Continuous Optimization Team

Akiko Takeda

連続最適化チーム　武田 朗子

RIKEN
AIP Center for Advanced Intelligence Project

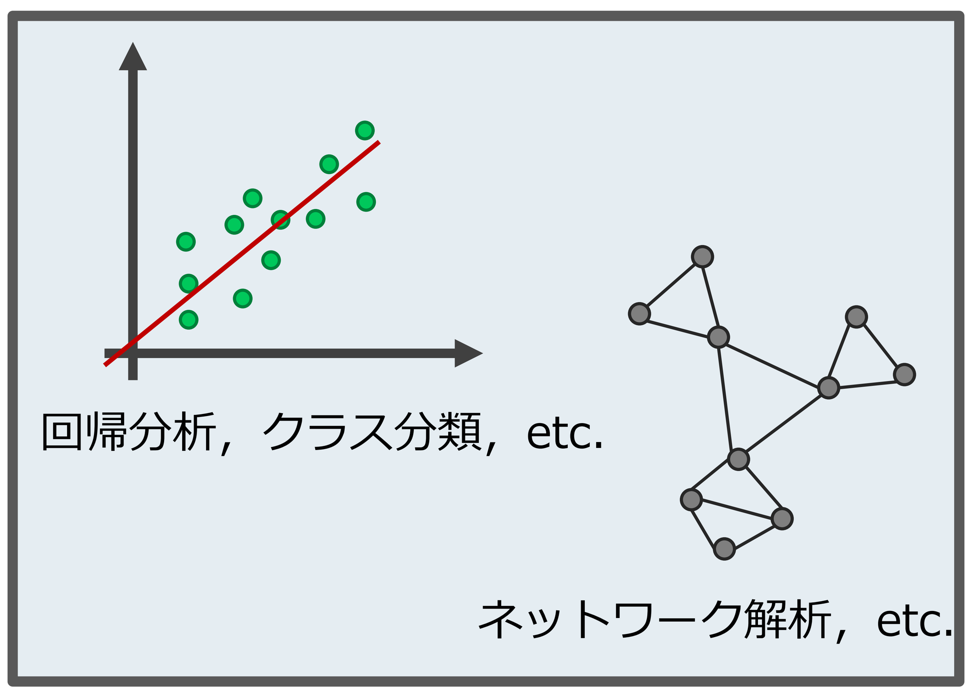
数理最適化（主に連続最適化）に関する研究

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

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

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

多数制約付き線形計画問題のためのランダム射影 [Poirion, Lourenço, Takeda, Lin. Alg. App., 2023]

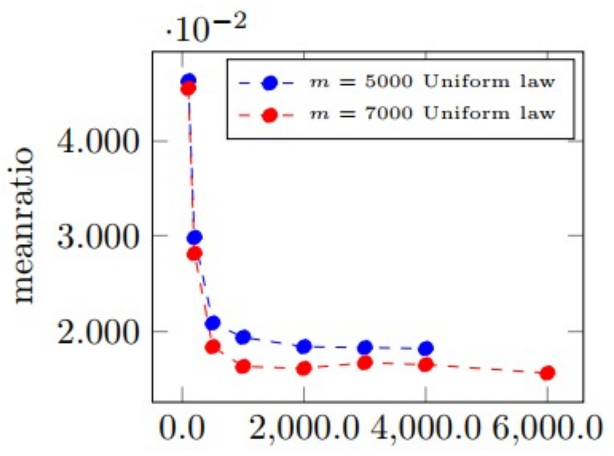
制約式の本数が多い問題 $\mathcal{P}$ の代わりにランダム行列で制約式が圧縮された緩和問題 $\mathcal{P}_S$ を解く。

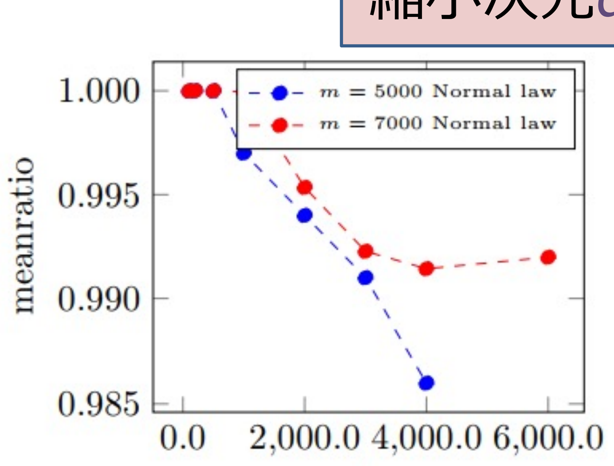
- $S_{ij}\sim\mathcal{N}(0,1/d)^2$
- $S\in\mathcal{R}^{d\times m}, d\ll m$

$$\mathcal{P}\left\{\begin{array}{l}\min_x\ c^\top x\\ Ax\ge b\\ x\in\mathbb{R}^n\end{array}\right.\quad(1)\qquad \mathcal{P}_S\left\{\begin{array}{l}\min_x\ c^\top x\\ SAx\ge Sb\\ x\in\mathbb{R}^n\end{array}\right.\quad(2)$$

$$v(\mathcal{P})(1-Error(d))\le v(\mathcal{P}_S)\le v(\mathcal{P})$$

最適値との差$Error(d)$と縮小次元dの関係を導出





ランダム射影を利用した非凸２次最適化法 [Fuji, Poirion, Takeda, SIAM J Optim, 2022]

