P-1

非平滑多目的最適化に対する近接勾配法
Proximal gradient methods for nonsmooth multiobjective optimization.

○宮崎裕貴 [1]
*Yuki Miyazaki[1]

Abstract: This study optimizes a multi-objective function whose components are determined by the sum of smooth and non-smooth convex functions. Previous research assumed Lipschitz continuity for the gradient of each smooth function. They constructed an algorithm to search for Pareto stationary points of the objective function and performed a convergence analysis of the method. In this study, we introduced a line search into this method which guarantees to find a weakly Pareto optimal points of the objective function even when the gradient is not Lipschitz continuous.

1. はじめに

多目的最適化とは, 複数の目的関数を同時に最適化する問題のことである. ほとんどの場合で, ある一つの点で全ての目的関数を最小化することは不可能である. そこで, 次に示す Pareto 解や弱 Pareto 解, Pareto 停留点という概念が重要になる [1].

m 個の目的関数 F_1 から $F_m(F_i:\mathbb{R}^n\to\mathbb{R})$ を並べたベクトル値関数
$F:\mathbb{R}^n\to\mathbb{R}^m$, $F(x)=(F_1(x),F_2(x),\cdots,F_m(x))^\top$ を考える. ある $x^\star\in\mathbb{R}^n$ に対して, $F(x)\le F(x^\star)$ かつ $F(x)\ne F(x^\star)$ を満たす $x\in\mathbb{R}^n$ が存在しないとき, $x^\star$ は F の **Pareto 解**という. ある $x^\star\in\mathbb{R}^n$ に対して, $F(x)<F(x^\star)$ を満たす $x\in\mathbb{R}^n$ が存在しないとき, $x^\star$ は F の**弱 Pareto 解**という. 任意の $d\in\mathbb{R}^n$ に対して $\min\limits_{i\in[1:m]}F_i'(\bar{x};d)\ge 0$ が成り立つとき, $\bar{x}$ は F の **Pareto 停留点**という (ただし $[1:m]:=\{1,2,\cdots,m\}$).

補題 1 $x^\star$ が F の Pareto 解ならば弱 Pareto 解であり, $x^\star$ が F の弱 Pareto 解ならば Pareto 停留点である. 加えて $F_1,\cdots,F_m$ が全て凸関数のとき, $x^\star$ が F の弱 Pareto 解であることの必要十分条件は $x^\star$ が Pareto 停留点であることである.

本研究では, F の各成分を $F_i(x)=f_i(x)+g_i(x)$, このうち f_i は連続微分可能で, g_i は閉真凸関数とする. 提案手法とその解析では, 先行研究 [2] でも用いられた多目的最適化に対する二つのメリット関数

$$w_\ell(x):=\max_{y\in\mathbb{R}^n}\left[\min_{i\in[1:m]}\{\nabla f_i(x)^\top(x-y)+g_i(x)-g_i(y)\}\right.$$
$$\left.-\frac{\ell}{2}\|y-x\|^2\right],\quad \ell>0$$
$$u_0(x):=\sup_{y\in\mathbb{R}^n}\min_{i\in[1:m]}(F_i(x)-F_i(y))$$

を用いる. この二つのメリット関数には次の事実が知られている [3].

補題 2 任意の $x\in\mathbb{R}^n$ に対して $w_\ell(x)$, $u_0(x)\ge 0$ であり, $w_\ell(\bar{x})=0$ と $\bar{x}$ が F の Pareto 停留点であること, $u_0(x^\star)=0$ と $x^\star$ が F の弱 Pareto 解であることがそれぞれ同値である.

この事実より, 多目的関数の最小化のため, $w_\ell(x)$ や $u_0(x)$ の ϵ 解を求めることにする.

先行研究 [2] では全ての f_i の勾配に Lipschitz 連続性 ($\exists M_i>0\,s.t.\,\forall x,y\in\mathbb{R}^n\ \|\nabla f_i(x)-\nabla f_i(y)\|\le M_i\|x-y\|$) を仮定して手法を提案した. その手法では各 Lipschitz 定数 M_i の最大値がわかっていることを仮定して計算に用いられ, 収束解析が行われた.

補題 3 スカラー値関数 $f:\mathbb{R}^n\to\mathbb{R}$ が連続微分可能で, $\nabla f(x)$ が M-Lipschitz 連続であるとき,

$$f(y)\le f(x)+\nabla f(x)^\top(y-x)+\frac{M}{2}\|y-x\|^2,\quad \forall x,y\in\mathbb{R}^n$$

が成り立つ.

一方我々は, 各 f_i の勾配の Lipschitz 定数を用いない手法を提案し, その解析を行った. さらに f の仮定から勾配の Lipschitz 連続性を外し, 新たに各 f_i そのものに Lipschitz 連続性を仮定したときも検討した.

補題 4 スカラー値関数 $f:\mathbb{R}^n\to\mathbb{R}$ が凸かつ $\frac{M}{2}$-Lipschitz 連続であるとき, 任意の劣勾配 $g\in\partial f(x)$ に対して

$$f(y)\le f(x)+g^\top(y-x)+M\|y-x\|,\quad \forall x,y\in\mathbb{R}^n$$

が成り立つ.

