# 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.)

*Proof.* Fix any value of $T$. It is straightforward to see that the algorithm resets for $n = \lfloor \log_2 T \rfloor$ times. For any $k = 0, 1, \ldots, n-1$, after the $k$-th reset, a new instance of Hedge with the optimal tunning is run for $2^k$ rounds (from $t = 2^k$ to $t = 2^{k+1} - 1$) before the next reset, and thus suffers at most $2\sqrt{2^k \ln N}$ regret: $\max_{p \in \Delta(N)} \sum_{t=2^k}^{2^{k+1}-1} \langle p_t - p, \ell_t \rangle \leq 2\sqrt{2^k \ln N}$.
After the last reset, a new instance of Hedge is run for no more than $2^n$ rounds, but one can imagine that it is still run for exactly $2^n$ rounds by feeding the $\mathbf{0}$ loss vector to the algorithm for the extra imaginary rounds, which has no effect to the regret. Therefore, the regret after the last reset is bounded by $2\sqrt{2^n \ln N}$: $\max_{p \in \Delta(N)} \sum_{t=2^n}^{T} \langle p_t - p, \ell_t \rangle \leq 2\sqrt{2^n \ln N}$.
Now we conclude that the total regret is bounded as

$$\max_{p \in \Delta(N)} \sum_{t=1}^{T} \langle p_t - p, \ell_t \rangle \leq \left( \sum_{k=0}^{n-1} \max_{p \in \Delta(N)} \sum_{t=2^k}^{2^{k+1}-1} \langle p_t - p, \ell_t \rangle \right) + \max_{p \in \Delta(N)} \sum_{t=2^n}^{T} \langle p_t - p, \ell_t \rangle$$

$$\leq \sum_{k=0}^{n} 2\sqrt{2^k \ln N} = 2\frac{\sqrt{2^{n+1}} - 1}{\sqrt{2} - 1}\sqrt{\ln N} \leq 2\frac{\sqrt{2T} - 1}{\sqrt{2} - 1}\sqrt{\ln N} = \mathcal{O}(\sqrt{T \ln N}),$$

which completes the proof. □

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 $i$, $[R_t(i)]_+^2 \le [R_{t-1}(i)]_+^2 + 2[R_{t-1}(i)]_+ r_t(i) + r_t^2(i)$.

*Proof.* Notice that (1) $[x]_+^2$ in non-decreasing in $x$ and (2) $[x]_+^2 \le x^2$. Therefore if $R_{t-1}(i) \le 0$, then

$$\begin{aligned}
[R_t(i)]_+^2 &= [R_{t-1}(i) + r_t(i)]_+^2 \\
&\le [r_t(i)]_+^2 \le r_t^2(i) && \text{(by (1) and (2))} \\
&= [R_{t-1}(i)]_+^2 + 2[R_{t-1}(i)]_+ r_t(i) + r_t^2(i). && ([R_{t-1}(i)]_+ = 0)
\end{aligned}$$

Otherwise if $R_{t-1}(i) \ge 0$, then by (2) $[R_t(i)]_+^2 \le R_t(i)^2 = R_{t-1}(i)^2 + 2R_{t-1}(i)r_t(i) + r_t^2(i) = [R_{t-1}(i)]_+^2 + 2[R_{t-1}(i)]_+ r_t(i) + r_t^2(i)$. □

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

*Proof.* Summing up the results of the last question over $i$, we get

$$\Phi_t \le \Phi_{t-1} + 2\sum_{i=1}^N [R_{t-1}(i)]_+ r_t(i) + \sum_{i=1}^N r_t^2(i) \le \Phi_{t-1} + 2\sum_{i=1}^N [R_{t-1}(i)]_+ r_t(i) + N.$$

It suffices to prove $\sum_{i=1}^N [R_{t-1}(i)]_+ r_t(i) = 0$, which is true because

$$\sum_{i=1}^N p_t(i) r_t(i) = \sum_{i=1}^N p_t(i) \langle p_t, \ell_t \rangle - \sum_{i=1}^N p_t(i)\ell_t(i) = \langle p_t, \ell_t \rangle - \langle p_t, \ell_t \rangle = 0,$$

and $p_t(i) \propto [R_{t-1}(i)]_+$. □

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

*Proof.* Using the conclusion from the last question we have $\Phi_T \le \Phi_0 + TN = TN$. Therefore, the regret can be bounded as

$$\mathcal{R}_T = \max_i R_T(i) \le \max_i [R_T(i)]_+ = \sqrt{\max_i [R_T(i)]_+^2} \le \sqrt{\Phi_T} \le \sqrt{TN},$$

finishing the proof. □

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 $p_{t+1}^j \in \Delta(M)$ such that for any combinatorial action $v_k \in S$:

