Research Article

Convergence of a Proximal Point Algorithm for Solving Minimization Problems

Abdelouahed Hamdi,[1] M. A. Noor,[2] and A. A. Mukheimer[1]

[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

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 convex nonquadratic distance-like functions. It can be seen as an extragradient iterative scheme. We prove the convergence rate of this new proximal point method under mild assumptions. Furthermore, it is shown that this estimate rate is better than the available ones.

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:

$$f_* := \min\{f(x) : x \in \mathbb{R}^n\}, \tag{2.1}$$

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 $\mathcal{Q}_c$

$$\mathcal{Q}_c(y) := \arg \min_{x \in \mathbb{R}^n} \left\{ f(x) + c^{-1} \mathbf{D_h}(x, y) \right\}, \tag{2.2}$$

where $\mathbf{D_h}$ is given by

$$\mathbf{D_h}(x, y) := h(x) - h(y) - \langle \nabla h(y), x - y \rangle, \tag{2.3}$$

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:

$$x^{k+1} := \arg \min_{x \in \mathbb{R}^n} \left\{ f(x) + c_k^{-1} \mathbf{D_h}\left(x, x^k\right) \right\}, \tag{2.4}$$

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

For $\mathbf{D_h}(x, y) = (1/2)\|x - y\|^2$, 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:

$$f(x^n) - f(x) = \mathcal{O}\left(\frac{1}{\sigma_n}\right), \quad \forall x \in \mathbb{R}^n. \tag{2.5}$$

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:

$$x^{k+1} := \arg \min_{x \in \mathbb{R}^n} \left\{ f(x) + c_k^{-1} \mathbf{D_h}(x, y_k) \right\}. \tag{2.6}$$

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

$$f(x^n) - f(x) = \mathcal{O}\left(\frac{1}{\gamma_n^2}\right), \quad \forall x, \ \gamma_n = \sum_{k=1}^{n} \sqrt{c_k}, \tag{2.7}$$

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 $h : \overline{S} \to \mathbb{R}$ be a finite-valued continuously differentiable function on S be and let $\mathbf{D_h}$ defined by

$$\mathbf{D_h}(x, y) := h(x) - h(y) - \langle \nabla h(y), x - y \rangle, \quad x \in \overline{S}, \ y \in S. \tag{2.8}$$

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 $L_1(y, \lambda) = \{x \in \overline{S} : \mathbf{D_h}(x, y) \le \lambda\}$ and $L_2(x, \lambda) = \{y \in S : \mathbf{D_h}(x, y) \le \lambda\}$ 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 $\mathbf{D_h}(y^*, y_k) \to 0$,

(e) if $\{x^k\}$ and $\{y_k\}$ are sequences such that $y_k \to y^* \in \overline{S}$, $\{x^k\}$ is bounded and $\mathbf{D_h}(x^k, y_k) \to 0$, 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) $\mathbf{D_h}(x, x) = 0$ *and* $\mathbf{D_h}(x, y) \ge 0$ *for $x \in \overline{S}$ and $y \in S$,*

(ii) *for all $a, b \in S$ and $c \in \overline{S}$,*

$$\mathbf{D_h}(c, a) + \mathbf{D_h}(a, b) - \mathbf{D_h}(c, b) = \langle \nabla h(b) - \nabla h(a), c - a \rangle, \tag{2.9}$$

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

$$\mathbf{D_h}(a, b) + \mathbf{D_h}(b, a) \le \|\nabla h(a) - \nabla h(b)\| \times \|a - b\|. \tag{2.10}$$

(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 $\mathbf{D_h}(x^*, x^k) \to 0$ and $\mathbf{D_h}(x^k, x^*) \to 0$.*

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

$$g \in C^2(\mathbb{R}^n), \qquad \lim_{\|x\| \to \infty} \frac{g(x)}{\|x\|} = +\infty, \tag{2.11}$$

then g is a Bregman function.

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

Remark 2.5. $\mathbf{D_h}(\cdot, \cdot)$ cannot be considered as a distance because of the lack of the triangle inequality and the symmetry property. $\mathbf{D_h}(\cdot, \cdot)$ 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:

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

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

$$v_{k+1} = v_k - \frac{x^{k+1} - y_k}{\alpha_k},$$

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

(3.1)

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 Güler'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: α_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:

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

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

$$v_{k+1} = \nabla h^* \left(\nabla h(v_k) + \frac{x^{k+1} - y_k}{\alpha_k} \right),$$

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

(4.1)

In this section we develop convergence results for the generalized Güler'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*

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

(4.2)

for all $\alpha_j \in [0, 1[$ and $A > 0$.

