Homework 1

Instructor: Haipeng Luo

1. Construct an example to show that Follow the Leader (FTL) could have $\Omega(T)$ regret in the worst case. (Hint: think about a very simple setting such as the expert problem with two experts.)

2. (**Doubling Trick**) We have seen that Hedge has regret bound $2\sqrt{T \ln N}$ with the optimal tuning $\eta = \sqrt{(\ln N)/T}$. What if T is unknown? One quick-and-dirty 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 with a learning rate tuned based on the new guess. The algorithm is outlined below (with $\mathbf{0}$ being the all-zero vector):

Algorithm 1: Doubling Trick with Hedge

```
Initialize: L_0 = 0 and η = sqrt((ln N)/T_0) where T_0 = 2
for t = 1, 2, ..., do
    if t > T_0 then
        double the guess: T_0 ← 2T_0
        reset the algorithm: L_{t-1} = 0 and η = sqrt((ln N)/T_0)
    compute p_t ∈ Δ(N) such that p_t(i) ∝ exp(−η L_{t−1}(i))
    play p_t and observe loss vector ℓ_t ∈ [0,1]^N
    update L_t = L_{t−1} + ℓ_t
```

(a) Prove that Algorithm 1 ensures that *for all* T, we have $R_T(i^\star) = \mathcal{O}(\sqrt{T \ln N})$. (Hint: consider how many times the algorithm resets and how large the regret can be between two resets.)

(b) In Lecture 3 we showed how to use time-varying learning rate to achieve "small-loss" bounds. Use the doubling trick to outline an algorithm that achieves the same bound $\mathcal{O}(\sqrt{L_T(i^\star) \ln N})$ without the knowledge of $L_T(i^\star)$.

(c) In Lecture 6 we showed an algorithm that ensures $R_T(i) \leq \frac{\ln N}{\eta} + \eta D_T(i)$ for $D_T(i) = \sum_{t=1}^{T}(\ell_t(i) - \ell_{t-1}(i))^2$. Can we again use the doubling trick to achieve $R_T(i^\star) = \mathcal{O}(\sqrt{D_T(i^\star) \ln N})$ without the knowledge of $D_T(i^\star)$?

(d) In general if we have an algorithm (with parameter η) that ensures $R_T(i^\star) \leq \frac{\ln N}{\eta} + \eta B_T$ for some quantity B_T, under what conditions can we use the doubling trick to achieve the bound $R_T(i^\star) = \mathcal{O}(\sqrt{B_T \ln N})$ without the knowledge of B_T?

3. (**Regret Matching**) Regret Matching is a suboptimal yet extremely simple algorithm for the expert problem. Specially, at round t Regret Matching predicts

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

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

(a) Prove that $[R_t(i)]_+^2 \leq [R_{t-1}(i)]_+^2 + 2[R_{t-1}(i)]_+ r_t(i) + r_t^2(i)$.

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

(c) Conclude that regret matching ensures $R_T(i^\star) \le \sqrt{TN}$.

4. (**Hedge is an FTPL**) Consider the following FTPL strategy for the expert problem: at time t, plays

$$i_t = \operatorname*{argmin}_i \left(L_{t-1}(i) - L_0(i)\right),$$

where $L_0(i)$ for $i = 1, \dots, N$ are N independent random variables with *Gumbel distribution*, that is, with CDF $\Pr[L_0(i) \le x] \le \exp(-\exp(-\eta x))$ for some parameter η.

(a) Prove that for any j, $\Pr[i_t = j] = \Pr\left[j = \operatorname{argmax}_i \frac{\exp(-\eta L_{t-1}(i))}{\exp(-\eta L_0(i))}\right]$.

(b) Prove that the random variable $v(i) = \exp(-\eta L_0(i))$ follows the standard exponential distribution, that is $\Pr[v(i) \le x] = 1 - e^{-x}$.

(c) For any positive numbers $a(1), \dots, a(N)$, prove that $\Pr\left[j = \operatorname{argmax}_i \frac{a(i)}{v(i)}\right] = \frac{a(j)}{\sum_{i=1}^N a(i)}$. Conclude that FTPL with Gumbel noise is equivalent to sampling an expert using Hedge's prediction.

