Homework 3

6501 Online Optimization and Learning in Games (Fall 2026)

Deadline: 11:59pm, November 12, 2026
Total points: 60

A LaTeX template is available at `https://www.overleaf.com/project/6ae965720502953383c3cefe/share#779c3730d081b7a20b162024eeae76ac6339aaaa7b4d1a57`.

1 A reduction from no-swap-regret to no-regret (Blum and Mansour, 2007)

This problem uses an approach different from the one discussed in class to develop a no-swap-regret algorithm. Consider online learning with $N \geq 2$ actions and loss vectors $\ell_t \in [0,1]^N$. For distributions $p_1, \ldots, p_T \in \Delta_N$, define the mixed-action swap regret by

$$\mathrm{SwapRegret} = \max_{\phi:[N]\to[N]} \sum_{t=1}^{T} \sum_{i=1}^{N} p_t(i)\big(\ell_t(i) - \ell_t(\phi(i))\big).$$

Consider the following algorithm: Run N no-regret algorithms $\mathcal{A}_1, \ldots, \mathcal{A}_N$. On round t, algorithm $\mathcal{A}_i$ outputs $q_t^{(i)} \in \Delta_N$. Form the transition matrix Q_t with entries $Q_t(i,j) = q_t^{(i)}(j)$, and choose $p_t \in \Delta_N$ satisfying

$$Q_t^\top p_t = p_t.$$

Since Q_t is a transition matrix, such p_t always exists.

After observing ℓ_t, feed the loss vector $g_t^{(i)} = p_t(i)\ell_t \in \mathbb{R}^N$ to $\mathcal{A}_i$.

(a) **(5 points)** Define the external regret of $\mathcal{A}_i$ on its supplied losses by

$$\mathrm{Regret}_i = \sum_{t=1}^{T} \langle q_t^{(i)}, g_t^{(i)} \rangle - \min_{j\in[N]} \sum_{t=1}^{T} g_t^{(i)}(j).$$

Prove that $\mathrm{SwapRegret} = \sum_{i=1}^{N} \mathrm{Regret}_i$.

(b) **(5 points)** Suppose each base algorithm $\mathcal{A}_i$ is exponential weights with learning rate $0 < \eta \leq 1$. Show that

$$\mathrm{SwapRegret} \leq \frac{N \log N}{\eta} + \eta T,$$

and choose η to obtain $\mathrm{SwapRegret} \leq O\left(\sqrt{NT \log N}\right)$. *Hint:* Use the regret bound of exponential weights on Page 7 here.

2 Groupwise swap regret

In class, we discussed groupwise regret and swap regret individually. In this problem, we will combine their ideas to design algorithms with vanishing *groupwise swap regret*, i.e., the swap regret on every group is $o(T)$.

Consider N actions and K possibly overlapping groups. On round t, observe memberships $c_t \in \{0,1\}^K$, choose $p_t \in \Delta_N$, then observe $\ell_t \in [0,1]^N$ and incur loss $\langle p_t, \ell_t \rangle$. Define the swap regret on group k by

$$\text{SwapRegret}_k = \max_{\phi:[N]\to[N]} \sum_{t=1}^{T} c_t(k) \sum_{i=1}^{N} p_t(i)\big(\ell_t(i) - \ell_t(\phi(i))\big).$$

Below, we use the no-regret robust-response framework from the multi-objective online learning lecture to design an algorithm controlling every group's swap regret simultaneously.

(a) **(5 points) Define the game.** For any $c \in \{0,1\}^K$ and $\ell \in [0,1]^N$, construct a game matrix $G_{c,\ell} \in \mathbb{R}^{N \times N^2 K}$ such that

$$\max_{q \in \Delta_{N^2K}} \sum_{t=1}^{T} p_t^\top G_t q \le o(T) \quad \Longrightarrow \quad \text{SwapRegret}_k = o(T) \text{ for every } k \in [K]$$

where $G_t \triangleq G_{c_t,\ell_t}$.

Hint: Index the entries of $G_{c,\ell}$ with index $(s,(i,j,k))$, where $s,i,j \in [N]$ and $k \in [K]$.

(b) **(10 points) Find a robust response.** For every $q \in \Delta_{N^2K}$ and every $c \in \{0,1\}^K$, construct a $p \in \Delta_N$, depending only on q and c, such that

$$p^\top G_{c,\ell} q \le 0 \qquad \text{for every } \ell \in [0,1]^N.$$

(c) **(5 points) Give the algorithm and regret bound.** Using parts (a) and (b), write down pseudocode for the full algorithm that achieves $o(T)$ groupwise swap regret, using exponential weights with learning rate η for the q-player. Derive a groupwise swap-regret bound for every group in terms of N, K, T, η.

3 Equilibrium

Consider a normal-form general-sum game, with action set $\mathcal{A}_i$ for player i. Recall that a distribution σ over joint actions is a coarse correlated equilibrium (CCE) if, for every player i and fixed action b_i,

$$\sum_a \sigma(a)\big(u_i(a) - u_i(b_i, a_{-i})\big) \geq 0.$$

It is a correlated equilibrium (CE) if, for every player i and pair of actions a_i, b_i,

$$\sum_{a_{-i}} \sigma(a_i, a_{-i})\big(u_i(a_i, a_{-i}) - u_i(b_i, a_{-i})\big) \geq 0.$$

A Nash equilibrium (NE) is a CE with product distribution $\sigma = p_1 \times \ldots \times p_N$.

(a) **(5 points)** Show that any CE is a CCE.

(b) **(5 points)** Show that if every player has *at most two* actions, then any CCE is a CE.

(c) **(5 points)** An action a_i is *strictly dominated* if there exists a mixed strategy $p_i \in \Delta_{\mathcal{A}_i}$ for player i such that

$$\mathbb{E}_{b_i \sim p_i}[u_i(b_i, a_{-i})] > u_i(a_i, a_{-i}) \qquad \text{for every } a_{-i}.$$

In other words, p_i yields strictly higher expected utility than a_i regardless of the other players' actions (for example, "Cooperate" in the Prisoner's Dilemma is strictly dominated by "Defect"). Show that any CE assigns zero probability to every strictly dominated action.

(d) **(5 points)** Continue from (c). CCE need not have the same property. Consider the following two-player game general-sum game:

	L	R
A	$(0,0)$	$(0,0)$
B	$(1,1)$	$(1,1)$
C	$(0,0)$	$(2,2)$

Let $\sigma(A, L) = \sigma(C, R) = 1/2$, with all other joint actions having probability zero. Show that σ is a CCE but assigns positive probability to a strictly dominated action.

(e) **(5 points)** Consider a two-player zero-sum game in which the row player's payoff matrix is G and the column player's payoff matrix is $-G$. Its game value is

$$v = \max_p \min_q p^\top G q = \min_q \max_p p^\top G q.$$

Let σ be a CCE of the game G, and let

$$w = \sum_{i,j} \sigma(i,j) G(i,j).$$

Prove that $w = v$.

(f) **(5 points)** Continue from (e). Let p and q be the marginal distributions of σ — that is, $p(i) = \sum_j \sigma(i,j)$ and $q(j) = \sum_i \sigma(i,j)$. Show that (p, q) is a NE.

4 Survey

Leave any feedback for the course.

References

Avrim Blum and Yishay Mansour. *From External to Internal Regret*. JMLR 8, 2007.
https://www.jmlr.org/papers/volume8/blum07a/blum07a.pdf.