CS364A: Exercise Set #9

Due by noon on Friday, December 6, 2013

Instructions:

(1) Turn in your solutions to all of the following exercises directly to one of the TAs (Kostas or Okke). Please type your solutions if possible and feel free to use the LaTeX template provided on the course home page. Email your solutions to `cs364a-aut1314-submissions@cs.stanford.edu`. If you prefer to hand-write your solutions, you can give it to one of the TAs in person at the start of the lecture.

(2) Your solutions will be graded on a "check/minus/zero" system, with "check" meaning satisfactory and "minus" meaning needs improvement.

(3) Solve these exercises and write up your solutions on your own. You may, however, discuss the exercises verbally at a high level with other students. You may also consult any books, papers, and Internet resources that you wish. And of course, you are encouraged to contact the course staff (via Piazza or office hours) to clarify the questions and the course material.

(4) No late assignments will be accepted.

Lecture 17 Exercises

Exercise 72

Prove that for every $\epsilon \in (0, \frac{1}{2}]$ and $x \in [0, 1]$,

$$(1 - \epsilon)^x \leq (1 - \epsilon x)$$

and

$$(1 + \epsilon)^x \leq (1 + \epsilon x).$$

[Hint: Here's one way you could structure your proof. For fixed ϵ, let $f_\epsilon(x)$ and $g_\epsilon(x)$ denote the functions on the left- and right-hand sides. Verify that $f_\epsilon(0) = g_\epsilon(0)$, $f_\epsilon(1) = g_\epsilon(1)$, $f'_\epsilon(0) < g'_\epsilon(0)$, and f is convex.]

Exercise 73

The multiplicative weights algorithm, as described in lecture, requires knowledge of the time horizon T to set the parameter ϵ. Modify the algorithm so that it doesn't need to know T in advance and continues to have time-averaged external regret $O(\sqrt{\ln n}/\sqrt{T})$.

[Hint: There is no need to redo the previous analysis from scratch. For example, you could consider restarting the algorithm each time to get to a day t of the form 2^i.]

Exercise 74

Suppose we modify the multiplicative weights algorithm so that the weight update step is $w^{t+1}(a) := w^t(a) \cdot (1 - \epsilon c^t(a))$ (rather than $w^{t+1}(a) := w^t(a) \cdot (1 - \epsilon)^{c^t(a)}$). Show that this algorithm also has regret $O(\sqrt{\ln n}/\sqrt{T})$.

[Hint: This time you should redo the analysis from class, making minor modifications as necessary.]

Exercise 75

Modify the online decision-making setting so that every time step t an adversary chooses a *payoff vector* $\pi^t : A \to [0,1]$ rather than a cost vector. The time-averaged external regret is now defined as $\frac{1}{T}\max_{a\in A}\sum_{t=1}^{T}\pi^t(a) - \frac{1}{T}\sum_{t=1}^{T}\pi^t(a^t)$.

Suppose we modify the multiplicative weights algorithm so that the weight update step is $w^{t+1}(a) := w^t(a)\cdot(1+\epsilon)^{\pi^t(a)}$ (rather than $w^{t+1}(a) := w^t(a)\cdot(1-\epsilon)^{c^t(a)}$). Show that this algorithm has regret $O(\sqrt{\ln n}/\sqrt{T})$.

[Hint: Redo the analysis from class, making minor modifications as necessary. Use Exercise 72.]

Exercise 76

We assumed in class that costs are bounded between 0 and 1. Suppose instead they are bounded between 0 and $c_{\max}$. Explain how to make minor modifications to the multiplicative weights algorithm so that it has time-averaged regret $O(c_{\max}\cdot\sqrt{\ln n}/\sqrt{T})$. (You can assume that T is known.) Note this implies $O(\frac{c_{\max}^2 \ln n}{\epsilon^2})$ steps to achieve regret ϵ.

Exercise 77

Consider a (λ,μ)-smooth cost-minimization game (see Lecture 14). Let $\mathbf{s}^1,\ldots,\mathbf{s}^T$ be an outcome sequence such that player i has time-averaged external regret R_i. Let σ denote the uniform distribution over this multi-set of T outcomes. Prove that the expected cost of σ is at most $\frac{\lambda}{1-\mu}$ times the cost of an optimal outcome, plus the error term $\frac{1}{1-\mu}\sum_{i=1}^{n} R_i$.

Exercise 78

Consider T iterations of no-regret dynamics in a k-player game. Let p_i^t denote the mixed strategy played by player i in iteration t, and $\sigma^t = \prod_{i=1}^{k} p_i^t$ the corresponding distribution over outcomes on day t.

In lecture we proved that the time-averaged joint distribution $\frac{1}{T}\sum_{t=1}^{T}\sigma^t$ is an approximate coarse correlated equilibrium, but did not claim anything about any individual distribution σ^t.

Prove that an individual distribution σ^t is an approximate coarse correlated equilibrium if and only if it is an approximate Nash equilibrium (with the same error term).

Lecture 18 Exercises

Exercise 79

Suppose we instantiate the Blum-Mansour reduction using n copies of the multiplicative weights algorithm, where n is the number of actions. What is the swap regret of the resulting "master" algorithm, as a function of n and T?

Exercise 80

Let A denote the matrix of row player payoffs of a two-player zero-sum game. Prove that a pair of $\hat{x},\hat{y}$ of mixed distributions is a mixed Nash equilibrium if and only if it is a *minimax pair*, meaning

$$\hat{x} \in \underset{x}{\operatorname{argmax}}\left(\min_y x^T A y\right)$$

and

$$\hat{y} \in \underset{y}{\operatorname{argmin}}\left(\max_x x^T A y\right).$$

Exercise 81

Prove that if (x_1, y_1) and (x_2, y_2) are mixed Nash equilibria of a two-player zero-sum game, then so are (x_1, y_2) and (x_2, y_1).