閉凸集合の不動点表現を用いた
最適化アルゴリズムと信号処理

山田功
東京工業大学 大学院理工学研究科 集積システム専攻

概要

ヒルベルト空間に定義された非拡大写像の不動点集合が閉凸集合になることはよく知られている．凸解析の分野では，凸最適化問題の解集合のように「陽に表現することが困難な閉凸集合」を「"計算可能な非拡大写像" の不動点集合」として表現する研究が急速に進んでいる．これらの成果に「非拡大写像の不動点集合中の 1 点への収束点列を生成するアルゴリズム (Krasnoselskii-Mann の不動点近似法)」や「"非拡大写像の不動点集合を制約集合とする凸最適化問題" のための逐次近似アルゴリズム (ハイブリッド最急降下法)」を応用することにより，これまで見通しのよい解法が知られていなかった「非可微分凸最適化問題」や「階層構造を持つ凸最適化問題」を解決する強力な最適化アルゴリズムが数多く生まれている．これらのアルゴリズムは信号処理分野を中心に各種の逆問題に広く応用され，重要な役割を演じている．本チュートリアルでは，これらの研究動向を筆者が関わった事例を中心に紹介する．

Optimization and Signal Processing
Using Fixed-Point Characterizations of Closed Convex Sets

Isao YAMADA
Department of Communications and Integrated Systems
Tokyo Institute of Technology

Abstract

In this tutorial lecture, we introduce powerful fixed-point theoretic strategies useful to develop many computational algorithms applicable to wide range of inverse problems especially in signal and image processing. These include fixed-point characterizations of the solution set of convex optimization problems, nonsmooth convex optimization algorithms based on Krasnoselskii-Mann iteration, and hierarchical convex optimization algorithms based on Hybrid steepest descent method.

1 はじめに

完備距離空間に定義された 1 未満のリプシッツ定数を持つ写像は縮小写像とよばれている．縮小写像は唯一の不動点を持ち，任意に選ばれた初期点に縮小写像を繰り返し適用することによって生成される点列が不動点に収束することが知られている (縮小写像の不動点定理: Banach 1922). この定理は単純かつ明快であるが故に強力であり，数理科学全体に登場する数多くの逐次近似アルゴズムがこの定理を基本原理としている．一方，1 をリプシッツ定数に持つ写像は非拡大写像とよばれており，その不動点の様相は一変する．例えば，写像 $T_1 : \mathbb{R} \to \mathbb{R} : x \mapsto x + 1$ は

$$|T_1(x) - T_1(y)| = |x - y| \quad (\forall x, y \in \mathbb{R})$$

を満たすので非拡大写像であるが不動点を持たない．ところが別の非拡大写像

$$T_2 : \mathbb{R} \to \mathbb{R} : x \mapsto \begin{cases} -1 & (x < -1) \\ x & (-1 \le x \le 1) \\ 1 & (1 < x) \end{cases}$$

は閉区間 $[-1, 1]$ に属す全ての点を不動点に持つ．これらの例は非拡大写像の不動点の性質が縮小写像の場合ほど単純でないことを示唆している． 幸い「ヒルベルト空間の非拡大写像が空でない不動点集合 (不動点全体として定義される集合) を持つとき, 不動点集合は閉凸集合になる」ことが知られている．さらに, 与えられた非拡大写像の不動点集合中の１点に弱収束する点列を生成するアルゴリズム (Krasnoselskii-Mann の不動点近似法) も知られている．

筆者は信号処理工学を専門にしている．信号処理工学に現れる多くの推定問題や逆問題は凸最適化問題として定式化されてきた．この事実は「凸最適化問題の解集合は極めて工学的利用価値の高い閉凸集合であること」を教えてくれる．筆者らは「凸最適化問題の解集合から得られる情報を最大限に利用した強力な信号処理技術」を確立することを目標にしており, 「凸最適化問題の解集合を精密に表現する方法」と「凸最適化問題の解集合上で更なる最適化を実現する方法」を模索してきた．標準的な凸最適化アルゴリズムの用途が「不特定の最適解の逐次近似」に限定されていることからも想像できるように, 一般に「凸最適化問題の解集合を陽に表現すること」は困難である． 一方， 凸 最適化問題の解集合のような非自明な閉凸集合も「“計算可能な非拡大写像” の不動点集合」になっていれば, このような非拡大写像は「非自明な閉凸集合を精密に表現する手段」として利用できるはずであり, さらには「非自明な閉凸集合から得られる情報を最大限に利用した強力な信号処理技術」の実現にも資することが期待される．実は, 非線形解析や凸解析の成果を駆使することにより, 非自明な閉凸集合を「“計算可能な非拡大写像” の不動点集合」として表現する問題は急速に解明されている．さらに, これらの不動点表現に Krasnoselskii-Mann の不動点近似法やハイブリッド最急降下法を適用することにより, これまで見通しのよい解法が知られていなかった最適化問題を解決する強力なアルゴリズムが実現されている．

