FY2023/2023年度

Continuous Optimization Team

Akiko Takeda

連続最適化チーム　武田 朗子

RIKEN Center for Advanced Intelligence Project

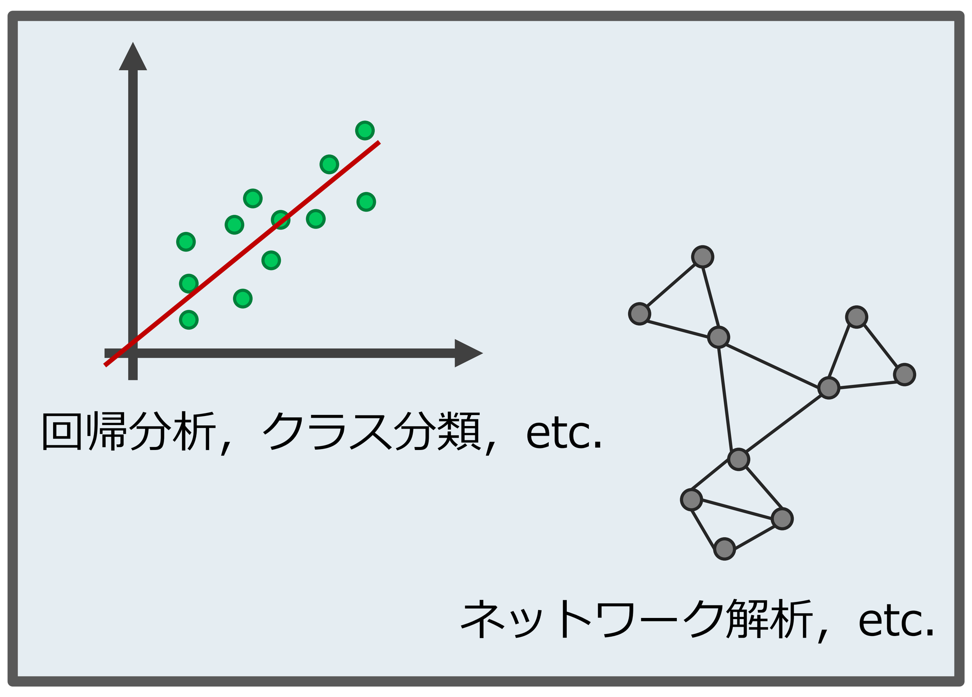
数理最適化（主に連続最適化）に関する研究

機械学習やデータマイニングから生じる最適化問題をうまく解き，実社会の問題解決に役立てたい

チームの強み：

- 大規模問題に対するランダム射影を用いた効率的解法
- 非凸最適化問題に対する理論保証つき解法
- 制約付き最適化問題に対する効率的解法

機械学習やデータマイニング



モデル化 →

数理最適化問題

$$\min.\ f(x)\quad \text{s.t.}\ g_1(x) \le 0,\ g_2(x) \le 0,\ \dots$$

↓

解法研究 (例):

- ランダム射影法
- 非平滑・非凸最適化法
- 制約付き最適化法

チームの得意とするところ

ランダム部分空間正則化 ニュートン 法 [Fuji, Poirion, Takeda, arXiv:2209.04170, 投稿中]

