構造的制約を持つ最適化問題に対する majorization-minimization 手法の構築

東京工業大学 *山川智也 YAMAKAWA Tomoya
05001310 東京工業大学 劉田香 LIU Tianxiang
01704970 東京工業大学 山下真 YAMASHITA Makoto

1. はじめに

本研究では、以下のような制約が特別な構造を持つ正則化付き最小化問題について考える。

$$\begin{aligned} \min_{x\in\mathbb{R}^n} &\quad F(x) := f(x) + \lambda P(x) \\ \text{s.t.} &\quad \mathcal{A}(x) \in D \end{aligned} \tag{1}$$

ただし、以下のような仮定を満たしているものとする。

- $\lambda > 0$ とし、$f : \mathbb{R}^n \to \mathbb{R}$ は微分可能な関数でその勾配 ∇f はリプシッツ定数 $L > 0$ を持つリプシッツ連続である。
- $\mathcal{A}$ は full row rank の線形写像であり、D は閉集合かつその射影 Proj_D は効率的に解くことができる一方で、$\mathrm{Proj}_{\mathcal{A}^{-1}(D)}$ は閉形式を持たない。
- $\inf F > -\infty$
- P は正則化項であり、例えば $\|x\|_1$ や $\|x\|_2^2$ などが挙げられる。

特に $\mathrm{Proj}_{\mathcal{A}^{-1}(D)}$ は閉形式を持たない仮定について、閉形式を持つ場合においても適用することはできるが、より簡便な既存手法があるため、このような仮定を導入する。これらの制約としては Fused lasso 信号モデルや dose-volume 制約、Hankel 構造制約問題などの様々な応用がある。一方で、これらの制約は基本的に非凸であるため、問題そのものは NP 困難である。さらに制約の特別な構造から、射影が陽に表せないという特徴を持つ。

本研究では、制約に関しては difference of convex (凸差分) 制約に表現した上で moving balls approximation (MBA) [1] を用い、目的関数に関しては majorization-minimization [3] 型の手法を取り入れたアルゴリズムを提案した。この手法を用いることによって、対象としている問題における制約の構造を活かしながら効率よく解くことを目指す。そして提案手法の部分問題が効率的に求解可能であること、アルゴリズムで生成された列の集積点が generalized Mangasarian-Fromovitz constraint qualifications (GMFCQ) という制約想定の下で Clarke critical point に収束することを示した。

2. 提案手法の概要

(1) の制約条件は非凸かつ射影が陽に表せないため解を求めるのが難しいことから、本研究においては制約条件を difference of convex (凸差分) 制約で表現する。具体的には以下の問題を解くことを考える。

$$\begin{aligned} \min_{x\in\mathbb{R}^n} &\quad F(x) := f(x) + \lambda P(x) \\ \text{s.t.} &\quad \frac{1}{2}\|\mathcal{A}(x)\|^2 - \underbrace{\max_{u\in D}\left\{\langle u, \mathcal{A}(x)\rangle - \frac{1}{2}\|u\|^2\right\}}_{Q(x)} \le \sigma. \end{aligned} \tag{2}$$

さらに、[2, section 3.1] より、Q の劣勾配は以下の計算式により効率的に計算できることが知られている。

$$\mathcal{A}^*(\mathrm{Proj}_D(\mathcal{A}(x))) \in \partial Q(x)$$

本研究においては、(2) を解くために、以下の部分問題 (3) を解くことを提案する。

$$\begin{aligned} \min_{x\in\mathbb{R}^n} &\quad \langle \nabla f(x^k),\, x - x^k\rangle + \frac{1}{2\gamma}\|x - x^k\|^2 + \lambda P(x) \\ \text{s.t.} &\quad \frac{1}{2}\|\mathcal{A}(x^k)\|^2 + \langle \mathcal{A}^*(\mathcal{A}(x^k)) - \zeta^k,\, x - x^k\rangle + \\ &\quad \frac{K}{2}\|x - x^k\|^2 - Q(x^k) - \sigma \le 0, \end{aligned} \tag{3}$$

ただし、γ はステップサイズであり、$\zeta^k = \mathcal{A}^*(\mathrm{Proj}_D(\mathcal{A}(x^k))) \in \partial Q(x^k), K = \|\mathcal{A}^*\mathcal{A}\|$ で

ある。部分問題に基づき、以下のアルゴリズム (NMMA) を提案する。

Algorithm 1 A non-monotone majorization-minimization algorithm (NMMA) for (2)

Step 0. Take any x^0 with $\mathcal{A}(x^0) \in D$. Let $\gamma_{\max} \geq \gamma_{\min} > 0$. Let $c > 0$, $\tau \in (0, 1)$ and integer $M \geq 0$. Set $k = 0$.

