Asymptotic convergence analysis of the proximal point algorithm for metrically regular mappings

Shin-ya Matsushita and Li Xu
Department of Electronics and Information Systems
Faculty of Systems Sciences and Technology
Akita Prefectural University
84-4 Ebinokuchi Tsuchiya, Yurihonjo City, Akita, 015-0055 JAPAN
email:matsushita@akita-pu.ac.jp (Shin-ya Matsushita)
email:xuli@akita-pu.ac.jp (Li Xu)

***Abstract*—This paper studies convergence properties of the proximal point algorithm when applied to a certain class of nonmonotone set-valued mappings. We consider an algorithm for solving an inclusion $0 \in T(x)$, where T is a metrically regular set-valued mapping acting from $\mathbb{R}^n$ into $\mathbb{R}^m$. The algorithm is given by the follwoing iteration: $x_0 \in \mathbb{R}^n$ and**

$$x_{k+1} = \alpha_k x_k + (1-\alpha_k) y_k, \quad \textbf{for} \quad k = 0, 1, 2, \ldots,$$

where $\{\alpha_k\}$ is a sequence in $[0,1]$ such that $\alpha_k \leq \bar{\alpha} < 1$, g_k is a Lipschitz mapping from $\mathbb{R}^n$ into $\mathbb{R}^m$ and y_k satisfies the following inclusion

$$0 \in g_k(y_k) - g_k(x_k) + T(y_k).$$

We prove that if the modulus of regularity of T is sufficiently small then the sequence generated by our algorithm converges to a solution to $0 \in T(x)$.

I. Introduction

We deal in this paper with methods for finding zeroes of set-valued mappings in Euclidean spaces, i.e., given Euclidean spaces $\mathbb{R}^n$ and $\mathbb{R}^m$ and a set-valued mapping $T : \mathbb{R}^n \to 2^{\mathbb{R}^m}$, we study the convergence of iterative method for solving the inclusion

$$0 \in T(x). \tag{I.1}$$

Our study is devoted to metrically regular mappings, and we present an algorithm to solve (I.1), which is constructed on the basis of the classical proximal point algorithm [15]. The proximal point was first proposed by Martinet [12] and attained its current formulation in the works of Rockafellar [15], where its connection with the augmented Lagrangian method for constrained nonlinear optimization. In particular, Rockafellar studied the proximal point algorithm for the case when H is a Hilbert space and T is a monotone set-valued mapping from H into itself and showed that when x_{k+1} is an approximate solution of the following proximal point iteration, i.e.,

$$0 \in \mu_k(x_{k+1} - x_k) + T(x_{k+1}) \quad \text{for} \quad k = 0, 1, 2, \ldots, \tag{I.2}$$

and T is maximal monotone, then for a sequence of positive scalars μ_k the iteration (I.2) produces a sequence x_k that is convergent to a solution to $0 \in T(x)$ for any starting point $x_0 \in H$. When T is monotone, i.e.

$$\langle x - y, u - v \rangle \geq 0,$$

for all $x, y \in H$, all $u \in T(x)$ and all $v \in T(y)$, and furthermore maximal monotone, i.e. $T = T^{'}$ whenever $T^{'} : H \to 2^H$ is monotone and $T(x) \subset T^{'}(x)$ for all $x \in H$, it follows from Minty's theorem (see [13]) that $(I + \gamma T)$ is onto and $(I + \gamma T)^{-1}$ is single valued for all positive $\gamma \in \mathbb{R}$, so that the sequence defined by (I.2) is well defined.

In the past three decades, a number of authors have considered generalizations and modifications of the proximal point algorithm and have also found applications of this method to specific variational problems (see, for examples, [3], [14], [9], [16], [8], [10], [11], [1], [2]). In particular, the convergence to a zero point of a maximal monotone set-valued mapping T of the sequence

$$x_{k+1} = \alpha_k x_k + (1-\alpha_k)(I + \gamma_k T)^{-1} x_k \quad \text{for} \quad k = 0, 1, 2, \ldots, \tag{I.3}$$

was observed by Eckstein and Bertsekas [3] (see also [9]), who showed that the sequence $\{x_k\}$ generated by (I.3) converges weakly to a solution $0 \in T(x)$ in the case that $\inf_k \alpha_k > -1$, $\sup_k \alpha_k < 1$ and $\inf_k \gamma_k > 0$.

