On splitting methods for monotone operators *

秋田県立大学・システム科学技術学部 知能メカトロニクス学科 † 松下　慎也
Shin-ya Matsushita
Department of Intelligent Mechatronics, Faculty of Systems Science and Technology
Akita Prefectural University

1　はじめに

次の問題を考える:

$$\text{find } u \in H \text{ s.t. } 0 \in (A+B)(u), \tag{1.1}$$

ただし、H はヒルベルト空間、$A, B\colon H \rightrightarrows H$ は極大単調 (2 章参照) とする。問題 (1.1) は、A と B を下半連続な真凸関数の劣微分に置き換えることで、二つの凸関数の和の最小化問題を表す数理モデルとなり ([1, 4] 参照)、情報学の分野におけるスパースモデリング等と密接に関係している ([8] 参照)。

本論文では、二つの極大単調作用素 A, B の和の零点問題 (1.1) に対する解法である inertial Douglas-Rachford 法 を考える。この方法は以下のように定義される [2]

$$\begin{cases} y_k = x_k + t_k(x_k - x_{k-1}) \\ x_{k+1} = ((1-\alpha_k)I + \alpha_k R_B R_A)(y_k) \quad (k = 0, 1, 2, \dots) \end{cases} \tag{1.2}$$

ただし、$x_{-1}, x_0 \in H$、$\{t_k\}, \{\alpha_k\} \subset (0, 1)$、$I$ は H 上の恒等写像、$R_A = 2J_A - I$ (resp. $R_B = 2J_B - I$))、J_A (resp. J_B) は集合値写像 A (resp. B) に対するリゾルベントと呼ばれ、$J_A(x) = \{u \in H : x \in (I + A)(u)\}$ と定義される。A が極大単調作用素の時、J_A は一価写像、つまり $J_A(x) = (I + A)^{-1}(x)$ となる ([1, 9, 10] 参照)。本論文を通して、A と B のリゾルベント J_A と J_B は容易に計算できると仮定する。

$t_k \equiv 0$ の時、点列 $\{x_k\}$ は Douglas-Rachford 法 [6] と呼ばれ、$\sum_{j=0}^{\infty} \alpha_j(1-\alpha_j) = \infty$ のとき、点列 $\{J_A(x_k)\}$ は問題 (1.1) の解に弱収束することが知られている (例えば、[1, Theorem 25.6] 参照)。また、$t_k \neq 0$ の時、$\{t_k\}$ と $\{\alpha_k\}$ が適切な条件を満たせば点列 $\{J_A(y_k)\}$ は問題 (1.1) の解に

This work was supported by the Ministry of Education, Culture, Sports, Science, and Technology [19K03639] and the Research Institute for Mathematical Sciences
〒015-0055 秋田県由利本荘市土谷字海老ノ口 84-4
http://www.akita-pu.ac.jp/system/elect/sce/matsushita/

弱収束することが知られている ([2, Theorem 8] 参照)。さらに、A または B が強単調を仮定すると、点列 $\{J_A(y_k)\}$ が強収束することも示されている。最近、Dong [5] は [2] で提案された条件とは異なる $\{t_k\}$ と $\{\alpha_k\}$ の条件を用いて $\{x_k\}$ が写像 $R_B R_A$ の不動点に弱収束することを示した。

本論文では、inertial Douglad-Rachford 法 (1.2) から生成された点列の収束性について考える。特に Dong の提案した条件を使うことで、点列 $\{J_A(x_k)\}$ と $\{J_A(y_k)\}$ が問題 (1.1) の解に弱収束することを示す。また、A または B が強単調の場合、$\{J_A(x_k)\}$ と $\{J_A(y_k)\}$ が解に強収束することを示す。

2 準備

本論文を通して H を実ヒルベルト空間とし、$\langle \cdot, \cdot \rangle$ と $\|\cdot\|$ をそれぞれ H の内積とノルムとする。作用素 $A\colon H \rightrightarrows H$ のグラフを $\mathrm{Gr}(A) = \{(x, x^*) | x^* \in A(x)\}$ と定義する。$A\colon H \rightrightarrows H$ が

