Practice Sessions Online Learning

Dirk van der Hoeven, Anand Kalvit, and Nicolò Cesa-Bianchi

Version of July 25, 2022

1 Practice Session One

In this session we will prove a bound on the regret of Follow The Leader (FTL) for squared loss and if time permits a first bound for Follow The Regularized Leader (FTRL) for linear losses. In one dimension, we define the squared loss in round t as

$$\ell_t(x) = (x - y_t)^2,$$

where $y_t \in [-1, 1]$.

Recall that the prediction of FTL in round t is defined as

$$x_t = \operatorname*{argmin}_{x \in V} \sum_{i=1}^{t-1} \ell_i(x),$$

where V is the domain of our predictions. In this example, we will work with $V = \mathbb{R}$. In round $t = 1$ we will choose $x_1 = 0$. In the first question we will compute x_t for $t \geq 2$.

Question 1 Compute x_t.

In the following questions we will show that we can bound the regret in terms of the (weighted) difference between x_t and x_{t+1}, summed over the rounds. In order to show that we will use that the regret can be written as

$$\begin{aligned} \sum_{t=1}^{T} \ell_t(x_t) - \min_{x \in V} \sum_{t=1}^{T} \ell_t(x) &= \sum_{t=1}^{T} \ell_t(x_t) - \sum_{t=1}^{T} \ell_t(x_{T+1}) \\ &= \sum_{t=1}^{T} \left(\sum_{i=1}^{t} \ell_i(x_t) - \sum_{i=1}^{t} \ell_i(x_{t+1}) \right). \end{aligned} \tag{1}$$

Question 2 Prove the equality in equation (1).

Hint: to ease notation a bit you may use $L_t(x) = \sum_{i=1}^{t} \ell_i(x)$.

Question 3 Argue that we can bound

$$\sum_{t=1}^{T} \left(\sum_{i=1}^{t} \ell_i(x_t) - \sum_{i=1}^{t} \ell_i(x_{t+1}) \right) \leq \sum_{t=1}^{T} \big(\ell_t(x_t) - \ell_t(x_{t+1}) \big).$$

Question 4 Show that for all t

$$\ell_t(x_t) - \ell_t(x_{t+1}) \leq 4|x_t - x_{t+1}|$$

Hint: $ab \leq |a||b|$

Question 5 Using the definition of x_t, x_{t+1}, show that for all t

$$|x_t - x_{t+1}| \leq \frac{2}{t}.$$

Hint: $|a - b| \leq |a| + |b|$ and $\frac{1}{t-1} - \frac{1}{t} = \frac{1}{t(t-1)}$.

Question 6 Combine the solutions to the first five questions and use the inequality $\sum_{t=1}^{T} \frac{1}{t} \leq 1 + \int_{2}^{T+1} \frac{1}{t-1} dt = 1 + \log(T)$ to bound the regret.

1.1 Adding Regularization to FTL

In the lecture we saw that FTL may suffer linear regret. Here we will remedy that by using regularization. The loss we consider in this section is $\ell_t(x) = xg_t$, where $g_t \in [-1, 1]$. The domain V in this example is the interval $[-1, 1]$. We will use the following prediction in round t:

$$x_t = \operatorname*{argmin}_{x \in V} \left(\sum_{i=1}^{t-1} x g_i + \frac{1}{2\eta} x^2 \right),$$

where $\eta > 0$ is a scalar called the learning rate, which we will choose at the end of this sequence of questions. As before, in round $t = 1$ we will use $x_1 = 0$. We will also use the following definition:

$$\psi_t(x) = \sum_{i=1}^{t-1} x g_i + \frac{1}{2\eta} x^2$$

so that $x_t = \operatorname{argmin}_{x \in V} \psi_t(x)$.

Question 7 Compute x_t for $t > 1$.

