Gradient mapping 型停留点評価基準に基づく近接勾配型アルゴリズムに関する一考察

A Note on Proximal Gradient Descent-type Algorithm Based on Gradient Mapping-type Stationarity Measure

久米 啓太 (Keita Kume)[1], 山田 功 (Isao Yamada)[1],
[1] 東京科学大学工学院情報通信系
(Dept. of Information and Communications Engineering, Institute of Science Tokyo)
e-mail : {kume, isao}@sp.ict.e.titech.ac.jp

1 はじめに

小文では，Euclid 空間 $\mathcal{X}$ で定式化された非平滑非凸最適化問題を考える:

$$\text{find } \boldsymbol{x}^\star \in \operatorname*{argmin}_{\boldsymbol{x}\in\mathcal{X}} (F+\phi)(\boldsymbol{x}). \tag{1}$$

ただし，$\phi : \mathcal{X} \to (-\infty, +\infty]$ と $F : \mathcal{X} \to \mathbb{R}$ は以下の条件を満たすものと仮定する．

1) ϕ は下半連続な真関数であり，$\mathrm{dom}(\phi) := \{\boldsymbol{x} \in \mathcal{X} \mid \phi(\boldsymbol{x}) \in \mathbb{R}\}$ 上で，連続かつ *prox-regular* [1, Def. 13.27] であり，ある $\gamma_\phi \in (0, +\infty]$ が存在して，任意の $\gamma \in (0, \gamma_\phi)$ に対し，$\phi(\cdot) + \frac{1}{2\gamma}\|\cdot\|^2$ は下に有界である．さらに，**近接写像** $\mathrm{prox}_{\gamma\phi} : \mathcal{X} \rightrightarrows \mathcal{X} : \bar{\boldsymbol{x}} \mapsto \operatorname*{argmin}_{\boldsymbol{x}\in\mathcal{X}} \left(\phi(\boldsymbol{x}) + \frac{1}{2\gamma}\|\boldsymbol{x} - \bar{\boldsymbol{x}}\|^2\right)$ は任意の $\gamma \in (0, \gamma_\phi)$ と $\bar{\boldsymbol{x}} \in \mathcal{X}$ で 少なくとも 1 点は容易に計算できるものとする．
2) F は $\mathrm{dom}(\phi)$ 上で，局所 Lipschitz 連続かつ $\partial_F F = \partial_L F$ を満たす．ただし，$\partial_F F$ と $\partial_L F$ は，F の *Fréchet (regular)* 劣微分と *Limiting (general)* 劣微分を表す [1, Def. 8.3].

限定的な条件下での最適化問題 (1) に対して，例えば，(i) F が連続微分可能かつ勾配 ∇F が Lipschitz 連続である場合，あるいは (ii) F が弱凸関数と可微分写像の合成であり，ϕ が凸関数である場合には，それぞれ近接勾配法 [2] と近接可変平滑化法 [3] が提案されている．一方，ϕ と F の凸性・微分可能性を仮定しない最適化問題 (1) の解法については，近年検討され始めている [4] が，既存解法は各反復で子問題を厳密に解く必要があり，計算機実装の観点から現実的な解法とはいえない．

小文では，最適化問題 (1) に対して，F の**平滑化関数** F_n $(n \in \mathbb{N})$ が利用可能である仮定 1.1の下で，「厳密解の逐次近似が必要な子問題」を持たない近接勾配型アルゴリズムを提案する．

