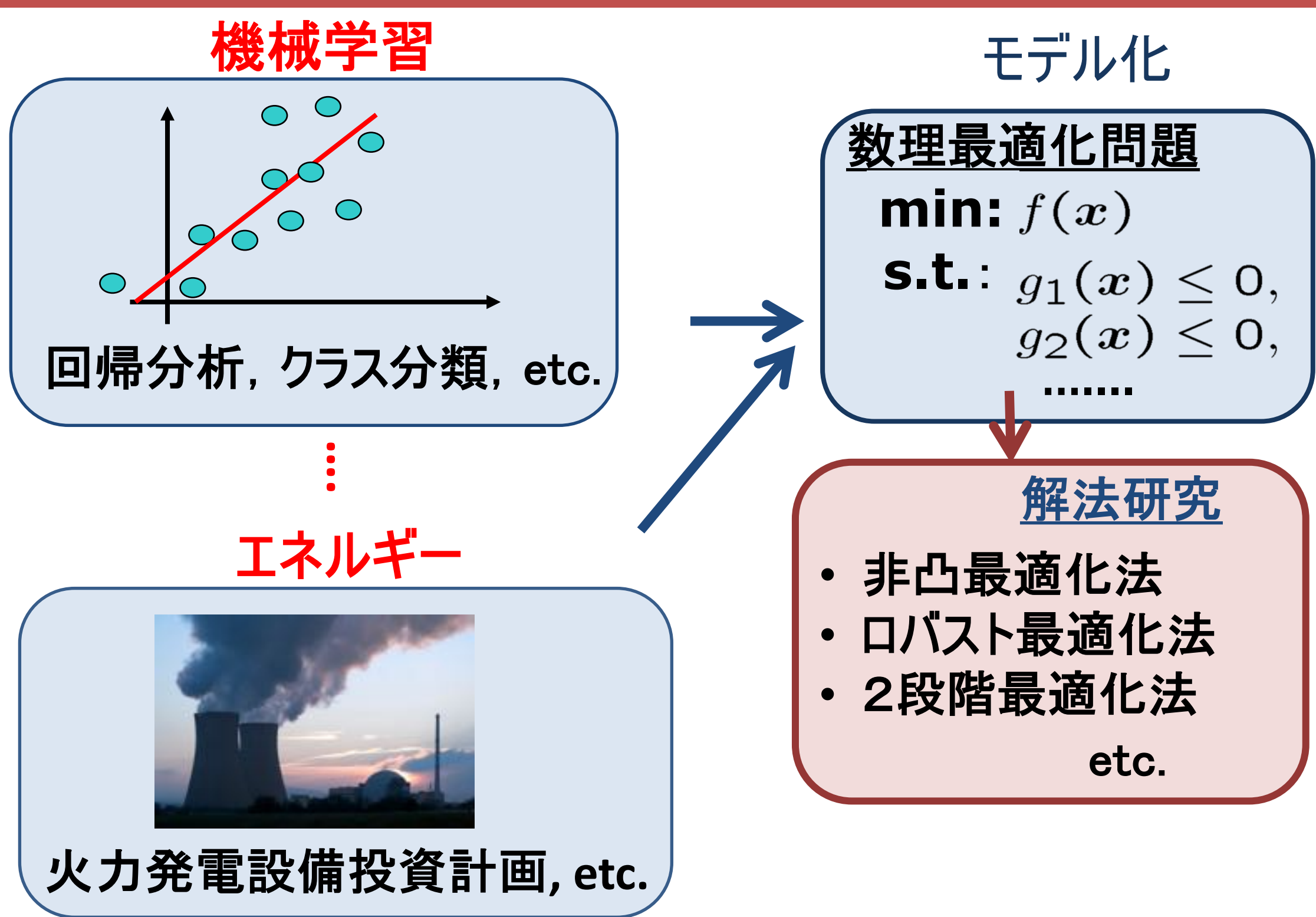


武田グループの主な研究成果 (2018年度)