Step 1. Pick any $\gamma \in [\gamma_{\min}, \gamma_{\max}]$.

1a) Let u^γ solve (3).

1b) If

$$F(u^\gamma) \leq \max_{[k-M]_+ \leq i \leq k} F(x^i) - \frac{c}{2}\|u - x^k\|^2,$$

go to **Step 2**; otherwise, update $\gamma \leftarrow \tau\gamma$ and go to **Step 1a)**.

Step 2. Set $x^{k+1} = u^\gamma$, update $k \leftarrow k+1$ and go to **Step 1**.

3. 収束性

まず、NMMA の収束に必要な制約想定である GMFCQ の定義について述べ、対象となる問題の制約式が GMFCQ を満たすことを示す。

定義 3.1 制約式 $q(x^*) \leq 0$ において、x^* が *generalized Mangasarian-Fromovitz constraint qualifications (GMFCQ)* を満たすとは、以下のことを指す。

$$q(x^*) = 0 \text{ ならば } 0 \notin \partial^\circ q(x^*)$$

命題 3.1 (2) の制約式 $q(x^*) = \frac{1}{2}\|\mathcal{A}(x)\|^2 - Q(x) - \sigma \leq 0$ は *GMFCQ* を満たす。

本研究では、NMMA で生成された部分列の集積点が Clarke critical point に収束することを示した。

定理 3.1 (NMMA の部分列の収束) *GMFCQ* が成り立ち、$\{(x^k, \zeta^k)\}$ は *NMMA* で生成された列とする。また、μ_k は (3) のラグランジュ乗数とし、$\{x^k\}$ は有界であるとする。そのとき、以下が成り立つ。

1. $\lim_{k\to\infty} \left\|x^{k+1} - x^k\right\| = 0$
2. $\{\mu_k\}, \{\zeta_k\}$ はともに有界である。
3. $\bar{x}$ を $\{x^k\}$ の集積点とする。そのとき、$\bar{x}$ は (2) の *Clarke critical point* である。

4. 数値実験

本研究においては fused lasso 問題を用いて NMMA と Ye らの split Bregman method for the fused lasso (SBFlasso) [4] と比較を行った。

表 1: 次元の数と平滑化項を変更した場合の様子

	SBFLasso		NMMA	
n	RecErr	Time(sec)	RecErr	Time(sec)
500	0.053	0.482	0.032	0.357
1000	0.048	2.165	0.027	0.845
2000	0.041	15.428	0.018	3.043
3000	0.042	43.381	0.016	5.842
4000	0.040	94.297	0.015	10.705
5000	0.040	175.631	0.014	14.736

表 4 にある通り、解の精度と速度の両方の面で同等以上の結果を生み出すことができた。特に次元数が多い状態においては提案手法はより短い時間で解くことができた。これは MBA と majorization-minimization を用いたことによって、解きにくい主問題をノルムの計算と soft threshold の計算で済む部分問題を反復計算に分解したため、それぞれの部分問題の計算時間が短くなり、大規模な問題に対しても高速に解くことができたと考えられる。

参考文献

[1] A. Auslender, R. Shefi, and M. Teboulle. A moving balls approximation method for a class of smooth constrained minimization problems. *SIAM Journal on Optimization*, 20(6):3232–3259, 2010.

[2] T. Liu, T. K. Pong, and A. Takeda. A successive difference-of-convex approximation method for a class of nonconvex nonsmooth optimization problems. *Mathematical Programming*, 176(1):339–367, 2019.

[3] Y. Sun, P. Babu, and D. P. Palomar. Majorization-minimization algorithms in signal processing, communications, and machine learning. *IEEE Transactions on Signal Processing*, 65(3):794–816, 2016.

[4] G.-B. Ye and X. Xie. Split bregman method for large scale fused lasso. *Computational Statistics & Data Analysis*, 55(4):1552–1569, 2011.