method solves the following proximal subproblem: given an initial point $x_0 \in dom\varphi$ and a sequence of positive parameters $(\lambda_k)_{k\in I\!N}$ such that $\liminf_{k\to+\infty} \lambda_k > 0$. calculate $(w_k)_{k\in I\!N}$ and compute $(x_k)_{k\in I\!N}$ by

$$\begin{cases} w_k \in \nabla c(x_k)^* \partial g(c(x_k)); \\ x_{k+1} = \arg\min_{x\in I\!R^d} \left(f(x) - \langle w_k, x - x_k\rangle + \frac{1}{2\lambda_k}\|x - x_k\|^2\right). \end{cases} \tag{1.3}$$

If $x_{k+1} = x_k$, stop. Otherwise, set $k := k+1$ and return the first step.

Note that if c is the identity function, then Algorithm (1.3) becomes exactly the linearized proximal point algorithm introduced in [14] for DC functions. The setting where g is the null function, Algorithm (1.3) is nothing else than the classical proximal point algorithm. The underlying assumption here is that the proximal subproblems can be solved efficiently.

In [9], we proved that the proposed algorithm provides a descent method and that every cluster point is a critical point of the functions under consideration provided that the boundedness of the sequences generated by the algorithm, see [9]-Theorem 3.3. In section 2, we will complete the convergence analysis by providing a linear convergence result under suitable assumptions.

On the other hand, remember that convex-composite optimization is a class of nonsmooth, nonconvex optimization problems that captures a wide variety of optimization models studied in modern optimization practice and theory including nonlinear programming, nonconvex minimax problems, nonconvex system identification, inverse problems and nonlinear filtering as well as most nonconvex problems in large-scale data analysis and machine learning, see [6], [7]. That is why in ection 3, we are interested in the following classe of composite optimization problems

$$\min_{x\in I\!R^d} \psi(x) := f(x) + g(c(x)), \tag{1.4}$$

where $g : I\!R^m \to I\!R \cup \{+\infty\}$ and $f : I\!R^d \to I\!R \cup \{+\infty\}$ are closed convex functions and $c : I\!R^d \to I\!R^m$ is a continuous map.

The function g encodes the modelling framework such as an nonlinear programming or an inverse problem and the function f is a regularizer used to induce further properties on the solution such as sparsity, smoothness, or a domain restriction. The function c is the functional data associated with a specific instance of the problem.

Recently there has been a resurgence of interest in the foundations of this problem class due to its importance for many problems in modern optimization, machine learning, and large scale data analysis with major contributions could be found in [2, 5, 8] and references therein. In order to go to the essential informations to share, we assume the reader has some basic knowledge of convex analysis as can be found, for example, in [13].

Throughout section 2, we will assume that the original function φ is bounded below. As it is well known that a necessary condition for $x \in dom\varphi$ to be a local minimizer of φ is in general hard to be reached, so, we will focus our attention on finding critical points of φ.

Before stating the definition of critical points, given a convex function $\varphi : I\!R^d \to I\!R \cup \{+\infty\}$, recall that a vector v is called a subgradient of φ at a point $x \in dom\varphi$ if the inequality

$$\varphi(y) \geq \varphi(x) + \langle v, y - x\rangle \ \forall y \in I\!R^d.$$

The set of all subgradients of φ at x is denoted by $\partial\varphi(x)$, and is called the subdifferential of φ at x. For any point $x \notin dom\varphi$, we define $\partial\varphi(x)$ to be the empty set. For any closed function φ, its Frechét subdifferential at x, $\hat{\partial}\varphi(x)$, is the collection of vectors v such that

$$\varphi(y) \geq \varphi(x) + \langle v, y - x\rangle + o(\|y - x\|) \ \forall y \in I\!R^d.$$

Unfortunately, $\hat{\partial}\varphi$ can be empty at certain points ever for Lipschitz continuous functions. To avoid this degeneracy, we arrive to the limiting subdifferential, $\partial\varphi$, defined as

$$v \in \partial\varphi(x) \Leftrightarrow \exists x_k \to x, \varphi(x_k) \to \varphi(x), v_k \in \hat{\partial}\varphi(x_k), v_k \to v.$$

