Fast Algorithms for Online Stochastic Convex Programming

SODA 2015

Online Stochastic Convex Programming

- Concave function f over a bounded domain $\subseteq \mathbb{R}^d$
- Convex set $S \subseteq [0,1]^d$ (i.e., constraints)
- Each time t, receive a set $A_t \subseteq [0,1]^d$
- Pick a vector $\boldsymbol{v}_t^\dagger \in A_t$

- Goal: $maximize\ f(\boldsymbol{v}^\dagger_{avg})\ subject\ to\ \boldsymbol{v}^\dagger_{avg} \in S$

Comparison

Bandits with Knapsack

- The available set of choices across time as the arms is persistent.

Online Packing

- The set of options in one time step is unrelated to the other time steps.

Stochastic Input Model

Random Permutation Model

- T sets, $X_1, ..., X_T$ in advance but unknown to the algorithm
- Set comes in an uniformly random order

IID Model

- There is a distribution $\mathcal{D}$ over subsets of $[0, 1]^d$
- For each time t, A_t is independently sample from $\mathcal{D}$

The RP model is stronger than the IID model

Benchmarks

$$\begin{aligned} \text{avg-regret}_1(T) &= \text{OPT} - f(\boldsymbol{v}^{\dagger}_{\text{avg}}), \text{ and} \\ \text{avg-regret}_2(T) &= d(\boldsymbol{v}^{\dagger}_{\text{avg}}, S). \end{aligned}$$

- where $d(\boldsymbol{v}^{\dagger}_{\text{avg}}, S)$ represents the distance between vector v and S.

$$d(A, B) = \inf_{x \in A, y \in B} d(x, y).$$

Feasibility Problem

- A special case of online stochastic convex problem
- No objective function f
- The aim is to have $\boldsymbol{v}^{\dagger}_{\mathrm{avg}}$ be in the set S
- Performance: the distance from the set S--- $d(\boldsymbol{v}^{\dagger}_{\mathrm{avg}}, S)$

Fenchel Duality

- Fenchel conjugate: $h^*(\boldsymbol{\theta}) := \max_{\boldsymbol{y}\in[0,1]^d}\{\boldsymbol{y}\cdot\boldsymbol{\theta} - h(\boldsymbol{y})\}$
- Dual function: $\boldsymbol{\theta}\cdot\boldsymbol{z} - h^*(\boldsymbol{\theta})$
- Dual optimum: $\max_{||\boldsymbol{\theta}||_*\le L}\{\boldsymbol{\theta}\cdot\boldsymbol{z} - h^*(\boldsymbol{\theta})\}$

- In the feasibility problem, we have
- $d(\boldsymbol{x}, S) = \max_{||\boldsymbol{\theta}||_*\le 1}\{\boldsymbol{\theta}\cdot\boldsymbol{x} - h_S(\boldsymbol{\theta})\}$
- where $h(\boldsymbol{x}) = d(\boldsymbol{x}, S)$ and $h^*(\boldsymbol{\theta}) = h_S(\boldsymbol{\theta}) := \max_{\boldsymbol{y}\in S}\boldsymbol{\theta}\cdot\boldsymbol{y}$

Connection to Online Learning

- In primal problem, the set A_t arrives in an online manner
- It is hard to use the online learning algorithms for making decisions in A_t (might predict an infeasible $\boldsymbol{v}_t^\dagger$)
- Instead of selecting $\boldsymbol{v}_t^\dagger$ in A_t, making decisions in dual variables $\boldsymbol{\theta}_t$ according to the dual function $\max_{||\boldsymbol{\theta}||_* \leq 1}\{\boldsymbol{\theta} \cdot \boldsymbol{x} - h_S(\boldsymbol{\theta})\}$