On the other hand, the situation becomes considerably more complicated when T fails to be monotone. A new approach to the subject was taken in [14], which deals with a class of nonmonotone mappings that, when restricted to a neighborhood of the solution set, are not far from being monotone. More recently, Arágon, Donchev and Geoffroy [1] considered the following proximal point algorithm for a certain class of a nonmonotone set-valued mappings.

$$0 \in g_k(x_{k+1} - x_k) + T(x_k), \quad \text{for} \quad k = 0, 1, 2, \ldots, \tag{I.4}$$

where g_k is a sequence of functions. They proved that if $\bar{x}$ is a solution of (I.1) and the mapping T is metrically regular at $\bar{x}$ for 0 with locally closed graph near $(\bar{x}, 0)$, then there exists a neighborhood O of $\bar{x}$ such that for each initial point $x_0 \in O$ one can find a sequence x_k satisfying (I.4) that is convergent to $\bar{x}$.

In this paper, motivated by (I.3) and (I.4), we will consider the following algorithm for finding zeroes of a metrically regular set-valued mapping. Given $x_0 \in \mathbb{R}^n$, find x_k such that

$$x_{k+1} = \alpha_k x_k + (1-\alpha_k) y_k, \quad \text{for} \quad k = 0, 1, 2, \ldots, \tag{I.5}$$

where $\{\alpha_k\}$ is a sequence in $[0, 1]$ such that $\alpha_k \le \bar{\alpha} < 1$, g_k is a sequence of Lipschitz mappings and y_k satisfies the following inclusion

$$0 \in g_k(y_k) - g_k(x_k) + T(y_k). \tag{I.6}$$

We show that if $\bar{x}$ is a solution of (I.1) and the mapping T is metrically regular at $\bar{x}$ for 0 with locally closed graph near $(\bar{x}, 0)$, then there exists a neighborhood O of $\bar{x}$ such that for each initial point $x_0 \in O$ one can find a sequence x_k satisfying (I.5) that is convergent to $\bar{x}$.

II. Preliminaries

Let $\mathbb{R}^n$ be a Euclidean space, let S be a set-valued mapping from $\mathbb{R}^n$ into the subsets of $\mathbb{R}^m$, denoted $S : \mathbb{R}^n \to 2^{\mathbb{R}^m}$. Let $(\bar{x}, \bar{y}) \in G(S)$. Here, $G(S) = \{(x, y) \in \mathbb{R}^n \times \mathbb{R}^m : y \in S(x)\}$ is the graph of S. Let $A, B \subset \mathbb{R}^n$ and $x \in \mathbb{R}^n$. The distance from a point x to a set A is defined by

$$d(x, A) = \inf_{y \in A} \rho(x, y)$$

and the Hausdorff semidistance from B to A is defined by

$$e(B, A) = \sup_{x \in B} d(x, A).$$

We denote by $B_r(a)$ the closed ball of radius r centered at a, and S^{-1} is the inverse of S defined as $x \in S^{-1}(y) \Leftrightarrow y \in S(x)$. We say that a set A is locally closed at $z \in A$ if there exists $\gamma > 0$ such that the set $A \cap B_\gamma(z)$ is closed.