As before, we will write the regret in terms of a (weighted) difference between x_t and x_{t+1} with some additional terms. Denote by $u = \operatorname{argmin}_{x \in V} \sum_{t=1}^{T} x g_t$. We can rewrite the regret as

$$\sum_{t=1}^{T} \big(\ell_t(x_t) - \ell_t(u)\big) = \sum_{t=1}^{T} \big(\psi_{t+1}(x_t) - \psi_{t+1}(x_{t+1})\big) + \psi_{T+1}(x_{T+1}) - \psi_{T+1}(u) + \frac{1}{2\eta} u^2. \quad (2)$$

Question 8 Prove the equality in equation (2).

Question 9 Argue that $\psi_{T+1}(x_{T+1}) - \psi_{T+1}(u) \leq 0$.

Question 10 Prove that $\psi_{t+1}(x_t) - \psi_{t+1}(x_{t+1}) \leq \ell_t(x_t) - \ell_t(x_{t+1}) \leq \eta g_t^2$.

By combining the above we can show that

$$\sum_{t=1}^{T} \big(\ell_t(x_t) - \ell_t(u)\big) \leq \frac{1}{2\eta} u^2 + \eta \sum_{t=1}^{T} g_t^2. \tag{3}$$

All that is left is to choose an appropriate learning rate.

Question 11 Prove that $f(\eta) = \frac{1}{2\eta} u^2 + \eta \sum_{t=1}^{T} g_t^2$ is convex in η.

Question 12 Compute the η that minimizes the regret bound in equation (3). What is the regret bound? What are the two reasons why we cannot use the optimal learning rate in practice?

2 Practice Session Two

In this practice session we will apply what we have learned in the lectures so far to online binary classification with surrogate losses. In this setting, in round t

1 the learner observes feature vector $\boldsymbol{z}_t \in \mathbb{R}^d$

2 the learner predicts $\widehat{y}_t = \text{sign}(\boldsymbol{x}_t^\top \boldsymbol{z}_t)$

3 the learner suffers the zero-one loss: $\mathbb{1}[\widehat{y}_t \neq y_t]$ and observes $y_t \in \{-1, 1\}$.

4 the learner updates $\boldsymbol{x}_t$ to $\boldsymbol{x}_{t+1}$.

Note that we do not make any assumptions on how y_t is generated.

The goal of online binary classification with surrogate losses is to bound the number of mistakes the learner makes in terms of a surrogate loss $\ell_t : V \mapsto \mathbb{R}_+$. A surrogate loss is a *convex* upper bound on the zero-one loss. Examples of surrogate losses are the hinge loss, the smooth hinge loss, and the logistic loss. A typical bound in online binary classification with surrogate losses is the following:

$$\sum_{t=1}^T \mathbb{1}[\widehat{y}_t \neq y_t] \leq \min_{\boldsymbol{x} \in V} \sum_{t=1}^T \ell_t(\boldsymbol{x}) + 4Z\sqrt{2 \max_{\boldsymbol{x} \in V} \{\psi(\boldsymbol{u}) - \psi(\boldsymbol{x}_1)\} \frac{T}{\mu}},$$

where $Z \geq \max_t \|\boldsymbol{z}_t\|_\star$, ψ, and $\|\cdot\|, \|\cdot\|_\star$ are dual norms.

In this practice session we will use one surrogate loss in particular, namely the smooth hinge loss. The smooth hinge loss is defined as

$$\ell_t(\boldsymbol{x}) = \begin{cases} \max\{1 - 2\boldsymbol{x}^\top \boldsymbol{z}_t y_t, 0\} & \text{if } \boldsymbol{x}^\top \boldsymbol{z}_t y_t \leq 0 \\ \max\{(1 - \boldsymbol{x}^\top \boldsymbol{z}_t y_t)^2, 0\} & \text{if } \boldsymbol{x}^\top \boldsymbol{z}_t y_t > 0, \end{cases} \tag{4}$$

Figure 1 below gives a visual representation of the smooth hinge and the zero-one loss as a function of $\boldsymbol{x}^\top \boldsymbol{z}_t y_t$.

