NUMERICAL MATHEMATICS
A Journal of Chinese Universities (English Series)

Vol. **15** No. 1, pp. 74-81
February 2006

A New Inexactness Criterion for Approximate Logarithmic-Quadratic Proximal Methods

Abdellah Bnouhachem[†]

School of Management Science and Engineering, Nanjing University, Nanjing 210093, China.

Received December 24, 2003; Accepted (in revised version) November 12, 2004

Abstract. Recently, a class of logarithmic-quadratic proximal (LQP) methods was introduced by Auslender, Teboulle and Ben-Tiba. The inexact versions of these methods solve the sub-problems in each iteration approximately. In this paper, we present a practical inexactness criterion for the inexact version of these methods.

Key words: Variational inequalities; maximal monotone operators; interior proximal methods.

AMS subject classifications: 90C33, 49J40

1 Introduction

Given an operator T, point to set in general, and a closed convex subset C of R^n, the variational inequality problem, denoted by (VI), consists of finding a vector $x^* \in C$ and $g^* \in T(x^*)$ such that

$$(x - x^*)^T g^* \geq 0, \quad \forall x \in C. \tag{1}$$

Our analysis will focus on the case where T is a maximal monotone mapping from R^n into itself and the constraint C is explicitly defined by

$$C := \{x \in R^n : Ax \leq b\},$$

where A is a $p \times n$ matrix, $b \in R^p$ and $p \geq n$. We suppose that the matrix A is of maximal rank, i.e., $\text{rank}A = n$ and that $\text{int}C = \{x : Ax < b\}$ is nonempty.

It is well known that the VI problem can be alternatively formulated as finding the zero point of a maximal monotone operator $\Pi = T + N_C$, i.e., find $x^* \in C$ such that $0 \in \Pi(x^*)$. A classical method to find the zero point of a maximal monotone operator Π is the proximal point algorithm (e.g., see [5, 8]). For given $x^{k-1} \in R^n$ and $\lambda_k \geq \lambda > 0$, the new iterate x^k is the solution of the following problem:

$$0 \in \Pi(x) + \lambda_k^{-1} \nabla q(x, x^{k-1}), \tag{2}$$

[*]Correspondence to: Abdellah Bnouhachem, School of Management Science and Engineering, Nanjing University, Nanjing 210093, China. Email: babedallah@yahoo.com
[†]The author was supported by NSFC grant 70371019.

where

$$q(x, x^{k-1}) = \frac{1}{2}\|x - x^{k-1}\|^2 \tag{3}$$

is a quadratic function of x. The proximal point algorithm can be seen as a regularization method in which the regularization parameter λ_k does not approach $+\infty$, thus avoiding the possible ill behavior of the regularized problems.

The recursion form of (2)-(3) can be written as

$$0 \in x^k - x^{k-1} + \lambda_k \Pi(x^k).$$

However, this ideal form of the method is often impractical, since the exact iteration (2) maybe in many cases require a computation as difficult as solving the original problem $0 \in T(x^*)$. In [8], Rockafellar has given an inexact variant of the method

$$e^k \in x^k - x^{k-1} + \lambda_k \Pi(x^k), \tag{4}$$

where $\{e^k\}$ is regarded as an error sequence. The method is called inexact proximal point algorithm. It was shown that if $e^k \to 0$ quickly enough such that

$$\sum_{k=1}^{+\infty} \|e^k\| < +\infty,$$

then $x^k \to z \in R^n$ with $0 \in \Pi(z)$.

Instead of using the quadratical function (3) as the proximal term, Eckstein [4] investigated the Bregman-function-based proximal method and has proved that the sequence $\{x^k\}$ generated by (4) converges to a root of Π under the following conditions:

$$\sum_{k=1}^{+\infty} \|e^k\| < +\infty \quad \text{and} \quad \sum_{k=1}^{+\infty} \langle e^k, x^k \rangle \text{ exists and is finite.} \tag{5}$$

On the other hand, for quadratic proximal method, Han and He [6] have proved the convergence for recursion (4) under the following accuracy criterion

$$\|e^k\| \le \eta_k \|x^k - x^{k-1}\| \quad \text{with} \quad \sum_{k=0}^{+\infty} \eta_k^2 < +\infty. \tag{6}$$

It seems that the accuracy criterion (6) can be checked and complemented in practice more easily than (5).

Recently, Auslender, Teboulle and Ben-Tiba [1] have proposed a new type of proximal interior methods replacing the quadratic function $q(x, x^{k-1})$, by the logarithmic-quadratic function $D(x, x^{k-1})$ (will be specified in Section 3), this method is called logarithmic-quadratic proximal method. In their inexact version[1], they have suggested to use the accuracy criterion of type (5). Since the accuracy criterion of type (6) is more useful in practice, in this paper, we proposed an accuracy criterion of type (6) for the logarithmic-quadratic proximal method.