Let $L > 0$. A mapping $g : \mathbb{R}^n \to \mathbb{R}^m$ is said to be Lipschitz continuous if

$$\|g(x) - g(y)\| \le L\|x - y\| \quad \text{for all} \quad x, y \in \mathbb{R}^n.$$

In this case, L is called the Lipschitz constant of g. The mapping S is said to be *metrically regular* at $\bar{x}$ for $\bar{y}$ if there exists a constant $\kappa > 0$ such that

$$d(x, S^{-1}(y)) \le \kappa d(y, S(x)), \quad \text{for all} \quad (x, y) \quad \text{close to} \quad (\bar{x}, \bar{y}). \tag{II.1}$$

The infimum of κ for which (II.1) holds is the *regularity modulus* denoted $\text{reg}S(\bar{x}|\bar{y})$; the case when S is not metrically regular at $\bar{x}$ for $\bar{y}$ corresponds to $\text{reg}S(\bar{x}|\bar{y}) = \infty$. The inequality (II.1) has direct use in providing an estimate for how far a point x is from being a solutionsh to the variational inclusion $y \in S(x)$; the expression $d(y, S(x))$ measures the residual when $y \notin S(x)$. Smaller values of κ correspond to more favorable behavior. For recent advances on metric regularity and applications to variational problems, see [7], [5] and [6].

We state the following set-valued generalization of the Banach fixed point theorem proved by Donchev and Hager [4] in a complete metric space that we employ to prove our main result (Theorem 3.2).

Lemma 2.1: (Donchev and Hager [4]) *Let (X, ρ) be a complete metric space and $\Phi : X \to 2^X$ be a set-valued mapping. Let $\bar{x} \in X$, $\alpha > 0$ and $0 \le \theta < 1$ such that $\Phi(x) \cap B_\alpha(\bar{x})$ is closed for all $x \in B_\alpha(\bar{x})$ and the following conditions hold:*

(i) $d(\bar{x}, \Phi(\bar{x})) \le \alpha(1 - \theta)$;

(ii) $e(\Phi(u) \cap B_\alpha(\bar{x}), \Phi(v)) \le \theta\rho(u, v)$ *for all* $u, v \in B_\alpha(\bar{x})$.

Then there exists $x_0 \in B_\alpha(\bar{x})$ such that $x_0 \in \Phi(x_0)$.

III. Convergence theorem

First, we recall the algorithm we consider to solve (I.1). Given a starting point x_0, find a sequence x_k by applying the iteration

$$x_{k+1} = \alpha_k x_k + (1-\alpha_k) y_k, \quad \text{for} \quad k = 0, 1, 2, \ldots,$$

where $\{\alpha_k\}$ is a sequence in $[0, 1]$ such that $\alpha_k \le \bar{\alpha} < 1$, g_k is a sequence of Lipschitz mappings and y_k satisfies the follwoing inclusion

$$0 \in g_k(y_k) - g_k(x_k) + T(y_k).$$

The main result of this section reads as follows:

Theorem 3.1: *Let $T : \mathbb{R}^n \to 2^{\mathbb{R}^m}$ be a set-valued mapping and $\bar{x} \in T^{-1}(0)$. Assume that $G(T)$ is locally closed at $(\bar{x}, 0)$ and T is metrically regular at $\bar{x}$ for 0. Choose a sequence of functions $g_k : \mathbb{R}^n \to \mathbb{R}^m$ with $g_k(0) = 0$ which are Lipschitz continuous in a neighborhood U of 0 and Lipschitz constants λ_k satisfying*

$$\sup_k \lambda_k < \frac{1}{2\text{reg}T(\bar{x}|0)}. \tag{III.1}$$

Then there exists a neighborhood O of $\bar{x}$ such that for any $x_0 \in O$ there exists a sequence $\{x_k\}$ generated by (I.5) and (I.6) is well-defined and $\{x_k\}$ converges to $\bar{x}$.

Proof: We first show that well-definedness of the sequence generated by our algorithm.

Let $\lambda = \sup_k \lambda_k$, then from (III.1) there exists $\kappa > \text{reg}T(\bar{x}|0)$ such that $\kappa\lambda < \frac{1}{2}$ and

$$d(x, T^{-1}(y)) \le \kappa d(y, T(x)) \tag{III.2}$$

for all (x, y) close to $(\bar{x}, 0)$. Let $\gamma > 0$ be such that $((\kappa\lambda)^{-1} - 1)^{-1} < \gamma < 1$. From (III.2), there exists $a > 0$ such that the mapping T is metrically regular on $B_a(\bar{x}) \times B_{2\lambda a}(0)$ with constant κ and $B_{2a}(0) \subset U$.

