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?

   Solution: For each $j \in [0, k]$, there are $\binom{T-1}{j}$ ways of choosing a change time t with $i_t \neq i_{t+1}$ since t is in $[1, T-1]$. There are N possible ways of choosing an expert in the first part of sequence, $N-1$ ways of choosing a different one in the next and similarly for other parts. Thus, the total number of ways of assigning an expert to each such sequence is $N(N-1)^j$. The total number of such experts is therefore

   $$M = \sum_{j=0}^{k} \binom{T-1}{j} N(N-1)^j. \qquad \square$$

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 $H(p) = -p\log(p) - (1-p)\log(1-p)$).

Solution: We can run RWM with the M experts. The regret of the algorithm is then

$$R_T \le 2\sqrt{T \log M}.$$

We can bound M as follows:

$$M = \sum_{j=0}^{k} \binom{T-1}{j} N(N-1)^j \le \sum_{j=0}^{k} \binom{T-1}{j} N^{k+1} \le e^{(T-1)H(\frac{k}{T-1})} N^{k+1}.$$

Thus,

$$\log M = (T-1)H\left(\frac{k}{T-1}\right) + (k+1)\log N.$$

Thus, the regret of the algorithm is bounded as follows:

$$R_T \le 2\sqrt{T\left[(k+1)\log N + (T-1)H\left(\frac{k}{T-1}\right)\right]}. \qquad \square$$

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

   Solution: The complexity of RWM run with M experts is in $O(MT)$, which is prohibitive. However, one can exploit the structure of the experts to design a more efficient algorithm. $\square$

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

$$\max_{y' \in \mathcal{Y}} \Psi(x, y') - \min_{x' \in \mathcal{X}} \Psi(x', y) \leq \begin{bmatrix} \delta\Psi_{\mathcal{X}}(x, y) \\ -\delta\Psi_{\mathcal{Y}}(x, y) \end{bmatrix} \cdot \begin{bmatrix} x - x' \\ y - y' \end{bmatrix}.$$

Solution: For any $y \in \mathcal{Y}$, let $x' = \operatorname{argmin}_{x \in \mathcal{X}} \Psi(x, y)$. Similarly, for any $x \in \mathcal{X}$, let $y' = \operatorname{argmax}_{y \in \mathcal{Y}} \Psi(x, y)$. Since for any y, $\Psi(\cdot, y)$ is a convex function of the first argument, the following inequality holds

$$\Psi(x, y) - \Psi(x', y) \leq \delta\Psi_{\mathcal{Y}}(x, y) \cdot (x - x'),$$

for any $(x, y) \in \mathcal{X} \times \mathcal{Y}$. By concavity of $\Psi(x, \cdot)$, we also have

$$\Psi(x, y') - \Psi(x, y) \leq -\delta\Psi_{\mathcal{Y}}(x, y) \cdot (y - y'),$$

for any $(x, y) \in \mathcal{X} \times \mathcal{Y}$. Therefore, we have that

$$\begin{aligned} \max_{y' \in \mathcal{Y}} \Psi(x, y') - \min_{x' \in \mathcal{X}} \Psi(x', y) &= \Psi(x, y') - \Psi(x, y) + \Psi(x, y) - \Psi(x', y) \\ &\leq \begin{bmatrix} \delta\Psi_{\mathcal{X}}(x, y) \\ -\delta\Psi_{\mathcal{Y}}(x, y) \end{bmatrix} \cdot \begin{bmatrix} x - x' \\ y - y' \end{bmatrix}, \end{aligned}$$

which concludes the proof. □

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.

Solution: We propose the following variant of the mirror descent algorithm. Starting with

$$\begin{aligned} x_1 &\leftarrow \operatorname*{argmin}_{x \in \mathcal{X}} \Phi_{\mathcal{X}}(x), \\ y_1 &\leftarrow \operatorname*{argmin}_{y \in \mathcal{Y}} \Phi_{\mathcal{Y}}(y), \end{aligned}$$

for each iteration $t = 1, \dots, T$, we let

$$x_{t+1} = \operatorname*{argmin}_{x \in \mathcal{X}} \left(\quad \delta\Psi_{\mathcal{X}}(x_t, y_t) \cdot x + \frac{1}{\eta} B_{\Phi_{\mathcal{X}}}(x \parallel x_t) \right),$$

$$y_{t+1} = \operatorname*{argmin}_{y \in \mathcal{Y}} \left(-\delta\Psi_{\mathcal{Y}}(x_t, y_t) \cdot y + \frac{1}{\eta} B_{\Phi_{\mathcal{Y}}}(y \parallel y_t) \right),$$

where B_Φ denotes the Bregman divergence associated to Φ and η is a learning rate which we specify below. The solution returned by this algorithm is $\frac{1}{T}\sum_{t=1}^{T}(x_t, y_t)$. □

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 $\alpha_{\mathcal{Y}}$.

