Research Article

Convergence of a Proximal Point Algorithm for Solving Minimization Problems

Abdelouahed Hamdi, M. A. Noor, and A. A. Mukheimer

1 Department of Mathematics, Prince Sultan University, P.O. Box 66833, Riyadh 11586, Saudi Arabia
2 Mathematics Department, COMSATS Institute of Information Technology, Islamabad, Pakistan

Correspondence should be addressed to Abdelouahed Hamdi, whamdi99@gmail.com

Received 16 January 2012; Accepted 4 February 2012

Academic Editor: Yonghong Yao

Copyright © 2012 Abdelouahed Hamdi et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

We introduce and consider a proximal point algorithm for solving minimization problems using the technique of Güler. This proximal point algorithm is obtained by substituting the usual quadratic proximal term by a class of strictly convex distance-like functions, called Bregman functions.

1. Introduction

The purpose of this paper is twofold. Firstly, it proposes an extension of the proximal point method introduced by Güler [1] in 1992, where the usual quadratic proximal term is substituted by a class of strictly convex distance-like functions, called Bregman functions. Secondly, it offers a general framework for the convergence analysis of the proximal point method of Güler. This framework is general enough to apply different classes of Bregman functions and still yield simple convergence proofs. The methods being analyzable in this context are called Güler’s generalized proximal point algorithm, and are closely related to the Bregman proximal methods [2–5]. The analysis, we develop is different from the works in [4, 5], since our method is based on Güler’s technique.

2. Preliminaries

To be more specific, we consider the minimization problem in the following form:

\[ f_* := \min \{ f(x) : x \in \mathbb{R}^n \}; \]  

(2.1)
where \( f : \mathbb{R}^n \to \mathbb{R} \cup \{\infty\} \) is a closed proper convex function. To solve the problem (2.1), Teboulle [6], Chen and Teboulle [2, 6], Eckstein [4] and Burachik [3] proposed a general scheme using the Bregman proximal mappings of the type \( \mathcal{J}_c \)

\[
\mathcal{J}_c(y) := \arg \min_{x \in \mathbb{R}^n} \left\{ f(x) + c^{-1} D_h(x, y) \right\},
\]

(2.2)

where \( D_h \) is given by

\[
D_h(x, y) := h(x) - h(y) - \langle \nabla h(y), x - y \rangle,
\]

(2.3)

with \( h \) is a strictly convex and continuously differentiable function.

Throughout this paper, \( \| \cdot \| \) denotes the \( l_2 \)-norm and \( \langle \cdot, \cdot \rangle \) denotes the Euclidean inner product in \( \mathbb{R}^n \). Let \( G \) be a continuous single-valued mapping from \( \mathbb{R}^n \) into \( \mathbb{R}^n \). The mapping \( G \) is Lipschitz continuous with Lipschitz constant \( L \), if for all \( x, y \in \mathbb{R}^n \), \( \| G(x) - G(y) \| \leq L \| x - y \| \). We denote also by \( \rho(x, X) \) the distance of \( x \) to the set \( X \) and it is given by \( \rho(x, X) = \min_{y \in X} \| x - y \| \). Further notations and definitions used in this paper are standard in convex analysis and may be found in Rockafellar’s book [7].

This type of kernels was introduced first by [8] in 1967. The corresponding algorithm using these Bregman proximal mappings is called the Generalized Proximal Point Method (GPPM) and known also under the terminology of Bregman Proximal Methods. These proximal method solve (2.1) by considering a sequence of unconstrained minimization problems, which can be summed as follows.

**Algorithm 2.1.** (1) Initialize \( x^0 \in \mathbb{R}^n : f(x^0) < \infty, \ c_0 > 0 \).

(2) Compute the solution \( x^{k+1} \) by the iterative scheme:

\[
x^{k+1} := \arg \min_{x \in \mathbb{R}^n} \left\{ f(x) + c_k^{-1} D_h(x, x^k) \right\},
\]

(2.4)

where \( \{c_k\} \) is a sequence of positive numbers and \( D_h(\cdot, \cdot) \) is defined by (2.3).