Let $x_0 \in B_a(\bar{x})$. For any $x \in B_a(\bar{x})$, we have

$$\begin{aligned} \|-(g_0(x) - g_0(x_0))\| &= \|g_0(x_0) - g_0(x)\| \\ &\le \lambda_0\|x_0 - x\| \\ &\le 2\lambda_0 a \\ &\le 2\lambda a. \end{aligned}$$

We will show that the mapping $\phi_0(y) = T^{-1}(-(g_0(y) - g_0(x_0)))$ satisfies the assumptions of the fixed point result in Lemma 2.1. First, by using the assumptions that T is metrically regular at $\bar{x}$ for 0, $0 \in T(\bar{x})$ and $g_k(0) = 0$, we have

$$\begin{aligned} d(\bar{x}, \phi_0(\bar{x})) &= d(\bar{x}, T^{-1}(-(g_0(\bar{x}) - g_0(x)))) \\ &\leq \kappa d(-(g_0(\bar{x}) - g_0(x_0)), T(\bar{x})) \\ &\leq \kappa \|g(x_0) - g(\bar{x})\| \\ &\leq \kappa\lambda_0 \|x_0 - \bar{x}\| \\ &\leq \kappa\lambda_0 a \\ &< a(1 - \kappa\lambda_0). \end{aligned}$$

Further, for any $u, v \in B_a(\bar{x})$, by the metric regularity of T,

$$\begin{aligned} & e(\phi_0(u) \cap B_a(\bar{x}), \phi_0(v)) \\ & = \sup_{x \in T^{-1}(-(g_0(u) - g_0(x_0))) \cap B_a(\bar{x})} d(x, T^{-1}(-(g_0(u) - g_0(x_0)))) \\ & \leq \sup_{x \in T^{-1}(-(g_0(u) - g_0(x_0))) \cap B_a(\bar{x})} \kappa d(-(g_0(u) - g_0(x_0)), T(x)) \\ & \leq \kappa \| - (g_0(u) - g_0(x_0)) - (-(g_0(v) - g_0(x_0)))\| \\ & \leq \kappa\lambda_0 \|u - v\|. \end{aligned}$$

To apply Lemma 2.1, it remains to see that the sets $\phi_0(y) \cap B_a(\bar{x})$ are closed for all $y \in B_a(\bar{x})$. Keeping in mind that T is locally closed graph, adjusting a needed, this can be easily shown. Hence by Lemma 2.1, there exists $y_0 \in \phi_0(y_0) \cap B_a(\bar{x})$, that is

$$y_0 \in B_a(\bar{x}) \quad \text{and} \quad 0 \in g_0(y_0) - g_0(x_0) + T(y_0).$$

Let

$$x_1 = \alpha_0 x_0 + (1 - \alpha_0) y_0.$$

For any $x \in B_a(\bar{x})$, we have

$$\begin{aligned} & \| - (g_1(x) - g_1(x_1))\| \\ \leq & \lambda_1 \|x_1 - x\| \\ = & \lambda_1 \|\alpha_0 x_0 + (1 - \alpha_0) y_0 - x\| \\ = & \lambda_1 \|\alpha_0 (x_0 - \bar{x}) + (1 - \alpha_0)(y_0 - \bar{x}) + \bar{x} - x\| \\ \leq & \lambda_1 \{\alpha_0 \|x_0 - \bar{x}\| + (1 - \alpha_0)\|y_0 - \bar{x}\| + \|\bar{x} - x\|\} \\ \leq & 2\lambda_1 a \\ \leq & 2\lambda a. \end{aligned}$$

