MLT Homework 8

Ana Borovac
Bas Haver

November 12, 2018

Question 1

Let $\psi(\lambda) = \frac{\lambda^2}{2}$. The Legendre-Fenchel transform of ψ is given by

$$\psi^*(\epsilon) = \sup_{\lambda \in \mathbb{R}} \lambda\epsilon - \psi(\lambda).$$

Subquestion 1.1

$\psi^*(\epsilon) = \frac{\epsilon^2}{2}$.

Solution

Let define $\psi_1(\lambda)$:

$$\psi_1(\lambda) = \lambda\epsilon - \frac{\lambda^2}{2}$$

Furthermore:

$$\psi_1'(\lambda) = \epsilon - \lambda$$

Since $\psi_1(\lambda)$ is a parabola it has only one extreme; particularly it has just a maximum (negative sign before λ^2). So, the maximum is reached at:

$$\lambda = \epsilon \quad \Rightarrow \quad \psi_1(\epsilon) = \epsilon \cdot \epsilon - \frac{\epsilon^2}{2} = \frac{\epsilon^2}{2}$$

We can conclude:

$$\psi^*(\epsilon) = \psi_1(\epsilon) = \frac{\epsilon^2}{2}$$

Subquestion 1.2

$(\psi^*)^{-1}(z) = \pm\sqrt{2z}$.

Solution

From previous point we know:

$$\psi^*(\epsilon) = \frac{\epsilon^2}{2}$$

It follows:

$$\begin{aligned} z &= \frac{\epsilon^2}{2} \\ 2z &= \epsilon^2 \\ \epsilon &= \pm\sqrt{2z} \end{aligned}$$

So:

$$(\psi^*)^{-1}(z) = \pm\sqrt{2z}$$

Question 2

The Blooper Reel

Subquestion 2.1

Deterministic fails for Adversarial Bandits *Show that any deterministic algorithm (UCB included) has linear regret in the adversarial bandit setting. Hint: you can use the argument on the top of page 23.*

Question 3

We consider an adversarial bandit model with K^2 arms indexed by $i \in [K]$ and $j \in [K]$. For each arm (i, j), the loss at time t is $a_t^i + b_t^j$, where $a_t^i \in [0, 1]$ and $b_t^j \in [0, 1]$ are chosen by the adversary before the start of the interaction. Then each round the learner picks an arm $(I_t, J_t) \in [K]^2$ and observes $a_t^{I_t}$ and $b_t^{J_t}$ separately (and incurs their sum as the loss).

Subquestion 3.1

Consider running a single instance of EXP3 on all K^2 arms (with loss range $[0, 2]$). Show that the expected pseudo-regret compared to the best arm (i^*, j^*) is bounded by

$$\bar{R}_n \leq 2\sqrt{2nK^2 \ln(K^2)}$$

Solution

Below we used the following facts:

- $\min x + y = \min x + \min y;\ \ x, y \geq 0$
- Linearity of expected value.
- Theorem from the lectures: $\bar{R}_n \leq \sqrt{2nK \ln K}$, where K in the number of arms; in our case we have K^2 arms.

$$
\begin{aligned}
\bar{R}_n &= \mathbb{E}_{I_1,\ldots,I_n,J_1,\ldots,J_n}\left\{\sum_{t=1}^{n} a_t^{I_t} + b_t^{J_t}\right\} - \min_k \sum_{t=1}^{n} a_t^k + b_t^k \\
&= \left(\mathbb{E}_{I_1,\ldots,I_n}\left\{\sum_{t=1}^{n} a_t^{I_t}\right\} - \min_k \sum_{t=1}^{n} a_t^k\right) + \left(\mathbb{E}_{J_1,\ldots,J_n}\left\{\sum_{t=1}^{n} b_t^{J_t}\right\} - \min_k \sum_{t=1}^{n} b_t^k\right) \\
&\leq \sqrt{2nK^2 \ln K^2} + \sqrt{2nK^2 \ln K^2} \\
&= 2\sqrt{2nK^2 \ln K^2}
\end{aligned}
$$

Subquestion 3.2

Now we will use the a_t^i and b_t^j observations separately. Consider running two K-arm instances of EXP3, one with $i \to a_t^i$ as the loss and one with $j \to b_t^j$ as the loss. Have the first algorithm control I_t and the second J_t. Show that the overall expected pseudo-regret is bounded by

$$\bar{R}_n \leq 2\sqrt{2nK \ln K}.$$

Solution

We do similar as before, just that now we choose arm in the set of K arms and not K^2.

$$
\begin{aligned}
\bar{R}_n &= \mathbb{E}_{I_1,\ldots,I_n,J_1,\ldots,J_n}\left\{\sum_{t=1}^{n} a_t^{I_t} + b_t^{J_t}\right\} - \min_k \sum_{t=1}^{n} a_t^k + b_t^k \\
&= \left(\mathbb{E}_{I_1,\ldots,I_n}\left\{\sum_{t=1}^{n} a_t^{I_t}\right\} - \min_k \sum_{t=1}^{n} a_t^k\right) + \left(\mathbb{E}_{J_1,\ldots,J_n}\left\{\sum_{t=1}^{n} b_t^{J_t}\right\} - \min_k \sum_{t=1}^{n} b_t^k\right) \\
&\leq \sqrt{2nK \ln K} + \sqrt{2nK \ln K} \\
&= 2\sqrt{2nK \ln K}
\end{aligned}
$$