Last week: video lecture on Multiplicative Weights Update (MWU) Algorithm

10 May 2019 09:40

Today: some applications of MWU

MWU:

n experts, T rounds

in each round, we have to pick an expert $i \in [n]$

loss when choosing expert i in round t is $f_i^t \in [-1, 1]$

goal: be competitive with best expert in hindsight

MWU Algorithm (parameter $\varepsilon \in (0, 1/2]$)

$w^1 = (1, \dots, 1)$

for $t = 1, \dots, T$:

$p^t := \frac{w^t}{\phi^t}$, $\phi^t = \sum_{i=1}^n w_i^t$

pick expert according to distr. p^t

$w_i^{t+1} = w_i^t \cdot (1 - \varepsilon f_i^t)$

Analysis of MWU:

$$\underset{\text{exp. loss of MWU alg.}}{\text{loss}} = \sum_{t=1}^T \underset{\langle p^t, f^t \rangle}{(p^t)^T f^t} = \sum_t \sum_i p_i^t f_i^t$$

$$\text{loss}_i = \sum_{t=1}^T f_i^t, \qquad \text{regret}_i = \text{loss} - \text{loss}_i$$

$$\text{regret} := \max_i \text{regret}_i = \text{loss} - \min_i \sum_t f_i^t = \text{loss} - \min_p \sum_{t=1}^T \langle p, f^t \rangle$$

$$\text{regret}_i \leq \varepsilon \cdot \sum^T (f_i^t)^2 + \frac{\ln n}{\varepsilon}$$

$$\leq \varepsilon \cdot \sum_{i=1}^T |f_i^t| + \frac{\ln n}{\varepsilon} \leq \varepsilon T + \frac{\ln n}{\varepsilon}$$

(choose $\varepsilon = \sqrt{\frac{\ln n}{T}}$

regret $\leq 2 \cdot \sqrt{T \ln n}$)

Using MWU to solve the Set Cover LP (and LPs more generally)

10 May 2019 09:40

[Plotkin, Shmoys, Tardos]

Goal:

solve the following kind of LP:

$$\text{find } \underbrace{x \in P}_{\substack{\text{easy constraints} \\ \text{(such as } x \geq 0)}} \text{ s.t. } \underbrace{Ax \geq b}_{\text{hard constraints}}$$

Example Set Cover

(variable x_i for each set

$$\min \sum_{i=1}^{n} w_i x_i$$
$$\text{s.t. } Ax \geq 1$$
$$x \geq 0$$

$a_{ij} = 1$ if element $i \in [m]$ is contained in set $j \in [n]$

A: $m \times n$ matrix

rephrase as feasibility problem

easy constraints

$$\text{find } x \text{ s.t. } \boxed{x \geq 0, \ \sum w_i x_i \leq \gamma}$$
$$\text{and } \boxed{Ax \geq 1}$$

hard constr.

binary search over γ to minimize cost

Define: $P := \{x \in \mathbb{R}^n : x \geq 0, x \leq 1, \sum w_i x_i \leq \gamma\}$

relax problem a little bit

goal: find $x \in P$ s.t. $Ax \geq 1 - \delta$ (for small $\delta > 0$)

(for set cover, we can set $x' := \frac{1}{1-\delta} \cdot x$
to obtain a feasible solution with objective value $\leq \frac{\gamma}{1-\delta}$)

General Idea

Use MWU to maintain a distribution p^t on the constraints
(we use constraints as experts)

as we go along, we produce a sequence of vectors x^t $\qquad A_i x \geq 1$
loss of expert/constraint i defined by $A_i x^t - 1$

Assume ORACLE to solve the following problem

10 May 2019 09:40

$Ax \geq 1$

Given a prob. distr. $p \in \mathbb{R}^n$ (on constr.)

find $x \in \mathcal{P}$ s.t. $p^T A x \geq 1$

(for general LPs, if no such x exists, original problem is infeasible)