Let

$$a_1 = \gamma \|x_1 - \bar{x}\|. \tag{III.3}$$

Since $\gamma < 1$, we have $a_1 < a$. We consider the mapping $\phi_1(y) = T^{-1}(-(g_1(y) - g_1(x_1)))$. By (III.3), the metric regularity of T and the choice of γ

$$\begin{aligned} d(\bar{x}, \phi_1(\bar{x})) &\leq d(\bar{x}, T^{-1}(-(g_1(\bar{x}) - g_1(x_1)))) \\ &\leq \kappa d(-(g_1(\bar{x}) - g_1(x_1)), T(\bar{x})) \\ &\leq \kappa \| - (g_1(\bar{x}) - g_1(x_1))\| \\ &\leq \kappa\lambda_1 \|x_1 - \bar{x}\| \\ &\leq a_1(1 - \kappa\gamma). \end{aligned}$$

For $u, v \in B_{a_1}(\bar{x})$, again by the metric regularity of T, we obtain

$$\begin{aligned} & e(\phi_1(u) \cap B_{a_1}(\bar{x}), \phi_1(v)) \\ & = \sup_{x \in T^{-1}(-(g_1(u) - g_1(x_1))) \cap B_{a_1}(\bar{x})} d(x, T^{-1}(-(g_1(u) - g_1(x_1)))) \\ & \leq \sup_{x \in T^{-1}(-(g_1(u) - g_1(x_1))) \cap B_{a_1}(\bar{x})} \kappa d(-(g_1(u) - g_1(x_1)), T(x)) \\ & \leq \kappa \| - (g_1(u) - g_1(x_1)) - (-(g_1(v) - g_1(x_1)))\| \\ & \leq \kappa\lambda_1 \|u - v\|. \end{aligned}$$

Because $\phi_1(y) \cap B_{a_1}(\bar{x})$ is closed for any $y \in B_{a_1}(\bar{x})$, by Lemma 2.1, there exists $y_1 \in \phi_1(y_1) \cap B_{a_1}(\bar{x})$, which by (III.3), satisfies

$$\|y_1 - \bar{x}\| \leq \gamma \|x_1 - \bar{x}\|.$$

Let

$$x_2 = \alpha_1 x_1 + (1 - \alpha_1) y_1.$$

It follows that

$$\begin{aligned} \|x_2 - \bar{x}\| &= \|\alpha_1 x_1 + (1 - \alpha_1) y_1 - \bar{x}\| \\ &\leq \alpha_1 \|x_1 - \bar{x}\| + (1 - \alpha_1)\|y_1 - \bar{x}\| \\ &\leq \alpha_1 \|x_1 - \bar{x}\| + \gamma(1 - \alpha_1)\|x_1 - \bar{x}\| \\ &= (\alpha_1 + \gamma(1 - \alpha_1))\|x_1 - \bar{x}\| \\ &\leq ((1 - \gamma)\bar{\alpha} + \gamma)\|x_1 - \bar{x}\| \end{aligned}$$

The induction step is now clear. Let $x_k \in B_a(\bar{x})$. Then for $\alpha_k = \gamma\|x_k - \bar{x}\|$, by applying Lemma 2.1 to $\phi_k : y \to T^{-1}(-(g_k(y) - g_k(x_k)))$ we obtain the existence of $y_k \in B_{a_k}(\bar{x})$ such that $0 \in g_k(y_k) - g_k(x_k) + T(y_k)$. And hence,

$$\|y_k - \bar{x}\| \leq \gamma \|x_k - \bar{x}\| \quad \text{for all} \quad k = 1, 2, \ldots.$$

Let

$$x_{k+1} = \alpha_k x_k + (1 - \alpha_k) y_k.$$

Thus, we establish that

$$\|x_{k+1} - \bar{x}\| \leq ((1 - \gamma)\bar{\alpha} + \gamma)\|x_k - \bar{x}\|.$$

Since $(1-\gamma)\bar{\alpha}+\gamma < 1-\gamma+\gamma = 1$, the sequence x_k converges to $\bar{x}$. ■

Note that if $\alpha_k = 0$ for all $k = 0, 1, 2, \ldots$, then we can consider the follwoing particular case of (I.5) and (I.6).

$$0 \in g_k(x_{k+1}) - g_k(x_k) + T(x_{k+1}), \quad \text{for} \quad k = 0, 1, 2, \ldots. \tag{III.4}$$

Now, we are able to state the following result.