2 Preliminaries

We list some important results on maximal monotone operator and some basic properties which will be needed in our following analysis. The domain of T and the graph of T are defined by

$$\text{dom}T := \{x | T(x) \ne \emptyset\} \qquad \text{and} \qquad G(T) := \{(x, y) \in R^n \times R^n : y \in T(x)\}.$$

An operator T is said to be monotone if

$$(x'-x)^T(y'-y)\geq 0 \qquad \forall y'\in T(x'), \qquad \forall y\in T(x), \qquad \forall x,x'\in \text{dom}T.$$

A monotone operator T is said to be maximal if its graph is not properly contained in the graph of any other monotone operator, in other words, if

$$(x-x')^T(y-y')\geq 0 \qquad \forall x'\in \text{dom}T, \qquad \forall y'\in T(x')\Rightarrow y\in T(x).$$

The normal cone operator associated with a closed convex set C is defined by

$$N_C(x)=\begin{cases}\{y: y^T(v-x)\leq 0, \quad \forall v\in C\} & \text{if } \ x\in C\\ \emptyset & \text{otherwise.}\end{cases}$$

Clearly, $\text{dom}N_C=C$ and we have $N_C=\{0\}$ when $C=R^n$ or when $x\in \text{int}C$. Since C is closed and convex then N_C is maximal monotone.

Lemma 2.1. *Let $T_i, i=1,2$, be maximal monotone. If $intdomT_1\cap domT_2\neq\emptyset$, then T_1+T_2 is also maximal monotone.*

Proof. See Chapter 12 of [9]. ■

In the analysis of the proposed algorithm we need the following result, which has already been stated in [1]. Here, for the completeness and the convenience of the readers, we include its proof as well.

Lemma 2.2. *For any $s>0, t>0$ and $u\geq 0$ we have*

$$(t-u)\left(2t-s-\frac{s^2}{t}\right)\geq \frac{3}{2}((u-t)^2-(u-s)^2)+\frac{1}{2}(t-s)^2. \tag{7}$$

Proof. Let δ be the left-hand side of (7), then developing and regrouping terms we obtain

$$\begin{aligned}\delta &= 2t^2-st-s^2-u(2t-s)+u\frac{s^2}{t}\\ &\geq 2t^2-st-s^2-u(2t-s)+u(2s-t).\end{aligned}$$

The above last inequality follows from the property $\frac{s^2}{t}\geq 2s-t$. By a simple manipulation we get

$$\begin{aligned}\delta &\geq s(t-s)+2t(t-s)-3u(t-s)\\ &= (s-t)(3u-2t-s) \qquad (8)\\ &= (t-s)(s-u)+2(t-s)(t-u). \qquad (9)\end{aligned}$$

Using in (9) the identities

$$\begin{aligned}2(t-s)(s-u)&=((u-t)^2-(u-s)^2-(t-s)^2),\\ 2(t-s)(t-u)&=((u-t)^2-(u-s)^2+(t-s)^2),\end{aligned}$$

we obtain (7). ■

3 The logarithmic-quadratic proximal method

Let $\nu > \mu > 0$ be any fixed parameters. For $v \in R^p_+$ define

$$d(u,v) = \begin{cases} \sum_{i=1}^{p} \frac{\nu}{2}(u_i - v_i)^2 + \mu(v_i^2 \log \frac{v_i}{u_i} + u_i v_i - v_i^2) & \text{if} \quad u \in R^p_+ \\ +\infty & \text{otherwise.} \end{cases}$$

It is easy to verify that $d(\cdot, v)$ is a closed proper convex function, nonnegative and $d(u,v) = 0$ if and only if $u = v$. One of the motivations behind the specific form of the function $d(\cdot,\cdot)$ is as follows: The first quadratic term is a usual regularization term used in a proximal method, while the second expression is added to enforce the method to become an interior one, i.e., to generate iterates staying in the positive orthant.