本チュートリアルでは, 工学的な観点から「閉凸集合の不動点表現の有用性」を明らかにするとともに, この種の表現を利用した凸最適化アルゴリズムの諸性質と信号処理問題への応用を筆者自身が関わった研究事例を中心に紹介する (注：紙数制限のため参考文献は最小限に留めているが, 文中に簡略化した関連文献情報を付しているので併せて御参照いただければ幸いである).

2 非拡大写像の不動点集合の基本性質と凸最適化アルゴリズム

不動点の存在を仮定する代わりに, 非拡大写像に限定せず, やや一般化された “準非拡大写像” の不動点集合の基本性質を併せて紹介しておく．以下, $\mathcal{H}$ は内積 $\langle \cdot, \cdot \rangle$ と誘導ノルム $\|\cdot\|$ が定義された実ヒルベルト空間とする．写像 $T : \mathcal{H} \to \mathcal{H}$ に対して, $T(z) = z$ を満足する $z \in \mathcal{H}$ が存在するとき, z を「T の不動点 (fixed point)」といい, $\mathrm{Fix}(T) := \{z \in \mathcal{H} \mid T(z) = z\}$ を T の不動点集合 (fixed point set) という．写像 $T : \mathcal{H} \to \mathcal{H}$ が少なくとも１つの不動点を持ち,

$$(\forall x \in \mathcal{H}, \forall z \in \mathrm{Fix}(T)) \quad \|T(x) - z\| \le \|x - z\|$$

を満たすとき, T は準非拡大写像 (quasi-nonexpansive mapping) であるという．準非拡大写像 T の不動点集合 $\mathrm{Fix}(T)$ は, 閉半空間の共通部分集合

$$\mathrm{Fix}(T) = \bigcap_{y \in \mathcal{H}} \left\{ x \in \mathcal{H} \mid \langle y - T(y), x \rangle \le \frac{\|y\|^2 - \|T(y)\|^2}{2} \right\}$$

として表現可能であり, $\mathcal{H}$ の閉凸集合になる. また,

$$(\forall x \notin \mathrm{Fix}(T), \forall z \in \mathrm{Fix}(T)) \quad \|T(x) - z\| < \|x - z\|$$

を満たす写像 T はアトラクト写像 (attracting mapping) であるという. 特に

$$(\exists \alpha > 0, \forall x \in \mathcal{H}, \forall z \in \mathrm{Fix}(T)) \quad \|x - z\|^2 - \|T(x) - z\|^2 \geq \alpha \|x - T(x)\|^2$$

を満たす写像 T は α-強アトラクト写像 (α-strongly attracting mapping) であるという [強アトラクト写像は高性能な適応フィルタリングを実現する鍵となっている (Yamada-Ogura'04a)]. 一方, 準非拡大写像 $T : \mathcal{H} \to \mathcal{H}$ がある $\beta \in [0,1)$ と準非拡大写像 N と恒等写像 $I : \mathcal{H} \to \mathcal{H}$ を用いて $T = (1-\beta)I + \beta N$ と表せるとき, T は β-平均写像 (β-averaged mapping) であるという [特に $\beta \in (0,1)$ のとき, $\mathrm{Fix}(T) = \mathrm{Fix}(N)$ となり, T は $(\frac{1-\beta}{\beta})$-強アトラクト写像になる (Yamada-Ogura'04b)]. 一方, 写像 $T : \mathcal{H} \to \mathcal{H}$ が

$$(\forall x, y \in \mathcal{H}) \quad \|T(x) - T(y)\| \leq \|x - y\|$$

を満たすとき, T は非拡大写像 (nonexpansive mapping) であるという. 不動点を持つ非拡大写像は準非拡大写像になる. 以下に準非拡大写像の不動点集合の基本的な性質を紹介しておく.