仮定 1.1 (F の平滑化関数)**.** 最適化問題 (1) を考える．総和不可能な 0 への単調非増加列 $(\mu_n)_{n=1}^\infty \subset \mathbb{R}_{++}$ と関数列 $F_n : \mathcal{X} \to \mathbb{R}$ $(n \in \mathbb{N})$ は，以下の条件を満たす: (i) $\lim_{n\to\infty} F_n(\bar{\boldsymbol{x}}) = F(\bar{\boldsymbol{x}})$ $(\bar{\boldsymbol{x}} \in \mathrm{dom}(\phi))$; (ii) 各 F_n は連続微分可能であり，勾配 $\nabla F_n : \mathcal{X} \to \mathcal{X}$ は $\mathrm{dom}(\phi)$ の凸包上で $L_{\nabla F_n}$-Lipschitz 連続である; (iii) ある定数 $\varpi_1, \varpi_2 \in \mathbb{R}_{++}$ が存在し，$L_{\nabla F_n}$ は $L_{\nabla F_n} := \varpi_1 + \varpi_2\mu_n^{-1}$ $(n \in \mathbb{N})$ と表せる; (iv) 任意の収束列 $(\boldsymbol{x}_n)_{n=1}^\infty \subset \mathrm{dom}(\phi)$ に対して，$(\nabla F_n(\boldsymbol{x}_n))_{n=1}^\infty$ は有界列であり，その任意の集積点 $\boldsymbol{v} \in \mathcal{X}$ は $\boldsymbol{v} \in \partial_L F\left(\lim_{n\to\infty} \boldsymbol{x}_n\right)$ を満たす; (v) ある定数 $M \in \mathbb{R}_{++}$ が存在して，$0 \leqslant F_{n+1}(\bar{\boldsymbol{x}}) - F_n(\bar{\boldsymbol{x}}) \leqslant M(\mu_n - \mu_{n+1})$ $(n \in \mathbb{N}, \bar{\boldsymbol{x}} \in \mathrm{dom}(\phi))$ が成立する (F_n の例は例 1.2参照).

さらに，*Gradient mapping* **型停留点評価基準** (ϕ が凸関数の場合については [3, 式 (10)] 参照)

$$(\gamma \in (0, \gamma_\phi)) \quad \mathcal{M}_\gamma^{F_n,\phi} : \mathcal{X} \to \mathbb{R} : \bar{\boldsymbol{x}} \mapsto \inf\left\{\left\|\frac{\bar{\boldsymbol{x}} - \boldsymbol{p}}{\gamma}\right\| \mid \boldsymbol{p} \in \mathrm{prox}_{\gamma\phi}(\bar{\boldsymbol{x}} - \gamma\nabla F_n(\boldsymbol{x}_n))\right\} \tag{2}$$

に基づく漸近的な評価 (定理 2.1参照) を行うことにより，提案法の生成列について，**停留点** $\boldsymbol{x}^\star \in \mathrm{dom}(\phi)$ $\left(\overset{\mathrm{Def}}{\Leftrightarrow} \partial_F(F+\phi)(\boldsymbol{x}^\star) \ni \mathbf{0}\right)$ に関する収束解析を与えている (定理 2.2参照).

例 1.2 (平滑化関数の構成法)**.** 平滑化関数の構成法についてはいくつか報告されている．例えば，Lipschitz 連続な $\eta(>0)$-弱凸関数 $g:\mathcal{Z}\to\mathbb{R}$ ($\mathcal{Z}$ は Euclid 空間) と可微分写像 $\mathfrak{S}:\mathcal{X}\to\mathcal{Z}$ からなる合成 $F:=g\circ\mathfrak{S}$ に対して，g の Moreau 包絡関数 ${}^{\mu}g(\bar{z}) := g(\mathrm{prox}_{\mu g}(\bar{z})) - \frac{1}{2\mu}\left\|\mathrm{prox}_{\mu g}(\bar{z})-\bar{z}\right\|^2$ によって定義される $F_n := {}^{\mu_n}g\circ\mathfrak{S}$ と $(\mu_n)_{n=1}^{\infty} := ((2\eta)^{-1}n^{-1/\alpha})$ $(\alpha \geqslant 1)$ を考える．このとき，$\mathfrak{S}$ とその Gâteaux 微分が $\mathrm{dom}(\phi)$ 上で Lipschitz 連続ならば，仮定 1.1を満たす [3]．また，g の近接写像が容易に計算できる場合には，F_n の関数値と勾配値も容易に計算できる．

2 Gradient mapping 型停留点評価基準に基づく提案法とその収束解析

$\mathcal{M}_{\gamma_n}^{F_n,\phi}$(式 (2) 参照) と最適化問題 (1) の停留点の間には，以下の関係が成立することを示す．

定理 2.1 ($\mathcal{M}_{\gamma_n}^{F_n,\phi}$ の漸近評価)**.** 最適化問題 (1) に対して，仮定 *1.1*を満たすものとする．このとき，$(\gamma_n)_{n=1}^{\infty} \subset (0,\gamma_\phi)$ と $\bar{\boldsymbol{x}} \in \mathrm{dom}(\phi)$ への収束列 $(\boldsymbol{x}_n)_{n=1}^{\infty} \subset \mathrm{dom}(\phi)$ に対して，$\liminf_{n\to\infty} \mathcal{M}_{\gamma_n}^{F_n,\phi}(\boldsymbol{x}_n) = 0$ が成立するならば，$\bar{\boldsymbol{x}}$ は最適化問題 (1) の停留点である *(*つまり $\partial_F(F+\phi)(\bar{\boldsymbol{x}}) \ni \mathbf{0}$ を満たす*)*.