For \( D_h(x, y) = (1/2)\|y - x\|^2 \), Algorithm 2.1 coincides with the classical proximal point algorithm (PPA) introduced by Moreau [9] and Martinet [10].

Under mild assumptions on the data of (2.1) ergodic convergence was proved [2, 5] when \( \sigma_n = \sum_{k=1}^n c_k \to \infty \) with the following global rate of convergence estimate:

\[
f(x^n) - f(x) = O \left( \frac{1}{\sigma_n} \right), \quad \forall x \in \mathbb{R}^n.
\]

(2.5)

Our purpose in this paper is to propose an algorithm of the same type as Algorithm 2.1 which has better convergence rate. To this goal, we propose to combine Güler’s scheme [1] and the Bregman proximal method. The main difference concerns the generation of an additional sequence \( \{y_k\} \subset \mathbb{R}^n \) in the unconstrained minimization (2.4) in such a way:

\[
x^{k+1} := \arg \min_{x \in \mathbb{R}^n} \left\{ f(x) + c_k^{-1} D_h(x, y_k) \right\}.
\]

(2.6)
We show (see Section 4) that this new proximal method possesses the following rate estimate

$$f(x^n) - f(x) = O\left(\frac{1}{\gamma_n^2}\right), \quad \forall x, \quad \gamma_n = \sum_{k=1}^{n} \sqrt{c_k},$$

(2.7)

which is faster than (2.5). Further, the convergence in terms of the objective values occurs when \(\gamma_n \to \infty\) which is weaker than \(\sigma_n \to \infty\).

We briefly recall here the notion of Bregman functions called also D-functions introduced by Brègman ([8], 1967), developed and used in the proximal theory by [4, 6, 11–13]. Let \(S\) be an open subset of \(\mathbb{R}^n\) and let \(h : S \to \mathbb{R}\) be a finite-valued continuously differentiable function on \(S\) be and let \(D_h\) defined by

$$D_h(x, y) := h(x) - h(y) - \langle \nabla h(y), x - y \rangle, \quad x \in \overline{S}, \ y \in S.$$  

(2.8)

**Definition 2.2.** \(h\) is called a Bregman function with zone \(S\) or a D-function if:

(a) \(h\) is continuously differentiable on \(S\) and continuous on \(\overline{S}\),
(b) \(h\) is strictly convex on \(\overline{S}\),
(c) for every \(\lambda \in \mathbb{R}\), the partial level sets \(L_1(y, \lambda) = \{x \in \overline{S} : D_h(x, y) \leq \lambda\}\) and \(L_2(x, \lambda) = \{y \in S : D_h(x, y) \leq \lambda\}\) are bounded for every \(y \in S\) and \(x \in \overline{S}\), respectively,
(d) if \(\{y_k\} \in S\) is a convergent sequence with limit \(y^*\), then \(D_h(y^*, y_k) \to 0\),
(e) if \(\{x_k\}\) and \(\{y_k\}\) are sequences such that \(y_k \to y^* \in \overline{S}\), \(\{x_k\}\) is bounded and \(D_h(x_k, y_k) \to 0\), then \(x_k \to y^*\).

From the above definition, we extract the following properties (see, for instance, [6, 13]).

**Lemma 2.3.** Let \(h\) be a Bregman function with zone \(S\). Then,

(i) \(D_h(x, x) = 0\) and \(D_h(x, y) \geq 0\) for \(x \in \overline{S}\) and \(y \in S\),
(ii) for all \(a, b \in S\) and \(c \in \overline{S}\),

$$D_h(c, a) + D_h(a, b) - D_h(c, b) = \langle \nabla h(b) - \nabla h(a), c - a \rangle,$$

(2.9)

(iii) for all \(a, b \in S\),

$$D_h(a, b) + D_h(b, a) \leq \|\nabla h(a) - \nabla h(b)\| \times \|a - b\|.$$  

(2.10)

(iv) for all \(x \in S\)

$$h^*(\nabla h(x)) = \langle x, \nabla h(x) \rangle - h(x),$$

(v) let \(\{x_k\} \in S\) such that \(x_k \to x^* \in S\), then \(D_h(x^*, x_k) \to 0\) and \(D_h(x_k, x^*) \to 0\).

