多目的最適化問題に対する加速付き近接勾配法

05001149　京都大学　*田辺広樹　TANABE Hiroki
05000235　京都大学　福田エレン秀美　FUKUDA Ellen Hidemi
01704632　京都大学　山下信雄　YAMASHITA Nobuo

1. はじめに

本発表では，次の多目的最適化問題について考える．

$$\min_{x \in \mathbf{R}^n} \quad F(x) \tag{1}$$

ここで，F は $\mathbf{R}^n$ から $(\mathbf{R} \cup \{\infty\})^m$ へのベクトル値関数であり，F の各成分 F_i は，連続的微分可能な凸関数 $f_i\colon \mathbf{R}^n \to \mathbf{R}$ と閉真凸関数 $g_i\colon \mathbf{R}^n \to \mathbf{R} \cup \{\infty\}$ によって，

$$F_i(x) = f_i(x) + g_i(x) \quad (i \in \{1, \ldots, m\})$$

と表されているとする．ここで，∇f_i は定数 $L_i > 0$ のリプシッツ連続であるとし，$L \coloneqq \max_{i \in \{1,\ldots,m\}} L_i$ とする．多目的最適化問題においては，すべての目的関数 F_i が同時に最小になるような解が存在しないことが多いため，**パレート最適性**という概念がよく用いられる．パレート解とは，ある目的関数値を改善しようとすると，少なくとも一つの他の目的関数値が改悪されてしまうような解のことをいう．

多目的最適化問題に対する主要な解法として**スカラー化法** [4] がある．これは，適当なパラメータを用いて記述された幾つかの単一目的最適化問題を解くことで元の問題の解を得る手法であるが，パラメータの値を適切に定めるのが難しいという問題点がある．

近年では，単一目的最適化に対するアルゴリズムの拡張手法である**降下法**が注目を集めている．例えば，最急降下法 [3] は，微分可能な多目的最適化問題に対してパレート解を生成する手法である．また，問題 (1) のような構造を持つ問題に対しては，近接勾配法 [6] を適用することができる．最急降下法や近接勾配法は，目的関数の勾配という一次の情報を用いることから**一次法**と呼ばれ，暫定解とパレート解の距離を見積もる**メリット関数** [8] と呼ばれる関数の値が $O(1/k)$ の速さでゼロに収束することが示されている [7]．

一方，単一目的最適化問題に対しては，通常の一次法を上回る収束率を保証する一次法である**加速勾配法**に対する研究も盛んである．加速勾配法は，Nesterov [5] によってはじめに確立されたのち，様々なバリエーションが提案されてきた．その中でも，近接勾配法の加速化手法である FISTA (Fast Iterative Shrinkage-Thresholding Algorithm) [1] は，画像・信号処理などの分野における加速勾配法の普及に大きく貢献した．多目的最適化問題に対しては，文献 [2] が，加速勾配法の拡張を試みているが，各反復における部分問題のラグランジュ乗数の点列が一点に収束するという強い仮定のもとでしか収束率の証明ができておらず，十分とは言えない．本発表では，問題 (1) に対し，近接勾配法 [6] の加速化手法である**加速付き近接勾配法**を提案し，合理的な仮定のもとで，その収束率を証明する．

2. 準備

本節では，本発表で用いる記号や用語を定義する．本発表では，$u, v \in \mathbf{R}^m$ に対して，$u_i \le (<) v_i \, (i = 1, \ldots, m)$ となるとき，$u \le (<) v$ と表す．また，問題 (1) に対して，集合 $\{x \in \mathbf{R}^n \mid F(x) < \infty\}$ を F の**実効定義域**といい，$\operatorname{dom} F$ で表す．さらに，定数 $\alpha \in \mathbf{R}^m$ に対して，集合 $\{x \in \mathbf{R}^n \mid F(x) \le \alpha\}$ を F の α に対する**レベル集合**といい，$\Omega_F(\alpha)$ と表す．加えて，集合 $A \subseteq \mathbf{R}^n$ に対する関数 F の像を $F(A) \coloneqq \{F(x) \in \mathbf{R}^m \mid x \in A\}$ とし，集合 $B \subseteq \mathbf{R}^m$ に対する関数 F の逆像を $F^{-1}(B) \coloneqq \{x \in \mathbf{R}^n \mid F(x) \in B\}$ とする．

次に，多目的最適化問題に対する最適性の概念を導入する．$F(x) < F(x^*)$ となる $x \in \mathbf{R}^n$ が存在しないような $x^* \in \mathbf{R}^n$ を，問題 (1) の**弱パレート最適解**という．さらに，本発表では，弱パレート最適解全体の集合を X^* と表す．

最後に，問題 (1) に対して，拡張実数値関数 $u_0\colon \mathbf{R}^n \to \mathbf{R} \cup \{\infty\}$ を，以下のように定義する．

$$u_0(x) = \sup_{y \in \mathbf{R}^n} \min_{i \in \{1,\ldots,m\}} \{F_i(x) - F_i(y)\}$$

u_0 は問題 (1) に対する**メリット関数**と呼ばれ，x が弱パレート最適解のときはゼロを，そうでないときにはゼロより大きい値を取ることが知られている [8]．ここで，$m = 1$ かつ問題 (1) に最適値 F^* が存在するとき，$u_0(x)$ は $F(x) - F^*$ と一致する．

3. 多目的最適化問題に対する近接勾配法

本節では，問題 (1) に対する既存手法である**近接勾配法** [6] を紹介する．近接勾配法のアルゴリズムは以下のように表される．

アルゴリズム 1. *1.* $x^0 \in \operatorname{dom} F$ とする．

2. $k \geq 1$ について，以下の更新を繰り返す．

$$x^k = \operatorname*{argmin}_{x \in \mathbf{R}^n} \max_{i \in \{1,\ldots,m\}} \left\{ \nabla f_i(x^{k-1})^\top (x - x^{k-1}) + g_i(x) - g_i(x^{k-1}) + \frac{L}{2}\|x - x^{k-1}\|^2 \right\}$$

ここで，$\{x^k\}$ は常に $\mathrm{dom}\, F$ の中で生成されるため，アルゴリズム 1 は well-defined である．アルゴリズム 1 に対しては，次のような収束率の結果が知られている [7].

