Pure and Applied Functional Analysis
Volume 8, Number 3, 2023, 847–854

Yokohama Publishers
ISSN 2189-3764 ONLINE JOURNAL

A REGULARIZATION OF DC OPTIMIZATION

ABDELLATIF MOUDAFI

Abstract. Numerous models of real world nonconvex optimization can be formulated as DC optimization problems which consist in minimizing a difference of two convex functions, see for instance [1]. A popular approach to address nonsmooth terms in convex optimization is to approximate them with their Moreau envelopes, see for example [7]. In the spirit of an Hiriart-Urruty's idea [2], we propose a complete smooth approximation of the original problem that relies on Moreau envelopes with eventually different regularization parameters. This would allow to enforcing the regularization of the convex or the concave part. A parallel proximal algorithm based on the classical gradient descent method is also proposed.

1. Introduction

Throughout, we will need few technical tools from variational analysis [4]. We equip $\mathbb{R}^n$ with the usual inner product $\langle \cdot, \cdot \rangle$ and the induced Euclidean norm $\| \cdot \|$ and we are interested in DC optimization problems of the form

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

where $g : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ and $h : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ are two proper, lower semi-continuous (lsc) convex functions. DC programming provided an extension of convex programming, not too large to still allow using convex analysis and convex optimization tools but suficientely wide to cover most real world nonconvex problems. Indeed, this problem has many applications such as multicommodity network, image restoration processing, discrete tomography, clustering and seems particularly well suited to model several nonconvex industrial problems (portfolio optimization, fuel mixture, molecular biology, phylogenetic analysis ...), see for example [1]. The convexity of the two DC components g and h of the objective function has been used to develop appropriate tools from both theoretical and algorithmic viewpoints. Besides, regularization has played a major role in recent development of statistical machine learning algorithms and other applications. Moreau envelopes have been prevalent due to their nice properties and computational convenience. In this paper, we address a regularization of both functions h and g with possibly different regularization parameters. This leads to

$$\min_{x \in \mathbb{R}^n} \varphi_{\lambda,\mu}(x) = g_\lambda(x) - h_\mu(x), \tag{1.2}$$

2020 *Mathematics Subject Classification.* Primary, 49J53, 65K10; Secondary, 49M37, 90C25.

Key words and phrases. DC minimization, regularization, Moreau-envelope, descent gradient iterations, proximity mapping, (approximate) stationary point.

$\lambda > 0, \mu > 0$, g_λ and h_μ standing for the Moreau envelopes of g and h, respectively.

This asymmetric regularization approach would give a flexibility and would allow to better regularize the convex or the concave part of φ by adapting the choice of the two parameters. The reason is that the functions g and h may have very distinct rough properties.

Remember that for a proper, lsc function f and parameter γ, its Moreau envelope is defined as

$$f_\gamma(x) = \inf_{w \in I\!R^n} \{f(w) + \frac{1}{2\gamma}\|w - x\|^2\}.$$

For instance, if f is the 1-norm, then f_γ is the celebrated Huber's function in robust statistics.

It is worth mentioning that (1.1) subsumes a wide spectrum of problems in optimization, see for instance [1] and the references therein. The major advantage of Problem (1.2) is its smoothness and hence can then be tackled via fast smooth optimization solvers.

Throughout, we will assume that the original function φ is bounded below.

Before stating the definition of exact and approximate stationary points, see [6], which are relaxed versions of the necessary condition, recall that the partial differential of a function f is defined as

$$\partial f(x) := \{u \in I\!R^n; f(w) \geq f(x) + \langle u, w - x\rangle \ \forall w \in I\!R^n\}.$$

Definition 1.1. $\rhd$ We will say that $x^* \in I\!R^n$ is a stationary point of φ if

$$\partial g(x^*) \cap \partial h(x^*) \neq \emptyset, \tag{1.3}$$

a relaxed version of the necessary condition $\partial h(x^*) \subset \partial g(x^*)$.

$\rhd$ Furthermore, we will say that $x^* \in I\!R^n$ is an ε-stationary point of φ if,

(1.4)

$\exists(\xi^*, y^*) \in I\!R^n \times I\!R^n$ such that $\xi^* \in \partial g(x^*) - \partial h(y^*)$ and $\max\{\|\xi^*\|, \|x^* - y^*\|\} \leq \varepsilon$.