**Lemma 2.4.** (i) Let \(g : \mathbb{R}^n \to \mathbb{R}\) be a strictly convex function such that

$$g \in C^2(\mathbb{R}^n), \quad \lim_{\|x\| \to \infty} \frac{g(x)}{\|x\|} = +\infty,$$

(2.11)

then \(g\) is a Bregman function.
(ii) If \( g \) is a Bregman function, then \( g(x) + c^T x + d \) for any \( c \in \mathbb{R}^n, \ d \in \mathbb{R} \), also is a Bregman function.

Remark 2.5. \( D_h(\cdot, \cdot) \) cannot be considered as a distance because of the lack of the triangle inequality and the symmetry property. \( D_h(\cdot, \cdot) \) is usually called an entropy distance.

The paper is organized as follows. In Section 3, we recall briefly the proximal point method of Güler. Section 4 will be devoted to the presentation and convergence analysis of the proposed algorithm. Finite convergence is shown in Section 5. Finally, in Section 6 we present an application of this method to solve variational inequalities problem.

3. Extragradient Algorithm

In 1992, Güler [1] has developed a new proximal point approach similar to the classical one (PPA) based on the idea stated by Nesterov [14].

Güler’s proximal point algorithm (GPPA) can be summed up as follows.

Algorithm 3.1. (i) Initialize \( x^0 \in \mathbb{R}^n : f(x^0) < \infty, \ c_0 > 0, \ A > 0 \).

Define \( v_0 := x^0, \ A_0 := A, \ k = 0 \).

(ii) Compute \( \alpha_k = (\sqrt{(A_k c_k)^2 + 4A_k c_k - A_k c_k})/2 \).

(iii) Compute the solution \( x^{k+1} \) by the iterative scheme:

\[
y_k = (1 - \alpha_k)x^k + \alpha_k v_k,
\]

\[
x^{k+1} := \arg \min_{z \in \mathbb{R}^n} \left\{ f(z) + (2c_k)^{-1}\|z - y_k\|^2 \right\},
\]

\[
\nu_{k+1} = \nu_k - \frac{x^{k+1} - y_k}{\alpha_k},
\]

\[
A_{k+1} = (1 - \alpha_k)A_k.
\]

For the convergence analysis, see Güler [1].

Remark 3.2. The GPPA can be seen as a suitable conjugate gradient type modification of the PPA of Rockafellar applied to (2.1).

4. Main Result

4.1. Introduction

The method that we are proposing is a modification of Güler’s new proximal point approach GPPA discussed in Section 3 and can be considered as a nonlinear (or a nonquadratic) version of GPPA with Bregman kernels. In this paper it is shown that this method, which we call BGPPA possesses the strong convergence results obtained by Güler [1] and therefore this new scheme provides faster (global) convergence rates than the classical Bregman proximal
point methods (cf. [2, 4–6, 11, 13, 15]). In this paper, we propose the following algorithm generalizing G"uler’s proximal point algorithm and summed up as follows.

**Algorithm 4.1.** (i) Initialize: \( x^0 \in \mathbb{R}^n : f(x^0) < \infty, \ c_0 > 0, \ A > 0. \)

Define \( v_0 := x^0, \ A_0 := A, \ k = 0 \)

(ii) Compute: \( \alpha_k \) such that \( \alpha_k^2 = (1 - \alpha_k)A_k c_k / LL^*. \)

(iii) Compute the solution \( x^{k+1} \) by the iterative scheme:

\[
\begin{align*}
y_k &= (1 - \alpha_k)x^k + \alpha_k v_k, \\
x^{k+1} &= \arg\min_{z \in \mathbb{R}^n} \left\{ f(z) + (c_k)^{-1}D_h(z, y_k) \right\}, \\
\nu_{k+1} &= \nabla h^* \left( \nabla h(\nu_k) + \frac{x^{k+1} - y_k}{\alpha_k} \right), \\
A_{k+1} &= (1 - \alpha_k)A_k.
\end{align*}
\]

In this section we develop convergence results for the generalized G"uler’s proximal point algorithm GGPPA presented in Section 4.2. Our analysis is basically based on the following lemma.

