数理解析研究所講究録
第 1972 巻 2015 年 1-22

非拡大写像の不動点表現が拓いた 凸最適化の進化と信号処理への応用

Recent Advances in Convex Optimization with Fixed Point Expression of Nonexpansive Operators and Signal Processing Applications *

山田 功 (Isao Yamada)

東京工業大学・大学院理工学研究科・通信情報工学専攻 / 学術国際情報センター †

Department of Communications and Computer Engineering /

Global Scientific Information and Computing Center (GSIC)

Tokyo Institute of Technology

概 要

近年, 様々な情報表現に現れるスパース性を柔軟に活用するための信号処理技術が急速に発展し, 目覚ましい効果を発揮している. 実はこれらの技術的課題の多くは必ずしも微分可能性を仮定しない特別なクラスの凸最適化問題として定式化されるため, 凸解析と不動点理論の最近の相互作用が生み出した新世代の凸最適化アルゴリズムによってはじめて解決されている. 小文では、筆者が開発に関わった最近の凸最適化アルゴリズムの中から特徴的な事例を不動点近似の立場から紹介すると共にいくつかの応用について簡単に紹介する.

キーワード : 凸最適化, 非拡大写像, 不動点, 近接写像, 主双対分解, Krasnosel'skii-Mann のアルゴリズム, ハイブリッド最急降下法, 画像復元

Keywords: convex optimization, nonexpansive operator, fixed point, proximity operator, primal-dual splitting, Krasnosel'skii-Mann algorithm, Hybrid steepest descent method, image recovery

2000 *Mathematics Subject Classification.* Primary: 47H10, 90C25; Secondary: 47H09

---

*本研究の一部は科学研究費補助金 (B-15H02752) の御支援で実施された.

†e-mail:isao@sp.ce.titech.ac.jp

1 はじめに

1.1 凸最適化と信号処理

適当な実ヒルベルト空間$\mathcal{X}$上に定義された凸関数$\theta:\mathcal{X}\to(-\infty,\infty](:=\mathbb{R}\cup\{\infty\})$を最小化する問題は凸最適化問題 (convex optimization) と総称されている[1]．凸最適化問題は “お椀型” 関数の最小化問題である．局所最適解が自動的に大域的最適解になることが保証されるため，最も健全な最適化問題のクラスとして広く応用されている[2]．実ヒルベルト空間$\mathcal{X}$には，応用対象ごとに所望情報を表現するのに相応しい舞台が選ばれる．また，目的関数θは未知の所望ベクトル$\bar{\boldsymbol{x}}\in\mathcal{X}$に関して利用できる知識に最も整合したベクトルが最小値を達成するように設計される．特に$\bar{\boldsymbol{x}}\in\mathcal{X}$に関するいくつかの知識がある閉凸集合$C\subset\mathcal{X}$を用いて, $\bar{\boldsymbol{x}}\in C$と表せる場合には, 推定値の$C$への所属も要請されるため，制約付き最適化問題

$$\left.\begin{array}{l}\text{Minimize } \theta(\boldsymbol{x})\\ \text{Subject to } \boldsymbol{x}\in C\end{array}\right\} \tag{1.1}$$

として定式化することが常套手段となる[3]．問題 (1.1) は, 関数

$$i_C:\mathcal{X}\ni\boldsymbol{x}\mapsto\begin{cases}0 & \text{if } \boldsymbol{x}\in C\\ \infty & \text{if } \boldsymbol{x}\notin C\end{cases}$$

を用いて，見かけ上制約条件の無い最適化問題

$$\text{Minimize } f(\boldsymbol{x}):=\theta(\boldsymbol{x})+i_C(\boldsymbol{x}) \tag{1.2}$$

として表現されることが多い．応用科学の課題が凸最適化問題に帰着され鮮やかに解決された例は枚挙にいとまがなく，これらの成功例が応用対象の一層の拡大を招いている．ところが，微分可能性すら仮定できない一般の凸関数fに対して，適用可能な実践的アルゴリズムを構築することは決して容易でなく，現在もなお，各問題の特徴を活かす優れた解法が探究されている．特に筆者が専門にしている信号処理の分野では，ガウスの最小二乗推定以来，信号処理問題を凸最適化に帰着させることは1つの定石となっており，凸最適化の新解法やその応用自体が重要なテーマとなっている[4]．

---

[1]関数解析や凸解析の用語に不慣れな読者のために，次節に最小限の準備を用意したので参考にされたい．小文の議論の多くは無限次元ヒルベルト空間を舞台にしても，そのまま成立するが，通常の信号画像処理への応用では有限次元に限定した議論で十分であるため，一部の議論を除き，舞台を有限次元空間に限定している．

[2]最適化問題の困難さの本質的なギャップが 「線形最適化問題 vs 非線形最適化問題」にあるのでなく,「凸最適化問題 vs 非凸最適化問題」にあることは 1960 年代から常識になっている．

[3]勿論，$\mathrm{dom}(\theta):=\{\boldsymbol{x}\in\mathcal{X}\mid\theta(\boldsymbol{x})<\infty\}$ が $\mathrm{dom}(\theta)\cap C\neq\emptyset$ を満たし，解集合 $\arg\min\theta(C):=\{\boldsymbol{x}\in C\mid\theta(\boldsymbol{x})\leq\theta(\boldsymbol{y})\ (\forall\boldsymbol{y}\in C)\}$ が空でないことが前提になる (Fact 1(d) 参照).

[4]IEEE Transactions on Signal Processing の EDICS には”OPT-CVXR Convex optimization and relaxation for SP”がリストアップされている.

例えば，雑音の重畳によって劣化した 256×256 のサイズのグレースケール画像から元画像を復元する信号処理問題を最適化問題に帰着させる場合には，各画像を 65536 個の濃淡値を成分とするベクトルに対応付け，議論の舞台 $\mathcal{X}$ として $\mathbb{R}^{65536}$ が採用される．元画像にある程度の滑らかさが仮定できる場合には急峻な変位にロバストな全変動 (total variation) を抑圧することにより，画像のエッジ情報を維持したまま効果的な雑音除去が可能となるし，画像の特徴抽出に使われる各種の変換域で ℓ_1 ノルムを抑圧することによって，元画像を特徴付けるスパース性が促進できる例が多々知られており，これらの知見は全て関数 θ の設計に利用できる [36]．勿論，各画素値のダイナミックレンジは制約条件を表現する閉凸集合 $C \subset \mathcal{X}$ の設計に利用できる．以上の例で目的関数 θ の設計に利用される全変動や ℓ_1 ノルムや i_C はいずれも大域的な微分可能性が仮定できない凸関数になっていることに注意されたい．もはや「微分がゼロとなるベクトルを探索する古典的戦略」は最適解の特徴付けに十分な役割を果たさないのである．一方，$\mathcal{X}$ を微分可能な小領域に分割する戦略も場合分けを要すため，実践的な解法に結びつくことは稀である．正に，「言うは易く行うは難し」である．このように信号処理への実践的応用に資する凸最適化問題を解決するには，周到に設計されたアルゴリズムによって点列を生成し，最適化問題 (1.1) の解集合 $\arg\min f(\mathcal{X}) (\subset \mathcal{X})$ に所属する 1 点 (1 つの画像に対応する) に収束させる戦略が有効と考えられており，凸解析や不動点理論の知見を駆使した工夫が必要となる [5]．

さらに，画像復元問題用に設計された最適化問題 (1.1) の解集合 $\arg\min f(\mathcal{X})$ に見栄えが顕著に異なる無限個の復元画像が含まれるときには，新たな凸関数 $J : \mathcal{X} \to (-\infty, \infty]$ と制約集合 $\arg\min f(\mathcal{X})$ を用意し「階層構造を持つ凸最適化問題」

$$\left.\begin{array}{l}\text{Minimize } J(\boldsymbol{x})\\ \text{Subject to } \boldsymbol{x} \in \arg\min f(\mathcal{X})\end{array}\right\} \tag{1.3}$$

を解決することが重要な目標になるだろう．ところが，$\arg\min f(\mathcal{X})$ に所属する不特定の 1 点を (点列の極限として) 逐次近似するのがやっとの状況でこの集合 (通常は無限集合) を陽に表現することは現実的でないため，問題 (1.3) の解決には更に非自明なアイディアが必要となる．

上記のような凸最適化問題の解決は，長い間困難視されてきたのであるが，最近，凸解析や不動点理論の知見を駆使することによって，近接写像 (proximity operator) と呼ばれる非線形写像の革新的な活用法が開発されたおかげで，実践的な解法が利用できるようになった．これにより，これまで未開拓だった「凸最適化の定式化」の中に応用価値の高い鉱脈 (例：画像復元，圧縮センシング，ロバスト統計，テンソルの低ランク復元等) が続々と発見されている [36, 19, 51, 29].

[5] 勿論，点列の生成に使われるアルゴリズムは計算可能な操作の有限個の組み合わせで実現されなければならない．

1.2 近接写像がもたらした凸最適化のブレークスルー

以下, 次節以降の準備も兼ねて，凸最適化アルゴリズムにブレークスルーをもたらした近接写像の斬新な活用法の背景を簡単に説明すると共に，問題 (1.1)[または問題 (1.2)] や問題 (1.3) の実用的な解法に至るまでの道筋を大まかに紹介しておく.

凸関数 $f : \mathcal{X} \to (-\infty, \infty]$ と $\boldsymbol{x}^\star \in \mathcal{X}$ に関して

$$(\forall \boldsymbol{x} \in \mathcal{X}) \quad f(\boldsymbol{x}^\star) \leq f(\boldsymbol{x}) \Leftrightarrow \mathbf{0} \in \partial f(\boldsymbol{x}^\star) \subset \mathcal{X} \tag{1.4}$$

が成立する．ただし，各 $\boldsymbol{x} \in \mathcal{X}$ に対して

$$\partial f(\boldsymbol{x}) := \{\boldsymbol{z} \in \mathcal{X} \mid f(\boldsymbol{x}) + \langle \boldsymbol{z}, \boldsymbol{y} - \boldsymbol{x} \rangle \leq f(\boldsymbol{y}) \quad (\forall \boldsymbol{y} \in \mathcal{X})\}$$

は $\boldsymbol{x}$ における f の劣微分 (subdifferential) といい，$(\boldsymbol{x}, f(\boldsymbol{x})) \in \mathcal{X} \times \mathbb{R}$ における f の接平面として定義される全ての関数の勾配の集合になる [3]. 特に f が $\boldsymbol{x}$ で微分可能である場合には，$\partial f(\boldsymbol{x}) = \{\nabla f(\boldsymbol{x})\}$ となるため，条件 (1.4) は，f の微分可能性が仮定できる場合に広く使われている「最適解 $\boldsymbol{x}^\star \in \mathcal{X}$ の特徴付け : $\nabla f(\boldsymbol{x}^\star) = \mathbf{0}$」の一般化になっている.

