Mehryar Mohri
Advanced Machine Learning 2016
Courant Institute of Mathematical Sciences
Homework assignment 2
April 11, 2016
Due: April 25, 2016

A. Non-stationary sequences.

The regret guarantees we discussed in class were typically with respect to the best fixed expert in hindsight. In many situations, any fixed expert admits a large loss. For example, if the problem consisted of predicting the US electricity usage out of N bins over a decade, no single bin value would be accurate since the power usage admits some seasonality with higher levels during July and August, or January and February. This leads us to considering instead non-stationary reference expert sequences $(i_1, \ldots, i_T)$, $i_t \in [1, N]$, with at most k changes, that is k indices $t \in [1, T]$ with $i_t \neq i_{t+1}$.

1. What is the total number M of such expert sequences?

2. Assume that $k \leq T/2$. Give a randomized on-line algorithm for this problem (you can assume that the algorithm knows T) and a bound on its expected regret expressed in terms of T and k (*hint*: you can express the bound in terms of the binary entropy function H using the inequality $\sum_{i=0}^{k} \binom{T}{i} \leq e^{TH(\frac{k}{T})}$ for $k \leq T/2$, where $H(p) = -p\log(p) - (1-p)\log(1-p)$).

3. What is the computational complexity of your algorithm? Can you suggest a way to improve it?

B. Mirror descent.

Let $\mathcal{X}$ and $\mathcal{Y}$ be compact and convex sets and $\Psi\colon \mathcal{X} \times \mathcal{Y} \to \mathbb{R}$ a function such that $\Psi(\cdot, y)$ is convex for any $y \in \mathcal{Y}$ and $\Psi(x, \cdot)$ concave for any $x \in \mathcal{X}$. Then, by the generalized von Neumann's theorem, there exists $(x^*, y^*) \in \mathcal{X} \times \mathcal{Y}$ such that

$$\Psi(x^*, y^*) = \inf_{x \in \mathcal{X}} \sup_{y \in \mathcal{Y}} \Psi(x, y) = \sup_{y \in \mathcal{Y}} \inf_{x \in \mathcal{X}} \Psi(x, y).$$

We seek an approximate solution (x, y) to (x^*, y^*) whose quality is measured by $\max_{y' \in \mathcal{Y}} \Psi(x, y') - \min_{x' \in \mathcal{X}} \Psi(x', y)$. For any x, we denote by $\delta\Psi_{\mathcal{X}}(x, y)$ an element of the subgradient of $\Psi(\cdot, y)$ at x and similarly for $\delta\Psi_{\mathcal{Y}}(x, y)$.

1. Prove that for any $(x, y) \in \mathcal{X} \times \mathcal{Y}$ there exists $(x', y') \in \mathcal{X} \times \mathcal{Y}$ such that
$$\max_{y' \in \mathcal{Y}} \Psi(x, y') - \min_{x' \in \mathcal{X}} \Psi(x', y) \leq \begin{bmatrix} \delta\Psi_{\mathcal{X}}(x, y) \\ -\delta\Psi_{\mathcal{Y}}(x, y) \end{bmatrix} \cdot \begin{bmatrix} x - x' \\ y - y' \end{bmatrix}.$$

2. Let $\Phi_{\mathcal{X}} \colon C_{\mathcal{X}} \to \mathbb{R}$ be a Legendre type function that is $\alpha_{\mathcal{X}}$-strongly convex with respect to a norm $\|\cdot\|_{\mathcal{X}}$, with $C_{\mathcal{X}}$ an open set containing $\mathcal{X}$ and similarly with $\Phi_{\mathcal{Y}} \colon C_{\mathcal{Y}} \to \mathbb{R}$ $\alpha_{\mathcal{Y}}$-strongly convex with respect to a norm $\|\cdot\|_{\mathcal{Y}}$. Assume that $\Psi_{\mathcal{X}}(\cdot, y)$ is $L_{\mathcal{X}}$-Lipschitz with respect to $\|\cdot\|_{\mathcal{X}}$, for any y, and similarly that $\Psi_{\mathcal{X}}(x, \cdot)$ is $L_{\mathcal{Y}}$-Lipschitz with respect to $\|\cdot\|_{\mathcal{Y}}$, for any x. Use the inequality of the previous question to derive a mirror descent solution for this problem.

3. Give a convergence guarantee for your algorithm in terms of $r_{\mathcal{X}}^2 = \sup_{x \in \mathcal{X}} \Phi_{\mathcal{X}}(x) - \Phi_{\mathcal{X}}(x_1)$, a similar quantity $r_{\mathcal{Y}}^2$ defined for $\mathcal{Y}$, $\alpha_{\mathcal{X}}$ and $\alpha_{\mathcal{Y}}$.

C. Continuous bandit.

Consider the problem where at each round $t \in [1, T]$, the player selects an *action* $x_t \in [0, 1]$ and incurs a loss $f_t(x_t)$ where $f_t \colon [0, 1] \to \mathbb{R}$ is L-Lipschitz.

1. Give an upper bound on the pseudo-regret of the algorithm that consists of running EXP3 with the action set $\mathcal{A} = \left\{\frac{1}{K}, \frac{2}{K}, \ldots, 1\right\}$ in terms of K, T, and L (note: here, the pseudo-regret of this algorithm is with respect to the best fixed action in hindsight in $[0, 1]$).

2. Choose $K = \left\lceil (\frac{T}{\log T})^{1/3} \right\rceil$ and give a big-O upper bound on the pseudo-regret of the algorithm.