(1) 単調とは、
$$\langle x - y, x^* - y^* \rangle \geq 0 \ \ ((x, x^*), (y, y^*) \in \mathrm{Gr}(A)); \tag{2.1}$$

(2) 強単調とは、ある定数 $\beta > 0$ が存在して $A - \beta I$ が単調となる。つまり、
$$\langle x - y, x^* - y^* \rangle \geq \beta \|x - y\|^2 \ \ ((x, x^*), (y, y^*) \in \mathrm{Gr}(A)); \tag{2.2}$$

(3) 極大単調とは、
$$A \text{ が単調、} B\colon H \rightrightarrows H \text{ が単調でかつ } \mathrm{Gr}(A) \subset \mathrm{Gr}(B) \ \Rightarrow \ A = B;$$

を満たす時をいう。$u \in H$ とする。$0 \in A(u)$ が成り立つとき、u を A の零点という。また、A の零点全体の集合を $A^{-1}(0)$、つまり $A^{-1}(0) = \{u \in H : 0 \in A(u)\}$ とする。

$f\colon H \to (-\infty, \infty]$ を下半連続な真凸関数とする。f に対する劣微分を以下のように定義する。
$$\partial f(x) = \{x^* \in H : f(y) \geq f(x) + \langle y - x, x^* \rangle \ (\forall x \in H)\}.$$
このとき、$\partial f\colon H \rightrightarrows H$ は極大単調になることが知られている ([1, 9, 10] 参照)。また、劣微分 ∂f に対するリゾルベント $J_{\partial f}$ には次の関係が成り立つ。
$$J_{\partial f}(x) = \operatorname*{argmin}_{y \in H} \left\{ \frac{1}{2}\|y - x\|^2 + f(y) \right\}.$$
このとき、リゾルベントを近接写像といい、prox_f とあらわす。つまり、$\mathrm{prox}_f = J_{\partial f}$。

C を H の空でない閉凸集合とする。集合 C の指示関数 i_C を
$$i_C(x) = \begin{cases} 0 & (x \in C) \\ \infty & (\text{その他}) \end{cases}$$
と定義する。このとき、i_C は下半連続な真凸関数となり、$J_{\partial i_C} = P_C$ が成り立つ。ここで、$P_C\colon H \to C$ は集合 C の上への距離射影といい、
$$P_C(x) = \operatorname*{argmin}_{y \in C} \|x - y\| \ (\forall x \in H)$$

で定義される。

写像 $T\colon H \to H$ と点 $v \in H$ に対して、$T(v) = v$ が成り立つとき、v を写像 T の不動点という。また、T の不動点全体の集合を $\mathrm{Fix}(T)$、つまり $\mathrm{Fix}(T) = \{v \in H : T(v) = v\}$ とする。写像 $T\colon H \to H$ が非拡大とは、

$$\|T(x) - T(y)\| \leq \|x - y\| \ (\forall x, y \in H)$$

が成り立つときをいう。本節で述べた基礎概念については、文献 [1, 9, 10] を参照するとよい。

主定理を得るため、次の結果は重要である。

命題 2.1. ([1]) $A, B\colon H \rightrightarrows H$ を極大単調作用素とする。このとき、以下が成り立つ。

(1) J_A, R_A は非拡大;
(2) $\frac{1}{2}(I + R_B R_A) = J_B(2J_A - I) + (I - J_A)$;
(3) $(A+B)^{-1}(0) = J_A\left(\mathrm{Fix}(R_B R_A)\right)$.