定理 2.1より，仮定 1.1の下で，$\liminf_{n\to\infty} \mathcal{M}_{\gamma_n}^{F_n,\phi}(\boldsymbol{x}_n) = 0$ を達成する点列 $(\boldsymbol{x}_n,\gamma_n) \subset \mathrm{dom}(\phi)\times(0,\gamma_\phi)$ を生成することが第一目標となる．この考察に基づき，以下の近接勾配型アルゴリズムを提案する．

提案法 ($\boldsymbol{x}_1 \in \mathrm{dom}(\phi), c \in (0,1/2), \rho \in (0,1), \widetilde{\gamma} \in (0,\gamma_\phi)$ は任意に与えられている)

各 $n\in\mathbb{N}$ において，以下の条件を満たす $m \in \mathbb{N}\cup\{0\}$ 全体の集合 $\mathfrak{M}_n \subset \mathbb{N}\cup\{0\}$ を考える:

$$\left(\exists \boldsymbol{x}^{\langle m\rangle} \in \mathrm{prox}_{\widetilde{\gamma}\rho^m\phi}(\boldsymbol{x}_n - \widetilde{\gamma}\rho^m \nabla F_n(\boldsymbol{x}_n))\right) \quad (F_n+\phi)(\boldsymbol{x}^{\langle m\rangle}) \leqslant (F_n+\phi)(\boldsymbol{x}_n) - c\widetilde{\gamma}\rho^m \left(\mathcal{M}_{\widetilde{\gamma}\rho^m}^{F_n,\phi}(\boldsymbol{x}_n)\right)^2.$$

$\bar{m} := \min \mathfrak{M}_n$ に対応する $\boldsymbol{x}^{\langle \bar{m}\rangle} \in \mathrm{dom}(\phi)$ を用いて，$\boldsymbol{x}_{n+1} := \boldsymbol{x}^{\langle \bar{m}\rangle}$，$\gamma_n := \widetilde{\gamma}\rho^{\bar{m}}$ とする．

仮定 1.1の下では，提案法内の各 $\mathfrak{M}_n$ は非空となり，$(\boldsymbol{x}_n,\gamma_n)_{n=1}^{\infty}$ は well-defined に生成される．提案法内の $\bar{m}$ は直線探索法によって有限回の反復で探索できるため，提案法は,「厳密解の逐次近似が必要な子問題」を持たない解法といえる．最後に，提案法の収束解析に関する定理を示す．

定理 2.2 (収束解析)**.** 仮定 *1.1*の下で提案法によって生成される $(\boldsymbol{x}_n,\gamma_n) \subset \mathrm{dom}(\phi)\times(0,\gamma_\phi)$ について，$\liminf_{n\to\infty} \mathcal{M}_{\gamma_n}^{F_n,\phi}(\boldsymbol{x}_n) = 0$ が成立する．さらに，$\lim_{j\to\infty} \mathcal{M}_{\gamma_{n_j}}^{F_{n_j},\phi}(\boldsymbol{x}_{n_j}) = 0$ を満たす部分列 $(\boldsymbol{x}_{n_j})_{j=1}^{\infty}$ の集積点 $\bar{\boldsymbol{x}} \in \mathcal{X}$ が $\bar{\boldsymbol{x}} \in \mathrm{dom}(\phi)$ を満たせば，定理 *2.1*により $\bar{\boldsymbol{x}}$ は最適化問題 (1) の停留点である．

謝辞　本研究は，日本学術振興会科学研究費助成事業 (24K23885) の助成を受けた．

参考文献

[1] R. Rockafellar and R. J.-B. Wets, *Variational Analysis*, 3rd ed.　Springer, 2010.

[2] E. Pauwels, “A note on stationarity in constrained optimization,” *arXiv:2402.09831*, 2024.

[3] K. Kume and I. Yamada, “A proximal variable smoothing for minimization of nonlinearly composite nonsmooth function – maxmin dispersion and MIMO applications,” *arXiv:2506.05974*, 2025.

