Stochastic Online Learning with Probabilistic Graph Feedback

Motivation

- Full information (complete graph)
- Bandit feedback (only self-loops)
- Graph feedback can cover full information and bandit settings
- Motivating examples:
  - Recommend island A would infer user’s preference of island B
  - Influence spread of seed A would infer influence ability of neighbor B
- Probabilistic graph

Lower Bounds

- Let \( i = 1 \) be the best arm
- An algorithm is consistent if \( R_i(T) = o(T^{-\delta}) \) for any \( \delta > 0 \)
- Let \( p_i^* \) be the probability that there is a directed path from \( i \) to \( j \) in a random realization of \( G \)

One-Step Triggering

\[
C(\mu) = \left\{ c \in [0, \infty)^v : \sum_{i \in V} p_i^c 1 \geq \frac{1}{KL(\mu, \mu_i)} \bigvee i \neq 1 \right\}
\]

Theorem. For any consistent algorithm, the regret satisfies

\[
\lim_{T \to \infty} \frac{R_i(T)}{\log T} \geq \inf_{\epsilon > 0} \left( \epsilon \right) \in \mathcal{C}(\mu, \Delta(\mu))
\]

Cascade Triggering

\[
C(\mu) = \left\{ c \in [0, \infty)^v : \sum_{i \in V} p_i^c 1 \geq \frac{1}{KL(\mu, \mu_i)} \bigvee i \neq 1 \right\}
\]

Theorem. For any consistent algorithm, the regret satisfies

\[
\lim_{T \to \infty} \frac{R_i(T)}{\log T} \geq \inf_{\epsilon > 0} \left( \epsilon \right) \in \mathcal{C}(\mu, \Delta(\mu))
\]

Results

- One-Step Triggering
  - \( \nu^* \): the set of exploration nodes that have the largest live probabilities among all incoming edges to some \( j \)
  - \( p_i^* \): the minimum exploration probability of \( i \)

Theorem. The regret satisfies for any \( \epsilon > 0 \),

\[
R(T) = O \left( \sum_{i=1}^v c_i(\epsilon) \Delta_i(\theta) + \log(T) \sum_{i=1}^v \Delta_i(\theta) \right)
\]

and

\[
\lim_{T \to \infty} \frac{R(T)}{\log T} \leq 4 \cdot \inf_{\epsilon > 0} \left( \epsilon \right) \in \mathcal{C}(\mu, \Delta(\theta))
\]

Cascade Triggering

- \( \nu^* = \{ p_i^* \geq 2 \max_{j \neq i} p_j^* \text{ for some } j \} \)
  - a relaxed version of \( \nu^* \)

Theorem. The regret satisfies for any \( \epsilon > 0 \),

\[
R(T) = O \left( \sum_{i=1}^v \Delta_i(\epsilon) \max_{c_i(\epsilon, \nu^*)} \log(T) \sum_{i=1}^v \Delta_i(\theta) \right)
\]

and

\[
\lim_{T \to \infty} \frac{R(T)}{\log T} \leq 4 \cdot \inf_{\epsilon > 0} \left( \epsilon \right) \in \mathcal{C}(\mu, \Delta(\theta))
\]

Experiments

- a. Cyclic feedback graph
  - \( \theta = (0.5 + \Delta, 0.5, ..., 0.5) \)

Contact

Shuai Li
weich@cs.cuhk.hk
Kwong-Sak Leung

References