性質 1 (準非拡大写像の不動点表現に関する基本性質)

(a) (Yamada-Ogura'04b) 写像 $T_i : \mathcal{H} \to \mathcal{H}$ $(i = 1,2)$ が各々α_i-強アトラクト写像で $\mathrm{Fix}(T_1) \cap \mathrm{Fix}(T_2) \neq \emptyset$ となるとき, T_1 と T_2 の合成写像: $T := T_1 T_2$ は $\left(\frac{\alpha_1 \alpha_2}{\alpha_1 + \alpha_2}\right)$ -強アトラクト写像になり, $\mathrm{Fix}(T) = \mathrm{Fix}(T_1) \cap \mathrm{Fix}(T_2)$ となる. さらに, 任意の $w \in (0,1)$ に対して, T_1 と T_2 の凸結合 $T := wT_1 + (1-w)T_2$ は $\left(\frac{(\alpha_1+1)(\alpha_2+1)}{(1-w)\alpha_1 + w\alpha_2 + 1} - 1\right)$ -強アトラクト写像になり, $\mathrm{Fix}(T) = \mathrm{Fix}(T_1) \cap \mathrm{Fix}(T_2)$ となる.

(b) (Ogura-Yamada'02) $T_i : \mathcal{H} \to \mathcal{H}$ $(i = 1,2)$ の各々が, ある $\alpha_i \in [0,1)$ に対して α_i-平均非拡大であるとき, 合成写像 $T := T_1 T_2$ は $\frac{\alpha_1 + \alpha_2 - 2\alpha_1\alpha_2}{1 - \alpha_1\alpha_2}$-平均非拡大になり, 任意の $w \in (0,1)$ に対して, 凸結合 $T := wT_1 + (1-w)T_2$ は $\{w\alpha_1 + (1-w)\alpha_2\}$-平均非拡大となる.

以下の定理は, その舞台をヒルベルト空間に限定しているものの, 「縮小写像の不動点定理」を平均非拡大写像に拡張した定理と見ることができる.

性質 2 (非拡大写像に対する Krasnoselskii-Mann の不動点近似法の収束定理: 例えば Groetsch'72) 非拡大写像$T : \mathcal{H} \to \mathcal{H}$の不動点集合 $\mathrm{Fix}(T)$ が空でないとき, $\sum_{n \geq 0} \alpha_n (1 - \alpha_n) = \infty$ を満たす任意の実数列 $(\alpha_n)_{n \geq 0} \subset [0,1]$ と任意の初期値 $x_0 \in \mathcal{H}$ を用いて「$x_{n+1} := (1 - \alpha_n)x_n + \alpha_n T(x_n)$ $(n = 0,1,2,\ldots)$」によって生成される点列 $(x_n)_{n=0}^{\infty}$ は, $\mathrm{Fix}(T)$ の 1 点に弱収束する.

閉凸集合 $C \subset \mathcal{H}$ と (Gâteaux) 微分可能な凸関数 $f : \mathcal{H} \to \mathcal{H}$ は $\arg\min_{\boldsymbol{x} \in C} f(x) \neq \emptyset$ を満たし, また, 勾配 (Gâteaux 微分)$\nabla f : \mathcal{H} \to \mathcal{H}$ がリプシッツ連続

$$(\exists \kappa > 0, \forall x, y \in \mathcal{H}) \ \|\nabla f(x) - \nabla f(y)\| \leq \kappa \|x - y\|$$

であると仮定する. このとき, C 上への距離射影 $P_C : \mathcal{H} \to C : x \mapsto \arg\min_{y \in C} \|x - y\|$ と任意の $\mu \in \left(0, \frac{2\gamma}{\kappa}\right] \subset \left(0, \frac{2}{\kappa}\right)$ に対して, $P_C (I - \mu \nabla f) : \mathcal{H} \to C$ は $\frac{1}{2-\gamma}$-平均非拡大写像となり [例えば (Yamada-Ogura-Shirakawa'02, Byrne'04, Yamada-Ogura'04b)],

$$\mathrm{Fix}\left(P_C\left(I - \mu \nabla f\right)\right) = \arg\min_{x \in C} f(\boldsymbol{x})$$