Question 1 Define $f_t(x) = \mathbb{1}[\text{sign}(\boldsymbol{x}^\top z_t) \neq y_t]$. Can we use the analysis of FTRL in the lecture to bound

$$\sum_{t=1}^T f_t(x_t) - \min_{x \in V} \sum_{t=1}^T f_t(x). \tag{5}$$

If not, what is the reason?

Since we can not use FTRL to control (5) we instead will use FTRL to control an upper bound on the number of mistakes, which is:

$$\sum_{t=1}^T \mathbb{1}[\widehat{y}_t \neq y_t] \leq \min_{\boldsymbol{x} \in V} \sum_{t=1}^T \ell_t(\boldsymbol{x}) + \sum_{t=1}^T \ell_t(\boldsymbol{x}_t) - \min_{\boldsymbol{x} \in V} \sum_{t=1}^T \ell_t(\boldsymbol{x}) \tag{6}$$

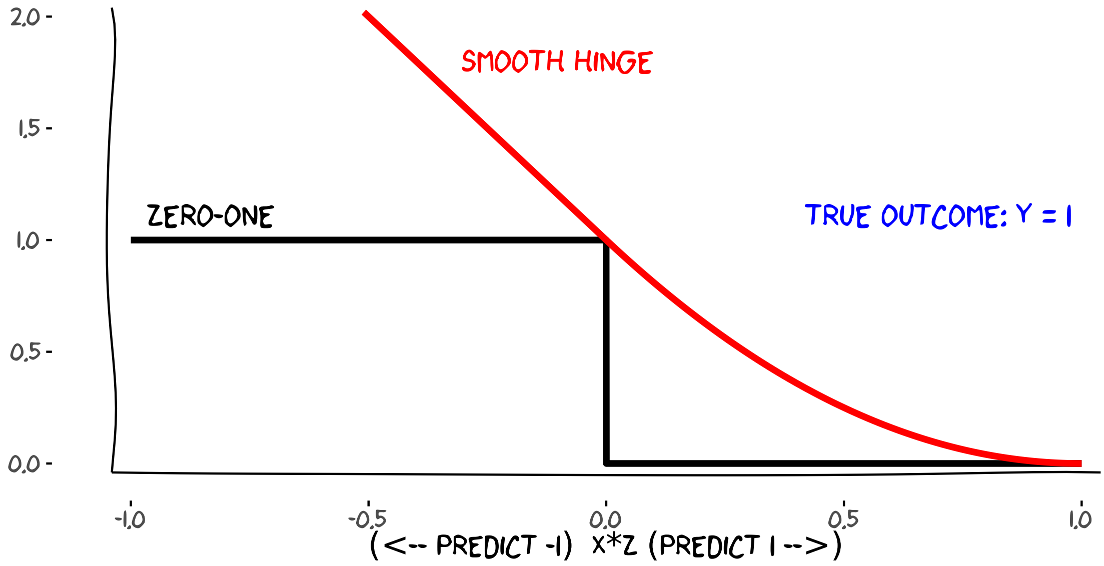

Figure 1: The smooth hinge and the zero-one loss as a function of $\boldsymbol{x}^\top \boldsymbol{z}_t y_t$.

Question 2 Prove the inequality in equation (6).

Suppose that we use

$$\boldsymbol{x}_t = \operatorname*{argmin}_{\boldsymbol{x} \in V} \sum_{i=1}^{t-1} \boldsymbol{x}^\top \nabla \ell_t(\boldsymbol{x}_t) + \frac{1}{\eta} \psi(\boldsymbol{x}),$$

which is the prediction of FTRL on linearised losses in round t. Denote by $\boldsymbol{u} = \operatorname{argmin}_{\boldsymbol{x} \in V} \sum_{t=1}^{T} \ell_t(\boldsymbol{x})$. Recall that if the regularization ϕ of FTRL is μ-strongly convex with respect to norm $\|\cdot\|$ we can bound

