Homework 1

(Revised on September 9)

6501 Online Optimization and Learning in Games (Fall 2026)
Deadline: 11:59pm, September 27, 2026
Total points: 80

A LaTeX template is available at `https://www.overleaf.com/read/dxwndthbvgjb#f24e32`.

1 Adaptive learning rates for exponential weights

Consider the experts problem with $N \geq 2$ experts and loss vectors $\ell_t \in [0,1]^N$. At round t, the learner chooses $p_t \in \Delta_N$, observes ℓ_t, and suffers loss $\langle p_t, \ell_t \rangle$. Let $L_t(i) = \sum_{s=1}^{t} \ell_s(i)$ denote the cumulative loss of expert i through round t, with $L_0(i) = 0$. Define Regret $= \sum_{t=1}^{T} \langle p_t, \ell_t \rangle - \min_{i \in [N]} L_T(i)$. We seek to adapt to the squared losses without knowing their total or T.

Initialize $p_1(i) = 1/N$ and set $\eta_0 = 1$. After observing ℓ_t, set

$$\eta_t = \sqrt{\frac{\ln N}{\ln N + \sum_{s=1}^{t} \sum_{i=1}^{N} p_s(i) \ell_s(i)^2}}$$

and use the exponential-weights update with time-varying temperature:

$$p_{t+1}(i) = \frac{\exp\bigl(-\eta_t L_t(i)\bigr)}{\sum_{j=1}^{N} \exp\bigl(-\eta_t L_t(j)\bigr)}.$$

For $L_t = (L_t(1), \ldots, L_t(N))$, define the potential

$$\Phi_t(\lambda) = \frac{1}{\lambda} \ln \left(\frac{1}{N} \sum_{i=1}^{N} \exp\bigl(-\lambda L_t(i)\bigr) \right), \qquad \text{for } \lambda > 0.$$

(a) **(5 points)** Prove the one-step inequality

$$\langle p_t, \ell_t \rangle \leq \Phi_{t-1}(\eta_{t-1}) - \Phi_t(\eta_{t-1}) + \eta_{t-1} \sum_{i=1}^{N} p_t(i) \ell_t(i)^2.$$

Then show that

$$\text{Regret} \leq \frac{\ln N}{\eta_T} + \sum_{t=1}^{T} \eta_{t-1} \sum_{i=1}^{N} p_t(i) \ell_t(i)^2 + \sum_{t=1}^{T} \bigl(\Phi_t(\eta_t) - \Phi_t(\eta_{t-1}) \bigr).$$

(Hint: Follow the proof on pages 8–9 here.)

(b) **(5 points)** Prove that $\Phi_t(\lambda)$ is nondecreasing in $\lambda > 0$. Deduce that the final sum in part (a) is nonpositive. *(Hint: You may first prove $\Phi_t(\lambda) = -\min_{p \in \Delta_N} \bigl\{ \langle p, L_t \rangle + \frac{1}{\lambda} \mathrm{KL}(p, \mathbf{1}/N) \bigr\}$, where $\mathbf{1}/N$ is the uniform distribution on $[N]$.)*

(c) **(5 points)** Combine parts (a)–(b) with the specified learning-rate schedule to prove that, for every $T \geq 1$,

$$\text{Regret} \leq C_0 \left(\ln N + \sqrt{(\ln N) \sum_{t=1}^{T} \sum_{i=1}^{N} p_t(i) \ell_t(i)^2} \right),$$

where C_0 is a universal constant.

2 Comparator-dependent regret bounds

Consider the experts problem with $N \geq 2$ experts and loss vectors $\ell_t \in [0,1]^N$. At each round t, the learner chooses $p_t \in \Delta_N$, observes ℓ_t, and suffers loss $\langle p_t, \ell_t \rangle$. As in Problem 1, let $L_t(i) = \sum_{s=1}^{t} \ell_s(i)$ denote the cumulative loss of expert i through round t, with $L_0(i) = 0$. For each expert $i \in [N] = \{1, \ldots, N\}$, define the instantaneous regret compared to expert i as

$$r_t(i) = \langle p_t, \ell_t \rangle - \ell_t(i).$$

The cumulative regret compared to expert i is

$$\text{Regret}(i) = \sum_{t=1}^{T} r_t(i) = \sum_{t=1}^{T} \langle p_t, \ell_t \rangle - L_T(i).$$

Thus, $\text{Regret}(i)$ is the learner's total loss minus the total loss of expert i, while $\text{Regret} = \max_{i \in [N]} \text{Regret}(i)$ is the regret compared to the best fixed expert in hindsight. In both parts, initialize $p_1(i) = 1/N$.