5. (**Perceptron**) Consider the following algorithm (called *Perceptron*) for the online binary classification problem:

Algorithm 2: Perceptron

Initialize: $w_1 = \mathbf{0} \in \mathbb{R}^d$
for $t = 1, 2, \dots, T$ **do**
  see example $x_t \in \mathbb{R}^d$
  predict $\widehat{y}_t = \textsc{sgn}(\langle w_t, x_t \rangle)$ where $\textsc{sgn}(y)$ is 1 if $y \ge 0$ and -1 otherwise
  receive true label $y_t \in \{-1, 1\}$
  if $\widehat{y}_t \ne y_t$ **then** $w_{t+1} = w_t + y_t x_t$
  else $w_{t+1} = w_t$

(a) Construct a loss function $f_t(w)$ so that Perceptron is equivalent to OGD with $\Omega = \mathbb{R}^d$ and $\eta = 1$, that is, $w_{t+1} = w_t - \nabla f_t(w_t)$. (For simplicity ignore the case when $\langle w_t, x_t \rangle = 0$.)

(b) Since the regret bound of OGD requires boundedness of Ω, it does not apply here. Instead, one can prove a *mistake bound* for Perceptron. For notation convenience, first ignore the rounds where the algorithm makes the correct prediction (that is, $\widehat{y}_t = y_t$). In other words, assume that the algorithm makes a mistake at every round so that T is exactly the number of mistakes. This is without loss of generality since the algorithm makes no changes unless it makes a mistake. Next, prove the following statements (all norms are L_2 norm).

   (i) $\|w_{t+1}\|^2 \le \|w_t\|^2 + 1$ if examples are normalized so that $\|x_t\| \le 1$.

   (ii) Assume there exists a perfect linear classifier $w^\star$ with $\|w^\star\| \le 1$ and *margin* $\gamma > 0$, that is, $y_t \langle w^\star, x_t \rangle \ge \gamma$ for all x_t. Prove $\|w_{T+1}\| \ge T\gamma$.

   (iii) Combine the above two statements to show that the number of mistakes T is bounded by $1/\gamma^2$.

6. (**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 function $\psi : \Omega \to \mathbb{R}$, the Bregman divergence (with respect to ψ) between two points w and u is defined as

$$D_\psi(w, u) = \psi(w) - \psi(u) - \langle \nabla \psi(u), w - u \rangle.$$

The update of OMD is then

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

for some parameter η. In other words, OMD tries to find a point that minimizes the loss at time t (recall all OCO problems can be reduced to a linear problem with loss function $\langle w, \nabla f_t(w_t) \rangle$ at time t) while being close to the previous point w_t.

(a) Let w'_{t+1} be such that $\nabla\psi(w'_{t+1}) = \nabla\psi(w_t) - \eta\nabla f_t(w_t)$ (assume it exists). Prove

$$w_{t+1} = \operatorname*{argmin}_{w\in\Omega} D_\psi(w, w'_{t+1}). \tag{1}$$

(b) Verify that for any $u \in \Omega$, the instantaneous regret can be written as

$$\langle w_t - u, \nabla f_t(w_t)\rangle = \tfrac{1}{\eta}\left(D_\psi(u, w_t) - D_\psi(u, w'_{t+1}) + D_\psi(w_t, w'_{t+1})\right)$$

(c) Use the generalized Pythagorean theorem, which states that $D_\psi(u, w_{t+1}) \leq D_\psi(u, w'_{t+1})$ for any $u \in \Omega$ if w_{t+1} is a projection of w'_{t+1} as in Eq. (1), to conclude the following regret bound of OMD:

$$\sum_{t=1}^{T} f_t(w_t) - f_t(u) \leq \frac{D_\psi(u, w_1)}{\eta} + \frac{1}{\eta}\sum_{t=1}^{T} D_\psi(w_t, w'_{t+1}). \tag{2}$$

(d) Show that Hedge is an instance of OMD and recover its regret bound using Eq. (2).

(e) Use $\psi(w) = \frac{1}{2}\|w\|_2^2$ to derive a different version of OGD. Prove its regret bound using Eq. (2).