On modified Douglas-Rachford algorithms

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:

$$\text{最小化 } f(x) + g(x), \tag{1}$$

ただし、$\mathcal{H}$ は実ヒルベルト空間、$f, g\colon \mathcal{H} \to \mathbb{R} \cup \{\infty\}$、は下半連続な真凸関数とする。問題 (1) は、凸関数の和の最小化問題を表す数理モデルであり、情報学の分野におけるスパースモデリングや画像処理における画像復元問題等と密接に関係している。

本論文では、凸最小化問題 (1) に対する解法である Douglas-Rachford の分割法について検討する。Douglas-Rachford の分割法は以下のように定義される ([6])。

$$v_{k+1} = \frac{1}{2}\left(I + R_g \circ R_f\right) x_k,\ k = 1, 2, \ldots. \tag{2}$$

ただし、$x_1 \in \mathcal{H}$、$R_f = I - 2\operatorname{prox}_f$、$\operatorname{prox}_f$ は関数 f に対する近接写像と呼ばれ、$\operatorname{prox}_f(x) = \operatorname{argmin}_{y\in\mathcal{H}} \left\{ f(y) + \frac{1}{2}\|y - x\|^2 \right\}$ と定義される。f が下半連続な真凸関数の時、prox_f は一価写像となる ([2, 11, 12] 参照)。本論文を通して、以下の条件を仮定する。

- f と g の近接写像 prox_f と prox_g は容易に計算できる。
- $$(\partial f + \partial g)^{-1}(0) \neq \emptyset^{*1} \tag{3}$$

(2) で生成された点列には以下の性質がある。

- $x^* \in \operatorname{Fix}(R_g \circ R_f)^{*2}$とする。この時以下の性質が成り立つ。
$$\|x_{k+1} - x^*\| \leq \|x_k - x^*\|\ (\forall k) \tag{4}$$
- $\{\operatorname{prox}_f x_k\}$ は問題 (1) の解に弱収束する

(例えば、[2, Theorem 26.11])。性質 (4) が成り立つとき、点列 $\{x_k\}$ は Fejér monotone であると言う [2, Definition 5.1]。

本論文では、Douglas-Rachford の分割法に焦点を当て、(2) の改良について検討する。具体的には、提案手法によって生成された点列の性質および収束性について考察する。

This work was supported by the Ministry of Education, Culture, Sports, Science, and Technology [19K03639] and the Research Institute for Mathematical Sciences
E-mail : matsushita@akita-pu.ac.jp　url : http://www.akita-pu.ac.jp/system/elect/sce/matsushita/

*1 制約想定 (例えば、$0 \in \operatorname{int}(\operatorname{dom} f - \operatorname{dom} g)$ [2, Proposition 15.24]) が成り立つとき、(3) は問題 (1) の解の存在と等価な条件となる。

*2 写像 $T : \mathcal{H} \to \mathcal{H}$ に対して、$\operatorname{Fix}(T)$ を写像 T の不動点全体の集合、つまり $\operatorname{Fix}(T) = \{u \in \mathcal{H} : Tu = u\}$ とする。

2　準備

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

(1) 単調とは、

$$\langle x - y, x^* - y^* \rangle \geq 0 \quad ((x, x^*), (y, y^*) \in \mathrm{Gr}(A)); \tag{5}$$

(2) 極大単調とは、

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

を満たす時をいう。集合値写像 A に対するリゾルベントは $J_A(x) = \{u \in \mathcal{H} : x \in (I + A)(u)\}$ 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 $u \in \mathcal{H}$ 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 $A^{-1}(0) = \{u \in \mathcal{H} : 0 \in A(u)\}$ とする。

$f\colon \mathcal{H} \to (-\infty, \infty]$ 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。

$$\partial f(x) = \{x^* \in \mathcal{H} : f(y) \geq f(x) + \langle y - x, x^* \rangle \ (\forall x \in \mathcal{H})\}.$$

このとき、$\partial f\colon \mathcal{H} \rightrightarrows \mathcal{H}$ は極大単調になることが知られている [11, 2]。また、$u \in (\partial f)^{-1}(0) \Leftrightarrow f(u) = \mathrm{argmin}_{y \in \mathcal{H}} f(y)$ が成り立つ。f に対する近接写像 $\mathrm{prox}_{\gamma f}$ $(\gamma > 0)$ を以下のように定義する。

$$\mathrm{prox}_{\gamma f}(x) = \operatorname*{argmin}_{y \in \mathcal{H}} \left\{ f(y) + \frac{1}{2\gamma} \|y - x\|^2 \right\}.$$