まず，最適解の集合 $\arg\min f(\mathcal{X})$ を特徴付ける条件 (1.4) を用いて，$\arg\min f(\mathcal{X})$ 中の 1 点の逐次近似アルゴリズムを構築するための大まかな方針とその背景にある考え方を見ていこう．条件 (1.4) に現れる $\partial f : \mathcal{X} \to 2^{\mathcal{X}}$ は ($\nabla f : \mathcal{X} \to \mathcal{X}$ と異なり) 集合値関数であるから，これをそのままアルゴリズム中の計算操作として利用することは難しい．そこで，$\mathcal{X}$ 上の恒等写像 $\mathrm{I}_{\mathcal{X}}$ を用いて集合値写像 $\mathrm{I}_{\mathcal{X}} + \partial f : \mathcal{X} \to 2^{\mathcal{X}}$ の逆写像 [∂f のレゾルベント (resolvent) という]

$$(\mathrm{I}_{\mathcal{X}} + \partial f)^{-1} : \mathcal{X} \to 2^{\mathcal{X}} : \boldsymbol{x} \mapsto \{\boldsymbol{y} \in \mathcal{X} \mid \boldsymbol{x} \in (\mathrm{I}_{\mathcal{X}} + \partial f)(\boldsymbol{y})\}$$

を考える．実は, 任意の $\boldsymbol{x} \in \mathcal{X}$ に対して, $(\mathrm{I}_{\mathcal{X}} + \partial f)^{-1}(\boldsymbol{x})$ は単集合となり，単価写像 (f の近接写像: proximity operator という)

$$\mathrm{prox}_f : \mathcal{X} \to \mathcal{X} : \boldsymbol{x} \mapsto (\mathrm{I}_{\mathcal{X}} + \partial f)^{-1}(\boldsymbol{x}) = \arg\min_{\boldsymbol{y} \in \mathcal{X}} \left[f(\boldsymbol{y}) + \frac{1}{2}\|\boldsymbol{y} - \boldsymbol{x}\|^2 \right] \tag{1.5}$$

が矛盾なく定義されることが知られている．簡単な議論から

$$\begin{aligned}
& (\forall \boldsymbol{x} \in \mathcal{X}) \quad f(\boldsymbol{x}^\star) \leq f(\boldsymbol{x}) \\
\Leftrightarrow \quad & \mathbf{0} \in \partial f(\boldsymbol{x}^\star) \Leftrightarrow \boldsymbol{x}^\star \in (\mathrm{I}_{\mathcal{X}} + \partial f)(\boldsymbol{x}^\star) \\
\Leftrightarrow \quad & \boldsymbol{x}^\star = (\mathrm{I}_{\mathcal{X}} + \partial f)^{-1}(\boldsymbol{x}^\star) \Leftrightarrow \boldsymbol{x}^\star = \mathrm{prox}_f(\boldsymbol{x}^\star) \\
\Leftrightarrow \quad & \boldsymbol{x}^\star \in \mathrm{Fix}(\mathrm{prox}_f) := \{\boldsymbol{z} \in \mathcal{X} \mid \mathrm{prox}_f(\boldsymbol{z}) = \boldsymbol{z}\}
\end{aligned} \tag{1.6}$$

が得られる．関係 (1.6) は，凸関数 f の最小点 (minimizer) が，写像 prox_f の不動点として完全に特徴付けられることを示している. さらに prox_f は $\frac{1}{2}$-平均非拡大写

像 (3.1 節参照) になっているため，Krasnosel'skii-Mann のアルゴリズム (Fact2(a)) が応用でき，任意の初期値 $\boldsymbol{x}_0 \in \mathcal{X}$ から

$$\boldsymbol{x}_{n+1} = \mathrm{prox}_f(\boldsymbol{x}_n) \quad (n = 0, 1, 2, \ldots) \tag{1.7}$$

によって生成される点列 $(\boldsymbol{x}_n)_{n=0}^{\infty}$ が $\mathrm{Fix}(\mathrm{prox}_f)$ の 1 点に弱収束することも示される．

以上は，Proximal Point Algorithm [32] の基本戦略に他ならないが，$\boldsymbol{x}_n$ が (1.7) に従って更新される度に $\mathrm{prox}_f(\boldsymbol{x}_n)$ の計算，すなわち，新たな狭義凸関数 $f(\cdot) + \frac{1}{2}\|\cdot - \boldsymbol{x}_n\|^2$ を (本来最小化したい関数 f の代わりに) 最小化することを要請している．残念ながら，以下の実例で見るように信号処理や逆問題への応用で最小化が望まれる目的関数 f の近接写像は，多くの場合簡単には計算できず，アルゴリズム (1.7) 自体には実践的応用価値は殆どない．ところが凸最適化アルゴリズムの最近の発展を調べてみると，それらの背景には共通に「計算可能な複数の近接写像を用いることにより，凸最適化問題を非拡大写像の不動点の逐次近似問題に帰着する大きな戦略」を見ることができる．アルゴリズム (1.7) はそのような発展の方向を決定づけた金字塔として，重要な意味を持っているのである．近接写像 prox_f の計算が容易にできる空間 $\mathcal{X}$ と関数 f の例を見ておこう (その他の有用な例については, 例えば，[11] を参照されたい).

例 1. (近接写像の例 [3, 11])

(a) (凸射影) 一般に実ヒルベルト空間 $\mathcal{H}$ の空でない閉凸集合 $C \subset \mathcal{H}$ と任意の $\boldsymbol{x} \in \mathcal{H}$ に対して，

$$\|\boldsymbol{x} - P_C(\boldsymbol{x})\| = \min_{\boldsymbol{z} \in C} \|\boldsymbol{x} - \boldsymbol{z}\|$$

を満たす唯一の点 $P_C(\boldsymbol{x}) \in C$ が存在する [49]．$\boldsymbol{x} \in \mathcal{H}$ に $P_C(\boldsymbol{x}) \in C$ を対応させる写像を C 上への凸射影 (あるいは距離射影) という．P_C は，$P_C(\boldsymbol{x}) = \mathrm{prox}_{\gamma i_C}(\boldsymbol{x})$ $(\forall \gamma \in (0, \infty), \forall \boldsymbol{x} \in \mathcal{H})$ を満たしており, 凸関数 i_C に対する近接写像となっている．凸射影の計算しやすさは C の形状に依存する．特に, 有限次元部分空間 (又はその直交補空間) 上への凸射影は計算可能な直交射影 (線形写像) になるため最小二乗推定法の基本原理になっている．

(b) (ソフト閾値法 [14]) m 次元実ヒルベルト空間 $\mathcal{H}$ の正規直交基底 $\{\boldsymbol{e}_k\}_{k=1}^{m}$ と正数 $\omega_k > 0$ $(k = 1, 2, \ldots, m)$ を用いて定義された凸関数

$$\psi : \mathcal{H} \to [0, \infty) : \boldsymbol{x} \mapsto \sum_{k=1}^{m} \omega_k |\langle \boldsymbol{x}, \boldsymbol{e}_k \rangle|$$

の近接写像はソフト閾値法 (soft-thresholding) の操作

$$\mathrm{prox}_{\psi}(\boldsymbol{x}) = \sum_{k=1}^{m} \mathrm{sgn}(\langle \boldsymbol{x}, \boldsymbol{e}_k \rangle) \max\{|\langle \boldsymbol{x}, \boldsymbol{e}_k \rangle| - \omega_k, 0\} \boldsymbol{e}_k \tag{1.8}$$

によって実現される[6]. 関数 ψ は「重み付き ℓ_1 ノルム」とよばれており，特に $\omega_k = 1$ $(k = 1, \ldots, m)$ とした ℓ_1 ノルムは「ℓ_0 準ノルム (ベクトル中の非ゼロ成分の数え上げ)」の下界の中で最大の凸関数になることから，多様な情報表現に現れるスパース性を積極的に活用する最近のデータ解析手法の中で重要な役割を担っている[7].

例 1 で見たような「prox_f が容易に計算できる凸関数 f」は特に「proximable な関数」と総称されており，最近の凸最適化アルゴリズムのブレークスルーに大きく貢献しているのであるが，残念ながら信号処理や逆問題への応用で proximable な関数自体の最小化が最終目標になることは殆どない．例えば, 例 1(b) で見た m 次元実ヒルベルト空間 $\mathcal{H}$ で ℓ_1 ノルムを最小化する問題の解は自明なゼロベクトルとなり，この最小化問題を考えても実用的な価値はない．ℓ_1 ノルムは他の凸関数 (例えば i_C) との和を最小化する問題の中ではじめて応用価値を持つのである．実際に，最近の信号処理や逆問題では「proximable な凸関数」や「線形写像と proximable な凸関数の合成」や「リプシッツ連続な勾配を持つ微分可能な凸関数」の有限和で表現される凸関数 f の最小化が要請されることが多い．また，以下の性質 1 はアルゴリズム (1.7) の応用可能性を著しく制限している．

性質 1. (a) 一般に, 凸関数 $\varphi_j : \mathcal{X} \to (-\infty, \infty]$ $(j = 1, 2, \ldots, p)$ の各々が，proximable な凸関数であったとしても，それらが互いに変数分離されていない限り，$\sum_{j=1}^{p} \varphi_j$ は proximable にならない．

(b) 2 つのヒルベルト空間 $\mathcal{X}$ と $\mathcal{Y}$ の間の有界線形写像 $A : \mathcal{X} \to \mathcal{Y}$ と proximable な凸関数 $\varphi : \mathcal{Y} \to (-\infty, \infty]$ を用いて定義された凸関数 $\varphi \circ A : \mathcal{X} \to (-\infty, \infty]$ は一般に proximable にならない．

幸い，最近になって「proximable な凸関数」や「線形写像と proximable な凸関数の合成」や「リプシッツ連続な勾配を持つ微分可能な凸関数」の有限和の最小化問題の解集合を「計算可能な非拡大写像」の不動点集合によって表現する革新的なアイディア (主双対分解 (primal-dual splitting)) が発明され，凸最適化にブレークスルーがもたらされた．実際に，この種の非拡大写像に Krasnosel'skii-Mann algorithm を適用することにより，広いクラスの問題 (1.2) が解決できるようになっており，信号処理や逆問題への実践的応用で目覚ましい効果を発揮している．さらに，Krasnosel'skii-Mann のアルゴリズムの代わりにハイブリッド最急降下法 (Fact2(b)) を応用することにより，広いクラスの「階層構造を持つ凸最適化問題 (1.3)」も解決できるようになっている[8].