補助定理 2.1. ([1, Corollary 25.5]) $A, B\colon H \rightrightarrows H$ を極大単調作用素とし、点列 $\{(x_k, u_k)\} \subset \mathrm{Gr}(A), \{(y_k, v_k)\} \subset \mathrm{Gr}(B)$ は $x_k \rightharpoonup x$、$u_k \rightharpoonup u$、$y_k \rightharpoonup y$、$v_k \rightharpoonup v$、$u_k + v_k \to 0$ かつ $x_k - y_k \to 0 \ \ (k \to \infty)$ を満たすとする。このとき、$x = y \in (A+B)^{-1}(0)$、$(x, u) \in \mathrm{Gr}(A)$ かつ $(y, v) \in \mathrm{Gr}(B)$ が成り立つ。

定理 2.1. ([5, Theorem 5]) $A, B\colon H \rightrightarrows H$ を極大単調作用素とし、零点問題 (1.1) の解が存在すると仮定する。また、点列 $\{x_k\}$ は (1.2) によって生成された点列とする。ただし、$\{\alpha_k\}$ と $\{t_k\}$ は以下の条件を満たす。

$$\alpha_k \in [0.5, 1-\epsilon], \ t_0 = 0, \ 0 \leq t_k \leq t_{k+1} \leq t \leq \frac{1 - (1+\epsilon)\alpha_k}{1 + \alpha_k}. \tag{2.3}$$

ただし、$\epsilon > 0$ は十分小さい値とする。このとき

(1) $\{x_k\}$ は有界;
(2) $\|x_{k+1} - x_k\| \to 0 \ (k \to \infty)$ が成り立つ;
(3) $\|y_k - R_B R_A(y_k)\| \to 0 \ (k \to \infty)$ が成り立つ;
(4) $\{x_k\}$ はある点 $u \in \mathrm{Fix}(R_B R_A)$ に弱収束する。

注意 2.1. 条件 (2.3) を満たす例として、$\alpha_k \equiv 1/2$、$t_0 = 0$、$0 \leq t_k \leq t_{k+1} \leq t < 1/3$ がある。

3 主結果

Dong によって提案された条件 (2.3) に焦点を当て、問題 (1.1) の解を近似する弱収束列の結果をはじめに示す。

定理 3.1. $A, B\colon H \rightrightarrows H$ を極大単調作用素とし、(1.1) の解が存在すると仮定する。また、点列 $\{x_k\}$ は (1.2) によって生成された点列とする。ただし、$\{t_k\}$ と $\{\alpha_k\}$ は以下の条件を満たす。

$$\alpha_k \in [0.5, 1-\epsilon],\ t_0 = 0,\ 0 \le t_k \le t_{k+1} \le t \le \frac{1-(1+\epsilon)\alpha_k}{1+\alpha_k}. \tag{3.1}$$

ただし、$\epsilon > 0$ は十分小さい値とする。このとき、$\{J_A(x_k)\}$ と $\{J_A(y_k)\}$ は (1.1) の解に弱収束する。

証明. 定理 2.1 (2), (4) と $\{y_k\}$ の定義より、$\{y_k\}$ はある $\overline{y} \in \mathrm{Fix}(R_B R_A)$ に弱収束する。命題 2.1 より、$J_A(\overline{y}) \in (A+B)^{-1}(0)$。命題 2.1 (2) より、

$$\frac{1}{2}(R_B R_A(y_k) - y_k) = J_B(2J_A - I)(y_k) - J_A(y_k). \tag{3.2}$$

また、

$$z_k := J_A(y_k),\ w_k := J_B(2J_A - I)(y_k),\ u_k := y_k - z_k,\ v_k := 2J_A(y_k) - y_k - w_k \tag{3.3}$$

とおく。(3.3) とリゾルベントの定義より

$$(z_k, u_k) \in \mathrm{Gr}(A), (w_k, v_k) \in \mathrm{Gr}(B) \text{ かつ } u_k + v_k = z_k - w_k \tag{3.4}$$