From now on, for simplicity of exposition, we will use $\nu = 2, \mu = 1$. Then, simple algebra shows that d given can be conveniently written as

$$d(u,v) = \begin{cases} \sum_{i=1}^{p} u_i^2 - u_i v_i + v_i^2 \log \frac{v_i}{u_i} & \text{if} \quad u \in R^p_+ \\ +\infty & \text{otherwise.} \end{cases}$$

Let a_i denote the rows of the matrix A, and define the following quantities

$$\begin{aligned} l_i(x) &= b_i - a_i^T x, \\ l(x) &= (l_1(x), l_2(x), \cdots, l_p(x))^T, \\ D(x,y) &= d(l(x), l(y)). \end{aligned}$$

For each $x \in \text{int}C, y \in \text{int}C$, we have

$$\nabla_x D(x,y) = -\sum_{i=1}^{p} a_i \left(2l_i(x) - l_i(y) - \frac{l_i(y)^2}{l_i(x)} \right). \tag{10}$$

Throughout this paper we assume that $\text{dom}T \cap \text{int}C \neq \emptyset$ and the solution set of (VI), denoted by S, is nonempty.

The inexact Logarithmic-Quadratic Proximal (LQP) method

- **Step 0** Given $0 < \lambda_L \leq \lambda_U < \infty$ and $\{\lambda_k\} \subset [\lambda_L, \lambda_U]$, a nonnegative sequence $\{\eta_k\}$ with

$$\sum_{k=0}^{\infty} \eta_k^2 < +\infty. \tag{11}$$

  Start with $x^0 \in \text{int}C$.

- **Step 1** For $k = 1, 2, \ldots$, if $x^{k-1} \notin S$, then generate a pair of $(x^k, e^k) \in \text{int}C \times R^m$ such that

$$g^k + \lambda_k^{-1} \nabla_x D(x^k, x^{k-1}) = e^k \qquad \text{with} \quad g^k \in T(x^k) \tag{12}$$

  and

$$\|e^k\| \leq \eta_k \|A(x^k - x^{k-1})\|. \tag{13}$$

The solvability of problem (12) can be found in [1]. Note that in the exact proximal point algorithm (2), x^k is the root of Π if and only if $x^k = x^{k-1}$. Hence, we can see the distance $\|A(x^k - x^{k-1})\|$ as an "error bound", which measures how much x^k fails to be in the roots set of Π.

We are now in position to establish the main result of this section.

Theorem 3.1. *Let $\{x^k\}$ be the sequence generated by LQP and $\{\eta_k\}, \{e^k\}$ be the sequences satisfy conditions (11) and (13). Then there exists an integer $k_0 \geq 1$ and a constant $c_0 > 0$, such that for all $k \geq k_0$.*

$$\|A(x^k - x^*)\|^2 \leq \left(1 + \frac{c_0\eta_k^2}{1 - c_0\eta_k^2}\right)\|A(x^{k-1} - x^*)\|^2 - \frac{1}{6}\|A(x^k - x^{k-1})\|^2. \tag{14}$$

Furthermore, $\{x^k\}$ is a bounded sequence and

$$\lim_{k\to\infty} \|x^k - x^{k-1}\| = 0.$$

Proof. Let $\Pi = T + N_C$, since $\mathrm{dom}T \cap \mathrm{int}C \neq \emptyset$, it follows from Lemma 2.1 that Π is maximal monotone and that

$$x^* \in S \qquad \Longleftrightarrow \qquad 0 \in \Pi(x^*).$$

Furthermore, since for $x^k \in \mathrm{int}C$, it holds that $N_C = \{0\}$ and we have

$$g^k \in \Pi(x^k).$$

From (10) and (12), we have

$$g^k - e^k = \lambda_k^{-1}\left(\sum_{i=1}^{p} a_i(2l_i(x^k) - l_i(x^{k-1}) - \frac{l_i(x^{k-1})^2}{l_i(x^k)})\right). \tag{15}$$

Using the definition of l_i and the monotonicity of Π, it follows from (15) that $\forall (x,g) \in G(\Pi)$,

$$\lambda_k(x - x^k)^T(g - e^k) \geq \sum_{i=1}^{p}(l_i(x^k) - l_i(x))\left(2l_i(x^k) - l_i(x^{k-1}) - \frac{l_i(x^{k-1})^2}{l_i(x^k)}\right). \tag{16}$$

Take now in (16) $(x,g) = (x^*, 0)$. Applying Lemma 2.2 with $s = l_i(x^{k-1})$, $t = l_i(x^k)$ and $u = l_i(x^*)$, and summing over $i = 1, \cdots, p$, we then obtain

$$\lambda_k(x^k - x^*)^T e^k \geq \frac{3}{2}(\|A(x^k - x^*)\|^2 - \|A(x^{k-1} - x^*)\|^2) + \frac{1}{2}\|A(x^k - x^{k-1})\|^2.$$

Then

$$\|A(x^k - x^*)\|^2 \leq \|A(x^{k-1} - x^*)\|^2 - \frac{1}{3}\|A(x^k - x^{k-1})\|^2 + \frac{2}{3}\lambda_k(x^k - x^*)^T e^k. \tag{17}$$