ORACLE for set cover

find x s.t. $x \geq 0$, $x \leq 1$, $\underline{\underline{p^T A x \geq 1}}$

and $w^T x$ is minimized

can be done greedily by setting $x_i > 0$ for most efficient coordinates

efficiency of coord. $i = \dfrac{(p^T A)_i}{w_i}$

Algorithm

MWU, n experts (constraints), initial distr. on experts $p^1 = \left(\frac{1}{n}, \dots, \frac{1}{n}\right)$

In round t:

1) use ORACLE to compute x^t s.t. $x^t \in \mathcal{P}$ and $\underline{\underline{(p^t)^T A x^t \geq 1}}$

2) define loss $f_i^t = \dfrac{A_i x^t - 1}{\rho}$ ← normalization factor

3) update $p^t \to p^{t+1}$ using MWU rule (with param. $\varepsilon \in [0, 1/2]$)

Normalization parameter ρ?

frequency of set cover instance

$\forall x \in \mathcal{P}$ and $\forall i \in [n]$, we have

$$-1 \leq A_i x - 1 \leq \rho = f - 1 \leq n - 1$$

$$\Rightarrow f_i^t \in \left[-\frac{1}{\rho}, 1\right]$$

Expected loss of alg. in round t

$$\langle p^t, f^t \rangle = \frac{1}{\rho} \langle p^t, Ax^t - 1 \rangle = \frac{1}{\rho}\big(\underbrace{\langle p^t, Ax^t \rangle - 1}_{\geq 0}\big) \geq 0$$

exp. loss of MWU alg. is ≥ 0

10 May 2019 09:40

Let us consider some constraint (expert) i

$$0 \leq \text{loss} = \text{loss}_i + \text{regret}_i$$

$$= \sum_{t=1}^{T} \frac{A_i x^t - 1}{\rho} + \text{regret}_i \qquad \left(\text{regret}_i \leq \varepsilon \sum_{t=1}^{T} |f_i^t| + \frac{\ln n}{\varepsilon} \right)$$

$$\leq \sum_{t=1}^{T} \frac{1}{\rho}(A_i x^t - 1) + \varepsilon \sum_{t=1}^{T} \frac{1}{\rho} |A_i x^t - 1| + \frac{\ln n}{\varepsilon}$$

$$= (1+\varepsilon) \cdot \sum_{t=1}^{T} \frac{1}{\rho}(A_i x^t - 1) + 2\varepsilon \cdot \sum_{t:\, A_i x^t - 1 < 0} \frac{1}{\rho} \underbrace{|A_i x^t - 1|}_{\leq 1} + \frac{\ln n}{\varepsilon}$$

$$\leq (1+\varepsilon) \sum_{t=1}^{T} \frac{1}{\rho}(A_i x^t - 1) + \frac{2\varepsilon T}{\rho} + \frac{\ln n}{\varepsilon}$$

set $\bar{x} := \frac{1}{T} \sum_{t=1}^{T} x^t$, $\varepsilon := \frac{\delta}{4}$, $T := \left\lceil \frac{8 \rho \ln n}{\delta^2} \right\rceil$

$$0 \leq (1+\varepsilon) \cdot (A_i \bar{x} - 1) + 2\varepsilon + \frac{\rho \cdot \ln n}{\varepsilon T}$$

$$A_i \bar{x} \geq 1 - \frac{2\varepsilon}{1+\varepsilon} - \frac{\rho \ln n}{\varepsilon T (1+\varepsilon)} = 1 - \frac{\delta/2}{1+\varepsilon} - \underbrace{\frac{\rho \ln n \cdot \delta}{2 \rho \ln n}}_{\delta/2}$$

$$\Rightarrow A_i \bar{x} \geq 1 - \delta$$

$\Rightarrow$ need $O\left(\frac{\rho \log n}{\delta^2}\right)$ repetitions