**Lemma 4.2** ([1, page 654]). One has

\[
\prod_{j=0}^{k-1} (1 - \alpha_j) \leq \frac{1}{1 + \left(\sqrt{A}/2\right) \sum_{j=0}^{k-1} \sqrt{\sigma_j}}^2
\]

for all \( \alpha_j \in [0, 1] \) and \( A > 0. \)

**Theorem 4.3.** For all \( x \in S \) such that \( f(x) < \infty, \) one has the following convergence rate estimate:

\[
f(x^k) - f(x) \leq \frac{4(f(x^0) - f(x) + AD_h(x, x^0))}{A \left(\sum_{j=0}^{k-1} \sqrt{\sigma_j} \right)^2}.
\]

**Proof.** Using the fact that \( \phi_0(x) := f(x^0) + AD_h(x, x^0), \ x^0 \in S, \) and Lemma 4.2, we obtain

\[
f(x^k) - f(x) \leq \prod_{j=0}^{k-1} (1 - \alpha_j) (\phi_0(x) - f(x))
\]

\[
\leq \frac{1}{1 + \left(\sqrt{A}/2\right) \sum_{j=0}^{k-1} \sqrt{\sigma_j}}^2 \left( f(x^0) + AD_h(x, x^0) - f(x) \right).
\]

Since \( [1 + \left(\sqrt{A}/2\right) \sum_{j=0}^{k-1} \sqrt{\sigma_j}]^2 \geq (\sum_{j=0}^{k-1} \sqrt{\sigma_j})^2 A/4, \) then (4.3) holds.
Theorem 4.4. Consider the sequence \( \{x^k\} \) generated by GGPP A and let \( x^* \) be a minimizer of \( f(x) \) on \( \mathbb{R}^n \). Assume that

1. \( h \) is a Bregman function with zone \( S \) such that \( \text{Dom}(\partial f) \subset S \),
2. \( \text{Im}(\nabla h) = \mathbb{R}^n \) or \( \text{Im}(\nabla h) \) is open,
3. \( \nabla h \) is Lipschitz continuous with coefficient \( L \),

then

(a) for all \( x^0 \in \text{Dom}(\nabla h) \), the sequence \( \{x^k\} \) is well defined,
(b) the GGPP A possesses this following convergence rate estimate:

\[
    f(x^k) - f_* \leq \frac{4(f(x^0) - f_* + \rho(x^0, X^*)^2(AL/2))}{A(\sum_{j=0}^{k-1} \sqrt{c_j})^2}
\]  \hspace{1cm} (4.5)

(c) \( f(x^k) - f_* \to 0 \), when \( \sum_{k=0}^{\infty} \sqrt{c_k} = \infty \),
(d) \[
    f(x^k) - f_* = O\left(\frac{1}{k^2}\right)
\]  \hspace{1cm} (4.6)

if \( c_k \geq c > 0 \).

Proof. (a) Follows from [4, Theorem 4].
(b) Uses assumption (2.3) in the following manner

\[
    D_h(x, x^0) = h(x) - h(x^0) - \langle \nabla h(x^0), x - x^0 \rangle
\]

\[
    \leq \left| \int_0^1 \left[ \nabla h(x^0 + t(x - x^0)) - \nabla h(x^0) \right] \|x - x^0\| dt \right|
\]

\[
    \leq L \|x - x^0\|^2 \int_0^1 t dt
\]

\[
    \leq \frac{L}{2} \|x - x^0\|^2,
\]  \hspace{1cm} (4.7)

and by taking \( x = x^* \) in (4.3), then we have

\[
    f(x^k) - f_* \leq \frac{4(f(x^0) - f_* + (AL/2)\|x^* - x^0\|^2)}{A(\sum_{j=0}^{k-1} \sqrt{c_j})^2}.
\]  \hspace{1cm} (4.8)

Since \( x^* \) is arbitrary, then (4.5) holds.
(c) Is obvious.
(d) It suffices to observe that if \( c_k \geq c > 0 \), we have

\[
\left( \sum_{j=0}^{k-1} \sqrt{c_j} \right)^2 \geq \left( \sum_{j=0}^{k-1} c \right)^2 = c^k.
\]

(4.9)

4.2. Finite Convergence

Note that the finite convergence property was established for the classical proximal point algorithm in the case of sharp minima, see, for example, [16]. Recently, Kiwiel [5] has extended this property to his generalized Bregman proximal method (BPM). In the following theorem we prove that Algorithm 3.1 has this property. Our proof is based on Kiwiel’s one [5, Theorem 6.1 page 1151].