が成り立つ。定理 2.1 (3)、(3.2) および (3.3) より

$$z_k - w_k \to 0\ (k \to \infty). \tag{3.5}$$

一方、J_A の非拡大性と定理 2.1 (1) より、$\{z_k\}$ は有界となる。このとき、ある $\{z_{k_j}\} \subset \{z_k\}$ が存在して $z_{k_j} \rightharpoonup \overline{z}\ \ (j \to \infty)$ となる。これと (3.5) より、$w_{k_j} \rightharpoonup \overline{z}\ \ (j \to \infty)$。また、(3.3) より、

$$u_{k_j} \rightharpoonup \overline{y} - \overline{z}, v_{k_j} \rightharpoonup \overline{z} - \overline{y}\ \ (j \to \infty) \tag{3.6}$$

が成り立つ。ここで補助定理 2.1 より、

$$\overline{z} \in (A+B)^{-1}(0), (\overline{z}, \overline{y} - \overline{z}) \in \mathrm{Gr}(A), (\overline{z}, \overline{z} - \overline{y}) \in \mathrm{Gr}(B). \tag{3.7}$$

これより、$\overline{z} = J_A(\overline{y})$ となる。$\{z_k\}$ の任意の弱収束部分列の収束先が $J_A(\overline{y})$ となることから、$J_A(y_k) \rightharpoonup J_A(\overline{y})\ \ (k \to \infty)$ が示せた。

一方、任意の $x^* \in H$ に対して、

$$\begin{aligned}
|\langle J_A(x_k) - J_A(\overline{y}), x^* \rangle| &\le |\langle J_A(x_k) - J_A(y_k), x^* \rangle| + |\langle J_A(y_k) - J_A(\overline{y}), x^* \rangle| \\
&\le \|x_k - y_k\| \|x^*\| + |\langle J_A(y_k) - J_A(\overline{y}), x^* \rangle| \\
&\le t\|x_k - x_{k-1}\| \|x^*\| + |\langle J_A(y_k) - J_A(\overline{y}), x^* \rangle|
\end{aligned}$$

となることから、$J_A(x_k) \rightharpoonup J_A(\overline{y})\ \ (k \to \infty)$ も示せる。 □

次に、強収束に関する結果を示す。

定理 3.2. $A, B\colon H \rightrightarrows H$ を極大単調作用素とし、(1.1) の解が存在するとする。また、点列 $\{x_k\}$ は (1.2) によって生成された点列とする。ただし、$\{t_k\}$ と $\{\alpha_k\}$ は (3.1) を満たす。A または B が強単調とする。このとき、$\{J_A(x_k)\}$ と $\{J_A(y_k)\}$ は零点問題の解に強収束する。

証明. A または B のどちらかが強単調のとき、零点問題の解 $(u \in (A+B)^{-1}(0))$ は一意である事が示せる [1, Proposition 23.35]。(i) A が強単調を仮定する。定理 3.1 の証明より、$(J_A(\overline{y}), \overline{y} - J_A(\overline{y})) \in \mathrm{Gr}(A)$、(3.3) と A の強単調性より、ある $\beta > 0$ が存在して

$$\beta \|z_k - J_A(\overline{y})\|^2 \leq \langle z_k - J_A(\overline{y}), u_k - (\overline{y} - J_A(\overline{y}))\rangle. \tag{3.8}$$

また、B は単調で $(J_A(\overline{y}), J_A(\overline{y}) - \overline{y}) \in B(v)$ より

$$0 \leq \langle w_k - J_A(\overline{y}), v_k - (J_A(\overline{y}) - \overline{y})\rangle = \langle w_k - J_A(\overline{y}), z_k - w_k - u_k - (J_A(\overline{y}) - \overline{y})\rangle \tag{3.9}$$

(3.8) と (3.9) を加えると、

$$\beta \|z_k - J_A(\overline{y})\|^2 \leq \langle z_k - w_k, u_k - (\overline{y} - J_A(\overline{y}))\rangle + \langle w_k - J_A(\overline{y}), z_k - w_k\rangle.$$