[4] D. Davis, D. Drusvyatskiy, and Z. Shi, “Stochastic optimization over proximally smooth sets,” *SIAM J. Optim.*, 2025.

Hölder 連続勾配をもつ多目的最適化問題に対する近接勾配法

Proximal gradient methods for multiobjective optimization problems with Hölder continuous gradients.

宮崎 裕貴 (Yuki Miyazaki)[1], 伊藤 勝 (Masaru Ito)[2],
柳下 翔太郎 (Shotaro Yagishita)[3]
[1,2] 日本大学 (Nihon University)
[3] 統計数理研究所 (The Institute of Statistical Mathematics)
e-mail : csyi25012@g.nihon-u.ac.jp

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, $F(x)\leqq F(x^\star)$ かつ $F(x)\neq F(x^\star)$ 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 $\min\limits_{i\in[1:m]}F_i'(\bar{x};d)>0$ が成り立つとき, $\bar{x}$ は F の **Pareto 停留点**という (ただし $[1:m]:=\{1,2,\cdots,m\}$).

$x^\star$ 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 Pareto 停留点ならば弱 Pareto 解である.

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 $\cdots$(1)

とする. 提案手法とその解析では, 先行研究でも用いられた多目的最適化に対して知られている二つのメリット関数

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

$$u_0(x):=\sup_{y\in\mathbb{R}^n}\min_{i\in[1:m]}(F_i(x)-F_i(y))$$

を用いる. 任意の $x\in\mathbb{R}^n$ に対して $w_\ell(x)$, $u_0(x)\geqq 0$ 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 解を求めることになる.

先行研究 [2] では全ての f_i の勾配に Lipschitz 連続性 ($\forall i\in[1:m]\,\exists L_i>0\,s.t.$ $\forall x,y\in\mathbb{R}^n\,\|\nabla f_i(x)-\nabla f_i(y)\|\leqq L_i\|x-y\|$) を仮定して手法の提案と解析が行われた. 一方我々は,

各 f_i の勾配の Hölder 連続性 $\cdots$(2)

($\forall i\in[1:m]\,\exists M_i>0,\exists\nu_i\in(0,1]\,s.t.\,\forall x,y\in\mathbb{R}^n\,\|\nabla f_i(x)-\nabla f_i(y)\|\leqq M_i\|x-y\|^{\nu_i}$) に一般化して解析を行った. 勾配に Hölder 連続性を仮定した多目的非凸関数への手法とその解析は [3] によって行われているが, 我々は新たに f_i と g_i すべてが凸関数のときに対応した手法を提案し, 解析も行った.

2　提案手法

本研究では, f_i が凸関数とは限らないときと, すべての f_i が凸関数のときにそれぞれアルゴリズムを提案する.

$$\text{条件 (P) を, 任意の } i \in [1:m] \text{ に対して } F_i(x^{k+1}) - F_i(x^k) \leqq -w_\ell(x^k) + \frac{\epsilon}{2\ell},$$

条件 (Q) を, 任意の $i \in [1:m]$ に対して

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

が成り立つこととする.

提案アルゴリズム. 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.

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

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) [$f_i$ が凸のとき (Q)] が成り立つものを採用する (バックトラッキング).

3　停止性の解析

定理 1　$x^0 \in D, \epsilon > 0$ とする. 仮定 (1)(2) を満たす F に対して条件 (P) を用いた提案手法の反復回数 K が $K \geqq \max_{i\in[1:m]} 4 \left(\frac{1-\nu_i}{1+\nu_i}\right)^{\frac{1-\nu_i}{2\nu_i}} \epsilon^{-\frac{1+\nu_i}{2\nu_i}} M_i^{\frac{1}{\nu_i}} \cdot \min_{j\in[1:m]} (F_j(x^0) - F_j^\star)$ を満たすとき $\min_{k\in[1:K]} w_1(x^k) \leqq \epsilon$ が成り立つ. さらに (1)(2) に加え, 全ての f_i が凸関数であり $\forall x \in \Omega_F(x^0)$ $\exists x^* \in X^*\, s.t.\, F(x^*) \leqq F(x)$ かつ $R := \sup_{F^*\in F(X^*\cup\Omega_F(F(x^0)))} \inf_{x\in F^{-1}(\{F^*\})} \|x - x^0\|^2 < \infty$ (ただし X^* は F の弱 Pareto 解集合, $\Omega_F(x^0)$ を $F(x^0)$ に対するレベル集合) と仮定する. 反復回数 K が, $K \geqq 2R \max_{i\in[1:m]} \left(\frac{1-\nu_i}{1+\nu_i}\right)^{\frac{1-\nu_i}{1+\nu_i}} M_i^{\frac{2}{1+\nu_i}} \epsilon^{-\frac{2}{1+\nu_i}}$ を満たすとき $u_0(x^K) \leqq \epsilon$.