数理最適化法(主に連続最適化)に関する研究
最近は、特に、非凸最適化問題に対する効率的解法

機械学習, データマイニング, 制御, エネルギー分野など
から生じる最適化問題をうまく解いて,
実社会の問題解決に役立てたい。



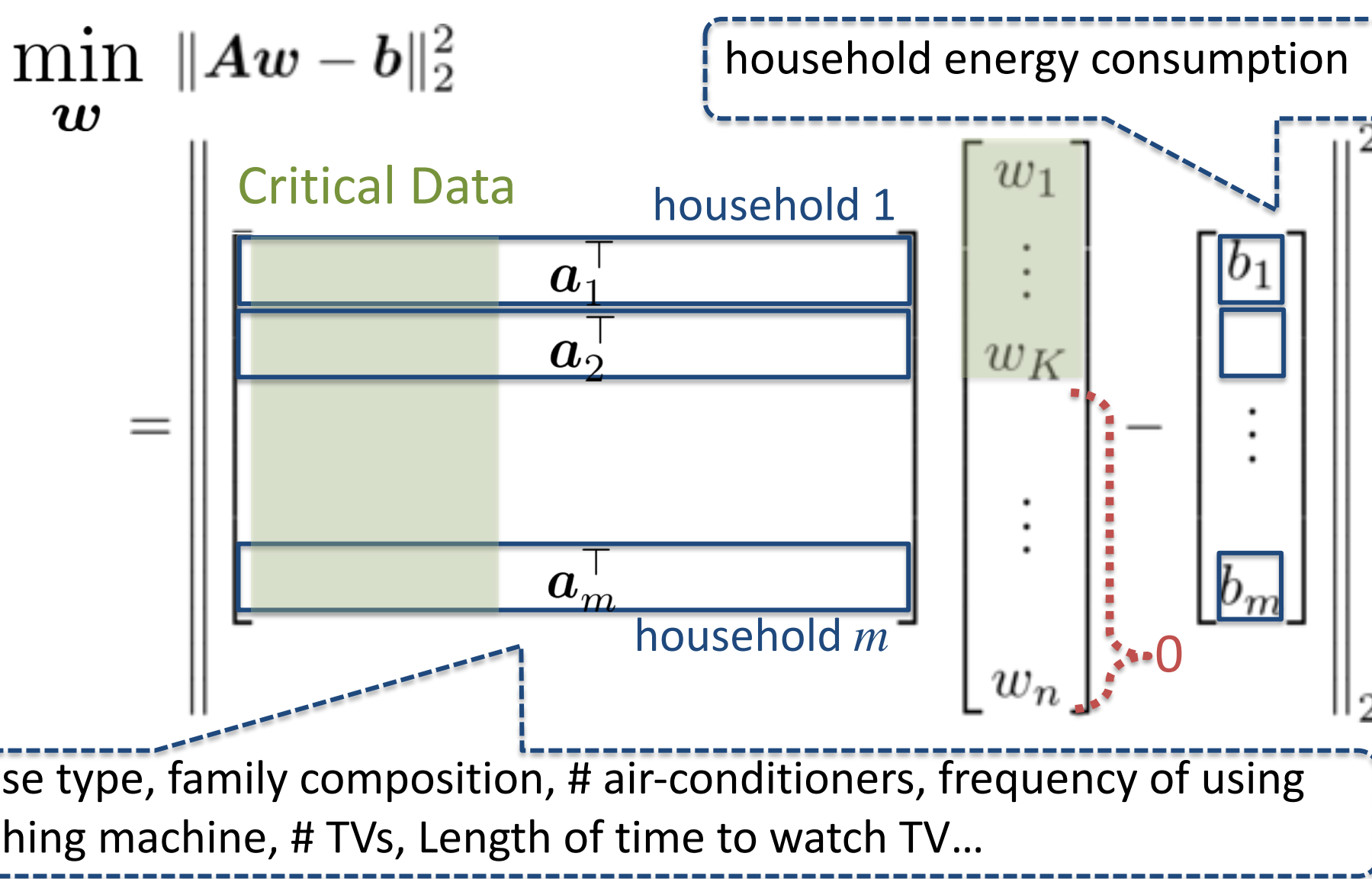
研究1: 効率的な非凸最適化アルゴリズム

Goto, Takeda, Tono [Math. Program '18]
Tono, Takeda, Gotoh [投稿中]

スパース最適化問題:

$$\min_w f(w) \\ \text{s.t. } \|w\|_0 \leq K$$

無駄な情報を落として疎な解
(変数選択)



- 今までは、非凸最適化問題を解かないですむような定式化が多かった
- 非凸最適化問題に対して高速な解法を考えた

- * 基数制約 $\|w\|_0 \leq K$ をうまく表現し、扱いやすい非凸最適化問題に再定式化
- * 遅いと言われていたDCアルゴリズムの問題点を、DC分解の工夫により解消

$$\min_w f(w) + \rho(\|w\|_1 - \|w\|_K)$$

$$\min_w \frac{L}{2} \|w\|_2^2 - \left(\frac{L}{2} \|w\|_2^2 - f(w) \right) + \rho(\|w\|_1 - \|w\|_K)$$