We will see that during the descent gradient method, the subdifferentials of g and h are not evaluate at the same point, so the ε-stationarity is then a natural relaxation of the exact one.

Proposition 1.2. *Let f be a proper, convex lsc, then, for every $\gamma > 0$, the Moreau envelope f_γ is $\frac{1}{\gamma}$-Lipschitz-continuously differentiable with gradient*

$$\nabla f_\gamma(x) = \frac{x - prox_\gamma^f(x)}{\gamma} = prox_{\gamma^{-1}}^{f^*}(x/\gamma). \tag{1.5}$$

f^* stands for the conjugate of f, namely $f^*(x) = \sup_{w \in I\!R^n}\{\langle x, w\rangle - f(w)\}$.

Note that, we also have

$$(f_\gamma)^*(x) = f^*(x) + \frac{\gamma}{2}\|x\|^2.$$

The Moreau envelope is an attractive regularization transform considering the facts that $f_\gamma(x)$ converges pointwise to $f(x)$ when $\gamma \to 0$ and that it shares the same critical points of f with

$$\inf f = \inf f_\gamma \text{ and } \mathrm{argmin} f = \mathrm{argmin} f_\gamma. \tag{1.6}$$

Note also that, it is very useful to associate $\varphi^{\diamond} = h^* - g^*$ with φ especially as $\inf \varphi = \inf \varphi^{\diamond}$.

A this stage, we want to mention that a study dealing with the same regularization idea, in the case $\lambda = \mu$, was developed in [6] with one of the two dc components assumed to be potentially weakly convex. This amounts to say that there exists $\beta > 0$ such that $g + \beta/2\|\cdot\|^2$ is convex.

Remark 1.3. Observe that in this last setting g_λ is $\max\left(\frac{1}{\lambda}, \frac{1}{1-\beta\lambda}\right)$-Lipschitz-continuously differentiable and the proximity mapping $prox_\lambda^g$ is $(1-\beta\lambda)$-cocoercive provided that $0 < \lambda < 1/\beta$, see for example [3]. We can also easily verify the relation $prox_\lambda^{g+\frac{\beta}{2}\|\cdot\|^2}(x) = prox^g_{\frac{\lambda}{1+\lambda\beta}}\left(\frac{x}{1+\lambda\beta}\right)$ according to the β-weak convexity of the function g together with the fact that $0 < \frac{\lambda}{1+\lambda\beta} < 1/\beta$.

Now, it is worth mentioning that a d.c. function φ has infinitely many d.c. decompositions, for example, $\varphi = (g + \psi) - (h + \psi)$ for every finite proper convex and lsc function ψ defined the whole space. It is then useful to find a suitable d.c. decomposition of φ, since it may have an important influence on the efficiency of algorithms for its solution. Consequently, we can relax conditions on the involving functions by assuming the functions g and/or h to be weakly convex. We can reach convexity of the two dc components by choosing an appropriate $\beta > 0$ and applying the following process

$$\varphi = g - h = (g + \beta/2\|\cdot\|^2) - (h + \beta/2\|\cdot\|^2).$$

In this case, the d.c. components in the dual problem will be continuously differentiable.

2. Some properties of $\varphi_{\lambda,\mu}$

$\triangleright$ In the light of Proposition 1.1, $\varphi_{\lambda,\mu}$ is differentiable and its gradient is given by

$$\begin{aligned}\nabla\varphi_{\lambda,\mu}(x) &= \frac{x - prox_\lambda^g(x)}{\lambda} - \frac{x - prox_\mu^h(x)}{\mu} \\ &= (\lambda^{-1} - \mu^{-1})x - (\lambda^{-1} prox_\lambda^g(x) - \mu^{-1} prox_\mu^h(x)).\end{aligned}$$

In the nice case $\lambda = \mu$, this reduces to

$$\nabla\varphi_{\lambda,\lambda}(x) = \frac{prox_\lambda^h(x) - prox_\lambda^g(x)}{\lambda}. \tag{2.1}$$

The optimality condition, in the general case, reads as

$$\nabla g_\lambda(x) = \nabla h_\mu(x) \Leftrightarrow prox_{\mu^{-1}}^{h^*}(x/\mu) = prox_{\lambda^{-1}}^{g^*}(x/\lambda), \tag{2.2}$$

and remains to solve the fixed-point problem

$$x = \kappa prox_\lambda^g(x) + (1-\kappa) prox_\mu^h(x), \text{with } \kappa = \frac{\mu}{\mu - \lambda}.$$