Definition 4.5. A closed proper convex function \( f : \mathbb{R}^n \to \mathbb{R} \) is said to have a sharp minimum on \( \mathbb{R}^n \) if and only if there exists \( \tau > 0 \) such that

\[
f(x) \geq \min_{\mathbb{R}^n} f + \tau \min_{z \in \text{Argmin}_f} ||x - z|| \quad \forall x.
\]

(4.10)

Theorem 4.6. Under the same hypothesis as in Theorem 4.4 and by considering GGPPA with \( f \) having a sharp minimum on \( \mathbb{R}^n \) and \( c_k \) being bounded, then there exists \( k \) such that \( 0 \in \partial f(x_k) \) and \( x_k \in X^* \).

Proof. Straightforward, using Theorem 4.4 and [5, Theorem 6.1, page 1151]. \( \square \)

5. Convergence Rate of GGPPA

If \( \{x^k\} \) is a sequence of points, one forms the sequence \( \{z^n\} \) of weighted averages given by

\[
z^n = \sum_{k=1}^{n} \frac{c_k}{\sigma_n} x^k, \quad \sigma_n = \sum_{k=1}^{n} c_k,
\]

(5.1)

where \( c_k > 0 \). If the sequence \( \{z^n\} \) converges, then \( \{x^k\}_k \) is said to converge ergodically.

Theorem 5.1. GGPPA possesses the following convergence rate:

\[
f(x^n) - f_* = o\left( \frac{1}{\sigma_n} \right),
\]

(5.2)

that is, \( \sigma_n(f(x^n) - f_*) \to 0 \). Furthermore, if \( c_k \geq c > 0 \), then one has

\[
f(x^n) - f_* = o\left( \frac{1}{n} \right).
\]

(5.3)
Proof. Let $x^*$ be a minimizer of $f$. For brevity, we denote $W_k = f(x^k) - f_{x^*}$, $V_k = f(y^k) - f$, and $\Delta = \nabla h(y^k) - \nabla h(x^{k+1})$. At optimality in the unconstrained minimization in GGPPA, we can write

$$\frac{\Delta}{c_k} \in \partial f(x^{k+1}),$$

(5.4)

and by the convexity of $f$, we have

$$c_k \left( f(x^{k+1}) - f(x) \right) \leq \left\langle x^{k+1} - x, \Delta \right\rangle.$$  

(5.5)

Setting $x = x^k$ in (5.5), we obtain

$$c_k (W_k - W_{k+1}) \geq \left\langle x^k - x^{k+1}, \Delta \right\rangle,$$

(5.6)

and for $x = y^k$, we have

$$c_k (V_k - W_{k+1}) \geq \left\langle y^k - x^{k+1}, \Delta \right\rangle.$$  

(5.7)

Or again, if we set $x = x^*$ in (5.5), and using the Cauchy-Schwartz inequality, we obtain

$$c_k \left( f(x^{k+1}) - f_* \right) \leq \left\langle x^{k+1} - x^*, \Delta \right\rangle,$$

(5.8)

that is,

$$c_k W_{k+1} \leq \left\langle x^{k+1} - y^k, \Delta \right\rangle + \left\langle y^k - x^*, \Delta \right\rangle.$$  

(5.9)

Since $h$ is convex, $\left\langle x^{k+1} - y^k, \Delta \right\rangle \leq 0$. Then we can write

$$c_k W_{k+1} \leq \left\langle y^k - x^*, \Delta \right\rangle,$$

(5.10)

that is,

$$c_k^2 W_{k+1}^2 \leq \left\| y^k - x^* \right\|^2 \left\| \Delta \right\|^2.$$  

(5.11)