を満たす (注：不動点の存在性は凸最適化問題の解の存在条件から保証される)．したがって, P_C が計算できるとき, 性質 2(Krasnoselskii-Mann の不動点近似法) を写像 $P_C\,(I-\mu\nabla f)$ に直接応用することにより「C 上で f を最小化する問題」の解への弱収束を保証するアルゴリズム (「射影勾配法 (Projected gradient method: 例えば Goldstein'64)」) が構成できる．実は, 射影 Landweber 法, 並列射影法, 交互射影法など, これまで信号画像復元問題に広く応用されてきた数多くの逐次近似アルゴリズムが性質 2 の特別な応用例になっている (例えば Yamada-Ogura-Shirakawa'02, Byrne'04 参照)．さらに「非可微分凸関数の和を最小化する問題」の解への弱収束を保証するいくつかのアルゴリズム (「前方-後方分離型近接点法 (Proximal Forward-Backward algorithm：例えば Mercier'79, Combettes-Wajs'05)」や「Douglas-Rachford 型近接点法 (Douglas-Rachford algorithm: 例えば Lions-Mercier'79, Combettes-Pesquet'07)」) も性質 2 を「近接点写像 (Proximity operator: Moreau'62) を用いて構成された非拡大写像」に応用した例として解釈できる．

筆者らは「非拡大写像の不動点集合を制約集合とする凸最適化問題」を解決するためにハイブリッド最急降下法 (Hybrid steepest descent method)[1] を提案している．幸いハイブリッド最急降下法はこれまで取扱いが困難視されてきた「魅力溢れる閉凸集合」を凸最適化問題の制約集合として利用可能にするため, 信号画像処理分野で応用が広がっている (例えば, Slavakis-Yamada'07, Zhang-Fadili-Starck'08)．以下の定理 1 はハイブリッド最急降下法の最も基本的な収束定理であるが, 筆者らの拡張 (例えば「凸関数 f に課せられた条件の緩和 (Ogura-Yamada'02)」や「T に課せられた条件の準非拡大写像への緩和 (Yamada-Ogura'04b)」) の他, 数多くの数理工学者や数学者によってバナッハ空間への拡張やアルゴリズムの一般化が試みられている (例えば [2] 参照)．

定理 1 (ハイブリッド最急降下法の強収束定理 [1]) 写像 $T:\mathcal{H}\to\mathcal{H}$ は $\mathrm{Fix}(T)\neq\emptyset$ を満たす非拡大写像とし, 凸関数 $f:\mathcal{H}\to\mathcal{H}$ の勾配 $\nabla f:\mathcal{H}\to\mathcal{H}$ は $T(\mathcal{H}):=\{T(x)\mid x\in\mathcal{H}\}$ 上で強単調：

$$(\exists\eta>0,\forall x,y\in T(\mathcal{H}))\ \ \langle\nabla f(x)-\nabla f(y),x-y\rangle\geq\eta\|x-y\|^2,$$

かつリプシッツ連続であるとする． また, ステップサイズ $(\lambda_n)_{n\geq1}\subset[0,\infty)$ は, (i) $\lim\limits_{n\to\infty}\lambda_n=0$, (ii) $\sum\limits_{n\geq1}\lambda_n=\infty$, (iii) $\sum\limits_{n\geq1}|\lambda_n-\lambda_{n+1}|<\infty$ を同時に満たすと仮定する [例えば, $\lambda_n:=\frac{1}{n}$ $(n=1,2,3,\ldots)$ は, 条件 (i)–(iii) を同時に満たす]．このとき, 任意の初期点 $u_0,v_0\in\mathcal{H}$ と

$$u_{n+1}:=T(u_n)-\lambda_{n+1}\nabla f(T(u_n))\ \text{または}\ v_{n+1}:=T\left(v_n-\lambda_{n+1}\nabla f(v_n)\right)$$

によって生成される 2 つの点列 $(u_n)_{n\geq0}$ および $(v_n)_{n\geq0}$ は, 共に $f(u^*)=\min\limits_{u\in\mathrm{Fix}(T)}f(u)$ を満たす唯一の点 $u^*\in\mathrm{Fix}(T)$ に強収束する．

参考文献

[1] I. Yamada, "The hybrid steepest descent method for the variational inequality problem over the intersection of fixed point sets of nonexpansive mappings," pp.473–504, in Inherently Parallel Algorithm for Feasibility and Optimization, (D. Butnariu, Y. Censor and S. Reich, Eds.) Elsevier 2001.

[2] C. Chidume, Geometric Properties of Banach Spaces and Nonlinear Iterations (*Chapter 7: Hybrid steepest descent method for variational inequalities*), Lecture Notes in Mathematics, vol.1965, Springer 2009.