Clearly, $\hat{\partial}\varphi(x) \subset \partial\varphi(x)$ for all x. It is known that the above subdifferential reduces to the classical subdifferential in convex analysis when φ is convex. In addition, if φ is continuously differentiable, then the limiting subdifferential reduces to the gradient, $\nabla\varphi$, of the function φ. From the definition, it follows that if $\bar{x}$ is a local minimizer, then $0 \in \hat{\partial}\varphi(\bar{x})$ and $0 \in \partial\varphi(\bar{x})$, which generalizes the familiar Fermat's rule. The latter condition is in general hard to be reached and we relax it to the following notion of critical point. Throughout, we also assume that the chain rule can be applied when needed which is always the case via qualification conditions as $N_{domg}(c(\bar{x})) \cap Ker(\nabla c(\bar{x})^*) = \{0\}$, N_{domg} being the normal cone to $domg$ or even better like $I\!R_+\big(domg - c(\bar{x})\big) - \nabla c(\bar{x})(I\!R^m) = I\!R^d$ are always satisfied, see [11].

Definition 1.1. We will say that $\bar{x} \in I\!R^d$ is a critical point of φ in (1.1) if

$$\partial f(\bar{x}) \cap \nabla c(\bar{x})^* \partial g(c(\bar{x})) \neq \emptyset. \tag{1.5}$$

Throughout section 2, Γ will denote the critical points set of the function φ.

2. Linear Convergence of a Proximal Linear Method

Throughout this section, we make the following assumptions on the functional components of the problem: $f : I\!R^d \to I\!R \cup \{+\infty\}$ is a proper, closed, convex function, $g : I\!R^m \to I\!R \cup \{+\infty\}$ is a convex differentiable function with a L-Lipschitz continuous gradient, i.e.,

$$\|\nabla g(x) - \nabla g(y)\| \leq L\|x - y\| \;\; \forall x, y \in I\!R^m$$

and $c : I\!R^d \to I\!R^m$ is a C^1-smooth mapping with a β-Lipschitz continuous Jacobian map

$$\|\nabla c(x) - \nabla c(y)\| \leq \beta\|x - y\| \;\; \forall x, y \in I\!R^d.$$

The proposed algorithm look for critical points of $f := g - h \circ c$ by constructing two sequences $(x_k)_{k\in I\!N}$ and $(w_k)_{k\in I\!N}$ by the following rules

$$\begin{cases} w_k \in \nabla c(x_k)^* \partial g(c(x_k)); \\ x_{k+1} = \arg\min_{x \in I\!R^d} \big(f(x) - \langle w_k, x - x_k\rangle + \frac{1}{2\lambda_k}\|x - x_k\|^2\big). \end{cases} \tag{2.1}$$

We showed in [9] that (2.1) is a descent method, that the sequence $\varphi(x_k)_{k\in I\!N}$ is convergent and that the sequence $(x_k)_{k\in I\!N}$ is asymptotically regular, namely $\lim\limits_{k\to+\infty} \|x_{k+1} - x_k\| = 0$. We proved also, [9]-Theorem 3.3, that every cluster-point $\bar{x}$ of $(x_k)_{k\in I\!N}$ is a critical point of the function φ provided that the sequence $(x_k)_{k\in I\!N}$ is bounded.

Here, we are interested in the case where φ is a DC function with h differentiable and ∇c bounded which is the case in the classical case of composite functions with linear operators.

Theorem 2.1. *Suppose that f is a strongly convex function (with constant $\rho > 0$), g a differantiable function with Lipschitz continuous gradient (with constant $L > 0$) and there exists a positive real γ such that $\sup_x \|\nabla c(x)\| \leq \gamma$. If $\rho > \frac{\kappa}{2}$, then there exists a constant $0 < r < 1$ such that*

$$\|x_{k+1} - \bar{x}\| \leq r\|x_k - \bar{x}\| \quad \forall \bar{x} \in \Gamma. \tag{2.2}$$

Therefore, the whole sequence $(x_k)_{k\in I\!N}$ converges linearly to a point $\bar{x} \in \Gamma$.

