

# 部分的フィードバックに基づくオンライン凸最適化 伊藤伸志 (NEC)

## —限られた情報で最適な意思決定—

### 本研究課題の目的

通常の最適化問題 (オフライン最適化) :  
事前に目的関数  $f$  が与えられた状況で解  $x$  を出力

$f$  を入力  $\rightarrow$   $x$  を出力 Goal: Minimize  $f(x)$

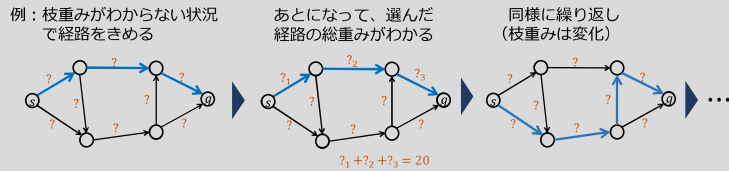


⊙ 実用上の制約: 評価指標を事前に知っていることが必要

評価指標が事前にわからない/時々刻々と変化する状況、例えば道の混雑度が常に変化する状況に対応するには?  
 $\rightarrow$  **オンライン最適化** の枠組みが有効

部分的 (バンディット) フィードバックオンライン最適化:  
解  $x_t$  を出力した後で目的関数値  $f_t(x_t)$  だけわかる

$x_1$  を出力  $\rightarrow$   $f_1$  を一部観測  $\rightarrow$   $x_2$  を出力  $\rightarrow$   $f_2$  を一部観測  $\rightarrow$  ...



⊙ 「試してみても初めて結果がわかる」状況にも対応

未来ビジョン: 不確定・流動的で情報が少ない環境における意思決定の最適化・効率化を実現したい  
本研究課題ではオンライン最適化によるアプローチを考察

オンライン最適化のゴール:  
リグレット  $R_T$  を小さくするアルゴリズムの開発

$$R_T := \sum_{t=1}^T f_t(x_t) - \min_{x^* \in A} \sum_{t=1}^T f_t(x^*)$$

$T$ : ラウンド数, 意思決定の回数

$A$ : 実行可能領域 (s-g 経路全体の集合)

$f_t$ : 目的関数 (経路長,  $f_t(x) := w_t^T x$ )

$x_t$ : アルゴリズムの出力 (s-g 経路)

$x^*$ : 最良の固定戦略 (全ラウンド平均の辺長さでの最短 s-g 経路)

$R_T$ : 最良戦略と比較したときの、後悔の大きさ

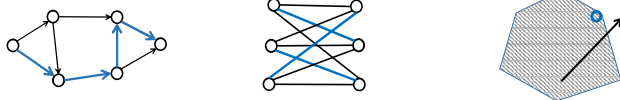
- $R_T = o(T)$  のとき, no-regret アルゴリズムという  $\rightarrow$  試行回数  $T$  を増やすと, 平均リグレット  $\frac{R_T}{T}$  は 0 に収束  $\rightarrow x_t$  の「良さ」が最良戦略  $x^*$  の「良さ」に漸近ある意味で、後悔が小さくなることを数学的に保障
- no-regret アルゴリズムの幅広い応用範囲:
  - オフライン最適化に対する 多項式時間近似スキーム
  - 確率的最適化アルゴリズム
  - ある種のゲーム均衡を求めるアルゴリズムなどの存在が従う

### 成果 1 : バンディット線形最適化

NeurIPS2019 に 2 件の論文が採択 [IHSKFKK2019a, IHSKFKK2019b]

目的関数  $f_t$  が線形関数のバンディット最適化問題を考察

応用例: 最短経路問題 最大重みマッチング 線形計画問題



|      | (i) 情報量: バンディット?     | (ii) 計算量: オラクル複雑性 | (iii) 最適性: リグレット上界  |
|------|----------------------|-------------------|---|
| 既存研究 | [KV2005]             | 完全情報              | $\checkmark O(T)$   |
|      | [CL2012]             | $\checkmark$      | No poly. bound  |
|      | [DH2006]             | $\checkmark$      | $\checkmark O(T^{2/3})$                                     |
|      | [IHSKFKK2019a (本研究)] | $\checkmark$      | $\checkmark \tilde{O}(T)$ (確率的設定では $\tilde{O}(\log T)$ に改善) |