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

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

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

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

   *Proof.* The reasoning of the first step is exactly the same as that in Lemma 5:

$$\begin{aligned}
p_t(j) &= \int_{\ell_0\in\mathbb{R}^N} \mathbf{1}\left[v_j = \operatorname*{argmin}_{w\in\Omega} \left\langle w, \sum_{s=0}^{t-1}\ell_s \right\rangle\right] h(\ell_0)d\ell_0 \\
&= \int_{\ell_0\in\mathbb{R}^N} \mathbf{1}\left[v_j = \operatorname*{argmin}_{w\in\Omega} \left\langle w, \sum_{s=0}^{t-1}\ell_s + v_j\odot\ell_t \right\rangle\right] h(\ell_0 + v_j\odot\ell_t)d\ell_0 \\
&\le \int_{\ell_0\in\mathbb{R}^N} \mathbf{1}\left[v_j = \operatorname*{argmin}_{w\in\Omega} \left\langle w, \sum_{s=0}^{t-1}\ell_s + v_j\odot\ell_t \right\rangle\right] h(\ell_0)e^{\eta\|v_j\odot\ell_t\|_1}d\ell_0 \\
&= e^{\eta\|v_j\odot\ell_t\|_1} p_{t+1}^j(j) = e^{\eta\langle v_j,\ell_t\rangle} p_{t+1}^j(j).
\end{aligned}$$

   It remains to argue $p_{t+1}^j(j) \le p_{t+1}(j)$. This is true because whenever $v_j$ minimizes the function $\left\langle w, \sum_{s=0}^{t-1}\ell_s + v_j\odot\ell_t\right\rangle$, it has to also minimize $\left\langle w, \sum_{s=0}^{t}\ell_s\right\rangle$, given that $v_j$ achieves the same value for these two functions, while all other $v_k$ ($k \ne j$) leads to a larger (if not equal) value for the second function. □

   (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 $\eta$.

   *Proof.* This also follows similar reasoning as the proof of Lemma 5:

$$\begin{aligned}
\mathbb{E}[\langle w_t - w_{t+1}, \ell_t\rangle] &= \sum_{j=1}^M (p_t(j) - p_{t+1}(j))\langle v_j, \ell_t\rangle \le \sum_{j=1}^M (1 - e^{-\eta\langle v_j,\ell_t\rangle}) p_t(j)\langle v_j, \ell_t\rangle \\
&\le \eta\sum_{j=1}^M p_t(j)\langle v_j,\ell_t\rangle^2 \le \eta m^2 \sum_{j=1}^M p_t(j) = \eta m^2,
\end{aligned}$$

   where the first inequality uses the result from the last question and the second inequality uses the fact $1 - e^{-z} \le z$ for all $z$. Finally, based on Lemma 3 of Lecture 2, we have

$$\mathbb{E}[\mathcal{R}_T] \le \frac{\mathbb{E}[\max_w\langle w,\ell_0\rangle - \min_w\langle w,\ell_0\rangle]}{\eta} + \sum_{t=1}^T \mathbb{E}[\langle w_t - w_{t+1}, \ell_t\rangle],$$

   where the first term is still bounded by $\frac{2m}{\eta}(1+\ln N)$ according to the proof of Theorem 2, and the second term is now improved to $\eta T m^2$. Picking the optimal $\eta$ then shows $\mathbb{E}[\mathcal{R}_T] = \mathcal{O}(m\sqrt{Tm\ln N})$. □

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)

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

where $\ell_0(i)$ for $i = 1, \dots, N$ 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 $\eta$.

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

*Proof.* This is by definition and rewriting:

$$\begin{aligned}
\Pr[i_t = j] &= \Pr[j = \operatorname*{argmin}_i \left( L_{t-1}(i) - \ell_0(i) \right)] \\
&= \Pr[j = \operatorname*{argmin}_i \exp\left( \eta L_{t-1}(i) - \eta \ell_0(i) \right)] \\
&= \Pr\left[ j = \operatorname*{argmax}_i \frac{\exp(-\eta L_{t-1}(i))}{\exp(-\eta \ell_0(i))} \right].
\end{aligned}$$

□

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

*Proof.* This is also by definition and direct calculation:

$$\begin{aligned}
\Pr[\beta(i) \le x] &= \Pr[\exp(-\eta \ell_0(i)) \le x] \\
&= \Pr[\ell_0(i) \ge -\tfrac{1}{\eta} \ln x] \\
&= 1 - \Pr[\ell_0(i) \le -\tfrac{1}{\eta} \ln x] \\
&= 1 - \exp(-\exp(-\eta(-\tfrac{1}{\eta} \ln x))) \\
&= 1 - e^{-x}.
\end{aligned}$$

□

(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.

*Proof.* Note that the density of the standard exponential distribution is $e^{-x}$. Direct calculation shows