[6]実際に計算してみると「$\mathrm{prox}_\psi(\boldsymbol{x})$ を計算する問題は成分毎に 1 変数関数を最小化する問題に分解でき，1 変数関数の最小化問題に帰着できる」ことが容易に確認できる．これは例 3(b) に述べる変数分離型関数の性質に他ならない．

[7]同様な考え方で「行列のランク関数」の下界の中で最大の凸関数が核ノルム (特異値の和) で達成されることも知られている．核ノルムの近接写像は特異値ベクトルのソフト閾値処理によって計算可能であり，広く逆問題に応用されている (例えば [19, 28] と参考文献を参照されたい).

[8]小文で説明できなかった多くの例を [51, 29] に紹介しているので，併せて参照されたい．

2 準備

2.1 ヒルベルト空間，ガトー微分

内積 (inner product) が定義された (実) ベクトル空間を (実) 内積空間 (inner product space) という．ベクトル空間 $\mathcal{X}$ の内積は初めから決まっているものでなく，一定の条件 [9] さえ満たせば，$\mathcal{X}$ を舞台として展開される議論ごとに都合のよい内積を定義しても一向に構わない [10]．完備な内積空間 [11] は (実) ヒルベルト空間であるという (ヒルベルト空間には頭文字を用いた表記 $\mathcal{H}$ がしばしば使われる)．2 つのヒルベルト空間 $(\mathcal{X}_1, \langle\cdot,\cdot\rangle_{\mathcal{X}_1}, \|\cdot\|_{\mathcal{X}_1})$ と $(\mathcal{X}_2, \langle\cdot,\cdot\rangle_{\mathcal{X}_2}, \|\cdot\|_{\mathcal{X}_2})$ の間に定義された有界線形作用素 $Q : \mathcal{X}_1 \to \mathcal{X}_2$ に対して [12]，「$(\forall \boldsymbol{x} \in \mathcal{X}_1, \forall \boldsymbol{y} \in \mathcal{X}_2)\ \langle Q\boldsymbol{x}, \boldsymbol{y}\rangle_{\mathcal{X}_2} = \langle \boldsymbol{x}, Q^*\boldsymbol{y}\rangle_{\mathcal{X}_1}$」を満足する唯一の有界線形作用素 $Q^* : \mathcal{X}_2 \to \mathcal{X}_1$ は Q の共役作用素であるという．特に，$\mathcal{X}_1 = \mathcal{X}_2 = \mathcal{X}$ で $Q^* = Q$ となるとき，Q は自己共役作用素であるという．以下によく使われる実ヒルベルト空間の構成法を紹介しておく [49]．

例 2. (a) (複数の実ヒルベルト空間の直積空間)
p 個の実ヒルベルト空間 $(\mathcal{X}_j, \langle\cdot,\cdot\rangle_j, \|\cdot\|_j)$ $(j = 1, \ldots, p)$ が与えられるとき，$\mathcal{X}_1 \times \cdots \times \mathcal{X}_p := \{(\boldsymbol{x}_1, \ldots, \boldsymbol{x}_p) \mid \boldsymbol{x}_j \in \mathcal{X}_j\ (j = 1, \ldots, p)\}$ は，成分毎の加算とスカラー倍演算の下でベクトル空間となり，$(\mathcal{X}_1 \times \cdots \times \mathcal{X}_p) \times (\mathcal{X}_1 \times \cdots \times \mathcal{X}_p)$ 上に定義された関数 $\langle(\boldsymbol{x}_1, \ldots, \boldsymbol{x}_p), (\boldsymbol{y}_1, \ldots, \boldsymbol{y}_p)\rangle := \sum_{j=1}^{p} \langle \boldsymbol{x}_j, \boldsymbol{y}_j\rangle_j$ は，$\mathcal{X}_1 \times \cdots \times \mathcal{X}_p$ 上の内積としての条件を満たすので，誘導ノルム $\|(\boldsymbol{x}_1, \ldots, \boldsymbol{x}_p)\| := \sqrt{\sum_{j=1}^{p} \|\boldsymbol{x}_j\|_j^2}$ も定義され，$(\mathcal{X}_1 \times \cdots \times \mathcal{X}_p, \langle\cdot,\cdot\rangle, \|\cdot\|)$ は新たな実ヒルベルト空間になる．

(b) (自己共役作用素を用いた内積を持つ実ヒルベルト空間)
実ヒルベルト空間 $(\mathcal{X}, \langle\cdot,\cdot\rangle, \|\cdot\|)$ に定義された自己共役作用素 $Q : \mathcal{X} \to \mathcal{X}$ に対して

$$(\exists \kappa > 0)(\forall \boldsymbol{x} \in \mathcal{X}) \quad \langle \boldsymbol{x}, Q\boldsymbol{x}\rangle \geq \kappa \|\boldsymbol{x}\|^2$$

[9] 実数体 $\mathbb{R}$ 上のベクトル空間 $\mathcal{X}$ の任意の元 $\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z} \in \mathcal{X}$ と任意の $\alpha, \beta \in \mathbb{R}$ に対して，以下の 3 条件を満足する関数 $\langle\cdot,\cdot\rangle : \mathcal{X} \times \mathcal{X} \to \mathbb{R}$ が定まるとき，$\langle \boldsymbol{x}, \boldsymbol{y}\rangle$ は $\boldsymbol{x}$ と $\boldsymbol{y}$ の内積であるという．

(i) $\langle \boldsymbol{x}, \boldsymbol{y}\rangle = \langle \boldsymbol{y}, \boldsymbol{x}\rangle$

(ii) $\langle \boldsymbol{x}, \boldsymbol{x}\rangle \geq 0$ であり，$\langle \boldsymbol{x}, \boldsymbol{x}\rangle = 0 \Leftrightarrow \boldsymbol{x} = \boldsymbol{0}$

(iii) $\langle \alpha\boldsymbol{x} + \beta\boldsymbol{y}, \boldsymbol{z}\rangle = \alpha\langle \boldsymbol{x}, \boldsymbol{z}\rangle + \beta\langle \boldsymbol{y}, \boldsymbol{z}\rangle$

さらに，$\|\boldsymbol{x}\| := \sqrt{\langle \boldsymbol{x}, \boldsymbol{x}\rangle}$ $(\boldsymbol{x} \in \mathcal{X})$ で定義された関数 $\|\cdot\| : \mathcal{X} \to [0, \infty)$ を内積 $\langle\cdot,\cdot\rangle$ から誘導されたノルムであるという．$(\mathcal{X}, \langle\cdot,\cdot\rangle, \|\cdot\|)$ を内積空間という．

[10] 実際に，凸最適化問題 (1.1) を考える際には，都合のよい「内積空間」を周到に設定することは極めて重要である．

[11] $\mathcal{X}$ の任意のコーシー列が $\mathcal{X}$ 中の 1 点に収束することが保証されるとき，$\mathcal{X}$ は完備であるという．有限次元内積空間は全て実ヒルベルト空間となる [49]．

[12] 線形作用素 Q が $\|Q\|_{\mathrm{op}} := \sup_{\|\boldsymbol{x}\|_{\mathcal{X}_1}=1} \|Q\boldsymbol{x}\|_{\mathcal{X}_2} < \infty$ を満たすとき，Q は有界線形作用素であるといい，$\|Q\|_{\mathrm{op}}$ を Q の作用素ノルムとよぶ．

が成り立つとき，

$$\langle\cdot,\cdot\rangle_Q : \mathcal{X}\times\mathcal{X}\to\mathbb{R} : (\boldsymbol{x},\boldsymbol{y})\mapsto\langle\boldsymbol{x},Q\boldsymbol{y}\rangle$$

はベクトル空間 $\mathcal{X}$ の内積の条件を満たし，誘導ノルム $\|\boldsymbol{x}\|_Q := \sqrt{\langle\boldsymbol{x},\boldsymbol{x}\rangle_Q}$ ($\boldsymbol{x}\in\mathcal{X}$) によって新しい実ヒルベルト空間 $(\mathcal{X},\langle\cdot,\cdot\rangle_Q,\|\cdot\|_Q)$ が定義される．

開集合 $U\subset\mathcal{X}$ 上の関数 $f:U\to\mathbb{R}$ と $\boldsymbol{x}\in U$ に対して

$$\lim_{\delta\to 0}\frac{f(\boldsymbol{x}+\delta\boldsymbol{h})-f(\boldsymbol{x})}{\delta}=\langle a(\boldsymbol{x}),\boldsymbol{h}\rangle\qquad(\forall\boldsymbol{h}\in\mathcal{X})$$

を満たす $a(\boldsymbol{x})\in\mathcal{X}$ が存在するとき，f は $\boldsymbol{x}$ でガトー微分可能 (Gâteaux differentiable) であるといい, $\nabla f(\boldsymbol{x}) := a(\boldsymbol{x})$ を $\boldsymbol{x}$ における f のガトー微分 (または勾配) という．ガトー微分 $\nabla f(\boldsymbol{x})$ の定義は $\mathcal{X}$ に定義される内積によって変わるので注意が必要である．

2.2 凸解析

Fact 1. (凸解析の基本事項 [3]) 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であるとする．

