This homework is due at 23:59, June 22, 2019.

Problem 1

In this problem, we will solve an online regression problem using online gradient descent. Consider the following game of *online regression*. For every $t \in \mathbb{N}$, the following happen sequentially.

1. REALITY announces $x_t \in \mathbb{R}^p$.
2. LEARNER announces $\gamma_t \in \mathbb{R}$.
3. REALITY announces $y_t \in [-L, L]$ for some $L > 0$.
4. LEARNER suffers the loss
$$f(y_t, \gamma_t) := \frac{1}{2}\left(y_t - \gamma_t\right)^2.$$

We aim at competing with linear regression functions. Then, the associated regret is given by

$$R_T(w) := \sum_{t=1}^{T} f(y_t, \gamma_t) - \sum_{t=1}^{T} f(y_t, \langle x_t, w\rangle), \quad \forall w \in \mathbb{R}^p, T \in \mathbb{N}.$$

1. (10 points) Consider the following algorithm. Let w_1 be the zero vector in $\mathbb{R}^p$. For every $t \in \mathbb{N}$, the algorithm computes

$$\begin{aligned}
\tilde{w}_{t+1} &\leftarrow w_t + \eta\left(y_t - \langle x_t, w_t\rangle\right) x_t, \\
w_{t+1} &\leftarrow \begin{cases} \tilde{w}_{t+1} & \text{, if } \|\tilde{w}_{t+1}\| \leq 1, \\ \frac{\tilde{w}_{t+1}}{\|\tilde{w}_{t+1}\|_2} & \text{, otherwise.} \end{cases} \\
\gamma_{t+1} &\leftarrow \langle x_{t+1}, w_{t+1}\rangle,
\end{aligned}$$

for some $\eta > 0$. **Show that the algorithm is an instance of online (projected) gradient descent. Specify the corresponding sequence of loss functions and constraint set.**

Solution. Define
$$f_t(w) := \frac{1}{2}\left(y_t - \langle x_t, w_t\rangle\right)^2.$$
Then, we can solve the online regression problem via solving online convex optimization with f_t on a constraint set to be defined and setting $\gamma_t = \langle x_t, w_t\rangle$. We have

$$\nabla f_t(w) = -\left(y_t - \langle x_t, w\rangle\right) x_t.$$

Let $\mathscr{B}$ be the unit 2-norm ball in $\mathbb{R}^p$ and $\Pi_{\mathscr{B}}$ be the projection onto $\mathscr{B}$, i.e.,

$$\Pi_{\mathscr{B}}(w) := \underset{w' \in \mathscr{B}}{\operatorname{argmin}} \frac{1}{2}\|w' - w\|_2^2.$$

Notice that
$$\Pi_{\mathscr{B}}(w) := \frac{w}{\|w\|_2}, \quad \forall w \in \mathbb{R}^p.$$
Then, we can write the algorithm equivalently as

$$w_{t+1} \leftarrow \Pi_{\mathscr{B}}\left(y_t - \eta \nabla f_t(w_t)\right),$$

showing the algorithm is simply online gradient descent (OGD) with respect to the constraint set $\mathscr{B}$.

2. (10 points) Let $\mathscr{B}$ be the unit 2-norm ball in $\mathbb{R}^p$. Suppose $x_t \in \mathscr{B}$ for all $t \in \mathbb{N}$. **Show the algorithm in the previous problem can achieve**

$$R_T(w) = O\left(L\sqrt{T}\right), \quad \forall w \in \mathscr{B}.$$

How should we set η?

Solution. By Cauchy-Schwarz inequality, we obtain

$$\|\nabla f_t(w)\|_2 \leq |y_t - \langle x_t, w\rangle| \leq |y_t| + |\langle x_t, w\rangle| \leq L+1, \quad \forall w \in \mathscr{B}.$$

By the triangle inequality, we have

$$\max_{w' \in \mathscr{B}} \|w - w'\|_2^2 \leq \left(\|w\|_2 + \|w'\|_2\right)^2 \leq 4, \quad \forall w \in \mathscr{B}.$$

Applying the regret bound for OGD in Lecture 8, we obtain

$$R_T(w) \leq \frac{2}{\eta} + \frac{\eta T (L+1)^2}{2}.$$

Choosing

$$\eta = \frac{2}{(L+1)\sqrt{T}},$$

the desired regret bound follows.

Problem 2

In this problem, we will show Hedge solves an *adversarial multi-armed bandit problem*, with minor modification. The bandit problem is as follows. Let $\mathscr{K} = \{1, \dots, K\}$ for some $K \in \mathbb{N}$. Let $T \in \mathbb{N}$. We consider the *oblivious adversary model*, in which REALITY chooses a sequence $(\omega_t)_{t \in \mathbb{N}}$ of vectors in $[0,1]^K$ before the first round. For each round t, $1 \leq t \leq T$, the following happen sequentially.