$$\begin{aligned} \sum_{t=1}^{T} \ell_t(\boldsymbol{x}_t) - \sum_{t=1}^{T} \ell_t(\boldsymbol{u}) &\leq \sum_{t=1}^{T} (\boldsymbol{x}_t - \boldsymbol{u})^\top \nabla \ell_t(\boldsymbol{x}_t) \\ &\leq \frac{1}{\eta}\big(\psi(\boldsymbol{u}) - \psi(\boldsymbol{x}_1)\big) + \sum_{t=1}^{T} \frac{2\eta}{\mu} \|\nabla \ell_t(\boldsymbol{x}_t)\|_\star^2. \end{aligned} \tag{7}$$

In the following sequence of questions we will appropriately bound (7).

Question 3 Compute $\nabla \ell_t(\boldsymbol{x}_t)$.

Question 4 Further bound the regret of FTRL using $Z \geq \max_t \|\boldsymbol{z}_t\|_\star$ and show that by choosing an appropriate learning rate η (which we can use in practice) we can bound

$$\sum_{t=1}^{T} \mathbb{1}[\widehat{y}_t \neq y_t] \leq \min_{\boldsymbol{x} \in V} \sum_{t=1}^{T} \ell_t(\boldsymbol{x}) + 4Z \sqrt{2 \max_{\boldsymbol{x} \in V} \{\psi(\boldsymbol{u}) - \psi(\boldsymbol{x}_1)\} \frac{T}{\mu}}.$$

2.1 Adding Randomness to the Predictions

So far we have ignored one issue: what happens if $\text{sign}(\boldsymbol{x}_t^\top \boldsymbol{z}_t) = 0$? While we could just say we always predict $\widehat{y}_t = 1$ or $\widehat{y}_t = -1$ in such cases, we can also flip an unbiased coin to choose $\widehat{y}_t$: if the coin flip gives us head we choose $\widehat{y}_t = 1$ and otherwise $\widehat{y}_t = -1$. If we flip a coin, the probability that $\widehat{y}_t \neq y_t$ is 0.5. If we assume that the sequence of $(y_t, \boldsymbol{z}_t)$ has been chosen before round 1, this implies that the expected zero-one loss of this prediction is also 0.5, i.e. $\mathbb{E}[\mathbb{1}[\widehat{y}_t \neq y_t]] = 0.5$. If we look at Figure 1 we can see that the smooth hinge loss is 1 if $\boldsymbol{x}_t^\top \boldsymbol{z}_t = 0$, which means that if $\boldsymbol{x}_t^\top \boldsymbol{z}_t = 0$ we have that

$$\mathbb{E}[\mathbb{1}[\widehat{y}_t \neq y_t]] = 0.5\ell_t(x_t).$$

This means that we can improve our bound in such cases! While this may seem like a minor improvement, we will generalize this idea and improve the bound we have shown before. In particular, we will show that if we add the perfect amount of randomness to our predictions our predictions will satisfy:

$$\mathbb{E}\left[\sum_{t=1}^T \mathbb{1}[\widehat{y}_t \neq y_t]\right] \leq \min_{\boldsymbol{x} \in V} \sum_{t=1}^T \ell_t(\boldsymbol{x}) +$$

Denote by $\ell_t^\star(\boldsymbol{x}_t) = (1 - \boldsymbol{x}_t^\top \boldsymbol{z}_t \text{sign}(\boldsymbol{x}_t^\top \boldsymbol{z}_t))^2$. Notice that $\ell_t^\star(\boldsymbol{x}_t) \in [0, 1]$ is an optimistic guess of the surrogate loss: if $\text{sign}(\boldsymbol{x}_t^\top \boldsymbol{z}_t) = y_t$ then $\ell_t^\star(\boldsymbol{x}_t) = \ell_t(\boldsymbol{x}_t)$.