(a) (下半連続関数) 関数 $f:\mathcal{H}\to(-\infty,\infty](:=\mathbb{R}\cup\{\infty\})$ と $a\in\mathbb{R}$ によって定義される集合

$$\mathrm{lev}_{\le a}f := \{x\in\mathcal{H}\mid f(\boldsymbol{x})\le a\}$$

が全ての $a\in\mathbb{R}$ に対して閉集合となるとき，f は $\mathcal{H}$ 上で下半連続 (lower semi-continuous) であるという．

(b) (凸関数, 真凸関数) 関数 $f:\mathcal{H}\to(-\infty,\infty]$ が任意の $\boldsymbol{x},\boldsymbol{y}\in\mathcal{H}$ と $\lambda\in[0,1]$ に対して

$$f(\lambda\boldsymbol{x}+(1-\lambda)\boldsymbol{y})\le\lambda f(\boldsymbol{x})+(1-\lambda)f(\boldsymbol{y})$$

を満たすとき，f は $\mathcal{H}$ 上の凸関数 (convex function) であるという．特に, $\mathrm{dom}(f)\neq\emptyset$ となるとき，f は $\mathcal{H}$ 上で真凸関数 (proper convex function) であるという．$\mathcal{H}$ 上で下半連続な真凸関数全体の集合を $\Gamma_0(\mathcal{H})$ と表わす．

(c) 関数 $f\in\Gamma_0(\mathcal{H})$ に対して定義される

$$f^* : \mathcal{H}\to(-\infty,\infty] : \boldsymbol{y}\mapsto\sup_{\boldsymbol{x}\in\mathcal{H}}\left[\langle\boldsymbol{y},\boldsymbol{x}\rangle-f(\boldsymbol{x})\right]$$

は $f^*\in\Gamma_0(\mathcal{H})$ となり，f の共役関数 (conjugate function/ Fenchel-Rockafellar conjugate function) であるといい，$f^{**} := (f^*)^* = f$ が成立する．

(d) (凸集合) C を $\mathcal{H}$ の部分集合とする．$\boldsymbol{x}_1, \boldsymbol{x}_2 \in C$ であるとき，全ての $\lambda \in [0,1]$ に対して，$\lambda \boldsymbol{x}_1 + (1-\lambda)\boldsymbol{x}_2 \in C$ となるとき，C は凸集合であるという．凸集合は凹みのない集合と考えてよい．凸集合 C が閉集合 (境界も自分自身に含まれる集合) であるとき，閉凸集合であるという．

(e) (凸最適化問題 (1.1) の解の存在性) 一般に，関数 $f \in \Gamma_0(\mathcal{H})$ は閉凸集合 $C(\subset \mathcal{H})$ 上に最小値を持つとは限らない．最小値の存在を保証するいくつかの十分条件が知られている．例えば，$\mathrm{dom}(f) \cap C \neq \emptyset$ であるとき，

$$\|\boldsymbol{x}\| \to \infty \Rightarrow f(\boldsymbol{x}) + i_C(\boldsymbol{x}) \to \infty$$

であれば，f は C の上で最小値を持つことが保証される (例えば [49] 参照)．

(f) (近接写像) 関数 $f \in \Gamma_0(\mathcal{H})$ と $\gamma \in (0,\infty)$ に対して，

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

を満足する $\mathrm{prox}_{\gamma f}(\boldsymbol{x}) \in \mathcal{H}$ が唯一存在する．写像 $\mathrm{prox}_{\gamma f} : \mathcal{H} \ni \boldsymbol{x} \mapsto \mathrm{prox}_{\gamma f}(\boldsymbol{x}) \in \mathcal{H}$ を f の (インデックス γ の) 近接写像 (proximity operator) という[13]．任意の $\boldsymbol{x} \in \mathcal{H}$ と $\gamma \in (0,\infty)$ に対して，$\mathrm{prox}_{\gamma f}(\boldsymbol{x})$ が容易に計算できる $f \in \Gamma_0(\mathcal{H})$ は proximable であるという．$f \in \Gamma_0(\mathcal{H})$ と $f^* \in \Gamma_0(\mathcal{H})$ の間には，

$$\mathrm{prox}_{\gamma f} + \gamma \mathrm{prox}_{\gamma^{-1} f^*} \circ \gamma^{-1}\mathrm{I}_{\mathcal{H}} = \mathrm{I}_{\mathcal{H}} \tag{2.1}$$

が成立する[14]．また，関数

$${}^{\gamma}f : \boldsymbol{x} \mapsto \min_{\boldsymbol{y}\in\mathcal{H}}\left(f(\boldsymbol{y}) + \frac{1}{2\gamma}\|\boldsymbol{x}-\boldsymbol{y}\|^2\right)$$

を f の「(インデックス γ の)Moreau エンベロープ」という．${}^{\gamma}f$ は ${}^{\gamma}f \leq f$ と $\lim_{\gamma\downarrow 0} {}^{\gamma}f(\boldsymbol{x}) = f(\boldsymbol{x})$ $(\forall \boldsymbol{x} \in \mathrm{dom}(f))$ を満たすばかりでなく，ガトー微分

$$\nabla\left({}^{\gamma}f\right)(\boldsymbol{x}) = \frac{1}{\gamma}(\boldsymbol{x} - \mathrm{prox}_{\gamma f}(\boldsymbol{x})) \quad (\forall \boldsymbol{x} \in \mathcal{H})$$

を持つ微分可能な凸関数となるため，f の理想的な平滑化を実現している．

例 3. (変数分離型凸関数の共役関数と近接写像) 例 2 のように，実ヒルベルト空間 $\mathcal{H}$ が実ヒルベルト空間 $\mathcal{X}_j$ $(j=1,\ldots,p)$ の直積空間となっており，$\varphi_j \in \Gamma_0(\mathcal{X}_j)$ を用いて $f := \sum_{j=1}^p \varphi_j \in \Gamma_0(\mathcal{H})$ を定義するとき，以下が成立する．

---

[13] 関数 $\gamma f \in \Gamma_0(\mathcal{H})$ の近接写像に一致することに注意されたい (1.2 節参照)．

[14] 極大単調作用素 ([3] 参照) とその逆写像の間に成立する関係 (inverse resolvent identity) の一例になっており，特に Moreau's decomposition と呼ばれることもある．式 (2.1) より，f が proximable であれば，f^* も proximable になることがわかる．

(a) 任意の $\boldsymbol{y}_j \in \mathcal{X}_j$ $(j = 1, \ldots, p)$ に対して

$$f^*(\boldsymbol{y}_1, \ldots, \boldsymbol{y}_p) = \sum_{j=1}^{p} \varphi_j^*(\boldsymbol{y}_j)$$

となる．

(b) 任意の $\boldsymbol{x}_j \in \mathcal{X}_j$ $(j = 1, \ldots, p)$ に対して

$$\mathrm{prox}_{\gamma f}(\boldsymbol{x}_1, \ldots, \boldsymbol{x}_p) = \left(\mathrm{prox}_{\gamma\varphi_1}(\boldsymbol{x}_1), \ldots, \mathrm{prox}_{\gamma\varphi_p}(\boldsymbol{x}_p)\right)$$

となる．

これより，$\varphi_j \in \Gamma_0(\mathcal{X}_j)$ $(j = 1, \ldots, p)$ が proximable であれば，$f \in \Gamma_0(\mathcal{H})$ も $f^* \in \Gamma_0(\mathcal{H})$ も proximable となることがわかる．

3 凸最適化問題と非拡大写像の不動点理論

3.1 非拡大写像と不動点

写像 $T : \mathcal{H} \to \mathcal{H}$ に対して [15]，$T(\boldsymbol{z}) = \boldsymbol{z}$ を満足する $\boldsymbol{z} \in \mathcal{H}$ が存在するとき，$\boldsymbol{z}$ を「T の不動点 (fixed point)」といい，$\mathrm{Fix}(T) := \{\boldsymbol{z} \in \mathcal{H} \mid T(\boldsymbol{z}) = \boldsymbol{z}\}$ を T の不動点集合 (fixed point set) という．写像 $T : \mathcal{H} \to \mathcal{H}$ に対して，ある定数 $\kappa > 0$ が存在し

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

が成立するとき，T はリプシッツ連続であるといい，κ を T のリプシッツ定数という．特に $\kappa < 1$ となるリプシッツ定数がとれるとき，T は縮小写像であるといい，T には唯一の不動点 $\boldsymbol{z} \in \mathcal{H}$ が存在する．T が縮小写像であるとき，任意の点 $\boldsymbol{x}_0 \in \mathcal{H}$ を初期値とする点列生成アルゴリズム「$\boldsymbol{x}_{n+1} := T(\boldsymbol{x}_n)$ $(n = 0, 1, 2, \ldots)$」によって $\lim_{n\to\infty} \|\boldsymbol{x}_n - \boldsymbol{z}\| = 0$ が保証されるため，縮小写像 T の不動点をいくらでも精度よく近似できること (バナッハ・ピカールの不動点定理) がよく知られている．一方，写像 T に $\kappa = 1$ となるリプシッツ定数がとれるとき，T は非拡大写像 (nonexpansive mapping) であるという．非拡大写像 $T : \mathcal{H} \to \mathcal{H}$ が不動点を持つとき (注：一般に非拡大写像は不動点の存在性が保証されない)，不動点集合 $\mathrm{Fix}(T) := \{\boldsymbol{z} \in \mathcal{H} \mid T(\boldsymbol{z}) = \boldsymbol{z}\}$ は閉凸集合になる [49]．特に非拡大写像 $T : \mathcal{H} \to \mathcal{H}$ に対してある $\alpha \in (0,1)$ と非拡大写像 $R : \mathcal{H} \to \mathcal{H}$ が存在し，$T = (1-\alpha)\mathrm{I}_{\mathcal{H}} + \alpha R$ とあらわせるとき，小文では T は α-平均非拡大であるといい，α を T の平均度 (averagedness constant) とよぶ，$\mathrm{Fix}(T) = \mathrm{Fix}(R)$ が成立する．$f \in \Gamma_0(\mathcal{H})$ と任意の $\gamma \in (0, \infty)$ に対して，近接写像 $\mathrm{prox}_{\gamma f}$: は $\frac{1}{2}$-平均非拡大写像となるため，$\nabla({}^{\gamma}f) : \mathcal{H} \to \mathcal{H}$ はリプシッツ連続 (リプシッツ定数 $\frac{1}{\gamma}$) になることが保証される．