4　今後の展望

今回の解析は点列の反復回数のみに着目して行ったが, 手法内の子問題に d^k を求める直線探索付き凸最適化があり, 今後この計算回数もふまえた計算量をさらに解析する. またこの子問題を別の計算量を減らせる手法に変更できるかを探っていきたい.

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] M. E. Pinheiro and G. N. Grapiglia, Universal nonmonotone line search method for nonconvex multiobjective optimization problems with convex constraints, *Comput. Optim. Appl.* **44**(2025) No.56.

Iterative Feedback Tuningにおけるバックトラッキングを用いた勾配法

Gradient Method with Backtracking in Iterative Feedback Tuning

豊田 充 (Mitsuru Toyoda)[1],
増田 士朗 (Shiro Masuda)[2]
[1] 長岡技術科学大学 (Nagaoka University of Technology), [2] 東京都立大学 (Tokyo Metropolitan University)
e-mail：toyoda@vos.nagaokaut.ac.jp

1 問題設定

下記の 1 入力 1 出力系にフィードバック制御器 $C(z,\boldsymbol{\rho})$ を適用した状況を考える:

$$y[k]=G(z)u[k],\ u[k]=C(z,\boldsymbol{\rho})(r[k]-y[k]),$$

上式中 $y[k]$ がシステムの出力，$u[k]$ が入力，$r[k]$ が参照信号であり，制御器 $C(z,\boldsymbol{\rho})$ は $C(z,\boldsymbol{\rho})=\boldsymbol{\rho}^\top\boldsymbol{\varphi}(\boldsymbol{z})$ なる線形な構造をもつとする．簡単のため，制御対象に加わる外乱は無視できるものとした．閉ループ系応答は閉ループ伝達関数 $T(z,\boldsymbol{\rho})=G(z)C(z,\boldsymbol{\rho})/\big(1+G(z)C(z,\boldsymbol{\rho})\big)$ を用いて $y[k]=T(z,\boldsymbol{\rho})r[k]$ と定まる．この応答と目標となる規範モデル $T_{\mathrm{d}}(z)$ による応答の 2 乗誤差を目的関数とした最小化問題を考える:

$$\underset{\boldsymbol{\rho}\in\mathbb{R}^{n_\rho}}{\operatorname{minimize}}\, J_y^N(\boldsymbol{\rho}):=\frac{1}{N}\sum_{k=0}^{N-1}(y[k]-y_{\mathrm{d}}[k])^2=\frac{1}{N}\sum_{k=0}^{N-1}\left((T(z,\boldsymbol{\rho})-T_{\mathrm{d}}(z))r[k]\right)^2. \tag{1}$$

制御対象つまり伝達関数 $G(z)$ の表式が既知であれば非線形最適化手法で求解が可能だが，データ駆動制御の問題設定では，制御対象の情報は用いず，観測データ系列 $\{y[k]\}$ のみを用いた形式の最適化アルゴリズムが対象となる．加えて，データの観測には新たな制御器パラメータの実装および応答の観測が必要となり工数の増加につながるため，多数の目的関数評価を伴うヒューリスティックやブラックボックス最適化手法は非現実的となる．

上記の背景から，あるパラメータを実装した状況での 1 回の実験および追加の 1 回の実験の計 2 回で勾配を厳密に求める方法論が IFT[1] であり，勾配ベースの手法が適用可能となる．一方で，目的関数 [2] が非凸となることから収束解析はいまだ挑戦的な課題である．先行研究 [2, 3] では制御分野でよく用いられてきたアルゴリズムの収束先つまり最適解との距離を用いてリアプノフ関数と対応するステップサイズを導出しているが，ステップサイズは制御対象の情報を必要としており実行可能性に難点がある．減衰するステップサイズ [4]，準ニュートン法つまり目的関数のヘッシアンの推定値 [5] と勾配法の組合せが先行研究で検討されており，数値計算によって優位性が検証されているが，やはり収束の理論的保証は困難である．