For $\eta_k > 0$, using Cauchy-Schwartz inequality we have

$$2\lambda_k(x^k - x^*)^T e^k \leq \frac{1}{2\eta_k^2}\|e^k\|^2 + 2\eta_k^2\lambda_k^2\|x^k - x^*\|^2. \tag{18}$$

Since A is of maximal rank and $\{\lambda_k\} \subset [\lambda_L, \lambda_U]$, we can find a constant $c_1 > 0$ such that

$$\lambda_k^2\|x^k - x^*\|^2 \leq c_1\|A(x^k - x^*)\|^2, \quad \forall k \geq 0.$$

Using (13), (18) becomes

$$\frac{2}{3}\lambda_k(x^k - x^*)^T e^k \le \frac{1}{6}\|A(x^k - x^{k-1})\|^2 + c_0\eta_k^2\|A(x^k - x^*)\|^2, \tag{19}$$

where $c_0 = 2c_1/3$. Since $\eta_k \to 0$, there exists $k_0 \ge 1$, such that for all $k \ge k_0, 1 - c_0\eta_k^2 > 0$. Substituting (19) in (17) we obtain

$$\begin{aligned} &\|A(x^k - x^*)\|^2 \\ \le\ & \left(1 + \frac{c_0\eta_k^2}{1 - c_0\eta_k^2}\right)\|A(x^{k-1} - x^*)\|^2 - \frac{1}{6(1 - c_0\eta_k^2)}\|A(x^k - x^{k-1})\|^2 \\ \le\ & \left(1 + \frac{c_0\eta_k^2}{1 - c_0\eta_k^2}\right)\|A(x^{k-1} - x^*)\|^2 - \frac{1}{6}\|A(x^k - x^{k-1})\|^2. \end{aligned}$$

The first part of the theorem is obtained and thus

$$\|A(x^k - x^*)\|^2 \le \left(1 + \frac{c_0\eta_k^2}{1 - c_0\eta_k^2}\right)\|A(x^{k-1} - x^*)\|^2, \qquad \forall k \ge k_0.$$

Since $\sum_{k=0}^{\infty}\eta_k^2 < +\infty$, it follows that

$$C_S := \sum_{k=k_0}^{\infty}\frac{c_0\eta_k^2}{1 - c_0\eta_k^2} < +\infty, \qquad \text{and} \qquad C_P := \prod_{k=k_0}^{\infty}\left(1 + \frac{c_0\eta_k^2}{1 - c_0\eta_k^2}\right) < +\infty,$$

and thus $\{x^k\}$ is bounded. Also from (14) we have

$$\begin{aligned} &\frac{1}{6}\sum_{k=k_0}^{\infty}\|A(x^k - x^{k-1})\|^2 \\ \le\ & \sum_{k=k_0}^{\infty}(\|A(x^{k-1} - x^*)\|^2 - \|A(x^k - x^*)\|^2) + \sum_{k=k_0}^{\infty}\frac{c_0\eta_k^2}{1 - c_0\eta_k^2}\|A(x^{k-1} - x^*)\|^2 \\ \le\ & \|A(x^{k_0-1} - x^*)\|^2 + \sum_{k=k_0}^{\infty}\frac{c_0\eta_k^2}{1 - c_0\eta_k^2}\left(\sup_{k_0\le k<\infty}\|A(x^{k-1} - x^*)\|^2\right) \\ \le\ & (1 + C_S C_P)\|A(x^{k_0-1} - x^*)\|^2 < +\infty. \end{aligned}$$

It follows that

$$\lim_{k\to\infty}\|A(x^k - x^{k-1})\| = 0 \tag{20}$$

and consequently (since A is of maximal rank)

$$\lim_{k\to\infty}\|x^k - x^{k-1}\| = 0.$$

The proof is complete. ■

We are ready to prove our main global convergence result.

Theorem 3.2. *Let $\{x^k\}$ be the sequence generated by LQP and $\{\eta_k\}, \{e^k\}$ be the sequences satisfy conditions (11) and (13). Then the sequence $\{x^k\}$ converges to some x^∞ with $0 \in \Pi(x^\infty)$.*

Proof. Now, let

$$c_{i,k}(x) := (l_i(x^k) - l_i(x))\left(2l_i(x^k) - l_i(x^{k-1}) - \frac{l_i(x^{k-1})^2}{l_i(x^k)}\right).$$