[15] T は線形写像でも非線形写像でもよい．

3.2 非拡大写像の不動点近似アルゴリズム

非拡大写像の不動点を逐次近似するために開発された簡潔で広く応用できるアルゴリズムを2つ紹介しておく．

Fact 2. 非拡大写像 $T:\mathcal{H}\to\mathcal{H}$ が空でない不動点集合 $\mathrm{Fix}(T)$ を持つとき，以下が成立する．

(a) (Krasnosel'skii-Mann のアルゴリズム (例えば [15, 22])) 任意の初期値 $\boldsymbol{x}_0\in\mathcal{H}$ と $\sum_{n\geq 0}\alpha_n(1-\alpha_n)=\infty$ を満たす実数列 $(\alpha_n)_{n\geq 0}\subset[0,1]$ を用いて，点列 $(\boldsymbol{x}_n)_{n\geq 0}\subset\mathcal{H}$ を

$$\boldsymbol{x}_{n+1}:=(1-\alpha_n)\boldsymbol{x}_n+\alpha_n T(\boldsymbol{x}_n) \tag{3.1}$$

によって生成すると，$(\boldsymbol{x}_n)_{n\geq 0}$ は $\mathrm{Fix}(T)$ 中の 1 点 (仮に $\boldsymbol{z}\in\mathrm{Fix}(T)$ とよぶ) に弱収束する [16](弱極限値 $\boldsymbol{z}$ は初期値に依存する)．

(b) (ハイブリッド最急降下法 [42]) 凸関数 $J:\mathcal{H}\to\mathbb{R}$ は $\arg\min_{\boldsymbol{x}\in\mathrm{Fix}(T)}J(\boldsymbol{x})\neq\emptyset$ を満たし，J のガトー微分は $T(\mathcal{H}):=\{T(\boldsymbol{x})\mid x\in\mathcal{H}\}$ 上で強単調 [17]：

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

かつリプシッツ連続であるとする．また，ステップサイズ $(\lambda_n)_{n\geq 1}\subset[0,\infty)$ は，(i) $\lim\limits_{n\to\infty}\lambda_n=0$, (ii) $\sum\limits_{n\geq 1}\lambda_n=\infty$, (iii) $\sum\limits_{n\geq 1}|\lambda_n-\lambda_{n+1}|<\infty$ を同時に満足するものと仮定する [例えば，$\lambda_n:=\frac{1}{n}$ $(n=1,2,3,\ldots)$]．このとき，任意の初期点 $\boldsymbol{u}_0\in\mathcal{H}$ を用いてハイブリッド最急降下法

$$\boldsymbol{u}_{n+1}:=T(\boldsymbol{u}_n)-\lambda_{n+1}\nabla J(T(\boldsymbol{u}_n)) \tag{3.2}$$

によって生成される点列 $(\boldsymbol{u}_n)_{n\geq 0}$ および、任意の初期点 $v_0\in\mathcal{H}$ を用いて

$$v_{n+1}:=T\left(v_n-\lambda_{n+1}\Theta'(v_n)\right),\quad n=0,1,2,\ldots \tag{3.3}$$

によって生成される点列 $(v_n)_{n\geq 0}$ は、共に $J(\boldsymbol{u}^*)=\min\limits_{\boldsymbol{u}\in\mathrm{Fix}(T)}J(\boldsymbol{u})$ を満足する唯一の点 $\boldsymbol{u}^*\in\mathrm{Fix}(T)$ に強収束する．

注意 1. (Krasnosel'skii-Mann のアルゴリズムについて)

---

[16]弱収束「$\lim_{n\to\infty}\langle\boldsymbol{x}_n-\boldsymbol{z},\boldsymbol{y}\rangle=0$ $(\forall\boldsymbol{y}\in\mathcal{H})$」であれば点列 $(\boldsymbol{x}_n)_{n\geq 0}$ はどの座標成分で見ても $\boldsymbol{z}$ に収束しているように見えるが，$\mathcal{H}$ が無限次元ヒルベルト空間の場合，強収束 $\lim\limits_{n\to\infty}\|\boldsymbol{x}_n-\boldsymbol{z}\|=0$ は保証されない．なお，有限次元ヒルベルト空間の場合には弱収束点列と強収束点列の定義は等価になる [49]．

[17]$J:\mathcal{H}\to\mathbb{R}$ がガトー微分可能なら，「J は凸関数」$\Leftrightarrow$「∇J は単調写像：$\langle\nabla J(\boldsymbol{x})-\nabla J(\boldsymbol{y}),\boldsymbol{x}-\boldsymbol{y}\rangle\geq 0$ $(\boldsymbol{x},\boldsymbol{y}\in\mathcal{H})$」が成立する [49]．

(a) 信号画像復元問題をはじめとする多様な逆問題の解法として広く応用されてきた数多くの凸最適化アルゴリズム，Krasnosel'skii-Mann のアルゴリズムの応用例として解釈可能である (例えば，[51, 5]).

(b) ある $\alpha \in (0,1)$ と非拡大写像 R を用いて $T = (1-\alpha)\mathrm{I}_{\mathcal{H}} + \alpha R$ (T が α-平均非拡大) と表せるとき，$\mathrm{Fix}(T) = \mathrm{Fix}(R)$ であることに注意すれば，任意の初期値 $\boldsymbol{y}_0 \in \mathcal{H}$ と $\sum_{n\geq 0} \beta_n \left(\frac{1}{\alpha} - \beta_n\right) = \infty$ を満たす実数列 $(\beta_n)_{n\geq 0} \subset \left[0, \frac{1}{\alpha}\right]$ を用いて，点列 $(\boldsymbol{y}_n)_{n\geq 0} \subset \mathcal{H}$ を

$$\boldsymbol{y}_{n+1} := (1-\beta_n)\boldsymbol{y}_n + \beta_n T(\boldsymbol{y}_n) \tag{3.4}$$

によって生成すると，$(\boldsymbol{y}_n)_{n\geq 0}$ は $\mathrm{Fix}(T)$ 中の 1 点に弱収束することが確かめられる．アルゴリズム (3.4) のステップサイズ β_n の選択域はアルゴリズム (3.1) のステップサイズ α_n の選択域に比べて，$\frac{1}{\alpha}$ 倍に拡がっていることに注意されたい．ステップサイズの選択は収束速度に影響を与えるので，平均非拡大性を特徴付ける平均度 α を可能な限り小さく評価することはアルゴリズムの収束性能を直接左右する．実は，多くの凸最適化問題は「2 つの平均非拡大写像の合成写像」の不動点を算出する問題に帰着できることが知られている (例えば，注意 3 参照).

(c) 筆者らは α_i-平均非拡大写像 $T_i : \mathcal{H} \to \mathcal{H}$ $(i = 1, 2)$ が与えられたとき，これらの合成写像 T_1T_2 の平均度が，$\frac{\alpha_1+\alpha_2-2\alpha_1\alpha_2}{1-\alpha_1\alpha_2}$ になることを明らかにしている [27]．この結果は，これまで広く応用されてきた「合成写像の平均度に関する従来の評価 (例えば [3, Prop.4.32])」に比べて小さくなるため，従来の評価に拠って開発されてきた多くの凸最適化アルゴリズムのステップサイズ選択域を拡大し，収束性能向上に活用できる [12].

注意 2. (ハイブリッド最急降下法について)

(a) (式 (3.3) について) 式 (3.2) から生成される点列 $(\boldsymbol{u}_n)_{n\geq 0}$ は、初期値 $\boldsymbol{u}_0$ の選び方によらず、$\lim_{n\to\infty} \boldsymbol{u}_n = \boldsymbol{u}^* \in \mathrm{Fix}(T)$ を満たすので、特に、$\boldsymbol{v}_0 := T(\boldsymbol{u}_0)$ から式 (3.3) によって生成される点列は、$\boldsymbol{v}_n = T(\boldsymbol{u}_n)$ $(n = 0, 1, 2, \ldots)$ となり、$\|\boldsymbol{v}_n - \boldsymbol{u}^*\| \leq \|\boldsymbol{u}_n - \boldsymbol{u}^*\| \to 0$ $(n \to \infty)$ が保証される。同様な議論から、$\boldsymbol{v}_0$ の選び方によらず、式 (3.3) によって生成される点列が $\lim_{n\to\infty} \boldsymbol{v}_n = \boldsymbol{u}^*$ を満たすことがわかる。

(b) (アンカー法: Anchor method[23, 25, 40, 2]) 任意に固定された $\boldsymbol{a} \in \mathcal{H}$ に対して、定義される凸関数 $J(\boldsymbol{x}) := \frac{1}{2}\|\boldsymbol{x} - \boldsymbol{a}\|^2$ にハイブリッド最急降下法 (3.2) を適用するとアンカー法：

$$\boldsymbol{u}_{n+1} := \lambda_{n+1}\boldsymbol{a} + (1-\lambda_{n+1})T(\boldsymbol{u}_n) \qquad (n = 0, 1, 2, \ldots) \tag{3.5}$$

を得る。このアルゴリズムの基本性質は、B. Halpern の報告以来、P. L. Lions (1994年フィールズ賞受賞者)、R. Wittmann、H. H. Bauschke らによって解明されてきた。

(c) 非拡大写像 T の計算に数値誤差が避けられない場合にも Fix(T) 上で J の最小化を保証するロバスト・ハイブリッド最急降下法も実現されている [43]．

(d) ハイブリッド最急降下法で用いる写像 T は準非拡大写像 (劣勾配射影など) の場合 [47] に，また J はガトー微分の強単調性を必要としない場合 [27] にも拡張されており，アルゴリズム (3.2) の応用対象は大きく広がっている．特に微分可能でない凸関数の「理想的な平滑化 (Moreau エンベロープ)」はリプシッツ連続になるので，このような凸関数の最小化問題に広く応用できる [51, 29, 30]．ハイブリッド最急降下法はバナッハ空間にも拡張されている (例えば [10] を参照されたい).

