CSCI 659 Homework 1

Fall 2022

Instructor: Haipeng Luo

This homework is due on **9/25, 11:59pm**. *See course website for more instructions on finishing and submitting your homework as well as the late policy. Total points: 60.*

1. (**Doubling Trick**) (6pts) We have seen that Hedge enjoys a regret bound $2\sqrt{T \ln N}$ with the optimal tuning $\eta = \sqrt{(\ln N)/T}$. What if T is unknown? One simple way to address this issue is the so-called "doubling trick". The idea is to make a guess on T, and once the actual horizon exceeds the guess, double the guess and restart the algorithm. This is outlined below (with $\mathbf{0}$ being the all-zero vector):

Algorithm 1: Hedge with a Doubling Trick

Initialize: $L_0 = \mathbf{0}$, initial guess $T_0 = 1$, and initial learning rate $\eta = \sqrt{(\ln N)/T_0}$
for $t = 1, 2, \ldots,$ **do**
  if $t \geq 2T_0$ **then**
    double the guess: $T_0 \leftarrow 2T_0$
    reset the algorithm: $L_{t-1} = \mathbf{0}$ and $\eta = \sqrt{(\ln N)/T_0}$
  compute $p_t \in \Delta(N)$ such that $p_t(i) \propto \exp(-\eta L_{t-1}(i))$
  play p_t and observe loss vector $\ell_t \in [0,1]^N$
  update $L_t = L_{t-1} + \ell_t$

Prove that Algorithm 1 ensures $\mathcal{R}_T = \mathcal{O}(\sqrt{T \ln N})$ *for all* T. (Hint: consider how many times the algorithm resets and how large the regret can be between two resets.)

2. (**Regret Matching**) Regret Matching is a suboptimal yet extremely simple and practical algorithm for the expert problem. Specifically, let $r_t \in [-1, 1]^N$ be such that $r_t(i) = \langle p_t, \ell_t \rangle - \ell_t(i)$ (that is, the instantaneous regret against expert i), and $R_t = \sum_{s \le t} r_s$. Then at round t, Regret Matching predicts $p_t \in \Delta(N)$ such that

$$p_t(i) \propto [R_{t-1}(i)]_+, \quad \text{where } [x]_+ = \max\{x, 0\}.$$

Prove the regret bound for this algorithm through the following steps.

(a) (4pts) Prove that for any i, $[R_t(i)]_+^2 \le [R_{t-1}(i)]_+^2 + 2[R_{t-1}(i)]_+ r_t(i) + r_t^2(i)$.

(b) (3pts) Define potential $\Phi_t = \sum_{i=1}^N [R_t(i)]_+^2$. Prove $\Phi_t \le \Phi_{t-1} + N$.

(c) (3pts) Conclude that Regret Matching ensures $\mathcal{R}_T \le \sqrt{TN}$.

3. (**Improved Analysis for FTPL**) In Lecture 2, we prove that for the combinatorial problem, FTPL achieves a suboptimal regret bound $\mathcal{O}(m\sqrt{TN\ln N})$. In this exercise, you need to prove that the exact same algorithm actually achieves a better bound $\mathcal{O}(m\sqrt{Tm\ln N})$. (See the lecture for all notations used here.)

   (a) (7pts) In the proof of Lemma 5 of Lecture 2, we prove $p_t(j) \le e^{\eta\|\ell_t\|_1} p_{t+1}(j)$. The key here is to improve this to

   $$p_t(j) \le e^{\eta\langle v_j, \ell_t\rangle} p_{t+1}(j).$$

   To show this, fix any j, and consider an auxiliary distribution $p^j_{t+1} \in \Delta(M)$ such that for any combinatorial action $v_k \in S$:

   $$p^j_{t+1}(k) = \Pr\left[v_k = \operatorname*{argmin}_{w\in\Omega}\left\langle w, \left(\sum_{s=0}^{t-1}\ell_s\right) + v_j \odot \ell_t\right\rangle\right]$$

   where $\odot$ denotes element-wise product. Follow the proof of Lemma 5 to show

   $$p_t(j) \le e^{\eta\langle v_j, \ell_t\rangle} p^j_{t+1}(j),$$

   and then conclude $p_t(j) \le e^{\eta\langle v_j, \ell_t\rangle} p_{t+1}(j)$.

