連続最適化チーム　武田 朗子
Continuous Optimization Team
Akiko Takeda

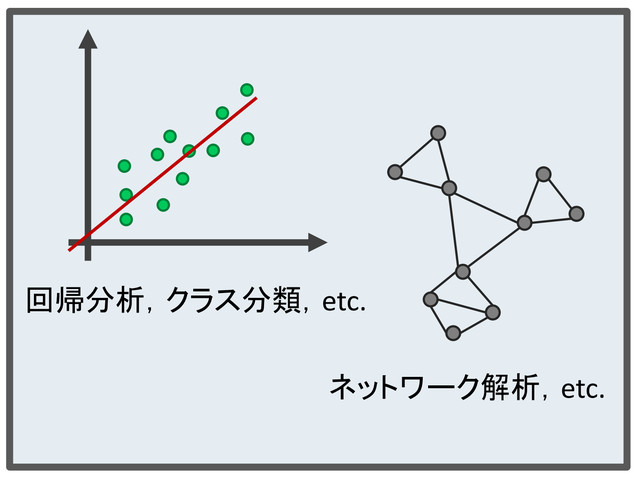
数理最適化（主に連続最適化）に関する研究

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：

- 複雑な制約をもつ問題に対する効率的解法
- 非平滑・非凸最適化問題に対する効率的解法
- 理論保証付き DC (Difference-of-Convex) アルゴリズム

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）：
- 非平滑・非凸最適化法
- DCアルゴリズム
- ロバスト最適化法

チームの得意とするところ

複数のランク制約をもつ問題に対する DC 近似アルゴリズム
[T. Liu, I. Markovsky, T. K. Pong, and A. Takeda, submitted, 2019]

複数のランク制約をもつ問題：

$$\min_{y=(y_1\cdots y_N)} f(y) \quad s.t.\ \ y \in \Omega_1 \cap \Omega_2.$$

$$\Omega_1 := \{y : rank(\mathcal{H}_{r_i+1}(y_i)) \le r_i \ \ \forall\, i\},$$

$$\Omega_2 := \{y : rank([\mathcal{H}_{r+1}(y_1)\cdots\mathcal{H}_{r+1}(y_N)]) = r\}.$$

- $y_i \in \mathbb{R}^n$
- $\mathcal{H}_s(\cdot)$ Hankel 行列
- r_i と r : 正整数
- f : level-bound かつ 勾配は Lipschitz 連続

提案解法 SDCAM:

$$\min_{y\in\mathbb{R}^{Nn}} F(y) := f(y) + \delta_\Omega(y) + \sum_{i=1}^{k} \delta_{C_i}(\mathcal{A}_i(y)).$$

ペア (Ω, C_i) は 3 通りの設定が可能

SDCAM は以下の問題を解く: $\min_{y\in\mathbb{R}^{Nn}} F_\lambda(y) := f(y) + \delta_\Omega(y) + \sum_{i=1}^{k} \frac{1}{2\lambda}\mathrm{dist}^2(\mathcal{A}_i(y), C_i).$

部分問題を解くには Ω への射影が必要だが，その計算は困難→他の方法は？

↓

$\hat{y}$ の Ω に対する擬似射影（$u \in \Omega$ に関して）

$$\mathcal{P}^s_\Omega(\hat{y}; u) := \{y \in \Omega : \hat{y} - y \in N_\Omega(y) \text{ and } \|y - \hat{y}\| \le \|u - \hat{y}\|\}.$$

← 既存のソルバー SLRA を用いることで得られる理論的な結果も存在

解の実行可能性は？

後処理：SDCAM の点から交互擬射影

$$z^{t+1} \in \mathcal{P}^s_{\Omega_1}(x^t; z^t) \text{ and } x^{t+1} \in \mathcal{P}^s_{\Omega_2}(z^{t+1}; x^t) \quad t = 0, 1, \cdots$$

理論的成果：

SDCAM の点列は以下を満たす：

$$\mathrm{dist}(\mathcal{A}_i(y^k), C_i) \le \sqrt{2\lambda_{k-1} f(y^{\mathrm{feas}})}$$

後処理の点列は，以下の条件を満たす $\bar{y} \in \Omega_1 \cap \Omega_2$ が存在すれば，局所 R-線形収束

$$rank([\mathcal{H}_{r+1}(\bar{y}_1)\cdots\mathcal{H}_{r+1}(\bar{y}_N)]) = r, \quad N_{\Omega_1}(\bar{y}) \cap N_{\Omega_2}(\bar{y}) = \{0\}.$$

Figure: Termination function value comparison.

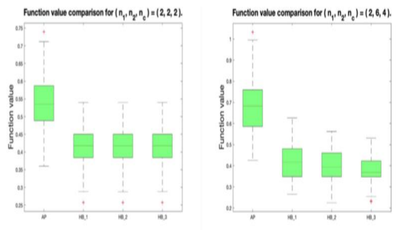
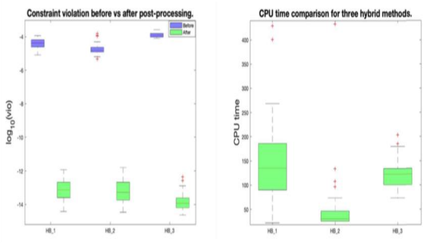


Figure: Constraint violation and CPU time comparison.