(e) ネットワーク上で定義された凸最適化問題を解決するために「並列計算可能なハイブリッド最急降下法」も開発されている [37]

3.3 凸最適化問題と不動点問題

実ヒルベルト空間 $(\mathcal{X}_1, \langle\cdot,\cdot\rangle_{\mathcal{X}_1}, \|\cdot\|_{\mathcal{X}_1})$ と $\left(\mathcal{X}_2^{(k)}, \langle\cdot,\cdot\rangle_{\mathcal{X}_2^{(k)}}, \|\cdot\|_{\mathcal{X}_2^{(k)}}\right)$ $(k = 1, \ldots, K)$ 上に定義された凸関数 $f, g \in \Gamma_0(\mathcal{X}_1)$, $h_k \in \Gamma_0(\mathcal{X}_2^{(k)})$ と有界線形作用素 $L_k : \mathcal{X}_1 \to \mathcal{X}_2^{(k)}$ $(k = 1, \ldots, K)$ が与えられるとき，凸最適化問題

$$\text{Find a point in } \mathcal{S}_{\mathrm{Pr}} := \arg\min_{\boldsymbol{x}\in\mathcal{X}_1} \left[f(\boldsymbol{x}) + g(\boldsymbol{x}) + \sum_{k=1}^{K} h_k(L_k\boldsymbol{x}) \right] \neq \emptyset \tag{3.6}$$

を考える．ただし，f は $\mathcal{X}_1$ 上で計算可能なガトー微分 $\nabla f : \mathcal{X}_1 \to \mathcal{X}_1$ を持ち，∇f はリプシッツ連続 (リプシッツ定数 $\beta > 0$)，すなわち

$$(\exists \beta > 0, \forall \boldsymbol{x}, \boldsymbol{y} \in \mathcal{X}_1) \quad \|\nabla f(\boldsymbol{x}) - \nabla f(\boldsymbol{y})\|_{\mathcal{X}_1} \le \beta \|\boldsymbol{x} - \boldsymbol{y}\|_{\mathcal{X}_1}$$

を満たしていると仮定する [18]．また，$g \in \Gamma_0(\mathcal{X}_1)$ と $h_k \in \Gamma_0(\mathcal{X}_2^{(k)})$ $(k = 1, \ldots, K)$ は各々が定義されている実ヒルベルト空間で proximable であるとする．最適化問題 (3.6) は現時点で信号処理や逆問題で必要とされる多種多様な凸最適化問題を広くカバーしているが，primal-dual splitting と総称される凸解析の革新的なアイディア (例えば [39, 13]) を不動点近似アルゴリズムに応用することによって，統一的に解決できるようになっている．以下では，[13] の議論に沿って，この問題の解集合 $\mathcal{S}_1$ が「計算可能な非拡大写像」の不動点集合を用いて精密に表現できることを紹介する．

[18] この場合，ガトー微分 ∇f はフレッシェ微分に一致する．

まず，直積空間 (例 2 参照) に内積とノルムを

$$\mathcal{X}_2 := \mathcal{X}_2^{(1)} \times \cdots \times \mathcal{X}_2^{(K)} := \left\{ (\boldsymbol{x}_1, \ldots, \boldsymbol{x}_K) \mid \boldsymbol{x}_k \in \mathcal{X}_2^{(k)} \quad (k = 1, \ldots, K) \right\}$$

$$\langle (\boldsymbol{x}_1, \ldots, \boldsymbol{x}_K), (\boldsymbol{y}_1, \ldots, \boldsymbol{y}_K) \rangle_{\mathcal{X}_2} := \sum_{k=1}^{K} \langle \boldsymbol{x}_k, \boldsymbol{y}_k \rangle_{\mathcal{X}_2^{(k)}}$$

$$\|(\boldsymbol{x}_1, \ldots, \boldsymbol{x}_K)\|_{\mathcal{X}_2} := \sqrt{\langle (\boldsymbol{x}_1, \ldots, \boldsymbol{x}_K), (\boldsymbol{x}_1, \ldots, \boldsymbol{x}_K) \rangle_{\mathcal{X}_2}}$$

のように与えることによって，新たな実ヒルベルト空間 $(\mathcal{X}_2, \langle \cdot, \cdot \rangle_{\mathcal{X}_2}, \| \cdot \|_{\mathcal{X}_2})$ を定義し，さらに，有界線形作用素 $L : \mathcal{X}_1 \to \mathcal{X}_2$ と $h \in \Gamma_0(\mathcal{X}_2)$ を

$$L : \boldsymbol{x} \mapsto (L_1 \boldsymbol{x}, \ldots, L_K \boldsymbol{x})$$

$$h : (\boldsymbol{x}_1, \ldots, \boldsymbol{x}_K) \mapsto \sum_{k=1}^{K} h_k(\boldsymbol{x}_k)$$

のように定義すると [19]，問題 (3.6) は

$$\text{Find a point in } \mathcal{S}_{\text{Pr}} := \arg \min_{\boldsymbol{x} \in \mathcal{X}_1} [f(\boldsymbol{x}) + g(\boldsymbol{x}) + h(L\boldsymbol{x})] \neq \emptyset \tag{3.7}$$

のように簡易表現できることがわかる．例 3 で見たように関数 h も $\mathcal{X}_2$ 上で proximable になることに注意されたい．問題 (3.7) を主問題 (primal problem) と呼ぶとき，

$$\text{Find a point in } \mathcal{S}_{\text{Du}} := \arg \min_{\boldsymbol{y} \in \mathcal{X}_2} [(f+g)^*(-L^* \boldsymbol{y}) + h^*(\boldsymbol{y})] \neq \emptyset \tag{3.8}$$

は双対問題 (dual problem) であるという (注：比較的緩い条件下で $\mathcal{S}_{\text{Du}} \neq \emptyset$ が保証される [3])．このとき，以下が成立する．

Fact 3. (凸最適化問題 (3.7) の解集合の不動点表現 [13, 29]) 直積空間 $\mathcal{X} := \mathcal{X}_1 \times \mathcal{X}_2$ 上に

$$\kappa := \frac{1}{\beta} \left(\frac{1}{\gamma_1} - \gamma_2 \|L\|_{\text{op}}^2 \right) > \frac{1}{2}$$

を満たす $\gamma_1, \gamma_2 > 0$ を用いて線形作用素

$$Q : \mathcal{X}_1 \times \mathcal{X}_2 \to \mathcal{X}_1 \times \mathcal{X}_2 : (\boldsymbol{x}, \boldsymbol{\xi}) \mapsto \left(\frac{1}{\gamma_1} \boldsymbol{x} - L^* \boldsymbol{\xi}, \frac{1}{\gamma_2} \boldsymbol{\xi} - L \boldsymbol{x} \right) \tag{3.9}$$

と非線形写像

$$\begin{aligned} & T_{\text{PDS}} : \mathcal{X}_1 \times \mathcal{X}_2 \to \mathcal{X}_1 \times \mathcal{X}_2 : (\boldsymbol{x}, \boldsymbol{\xi}) \mapsto (\hat{\boldsymbol{x}}, \hat{\boldsymbol{\xi}}), \\ & \text{ただし，} \left[\begin{array}{l} \hat{\boldsymbol{x}} := \text{prox}_{\gamma_1 g} (\boldsymbol{x} - \gamma_1 (\nabla f(\boldsymbol{x}) + L^* \boldsymbol{\xi})) \\ \hat{\boldsymbol{\xi}} := \text{prox}_{\gamma_2 h^*} (\boldsymbol{\xi} + \gamma_2 L (2\hat{\boldsymbol{x}} - \boldsymbol{x})) \end{array} \right. \end{aligned} \tag{3.10}$$

を定義する [20]．このとき，以下が成立する．

[19] $\|L\|_{\text{op}} < \infty$ は L の定義から容易に確かめられる．

[20] 例 3(a),(b) に述べた変数分離型凸関数の性質と関係 (2.1) を使うと，$h^* \in \Gamma_0(\mathcal{X}_2)$ が proximable であり，T_{PDS} が計算可能であることが確かめられる．

(a) 任意の $(\boldsymbol{x}^\star, \boldsymbol{\xi}^\star) \in \mathcal{X}_1 \times \mathcal{X}_2$ に対して

$$(\boldsymbol{x}^\star, \boldsymbol{\xi}^\star) \in \mathcal{S}_{\mathrm{Pr}} \times \mathcal{S}_{\mathrm{Du}} \Leftrightarrow (\boldsymbol{x}^\star, \boldsymbol{\xi}^\star) \in \mathrm{Fix}(T_{\mathrm{PDS}}) \tag{3.11}$$

が成立する．関係 (3.11) は，主問題と双対問題の解集合が,「f の勾配と prox-imable な関数 g, h の近接写像の計算だけで構成された T_{PDS} の不動点集合」として精密に表現できることを示しており，主双対分解 (primal-dual splitting) とよばれている．

(b) 例 2(a)[$p=2$] のように内積 $\langle\cdot,\cdot\rangle$ とノルム $\|\cdot\|$ が与えられた実ヒルベルト空間 $(\mathcal{X}, \langle\cdot,\cdot\rangle, \|\cdot\|)$ 上で Q は有界な自己共役作用素となり，例 2(b) の条件を満たす．したがって，Q は新しい実ヒルベルト空間 $(\mathcal{X}_Q := \mathcal{X}_1 \times \mathcal{X}_2, \langle\cdot,\cdot\rangle_Q, \|\cdot\|_Q)$ を定義する．

(c) 写像 T_{PDS} を (c) で定義した $(\mathcal{X}_Q := \mathcal{X}_1 \times \mathcal{X}_2, \langle\cdot,\cdot\rangle_Q, \|\cdot\|_Q)$ 上の写像と見れば，$\frac{2\kappa}{4\kappa-1}$-平均非拡大写像になる．

注意 3. (主双対分解について)