In the interesting case $\lambda = \mu$, relation (2.2) amounts to

$$prox_\lambda^h(x) = prox_\lambda^g(x) \Leftrightarrow prox_{\lambda^{-1}}^{h^*}(x/\lambda) = prox_{\lambda^{-1}}^{g^*}(x/\lambda). \tag{2.3}$$

Since, we always have

$$f_\gamma(x) = f(prox_\gamma^f(x)) + \frac{1}{2\gamma}\|prox_\gamma^f(x) - x\|^2 \le f(w) + \frac{1}{2\gamma}\|w - x\|^2 \ \forall w \in I\!R^n,$$

we can write

$$\begin{aligned} g_\lambda(x) - \frac{1}{2\mu}\|prox_\lambda^g(x) - x\|^2 - h(prox_\lambda^g(x)) &\le \varphi_{\lambda,\mu}(x) = g_\lambda(x) - h_\mu(x) \\ &\le g(prox_\mu^h(x)) \\ &\quad + \frac{1}{2\lambda}\|prox_\mu^h(x) - x\|^2 - h_\mu(x). \end{aligned}$$

Therefore

$$\begin{aligned} \varphi(prox_\lambda^g(x)) + \frac{1}{2\lambda\mu}(\mu - \lambda)\|prox_\lambda^g(x) - x\|^2 &\le \varphi_{\lambda,\mu}(x) \\ &\le \varphi(prox_\mu^h(x)) \\ &\quad + \frac{1}{2\lambda\mu}(\mu - \lambda)\|prox_\mu^h(x) - x\|^2. \end{aligned}$$

In other words

$$\begin{aligned} \varphi(prox_\lambda^g(x)) + \frac{\lambda}{2\mu}(\mu - \lambda)\|\nabla g_\lambda(x)\|^2 &\le \varphi_{\lambda,\mu}(x) \\ &\le \varphi(prox_\mu^h(x)) + \frac{\mu}{2\lambda}(\mu - \lambda)\|\nabla h_\mu(x)\|^2. \end{aligned}$$

The former inequality, in the nice case $\lambda = \mu$, reduces to

$$\varphi(prox_\lambda^g(x)) \le \varphi_{\lambda,\lambda}(x) \le \varphi(prox_\lambda^h(x)). \tag{2.4}$$

In this optimal case the smooth function $\varphi_{\lambda,\lambda}$ shares minimizers and stationary points with the nonsmooth function φ, see the interesting paper by Hiriart-Urruty [2].

$\triangleright$ Now, let us show that the gradient $\nabla\varphi_{\lambda,\mu}$ is $\frac{1}{\mu}$-weakly monotone (which amounts to saying that $\varphi_{\lambda,\mu} + \frac{1}{\mu}\|\cdot\|^2$ is convex) and that it is $\frac{1}{\max(\lambda,\mu)}$-smooth.

$$\begin{aligned} \langle\nabla\varphi_{\lambda,\mu}(x) - \nabla\varphi_{\lambda,\mu}(y), x - y\rangle &= \langle(\frac{x - prox_\lambda^g(x)}{\lambda} - \frac{x - prox_\mu^h(x)}{\mu}) \\ &\quad -(\frac{y - prox_\lambda^g(y)}{\lambda} - \frac{y - prox_\mu^h(y)}{\mu}), x - y\rangle \\ &= (\lambda^{-1} - \mu^{-1})\|x - y\|^2 \\ &\quad -\frac{1}{\lambda}\langle prox_\lambda^g(x) - prox_\lambda^g(y), x - y\rangle \\ &\quad +\frac{1}{\mu}\langle prox_\mu^h(x) - prox_\mu^h(y), x - y\rangle. \end{aligned}$$

According to the fact that the proximal mappings are firmly nonexpansive, this leads to

$$\begin{aligned} -\frac{1}{\mu}\|x-y\|^2+\frac{1}{\mu}\|prox_\mu^h(x)-prox_\mu^h(y)\|^2 &\leq \langle \nabla\varphi_{\lambda,\mu}(x)-\nabla\varphi_{\lambda,\mu}(y), x-y\rangle \\ &\leq \frac{1}{\lambda}\|x-y\|^2 \\ &\quad -\frac{1}{\lambda}\|prox_\lambda^g(x)-prox_\lambda^g(y)\|^2. \end{aligned}$$