非平滑・非凸最適化に対する確率的手法
[M. Metel, A. Takeda, ICML, 2019]
[M. Metel, A. Takeda, arXiv:1905.10188, 2019]

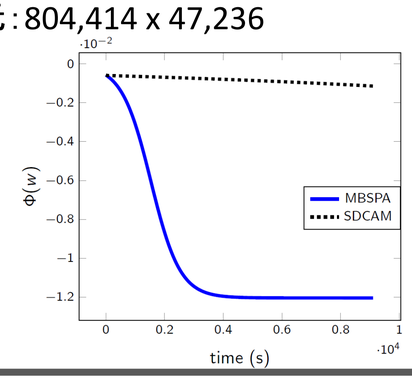
問題設定：

$$\min_{w\in\mathbb{R}^d} \Phi(w) := f(w) + g(w) + h(w)$$

where $f(w) := \mathbb{E}_\xi[F(w, \xi)]$

- $f(w)$: 損失関数 (勾配は Lipschitz 連続)
- $g(w)$: スパース正則化項 (Lipschitz 連続)
- $h(w)$: 凸制約 (閉真凸関数)

非負スパース PCA: RCV1 dataset
次元：804,414 x 47,236



劣微分の意味での収束：

$$\mathcal{G}^\Phi_\gamma(w) := \left\{\frac{1}{\gamma}\left(w - \mathrm{prox}_{\gamma h}(w - \gamma s)\right) : s \in \nabla f(w) + \partial g(w)\right\}$$

収束基準：期待値評価による ε-accurate な解

$$\mathbb{E}\left[\mathrm{dist}(0, \mathcal{G}^\Phi_\gamma(w))\right] \le \epsilon$$

$\mathrm{dist}(0, \mathcal{G}^\Phi_\gamma(w)) = 0 \to 0 \in \nabla f(w) + \partial g(w) + \partial h(w)$ が成り立つ

ミニバッチ確率的近接法 (Mini-batch stochastic proximal algorithm; MBSPA):

$f(w) + g(w)$ の平滑近似に対する確率的勾配

$$\nabla A^k_{\lambda,M}(w^k, \xi^k) = \frac{1}{M}\sum_{j=1}^{M} \nabla F(w^k, \xi^k_j) + \frac{1}{\lambda}(w^k - \zeta^\lambda(w^k))$$

where $\zeta^\lambda(w^k) \in \mathrm{prox}_{\lambda g}(w^k)$

アルゴリズムでの主な計算：$w^{k+1} = \mathrm{prox}_{\gamma h}(w^k - \gamma \nabla A^k_{\lambda M}(w^k, \xi^k))$

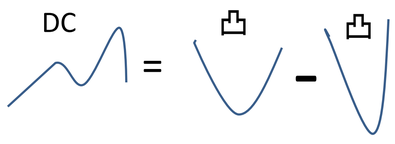
計算複雑度：

勾配の計算回数：$O(\epsilon^{-5})$　近接作用素の計算回数：$O(\epsilon^{-3})$

混合整数DC最適化に対する理論保証付きDCアルゴルズム [T. Okuno and Y. Ikebe, Optimization, 2019]

混合DC最適化問題

Difference-of -Convex functions



$$\begin{aligned} \min\ & g(x_1, x_2) - h(x_1, x_2) \\ \text{s.t.}\ & x \in S,\ x_1 \in \mathbb{Z}^m,\ x_2 \in \mathbb{R}^n \end{aligned}$$

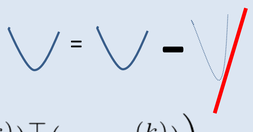
- $g, h : \mathbb{R}^{m+n} \to \mathbb{R}$: 微分可能な閉真凸関数
- $S \subseteq \mathbb{R}^{m+n}$: 空ではない閉凸集合
- $\mathbb{Z}$: 整数集合，$\mathbb{R}$: 実数集合

- 離散変数と連続変数を併せもつ最適化問題 (混合整数最適化問題) の一つのクラス
- 目的関数が非凸であるため求解は非常に難しい

提案手法

以下を収束するまで繰り返す

以下の子問題の最適解を $x^{(k+1)}$ とする



$$\begin{aligned} \min\ & g(x_1, x_2) - \left(h(x_1, x_2^{(k)}) + (y^{(k)})^\top (x - x^{(k)})\right) \\ \text{s.t.}\ & x \in S,\ x_1 \in \mathbb{Z}^m,\ x_2 \in \mathbb{R}^n \\ & (\text{但し，劣勾配} y^{(k)} \in \partial h(x^{(k)})) \end{aligned}$$

$h(x)$ を線形近似して凸化

子問題は混合整数凸最適化問題なので既存のソルバーで解くことができる

- 離散 DC 最適化における Maehara et al. (2017) の拡張
- 混合整数最適化で理論的保証がついた手法は初
- 既存ソルバーで予め出力した解を改善，でなければ最適性の保証ができるので実用面においても優れている

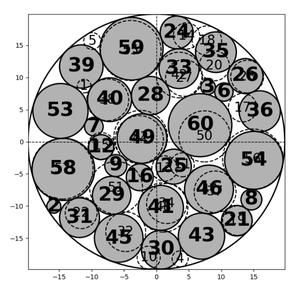
主結果

1. 考えている問題と或等価な連続DC最適化問題の DC 停留点と呼ばれる最適性条件を満たす点への大域的収束性を証明
2. 平滑化法と呼ばれる非平滑最適化に対する手法と組み合わせて，非平滑な関数をもつ混合整数 DC に対しても同様の理論結果を得た

応用例）円を円の中に配置する問題（2次元）
[Y. Ikebe, S. Masuda, T. Okuno, submitted]

入力：n 個の小円と外円
出力：外円の中で総面積が最大かつ，互いに重ならないような小円の配置

既存手法よりも
より高速に，より良い解



提案手法に基づく出力結果例