Solution: Observe that by convexity of $\Psi_{\mathcal{X}}$ the following holds:

$$\max_{y \in \mathcal{Y}} \Psi\left(\frac{1}{T}\sum_{t=1}^{T} x_t, y\right) \leq \max_{y \in \mathcal{Y}} \frac{1}{T}\sum_{t=1}^{T} \Psi(x_t, y).$$

Similarly, by concavity of $\Psi_{\mathcal{Y}}$ the following holds:

$$-\min_{x \in \mathcal{X}} \Psi\left(x, \frac{1}{T}\sum_{t=1}^{T} y_t\right) \leq -\min_{x \in \mathcal{Y}} \frac{1}{T}\sum_{t=1}^{T} \Psi(x, y_t).$$

Therefore, combining these inequalities with part 1 leads to the following bound:

$$\begin{aligned}
&\max_{y \in \mathcal{Y}} \Psi\left(\frac{1}{T}\sum_{t=1}^{T} x_t, y\right) - \min_{x \in \mathcal{X}} \Psi\left(x, \frac{1}{T}\sum_{t=1}^{T} y_t\right) \\
&\quad \leq \max_{y \in \mathcal{Y}} \frac{1}{T}\sum_{t=1}^{T} \Psi(x_t, y) - \min_{x \in \mathcal{Y}} \frac{1}{T}\sum_{t=1}^{T} \Psi(x, y_t) \\
&\quad \leq \max_{(x,y) \in \mathcal{X} \times \mathcal{Y}} \frac{1}{T}\sum_{t=1}^{T} \Psi(x_t, y) - \Psi(x, y_t) \\
&\quad \leq \max_{(x,y) \in \mathcal{X} \times \mathcal{Y}} \left(\frac{1}{T}\sum_{t=1}^{T} \delta\Psi_{\mathcal{X}}(x_t, y_t) \cdot (x_t - x) \right. \\
&\qquad\qquad \left. - \frac{1}{T}\sum_{t=1}^{T} \delta\Psi_{\mathcal{Y}}(x_t, y_t) \cdot (y_t - y) \right).
\end{aligned}$$

Applying, the same arguments as in the proof of convergence of the mirror descent algorithm with $f_t(x) = \delta\Psi_{\mathcal{X}}(x_t, y_t) \cdot x$, we have that

$$\sum_{t=1}^{T} \delta\Psi_{\mathcal{X}}(x_t, y_t) \cdot (x_t - x) \leq \frac{r_{\mathcal{X}}^2}{\eta} + \frac{\eta L_{\mathcal{X}}^2 T}{2\alpha_{\mathcal{X}}},$$

for any $x \in \mathcal{X}$. Note that above result is using the fact that $\Phi_{\mathcal{X}}$ is a $\alpha_{\mathcal{X}}$-strongly convex Legendre type function and $\Psi_{\mathcal{X}}$ is $L_{\mathcal{X}}$-Lipschitz. Similarly,

$$\sum_{t=1}^{T} -\delta\Psi_{\mathcal{X}}(x_t, y_t) \cdot (y_t - y) \leq \frac{r_{\mathcal{Y}}^2}{\eta} + \frac{\eta L_{\mathcal{Y}}^2 T}{2\alpha_{\mathcal{Y}}},$$

for any $y \in \mathcal{Y}$. Therefore, solving for the optimal η leads to the following convergence guarantee

$$\max_{y \in \mathcal{Y}} \Psi\left(\frac{1}{T}\sum_{t=1}^{T} x_t, y\right) - \min_{x \in \mathcal{X}} \Psi\left(x, \frac{1}{T}\sum_{t=1}^{T} y_t\right) \leq \sqrt{\frac{2(r_{\mathcal{X}}^2 + r_{\mathcal{Y}}^2)(\frac{L_{\mathcal{X}}^2}{\alpha_{\mathcal{X}}} + \frac{L_{\mathcal{Y}}^2}{\alpha_{\mathcal{Y}}})}{T}},$$

which provides a convergence guarantee for the algorithm in part 2.

□

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 $[0, 1]$).

   Solution: The pseudo-regret of EXP3 reduced to $\mathcal{A}$ is bounded by $\sqrt{2KT \log K}$. For any $x \in [0, 1]$, there exists $a(x) \in \mathcal{A}$ with $|x - a(x)| \leq$

$\frac{1}{K}$. Thus, since f_ts are L-Lipschitz,

$$\begin{aligned}
\mathrm{E}\left[\sum_{t=1}^{T} f_t(x_t) - f_t(x)\right] &= \mathrm{E}\left[\sum_{t=1}^{T} f_t(x_t) - f_t(a(x)) + f_t(a(x)) - f(x)\right] \\
&\leq \mathrm{E}\left[\sum_{t=1}^{T} f_t(x_t) - f_t(a(x))\right] + \frac{LT}{K} \\
&\leq \sqrt{2KT\log K} + \frac{LT}{K}. \qquad \square
\end{aligned}$$

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

Solution: For that choice of K, the regret of the algorithm is bounded by

$$\begin{aligned}
O\left(\sqrt{KT\log K} + \frac{L}{K}\right) &= O\left(\left[\frac{T^{4/3}}{(\log T)^{1/3}}\log T\right]^{1/2} + T^{2/3}(\log T)^{1/3}\right) \\
&= O\left(\left[T^{2/3}(\log T)^{1/3}\right] + T^{2/3}(\log T)^{1/3}\right) \\
&= O\left(T^{\frac{2}{3}}(\log T)^{\frac{1}{3}}\right). \qquad \square
\end{aligned}$$