(a) **(5 points) (Refined analysis for exponential weights)** Recall that exponential weights uses

$$p_{t+1}(i) = \frac{\exp\big(-\eta L_t(i)\big)}{\sum_{j=1}^{N} \exp\big(-\eta L_t(j)\big)}.$$

For any fixed $0 < \eta \leq 1$, prove that this algorithm ensures

$$\text{Regret} \leq \frac{\ln N}{\eta} + \eta \sum_{t=1}^{T} \sum_{i=1}^{N} p_t(i) r_t(i)^2.$$

(Hint: Does subtracting the same scalar from every expert's loss change the exponential-weights distributions? Apply the analysis from class to suitably shifted losses.)

(b) **(10 points) (A comparator-dependent bound)** The second-order term in the regret bound in part (a) averages over the learner's distribution. We now seek a bound depending only on the deviations of the comparator expert. Define

$$C_t(i) = \sum_{s=1}^{t} r_s(i)^2, \qquad C_0(i) = 0.$$

For any fixed $0 < \eta \leq \frac{1}{2}$, consider the update

$$p_{t+1}(i) = \frac{\exp\big(-\eta L_t(i) - \eta^2 C_t(i)\big)}{\sum_{j=1}^{N} \exp\big(-\eta L_t(j) - \eta^2 C_t(j)\big)}.$$

Prove that this algorithm ensures, simultaneously for all $i \in [N]$,

$$\text{Regret}(i) \leq \frac{\ln N}{\eta} + \eta \sum_{t=1}^{T} r_t(i)^2.$$

(Hint: You may use $\exp(-x - x^2) \leq 1 - x$ *for* $|x| \leq \frac{1}{2}$*.)*

Remark. These results prepare for topics to be discussed later in the course. Combining Problem 2(a) with the learning-rate scheduling idea in Problem 1 gives a bound of order

$$\text{Regret} \lesssim \ln N + \sqrt{(\ln N) \sum_{t=1}^{T} \sum_{i=1}^{N} p_t(i) r_t(i)^2}.$$

For Problem 2(b), however, it is unclear how a single learning-rate schedule can give the comparator-dependent bound of

$$\text{Regret}(i) \lesssim \ln N + \sqrt{(\ln N) \sum_{t=1}^{T} r_t(i)^2} \tag{1}$$

for all experts i simultaneously, because the best learning rate suggested by the bound can differ across comparators. It turns out that with a more involved algorithm design, (1) is indeed achievable for all i simultaneously, up to additional logarithmic factors. Such guarantees will be crucial for the adaptive regret bounds we discuss later in the course.

3 Online Mirror Descent

Let $\mathcal{X} \subseteq \mathbb{R}^d$ be a nonempty closed convex set, and let each f_t be convex and differentiable. At round t, the learner chooses $x_t \in \mathcal{X}$, suffers loss $f_t(x_t)$, and observes $g_t = \nabla f_t(x_t)$. For a convex function h differentiable at x, define its Bregman divergence by

$$D_h(u, x) = h(u) - h(x) - \langle \nabla h(x), u - x \rangle.$$

Given a convex regularizer ψ and a fixed learning rate $\eta > 0$, Online Mirror Descent (OMD) uses

$$x_{t+1} \in \operatorname*{argmin}_{x \in \mathcal{X}} \left\{ \langle x, g_t \rangle + \frac{1}{\eta} D_\psi(x, x_t) \right\}. \tag{2}$$

Assume $x_1 \in \mathcal{X}$, all updates attain their minima, and ψ is differentiable at every iterate.