Proof. According again to (2.1) together with the subdifferential chain rule, we can write

$$w_k + \frac{x_k - x_{k+1}}{\lambda_k} \in \partial f(x_{k+1}) \quad w_k = \nabla c(x_k)^* y_k \text{ with } y_k = \nabla g(c(x_k)). \tag{2.3}$$

Let $\bar{x} \in \Gamma$ be a critical point of f, namely,

$$\bar{w} \in \partial f(\bar{x}) \quad \bar{w} = \nabla c(\bar{x})^* \bar{y} \text{ with } \bar{y} = \nabla g(c(\bar{x})). \tag{2.4}$$

By strongly monotonicity of g, (2.3) and (2.4), we have

$$\begin{array}{rcl} 0 & \leq & \langle w_k - \bar{w} + \frac{x_k - x_{k+1}}{\lambda_k}, x_{k+1} - \bar{x} \rangle - \rho \|x_{k+1} - \bar{x}\|^2 \\ & = & \langle w_k - \bar{w}, x_{k+1} - \bar{x} \rangle + \frac{1}{\lambda_k} \langle x_k - x_{k+1}, x_{k+1} - \bar{x} \rangle - \rho \|x_{k+1} - \bar{x}\|^2. \end{array}$$

Having in mind that there exists a positive real γ such that $\sup_x \|\nabla c(x)\| \leq \gamma$, using both the Cauchy-Schwarz inequality and the mean value theorem together with both Lipschitz continuity of ∇g and ∇c, we successively have

$$\begin{array}{rcl} \langle w_k - \bar{w}, x_{k+1} - \bar{x} \rangle & = & \langle \nabla c(x_k)^* y_k - \nabla c(\bar{x})^* \bar{y}, x_{k+1} - \bar{x} \rangle \\ & \leq & \|\nabla c(x_k)^* y_k - \nabla c(\bar{x})^* \bar{y}\| \|x_{k+1} - \bar{x}\| \\ & \leq & \big(\|\nabla c(x_k)^*\| \|y_k - \bar{y}\| + \|\bar{y}\| \|\nabla c(x_k)^* - \nabla c(\bar{x})^*\| \big) \|x_{k+1} - \bar{x}\| \\ & \leq & \big(L \|\nabla c(x_k)^*\| \|c(x_k) - c(\bar{x})\| + \beta \|\bar{y}\| \|x_k - \bar{x}\| \big) \|x_{k+1} - \bar{x}\| \\ & \leq & \big(\gamma^2 L + \beta \|\nabla g(c(\bar{x}))\| \big) \|x_k - \bar{x}\| \|x_{k+1} - \bar{x}\|. \end{array}$$

In view of the identity $\langle x_k - x_{k+1}, x_{k+1} - \bar{x} \rangle = \|x_k - \bar{x}\|^2 - \|x_{k+1} - \bar{x}\|^2 - \|x_{k+1} - x_k\|^2$, by setting $\kappa = \big(\gamma^2 L + \beta \|\nabla g(c(\bar{x}))\| \big)$, we can write

$$\begin{array}{rcl} 0 & \leq & 2\lambda_k \kappa \|x_k - \bar{x}\| \|x_{k+1} - \bar{x}\| - \|x_{k+1} - \bar{x}\|^2 - \|x_{k+1} - x_k\|^2 \\ & + & \|x_k - \bar{x}\|^2 - 2\lambda_k \rho \|x_{k+1} - \bar{x}\|^2 \\ & \leq & \lambda_k \kappa (\|x_k - \bar{x}\|^2 + \|x_{k+1} - \bar{x}\|^2) - \|x_{k+1} - \bar{x}\|^2 - \|x_{k+1} - x_k\|^2 \\ & + & \|x_k - \bar{x}\|^2 - 2\lambda_k \rho \|x_{k+1} - \bar{x}\|^2 \\ & = & (1 + \lambda_k \kappa) \|x_k - \bar{x}\|^2 - \big(1 + \lambda_k (2\rho - \kappa) \big) \|x_{k+1} - \bar{x}\|^2 - \|x_{k+1} - x_k\|^2. \end{array}$$