問題設定：非凸２次最適化問題

$$\mathcal{P}\equiv\min_{x\in\mathbb{R}^n}\{x^\top Qx+c^\top x\mid Ax\le b\},$$

$$\mathcal{P}'\equiv\min_{u\in\mathbb{R}^d}\{u^\top\boxed{\mathcal{F}^+(PQP^\top)}u+(Pc)^\top u\mid AP^\top u\le b\}$$

$\mathcal{F}^+$：半正定値錐への射影

ランダム射影P+凸化で変数の数が小（n≪d），かつ凸な問題に帰着

効率的に解ける

不定値カーネルSVM

- データはmnistの「０」「１」画像
- 1,000訓練データ
- シンプソンカーネル

約半分に次元縮小してもよい性能

表1　$\mathcal{P}$ achieved 66.70% training-accuracy and 53.00% test-accuracy

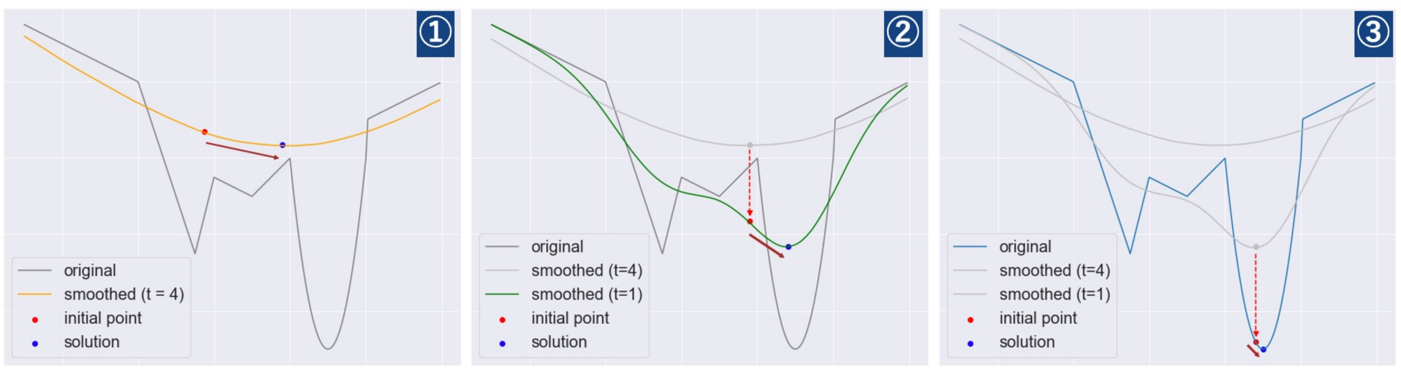
d	Training Accuracy (%)	Test Accuracy (%)
300	55.09 ± 14.40	53.43 ± 12.80
400	64.88 ± 15.64	62.80 ± 14.14
500	89.00 ± 12.61	83.30 ± 11.56
600	95.00 ± 0.85	90.23 ± 1.54
700	92.56 ± 1.60	86.53 ± 2.95
800	84.26 ± 3.09	76.60 ± 4.27
900	72.44 ± 3.20	60.83 ± 4.19
1000	66.96 ± 0.63	53.60 ± 0.78

非凸最適化問題のよりよい解を目指した解法 [Iwakiri, Wang, Ito, Takeda, NeurIPS, 2022]

◻ ガウシアンホモトピー (GH) 法

非凸最適化において，初期値の選択に依存せず，
よい解を見つけることを目指した手法

①: 目的関数を平滑化して凸関数に変形し，最小化
②, ③: 平滑化パラメータ t を小さくしながら，解を更新



◻ 既存の GH 法の問題点: 二重ループ構造

ループ1: 平滑化パラメータ t を更新
ループ2: t を固定したもとで，決定変数 x を更新

- ループ2で収束を待つ必要があり，GD と比べて非常に低速
- また，一般の非凸関数に対する収束レートの理論解析もなし

◻ 本研究の貢献

- x に加えて t を決定変数と見なし，これらを同時に更新する一重ループの手法を提案

Sample u_k from $\mathcal{N}(0,\mathrm{I}_d)$

$$x_{k+1}=x_k-\beta\bar{G}_{x,u},\ \bar{G}_{x,u}=\left\{\begin{array}{l}\tilde{g}_x(x_k,t_k;u_k)\ \text{(determ.)}\\ \tilde{G}_x(x_k,t_k;\xi_k,u_k),\ \xi_k\sim P\ \text{(stoc.)}\end{array}\right.$$