この場合も同じ手法で弱 Pareto 解の近似解を求められることを保証し, 計算量の解析をおこなった.

1: 日大理工・院 (前)・数学

2. 提案手法

初期点 $x^0 \in D := \bigcap_{i\in[1:m]} \mathrm{dom}(F_i)$ と $\epsilon > 0, \ell_{-1} \geq 2$ をとり, 各 $k = 0, 1, 2, \ldots$ に対して, 以下の反復によって $\{x^k\}$ を生成する.

$$d^k \leftarrow \operatorname*{argmin}_{d\in\mathbb{R}^n} \left[\max_{i\in[1:m]} \{\nabla f_i(x^k)^\top d + g_i(x^k + d) - g_i(x^k)\} + \frac{\ell_k}{2}\|d\|^2 \right],$$

$$x^{k+1} \leftarrow x^k + d^k$$

ただし, ℓ_k は等比数列 $\{2^i \ell_k^0\}_{i\geq 0}$ ($\ell_k^0 = \max\{1, \ell_{k-1}/2\}$) から選び,
条件 (P): 任意の $i \in [1:m]$ に対して

$$F_i(x^{k+1}) - F_i(x^k) \leq \nabla f_i(x^k)^\top (x^{k+1} - x^k) + g_i(x^{k+1}) - g_i(x^k) + \frac{\ell_k}{2}\|d^k\|^2 + \frac{\epsilon}{2}$$

が成り立つものを採用する (バックトラッキング).

3. 計算量の解析（主結果）

計算量の解析に必要な次の仮定 5 を課す.

仮定 5 全ての f_i が凸関数, g_i が閉真凸関数であり $\forall x \in \Omega_F(x^0)\, \exists x^* \in X^*\, s.t.\, F(x^*) \leqq F(x)$ かつ

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

(ただし X^* は F の弱 Pareto 解集合,
$\Omega_F(F(x^0)) := \{x \in \mathbb{R}^n | F(x) \leq F(x^0)\}$ は $F(x^0)$ に対するレベル集合) と仮定する.

定理 6 $x^0 \in D, \epsilon > 0$ とする. 各 f_i が連続微分可能で ∇f_i が M_i-Lipschitz 連続かつ仮定 5 を満たす F に対して提案手法の反復回数 K が,

$$K \geqq 2R\epsilon^{-1} \max_{i\in[1:m]} M_i$$

を満たすとき

$$u_0(x^K) \leqq \epsilon.$$

定理 7 $x^0 \in D, \epsilon > 0$ とする. 各 f_i が $\frac{M_i}{2}$-Lipschitz 連続かつ仮定 5 を満たす F に対して提案手法の反復回数 K が,

$$K \geqq 2R\epsilon^{-2} \max_{i\in[1:m]} M_i^2$$

を満たすとき

$$u_0(x^K) \leqq \epsilon.$$

4. 先行研究との比較

	f_i の仮定	∇f_i の仮定	g_i の仮定	手法で用いる定数	収束性	収束率
Tanabe et al. [2]	凸	Lipschitz 連続	閉真凸	f_i の Lipschitz 定数	$u_0(x^K) \leq \epsilon$	$O(\epsilon^{-1})$
提案手法, 定理 5	凸	Lipschitz 連続	閉真凸	**不要**	$u_0(x^K) \leq \epsilon$	$O(\epsilon^{-1})$
提案手法, 定理 6	凸, Lipschitz 連続	なし	閉真凸	**不要**	$u_0(x^K) \leq \epsilon$	$O(\epsilon^{-2})$

5. まとめと今後の展望

本研究では, 直線探索付きの多目的近接勾配法を提案した. 提案手法は f が平滑のときに限らず, f が非平滑かつ Lipschitz 連続性をもつときも f の弱 Pareto 解の近似解を求められることを保証した. またその近似解を求める計算量を解析した.

提案手法の中で d^k を求めるには部分問題の凸最適化を繰り返すが, その部分問題が 1 回で済む手法はあるかを探る ([2] では探索方向を 1 回求めて, ステップサイズのみ調整されている). また提案手法の計算は入力される ϵ に依存しているため, メリット関数 $w_1(x), u_0(x)$ が 0 に収束するという保証がない. メリット関数を収束させられる手法にできるかを検討していく.

6. 参考文献

[1] H. Tanabe, E. H. Fukuda, and N. Yamashita, Proximal gradient methods for multiobjective optimization and their applications, *Comput. Optim. Appl.* **72**(2019) 339–361.

[2] H. Tanabe, E. H. Fukuda, and N. Yamashita, Convergence rates analysis of a multiobjective proximal gradient method, *Optim. Lett.* **17**(2023) 333–350.

[3] H. Tanabe, E. H. Fukuda, and N. Yamashita, New merit functions for multiobjective optimization and their properties. *Optim.* **73**(2024) 3821-3858.