It follows that

$$\big(1 + \lambda_k (2\rho - \kappa) \big) \|x_{k+1} - \bar{x}\|^2 \leq (1 + \lambda_k \kappa) \|x_k - \bar{x}\|^2,$$

which complete the proof by setting $0 < r := \sqrt{\frac{1+\lambda_k \kappa}{1+\lambda_k(2\rho - \kappa)}} < 1$ and by virtue of [9]-Theorem 3.3 together with [12]-Opial's Lemma. □

If c is the identity function, then Algorithm (1.3) reduces to the linearized proximal point algorithm introduced in [14] for DC functions and we retrieve their main Theorem 3 with a weaker condition on the strong convexity modulus. We foresee further progress in this topics in the near future beginning with global and Linear convergences of the sequences (1.3) under suitable additional assumptions such as Kurdyka-Lojasiewicz property, which is satisfied by a wide variety of functions such as proper closed semi algebraic functions, and plays an important role in the convergence analysis of many first-order methods.

3. New on Split Convex Composite Minimization

Convex composite optimization is a class of nonsmooth, nonconvex optimization problems that captures a wide variety of optimization models studied in modern optimization practice and theory including nonlinear programming, nonconvex minimax problems, nonconvex system identification, inverse problems and nonlinear filtering as well as most nonconvex problems in large-scale data analysis and machine learning, see for example [6] and [7].

In this section, we are interested in the following class of composite optimization problems

$$\min_{x\in\mathbb{R}^d} f(x) + g(c(x)), \tag{3.1}$$

where $g : \mathbb{R}^m \to \mathbb{R}\cup\{+\infty\}$ and $f : \mathbb{R}^d \to \mathbb{R}\cup\{+\infty\}$ are closed convex functions and $c : \mathbb{R}^d \to \mathbb{R}^m$ is a given map.

Our goal is to propose an alternative formulation that relies on an infimal post-composition. To begin with, let us introduce the following notion,

Definition 3.1. Given $f : \mathbb{R}^d \to \mathbb{R} \cup \{+\infty\}$ and a map $c : \mathbb{R}^d \to \mathbb{R}^m$, the infimal post-composition of f by c is defined by

$$c \triangleright f : \mathbb{R}^m \to \mathbb{R} \cup \{+\infty\} : u \mapsto \inf_{x\in\mathbb{R}^d; c(x)=u} f(x). \tag{3.2}$$

$c \triangleright f$ being exact at $u \in \mathbb{R}^m$, if $arg\min_{x\in\mathbb{R}^d;c(x)=u} f(x) \neq \emptyset$. Note that $dom(c \triangleright f) \subset ran(c)$.

Relying on this notion, we provide in what follows an alternative formulation of (3.3).

Problem. Let $f : \mathbb{R}^d \to \mathbb{R} \cup \{+\infty\}$ be a proper convex lower semicontinuous function, and let $c : \mathbb{R}^d \to \mathbb{R}^m$ be a given map, we consider the problems

$$\inf_{x\in\mathbb{R}^d} f(x) + g(c(x)), \tag{3.3}$$

and

$$\inf_{u\in\mathbb{R}^m} g(u) + (c \triangleright f)(u). \tag{3.4}$$

The solution sets of (3.3) and (3.4) will be denoted by $S((3.3))$ and $S((3.4))$, respectively.

The following proposition gives a connection between problems (3.3) and (3.4).

Proposition 3.2. *The following statements hold true*

(i) $\inf_{x\in\mathbb{R}^d} f(x) + g(c(x)) = \inf_{u\in\mathbb{R}^m} g(u) + (c \triangleright f)(u)$.