$L : (A^T A)$ の最大固有値

Pham Dinh & Le Thi ('14)

n	m	method	iter.	time	card.	residual
100	1000	MIP-DCA	22.0	0.719 (0.045)	10	8.304e-01 (2.807e-02)
		標準-DCA	21.0	0.509 (0.025)	10	6.827e-01 (1.802e-02)
		解析的 DCA	26.6	0.024 (0.002)	10	6.986e-01 (2.263e-02)
500	1000	MIP-DCA	24.1	31.310 (3.109)	50	7.950e-01 (1.031e-01)
		標準-DCA	24.0	23.604 (1.939)	50	5.282e-01 (1.358e-02)
		解析的 DCA	32.9	0.139 (0.024)	50	5.374e-01 (1.333e-02)
500	5000	MIP-DCA	24.5	32.125 (1.511)	50	9.137e-01 (1.881e-02)
		標準-DCA	24.0	21.410 (0.755)	50	6.794e-01 (8.984e-03)
		解析的 DCA	28.4	0.118 (0.016)	50	6.874e-01 (1.169e-02)
1000	5000	MIP-DCA	24.6	297.556 (19.745)	100	8.615e-01 (4.740e-02)
		標準-DCA	24.0	233.123 (12.390)	100	6.332e-01 (7.270e-03)
		解析的 DCA	30.3	0.502 (0.026)	100	6.386e-01 (7.842e-03)

扱いやすい問題への
再定式化の効果

DC分解の工夫の効果

提案手法が
非常に速い

研究2: 非平滑・非凸最適化問題に対する効率的解法

Liu, Pong, Takeda [Math. Program., '18]
Liu, Pong, Takeda [Comp. Optim. Appl., '19]

以下の非平滑・非凸最適化問題に着目:

$$\min_{x \in \mathbb{R}^n} F(x) = f(x) + P_0(x) + \sum_{i=1}^m P_i(A_i x)$$

- $F: L$ -smooth 関数, A_i : 線形写像
- P_i : 近接写像を容易に計算できる関数

$m > 0$ かつ一般の線形写像に対して、既存手法は収束の保証をもたない

Question:

収束の保証をもつアルゴリズムで、 P_i の近接写像と ∇f の計算のみ
にもとづくものをどのように設計するか?

重要な性質: Moreau包絡関数 $e_\lambda P_i(x)$ は difference-of-convex (DC) 関数

$$e_\lambda P_i(x) = \frac{1}{2\lambda} \|x\|^2 - \underbrace{\sup_{y \in \text{dom } P_i} \left\{ \frac{1}{\lambda} \langle x, y \rangle - \frac{1}{2\lambda} \|y\|^2 - P_i(y) \right\}}_{D_{\lambda, P_i}(x)}$$

解法 SDCAM: $\lambda \downarrow 0$ に対して、以下の問題を $\text{NPG}_{\text{major}}$ (近接写像 P_i の計算に
基づく解法) により逐次的に解く

$$F_\lambda(x) = f(x) + P_0(x) + \sum_{i=1}^m e_\lambda P_i(A_i x) = \underbrace{f(x) + \sum_{i=1}^m \frac{1}{2\lambda} \|A_i x\|^2}_{h(x)} + \underbrace{P_0(x)}_{P(x)} - \underbrace{\sum_{i=1}^m D_{\lambda, P_i}(A_i x)}_{g(x)}$$

理論保証:

- SDCAM が生成する列は有界
- ある制約想定のもとで、集積点が停留点

$$0 \in \nabla f(x^*) + \partial P_0(x^*) + \sum_{i=1}^m A_i^* \partial P_i(A_i x^*).$$

応用:

非凸 fused 正則化問題:

$$\min_{x \in \mathbb{R}^n} \frac{1}{2} \|x - b\|^2 + c_1 \|x\|_1 + c_2 \sum_{i=1}^{n-1} |x_{i+1} - x_i|^{\frac{1}{2}}$$

