Lecture 02:
Follow-the-Regularized Leader – No-regret via Regularization

Lecturer: John Lazarsfeld

September 16, 2025

Abstract

Family of leader-based algorithms, analysis of Follow-the-Regularized-Leader (FTRL) via coupling with Be-the-Leader and Follow-the-Leader, Multiplicative Weights Update (MWU) as FTRL with entropic regularizaation, and lower bounds for online learning.

1 Warmup: Multiplicative Weights Update

Recap from Lecture 01. In the previous lecture, we introduced the general Online Convex Optimization setting (as well as the special *experts* case) and motivated using external regret as a performance metric. We also introduced the fundamental Online Gradient Descent (OGD) algorithm and proved that it obtained the sublinear regret bound $\text{Reg}_{OGD}(T) \leq O(\sqrt{T})$. In this lecture, we begin by introducing a second, canonical online learning algorithm called *Multiplicative Weights Update* (MWU), which we describe for the experts setting.

MWU Algorithm. In the experts setting, the MWU algorithm chooses its iterates by updating the n coordinates of its next distribution choice x_{t+1} in a *multiplicative* fashion, with multiplicative factors depending on the most recent loss value of the n coordinates.

Algorithm 1 Multiplicative Weights Update (MWU) for Experts Setting

Input: Initial distribution $x_1 = (1/n, \dots, 1/n) \in \Delta_n$; stepsize parameter $\eta > 0$.
for $t = 1, \dots, T$ **do**:
1. Play action $x_t \in \Delta_n$, and incur cost $\langle \ell_t, x_t \rangle$. Observe the feedback $\ell_t \in \mathbb{R}^n$.
2. Update $x_{t+1} \in \Delta_n$ by

$$x_{t+1}(i) \;:=\; \frac{x_t(i) \cdot \exp(-\eta \ell_t(i))}{\sum_{j=1}^n x_t(j) \cdot \exp(-\eta \ell_t(j))} \quad \text{for } i \in [n] \;. \tag{1}$$

end for

Remark 1. We make several remarks about the MWU algorithm:

- **Simlarities with OGD**:
  While at first glance the update rule of MWU appears much different than that of Online Gradient Descent, both methods share the same algorithmic principal: the next action choice x_{t+1} should update its coordinates in a *greedy* manner based on the most recent feedback.

  Similar to OGD, for MWU, the aggressiveness of this greedy strategy is controlled by the stepsize parameter η. Larger settings of η correspond to more aggressive updates (more *exploitation*) and smaller settings place less influence on the most recent feedback (more *exploration*).

- **Historical notes**:
  The MWU algorithm (or variants) have been (re)-discovered in many related contexts. The variant stated in Algorithm 1 is sometimes known as *Hedge* and originates with Freund and Schapire (1997). The algorithm is also known as the *Exponential Weights* method, and has had many applications across online learning, and ML theory more broadly. See also Arora et al. (2012) for more historical notes.

- **Regret guarantees and analysis**:
  Given the previous point, it is no surprise that MWU has strong performance guarantees in the online learning setting. Informally, MWU has the following regret guarantee:

  Theorem. *In the experts setting, assuming bounded loss vectors and with an appropriate setting of η, MWU obtains regret* $\mathrm{Reg}_{MWU}(T) \le O(\sqrt{\log n \cdot T})$.

  There exist multiple proofs of this result, many of which are based on the use of a suitably-defined potential function. However, these more direct proofs can feel opaque with building a better intuition for why MWU should be an effective online learning algorithm.

Focus of lecture: MWU as an instantiation of Follow-the-Regularized-Leader. In light of the final point of Remark 1, note that it turns out that MWU is a special instantiation of a more general family of algorithms called *Follow-the-Regularized-Leader* (FTRL). Deriving regret bounds for the FTRL family gives a more intuitive understanding for MWU's behavior, and thus in this lecture the goal is to introduce and analyze this more general family. The proof of MWU's sublinear regret bound will follow as a corollary.