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

$$f\left(x^k\right) - f(x) \leq \frac{4 \left(f(x^0) - f(x) + A \mathbf{D_h}(x, x^0) \right)}{A \left(\sum_{j=0}^{k-1} \sqrt{c_j} \right)^2}.$$

(4.3)

Proof. Using the fact that $\phi_0(x) := f(x^0) + A\mathbf{D_h}(x, x^0), \; x^0 \in \mathcal{S}$, and Lemma 4.2, we obtain

$$f\left(x^k\right) - f(x) \leq \prod_{j=0}^{k-1} (1 - \alpha_j)(\phi_0(x) - f(x))$$

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

(4.4)

Since $\left[1 + \left(\sqrt{A}/2 \right) \sum_{j=0}^{k-1} \sqrt{c_j} \right]^2 \geq \left(\sum_{j=0}^{k-1} \sqrt{c_j} \right)^2 A/4$, then (4.3) holds. □

Theorem 4.4. *Consider the sequence $\{x^k\}$ generated by GGPPA and let x^* be a minimizer of $f(x)$ on $\mathbb{R}^n$. Assume that*

(1) *h is a Bregman function with zone $\mathcal{S}$ such that $\overline{\mathrm{Dom}(\partial f)} \subset \mathcal{S}$,*

(2) *$\mathrm{Im}(\nabla h) = \mathbb{R}^n$ or $\mathrm{Im}(\nabla h)$ is open,*

(3) *∇h is Lipschitz continuous with coefficient L,*

then

(a) *for all $x^0 \in \mathrm{Dom}(\nabla h)$, the sequence $\{x^k\}$ is well defined,*

(b) *the GGPPA possesses this following convergence rate estimate:*

$$f\left(x^k\right) - f_* \leq \frac{4\left(f(x^0) - f_* + \rho(x^0, X^*)^2 (AL)/2\right)}{A\left(\sum_{j=0}^{k-1} \sqrt{c_j}\right)^2}, \tag{4.5}$$

(c) *$f(x^k) - f_* \to 0$, when $\sum_{k=0}^{\infty} \sqrt{c_k} = \infty$,*

(d)

$$f\left(x^k\right) - f_* = \mathcal{O}\left(\frac{1}{k^2}\right) \tag{4.6}$$

if $c_k \geq c > 0$.

Proof. (a) Follows from [4, Theorem 4].

(b) Uses assumption (2.3) in the following manner

$$\begin{aligned}
\mathbf{D_h}\left(x, x^0\right) &= h(x) - h\left(x^0\right) - \left\langle \nabla h\left(x^0\right), x - x^0 \right\rangle \\
&\leq \left| \int_0^1 \left[\nabla h\left(x^0 + t\left(x - x^0\right)\right) - \nabla h\left(x^0\right) \right] \left\| x - x^0 \right\| dt \right| \\
&\leq L \left\| x - x^0 \right\|^2 \int_0^1 t\, dt \\
&\leq \frac{L}{2} \left\| x - x^0 \right\|^2,
\end{aligned} \tag{4.7}$$

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

$$f\left(x^k\right) - f_* \leq \frac{4\left(f(x^0) - f_* + (AL/2) \left\| x^* - x^0 \right\|^2\right)}{A\left(\sum_{j=0}^{k-1} \sqrt{c_j}\right)^2}. \tag{4.8}$$

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

$$\left(\sum_{j=0}^{k-1}\sqrt{c_j}\right)^2 \geq \left(\sum_{j=0}^{k-1}\sqrt{c}\right)^2 = ck^2. \tag{4.9}$$

$\square$

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

$$f(x) \geq \min_{\mathbb{R}^n} f + \tau \min_{z \in \text{Argmin} f} \|x - z\| \quad \forall x. \tag{4.10}$$

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

$$z^n = \sum_{k=1}^{n} \frac{c_k}{\sigma_n} x^k, \qquad \sigma_n = \sum_{k=1}^{n} c_k, \tag{5.1}$$

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:*

$$f(x^n) - f_* = o\left(\frac{1}{\sigma_n}\right), \tag{5.2}$$

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

$$f(x^n) - f_* = o\left(\frac{1}{n}\right). \tag{5.3}$$

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

$$\frac{\Delta}{c_k} \in \partial f\left(x^{k+1}\right), \tag{5.4}$$

and by the convexity of f, we have

$$c_k\left(f\left(x^{k+1}\right) - f(x)\right) \le \left\langle x^{k+1} - x, \Delta \right\rangle. \tag{5.5}$$