Which ensures that

$$-\frac{1}{\mu}\|x-y\|^2 \leq \langle \nabla\varphi_{\lambda,\mu}(x)-\nabla\varphi_{\lambda,\mu}(y), x-y\rangle \leq \frac{1}{\lambda}\|x-y\|^2, \tag{2.5}$$

and therefore

$$|\langle \nabla\varphi_{\lambda,\mu}(x)-\nabla\varphi_{\lambda,\mu}(y), x-y\rangle| \leq \frac{1}{\max(\lambda,\mu)}\|x-y\|^2. \tag{2.6}$$

Clearly, $\varphi_{\lambda,\mu}$ is then $\frac{1}{\mu}$-weakly monotone and $\frac{1}{\max(\lambda,\mu)}$-smooth.

The key properties of the Moreau envelope naturally lead to consider gradient methods. It should be mentioned that algorithms that aim to finding critical points of dc functions were presented, for exemple, in [1] and [5]. The DCA which interleaves subgradient evaluations of the second function and the first one, and the proximal-gradient method that combines an ascent subgradient step on the second function with a proximal step on the first one.

3. Gradient descent method

We consider the following algorithm based on the Classical Gradient Descent Method and we assume that g and h are two proximable functions (i.e., which loosely speaking means that their proximal operator can be efficiently computed for every $\lambda > 0$ and $\mu > 0$, respectively):

Parallel Proximal DC Algorithm: Starting from $x_0 \in I\!R^n$, it generates the iterates $(x_k, y_k, z_k)_{k\in I\!N}$ by the following rules:
Select $\mu, \lambda > 0$ and $0 < \gamma < 2$ and compute
$\triangleright$ $y_k = \nabla h_\mu(x_k) = \mu^{-1}(x_k - prox_\mu^h(x_k))$,
$\triangleright$ $z_k = \nabla g_\lambda(x_k) = \lambda^{-1}(x_k - prox_\lambda^g(x_k))$.
Next, set:

$$\triangleright\ x_{k+1} = x_k - \gamma\max(\lambda,\mu)\nabla\varphi_{\lambda,\mu}(x_k) = x_k - \gamma\max(\lambda,\mu)(z_k - y_k). \tag{3.1}$$

Theorem 3.1. *Starting from $x_0 \in I\!R^n$, we consider the iterates $(x_k, y_k, z_k)_{k\in I\!N}$ generated by the above algorithm. Then, for every $k \in I\!R^n$, we have*

$$\varphi_{\lambda,\mu}(x_{k+1}) \leq \varphi_{\lambda,\mu}(x_k) - \frac{2-\gamma}{2\gamma\max(\lambda,\mu)}\|x_{k+1}-x_k\|^2 \text{ and } \min_{1\leq i\leq k}\|y_k - z_k\| = o\Big(\frac{1}{\sqrt{k}}\Big).$$

$\triangleright$ *If the objective function φ is bounded below, then the sequence $(x_k)_{k\in I\!N}$ has a finite length and the sequences $(y_k)_{k\in I\!N}$ and $(z_k)_{k\in I\!N}$ have the same set of cluster points, says Γ.*

$\triangleright$ *If* $(x_k)_{k\in\mathbb{N}}$ *is bounded,* x^* *a cluster point of* $(x_k)_{k\in\mathbb{N}}$*, then every* $y^* \in \Gamma$ *is* ε*-stationary for* $\varphi^\diamond$*, and* $prox^g_\lambda(x^*)$ *and* $prox^h_\mu(x^*)$ *are both* ε*-stationary points for* φ*, provided that* λ *and* μ *are very close.*
$\triangleright$ *In the optimal case* $\lambda = \mu$*, we have that*

$$prox^g_\lambda(x^*) = prox^h_\lambda(x^*) \text{ and } \varphi_{\lambda,\lambda}(x^*) = \varphi(prox^g_\lambda(x^*)) = \varphi(prox^h_\lambda(x^*)).$$

