Online Learning for Global Cost Functions

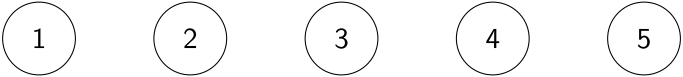
Even-Dar et al. *COLT* 2009

Jonay Trénous

March 23, 2016

Overview

Motivation: Load Balancing / Job Scheduling



- Distribute Jobs over the Machines
- Objective: Minimize maximum Load
- Cost not additive over timesteps

Learning Setup

- Finite Set of Alternatives

$$K = \{1, ..., k\}$$

- At each time step t, Algorithm A chooses a distribution

$$p_t \in \Delta(K)$$

- Incur Vector of Losses

$$l_t \in [0,1]^k$$

- Vector of average losses after T Timesteps:

$$L_T = (L_T(1), ..., L_T(k)) = \left(\frac{1}{T} \sum_{t=1}^{T} l_t(1), ..., \frac{1}{T} \sum_{t=1}^{T} l_t(k) \right)$$

- Vector of average losses incurred by Algorithm

$$L_T^A = (L_T^A(1), ..., L_T^A(k)) = \left(\frac{1}{T} \sum_{t=1}^{T} p_t(1) l_t(1), ..., \frac{1}{T} \sum_{t=1}^{T} p_t(k) l_t(k) \right)$$

Global Cost Functions

- Standard Setup: Cost per round Additive in $L_T^A(i)$

$$C(L_T^A) = \sum_{i=1}^{k} L_T^A(i) = \|L_T^A\|_1$$

- Generalization: $L_d, d \geq 1$ norm

$$C(L_T^A) = \|L_T^A\|_d$$

- L_∞ norm for Load Balancing example

Regret

- Regret measured against best static allocation

$$p^* = \text{argmin}_{p \in \Delta(K)} C(L_T^p)$$

-

$$R_T(A) = C(L_T^A) - C^*(L_T) = C(L_T^A) - C(L_T^{p^*})$$

- C convex, C^* concave
- Goal: Algorithm A that guarantees vanishing Regret

Vector Valued Games

- Strategy Sets

$$A \subseteq \mathbb{R}^a, B \subseteq \mathbb{R}^b$$

- (Bilinear) Payoff Function

$$m \colon A \times B \to \mathbb{R}^c$$

- Target Set $V \subseteq \mathbb{R}^c$
- Goal of player A: Achieve average payoff close to V

Approachability Theory

- Target Set V is *approachable* with convergence rate $r(t)$ if

$$\exists a_1, a_2, \ldots \in A \text{ s.t. } \forall b_1, b_2, \ldots \in B \colon \operatorname{dist}(\hat{m}_t, V) \leq r(t)$$

- (Blackwell, 1956). *If V is a closed and convex set and for every $b \in B$, there is an $a \in A$ such that $m(a, b) \in V$, then V is approachable with a convergence rate of*

$$r(t) = \frac{2D}{\sqrt{t}}$$

Where D is the diameter of the image of the Payoff function m.

Regret minimization as a Game

- Strategy Sets:

$$A = \Delta(K), B = [0,1]^k$$

- Payoff Function

$$m_t = m(p_t, l_t) = (p_t(1)l_t(1), ..., p_t(k)l_t(k), l_t(1), ..., l_t(k))$$

$\Rightarrow$ Average Payoff

$$\hat{m}_T = (L_T^A, L_T)$$

- Target Set

$$S = \{(x,y) \| C(x) \leq C^*(y), (x,y) \geq 0\}$$

Convexity of Target Set

Let $\beta \in [0,1], (x_1, y_1), (x_2, y_2) \in S$.

$$
\begin{aligned}
& C(\beta x_1 + (1-\beta)x_2) \\
& \leq \beta C(x_1) + (1-\beta) C(x_2) && \text{(Conv. of } C\text{)} \\
& \leq \beta C^*(y_1) + (1-\beta) C^*(y_2) && \text{(Def. } S\text{)} \\
& \leq C^*(\beta y_1 + (1-\beta) y_2) && \text{(Conc. of } C^*\text{)}
\end{aligned}
$$

Realizability

- Let $l \in [0,1]^k$
- Choose optimal static allocation $p^* = \text{argmin}_{p \in \Delta(K)} C(p(1)l(1), ..., p(k)l(k))$
- By definition of p^* and S

$$m(p^*, l) \in S$$

Approachability Result

- By Blackwells Theorem S approachable with convergence rate

$$r(t) = \frac{2D}{\sqrt{t}}$$

- $m(\Delta(K) \times [0,1]^k) \subseteq [0,1]^{2k} \Rightarrow D \leq \sqrt{2k}$
- Therefore S is approachable with convergence rate

$$r(t) = \sqrt{\frac{8k}{t}}$$

Regret Bounds

Assume C, C^* are Λ-Lipschitz. After T rounds of Alg. A, choose $(x, y) \in S$ such that

$$\|(L_T^A, L_T) - (x, y)\| \leq \sqrt{\frac{8k}{T}}$$

$$R_T(A) = C(L_T^A) - C^*(L_T)$$

$$= \underbrace{[C(x) - C^*(y)]}_{\leq 0} + [C(L_T^A) - C(x)] + [C^*(y) - C^*(L_T)] \qquad \text{(add zero)}$$

$$\leq \Lambda \left(\|L_T^A - x\| + \|L_T - y\| \right) \qquad \text{By Lipschitz}$$

$$\leq \Lambda \sqrt{2\|L_T^A - x\|^2 + 2\|L_T - y\|^2} \qquad (a^2 + b^2) \geq 2ab$$

$$= \Lambda \sqrt{2\|\hat{m}_t - (x, y)\|^2} \leq 4\Lambda \sqrt{\frac{k}{T}}$$

Conclusions

- Generalization of Loss Functions with Vanishing Regret
- Optimality Result: For any L_d norm, Regret of $\frac{1}{\sqrt{T}}$ is optimal
- Further results: Bound Logarithmic in number of Alternatives for L_∞ norm