(ii) *Suppose that* $\bar{x} \in S((3.3))$, *then* $c(\bar{x}) \in S((3.4))$.

(iii) *Suppose that* $\bar{u} \in S((3.4))$, *then* $arg\min_{x\in\mathbb{R}^d;c(x)=\bar{u}} f(x) \subset S((3.3))$.

Proof. (i)

$$\begin{aligned} \inf_{u\in\mathbb{R}^m} g(u) + (c \triangleright f)(u) &= \inf_{u\in\mathbb{R}^m} \big(g(u) + \inf_{x\in\mathbb{R}^d;c(x)=u} f(x)\big) \\ &= inf_{x\in\mathbb{R}^d}\big(f(x) + \inf_{u\in\mathbb{R}^m;c(x)=u} g(u)\big) \\ &= inf_{x\in\mathbb{R}^d}\big(f(x) + g(c(x))\big). \end{aligned}$$

(ii) Suppose $\bar{x} \in S((3.3))$, we successively have,

$$\begin{aligned} f(\bar{x}) + g(c(\bar{x})) &= \inf_{x\in\mathbb{R}^d} \big(f(x) + g(c(x))\big) \\ &= \inf_{u\in\mathbb{R}^m} g(u) + (c \triangleright f)(u) \\ &\leq g(c(\bar{x})) + (c \triangleright f)(c(\bar{x})) \\ &= g(c(\bar{x})) + f(c(\bar{x})). \end{aligned}$$

Which leads to

$$g(c(\bar{x})) + (c \triangleright f)(c(\bar{x})) = \min_{u\in\mathbb{R}^m} g(u) + (c \triangleright f)(u),$$

in other words $\bar{u} = c(\bar{x})$ is a solution of (3.4).

(iii) Suppose $\bar{u} \in S((3.4))$ and let $\bar{x}$ such that $c(\bar{x}) = \bar{u}$ and $f(\bar{x}) = (c \triangleright f)(\bar{u})$. Using (i), we can write

$$\begin{aligned} f(\bar{x}) + g(c(\bar{x})) &= g(\bar{u}) + (c \triangleright f)(\bar{u}) \\ &= \min_{u\in\mathbb{R}^m} g(u) + (c \triangleright f)(u) \\ &= \inf_{x\in\mathbb{R}^d} \big(f(x) + g(c(x))\big). \end{aligned}$$

from which we infer that any element $arg\min_{x\in I\!R^d; c(x)=\bar{u}} f(x)$ solves (3.3). □

Solving the alternative formulation of (3.3) by means of proximal splitting algorithms needs the computation of the proximal mapping of infimal post-composition $c\triangleright f$. This does not of course preclude the possibility of solving (3.4) using other algorithms. But, if we want to follow exactly the approach developed in [3], $c\triangleright f$ must be proper, convex and lower semicontinuous. Note that in the case when $c = L$ a linear operator, the key properties have been found by means of the notion of conjugate and especially the nice formula $(L\triangleright f)^* = f^*\circ L^*$ see [1]. To that end, we could define, for every $f : I\!R^d \to I\!R\cup\{+\infty\}$ and a map $c : I\!R^d \to R^m$, the following similar notions

$$f_c^*(u) = \sup_{x\in I\!R^d}(\langle c(x), u\rangle - f(x)) \text{ and } prox_{c,f}(u) = \inf_{x\in I\!R^d}(f(x) + \frac{1}{2}\|c(x)-u\|^2),$$

and use the analysis developed in [4] even [2]. Indeed in [4] was studied the case where the composition is convex. The convex-composite setting was of course proposed as a structure approach to nonconvex problems. We refer to some relevant results of the analysis of convex convex-composite functions developed in [4], see also [2]-Theorem 1 for a more general result together with the conjugate duality proposed for other possible results similar to those developed in [1] for the classical infimal post-composition.