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

$$c_k(W_k - W_{k+1}) \ge \left\langle x^k - x^{k+1}, \Delta \right\rangle, \tag{5.6}$$

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

$$c_k(V_k - W_{k+1}) \ge \left\langle y^k - x^{k+1}, \Delta \right\rangle. \tag{5.7}$$

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

$$c_k\left(f\left(x^{k+1}\right) - f_*\right) \le \left\langle x^{k+1} - x^*, \Delta \right\rangle, \tag{5.8}$$

that is,

$$c_k W_{k+1} \le \left\langle x^{k+1} - y^k, \Delta \right\rangle + \left\langle y^k - x^*, \Delta \right\rangle. \tag{5.9}$$

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

$$c_k W_{k+1} \le \left\langle y^k - x^*, \Delta \right\rangle, \tag{5.10}$$

that is,

$$c_k^2 W_{k+1}^2 \le \left\| y^k - x^* \right\|^2 \|\Delta\|^2. \tag{5.11}$$

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

$$c_k^2 W_{k+1}^2 \le Lc_k(V_k - W_{k+1}) \left\| y^k - x^* \right\|^2. \tag{5.12}$$

For short we denote $M_k = \|y^k - x^*\|$ thus, (5.12) becomes

$$c_k W_{k+1}^2 + LM_k^2 W_{k+1} \leq LM_k^2 V_k. \tag{5.13}$$

Then by dividing both terms by $LM_k^2 > 0$, we get

$$W_{k+1}\left(\frac{c_k}{LM_k^2} W_{k+1} + 1\right) \leq V_k. \tag{5.14}$$

Since the left-side term is positive, then

$$V_k^{-1} \leq W_{k+1}^{-1}\left(\frac{c_k}{LM_k^2} W_{k+1} + 1\right)^{-1}. \tag{5.15}$$

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)):

$$c_k W_{k+1} \leq \langle \Delta, x^{k+1} - x^* \rangle = \mathbf{D_h}\left(x^*, y^k\right) - \mathbf{D_h}\left(x^*, x^{k+1}\right) - \mathbf{D_h}\left(x^{k+1}, y^k\right). \tag{5.16}$$

Indeed, since $\mathbf{D_h}(\cdot, \cdot) \geq 0$, then (the proof of this next inequality can be found in the proof of Theorem 4.4-(b))

$$c_k W_{k+1} \leq \mathbf{D_h}\left(x^*, y^k\right) \leq \frac{L}{2}\left\|x^* - y^k\right\|^2 = \frac{LM_k^2}{2}. \tag{5.17}$$

Therefore, $0 < (c_k/LM_k^2)W_{k+1} \leq 1/2$ and we obtain

$$0 \leq V_k^{-1} \leq W_{k+1}^{-1}\left(1 - \frac{2c_k}{3LM_k^2} W_{k+1}\right). \tag{5.18}$$

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

Case 1. If $\underline{f(x^{k+1}) \leq f(y^k) \leq f(x^k)}$. Then $W_{k+1} \leq V_k \leq W_k$ and we have $V_k^{-1} \geq W_k^{-1}$. Thus, (5.18) becomes

$$W_k^{-1} \leq V_k^{-1} \leq W_{k+1}^{-1}\left(1 - \frac{2c_k}{3LM_k^2} W_{k+1}\right), \tag{5.19}$$

and by summation from $k = 0$ to $k = n$, we get

$$0 \leq \sum_{k=0}^{n} W_k^{-1} \leq \sum_{k=0}^{n} W_{k+1}^{-1} - \sum_{k=0}^{n} \frac{2c_k}{3LM_k^2}, \tag{5.20}$$

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 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

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

which implies

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

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

$$W_{n+1}^{-1} \geq W_{k+1}^{-1} \geq \frac{2c_k}{3LM_k^2}, \tag{5.28}$$

and by summation from $k = 0$ to $k = n$, we get

$$(n + 1)W_{n+1}^{-1} \geq \sum_{k=0}^{n} \frac{2c_k}{3LM_k^2}. \tag{5.29}$$

Then,

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

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

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

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

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

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 $\rho(x^{k+1}, X^*) \rightarrow 0$, then $\rho(x^{k+1}, X^*)^{-2} \rightarrow \infty$; hence, we obtain

$$0 \leq \sigma_{n+1}\left(f\left(x^{n+1}\right) - f_*\right) \longrightarrow 0, \tag{5.33}$$

which implies,

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

that is,

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

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. $\qquad \square$

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*

$$f(z^n) - f_* = o\left(\frac{1}{n}\right). \tag{5.36}$$

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

References

[1] O. Güler, "New proximal point algorithms for convex minimization," *SIAM Journal on Optimization*, vol. 2, no. 4, pp. 649–664, 1992.