$$\begin{aligned}
\Pr\left[ j = \operatorname*{argmax}_i \frac{a(i)}{\beta(i)} \right] &= \int_0^\infty e^{-\beta(j)} \Pr\left[ \frac{a(i)}{\beta(i)} \le \frac{a(j)}{\beta(j)}, \quad \forall i \ne j \right] d\beta(j) \\
&= \int_0^\infty e^{-\beta(j)} \prod_{j \ne i} \Pr\left[ \frac{a(i)}{\beta(i)} \le \frac{a(j)}{\beta(j)} \right] d\beta(j) \quad \text{(by independence)} \\
&= \int_0^\infty e^{-\beta(j)} \prod_{j \ne i} \Pr\left[ \frac{a(i)}{a(j)} \beta(j) \le \beta(i) \right] d\beta(j) \\
&= \int_0^\infty e^{-\beta(j)} \prod_{j \ne i} \exp\left( -\frac{a(i)}{a(j)} \beta(j) \right) d\beta(j) \qquad \text{(result from (b))} \\
&= \int_0^\infty \exp\left( \frac{-\sum_{i=1}^N a(i)}{a(j)} \beta(j) \right) d\beta(j) \\
&= \frac{-a(j)}{\sum_{i=1}^N a(i)} \exp\left( \frac{-\sum_{i=1}^N a(i)}{a(j)} \beta(j) \right) \Bigg|_0^\infty \\
&= \frac{a(j)}{\sum_{i=1}^N a(i)}.
\end{aligned}$$

Combining all the results shows that $\Pr[i_t = j] \propto \exp(\eta L_{t-1}(j))$, same as 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.)

*Proof.* Let $F(w) = \langle w, \ell_t \rangle + \frac{1}{\eta} D_\psi(w, w_t)$. Since $w_{t+1}$ minimizes $F$, by applying Lemma 1 we have for any $u \in \Omega$:

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

Note that the only non-linear term in $F$ is $\frac{1}{\eta}\psi(w)$, and thus $D_F = \frac{1}{\eta} D_\psi$. Rearranging then proves the first statement. The second statement is simply by adding $\langle w_t, \ell_t \rangle$ to both sides, summing over $t$, rearranging, telescoping, and the fact $D_\psi(u, w_{T+1}) \ge 0$. ☐

(b) (5pts) Suppose that $\psi$ 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

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

*Proof.* Setting $u = w_t$ in Eq. (1) and using strong convexity gives

$$\eta \langle w_{t+1} - w_t, \ell_t \rangle \le -D_\psi(w_t, w_{t+1}) - D_\psi(w_{t+1}, w_t) \le - \|w_t - w_{t+1}\|^2 .$$

Rearranging and using Hölder's inequality, we arrive at

$$\|w_t - w_{t+1}\|^2 \le \eta \langle w_t - w_{t+1}, \ell_t \rangle \le \eta \|w_t - w_{t+1}\| \|\ell_t\|_\star .$$

Dividing both sides by $\|w_t - w_{t+1}\|$ finishes the proof for the first statement. The second statement is a direct application of Eq. (2), Hölder's inequality, and the stability property we just proved. ☐

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

*Proof.* Let the regularizer be the (negative) entropy $\psi(p) = \sum_i p(i) \ln p(i)$. Then $D_\psi(p, q)$ is exactly the KL divergence $\sum_{i=1}^{N} p(i) \ln \frac{p(i)}{q(i)}$, and OMD becomes

$$p_{t+1} = \operatorname*{argmin}_{p \in \Delta(N)} \langle p, \ell_t \rangle + \frac{1}{\eta} \sum_{i=1}^{N} p(i) \ln \frac{p(i)}{p_t(i)}.$$

Direct calculation (by writing down Lagrangian and setting the gradient to zero) shows $p_{t+1}(i) \propto p_t(i) \exp(-\eta \ell_t(i))$. Expanding this definition recursively shows $p_{t+1}(i) \propto \exp\left(-\eta \sum_{s \le t} \ell_s(i)\right)$, exactly the same as Hedge.

As for the regret bound, we already know that $\psi$ is strongly convex with respect to the $L_1$ norm, so Eq. (3) implies

$$\mathcal{R}_T \leq \frac{\max_{p\in\Delta(N)} \sum_{i=1}^N p(i)\ln(Np(i))}{\eta} + \eta T \leq \frac{\ln N}{\eta} + \eta T,$$

which is $2\sqrt{T \ln N}$ with the optimal $\eta$, recovering the same bound we proved before. □

(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 $\eta$ 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 $\Omega$ and $G$ is such that $\max_t \|\ell_t\|_2 \leq G$.

*Proof.* In this case, we have $D_\psi(w,u) = \frac{1}{2}\|w-u\|_2^2$ and thus

$$w_{t+1} = \operatorname*{argmin}_{w\in\Omega} \langle w, \ell_t\rangle + \tfrac{1}{2\eta}\|w-w_t\|_2^2 = \operatorname*{argmin}_{w\in\Omega} \|w-(w_t-\eta\ell_t)\|_2^2,$$

which is equivalent to the non-lazy version of OGD:

$$u_{t+1} = w_t - \eta\ell_t; \quad w_{t+1} = \operatorname*{argmin}_{w\in\Omega}\|w-u_{t+1}\|_2.$$

As for the regret bound, recall that $\psi$ is strongly convex with respect to $L_2$ norm. Applying Eq. (3) thus proves

$$\mathcal{R}_T \leq \frac{\max_{u\in\Omega}\|u-w_1\|_2^2}{2\eta} + \eta\sum_{t=1}^T\|\ell_t\|_2^2 \leq \frac{\text{diam}(\Omega)^2}{2\eta} + \eta T G^2.$$

Picking the optimal $\eta$ finishes the proof. □