特に、$\mathrm{prox}_{\gamma f} = (I + \gamma \partial f)^{-1}$ が成り立つ。本節で述べた概念については、文献 [11, 2] を参照するとよい。

3　主結果

以下の最適化手法を考える。

$$\begin{cases} \overline{x}_k = x_k + \gamma_k (x_k - x_{k-1}) \\ x_{k+1} = \dfrac{1}{2} (I + R_g \circ R_f) \, \overline{x}_k \end{cases} \tag{6}$$

ここで、$x_0, x_1 \in \mathcal{H}$ で $\{\gamma_k\} \subset (0, \infty)$ とする。

注意 3.1. $\gamma_k \equiv 0$ の時、(6) は Douglas-Rachford の分割法である。(6) は文献 [3, 5] 等で研究された解法で

$$0 = \gamma_0 \leq \gamma_k \leq \gamma_{k+1} \leq \gamma < \frac{1}{3} \ (\forall k)$$

が成り立つ時、$\{x_k\}$ は $R_g \circ R_f$ の不動点、つまり $R_g \circ R_f u = u$ となる $u \in \mathcal{H}$ に弱収束することが知られている ([5, Theorem 3.1])。また、$\{\mathrm{prox}_f\, x_k\}$ は問題 (1) の解に弱収束する ([7])。一方、(6) で生成された点列 $\{x_k\}$ について、性質 (4) を満たさない例が知られている ([9])。

本論文では性質 (4) を満たす (6) の改良について検討する。主定理を紹介する。

定理 3.1. $f, g\colon \mathcal{H} \to \mathbb{R} \cup \{\infty\}$ は下半連続な真凸関数とする。点列 $\{x_k\}$ は以下の方法で生成された点列とする。

$$\begin{cases} \overline{x}_k = \begin{cases} x_k & (k\ :\ \text{偶数}) \\ x_k + \gamma_k(x_k - x_{k-1}) & (k\ :\ \text{奇数}) \end{cases} \\ x_{k+1} = \dfrac{1}{2}\left(I + R_g \circ R_f\right)\overline{x}_k, \end{cases} \tag{7}$$

ただし、

$$0 \le \gamma_k \le \gamma < 1\ (\forall k). \tag{8}$$

このとき、以下が成り立つ。

- 点列 $\{x_{2k}\}$ は Fejér monotone である。
- $\{\operatorname{prox}_f x_k\}$ は問題 (1) の解に弱収束する。

証明の概略. 仮定 (3) より $\operatorname{Fix}(R_g \circ R_f) \neq \emptyset$ となる [2, Proposition 26.1]。$x^* \in \operatorname{Fix}(R_g \circ R_f)$ とする。この時、

$$\left\|\frac{1}{2}\left(I + R_g \circ R_f\right)x - x^*\right\|^2 \le \|x - x^*\|^2 - \left\|\frac{1}{2}\left(I + R_g \circ R_f\right)x - x\right\|^2\ \ (\forall x \in \mathcal{H}) \tag{9}$$

が成り立つ [2, Proposition 4.31]。

一方、$w_k := x_{2k}\ (\forall k)$ とおく。x_k の定義より、

$$\begin{aligned} w_{k+1} = x_{2(k+1)} &= \frac{1}{2}\left(I + R_g \circ R_f\right)(\overline{x}_{2k+1}) \\ &= \frac{1}{2}\left(I + R_g \circ R_f\right)(x_{2k+1} + \gamma_{2k+1}(x_{2k+1} - x_{2k})) \\ &= \frac{1}{2}\left(I + R_g \circ R_f\right)\left(\frac{1}{2}\left(I + R_g \circ R_f\right)x_{2k} + \gamma_{2k+1}\left(\frac{1}{2}\left(I + R_g \circ R_f\right)x_{2k} - x_{2k}\right)\right) \\ &= \frac{1}{2}\left(I + R_g \circ R_f\right)\left(\frac{1}{2}\left(I + R_g \circ R_f\right)w_k + \gamma_{2k+1}\left(\frac{1}{2}\left(I + R_g \circ R_f\right)w_k - w_k\right)\right) \end{aligned}$$

が成り立つ。これと (9) およびノルムの性質[*3]を用いると以下の不等式が得られる。

$$\|w_{k+1} - x^*\|^2 \le \|w_k - x^*\|^2 - (1 + \gamma_{2k+1})(1 - \gamma_{2k+1})\left\|\frac{1}{2}\left(I + R_g \circ R_f\right)w_k - w_k\right\|^2.$$