2 主な結果

従来の勾配法と同じく，バックトラッキングによるステップサイズの調整を行うことで，いわゆる gradient descent 条件

$$J_y(\boldsymbol{\rho}^{(i+1)})-J_y(\boldsymbol{\rho}^{(i)})\le-\frac{1}{2L}\|\nabla J_y(\boldsymbol{\rho}^{(i)})\|_2^2$$

が満たされることとなり，停留点への収束がいえることを指摘した．目的関数のリプシッツ平滑性は目的関数の直接的な微分計算から従う．バックトラッキングは目的関数値の観測つまり制御パラメータを実装した上での実験のみで実行できることから，完全なモデルフリーの形で収束性をいうことができる．

加えて，完全なモデルマッチングを達成する制御器は閉ループ伝達関数の表式から逆算することで $C_{\mathrm{d}}(z) = T_{\mathrm{d}}(z)/(G(z)(1 - T_{\mathrm{d}}(z)))$ となるが，この理想的な制御器 $C_{\mathrm{d}}(z)$ が $C(z, \boldsymbol{\rho}) = \boldsymbol{\rho}^\top \boldsymbol{\varphi}(\boldsymbol{z})$ の構造で表されることは現実的に考えにくい．$C_{\mathrm{d}}(z) = \boldsymbol{\rho}_{\mathrm{d}}^\top \boldsymbol{\varphi}(z) + \boldsymbol{\theta}_{\mathrm{d}}^\top \boldsymbol{\psi}(z)$ と表現できる場合を考える，つまり制御器

$$\widetilde{C}(z, \boldsymbol{\rho}, \boldsymbol{\theta}) = C(z, \boldsymbol{\rho}) + D(z, \boldsymbol{\theta}) = \boldsymbol{\rho}^\top \boldsymbol{\varphi}(z) + \boldsymbol{\theta}^\top \boldsymbol{\psi}(z)$$

のクラスに拡張し検討を行った．このテクニックは別のデータ駆動制御アルゴリズムで検討された例 [6] があるが，IFT に関しては著者の知る限り見当たらない．具体的な収束レートの導出を目的として，Polyak–Łojasiewicz (PŁ) 不等式と呼ばれる，収束解析で重要な役割を果たす不等式が近似的に成りたつ．特にマッチングが達成される場合には厳密に PŁ 不等式が成りたち，勾配法による 1 次収束が従う．

3 まとめ

完全なモデルフリーでの制御器パラメータ調整を意図し，バックトラッキングを用いた勾配法の適用を本研究では検証した．ステップサイズ探索における目的関数値の評価を抑えるための，データ駆動予測や FRIT，VRFT をはじめとする別のデータ駆動制御関連手法の援用の検討が将来的な課題である．

謝辞　本研究は JSPS 科研費 JP22K14279 の助成を受けたものです．

参考文献

[1] H. Hjalmarsson, M. Gevers, S. Gunnarsson, and O. Lequin, “Iterative feedback tuning: theory and applications,” *IEEE Control Systems Magazine*, vol. 18, no. 4, pp. 26–41, 1998.

[2] A. S. Bazanella, M. Gevers, L. Mišković, and B. D. Anderson, “Iterative minimization of H_2 control performance criteria,” *Automatica*, vol. 44, no. 10, pp. 2549–2559, 2008.

[3] D. Eckhard and A. S. Bazanella, “Optimizing the convergence of data-based controller tuning,” in *Proceedings of the 2009 European Control Conference (ECC)*, pp. 910–915, 2009.

[4] J. K. Huusom, N. K. Poulsen, and S. B. Jørgensen, “Improving convergence of iterative feedback tuning,” *Journal of Process Control*, vol. 19, no. 4, pp. 570–578, 2009.

[5] D. Tesch, D. Eckhard, and A. S. Bazanella, “Iterative feedback tuning for cascade systems,” in *Proceedings of the 2016 European Control Conference (ECC)*, pp. 495–500, 2016.

[6] M. Campi, A. Lecchini, and S. Savaresi, “Virtual reference feedback tuning: a direct method for the design of feedback controllers,” *Automatica*, vol. 38, no. 8, pp. 1337–1346, 2002.