- LEARNER announces $\gamma_t \in \mathscr{K}$.
- REALITY announces $\omega_t(\gamma_t)$, the γ_t-th entry of ω_t (while keeping the entire vector ω_t secret).
- LEARNER suffers loss $\lambda(\omega_t, \gamma_t) := \omega_t(\gamma_t)$.

The regret is now given by

$$R_T(k) := \mathsf{E}\left[\sum_{t=1}^{T} \lambda(\omega_t, \gamma_t) - \sum_{t=1}^{T} \lambda(\omega_t, k)\right], \quad \forall k \in \mathscr{K},$$

where the expectation is with respect to the possible randomness of LEARNER's algorithm.

Consider the following algorithm. Let π_1 be the uniform probability distribution on $\mathscr{K}$. For each round t, LEARNER chooses $\gamma_t \in \mathscr{K}$ randomly following π_t, and computes π_{t+1} such that

$$\pi_{t+1}(k) \propto \pi_t(k) \mathrm{e}^{-\eta \tilde{\lambda}(\omega_t, k)}, \quad \forall k \in \mathscr{K},$$

where

$$\tilde{\lambda}(\omega_t, k) := \frac{\lambda(\omega_t, k)}{\pi_t(k)} \mathbb{1}_{\{\gamma_t = k\}}, \quad \forall k \in \mathscr{K}.$$

Recall that $\mathbb{1}_{\{\gamma_t = k\}}$ denotes the indicator function of the event $\{\gamma_t = k\}$.

1. (10 points) **Show that**

$$\mathsf{E}_{\gamma_t \sim \pi_t} \tilde{\lambda}(\omega_t, k) = \lambda(\omega_t, k), \quad \forall k \in \mathscr{K}.$$

That is, $\tilde{\lambda}(\omega_t, k)$ is an unbiased estimate of $\lambda(\omega_t, k)$.

2. (10 points) The following theorem provides an upper bound of the mixability gap.

   Theorem 1 ([1]). *Let ξ be a non-negative random variable and $\eta > 0$. Then, it holds that*

   $$\log \mathsf{E}\, \mathrm{e}^{-\eta(\xi - \mathsf{E}\,\xi)} \le \frac{\eta^2}{2} \mathsf{E}\,\xi^2.$$

   Use the theorem to show that

   $$\sum_{t=1}^{T} \lambda(\omega_t, \gamma_t) \le \frac{\eta}{2} \sum_{t=1}^{T} \sum_{k \in \mathcal{K}} \pi_t(k) \left[\tilde{\lambda}(\omega_t, k)\right]^2 - \frac{1}{\eta} \log \sum_{k \in \mathcal{K}} \pi_1(k) \mathrm{e}^{-\eta \sum_{t=1}^{T} \tilde{\lambda}(\omega_t, k)}.$$

3. (10 points) **Show that the algorithm can achieve**

   $$R_T = O\left(\sqrt{TK \log K}\right).$$

   How should we choose η?

Solution. See Peter Bartlett's lecture slides `https://www.stat.berkeley.edu/~bartlett/courses/2014fall-cs294stat260/lectures/bandit-adversarial-notes.pdf`.

Problem 3

In this problem, we will derive Fixed Share as a special case of the aggregating algorithm. Consider the following protocol. Let $\mathcal{K} := \{1, \dots, K\}$ be the *pool of predictors* for some $K \in \mathbb{N}$, and let π be the uniform probability distribution on $\mathcal{K}$. Let $k^\star \in \mathcal{K}$ be a random variable following π. Let Γ be the prediction space, and Ω be the outcome space. For every $t \in \mathbb{N}$, the following happen in order.

1. STOCHASTIC EXPERT announces a function $\Phi_t : \mathcal{K} \to \Gamma$ and the corresponding random prediction

   $$\xi_t := \Phi_t(k^\star).$$

2. LEARNER announces $\gamma_t \in \Gamma$.
3. REALITY announces $\omega_t \in \Omega$.
4. LEARNER suffers the loss $\lambda(\omega_t, \gamma_t)$, for some η-mixable loss function $\lambda : \Omega \times \Gamma \to \mathbb{R}$.

Assume the conditions on Ω, Γ, and λ for the aggregating algorithm hold in this protocol.