Proof. We successively have

$$\begin{aligned}
\varphi_{\lambda,\mu}(x_{k+1}) &\leq \varphi_{\lambda,\mu}(x_k) + \langle \nabla\varphi_{\lambda,\mu}(x_k), x_{k+1} - x_k\rangle + \frac{1}{2\max(\lambda,\mu)}\|x_{k+1} - x_k\|^2 \\
&= \varphi_{\lambda,\mu}(x_k) - \frac{1}{\gamma\max(\lambda,\mu)}\|x_{k+1} - x_k\|^2 + \frac{1}{2\max(\lambda,\mu)}\|x_{k+1} - x_k\|^2 \\
&= \varphi_{\lambda,\mu}(x_k) - \frac{\gamma - 2}{2\gamma\max(\lambda,\mu)}\|x_{k+1} - x_k\|^2 \\
&= \varphi_{\lambda,\mu}(x_k) - \frac{(2-\gamma)\gamma\max(\lambda,\mu)}{2}\|z_k - y_k\|^2.
\end{aligned}$$

The sequence $(\varphi_{\lambda,\mu}(x_k))_{k\in\mathbb{N}}$ is a monotonically decreasing sequence. Furthermore, it converges to some limit φ^* provided that the objective function φ is bounded below. This ensurses that the sequences $\big(\|x_{k+1} - x_k\|^2\big)_{k\in\mathbb{N}}$ and $\big(\|y_k - z_k\|^2\big)_{k\in\mathbb{N}}$ are summable, which in turn imply that

$$(x_k)_{k\in\mathbb{N}} \text{ is asymptotically regular and } \min_{1\leq i\leq k}\|y_k - z_k\| = o\Big(\frac{1}{\sqrt{k}}\Big). \tag{3.2}$$

On the other hand, if the sequence $(x_k)_{k\in\mathbb{N}}$ is bounded so are $(y_k)_{k\in\mathbb{N}}$ and $(z_k)_{k\in\mathbb{N}}$ since the gradients of both g_λ and h_μ are Lipchitz continuous. Moreover the sequences $(y_k)_{k\in\mathbb{N}}$ and $(z_k)_{k\in\mathbb{N}}$ have the same set of cluster points. Now, for any cluster points x^* and y^* of the sequences $(x_k)_{k\in\mathbb{N}}$ and $(y_k)_{k\in\mathbb{N}}$, *up to possibly extracting new subsequences*, we obtain at the limit

$$y^* = \nabla g_\lambda(x^*) = \nabla h_\mu(x^*).$$

According to

$$\nabla g_\lambda(x_k) \in \partial g(prox^g_\lambda(x_k)) \Leftrightarrow prox^g_\lambda(x_k) \in \partial g^*(\nabla g_\lambda(x_k)),$$

and passing to the limit, on a subsequence, we get

$$y^* \in \partial g(prox^g_\lambda(x^*)) \Leftrightarrow prox^g_\lambda(x^*) \in \partial g^*(y^*).$$

Similarly, from

$$\nabla h_\mu(x_k) \in \partial h(prox^h_\mu(x_k)) \Leftrightarrow prox^h_\mu(x_k) \in \partial h^*(\nabla h_\mu(x_k)),$$

we derive

$$y^* \in \partial h(prox^h_\mu(x^*)) \Leftrightarrow prox^h_\mu(x^*) \in \partial h^*(y^*).$$

Consequently,

$$prox^g_\lambda(x^*) = x^* - \lambda y^* \text{ and } prox^h_\mu(x^*) = x^* - \mu y^* \text{ are } \varepsilon-\text{stationary points of } \varphi,$$

when

$$|\lambda - \mu|\|y^*\| \leq \varepsilon, \text{ in other words } \lambda \text{ is close enough to } \mu.$$

Because,

$$0 \in \partial g(prox_\lambda^g(x^*)) - \partial(prox_\mu^h(x^*)) \text{ and } \|prox_\mu^h(x^*) - prox_\lambda^g(x^*)\| = |\lambda-\mu|\|y^*\|. \tag{3.3}$$

Moreover, in this case, y^* is an ε-stationary points of $\varphi^\diamond$.