問題設定: 大規模変数をもつ非凸最適化問題

研究成果：次元縮小した逆行列計算を用いた正則化ニュートン法を構築．ランダム部分空間法では，最悪反復回数（反復計算量）と局所１次収束性の両立を達成は初

提案手法のアイディア

横長($d \ll n$)ランダム行列 P_k を各反復で取り直して作った次元縮小最適化問題を２次近似して，探索方向を計算

$$\min_{x \in R^n} f(x) \Rightarrow \min_{u \in R^d} f(x_k + P_k^\top u)$$

ランダム部分空間正則化ニュートン法の更新式：

$$x_{k+1} = x_k + t_k d_k$$

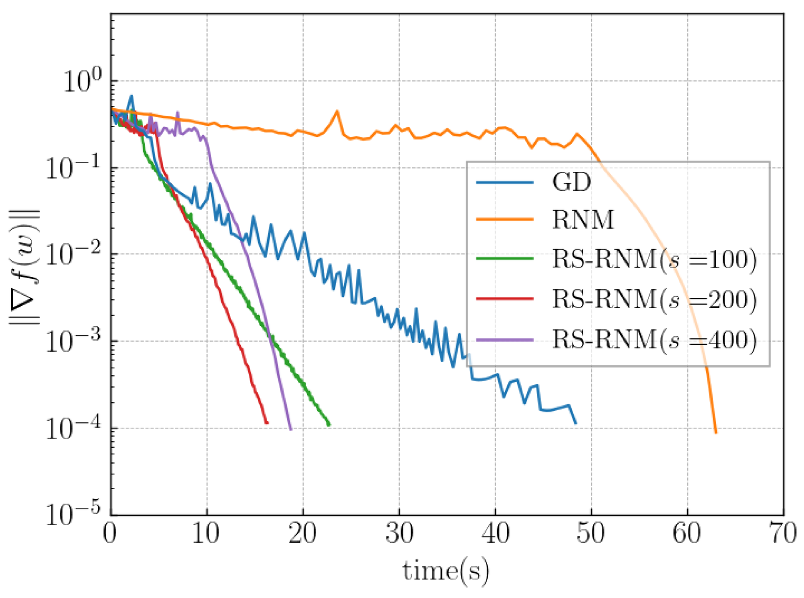
$$d_k = -P_k^\top \left(P_k \nabla^2 f(x_k) P_k^\top + \eta_k I_d\right)^{-1} P_k \nabla f(x_k)$$

正則化(凸化)+2次近似して得られた解を元の空間に戻す

次元縮小された行列の逆行列計算により，探索方向 d_k の計算可能，かつ，次元縮小前と同じオーダーの反復計算量を達成

局所的収束性の解析

- 一般的な場合: １ 次収束
- 局所的最適解におけるヘッセ行列がランク落ちしている場合: 超 1 次収束



非凸最適化問題に対するランダム部分空間法との比較

Random Subspace Algo.		Iter Comp.	Local
Direct Search	[Roberts and Royer, 2022]	$O(\epsilon^{-2})$	-
Adaptive Cubic Reg.	[Shao, 2022]	$O(\epsilon^{-3/2})$	-
Gauss Newton	[Cartis et al., 2022]	$O(\epsilon^{-2})$	-
Regularized Newton	[Fuji et al., 2022]	$O(\epsilon^{-2})$	1+

非線形 SDP の 2 次停留点へ収束する内点法の開発 [Arahata, Okuno, Takeda, Comp. Opt. & Appl., 2023]

非線形半正定値最適化問題（NSDP）

$$\min_{x \in \mathbb{R}^n} f(x)\ \text{s.t.}\ X(x) \in \mathbb{S}^m_+$$

$\mathbb{S}^m_+$ ： m 次実半正定値対称行列の集合

f, X ：ともに 2 回連続的微分可能

NSDPの応用例

- 固有値最適化
- 半正定値行列分解
- ランク最適化, etc

X の形を対角行列に限定すれば，NSDP は標準的な非線形最適化

- **研究背景:** NSDPに対する既存アルゴリズムの理論保証は全て局所最適性の 1 次の必要条件であるカルーシュ・キューン・タッカー条件を満たす点 (KKT 点) への収束のみだった
- **研究成果 1:** KKT 条件より強い 2 次の最適性条件を満たす点 (SOSP) への大域的収束性をもつアルゴリズムを初めて開発
- **研究成果 2:** SOSP への収束に要する最悪反復回数（反復計算量）を評価