Algorithm for the feasibility problem

Initialize $\boldsymbol{\theta}_1$.
for all $t = 1, ..., T$ **do**
  Set $\boldsymbol{v}_t^\dagger = \arg\min_{\boldsymbol{v} \in A_t} \boldsymbol{\theta}_t \cdot \boldsymbol{v}$
  Choose $\boldsymbol{\theta}_{t+1}$ by doing an OCO update with $g_t(\boldsymbol{\theta}) = \boldsymbol{\theta} \cdot \boldsymbol{v}_t^\dagger - h_S(\boldsymbol{\theta})$, and domain $W = \{||\boldsymbol{\theta}||_* \leq 1\}$.
end for

- Updating rules (Online Mirror Descent):

$$\boldsymbol{\theta}_{t+1,j} = \frac{w_{t,j}}{\sum_j w_{t,j}}, \text{ where } w_{t,j} = w_{t-1,j}(1+\epsilon)^{g_t(\mathbf{e}_j)/M}$$

OMD: http://web.stanford.edu/~takapoui/omd.pdf

Performance

$$\begin{aligned}\mathbb{E}[\mathit{avg\text{-}regret}_2(T)] &:= \mathbb{E}[d(\boldsymbol{v}^{\dagger}_{avg}, S)] \\ &\leq O\left(\frac{\mathcal{R}(T)}{T} + ||\mathbf{1}_d||\sqrt{\frac{s\log(d)}{T}}\right)\end{aligned}$$

$$\mathcal{R}(T) \leq O(L\sqrt{dT})$$

- where R(T) represents the regret for OCO with $g_t(\boldsymbol{\theta})$; and d and s represents the number of dimension and the coordinate-wise largest value of one vector in S

Online Stochastic Problem

- Reduce the feasibility problem to the online stochastic problem:

$$S' = \{\boldsymbol{v} : f(\boldsymbol{v}) \geq \mathrm{OPT}, \boldsymbol{v} \in S\}$$

- Require the estimation of OPT with $\tilde{O}(\frac{1}{\sqrt{t}})$ per step errors.

Involve a parameter Z

- Define OPT^{δ} as the optimal value of f with feasibility constraints relaxed to $d(\frac{1}{T}\sum_t \boldsymbol{v}_t, S) \leq \delta$
- Then, *given a* $Z \geq 0$ *such that that for all* $\delta \geq 0$ have

$$OPT^{\delta} \leq OPT + Z\delta$$

Online Stochastic Problem

- Combine the objective and the constraints together with a constant factor Z.

Initialize $\boldsymbol{\theta}_1, \boldsymbol{\phi}_1$.
for all $t = 1, ..., T$ **do**
Choose option
$\boldsymbol{v}_t^\dagger = \arg\max_{\boldsymbol{v} \in A_t} -\boldsymbol{\phi}_t \cdot \boldsymbol{v} - 2(Z+L)\boldsymbol{\theta}_t \cdot \boldsymbol{v}$.
Choose $\boldsymbol{\theta}_{t+1}$ by doing an OCO update for $g_t(\boldsymbol{\theta}) = \boldsymbol{\theta} \cdot \boldsymbol{v}_t^\dagger - h_S(\boldsymbol{\theta})$ over domain $W = \{\|\boldsymbol{\theta}\|_* \leq 1\}$.
Choose $\boldsymbol{\phi}_{t+1}$ by doing an OCO update for $\psi_t(\boldsymbol{\phi}) = \boldsymbol{\phi} \cdot \boldsymbol{v}_t^\dagger - (-f)^*(\boldsymbol{\phi})$ over domain $U = \{\|\boldsymbol{\phi}\|_* \leq L\}$.
end for

f is the objective, and L is the parameter of L-Lipschitz

Performance

- For online stochastic problem, the regrets are

$$\mathbb{E}[\textit{avg-regret}_1(T)] = (Z+L)\cdot O\left(\sqrt{\frac{C}{T}}\right)$$

$$\mathbb{E}[\textit{avg-regret}_2(T)] = O\left(\sqrt{\frac{C}{T}}\right)$$

Take-aways

- The online learning algorithms usually require a fixed set over all rounds for making decisions, but these feasible sets among different rounds are hard to be fixed.
- However, online learning algorithms can work well on the dual problem---i.e., the feasible sets of dual variables could be fixed.

- The analysis is still hard to follow.

Thanks!