Our predictions will be the following. With probability $1 - \ell_t^\star(\boldsymbol{x}_t)$ predict $\widehat{y}_t = \text{sign}(\boldsymbol{x}_t^\top \boldsymbol{z}_t)$ and with probability $\ell_t^\star(\boldsymbol{x}_t)$ predict $\widehat{y}_t = 1$ with probability 0.5 and $\widehat{y}_t = -1$ with probability 0.5.

Question 5 Show that

$$\mathbb{E}[\mathbb{1}[\widehat{y}_t \neq y_t]] \leq \frac{1}{2}\ell_t(\boldsymbol{x}_t) \tag{8}$$

Hint: split the proof into two cases: $\text{sign}(\boldsymbol{x}_t^\top \boldsymbol{z}_t) = y_t$ and $\text{sign}(\boldsymbol{x}_t^\top \boldsymbol{z}_t) \neq y_t$.

While we could use equation (8) to improve the bound in question 4 by a factor 2, we will instead rewrite equation (8) as

$$\mathbb{E}[\mathbb{1}[\widehat{y}_t \neq y_t]] \leq \ell_t(\boldsymbol{x}_t) - \frac{1}{2}\ell_t(\boldsymbol{x}_t) \tag{9}$$

While this seems like an inconsequential rewrite of equation (8) it turns out to be a very useful inequality. In order to use its potential we will first prove a slightly different bound in the analysis of FTRL.

Question 6 Prove that

$$\sum_{t=1}^T \ell_t(\boldsymbol{x}_t) - \sum_{t=1}^T \ell_t(\boldsymbol{u}) \leq \frac{1}{\eta}\big(\psi(\boldsymbol{u}) - \psi(\boldsymbol{x}_1)\big) + \eta \sum_{t=1}^T \frac{8Z^2}{\mu}\ell_t(\boldsymbol{x}_t)$$

With the above inequalities in hand we will prove the regret bound.

Question 7 Prove that

$$\mathbb{E}\left[\sum_{t=1}^{T} \mathbb{1}[\widehat{y}_t \neq y_t]\right] \leq \sum_{t=1}^{T} \ell_t(\boldsymbol{u}) + \frac{1}{\eta}\big(\psi(\boldsymbol{u}) - \psi(\boldsymbol{x}_1)\big) + \left(\eta\frac{8Z^2}{\mu} - \frac{1}{2}\right)\sum_{t=1}^{T} \ell_t(\boldsymbol{x}_t).$$

Question 8 Choose η such that

$$\mathbb{E}\left[\sum_{t=1}^{T} \mathbb{1}[\widehat{y}_t \neq y_t]\right] \leq \sum_{t=1}^{T} \ell_t(\boldsymbol{u}) + \frac{16Z^2}{\mu}\big(\psi(\boldsymbol{u}) - \psi(\boldsymbol{x}_1)\big).$$

Can we choose this η in practice? Can we use $V = \mathbb{R}^d$?

We can also prove a bound that holds with high-probability, not only in expectation. In order to do so we will use a version of Friedman's inequality. With probability at least $1-\delta$, we have that for any $\lambda \in [0,1]$

$$\sum_{t=1}^{T} \big(\mathbb{1}[\widehat{y}_t \neq y_t] - \mathbb{E}\left[\mathbb{1}[\widehat{y}_t \neq y_t]\right]\big) \leq \lambda \sum_{t=1}^{T} \mathbb{E}\left[\mathbb{1}[\widehat{y}_t \neq y_t]\right] + \frac{\log(1/\delta)}{\lambda} \tag{10}$$