リーマン多様体上の不等式制約付き最適化問題に対する信頼領域内点法

A trust region-type interior point method for Riemannian optimization problem with inequality constraints

小原 光暁 (Mitsuaki Obara)[1], 奥野 貴之 (Takayuki Okuno)[2],
武田 朗子 (Akiko Takeda)[1,3]
[1] 東京大学 (University of Tokyo), [2] 成蹊大学 (Seikei University), [3] 理研 AIP(RIKEN AIP)
e-mail : takayuki-okuno@st.seikei.ac.jp

1 はじめに

本研究では, 次のようなリーマン多様体上の不等式制約付き最適化問題を考える.

$$\begin{aligned} &\underset{x \in \mathcal{M}}{\text{minimize}} && f(x) \\ &\text{subject to} && g_i(x) \geq 0, \quad i \in \mathcal{I} \coloneqq \{1, \ldots, m\}. \end{aligned} \tag{1}$$

ここで $\mathcal{M}$ は d 次元の連結な完備リーマン多様体であり, $f\colon \mathcal{M} \to \mathbb{R}$, $g_i\colon \mathcal{M} \to \mathbb{R}$ $(i \in \mathcal{I})$ は $\mathcal{M}$ 上の 2 回連続的微分可能な実数値関数である. リーマン多様体上の無制約最適化が集中的に研究されてきた中, この数年で制約付きの問題が注目を集めてきており, ユークリッド空間上の制約付き最適化のための代表的手法である拡張ラグランジュ法, 逐次 2 次最適化法, 主双対内点法がそれぞれリーマン多様体へと拡張されている. 本研究の主貢献は, 信頼領域法を組込むことで問題 (1) の 2 次の最適性条件を満たす点へ大域的収束をする主双対内点法を開発したことであり, 加えて同手法が局所最適解への局所的概 2 次収束性をもつことを示したことにある. 以下では, 提案手法の概要のみを示すが, 詳細は大域的収束性については [2] を, 概 2 次収束性については [1] を参照されたい.

2 提案手法の概略

提案手法は, 問題 (1) の局所最適解が満たすべき, 以下のような 2 次の最適性条件を満たす点 $\omega^* := (x^*, (y_i^*)_{i \in \mathcal{I}}) \in \mathcal{M} \times \mathbb{R}^m$ を見つけることを目的とする.

$$\begin{aligned} &\mathrm{grad}_x \mathcal{L}(\omega^*) = 0_{x^*},\ y_i^* \geq 0,\ g_i(x^*) \geq 0,\ y_i^* g_i(x^*) = 0\ (i \in \mathcal{I}), \\ &\langle \mathrm{Hess}_x \mathcal{L}(\omega^*)[\xi_{x^*}], \xi_{x^*} \rangle_{x^*} \geq 0\ (\xi_{x^*} \in \mathcal{C}(\omega^*)). \end{aligned} \tag{2}$$

ただし, $\mathcal{L}(\omega) \coloneqq f(x) - \sum_{i \in \mathcal{I}} y_i g_i(x)$ はラグランジュ関数であり, grad_x, Hess_x は $x \in \mathcal{M}$ におけるリーマン多様体の意味での勾配とヘッシアンを表し, $\mathcal{C}(\omega^*) \subseteq \mathcal{M} \times \mathbb{R}^m_+$ は ω^* における臨界錐 (critical cone) である. 条件 (2) の第 1 行はカルーシュ・キューン・タッカー（KKT）条件に他ならず, 既存手法たちはこの KKT 条件のみを満たす点への収束保証がある.

2.1 外部反復

提案手法は外部反復と内部反復より構成される. まず外部反復では, バリアパラメータと呼ばれる $\lim_{k \to \infty} \mu_k = 0$ であるような点列 $\{\mu_k\} \subseteq \mathbb{R}_{++}$ と原点で 0 をとるような連続関数 $\varepsilon_{\mathcal{L}}, \varepsilon_C, \varepsilon_S\colon \mathbb{R}_+ \to \mathbb{R}_+$ を導入して, 各反復で条件 (2) に対する近似条件である (3)-(4) を満足する $\omega_{k+1} := (x_{k+1}, y_{k+1}) \in \mathcal{M} \times \mathbb{R}^m_+$ を求める. $k \to \infty$ とすれば, これらの条件式は極限において (2)

と同値である.

