Exercise 1: Multiplicative Weights Update with Switches

Given \( n \) experts, in each round one must choose (the advice of) one expert. There are \( T \) rounds in total and \( T \) is know in advance. Choosing expert \( i \) in round \( t \) causes a loss of \( f^t_i \) and \( |f^t_i| \leq 1 \). We saw how to use multiplicative weight updates (MWU) to select experts, such that the regret, defined as difference between the expected loss of MWU and the optimal strategy, is in \( O(\sqrt{T \ln n}) \).

Assume that we want to compare MWU to the best strategy which is allowed to switch from one expert to another at most \( k \) times (instead of comparing MWU to just the best expert). In the rounds between switches the latest chosen expert is reused. The minimal loss with \( k \) switches is given by

\[
L_k := \min_{1=t_1<...<t_k<T+1} \left( \sum_{j=1}^{k} \min_{1 \leq i \leq n} \sum_{t_i=t_{j-1}}^{t_{j+1}} f_i^t \right).
\]

Adapt the MWU strategy (where we are allowed to switch each round), such that it has a regret of at most \( O(\sqrt{Tk \ln(nT)}) \) compared to the best strategy with at most \( k \) expert switches. Then prove your claim. For which range of \( k \) does the average regret converge to zero, as \( T \to \infty \)?

Sample Solution

Solution A

We increase the number of experts to simulate the optimization of experts for the (integer) intervals \( [x, y] \subseteq [1..T] \). Let the set of experts be \( E := \{(i, x, y) \in [n] \times [T]^2 \mid x < y \} \). For \( (i, x, y) \in E \) define the loss at time \( t \) as

\[
f^t_{i,x,y} = \begin{cases} f^t_i, & t \in [x, y] \\ 0, & \text{else} \end{cases}
\]

We can simulate each of these experts during the runtime of the MWU algorithm, since we only need to know the previous loss \( f^t_i \) to compute the losses \( f^t_{i,x,y} \). On top of that, we (figuratively speaking) “stretch” each round \( t \) into \( k \) rounds \( t_1, \ldots, t_k \), whereas each expert has exactly the same losses in rounds \( t_1, \ldots, t_k \) as in round \( t \). To simulate said “time stretch” in the MWU algorithm we simply update our weights \( k \) times per round with the same losses \( f^t_e, e \in E \). Then, for any given round \( t \), we obtain probability distributions \( q^1, \ldots, q^k \) on the set \( E \). We choose an expert from \( E \) for round \( t \) according to the probability distribution

\[
p^t := \frac{1}{k} \sum_{j=1}^{k} q_j^t. \tag{1}
\]

If \( (i, x, y) \in E \) is the expert we chose for round \( t \) according to \( p^t \), we simply map it to expert \( i \). That is, for each round \( t \) we obtain one expert in the original set of experts \( [n] \), which has the claimed regret with respect to the best solution with \( k \) switches. It remains to prove the claim.
From the lecture we know that the probability distribution $q^t_j$ with $t \in [T], j \in [k]$, which we get for the simulation of $|E| \leq nT^2$ experts in $kT$ rounds, generates an expected regret of $O(\sqrt{Tk \ln(nT^2)}) = O(\sqrt{Tk \ln(nT)})$ in comparison to the loss $L(e^*)$ of the best expert $e^* \in E$. More formally:

$$\sum_{t=1}^{T} \sum_{j=1}^{k} \langle q^t_j, f^t \rangle - L(e^*) \in O(\sqrt{Tk \ln(nT)})$$

Let $i^*_1, \ldots, i^*_k$ be the experts minimizing the loss with $k$ switches in rounds $t^*_1, \ldots, t^*_k \in [T]$. Each expert $i^*_j$ corresponds to an expert $E \ni e_j = (i^*_j, t^*_j, t^*_j+1-1)$. Let $L(e_j)$ be the sum of losses of expert $e_j$ with simulated “time stretch” (meaning that we incur losses $k$ times). Let $L(i^*_j)$ be the losses of $i^*_j$ in its respective time frame $[t^*_j, t^*_j+1-1]$. Then we have

$$L(e_j) = k \cdot L(i^*_j)$$  \hspace{1cm} (2)

(recall the definition of the loss $f^t_{i^*_j}$). In round $t$ which is zero for $t \notin [t^*_j, t^*_j+1-1]$). Obviously, the losses of some expert $e_j$ are bigger than for the best expert $e^*$:

$$L(e_j) \geq L(e^*)$$  \hspace{1cm} (3)