これより、$\{x_{2k}\}$ は性質 (4) を満たす。この時、$\{x_{2k}\}$ は $\operatorname{Fix}(R_g \circ R_f)$ のある点に弱収束する [2, Theorem 5.5]。また、$x_{2k+1} = \frac{1}{2}\left(I + R_g \circ R_f\right)(x_{2k})$ より $\|x_{2k+1} - x_{2k}\| \to 0$ が成り立つので、$\{x_k\}$ は $\operatorname{Fix}(R_g \circ R_f)$ のある点に弱収束する。

一方、[1] の結果を利用すれば $\{\operatorname{prox}_f x_k\}$ が問題 (1) の解に弱収束することが示せる。

□

注意 3.2.

- $A, B\colon \mathcal{H} \rightrightarrows \mathcal{H}$ を極大単調作用素とする。今回提案した最適化手法 (7) は、次の問題に対する解法として一般化できる。

$$\text{find } u \in \mathcal{H} \text{ such that } 0 \in (A + B)(u). \tag{10}$$

その際、近接写像の代わりに極大単調作用素のリゾルベント (J_A と J_B) を利用する。

[*3] $x, y \in \mathcal{H}, \alpha \in \mathbb{R}$ に対して、$\|(1-\alpha)x + \alpha y\|^2 = (1-\alpha)\|x\|^2 + \alpha\|y\|^2 - \alpha(1-\alpha)\|x - y\|^2$ が成り立つ。

- 問題 (10) は次の問題を含む ([4])。

$$\min_{x\in\mathcal{H}}\left\{f(x)+\sum_{i=1}^{m} g_i\circ L_i(x)\right\} \tag{11}$$

ここで、$\mathcal{H}, \{\mathcal{G}_i\}_{i=1}^m$ は実ヒルベルト空間，$f\colon\mathcal{H}\to\mathbb{R}\cup\{\infty\}$ と $g_i\colon\mathcal{G}_i\to\mathbb{R}\cup\{\infty\}$ は下半連続な真凸関数、$L_i\colon\mathcal{H}\to\mathcal{G}_i$ は有界線形作用素とする。問題 (11) はスパースモデリング ([10]) や画像復元問題 ([4, 8]) で現れる問題を表現できる。

参考文献

[1] H. H. Bauschke. New demiclosedness principles for (firmly) nonexpansive operators. In *Computational and Analytical Mathematics, D. H. Bailey, H. H. Bauschke, P. Borwein, F. Garvan, M. Théra, J. D. Vanderwerff, and H. Wolkowicz, eds.*, pp. 19–28. Springer, Berlin, 2013.

[2] H. H. Bauschke and P. L. Combettes. *Convex Analysis and Monotone Operator Theory in Hilbert Spaces*. Springer, New York, 2nd, edition, 2017.

[3] R. I. Boţ, E. R. Csetnek, and C. Hendrich. Inertial Douglas-Rachford splitting for monotone inclusion problems. *Applied Math. Comp.*, Vol. 256, pp. 472–487, 2015.

[4] R. I. Boţ and C. Hendrich. A Douglas–Rachford type primal-dual method for solving inclusions with mixtures of composite and parallel-sum type monotone operators. *SIAM J Optim.*, Vol. 28, pp. 2541–2565, 2013.

[5] Y. Dong. New inertial factors of the Krasnoselskiĭ-Mann iteration. *Set-Valued Var. Anal.*, Vol. 29, pp. 145–161, 2021.

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

[7] S. Matsushita. On splitting methods for monotone operators. 数理解析研究所講究録, Vol. 2190, pp. 1–7, 2021.

[8] S. Matsushita. Primal-dual splitting algorithms and its applications. 数理解析研究所講究録, Vol. 2194, pp. 89–96, 2021.

[9] Z. Mu and Y. Peng. A note on the inertial proximal point method. *Stat. Optim. Inf. Comput.*, Vol. 3, pp. 241–248, 2015.

[10] M. Nagahara. *Sparsity Methods for Systems and Control*. now publishers, 2020.

[11] W. Takahashi. *Nonlinear Functional Analysis. Fixed Points Theory and its Applications*. Yokohama Publishers, Yokohama, 2000.

[12] 高橋涉. 非線形・凸解析学入門. 横浜図書, 2005.