In the optimal case $\lambda = \mu$, we have $prox_\lambda^g(x^*) = prox_\lambda^h(x^*)$, hence the latter satisfy the exact stationary condition (1.3). Regarding the coincidence of the values of φ and $\varphi_{\lambda,\lambda}$ at $prox_\lambda^g(x^*)$ and x^*, this follows from relation (2.4) together with continuity of $\varphi_{\lambda,\lambda}$ and the relation $\inf \varphi_{\lambda,\lambda} = \inf \varphi$, see [2]. Note that, in this case, y^* is a stationary point of $\varphi^\diamond$. □

Remark 3.2. In the weakly convex case, having in mind Remark 1.1, following the same lines as in the proof of (2.5)-(2.6) and as long as $\beta \max(\lambda, \mu) < 1$, we derive

$$\frac{-1+\beta(\lambda-\mu)}{\mu(1-\lambda\beta)}\|x-y\|^2 \leq \langle \nabla\varphi_{\lambda,\mu}(x) - \nabla\varphi_{\lambda,\mu}(y), x-y\rangle \leq \frac{1+\beta(\lambda-\mu)}{\lambda(1-\mu\beta)}\|x-y\|^2, \tag{3.4}$$

and therefore in the case $\lambda = \mu$, clearly $\varphi_{\lambda,\mu}$ is $\frac{1}{\lambda(1-\lambda\beta)}$-smooth.

The Parallel Proximal DC Algorithm is then applicable with the following decomposition

$$\varphi = \tilde{g} - \tilde{h}, \text{ where } \tilde{g} = g + \beta/2\|\cdot\|^2 \text{ and } \tilde{h} = h + \beta/2\|\cdot\|^2,$$

and its related convergence results hold true provided that $0 < \lambda < 1/\beta$ and $0 < \gamma < 2$. Further, according to a formula in Remark 1.1, the algorithm can be written as

Parallel Proximal DC Algorithm: Starting from $x_0 \in I\!R^n$, it generates the iterates $(x_k, y_k, z_k)_{k\in I\!N}$ by the following rules:

Select $0 < \lambda < 1/\beta$ and $0 < \gamma < 2$ and compute

$\rhd$ $y_k = prox_\lambda^{\tilde{h}}(x_k) = prox^h_{\frac{\lambda}{1+\lambda\beta}}\left(\frac{x_k}{1+\lambda\beta}\right)$.

$\rhd$ $z_k = prox_\lambda^{\tilde{g}}(x_k) = prox^g_{\frac{\lambda}{1+\lambda\beta}}\left(\frac{x_k}{1+\lambda\beta}\right)$.

Next, set:

$$\rhd\ x_{k+1} = x_k - \gamma\lambda(1-\lambda\beta)\nabla\tilde{\varphi}_{\lambda,\lambda}(x_k) = x_k - \gamma(1-\lambda\beta)(z_k - y_k). \tag{3.5}$$

It is worth mentioning that accelerated versions of these approaches can also been considered, but their use will make the analysis more complicated than for the classical gradient case.

Acknowledgment

We wish to thank the anonymous referee for his careful reading of the paper and for constructive comments which helped us to improve the first version.

References

[1] L. T. H. An and D. T. Pham, *The DC programming and DCA revised with DC Models of real world nonconvex optimization problems*, Annals of operations research, **133** (2005), 23–46.

[2] J.-B. Hiriart-Urruty, *How to regularize a difference of convex functions*, Journal of Mathematical Analysis and Applications **162** (1991), 196–209.

[3] T. Hoheisel, M. Laborde, A. Oberman, A regularization interpretation of the proximal point method for weakly convex functions, Journal of Dynamics & Games **7** (2020), 79–96.
[4] R. T. Rockafellar and R. J-B Wets, *Variational Analysis*, Springer-Verlag, 1998.
[5] W. Y. Sun, R. J. B. Sampaio and M. A. B Candido, *Proximal point algorithm for minimization of DC functions*, Journal of Computational mathematics **21** (2003), 451–462.
[6] K. Sun and X-A. Sun, *Algorithms for difference of Convex (DC) programs based on difference of Moreau envelopes smoothing.* (2021) arXiv:2104.01470.
[7] I. Yamada , M. Yukawa and M. Yamagishi, *Minimizing the Moreau Envelope of Nonsmooth Convex Functions over the Fixed Point Set of Certain Quasi-Nonexpansive Mappings*, Fixed-Point Algorithms for Inverse Problems in Science and Engineering (2011), 345–390.

Manuscript received September 29 2021
revised February 15 2022

A. Moudafi
Aix-Marseille Université, Laboratoire d'Informatique et Systèmes (LIS UMR 7020 CNRS / AMU /UTLN) Marseille, France
E-mail address: abdellatif.moudafi@univ-amu.fr