Using the relation $\left\| \Delta \right\|^2 \leq L \left\langle \Delta, y^k - x^{k+1} \right\rangle$ and the inequality (5.7), we get the relation

$$c_k^2 W_{k+1}^2 \leq L c_k (V_k - W_{k+1}) \left\| y^k - x^* \right\|^2.$$  

(5.12)
For short we denote \( M_k = \| y^k - x^* \| \) thus, (5.12) becomes
\[
c_k W_{k+1}^2 + LM_k^2 W_{k+1} \leq LM_k^2 V_k.
\] (5.13)

Then by dividing both terms by \( LM_k^2 > 0 \), we get
\[
W_{k+1} \left( \frac{c_k}{LM_k^2} W_{k+1} + 1 \right) \leq V_k.
\] (5.14)

Since the left-side term is positive, then
\[
V_k^{-1} \leq W_{k+1}^{-1} \left( \frac{c_k}{LM_k^2} W_{k+1} + 1 \right)^{-1}.
\] (5.15)

Now following Güler [17, page 410], we use the fact that \((1 + x)^{-1} \leq 1 - 2x/3\), for all \( x \in [0,1/2] \). To apply this inequality, it suffices to show that \( (c_k/ LM_k^2) W_{k+1} \) is less than or equal to 1/2. This can be deduced from this relation (see Lemma 2.3 (ii)):
\[
c_k W_{k+1} \leq \langle \Delta, x^{k+1} - x^* \rangle = D_h(x^*, y^k) - D_h(x^*, x^{k+1}) - D_h(x^{k+1}, y^k).
\] (5.16)

Indeed, since \( D_h(\cdot, \cdot) \geq 0 \), then (the proof of this next inequality can be found in the proof of Theorem 4.4-(b))
\[
c_k W_{k+1} \leq \frac{L}{2} \| x^* - y^k \|^2 = \frac{LM_k^2}{2}.
\] (5.17)

Therefore, \( 0 < (c_k/ LM_k^2) W_{k+1} \leq 1/2 \) and we obtain
\[
0 \leq V_k^{-1} \leq W_{k+1}^{-1} \left( 1 - \frac{2c_k}{3LM_k^2} W_{k+1} \right).
\] (5.18)

To continue the proof, we will separate some different cases.

**Case 1.** If \( f(x^{k+1}) \leq f(y^k) \leq f(x^k) \). Then \( W_{k+1} \leq V_k \leq W_k \) and we have \( V_k^{-1} \geq W_k^{-1} \). Thus, (5.18) becomes
\[
W_k^{-1} \leq V_k^{-1} \leq W_{k+1}^{-1} \left( 1 - \frac{2c_k}{3LM_k^2} W_{k+1} \right),
\] (5.19)

and by summation from \( k = 0 \) to \( k = n \), we get
\[
0 \leq \sum_{k=0}^{n} W_k^{-1} \leq \sum_{k=0}^{n} W_{k+1}^{-1} - \sum_{k=0}^{n} \frac{2c_k}{3LM_k^2}.
\] (5.20)
that is,

$$0 \leq W_0^{-1} \leq W_{n+1}^{-1} - \sum_{k=0}^{n} \frac{2c_k}{3LM_k^2}. \tag{5.21}$$

Then,

$$0 \leq W_{n+1} \leq \frac{3L}{2 \sum_{k=0}^{n} c_k \|y^k - x^*\|^2}. \tag{5.22}$$

Since $x^*$ is an arbitrary solution, we can write

$$0 \leq W_{n+1} \leq \frac{3L}{2 \sum_{k=0}^{n} c_k \rho(y^k, X^*)^{-2}}. \tag{5.23}$$

and by multiplying both terms by $\sigma_{n+1} = \sum_{k=0}^{n} c_k$, we obtain

$$0 \leq \sigma_{n+1} W_{n+1} \leq \frac{3L}{2 \sum_{k=0}^{n} (c_k / \sigma_{n+1}) \rho(y^k, X^*)^{-2}}. \tag{5.24}$$