Question 9 Use equation (10) to show for $\lambda \in [0,1]$ that with probability at least $1-\delta$

$$\begin{aligned}\sum_{t=1}^{T} \mathbb{1}[\widehat{y}_t \neq y_t] \leq \sum_{t=1}^{T} \ell_t(\boldsymbol{u}) + \frac{1}{\eta}\big(\psi(\boldsymbol{u}) - \psi(\boldsymbol{x}_1)\big) + \frac{\log(1/\delta)}{\lambda} \\ + \left(\eta\frac{8Z^2}{\mu} + \lambda - \frac{1}{2}\right)\sum_{t=1}^{T} \ell_t(\boldsymbol{x}_t).\end{aligned}$$

Question 10 Show that for appropriate η and λ we have that, with probability at least $1-\delta$

$$\sum_{t=1}^{T} \mathbb{1}[\widehat{y}_t \neq y_t] \leq \sum_{t=1}^{T} \ell_t(\boldsymbol{u}) + \frac{32Z^2}{\mu}\big(\psi(\boldsymbol{u}) - \psi(\boldsymbol{x}_1)\big) + 4\log(1/\delta).$$

3 Practice Session Three

In this practice session we will learn how to sell add spaces with sublinear regret. In online advertising, publishers sell their online ad space to advertisers through second-price auctions managed by ad exchanges. For each impression (ad display) created on the publisher's website, the ad exchange runs an auction on the fly. Empirical evidence shows that an informed choice of the seller's reserve price, disqualifying any bid below it, can indeed have a significant impact on the revenue of the seller. We assume the seller is also observing the highest bid together with the revenue. This richer feedback, which is key to proving the results in these notes, is made available by some ad exchanges such as AppNexus.

The seller's revenue in a second-price auction is computed as follows: if the reserve price r is not larger than the second-highest bid $b(2)$, then the item is sold to the highest bidder and the seller's revenue is $b(2)$. If r is between $b(2)$ and the highest bid $b(1)$, then the item is sold to the highest bidder and the seller's revenue is the reserve price. Finally, if r is bigger than $b(1)$, then the item is not sold and the seller's revenue is zero. Formally, the seller's revenue is

$$g\big(r, b(1), b(2)\big) = \max\big\{r, b(2)\big\}\mathbb{1}[r \leq b(1)]$$

Note that the revenue only depends on the reserve price r and on the two highest bids $b(1) \geq b(2)$, which —by assumption— all belong to the unit interval $[0, 1]$.

At the beginning of each auction $t = 1, 2, \ldots,$ the seller computes a reserve price $r_t \in [0, 1]$. Then, bids $b_t(1), b_t(2)$ are collected by the auctioneer, and the seller (which is not the same as the auctioneer) observes his revenue $g_t(r_t) = g\big(r_t, b_t(1), b_t(2)\big)$, together with the highest bid $b_t(1)$. Crucially, knowing $g_t(r_t)$ and $b_t(1)$ allows to compute $g_t(r)$ for all $r \geq r_t$. For technical reasons, we use losses $\ell_t(r_t) = 1 - g_t(r_t)$ instead of revenues, see Figure 2 for a pictorial representation.

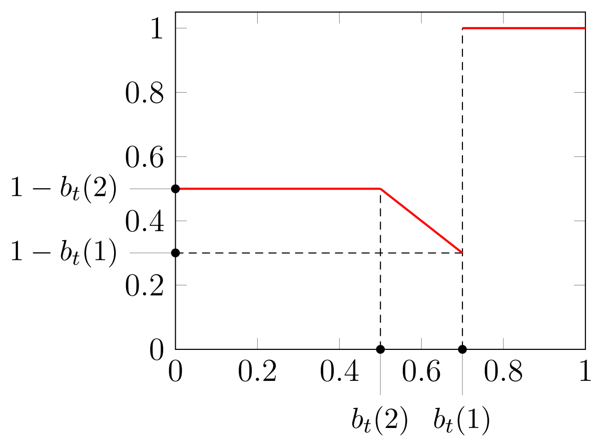

Figure 2: The loss function $\ell_t(r_t) = 1 - \max\{r_t, b_t(2)\}\mathbb{1}[r_t \leq b_t(1)]$ when $b_t(1) = 0.7$ and $b_t(2) = 0.5$.