(a) **(3 points)** Let $F : \mathcal{X} \to \mathbb{R}$ be convex and differentiable, and let $x^\star \in \operatorname{argmin}_{x \in \mathcal{X}} F(x)$. Prove that for every $u \in \mathcal{X}$,

$$F(x^\star) \le F(u) - D_F(u, x^\star).$$

(b) **(7 points)** Apply part (a) to the OMD objective (2) to show that, for every $u \in \mathcal{X}$,

$$\eta \langle x_{t+1} - u, g_t \rangle \le D_\psi(u, x_t) - D_\psi(u, x_{t+1}) - D_\psi(x_{t+1}, x_t). \tag{3}$$

Derive the following bound on the regret compared to u:

$$\begin{aligned} \sum_{t=1}^{T} \big(f_t(x_t) - f_t(u) \big) &\le \sum_{t=1}^{T} \langle x_t - u, g_t \rangle \\ &\le \frac{D_\psi(u, x_1)}{\eta} + \sum_{t=1}^{T} \left(\langle x_t - x_{t+1}, g_t \rangle - \frac{D_\psi(x_{t+1}, x_t)}{\eta} \right). \end{aligned} \tag{4}$$

(c) **(5 points)** Suppose ψ is μ-strongly convex for some $\mu > 0$ with respect to a norm $\|\cdot\|$, and $\|\cdot\|_\star$ is its dual norm. Use (4) to establish, for every $u \in \mathcal{X}$,

$$\sum_{t=1}^{T} \big(f_t(x_t) - f_t(u) \big) \le \frac{D_\psi(u, x_1)}{\eta} + \frac{\eta}{2\mu} \sum_{t=1}^{T} \|g_t\|_\star^2. \tag{5}$$

(d) **(5 points)** Specialize to the expert setting where $\mathcal{X} = \Delta_N$ and $f_t(x) = \langle x, \ell_t \rangle$. Initialize $x_1(i) = 1/N$ and take the negative-entropy regularizer

$$\psi(x) = \sum_{i=1}^{N} x(i) \ln x(i).$$

Verify that OMD with this regularizer gives the exponential-weights update $x_{t+1}(i) \propto x_t(i) \exp(-\eta \ell_t(i))$.

(e) **(5 points)** Let $\psi(x) = \frac{1}{2}\|x\|_2^2$. Verify that OMD with this regularizer recovers projected gradient descent:

$$x_{t+1} = \Pi_{\mathcal{X}}(x_t - \eta g_t), \qquad \Pi_{\mathcal{X}}(y) = \operatorname*{argmin}_{x \in \mathcal{X}} \|x - y\|_2^2.$$

Here $\Pi_{\mathcal{X}}$ denotes Euclidean projection onto $\mathcal{X}$. Suppose $\|g_t\|_2 \le G$ for all t, and let $D = \sup_{x,u \in \mathcal{X}} \|x - u\|_2 < \infty$ be the diameter of $\mathcal{X}$. Use (5) and choose η to obtain Regret $:= \sum_{t=1}^{T} f_t(x_t) - \min_{u \in \mathcal{X}} \sum_{t=1}^{T} f_t(u) \le DG\sqrt{T}$. You may assume $D, G > 0$; the zero cases are trivial.

4 FTPL with Gaussian perturbations

Consider online linear optimization over a nonempty compact convex set $\mathcal{X} \subseteq \mathbb{R}^d$. Let $\|\cdot\|$ be a norm with dual norm $\|\cdot\|_\star$, and let $R = \sup_{x,y\in\mathcal{X}} \|x - y\|$ be the diameter of $\mathcal{X}$. We consider an *oblivious adversary*: the loss vectors $\ell_1, \ldots, \ell_T \in \mathbb{R}^d$ are fixed before any interaction with the learner. All expectations below are over the learner's perturbations, with the loss sequence held fixed. At round t, the learner chooses $x_t \in \mathcal{X}$, suffers loss $\langle x_t, \ell_t\rangle$, and observes ℓ_t. Write $L_t = \sum_{s=1}^t \ell_s$, with $L_0 = 0$.

For a fixed $\eta > 0$, Follow the Perturbed Leader (FTPL) draws an independent $Z_t \sim \mathcal{D}$ on each round and chooses

$$x_t = \operatorname*{argmin}_{x\in\mathcal{X}} \left\{ \left\langle x, \sum_{s=1}^{t-1} \ell_s + \frac{Z_t}{\eta} \right\rangle \right\},$$

with a fixed tie-breaking rule. Define x_{T+1} by the same update using L_T and an independent $Z_{T+1} \sim \mathcal{D}$. Assume $\mathcal{D}$ has a density and $\mathbb{E}_{Z\sim\mathcal{D}}[\|Z\|_\star] < \infty$. Define

$$\text{Regret} = \sum_{t=1}^T \langle x_t, \ell_t\rangle - \min_{x\in\mathcal{X}} \sum_{t=1}^T \langle x, \ell_t\rangle.$$