Let $T \in \mathbb{N}$. Define the cumulative losses of LEARNER and STOCHASTIC EXPERT respectively as

$$L_T(\mathrm{L}) := \sum_{t=1}^{T} \lambda(\omega_t, \gamma_t), \quad L_T(\mathrm{SE}) := \sum_{t=1}^{T} \lambda(\omega_t, \xi_t).$$

Notice that $L_T(\mathrm{SE})$ is a random variable, while $L_T(\mathrm{L})$ is a deterministic number.

Consider the following strategy for LEARNER. Let $\pi_1 = \pi$. For every $t \in \mathbb{N}$, announce any prediction $\gamma_t \in \Gamma$ such that

$$\lambda(\omega, \gamma_t) \le \frac{-1}{\eta} \log \sum_{k \in \mathcal{K}} \pi_t(k) \mathrm{e}^{-\eta \lambda(\omega, \Phi_t(k))}, \quad \forall \omega \in \Omega,$$

and then compute π_{t+1} as

$$\pi_{t+1}(k) = \frac{\pi_t(k) \mathrm{e}^{-\eta \lambda(\omega_t, \Phi_t(k))}}{\sum_{k \in \mathcal{K}} \pi_t(k) \mathrm{e}^{-\eta \lambda(\omega_t, \Phi_t(k))}}, \quad \forall k \in \mathcal{K}.$$

Problem 3 continued on the next page...

1. (10 points) **Show that**
$$L_T(\mathbf{L}) \leq \frac{-1}{\eta} \log \mathsf{E}\, \mathrm{e}^{-\eta L_T(\mathbf{SE})},$$
where the expectation is with respect to π.

2. (10 points) Suppose that
$$\mathsf{P}\left(L_T(\mathrm{SE}) \leq L\right) \geq p,$$
for some $L > 0$ and $p \in]0,1]$. **Show that**
$$L_T(\mathbf{L}) \leq L + \frac{1}{\eta} \log \frac{1}{p}. \tag{1}$$

3. (10 points) Consider the standard formulation of learning with expert advice (where there does not exist a stochastic expert) with the same outcome space Ω, prediction space Γ, and loss function λ. Suppose there are K experts and we use the uniform distribution as the prior to run the aggregating algorithm. **Use** (1) **to show that**
$$\sum_{t=1}^{T} \lambda(\omega_t, \gamma_t) \leq \sum_{t=1}^{T} \lambda(\omega_t, \gamma_t(k)) + \frac{1}{\eta} \log K, \quad \forall k \in \mathscr{K},$$
where $\gamma_t(k)$ denotes the prediction of EXPERT-k for the t-th round.

4. (10 points) Let us modify the behavior of STOCHASTIC EXPERT in the protocol as follows. Let $\alpha \in [0,1]$.
   - Before the first round, STOCHASTIC EXPERT chooses some $k_1 \in \mathscr{K}$ following the uniform distribution.
   - For every $t \in \mathbb{N}$, STOCHASTIC EXPERT announces $\xi_t = \Phi_t(k_t)$; then, independent of the history, with probability $(1-\alpha)$ STOCHASTIC EXPERT sets $k_{t+1} = k_t$, and with probability α STOCHASTIC EXPERT sets $k_{t+1} \in \mathscr{K} \setminus \{k_t\}$ randomly following the uniform distribution.

   Show how this setting can be equivalently formulated as learning with a time-invariant stochastic expert, with an enlarged pool set $\tilde{\mathscr{K}}$ of cardinality K^T and another probability distribution $\tilde{\pi}$ on $\tilde{\mathscr{K}}$.

5. (10 points) Consider the standard formulation of learning with expert advice (where there does not exist a stochastic expert) with the same outcome space Ω, prediction space Γ, and loss function λ. Suppose there are K experts and we use the uniform distribution as the prior to run the aggregating algorithm. Let $(k_t)_{1 \leq t \leq T}$ be an arbitrary piecewise constant sequence of integers in $\mathscr{K}$; let m be the number of changes in the sequence. **Use the results above to show for any $\alpha \in [0,1]$, there is an algorithm (ignoring its computational complexity) achieving the *shifting regret bound*:**
$$\sum_{t=1}^{T} \lambda(\omega_t, \gamma_t) \leq \sum_{t=1}^{T} \lambda(\omega_t, \gamma_t(k_t)) + \frac{1}{\eta}\left[\log K + m \log(K-1) + m \log \frac{1}{\alpha} + (T-m-1) \log \frac{1}{1-\alpha}\right].$$

6. (Self study, 0 points) **Show the algorithm for the shifting regret bound above is indeed equivalent to Fixed Share.**

Solution. See Section 3.1 & 3.3 of the paper “Derandomizing stochastic prediction strategies” by V. Vovk.

References

[1] STOLTZ, G. *Information incomplète et regret interne en prédiction de suites individuelles.* PhD thesis, Université Paris-XI, 2005.