From (16), we have

$$(x - x^k)^T(g - e^k) \geq \lambda_k^{-1}\sum_{i=1}^{p} c_{i,k}(x). \tag{21}$$

Using (8) with $s = l_i(x^{k-1}), t = l_i(x^k)$ and $u = l_i(x)$, for $x \in C$ we obtain that

$$c_{i,k}(x) \geq \{l_i(x^{k-1}) - l_i(x^k)\}\{3l_i(x) - [l_i(x^{k-1}) + 2l_i(x^k)]\}.$$

It follows from (20) that

$$\lim_{k\to\infty}(l_i(x^{k-1}) - l_i(x^k)) = 0, \quad i = 1, \ldots, p.$$

Since $\{\lambda_k\}$ is bounded, we obtain for each x

$$\liminf_{k\to\infty} \lambda_k^{-1}\sum_{i=1}^{p} c_{i,k}(x) \geq 0. \tag{22}$$

From Theorem 3.1, $\{x^k\}$ is bounded and has at least a cluster point. Let x^∞ be a cluster point of the sequence $\{x^k\}$ and the subsequence $\{x^{k_j}\}$ converges to x^∞. Taking $k = k_j$ in (21), using (22), since $\{e^k\}$ converges to zero, passing to the limit in (21), it follows that

$$(x - x^\infty)^T g \geq 0, \qquad \forall (x, g) \in G(\Pi).$$

Since Π is a maximal monotone operator, the above inequality implies that $0 \in \Pi(x^\infty)$, i.e., $x^\infty \in S$. Note that (14) is true for all roots of Π, hence we have

$$\|A(x^k - x^\infty)\|^2 \leq \left(1 + \frac{c_0\eta_k^2}{1 - c_0\eta_k^2}\right)\|A(x^{k-1} - x^\infty)\|^2, \qquad \forall k \geq k_0. \tag{23}$$

Since $\{x^{k_j}\} \to x^\infty$ and

$$\prod_{k=k_0}^{\infty}\left(1 + \frac{c_0\eta_k^2}{1 - c_0\eta_k^2}\right) < +\infty,$$

for any given $\epsilon > 0$, there is an $l > 0$, such that

$$\|A(x^{k_l} - x^\infty)\| < \frac{\epsilon}{2} \qquad \text{and} \qquad \sqrt{\prod_{k=k_l}^{\infty}\left(1 + \frac{c_0\eta_k^2}{1 - c_0\eta_k^2}\right)} < 2. \tag{24}$$

Therefore, for any $k \geq k_l$, it follows from (23) and (24) that

$$\|A(x^k - x^\infty)\| \leq \sqrt{\prod_{t=k_l}^{k-1}\left(1 + \frac{c_0\eta_t^2}{1 - c_0\eta_t^2}\right)}\left(\|A(x^{k_l} - x^\infty)\|\right) < \epsilon.$$

and the sequence $\{x^k\}$ converges to x^∞. ∎

4 Concluding remarks

The present study deals with the logarithmic-quadratic proximal method for solving variational inequality over polyhedral set. It assesses and reconsiders this method in a different way from the one indicated by [1]. The innovative contribution of this study is to show that the sequence $\{x^k\}$ generated by LQP converges to a root of the maximal monotone operator Π under conditions (11) and (13) which can be regarded as a weaker criterion yet can be easily enforced in practice.

References

[1] Auslender A, Teboulle M, Ben-Tiba S. A logarithmic-quadratic proximal method for variational inequalities. Comput. Optim. Appl., 1999, 12: 31-40.

[2] Auslender A, Haddou M. An interior proximal methods for convex linearly constrained problems and its extension to variational inequalities. Math. Prog., 1995, 71: 77-100.

[3] Burachik R S, Iusem A N. A generalized proximal point algorithm for the variational inequality problem in a Hilbert space. SIAM J. Optim., 1998, 8: 197-216.

[4] Eckestein J. Approximate iterations in Bregman-function-based proximal algorithms. Math. Prog., 1998, 83: 113-123.

[5] Guler O. On the convergence of the proximal point algorithm for convex minimization. SIAM J. Control Optim., 1991, 29: 403-419.

[6] Han D R, He B S. A new accuracy criterion for approximate proximal point algorithms. J. Math. Anal. Appl., 2001, 263: 343-256.

[7] He B S. Inexact implicit methods for monotone general variational inequalities. Math. Prog., 1999, 86: 199-217.

[8] Rockafellar R T. Monotone operators and the proximal point algorithm. SIAM J. Control Optim., 1976, 14: 877-898.

[9] Rockafellar R T, Wets R J B. Variational analysis. Springer-Verlag, New York, 1998.

[10] Teboulle M. Convergence of proximal-like algorithms. SIAM J. Optim., 1997, 7: 1069-1083.