Theorem 3.2: *Let $T : \mathbb{R}^n \to 2^{\mathbb{R}^m}$ be a set-valued mapping and $\bar{x} \in T^{-1}(0)$. Assume that $G(T)$ is locally closed at $(\bar{x}, 0)$ and T is metrically regular at $\bar{x}$ for 0. Choose a sequence of functions $g_k : \mathbb{R}^n \to \mathbb{R}^m$ with $g_k(0) = 0$ which are Lipschitz continuous in a neighborhood U of 0 and Lipschitz constants λ_k satisfying*

$$\sup_k \lambda_k < \frac{1}{2\text{reg}T(\bar{x}|0)}.$$

Then there exists a neighborhood O of $\bar{x}$ such that for any $x_0 \in O$ there exists a sequence $\{x_k\}$ generated by (III.4) is well-defined and $\{x_k\}$ converges to $\bar{x}$.

ACKNOWLEDGMENT

The first author is supported by the Japan Society for the Promotion of Science as Grand-in-Aid for Young Scientists (B).

REFERENCES

[1] F. J. Aragón Artacho, A. L. Dontchev and M. H. Geoffroy, Convergence of the proximal point method for metrically regular mappings, ESAIM Proc., Vol. 17, pp.1-8, 2007.

[2] F. J. Aragón Artacho and M. H. Geoffroy, Uniformity and inexact version of a proximal method for metrically regular mappings, J. Math. Anal. Appl., Vol. 335, pp.168-183, 2007.

[3] J. Eckstein and D. P. Bertsekas, On the Douglas-Rachford splitting method and the proximal point algorithm for maximal monotone operators, Math. Program., Vol. 55, pp.293-318 1992.

[4] A. L. Donchev and W. W. Hager, An inverse mapping theorem for set-valued maps, Proc. Amer. Math. Soc., Vol.121, pp.481-489, 1994.

[5] A. L. Donchev, A. S. Lewis and R. T. Rockafellar, The radius of metric regularity, Trans. Amer. Math. Soc., Vol.355, pp.493-517, 2002.

[6] A. L. Donchev and R. T. Rockafellar, Regularity and conditioning of solution mappings in variational analysis, Set-Valued Anal., Vol.12, pp.79-109, 2004.

[7] A. D. Ioffe, Metric regularity and subdifferential calculus, Russian Math. Surveys, Vol. 55, pp.501-558, 2000.

[8] A. N. Iusem, T. Pennanen and B. F. Svaiter, Inexact variants of the proximal point algorithm without monotonicity, SIAM J. Optim., Vol.13, pp.1080-1097, 2003.

[9] S. Kamimura and W. Takahashi, Approximating solutions of maximal monotone operators in Hibert spaces, J. Approx. Theory, Vol. 106 pp.226-240, 2000.

[10] S. Kamimura and W. Takahashi, Strong Convergence of a Proximal-Type Algorithm in a Banach Space, SIAM J. Optim., Vol.13, pp.938-945 2002.

[11] S. Kamimura, F. Kohsaka and W. Takahashi, Weak and strong convergence theorems for maximal monotone operators in a Banach space, Set-Valued Anal., Vol.12, pp.417-429, 2004.

[12] B. Martinet, Regularisation, d'inéquations variationelles par approximations succesives, Revue Francaise d 'Informatique et de Recherche Operationelle, Vol.4, pp.154-159, 1970.

[13] G. A. Minty, A theorem on monotone sets in Hilbert spaces, J. Math. Anal. Appl., Vol.11, pp.434-439, 1967.

[14] T. Pennanen, Local convergence of the proximal point algorithm and multiplier methods without monotonicity, Math. Oper. Res., Vol.27, pp.170-191, 2002.

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

[16] M.V. Solodov and B.F. Svaiter, Forcing strong convergence of proximal point iterations in a Hilbert space, Math. Program., Vol.87, pp.189-202, 2000.

[17] W. Takahashi, Convex Analysis and Approximation Fixed Points, Yokohama Publishers, Yokohama, 2000.