※オラクル複雑性: 対応するオフライン最適化問題を解く回数

- バンディットフィードバックの問題で、計算効率と最適性を両立できるかは未解決の問題だった
- 本研究で、少情報性・計算効率性・最適性を両立する初のアルゴリズムを提案
- リグレット上界はこれ以上はほぼ改善不可能 [IHSKFKK2019b]

### 参考文献

[CL2012]: N. Cesa-Bianchi and G. Lugosi. Combinatorial bandits. *Journal of Computer and System Sciences*, 78(5):1404–1422, 2012.  
 [DH2006]: V. Dani and T. P. Hayes. Robbing the bandit: Less regret in online geometric optimization against an adaptive adversary. In *Proceedings of the seventeenth annual ACM-SIAM symposium on Discrete algorithms*, pages 937–945, 2006.  
 [IF2016]: S. Ito and R. Fujimaki. Large-scale price optimization via network flow. In *Advances in Neural Information Processing Systems*, pages 3855–3863, 2016.  
 [I2019]: S. Ito. Submodular function minimization with noisy evaluation oracle. In *Advances in Neural Information Processing Systems*, 2019.  
 [I2020]: An optimal algorithm for bandit convex optimization with strongly-convex and smooth loss. In *International Conference on Artificial Intelligence and Statistics*, page to appear, 2020.  
 [IHSKFKK2019a]: S. Ito, D. Hatano, H. Sumita, K. Takemura, T. Fukunaga, N. Kakimura, and K. Kawarabayashi. Improved regret bounds for bandit combinatorial optimization. In *Advances in Neural Information Processing Systems*, 2019.  
 [IHSKFKK2019a]: S. Ito, D. Hatano, H. Sumita, K. Takemura, T. Fukunaga, N. Kakimura, and K. Kawarabayashi. Oracle-efficient algorithms for online linear optimization with bandit feedback. In *Advances in Neural Information Processing Systems*, 2019.  
 [KV2005]: A. Kraljic and S. Vempala. Efficient algorithms for online decision problems. *Journal of Computer and System Sciences*, 71(3):291–307, 2005.

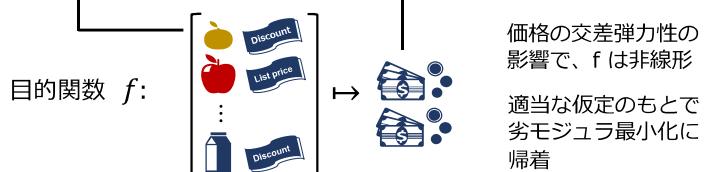
### 成果 2 : バンディット非線形最適化

NeurIPS2019 と AISTATS2020 にそれぞれ 1 件論文が採択 [i2019], [i2020]

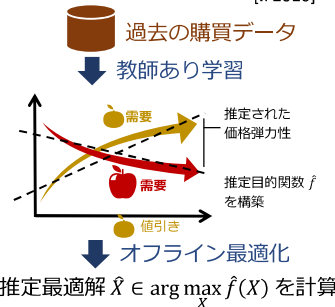
バンディット型の劣モジュラ最小化・凸最適化を考察

劣モジュラ最小化の応用例: 小売り価格最適化

- 複数の商品の価格をうまく選んで、総利益を最大化したい

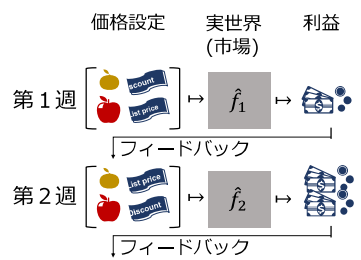


[既存] モデルベース最適化 [IF2016]



- ⊙ 十分な過去データが必要
- ⊙ 定常的な価格弾力性を仮定

[本研究] バンディット型最適化



- ⊙ 過去データなしでも適用可能
- ⊙ 需要特性が変動する状況にも対応

バンディット劣モジュラ最小化 [i2019]・バンディット凸最適化 [i2020] のそれぞれに対し、ほぼ最適リグレットを達成する効率的なアルゴリズムを提案