   (b) (5pts) Based on the result from last question, prove $\mathbb{E}[\langle w_t - w_{t+1}, \ell_t\rangle] \le \eta m^2$. Then further conclude the regret bound $\mathcal{O}(m\sqrt{Tm\ln N})$ when using the optimal η.

4. (**Hedge is an FTPL**) Consider the following FTPL strategy for the expert problem: at time t, select expert (recall $L_t = \sum_{s \le t} \ell_s$ is the cumulative loss vector)

$$i_t = \underset{i}{\operatorname{argmin}} \left(L_{t-1}(i) - \ell_0(i) \right),$$

where $\ell_0(i)$ for $i = 1, \ldots, N$ are N independent random variables with *Gumbel distribution*, that is, with CDF: $\Pr[\ell_0(i) \le x] = \exp(-\exp(-\eta x))$ for some parameter η.

(a) (3pts) Prove that for any j, $\Pr[i_t = j] = \Pr\left[j = \operatorname{argmax}_i \frac{\exp(-\eta L_{t-1}(i))}{\exp(-\eta \ell_0(i))} \right]$.

(b) (3pts) Prove that the random variable $\beta(i) = \exp(-\eta \ell_0(i))$ follows the standard exponential distribution, that is $\Pr[\beta(i) \le x] = 1 - e^{-x}$.

(c) (6pts) For any $a \in \mathbb{R}^N_{>0}$, prove that for any j, $\Pr\left[j = \operatorname{argmax}_i \frac{a(i)}{\beta(i)} \right] = \frac{a(j)}{\sum_{i=1}^N a(i)}$. Conclude that FTPL with Gumbel noise is equivalent to Hedge.

5. (**Online Mirror Descent**) Besides FTRL and FTPL, *Online Mirror Descent* (OMD) is yet another general framework to derive online learning algorithm for OCO. For a convex regularizer function $\psi : \Omega \to \mathbb{R}$ (also called mirror map) and a learning rate $\eta > 0$, the update of OMD is

$$w_{t+1} = \operatorname*{argmin}_{w \in \Omega} \langle w, \ell_t \rangle + \frac{1}{\eta} D_\psi(w, w_t),$$

starting from an arbitrary $w_1 \in \Omega$. In other words, OMD tries to find a point that minimizes the loss at time t while being close to the previous point w_t (in terms of their Bregman divergence). In this exercise, you will prove a regret bound for OMD similar to that of FTRL and instantiate OMD in two examples.

(a) (5pts) Use Lemma 1 from Lecture 2 to prove for any $u \in \Omega$:

$$\eta \langle w_{t+1} - u, \ell_t \rangle \le D_\psi(u, w_t) - D_\psi(u, w_{t+1}) - D_\psi(w_{t+1}, w_t), \tag{1}$$

then further conclude that OMD's regret against any u is bounded as:

$$\sum_{t=1}^{T} \langle w_t - u, \ell_t \rangle \le \frac{D_\psi(u, w_1)}{\eta} + \sum_{t=1}^{T} \langle w_t - w_{t+1}, \ell_t \rangle - \frac{1}{\eta} \sum_{t=1}^{T} D_\psi(w_{t+1}, w_t). \tag{2}$$

(Note the similarity of this bound compared to that in Lemma 3 of Lecture 2 for FTRL.)

(b) (5pts) Suppose that ψ is strongly convex with respect to some norm $\|\cdot\|$. By setting $u = w_t$ in Eq. (1), prove the stability of OMD: $\|w_t - w_{t+1}\| \le \eta \|\ell_t\|_\star$ (the same stability property that FTRL enjoys), then conclude the regret bound

$$\mathcal{R}_T \le \frac{\max_{u \in \Omega} D_\psi(u, w_1)}{\eta} + \eta \sum_{t=1}^{T} \|\ell_t\|_\star^2 . \tag{3}$$

(c) (5pts) Show that Hedge is an instance of OMD with a specific ψ, then recover its regret bound using Eq. (3) (assuming w_1 is the uniform distribution).

(d) (5pts) Use $\psi(w) = \frac{1}{2} \|w\|_2^2$ to derive the non-lazy version of OGD we discussed in Lecture 2. Then apply Eq. (3) to show that with the optimal η OMD enjoys $\mathcal{R}_T = \mathcal{O}(\text{diam}(\Omega) G \sqrt{T})$ where $\text{diam}(\Omega) = \max_{w,u \in \Omega} \|w - u\|_2$ is the diameter of Ω and G is such that $\max_t \|\ell_t\|_2 \le G$.