Hence

$$\sum_{t=1}^{T} \sum_{j=1}^{k} \langle q^t_j, f^t \rangle - L(e^*) \in O(\sqrt{Tk \ln(nT)}) \hspace{1cm} \text{(regret w.r.t. } e^*)$$

$$\sum_{t=1}^{T} \sum_{j=1}^{k} \langle q^t_j, f^t \rangle - L(e_j) \in O(\sqrt{Tk \ln(nT)}) \hspace{1cm} \text{(regret w.r.t. } e_j)$$

$$\sum_{t=1}^{T} k \langle p^t_j, f^t \rangle - kL(i^*_j) \in O(\sqrt{Tk \ln(nT)}) \hspace{1cm} \text{(regret w.r.t. } i^*_j \text{ in interval } [t^*_j, t^*_j+1-1])$$

$$\sum_{t=1}^{T} \sum_{j=1}^{k} \langle p^t_j, f^t \rangle - \sum_{j=1}^{k} L(i^*_j) \in O(\sqrt{Tk \ln(nT)}) \hspace{1cm} \text{(regret w.r.t. best solution with } k \text{ switches)}$$

**Solution B (Sketch)**

We use an (even bigger) set of experts $E'$ that reflects all possible strategies with $k$ switches. Each expert $e \in E'$ defines a choice of $t_1, \ldots, t_k \in [T]$ with $t_1 < \ldots < t_k$ and an assignment of experts from the set $[n]$ to the intervals $[t_j, t_{j+1}-1]$ with $j \in [k]$ (we formally set $t_{k+1} := T + 1$).

We have $T \leq \binom{k}{k-1} \leq T^k$ possible choices of $t_2, \ldots, t_k \in [T]$ with $1 = t_1 < \ldots < t_k$ and $n^k$ possible assignments of an expert from $[n]$ to each interval $[t_j, t_{j+1}-1]$. Thus $E' \subseteq [n]^k \times [T]^k$ and $|E'| \leq n^k \cdot T^k$.

We define the loss of some expert $e = (i_1, \ldots, i_k, t_1, \ldots, t_k) \in E'$ as follows. If the current round $t$ is within interval $t \in [t_j, t_{j+1}-1]$ we set $f^t_e := f^t_{i_j}$ (where $f^t_{i_j}$ denotes the loss of $i_j \in [n]$ in round $t$).

We conduct MWU with $E'$, which gives us some probability distribution $q^t_1, \ldots, q^t_T$ on $E'$. More precisely for any given $t \in T$ we have a probability $q^t_e$ for each $e \in E'$, whereas $\sum_{e \in E'} q^t_e = 1$.

We recover a distribution $p^1, \ldots, p^T$ on $[n]$ from $q^1, \ldots, q^T$ as follows. Let $p^t_i$ be the probability of expert $i$ in round $t$, then we define

$$p^t_i := \sum_{e=(i_1, \ldots, i_k, t_1, \ldots, t_k) \in E'} q^t_e,$$

where $i \in [t_j, t_{j+1}-1]$ and $i_j = i$.

The loss of the best expert $e^* = (i^*_1, \ldots, i^*_k, t^*_1, \ldots, t^*_k) \in E'$ equals the loss of the best choice of $i^*_1, \ldots, i^*_k$ with $k$ switches in rounds $t^*_1, \ldots, t^*_k \in [T]$. 


Furthermore, the (expected) regret of choosing experts according to $q_1, \ldots, q_T$ over the best expert $e^* \in E'$ is $O(\sqrt{T \log(n^k T^k)}) = O(\sqrt{T k \log(nT)})$ according to what we know from the lecture.

Finally, due to the definition of the loss function $f_e^t$ on $E'$ (see above), the loss of choosing experts from $E'$ according to $q_1, \ldots, q_T$ equals the loss of choosing experts from $[n]$ according to $p_1, \ldots, p_T$ (w.r.t. their respective loss functions). Thus the computed distribution $p_1^*, \ldots, p_T^*$ incurs a regret of $O(\sqrt{T k \log(nT)})$ over the best strategy $e^* = (i_1^*, \ldots, i_k^*, t_1^*, \ldots, t_k^*)$ with $k$ switches.