定理 1. *アルゴリズム 1 によって生成される点列を* $\{x^k\}$ *とする．このとき，*

$$R := \sup_{F^* \in F(X^* \cap \Omega_F(F(x^0)))} \min_{x \in F^{-1}(\{F^*\})} \|x - x^0\|^2 < \infty$$

ならば，任意の $k \geq 1$ *について次式が成り立つ．*

$$u_0(x^k) \leq \frac{LR}{2k}$$

$m = 1$ のとき，問題 (1) に最適値 F^* が存在するならば，$R < \infty$ であるため，定理 1 より，$F(x) - F^*$ が $O(1/k)$ でゼロに収束することが言える．よって，定理 1 は，単一目的最適化問題に対する近接勾配法の収束率の定理 [1] の拡張であることがわかる．

4. 多目的最適化問題に対する加速付き近接勾配法

本節では，近接勾配法の加速化手法である**加速付き近接勾配法**を提案する．加速付き近接勾配法のアルゴリズムは以下のように表される．

アルゴリズム 2. *1.* $y^1 = x^0 \in \mathrm{dom}\, F, t_1 = 1$ とする．

2. $k \geq 1$ について，以下の更新を繰り返す．

$$\begin{cases} x^k = \operatorname*{argmin}_{x \in \mathbf{R}^n} \max_{i \in \{1,\ldots,m\}} \left\{ \nabla f_i(y^k)^\top (x - y^k) + g_i(x) + f_i(y^k) - F_i(x^{k-1}) + \frac{L}{2}\|x - y^k\|^2 \right\} \\ t_{k+1} = \dfrac{1 + \sqrt{1 + 4t_k^2}}{2} \\ y^{k+1} = x^k + \dfrac{t_k - 1}{t_{k+1}}(x^k - x^{k-1}) \end{cases}$$

$m = 1$ のとき，このアルゴリズムは単一目的最適化問題に対する FISTA [1] と一致する．アルゴリズム 1 と同様に，$\{x^k\}$ は常に $\mathrm{dom}\, F$ の中で生成されるため，アルゴリズム 2 は well-defined である．ここで，必ずしも $y^k \in \mathrm{dom}\, F$ ではないことに注意する．また，多目的最適化問題に対する既存の加速化手法である [2] では，部分問題に含まれる項 $F_i(x^{k-1})$ を，$F_i(y^k)$ としている．この違いとメリット関数 u_0 の導入により，アルゴリズム 2 に対しては，次のように，文献 [2] では厳密になされていなかった計算量の解析が可能となる．

定理 2. *アルゴリズム 2 によって生成される点列を* $\{x^k\}, \{y^k\}$ *とする．このとき，*

$$R := \sup_{F^* \in F(X^* \cap \Omega_F(F(x^0)))} \min_{x \in F^{-1}(\{F^*\})} \|x - x^0\|^2 < \infty$$

ならば，任意の $k \geq 1$ *について次式が成り立つ．*

$$u_0(x^k) \leq \frac{2LR}{(k+1)^2}$$

定理 1 と同様に，$m = 1$ のとき，問題 (1) に最適値 F^* が存在するならば，$R < \infty$ であるため，定理 2 より，$F(x) - F^*$ が $O(1/k^2)$ でゼロに収束することが言える．よって，定理 2 は，単一目的最適化問題に対する FISTA の収束率の定理 [1] の拡張であることがわかる．

参考文献

[1] Beck, A. and Teboulle, M.: A fast iterative shrinkage-thresholding algorithm for linear inverse problems, *SIAM Journal on Imaging Sciences*, Vol. 2 (2009), 183–202.

[2] El Moudden, M. and El Mouatasim, A.: Accelerated diagonal steepest descent method for unconstrained multiobjective optimization, *Journal of Optimization Theory and Applications*, (2020), doi: 10.1007/s10957–020–01785–9.

[3] Fliege, J. and Svaiter, B. F.: Steepest descent methods for multicriteria optimization, *Mathematical Methods of Operations Research*, Vol. 51 (2000), 479–494.

[4] Johannes, J.: Scalarization in vector optimization, *Mathematical Programming*, Vol. 29 (1984), 203–218.

[5] Nesterov, Y.: A method for solving the convex programming problem with convergence rate $O(1/k^2)$, *Dokl. Akad. Nauk SSSR*, Vol. 269 (1983), 543–547, (In Russian).

[6] Tanabe, H., Fukuda, E. H. and Yamashita, N.: Proximal gradient methods for multiobjective optimization and their applications, *Computational Optimization and Applications*, Vol. 72 (2019), 339–361.

[7] Tanabe, H., Fukuda, E. H. and Yamashita, N.: Convergence rates analysis of multiobjective proximal gradient methods, arXiv: 2010.08217, 2020.

[8] Tanabe, H., Fukuda, E. H. and Yamashita, N.: New merit functions and error bounds for non-convex multiobjective optimization, arXiv: 2010.09333, 2020.