$z_k - w_k \to 0$、$u_k \rightharpoonup \overline{y} - J_A(\overline{y})$、$w_k \rightharpoonup J_A(\overline{y})$ より、$z_k \to J_A(\overline{y})$ $(k \to \infty)$ が得られる。

一方、

$$\begin{aligned}\|J_A(x_k) - J_A(\overline{y})\| &\leq \|J_A(x_k) - J_A(y_k)\| + \|J_A(y_k) - J_A(\overline{y})\| \\ &\leq t\|x_k - x_{k-1}\| + \|J_A(y_k) - J_A(\overline{y})\|.\end{aligned}$$

これより、$J_A(x_k) \to J_A(\overline{y})\ \ (k \to \infty)$ が示せる。

(ii) B が強単調の時。A が強単調の時と同様な議論により強収束が示せる。 □

4 応用

3 章で得られた結果の応用を与える。

4.1 最小化問題

$f, g : H \to (-\infty, \infty]$ を下半連続な真凸関数とする。次の問題を考える:

$$\text{minimize } f(x) + g(x). \tag{4.1}$$

問題 (4.1) は二つの関数の和の最小化問題であり、スパースモデリングにおける正則化項を含む最適化問題 ([8] 参照) を含む一般的な数理モデルである。

問題 (4.1) は以下の劣微分を用いた包含の解を見つける問題と等価になる

$$\text{find } \ u \in H \ \text{ s.t. } \ 0 \in \partial(f+g)(u) \tag{4.2}$$

一方、∂f と ∂g はそれぞれ極大単調作用素であり、劣微分の性質より $(\partial f + \partial g)(x) \subset \partial(f + g)(x) \quad (\forall x \in H)$ が成り立つ。したがって、$A := \partial f$ と $B := \partial g$ と置くことで問題 (1.1) の解は問題 (4.1) の解となる。

3 章の結果を応用して、以下の定理を得ることができる。

定理 4.1. $f, g : H \to (-\infty, \infty]$ を下半連続な真凸関数とし、零点問題 $0 \in (\partial f + \partial g)(u)$ の解 u が存在すると仮定する。点列 $\{x_k\}$ を以下の方法によって生成する。

$$\begin{cases} y_k = x_k + t_k(x_k - x_{k-1}) \\ x_{k+1} = ((1-\alpha_k)I + \alpha_k R_g R_f)(y_k) \quad (k = 0, 1, 2, \ldots) \end{cases} \tag{4.3}$$

ただし、$x_{-1}, x_0 \in H$、$\{t_k\}$ と $\{\alpha_k\} \subset (0, 1)$ は (3.1) を満たし、$R_f = 2\mathrm{prox}_f - I$ (resp. $R_g = 2\mathrm{prox}_f - I$)) とする。このとき、$\{\mathrm{prox}_f(x_k)\}$ と $\{\mathrm{prox}_f(y_k)\}$ は (4.1) の解に弱収束する。

注意 4.1. 零点問題 $0 \in (\partial f + \partial g)(u)$ の解 u の存在は、非拡大写像 $R_g R_f$ の不動点の存在と等価となる。非拡大写像の不動点の存在については、[1, 9, 10] を参照するとよい。

4.2 最良近似問題

C, D を H の空でない閉凸集合、$z \in H$ とする。次の問題を考える:

$$\text{minimize } \frac{1}{2}\|x - z\|^2 + i_C(x) + i_D(x). \tag{4.4}$$

問題 (4.4) は与えられた点 z から最も近い集合 C と集合 D の共通部分 $C \cap D$ に含まれる点を見つける問題 (最良近似問題) である。劣微分を用いることで問題 (4.4) は以下の包含を満たす点 u を見つける問題と等価になる。