[2] G. Chen and M. Teboulle, "Convergence analysis of a proximal-like minimization algorithm using Bregman functions," *SIAM Journal on Optimization*, vol. 3, no. 3, pp. 538–543, 1993.

[3] R. S. Burachik, *Generalized proximal point methods for the variationnal inequality problem*, Ph.D. thesis, Instituto de Matemàtica Pura e Aplcada, Rio de Janeiro, Brazil, 1995.

[4] J. Eckstein, "Nonlinear proximal point algorithms using Bregman functions, with applications to convex programming," *Mathematics of Operations Research*, vol. 18, no. 1, pp. 202–226, 1993.

[5] K. C. Kiwiel, "Proximal minimization methods with generalized Bregman functions," *SIAM Journal on Control and Optimization*, vol. 35, no. 4, pp. 1142–1168, 1997.

[6] M. Teboulle, "Entropic proximal mappings with applications to nonlinear programming," *Mathematics of Operations Research*, vol. 17, no. 3, pp. 670–690, 1992.

[7] R. T. Rockafellar, *Convex Analysis*, vol. 28 of *Princeton Mathematical Series*, Princeton University Press, Princeton, NJ, USA, 1970.

[8] L. M. Brègman, "A relaxation method of finding a common point of convex sets and its application to the solution of problems in convex programming," *USSR Computational Mathematics and Mathematical Physics*, vol. 7, no. 3, pp. 620–631, 1967.

[9] J.-J. Moreau, "Proximitéet dualité dans un espace hilbertien," *Bulletin de la Société Mathématique de France*, vol. 93, pp. 273–299, 1965.

[10] B. Martinet, *Algorithmes pour la résolution de problèmes d'optimisation et de minimax*, Thèse d'Etat, Université de Grenoble, 1972.

[11] Y. Censor and S. A. Zenios, "Proximal minimization algorithm with *D*-functions," *Journal of Optimization Theory and Applications*, vol. 73, no. 3, pp. 451–464, 1992.

[12] O. Güler, "Ergodic convergence in proximal point algorithms with Bregman functions," in *Advances in Optimization and Approximation*, vol. 1 of *Nonconvex Optimization and Its Applications*, pp. 155–165, Kluwer Academic, Dordrecht, The Netherlands, 1994.

[13] Y. Censor and S. A. Zenios, *Parallel Optimization*, Numerical Mathematics and Scientific Computation, Oxford University Press, New York, NY, USA, 1997.

[14] Yu. E. Nesterov, "An approach to constructing optimal methods for minimization of smooth convex functions," *Èkonomika i Matematicheskie Metody*, vol. 24, no. 3, pp. 509–517, 1988, Translated for O. Güler by V. Rachmanov, 1988.

[15] A. N. Iusem, B. F. Svaiter, and M. Teboulle, "Entropy-like proximal methods in convex programming," *Mathematics of Operations Research*, vol. 19, no. 4, pp. 790–814, 1994.

[16] M. C. Ferris, "Finite termination of the proximal point algorithm," *Mathematical Programming*, vol. 50, no. 3, pp. 359–366, 1991.

[17] O. Güler, "On the convergence of the proximal point algorithm for convex minimization," *SIAM Journal on Control and Optimization*, vol. 29, no. 2, pp. 403–419, 1991.

[18] M. Aslam Noor, "Some developments in general variational inequalities," *Applied Mathematics and Computation*, vol. 152, no. 1, pp. 199–277, 2004.

[19] Y. Yao, M. A. Noor, Y.-C. Liou, and S. M. Kang, "Iterative algorithms for general multivalued variational inequalities," *Abstract and Applied Analysis*, vol. 2012, Article ID 768272, 10 pages, 2012.

[20] M. A. Noor, K. I. Noor, and E. Al-Said, "Iterative methods for solving nonconvex equilibrium variational inequalities," *Applied Mathematics and Information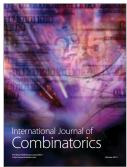 Sciences*, vol. 6, no. 1, pp. 65–69, 2012.