$$\left\|\mathrm{grad}_x \mathcal{L}(\omega_{k+1})\right\|_{x_{k+1}} \leq \varepsilon_{\mathcal{L}}(\mu_k),\ \|Y_{k+1} g(x_{k+1}) - \mu_k \mathbf{1}\| \leq \varepsilon_C(\mu_k), \tag{3}$$

$$\lambda^{\min}\left[H_{k+1}\right] \geq -\varepsilon_S(\mu_k),\ g(x_{k+1}) > 0,\ y_{k+1} > 0. \tag{4}$$

ただし, ここで $\lambda^{\min}$ は最小固有値を, $Y_{k+1} := \mathrm{diag}\big(y_{k+1}\big)$, $H_{k+1} := H(\omega_{k+1})$, $H(\omega) := \mathrm{Hess}_x \mathcal{L}(\omega) + \mathcal{G}_x Y G(x)^{-1} \mathcal{G}_x^*$ であり, $G(x) \coloneqq \mathrm{diag}(g(x)) \in \mathbb{R}^{m\times m}, Y \coloneqq \mathrm{diag}(y) \in \mathbb{R}^{m\times m}, \mathcal{G}_x[v] \coloneqq \sum_{i\in\mathcal{I}} v_i \mathrm{grad} g_i(x) \in T_x\mathcal{M}$ である. 適当な条件の下で $\{\omega_k\}$ の集積点は条件 (2) を満足することを証明した [2, Theorem 4.2].

2.2 内部反復

内部反復では, 外部反復における ω_{k+1} を内点法と信頼領域法により求める. 現在点 $\omega = (x, y)$ における探索方向 $(d_x, d_y) \in T_x\mathcal{M} \times \mathbb{R}^m$ を次のように生成する. まずは, $Y := \mathrm{diag}(y)$ として (3) において二つの右辺項を 0 とおいた式と等価な方程式である $\Psi(\omega;\mu) \coloneqq \begin{bmatrix} \mathrm{grad}_x \mathcal{L}(\omega) \\ Yg(x) - \mu\mathbf{1} \end{bmatrix} = 0$ を考え, これに対するニュートン方程式は次のように表現できる.

$$H(\omega)[d_x] = -c_\mu(x) := -\mathrm{grad} f(x) + \mu \mathcal{G}_x \left[G(x)^{-1}\mathbf{1}\right], \tag{5}$$

$$d_y = -y + \mu G(x)^{-1}\mathbf{1} - YG(x)^{-1}\mathcal{G}_x^*[d_x]. \tag{6}$$

式 (5) が $m_{\omega,\mu}(d) \coloneqq \frac{1}{2}\langle H(\omega)[d], d\rangle_x + \langle c_\mu(x), d\rangle_x$ の最小化問題の停留条件であることに注意すると, これに $\Delta > 0$ を信頼半径として, 信頼領域制約をつけた次の問題を考えることができる.

$$\underset{d_x \in T_x\mathcal{M}}{\mathrm{minimize}}\ m_{\omega,\mu_x}(d_x) \text{ subject to } \|d_x\|_x \leq \Delta. \tag{7}$$

$m_{\omega,\mu_x}(d_x)$ が (1) に対する対数障壁関数 $P_\mu(x) \coloneqq f(x) - \mu \sum_{i\in\mathcal{I}} \log g_i(x)$ の 2 次近似関数であり, この $P_\mu(x)$ の局所最適解は条件 (3) と (4) を厳密に満たすことから, 信頼領域法のアイディアに基づき, 信頼半径 Δ を適宜調節しながら (7) の大域的最小解を求め, $P_\mu(x)$ の値をうまく減少させるような d_x を求める. y 空間における探索方向 d_y については (6) と d_x から求める. 探索方向 (d_x, d_y) を決定後, $\mathcal{M} \times \mathbb{R}^m$ のレトラクションを用いて, 内点制約 $g(x) > 0, y > 0$ を違反しない範囲内で次の点を決定する. いくつかの条件の下で本手法が実際に (3),(4) を満たす点を有限回の反復で求めることができることを証明した [2, Theorem 4.18].

参考文献

[1] M. Obara, T. Okuno, and A. Takeda, *Local near-quadratic convergence of Riemannian interior point methods.* arXiv:2505.19724, 2025.

[2] M. Obara, T. Okuno, and A. Takeda, *A primal-dual interior point turst region method for second-order stationary points of Riemanian inequality-constrained optimization problems.* arXiv:2501.15419v2, 2025.