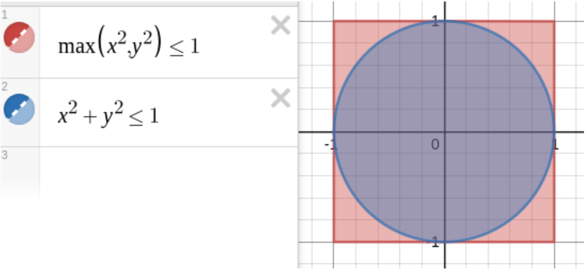
Exploiting Structure of Uncertainty for Efficient Matroid Semi-Bandits

Pierre Perrault (INRIA Lille — CMLA, ENS PS)
Vianney Perchet (CMLA, ENS PS — Criteo AI Lab)
Michal Valko (INRIA Lille)

Semi-bandits confidence regions



	
$\max_i \{ (\delta_i - \overline{\mu}_{i,t-1})^2 N_{i,t-1} \} \leq \log(t)$	$\sum_i (\delta_i - \overline{\mu}_{i,t-1})^2 N_{i,t-1} \leq \log(t)$
Not very accurate	Accurate

Efficiency

Algorithms use the OFU principle:

$$A_t \in \underset{A\in\mathcal{A},\ \boldsymbol{\mu}\in\mathcal{C}_t}{\arg\max}\ \mathbf{e}_A^\mathsf{T}\boldsymbol{\mu} = \underset{A\in\mathcal{A}}{\arg\max}\ \underbrace{\mathbf{e}_A^\mathsf{T}\overline{\boldsymbol{\mu}}_{t-1}}_{L(A)} + \underbrace{\max_{\boldsymbol{\mu}\in\mathcal{C}_t-\overline{\boldsymbol{\mu}}_{t-1}} \mathbf{e}_A^\mathsf{T}\boldsymbol{\mu}}_{F(A)}.$$

Theorem (Perrault et al.)

 *: F linear,*  *: F submodular.*

	
$\max_i \left\{(\delta_i - \overline{\mu}_{i,t-1})^2 N_{i,t-1}\right\} \leq \log(t)$	$\sum_i (\delta_i - \overline{\mu}_{i,t-1})^2 N_{i,t-1} \leq \log(t)$
Not very accurate	Accurate
Efficient	**Inefficient**

NEW: Approximation for matroid

Assume non-negative rewards. $\mathcal{A}$ is the family of independent sets.

GREEDY:

$$\frac{L(S) + F(S)}{1 - 1/e} \geq L(O) + F(O), \quad \forall O \in \mathcal{A}.$$

Gives **linear regret.** We expect a constant close to 1 for F small.

NEW: Approximation for matroid

Assume non-negative rewards. $\mathcal{A}$ is the family of independent sets.

GREEDY:

$$\frac{L(S) + F(S)}{1 - 1/e} \geq L(O) + F(O), \;\; \forall O \in \mathcal{A}.$$

Gives **linear regret.** We expect a constant close to 1 for F small.

$$L(S_1) \geq L(O), \;\; \forall O \in \mathcal{A}.$$

$$\frac{F(S_2)}{1 - 1/e} \geq F(O), \;\; \forall O \in \mathcal{A}.$$

NEW: Approximation for matroid

Assume non-negative rewards. $\mathcal{A}$ is the family of independent sets.

GREEDY:

$$\frac{L(S) + F(S)}{1 - 1/e} \geq L(O) + F(O), \ \ \forall O \in \mathcal{A}.$$

Gives **linear regret.** We expect a constant close to 1 for F small.

$$L(S_1) \geq L(O), \ \ \forall O \in \mathcal{A}.$$

$$\frac{F(S_2)}{1 - 1/e} \geq F(O), \ \ \forall O \in \mathcal{A}.$$

Theorem (Perrault et al.)

GREEDY *for maximizing* $L + F$ *gives* S *such that*

$$L(S) + 2F(S) \geq L(O) + F(O), \ \ \forall O \in \mathcal{A}.$$

When reward can be negative...

LOCAL SEARCH based algorithm.

Theorem (Perrault et al.)

$$L(S) + 2(1+\varepsilon)F(S) \geq L(O) + F(O), \quad \forall O \in \mathcal{A}.$$

Time complexity per round: $\mathcal{O}\big(m^2 n \log(mt)/\varepsilon\big)$

- Start from the greedy solution $S_{\text{init}} \in \arg\max_A L(A)$.
- Then, repeatedly try three basic operations in order to improve the current solution.
- Improvements greater than $\frac{\varepsilon}{m}F(S)$.

Thank you! Poster: Pacific Ballroom #53

Extension to budgeted bandit, where we want to minimize

$$\left(\frac{L_1 - F_1}{L_2 + F_2}\right)^+.$$

Solution uses NEW concept: *Approximation Lagrangian*

$$\mathcal{L}_\kappa(\lambda, S) \triangleq L_1(S) - \kappa F_1(S) - \lambda(L_2(S) + \kappa F_2(S)),$$

Experiments