The learner's regret is defined by

$$R_T = \mathbb{E}\left[\sum_{t=1}^{T} \ell_t(r_t)\right] - \inf_{0 \leq y \leq 1} \sum_{t=1}^{T} \ell_t(y) \ ,$$

where the expectation is with respect to the randomness in the reserves r_t. We introduce the Exp3-RTB algorithm (where RTB stands for Real Time Bidding), a variant of Exp3 exploiting the richer feedback $\big\{\ell_t(r) \,:\, y \geq r_t\big\}$. The algorithm uses a discretization of the action space $[0,1]$ in $K = \lceil 1/\gamma \rceil$ actions $y_k := (k-1)\gamma$ for $k = 1,\ldots,K$.

Algorithm 1 (Exp3-RTB)

Require: Exploration parameter $\gamma > 0$ and learning rate $\eta > 0$.

1: Set uniform distribution p_1 over $\{1,\ldots,K\}$ where $K = \lceil 1/\gamma \rceil$

2: **for** $t = 1, 2, \ldots$ **do**

3: compute distribution $q_t(k) = (1-\gamma)p_t(k) + \gamma \mathbb{1}[k=1]$ for $k = 1,\ldots,K$;

4: draw $I_t \sim q_t$ and choose $r_t = y_{I_t} = (I_t - 1)\gamma$;

5: for each $k = 1,\ldots,K$, compute the estimated loss

$$\widehat{\ell}_t(k) = \frac{\ell_t(y_k)}{\sum_{j=1}^k q_t(j)} \mathbb{1}[I_t \leq k]$$

6: for each $k = 1,\ldots,K$, compute the new probability assignment

$$p_{t+1}(k) = \frac{\exp\big(-\eta \sum_{i=1}^t \widehat{\ell}_i(k)\big)}{\sum_{j=1}^K \exp\Big(-\eta \sum_{i=1}^t \widehat{\ell}_i(j)\Big)}$$

7: **end for**

In the coming sequence of questions we will analyse Exp3-RTB and prove that the predictions of Exp3-RTB satisfy

$$R_T = \widetilde{\mathcal{O}}\big(\sqrt{T}\big),$$

where $\widetilde{\mathcal{O}}$ hides constants and logarithmic factors.

The analysis follows the same lines as the regret analysis of Exp3. The key change is a tighter control of the variance term allowed by the richer feedback.

Pick any reserve price $y_k = (k-1)\gamma$. We first control the regret associated with actions drawn from p_t (the regret associated with q_t will be studied as a direct consequence). A standard analysis of the regret associated with actions drawn from p_t shows that

$$\sum_{t=1}^T \sum_{j=1}^K p_t(j)\widehat{\ell}_t(j) - \sum_{t=1}^T \widehat{\ell}_t(k) \leq \frac{\log K}{\eta} + \sum_{j=1}^K p_t(j)\widehat{\ell}_t(j) + \frac{1}{\eta} \log\left(\sum_{j=1}^K p_t(j)\exp(-\eta \widehat{\ell}_t(j))\right) \tag{11}$$

Question 1 Use $1 + x \leq \exp(x)$ for $x \in \mathbb{R}$ and $\exp(x) \leq 1 + x + \frac{1}{2}x^2$ for $x \leq 0$ to show that

$$\sum_{t=1}^T \sum_{j=1}^K p_t(j)\widehat{\ell}_t(j) - \sum_{t=1}^T \widehat{\ell}_t(k) \leq \frac{\log K}{\eta} + \frac{\eta}{2} \sum_{t=1}^T \sum_{j=1}^K p_t(j)\widehat{\ell}_t(j)^2. \tag{12}$$

Denote by $\mathbb{E}_{t-1}[\cdot]$ the expectation conditioned on $I_1,\ldots,I_{t-1}$.