提案手法のアイディア

$\mu \to 0$ としながら主双対バリア関数の SOSP を勾配法 + 負曲率方向で求める

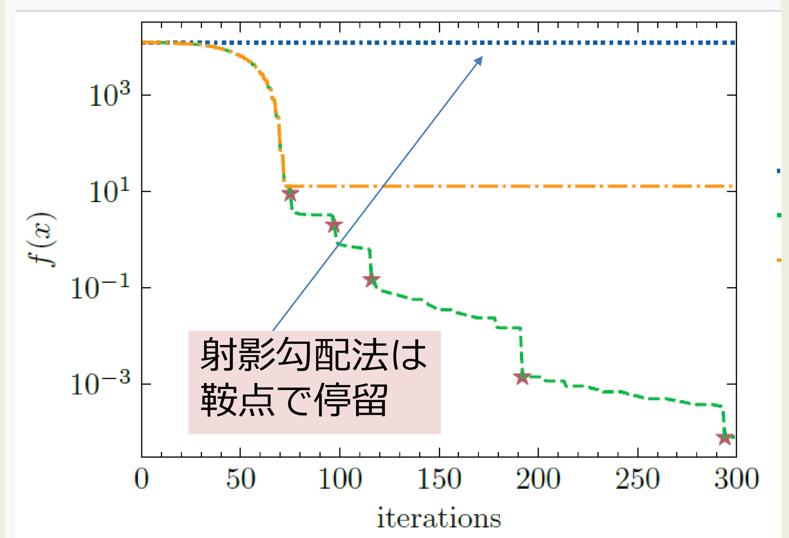
$$F_{\rm PB}(x;\mu) + F_{\rm BC}(x,\Lambda;\mu)$$

$$F_{\rm PB}(x;\mu) := f(x) - \mu \log\det X(x)$$

$$F_{\rm BC}(x,\Lambda;\mu) = {\rm Tr}(X(x)\Lambda) - \mu \log\det X(x)\Lambda$$

$\mu \to 0$ によって中心パスは NSDP の SOSP に収束 (本研究の結果の一つ)

数値実験 (射影勾配法との比較)



- 射影勾配法
- 提案法 (負曲率方向なし)
- 提案法
- 負曲率方向使用

制約付き最小二乗問題に対する Levenberg–Marquardt 法 [Marumo, Okuno, Takeda, Comp. Opt. & Appl., 2022]

問題設定: 凸制約付き最小二乗問題

$$\min_{x \in \mathbb{R}^d} \left\{ F(x) := \sum_{i=1}^n (f_i(x))^2 \right\} \quad \text{s.t.}\ x \in C$$

- $f_i : \mathbb{R}^d \to \mathbb{R}$: 滑らかな関数
- $C \subseteq \mathbb{R}^d$: 　凸集合，射影が計算可能

LM 法 [Levenberg, 1944], [Marquardt, 1963]

$$x_{k+1} = \operatorname*{argmin}_{x \in C} \left\{ \sum_{i=1}^n \Big(\underbrace{f_i(x_k) + \nabla f_i(x_k)^\top (x - x_k)}_{f_i(x)\text{ の一次近似}} \Big)^2 + \mu_k \|x - x_k\|^2 \right\}$$

元々は無制約問題（$C = \mathbb{R}^d$）のための手法

本研究の貢献: 以下の 2 つを示した

1. $\mu_k \propto \sqrt{F(x_k)}$ と定め，子問題を「ある近似条件」を満たすように近似的に解くと，
   a) ε-停留点が $O(\varepsilon^{-2})$ 反復で求まる
   b) 局所 2 次収束を達成
2. a) を達成するだけなら，子問題をより不正確に解いてもよい

既存研究との比較

- 制約付き問題に対する LM 法に反復計算量保証をつけたのは初 [Ueda & Yamashita, 2010], [Zhao & Fan, 2016] 等は無制約問題に対し $O(\varepsilon^{-2})$
- 反復計算量 $O(\varepsilon^{-2})$ と局所 2 次収束の両立は初 [Bergou+, 2020]: 無制約問題に対し $O(\varepsilon^{-2}\log\varepsilon^{-1})$ と局所 2 次

数値実験

十数種類の問題例で既存手法と性能比較．提案 LM 法は安定して高速であることを確認

黒: 提案LM, 紫: 近接勾配法, 他: 既存LM