ただし、 b は区分定数スパース信号の(雑音を含む)観測値

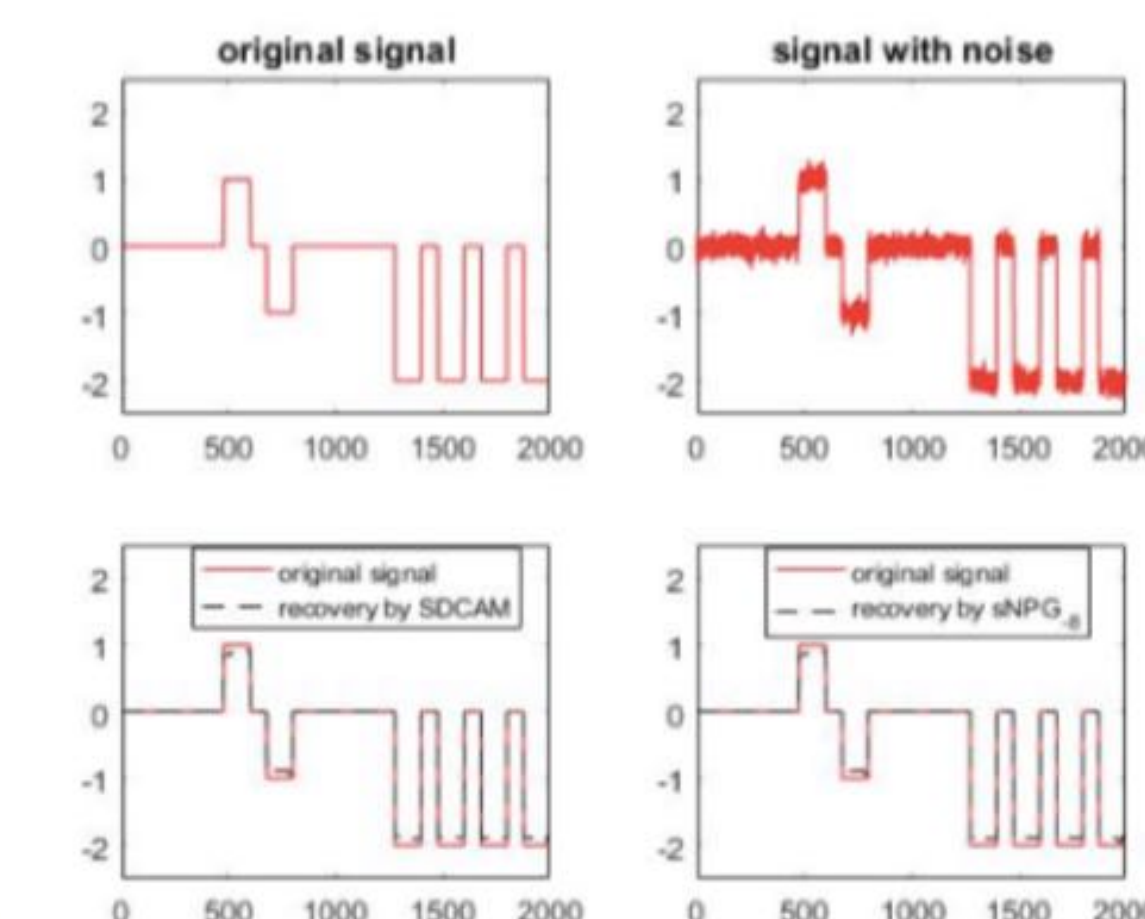


Figure 1: Recovery comparison for noisy signal.

研究3: 微分不可能な下位レベル最適化問題をもつ2段階最適化問題に対するアルゴリズム開発と ハイパーパラメータ最適化への応用

Okuno, Takeda, Kawana [投稿中]

2段階最適化問題 (bilevel optimization problem)

$$\min_{w, \lambda} f(w) \\ \text{s.t. } w \in \underset{\hat{w}}{\text{argmin}} R(\hat{w}, \bar{\lambda}) + \lambda_1 \sum_{i=1}^n |\hat{w}_i|^p \\ \lambda = (\lambda_1, \bar{\lambda})^T \geq 0$$

- 制約条件の中に別の最適化問題(下位レベル最適化問題)がある
- 下位レベル最適化問題が微分不可能で非常に取り扱いにくい。
- 既存研究は少なく、理論的な既存結果は皆無

提案法(平滑化法)

下位レベルの問題を下で置き換えた2段階
最適化問題を $\mu \rightarrow 0+$ としながら繰り返し解く。

$$w \in \underset{\hat{w}}{\text{argmin}} R(\hat{w}, \bar{\lambda}) + \lambda_1 \sum_{i=1}^n (\hat{w}_i^2 + \mu^2)^{\frac{p}{2}}$$

$$w := (w_1, \dots, w_n)^T \\ \bar{\lambda} := (\lambda_2, \dots, \lambda_m)^T$$

$$0 < p \leq 1$$

応用例
ハイパーパラメータ選択

スパース最適化(研究1参照):

$$\min_{\hat{w}} \|A\hat{w} - b\|^2 + \lambda_1 \sum_{i=1}^n |\hat{w}_i|^p \\ R(\hat{w}, \bar{\lambda}) \quad \text{ハイパーパラメータ}$$

- ハイパーパラメータの値によって解 $\hat{w}$ も異なる
- テスト誤差 $f(w)$ を小さくするハイパーパラメータが望ましい

2段階最適化への定式化

$$\sum_{i=1}^n |\hat{w}_i|^p \implies \sum_{i=1}^n (\hat{w}_i^2 + \mu^2)^{\frac{p}{2}} \text{ に置き換えて}$$

得た最適化問題は微分可能なので、
標準的な手法で解ける。

主結果

微分不可能な関数を含んだ
2段階最適化問題の理論的解析は本研究が初

- 上の最適化問題の最適解が満たすべき理論的性質 (SB-KKT条件) を発見
- 提案手法によって生成した点列が適当な仮定の下で SB-KKT条件を満たす点に収束することを数学的に証明
- 実データを用いた数値実験で、バイズ最適化に比べて提案手法は精度も高く、高速であることを確認