Question 2 Show that the estimator of the loss is unbiased, i.e. show that

$$\mathbb{E}_{t-1}\left[\widehat{\ell}_t(j)\right] = \ell_t(y_j).$$

Question 3 Show that

$$\mathbb{E}_{t-1}\left[p_t(j)\widehat{\ell}_t(j)^2\right] \le \frac{q_t(j)}{(1-\gamma)\sum_{s=1}^{j} q_t(s)}.$$

By using the (in)equalties in questions 2 and 3 we can see that taking expectation on both sides of (12) implies, similarly to what is done in the analysis of Exp3,

$$\mathbb{E}\left[\sum_{t=1}^{T}\sum_{j=1}^{K} p_t(j)\ell_t(j)\right] - \sum_{t=1}^{T}\ell_t(y_k) \le \frac{\log K}{\eta} + \frac{\eta}{2(1-\gamma)}\sum_{t=1}^{T}\mathbb{E}\left[\sum_{j=1}^{K}\frac{q_t(j)}{\sum_{s=1}^{j} q_t(s)}\right]. \tag{13}$$

For our next step we will bound $\sum_{j=1}^{K}\frac{q_t(j)}{\sum_{s=1}^{j} q_t(s)}$. To do so we will first need the following technical result due to Gaillard et al. [2014].

Lemma 1 *Let $a_0 > 0$ and $a_1, \ldots, a_m \in [0,1]$ and let $f : (0,+\infty) \mapsto [0,+\infty)$ be a nonincreasing function. Then*

$$\sum_{n=1}^{m} a_n f(a_0 + \ldots + a_n) \le \int_{a_0}^{a_0+a_1+\ldots+a_m} f(u)\mathrm{d}u.$$

Let us abbreviate $A_n = a_0 + \ldots + a_n$. Before we use Lemma 1, let us prove it first.

Question 4 Prove Lemma 1.
Hint: use that $a_n f(A_n) = \int_{A_{n-1}}^{A_n} f(A_n)\mathrm{d}u$ and the assumption that f in nonincreasing.

Question 5 Use Lemma 1 and the fact that $q_t(1) \ge \gamma$ to show that

$$\sum_{j=1}^{K}\frac{q_t(j)}{\sum_{s=1}^{j} q_t(s)} \le 1 + \log\left(\frac{1}{\gamma}\right). \tag{14}$$

If we substitute equation (14) into (13), we get

$$\mathbb{E}\left[\sum_{t=1}^{T}\sum_{i=1}^{K} p_t(i)\ell_t(i)\right] - \sum_{t=1}^{T}\ell_t(y_k) \le \frac{\eta T \log(e/\gamma)}{2(1-\gamma)} + \frac{\log K}{\eta} \tag{15}$$

For our next step we will control the regret of the reserve prices $r_t = y_{I_t}$, where I_t is drawn from $q_t = (1-\gamma)p_t + \gamma\delta_1$.

Question 6 Show that

$$\mathbb{E}\left[\sum_{t=1}^{T} \ell_t(r_t)\right] - \sum_{t=1}^{T} \ell_t(y_k) \leq \frac{\eta T \log(e/\gamma)}{2} + \frac{\log K}{\eta} + \gamma T \tag{16}$$

To conclude the proof, we need to upper bound the regret against any fixed $y \in [0, 1]$ rather than the regret against a fixed y_k.

Question 7 Show that

$$\min_{k=1,\dots,K} \sum_{t=1}^{T} \ell_t(y_k) \leq \min_{0 \leq y \leq 1} \sum_{t=1}^{T} \ell_t(y) + \gamma\, T. \tag{17}$$

Question 8 Combine the preceding inequalities and choose η and γ such that

$$R_T \leq 2\sqrt{\tfrac{1}{2} T \log(e\sqrt{T})} + \sqrt{T}$$

References

Pierre Gaillard, Gilles Stoltz, and Tim van Erven. A second-order bound with excess losses. In *Conference on Learning Theory*, 2014.