(a) 集合値写像 $A : \mathcal{X} \to 2^{\mathcal{X}} : (\boldsymbol{x}, \boldsymbol{\xi}) \mapsto (\partial g(\boldsymbol{x}) + L^* \boldsymbol{\xi}) \times (-L\boldsymbol{x} + \partial h^*(\boldsymbol{\xi}))$ と単価写像 $B : \mathcal{X} \to \mathcal{X} : (\boldsymbol{x}, \boldsymbol{\xi}) \mapsto (\nabla f(\boldsymbol{x}), \mathbf{0})$ を定義すると，まず

$$\begin{aligned}(\boldsymbol{x}^\star, \boldsymbol{\xi}^\star) \in \mathcal{S}_{\mathrm{Pr}} \times \mathcal{S}_{\mathrm{Du}} \quad &\Leftrightarrow \quad (\mathbf{0}, \mathbf{0}) \in (A + B)(\boldsymbol{x}^*, \boldsymbol{\xi}^*) \\ &\Leftrightarrow \quad (\mathbf{0}, \mathbf{0}) \in \left(Q^{-1} \circ A + Q^{-1} \circ B\right)(\boldsymbol{x}^*, \boldsymbol{\xi}^*)\end{aligned}$$

が確かめられる．次に，$Q^{-1} \circ A : \mathcal{X} \to 2^{\mathcal{X}}$ は Fact2(c) で定義されたヒルベルト空間 $\mathcal{X}_Q$ 上で極大単調作用素 (maximally monotone operator) になり，そのレゾルベント $J_{Q^{-1} \circ A} := (\mathrm{I}_{\mathcal{X}} + Q^{-1} \circ A)^{-1} : \mathcal{X}_Q \to \mathcal{X}_Q$(単価写像になる) が $\frac{1}{2}$-平均非拡大写像となることが確かめられる．また，[1] の結果から $\mathrm{I}_{\mathcal{X}} - Q^{-1} \circ B : \mathcal{X}_Q \to \mathcal{X}_Q$ が $\frac{1}{2\kappa}$-平均非拡大写像になることも確かめられる．さらに,「2 つの極大単調作用素 (1 つは単価写像) の和の零点」を「単価写像の不動点」として表現する代表的なアイディア (Forward-Backward Splitting [3]) を用いれば

$$\begin{aligned}&(\mathbf{0}, \mathbf{0}) \in \left(Q^{-1} \circ A + Q^{-1} \circ B\right)(\boldsymbol{x}^*, \boldsymbol{\xi}^*) \\ \Leftrightarrow \quad &(\boldsymbol{x}^*, \boldsymbol{\xi}^*) \in \mathrm{Fix}\left(J_{Q^{-1} \circ A} \circ \left(\mathrm{I}_{\mathcal{X}} - Q^{-1} \circ B\right)\right)\end{aligned}$$

も直ちに確かめられる．実は Fact3 で定義された T_{PDS} は

$$T_{\mathrm{PDS}} = J_{Q^{-1} \circ A} \circ \left(\mathrm{I}_{\mathcal{X}} - Q^{-1} \circ B\right) \tag{3.12}$$

を満たし，2 つの平均非拡大写像の合成写像になることも確かめられる．T_{PDS} の平均度評価 $\frac{2\kappa}{4\kappa-1}$ には注意 1(c) で紹介した結果が応用されている．

(b) 凸最適化問題の解集合が非拡大写像の不動点集合として表現できる事実は，Fact3 の他にも数多くの例が知られており (例えば [51])，Fact2(a) を適用することによって, 極めて多くの凸最適化アルゴリズムが統一的に導出できる．例えば,「$f \equiv 0$, $L^*L = \rho I_{\mathcal{X}_1}$ $(\exists \rho > 0)$」という特別な条件下で問題 (3.7) に応用可能な古典的アルゴリズムとして *Alternating Directions Methods of Multipliers*(ADMM)[18, 17, 16] が広く知られているが，[13, Algoorithm3.2] は ADMM の一般化となっており，Fact3 と同様な不動点表現に Fact2(a) を適用することによって導かれている．

(c) 3.2 節と 3.3 節の議論から凸最適化問題 3.6 を解決するには Fact3 で得られた非拡大写像 $T_{\rm PDS}$ に Krasnosel'skii-Mann のアルゴリズム (Fact2(a)) を適用すればよいことがわかる．勿論，$T_{\rm PDS}$ の平均度を活かし，ステップサイズの選択域を拡大したアルゴリズム (3.4) が適用できる．

(d) 筆者らは収束性能の更なる改善のために，式 (3.12) の右辺に現れた写像に可変パラメータを導入し,「最初から最後まで固定された写像 $T_{\rm PDS}$ を使う枠組み (Krasnosel'skii-Mann のアルゴリズム)」に留まらない一般化も与えている [12].

(e) 「問題 (3.6) の解集合上の凸最適化問題 (階層構造を持つ問題 (1.3)」を解決するには $T_{\rm PDS}$ にハイブリッド最急降下法 (Fact2(b)) を適用すればよいことがわかる [51, 29].

4 おわりに

小文では信号処理技術の発展に大きな影響を与えている凸最適化のブレークスルーについて，その一端を不動点理論の視点から紹介したが，関連するその他のいくつかの話題についても簡単に触れておく [21].

(a) (適応射影劣勾配法とオンライン学習問題への応用) 筆者らは凸関数列 $(\Theta_n)_{n\in\mathbb{N}}$ を漸近的に最小化するためのアルゴリズム適応射影劣勾配法 (Adaptive projected subgradient method) [45, 46, 35] を考案し，これを音響エコー消去問題 [54], 無線通信システムの干渉抑圧問題 [53, 6], オンラインパターン認識問題 [34], 分散ネットワーク上の適応学習問題 [7, 8] など, 情報通信システムに登場する多くのオンライン学習問題に応用し, その有効性を実証してきた．適応射影劣勾配法の考え方と応用事例を解説論文 [38][22] にやさしく紹介している．適応射影劣勾配法は, 超複素数 (Cayley-Dickson 数) で表現されたオンライン学習問題にも拡張されている [26].

[21] 数理的な研究対象として眺めたとき，信号処理と最適化と逆問題は殆ど一体である．このような視点で筆者らが取り組んできた研究事例を [50] にやさしく紹介している．

[22] 2014 IEEE Signal Processing Magazine Best Paper Award を受賞している．

(b) (低階数最小分散擬似不偏推定法と悪条件逆問題への応用) 悪条件 (Ill-condition) の線形逆問題に「(出力誤差の 2 乗平均値の最小化を指導原理とする) 線形最小 2 乗法」を適用すると「パラメータ推定値と真値との 2 乗平均誤差」は線形モデルの係数行列が持つ"非零特異値"の逆数の 2 乗のオーダとなり, 満足のいく結果を与えないことが知られている. 筆者らは「低階数行列の中で最適な推定行列とは何か」を検討するために,「最小分散不偏推定法の方針」と「低階数推定法の方針」を融合した「最小分散低階数擬似不偏推定法 (MV-PURE)」を提案している [48, 31]. MV-PURE は, バイアス生成行列の全てのユニタリ不変ノルムを同時最小化した上で推定値の分散の最小化を達成する最良低階数推定行列として定義されており, 低階数行列を第 1 層の制約集合に持つ「階層構造を持つ非凸最適化問題」の解となっている. 幸い, 筆者らは, この解を陽に与えることに成功している. MV-PURE は線形制約条件を課した推定法であり、ガウス・マルコフの最小分散不偏推定法や Marquardt の推定法 (D.W.Marquardt, Technometrics, vol. 12, 1970) や Chipman の推定法 (Econometrica, vol. 32, 1964, Linear Algebra Appl., vol. 289, 1999) を特別な場合に含む統一的な推定法になっている.

(c) (代数的位相アンラップと 2 次元スプライン近似) 複素数の偏角 (位相とよばれる) は, 2π の整数倍の任意性を持つことは周知のとおりであるが, $\mathbb{R}^2$ 中の有限の離散標本点上に与えられた「(2π の整数倍の任意性を持つ) 位相」を補間する全ての関数の中からできるだけ振動が小さく抑えられているものを決定し, 各標本点にあった「位相の任意性を解消する問題」は 2 次元位相アンラップ (Tow dimensional phase unwrapping) とよばれており, 合成開口レーダ-による地形情報推定 [33] や MRI 画像処理 [21] の成否を決定する重要な問題である [20]. この問題の意味を理解することは容易であるが, 本来, 連続的な領域で定義されるべき振動量を最小にする離散情報 (2π の整数倍の組み合わせ) を特定することが求められるため, 長年のボトルネックになっている. 実際には, この問題の背景にある関数近似的な性質を無視することにより, 形式的な離散最適化問題 (NP 困難になる [9]) に置き換えて取り組むことが多いのであるが, 結局は最適性の保証のないヒューリスティックな解法しか生まれていない. 筆者は 1990 年代後半に代数的なアプローチ (スツルムの定理の一般化) によって位相アンラップが原理的に解決できるのではないかと考え, このアイディア (代数的位相アンラップ) を結実させるべく, 検討を進めてきた [41, 44, 52]. 最近, 部分終結式の理論 [4] を応用することにより, 多項式の除算演算に伴う代数的位相アンラップの数値的不安定性が解消されることが明らかとなり, また, ポアンカレの補題を応用することによって, 2 次元位相アンラップの問題が「複素数の実部と虚部の同時関数補間問題」(2 次元スプライン関数空間の凸最適化問題に緩和される) に帰着できることも明らかになっている [24].

参考文献

[1] J. -B. Baillon and G. Haddad, *Quelques propriétés des opérateurs angle-bornés et n-cycliquement monotones*, Isr. J. Math., **26**, pp.137-150, 1977.

[2] H. H. Bauschke, *The approximation of fixed points of compositions of nonexpansive mappings in Hilbert space*, J. Math. Anal. Appl., **202**, pp.150-159, 1996.

[3] H. H. Bauschke and P. L. Combettes, *Convex analysis and monotone operator theory in Hilbert spaces*, Springer-Verlag, 2011.

[4] W. S. Brown and J. F. Traub, *On Euclid's algorithm and the theory of subresultants*, J. ACM, **18**(4), pp.505-514, 1971.

[5] C. L. Byrne, *A unified treatment of some iterative algorithms in signal processing and image reconstruction*, Inverse Problems, **20**, pp.103-120, 2004.

[6] R. L. G. Cavalcante and I. Yamada, *Multiaccess interference suppression in orthogonal space-time block coded MIMO systems by the Adaptive Projected Subgradient Method*,IEEE Trans. Signal Process., **56**(3), pp.1028-1042,2008.