Since $y^k$ and $x^{k+1}$ converge to the same point (indeed, we can see it via the formula giving $\nu_{k+1}$ in the algorithm GGPPA and $\rho(x^{k+1}, X^*) \to 0$, then $\rho(x^{k+1}, X^*)^{-2} \to \infty$; hence, we obtain

$$0 \leq \sigma_{n+1} \left( f(x^{n+1}) - f_* \right) \to 0, \tag{5.25}$$

which implies

$$f(x^{n+1}) - f_* = o\left( \frac{1}{\sigma_{n+1}} \right), \tag{5.26}$$

that is,

$$f(x^n) - f_* = o\left( \frac{1}{\sigma_n} \right). \tag{5.27}$$

Case 2. If $f(x^{k+1}) \leq f(x^k) \leq f(y^k)$. Then $W_{k+1} \leq W_k \leq V_k$ and we have $W_{k+1}^{-1} \geq W_k^{-1}$ therefore; for $n \geq k$ we have $W_{n+1}^{-1} \geq W_k^{-1}$. Thus, using inequality (5.18), we write

$$W_{n+1}^{-1} \geq W_k^{-1} \geq \frac{2c_k}{3LM_k^2}. \tag{5.28}$$
and by summation from $k = 0$ to $k = n$, we get

$$(n + 1)W_{n+1}^{-1} \geq \frac{2c_k}{3LM_k^2}.$$  \hfill (5.29)

Then,

$$0 \leq W_{n+1} \leq \frac{3L(n + 1)}{2 \sum_{k=0}^{n} c_k \|y^k - x^*\|^2}.$$  \hfill (5.30)

Since $x^*$ is an arbitrary solution, we can write

$$0 \leq W_{n+1} \leq \frac{3L(n + 1)}{2 \sum_{k=0}^{n} c_k \rho (y^k, X^*)^{-2}},$$  \hfill (5.31)

and by multiplying both terms by $\sigma_{n+1} = \sum_{k=0}^{n} c_k$, we obtain

$$0 \leq \sigma_{n+1} W_{n+1} \leq \frac{3L}{2 \sum_{k=0}^{n} (c_k / (n + 1)(n + 1)\sigma_{n+1}) \rho (y^k, X^*)^{-2}}.$$  \hfill (5.32)

Since $y^k$ and $x^{k+1}$ converge to the same point (indeed, we can see it via the formula giving $v_{k+1}$ in the algorithm GGPPA and $\rho(x^{k+1}, X^*) \to 0$, then $\rho(x^{k+1}, X^*)^{-2} \to \infty$; hence, we obtain

$$0 \leq \sigma_{n+1} (f(x^{n+1}) - f^*) \to 0,$$  \hfill (5.33)

which implies,

$$f(x^{n+1}) - f^* = O\left(\frac{1}{\sigma_{n+1}}\right),$$  \hfill (5.34)

that is,

$$f(x^n) - f^* = O\left(\frac{1}{\sigma_n}\right).$$  \hfill (5.35)

Case 3. If $f(x^k) \leq f(x^{k+1}) \leq f(y^k)$. In this case we observe that sequence $\{f(x^k)\}$ is increasing, which may imply a divergence of the approach.

Since $f$ is convex, then the following convergence rate estimate can be derived directly.

**Corollary 5.2.** If one assumes that $c_k \geq c > 0$ for all $k$, then

$$f(x^n) - f^* = O\left(\frac{1}{n}\right).$$  \hfill (5.36)
6. Conclusion

We have introduced an extragradient method to minimize convex problems. The algorithm is based on a generalization of the technique originally proposed by Nesterov [14] and readapted by Güler in [1, 17], where the usual quadratic proximal term was substituted by a class of convex nonquadratic distance-like functions. The new algorithm has a better theoretical convergence rate compared to the available ones. This motivates naturally the study of the numerical efficiency of the new algorithm and its application to solve variational inequality problems [18, 19]. Also, further efforts are needed to consider the given study for nonconvex situations and apply it to solve nonconvex equilibrium problems [20].

Acknowledgments

The authors would like to thank Dr. Osman Güler for providing them with reference [12] and the English translation of the original paper of Nesterov [14].

References


Submit your manuscripts at http://www.hindawi.com