$$0 \in (I - z + \partial(i_C + i_D))(u).$$

一方、∂i_C と ∂i_D はそれぞれ極大単調であり、劣微分の性質より $(\partial i_C + \partial i_D)(x) \subset \partial(i_C + i_D)(x)$ $(\forall x \in H)$ が成り立つ。したがって、$A := I - z + \partial i_C$ と $B := \partial i_D$ とすれば、問題 (1.1) の解、つまり

$$0 \in (I - z + \partial i_C + \partial i_D)(u) \tag{4.5}$$

を満たす点 u は問題 (4.4) の解となることがわかる。

3 章の結果を応用して、以下の定理を得ることができる。

定理 4.2. C, D を H の空でない閉凸集合とし、$z \in H$ とする。零点問題 $0 \in (I - z + \partial i_C + \partial i_D)(u)$ の解 u が存在すると仮定する。点列 $\{x_k\}$ を以下の方法によって生成する。

$$\begin{cases} y_k = x_k + t_k(x_k - x_{k-1}) \\ x_{k+1} = \left((1-\alpha_k)I + \alpha_k(2P_D - I)\left(2P_C\left(\frac{1}{2}(\cdot + z)\right) - I\right)\right)(y_k) \quad (k = 0, 1, 2, \ldots) \end{cases} \tag{4.6}$$

ただし、$x_{-1}, x_0 \in H$、$\{t_k\}$ と $\{\alpha_k\} \subset (0, 1)$ は (3.1) を満たし、P_C (resp. P_D)) は C (resp. D) の上への距離射影とする。このとき、$\{P_C\left(\frac{1}{2}(x_k + z)\right)\}$ と $\{P_C\left(\frac{1}{2}(y_k + z)\right)\}$ は (4.4) の解に強収束する。

証明. $I - z + \partial i_C$ が強単調となる事は比較的容易に示す事ができる。また、

$$J_{I-z+\partial i_C}(x) = P_C\left(\frac{1}{2}(x+z)\right) \quad (x \in H)$$

が成り立つ (例えば、[7] 参照)。定理 3.2 より結果を示す事ができる。 $\square$

注意 4.2. 零点問題 $0 \in (I - z + \partial i_C + \partial i_D)(u)$ の解 u の存在性は、$\partial i_C + \partial i_D$ の極大単調性と等価である。単調作用素の和の極大性を保障する条件については、[1, 24 章] を参照するとよい。

参考文献

[1] H. H. Bauschke and P. L. Combettes, *Convex Analysis and Monotone Operator Theory in Hilbert Spaces*, Springer, New York, 2011.

[2] R. I. Bot, E. R. Csetnek and C. Hendrich, *Inertial Douglas-Rachford splitting for monotone inclusion problems*, Applied Math. Comp. 256 472-487 (2015).

[3] P. L. Combettes, *Iterative construction of the resolvent of a sum of maximal monotone operators*, J. Convex Anal. **16** (2009), 727-748.

[4] P. L. Combettes and J.-C. Pesquet, *Proximal splitting methods in signal processing.* in Fixed-Point Algorithms for Inverse Problems in Science and Engineering, Springer Optim. Appl. 49, Springer, New York 185-212 (2011)

[5] Y. Dong, *New inertial factors of the Krasnosel'skii-Mann iteration*, http://www.optimization-online.org/DB_FILE/2019/05/7191.pdf.

[6] P.-L. Lions and B. Mercier, *Splitting algorithms for the sum of two nonlinear operators*, SIAM J. Numer. Anal. **16** (1979), 964-979.

[7] S. Matsushita, *On the resolvent of the sum of maximal monotone operators*, 数理解析研究所講究録, **2112** (2019) 208-212.

[8] 永原正章, スパースモデリングのための凸最適化-近接勾配法による高速アルゴリズム, システム/制御/情報 **61** (2017) 20-28.

[9] W. Takahashi, *Nonlinear functional analysis. fixed points theory and its applications*, Yokohama Publishers, Yokohama 2000.

[10] 高橋渉, 非線形・凸解析学入門, 横浜図書, 2005.