[7] R. Cavalcante, I. Yamada, and B. Mulgrew, *An adaptive projected subgradient approach to learning in diffusion networks*, IEEE Trans. Signal Process., **57**(7), pp.2762-2774, 2009.

[8] R. Cavalcante, A. Rogers, N. Jennings and I. Yamada, *Distributed asymptotic minimization of sequences of convex functions by a broadcast adaptive subgradient method*, IEEE J. Sel. Top. Signal Process., **5**, 2011.

[9] C. W. Chen and H. A. Zebker, *Network approaches to two-dimensional phase unwrapping: intractability and two new algorithms*, J. Opt. Soc. Am. A, **17**(3), pp.401-414,2000.

[10] 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.

[11] P. L. Combettes and J.-C. Pesquet, *Proximal splitting methods in signal processing*, pp.185-212, In: Fixed Point Algorithms for Inverse Problems in Science and Engineering (Bauschke, Burachik, Combettes, Elser, Luke, Wolkowicz, eds.) , Springer-Verlag, 2011.

[12] P. L. COMBETTES AND I. YAMADA, *Compositions and convex combinations of averaged nonexpansive operators*, Journal of Mathematical Analysis and Applications, **425**(1), pp.55-70, 2015.

[13] L. CONDAT, *A primal-dual splitting method for convex optimization involving Lipschitzian,proximable and linear composite terms*, J. Optim. Theory Appl., **158**(2), pp.460-479, 2013.

[14] D. L. DONOHO, *De-noising by soft-thresholding*, IEEE Trans. Information Theory, **41**(3), pp.613-627, 1995.

[15] W. G. DOTSON, *On the Mann iterative process*, Trans. Amer. Math. Soc., **149**, pp.65-73, 1970.

[16] J. ECKSTEIN AND D. BERTSEKAS, *On the Douglas-Rachfold splitting method and proximal point algorithm for maximal monotone operators*, Math. Program., **55**, pp.293-318, 1992.

[17] M. FORTIN AND R. GLOWINSKI, *Augmented Lagrangian Methods: applications to the numerical solution of boundary-value problems*, Elsevier, 1983.

[18] D. GABAY AND B. MERCIER, *A dual algorithm for the solution of nonlinear variational problems via finite elements approximations*, Comput. Math. Appl., **2**(1), pp.17-40, 1976.

[19] S. GANDY, B. RECHT AND I. YAMADA, *Tensor completion and low-n-rank tensor recovery via convex optimization*, Inverse Probl., **27**(2), 025010, 2011.

[20] D. C. GHIGLIA AND M. D. PRITT, *Two-Dimensional Phase Unwrapping: Theory, Algorithms, and Software*, New York: Wiley, 1998.

[21] G. H. GLOVER AND E. SCHNEIDER, *Three-point Dixon technique for true water/fat decomposition with B_0 inhomogeneity correction*, Magnetic Resonance in Medicine, **18**(2), pp.371-383, 1991.

[22] C. W. GROETSCH, *A note on segmenting Mann iterates*, J. Math. Anal. Appl. **40**, pp.369-372, 1972.

[23] B. HALPERN, *Fixed points of nonexpanding maps*, Bull. Amer. Math. Soc.,**73**, pp.957-961, 1967.

[24] D. KITAHARA AND I. YAMADA, *Algebraic phase unwrapping along the real axis: extensions and stabilizations*, Multidimens. Syst. Signal Process.,**26**(1), pp.3-45, 2015.

[25] P. L. Lions, *Approximation de points fixes de contractions*, C. R. Acad. Sci. Paris Sèrie A-B, **284**, pp.1357-1359, 1977.

[26] T. Mizoguchi and I. Yamada, *An algebraic translation of Cayley-Dickson linear systems and its applications to online Learning*, IEEE Trans on Signal Process.,**62**(6), pp.1438-1453, 2014.

[27] N. Ogura and I. Yamada, *Non-strictly convex minimization over the bounded fixed point set of nonexpansive mapping*, Numer. Funct. Anal. Optim., **24**, pp.129-135, 2003.

[28] S. Ono, T. Miyata and I. Yamada, *Cartoon-texture image decomposition using blockwise low-rank texture characterization*, IEEE Trans.Image Process., **23**(3), pp.1128-1142, 2014.

[29] S. Ono and I. Yamada, *Hierarchical convex optimization with primal-dual splitting*, IEEE Trans Signal Process., **63**(2), pp.373-388, 2015.

[30] S. Ono and I. Yamada, *Signal recovery with certain involved convex data-fidelity constraints*, IEEE Trans Signal Process., **63**(22), pp.6149-6163, 2015.

[31] T. Piotrowski and I. Yamada, *MV-PURE estimator : Minimum-variance pseudo-unbiased reduced-rank estimator for linearly constrained ill-conditioned inverse problems*, IEEE Trans. Signal Process., **56**(8), pp.3408–3423, 2008.

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

[33] P. A. Rosen, S. Hensley, I. R. Joughin, F. K. Li, S. N. Madsen, E. Rodriguez, and R. M. Goldstein, “Synthetic Aperture Radar Interferometry,” Proceedings of the IEEE, vol.88, no.3, pp. 333-382, 2000.

[34] K. Slavakis, S. Theodoridis and I. Yamada, *Online kernel-based classification using adaptive projection algorithms*, IEEE Trans. Signal Process., **56**(7), pp.2781-2796, 2008.

[35] K. Slavakis and I. Yamada, *The adaptive projected subgradient method constrained by families of quasi-nonexpansive mappings and its application to online learning*, SIAM Journal on Optimization, **23**(1), pp.126-152, 2013.

[36] J.-L. Starck, F. Murtagh and J. M. Fadili, *Sparse image and signal processing: wavelets, curvelets, morphological diversity*, Cambridge University Press, 2010.

[37] N. Takahashi and I. Yamada, *Parallel algorithms for variational inequalities over the Cartesian product of the intersections of the fixed point sets of nonexpansive mappings*, J. Approx. Theory, **153**, pp.139-160, 2008.

[38] S. Theodoridis, K. Slavakis and, I. Yamada, *Adaptive learning in a world of projections: a unifying framework for linear and nonlinear classification and regression tasks*," IEEE Signal Process. Mag., **21**(1), pp.97-123, 2011.

[39] B. C. Vu, *A splitting algorithm for dual monotone inclusions involving cocoercive operators*, Adv.Comput.Math., **38**(3), pp.667-681, 2013.

[40] R. Wittmann, *Approximation of fixed points of nonexpansive mappings*, Arch. Math.,**58**, pp.486-491, 1992.

[41] I. Yamada, K. Kurosawa, H. Hasegawa and K. Sakaniwa, *Algebraic multidimensional phase unwrapping and zero distribution of complex polynomials - characterization of multivariate stable polynomials*, IEEE Trans. Signal Proces., vol.46, no.6 pp.1639–1664, 1998.

[42] I. Yamada, *The hybrid steepest descent method for the variational inequality problem over the intersection of fixed point sets of nonexpansive mappings*, In: Inherently parallel algorithms in feasibility and optimization and their applications (D. Butnariu, Y. Censor and S. Reich, eds.), Studies in Computational Mathematics, **8**, pp.473-504, Elsevier, 2001.

[43] I. Yamada, N. Ogura and N. Shirakawa, *A numerically robust hybrid steepest descent method for the convexly constrained generalized inverse problems*,pp.269-305, In: Inverse Problems, Image Analysis, and Medical Imaging (Z. Nashed and O. Scherzer, eds.), Contemporary Mathematics, **313**, Amer. Math. Soc.,2002.

[44] I. Yamada and N. K. Bose, *Algebraic phase unwrapping and zero distribution of polynomial for continuous-time systems*, IEEE Trans. Circuits SystI, Fundam. Theory Appl. vol.49, no.3, pp.298-304, 2002.

[45] 山田功, 射影型適応アルゴリズムの新展開—射影劣こう配法による統一的視点とその応用, 電子情報通信学会学会誌, **86**(8), pp.654-658,2003.

[46] I. Yamada and N. Ogura, *Adaptive projected subgradient method for asymptotic minimization of sequence of nonnegative convex functions*, Numer. Funct. Anal. Optim., **25**, pp.593-617, 2004.

[47] I. Yamada and N. Ogura, *Hybrid steepest descent method for variational inequality problem over the fixed point set of certain quasi-nonexpansive mappings*, Numer. Funct. Anal. Optim., **25**(7&8), pp.619-655, 2004.

[48] I. Yamada and J. Elbadraoui,*Minimum-variance pseudo-unbiased low-rank estimator for ill-conditioned inverse problems*, Proc. 2006 IEEE International Conference on Acoustics, Speech and Signal Processing, .III, pp.325-328, Toulouse, May 2006.

[49] 山田功, 工学のための関数解析, 数理工学社 (サイエンス社), 2009.

[50] 山田功, 信号処理・最適化・逆問題 — 学際的自由研究のたのしみ, Fundamentals Reviews, **5**(1), pp.68-79, 2011.

[51] I. Yamada, M. Yukawa,and M. Yamagishi, *Minimizing the Moreau envelope of nonsmooth convex functions over the fixed point set of certain quasi-nonexpansive mappings*,pp.345-390, In: Fixed Point Algorithms for Inverse Problems in Science and Engineering (Bauschke, Burachik, Combettes, Elser, Luke and Wolkowicz, eds.) , Springer-Verlag, 2011.

[52] I. Yamada and K. Oguchi, *High-resolution estimation of the directions-of-arrival distribution by algebraic phase unwrapping*, Multidimens. Syst. Signal Process., **22**(1-3), pp.191-211, 2011.

[53] M. Yukawa, R. L. G. Cavalcante and I. Yamada, *Efficient blind MAI suppression in DS/CDMA systems by embedded constraint parallel projection techniques*,IEICE Trans. Fundamentals, **E88-A**(8), pp.2062-2071, 2005.

[54] M. Yukawa, K. Slavakis, and I. Yamada, *Adaptive parallel quadratic-metric projection algorithms*, IEEE Trans. Audio, Speech and Language Processing, **15**(5), pp.1665-1680, 2007.