For distributions P, Q with densities p, q, define their total-variation distance:

$$\mathrm{TV}(P, Q) = \frac{1}{2} \int_{\mathbb{R}^d} |p(x) - q(x)|\, dx.$$

Assume the distribution $\mathcal{D}$ is κ-smooth:

$$\mathrm{TV}(\mathcal{D}, \mathcal{D} + u) \le \kappa \|u\|_\star \qquad \text{for every } u \in \mathbb{R}^d, \tag{6}$$

where $\mathcal{D} + u$ denotes the distribution of $Z + u$ for $Z \sim \mathcal{D}$.

(a) **(5 points)** Define the potential

$$\Phi_t = \mathbb{E}_{Z\sim\mathcal{D}} \left[\min_{x\in\mathcal{X}} \left\langle x, \sum_{s=1}^t \ell_s + \frac{Z}{\eta} \right\rangle \right], \qquad t = 0, \ldots, T,$$

where the sum is zero when $t = 0$. Prove the one-step inequality

$$\mathbb{E}[\langle x_t, \ell_t\rangle] \le \Phi_t - \Phi_{t-1} + \mathbb{E}[\langle x_t - x_{t+1}, \ell_t\rangle].$$

(Hint: Similar to the proof of the "Be-The-Leader Lemma" discussed in class.)

(b) **(5 points)** Sum the inequality in part (a) and bound the endpoint potentials to prove

$$\mathbb{E}[\text{Regret}] \le \frac{R\mathbb{E}_{Z\sim\mathcal{D}}[\|Z\|_\star]}{\eta} + \sum_{t=1}^T \mathbb{E}[\langle x_t - x_{t+1}, \ell_t\rangle].$$

(Hint: Compare Φ_T with the loss of the best fixed action.)

(c) **(5 points)** Let $p : \mathbb{R}^d \to [0, \infty)$ be the density of $\mathcal{D}$. Define $h : \mathbb{R}^d \to \mathcal{X}$ by

$$h(L) = \operatorname*{argmin}_{x\in\mathcal{X}} \langle x, L\rangle,$$

using the same tie-breaking rule as FTPL. Show that

$$\mathbb{E}[\langle x_t - x_{t+1}, \ell_t\rangle] = \int_{\mathbb{R}^d} (p(z) - p(z - \eta\ell_t)) \left\langle h\left(L_{t-1} + \frac{1}{\eta} z\right), \ell_t \right\rangle dz.$$

(Hint: By the decision rule, we can write $x_t = h\left(L_{t-1} + \frac{Z_t}{\eta}\right)$ and similarly for x_{t+1}. Express the expectation of $\langle x_t - x_{t+1}, \ell_t\rangle$ using the density function of Z_t and Z_{t+1}.)

(d) **(5 points)** Use (6) and part (c) to show that

$$\mathbb{E}[\langle x_t - x_{t+1}, \ell_t \rangle] \leq \eta \kappa R \|\ell_t\|_\star^2.$$

Combine this with part (b) to conclude

$$\mathbb{E}[\text{Regret}] \leq \frac{R\mathbb{E}_{Z\sim\mathcal{D}}[\|Z\|_\star]}{\eta} + \eta\kappa R \sum_{t=1}^{T} \|\ell_t\|_\star^2.$$

(e) **(5 points)** Specialize to the Euclidean norm and standard Gaussian perturbations $\mathcal{D} = \mathcal{N}(0, I_d)$. Show that

$$\mathrm{TV}(\mathcal{D}, \mathcal{D} + u) \leq \frac{1}{2}\|u\|_2.$$

You may directly use $\mathbb{E}_{Z\sim\mathcal{D}}[\|Z\|_2] \leq \sqrt{d}$ without proof. Suppose $\|\ell_t\|_2 \leq G$ for all t, where $G > 0$. Apply part (d) to obtain

$$\mathbb{E}[\text{Regret}] \leq \frac{R\sqrt{d}}{\eta} + \frac{\eta R G^2 T}{2},$$

and choose η to conclude $\mathbb{E}[\text{Regret}] = \mathcal{O}(RGd^{1/4}\sqrt{T})$.

5 Survey

Leave any feedback for the course.