4. Conclusion

To conclude, we first propose a linear convergence result for a proximal linear method for solving DC-composite minimization problems. This generalizes, in the special case of DC optimization, the main Theorem in [14], namely Theorem 3 and improves the condition on the strong convexity modulus. The result still holds true for inexact versions of the algorithm. Indeed, for the method to be practical, it is important to handle approximate solutions of subproblems. We also provide an alternative formulation of split convex composite minimization problems relying on a notion of infimal post-composition. We expect that this new approach becomes a further step towards developing other algorithms for solving split composite minimization problems. This will be the subject of further research.

Statements and Declarations

The author declares that he has no conflict of interest, and the manuscript has no associated data.

Acknowledgments

I would like to thank my research laboratory, Laboratoire d'Informatique et Systèmes, and the I&M team for giving me complete freedom to choose my research topics. Research is a craft, not an industry.

References

[1] H. H. Bauschke and P. L. Combettes. *Convex Analysis and Monotone Operator Theory in Hilbert Spaces*. Second Edition, CMS Books in Mathematics, Springer, Cham, 2017.

[2] R. I. Bot, S. M. Grad, and G. Wanka. A new constraint qualification for the formula of the subdifferential of composed convex functions in infinite dimensional spaces. *Mathematische Nachrichten*, 281: 1088-1107, 2008.

[3] L. M. Briceno-Arias and N. Pustelnik. Infimal post-composition approach for composite convex optimization applied to image restoration. *Signal Processing*, 223: 109549, 2024.

[4] J. V. Burke, T. Hoheisel, and Q. V. Nguyen. A study of convex convex-composite functions via infimal convolution with applications. *Mathematics of Operations Research*, 46 (4): 1324-1348, 2021.

[5] D. Drusvyatskiy and A. Lewis. Error bounds, quadratic growth, and linear convergence of proximal methods. *Mathematics of Operations Research*, 43(3): 919-948, 2018.

[6] D. Drusvyatskiy and C. Paquette. Efficiency of minimizing compositions of convex functions and smooth maps. *Mathematical Program, Series A and B*, 178 (1-2): 503-558, 2019.

[7] J. C. Duchi and F. Ruan. Solving (most) of a set of quadratic equalities: Composite optimization for robust phase retrieval. *Information and Inference: A Journal of the Institute of Mathematics and its Applications*, 8(3): 471-529, 2019. .

[8] A. S. Lewis and S. J. Wright. A proximal method for composite minimization. *Mathematical Programming*, 158(1-2), Series B, 501-546, 2016.

[9] A. Moudafi. Proximal linear methods for DC composite minimization problems. *Journal of Applied and Numerical Optimization*, 5(3): 391-398, 2023.

[10] A. Moudafi and P.-E. Maingé. On the convergence of an approximate proximal method for D.C. functions. *Journal of Computational Mathematics*, 24(4): 475-480, 2006.

[11] B. S. Mordukhovich. *Variational Analysis and Generalized Differentiation I.* Volume 330 of Grundlehren der Mathematischen Wissenschaften, Fundamental Principles of Mathematical Sciences. Springer, Berlin, 2006.

[12] Z. Opial. Weak convergence of the sequence of successive approximations for nonexpansive mappings. *Bulletin of the American Mathematical Society*, 73: 591-597, 1967.

[13] R. T. Rockafellar and R. J.-B. Wets. *Variational Analysis*. Springer-Verlag, Berlin Heidelberg, 1998.

[14] J. C. O. Souza, P. R. Oliveira, and A. Soubeyran. Global convergence of a proximal linearized algorithm for difference of convex functions. *Optimization Letters*, 10: 1529-1539, 2016.

[15] W. Sun, R. J. B. Sampaio, and M. A. B. Candido. Proximal point algorithm for minimization of DC Functions. *Journal of Computational Mathematics* 21: 451-462, 2003.

[16] P. D. Tao and L. T. H. An. Convex analysis approach to dc programming: Theory, algorithms and applications. *Acta Mathematica Vietnamica*, 22: 289-355, 1997.