Sample v_k from $\mathcal{N}(0,\mathrm{I}_d)$

$$t_{k+1}=\left\{\begin{array}{l}\gamma t_k\ (\mathrm{SLGH_r})\\ \max\{\min\{t_k-\eta\bar{G}_{t,v},\gamma t_k\},\epsilon'\}\ (\mathrm{SLGH_d})\end{array}\right.,\ \bar{G}_{t,v}=\left\{\begin{array}{l}\tilde{g}_t(x_k,t_k;v_k)\ \text{(determ.)}\\ \tilde{G}_t(x_k,t_k;\xi_k,v_k),\ \xi_k\sim P\ \text{(stoc.)}\end{array}\right.$$

- 理論解析を行い，停留点への収束を保証した上，収束レートが GD と同じオーダーであることを証明
- ブラックボックスな敵対的攻撃にて，GD を精度で上回り，かつ，既存の GH 法よりも高速なことを確認

提案手法の攻撃成功率が最も高い

Methods	Succ. rate	Avg. iters to 1st succ.	Avg. L_2 (succ.)	Avg. total loss
ZOSGD	88%	**835** ± 1238	0.076 ± 0.085	27.70 ± 74.80
ZOAdaMM	85%	3335 ± 2634	**0.050** ± 0.055	20.24 ± 62.48
ZOGradOpt	65%	6789 ± 1901	0.249 ± 0.159	41.45 ± 76.04
ZOSLGH$_r$ ($\gamma=0.999$)	**93%**	4979 ± 756	0.246 ± 0.178	**14.26** ± 54.61
ZOSLGH$_d$ ($\gamma=0.999$)	**92%**	4436 ± 805	0.150 ± 0.084	**16.49** ± 58.69

既存の GH 法は提案手法より低速

リーマン多様体上の制約付き最適化問題に対する逐次２次最適化 [Obara, Okuno, Takeda, SIAM J Optim, 2022]

考える問題：リーマン多様体上の制約付き最適化問題（RNLO）

$$\begin{array}{ll}\underset{x\in\mathcal{M}}{\text{minimize}} & f(x)\\ \text{subject to} & g_i(x)\le 0,\quad i\in\mathcal{I}:=\{1,\ldots,m\},\\ & h_j(x)=0,\quad j\in\mathcal{E}:=\{1,\ldots,n\},\end{array}$$

$\mathcal{M}$：リーマン多様体

ユークリッド空間を一般化した位相空間

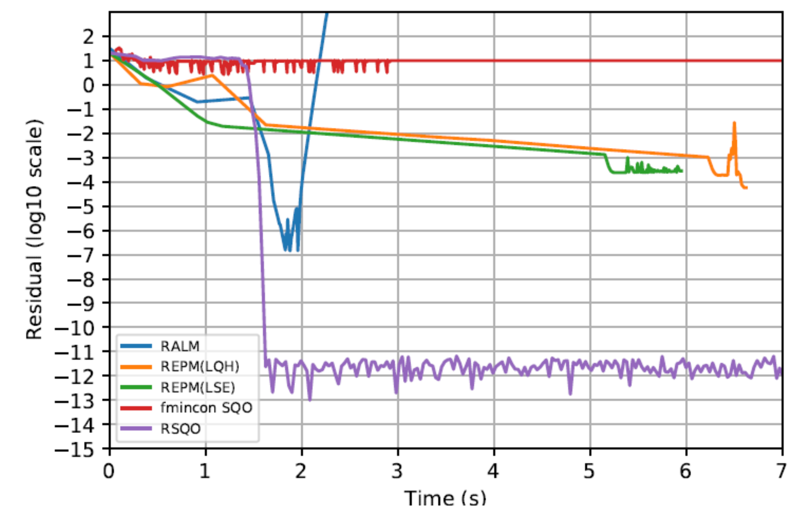
f,g_i,h_j ： $\mathcal{M}$ 上で定義された連続的微分可能な実数値関数

RNLOの背景

1. 無制約RNLOの研究の成熟を受けて，この数年で理論面や応用面の研究は活発化
2. 制御における応用例：**安定なシステム同定**においてRNLOへのモデリングが有効であることが示された

本研究の貢献

1. 新手法の提案：元問題の近似問題である**接空間上で定義された制約付き２次最適化問題**を繰り返し解いて点列を生成
2. **大域的収束性と局所２次収束性**を証明（２種の収束性保証のある手法はRNLOに対して は初めて）



数値実験

低ランク非負行列補完などを対象に，時間（横軸），最適性残差（縦軸対数目盛）で提案手法（紫）と他の数種の手法を比較．**精度**に関して大きな優位性を確認

値が小さいほど解に近い