2 The Follow-the-Regularized-Leader Family

In this section, we now introduce and analyze the *Follow-the-Regularized-Leader* (FTRL) family of online learning algorithms. FTRL algorithms are one of the workhorse methods of online learning, and under suitable parameter setings, this family obtains sublinear regret guarantees. The key algorithmic component used by FTRL is *regularization*, which ensures the algorithm's iterates are sufficiently stable.

Overview of remainder of the lecture. We will proceed to introduce FTRL and derive a full proof of its regret bound that emphasizes the importance of such *stability* for obtaining sublinear regret in the adversarial OCO setting. In the subsequent section, we will then show how MWU is a particular instantiation of FTRL using entropic regularization. This provides an alternative and perhaps more intuitive proof of MWU's $O(\sqrt{T})$ regret bound. In Section 3, we conclude the lecture by presenting information theoretic *lower bounds* for online learning that scale like $\Omega(\sqrt{T})$. Together, this implies for general online convex optimization that the regret guarantees of FTRL, MWU, and OGD are tight in their dependence on T.

2.1 FTRL Update Rule.

Reminders on assumptions of OCO setting. Throughout, we will assume the general OCO setting introduced from Lecture 01. In particular, we will assume a compact convex decision set $\mathcal{X} \subset \mathbb{R}$, and we assume each loss function $f_t : \mathcal{X} \to \mathbb{R}$ is convex and differentiable. We assume the *full feedback* setting where the learner observes f_t after choosing its action x_t.

FTRL is a *family* of algorithms instantiated with a strictly convex *regularizer* $R : \mathcal{X} \to \mathbb{R}$. At each timestep, the FTRL update rule chooses the action that minimizes the sum of the previously-observed loss functions up to the current round, plus the addition of the regularizer. Stated formally:

Remark 2. We make several remarks on the regularizer R in the FTRL family:

- **On the role of the regularizer**: In FTRL, the role of the regularizer is to *stabilize* the algorithm's iterates: the regularization term in the objective of (2), weighted by a stepsize parameter η, balances the dependence on the next iterate's dependence on the loss functions from previous rounds. As the stepsize η grows smaller, the influence of the regularizer

Algorithm 2 Follow-the-Regularized-Leader (FTRL) for Online Convex Optimization

Input: Initialization $x_1 \in \mathcal{X}$; strictly convex regularizer $R : \mathcal{X} \to \mathbb{R}$; stepsize $\eta > 0$.
for $t = 1, \dots, T$ **do**:
1. Play action $x_t \in \mathcal{X}$ and incur cost $f_t(x_t)$. Observe function $f_t : \mathcal{X} \to \mathbb{R}$.
2. Perform the update:

$$x_{t+1} := \operatorname*{argmin}_{x \in \mathcal{X}} \left\{ \sum_{k=1}^{t} f_k(x) + \frac{1}{\eta} R(x) \right\} . \tag{2}$$

end for

relative to the history of loss functions from prior rounds grows larger, increasing the role of exploration as opposed to exploitation (c.f., the discussion of the MWU algorithm).

- **Example regularizers**: Usually, we will wish to instantiate FTRL with a regularizer R that has nice analytic properties with respect to the geometry of the decision set $\mathcal{X}$. For example, the Euclidean regularizer $\|x\|_2^2$ when $\mathcal{X}$ is an ℓ_2 ball, or the negative entropy function $\sum_{i=1}^n x_i \log x_i$ when $\mathcal{X}$ is the simplex Δ_n.

  Convexity of the regularizer: As a baseline, we assume in Algorithm 2 that the regularizer R is *strictly* convex. Given the convexity of the loss functions $\{f_t\}$ and the compactness of $\mathcal{X}$, this ensures that the argmin operator in (2) exists and is unique. However, in the main regret bound we will derive for FTRL, we make the stronger assumption of *strong convexity*.

  For this, we briefly recall below the notion of strong convexity as well as some other tools from convex analysis that will be used in the remainder of the lecture.

Refresher on strong convexity and dual norms. We recall the notion of strong-convexity, dual norms, and the generalized Cauchy-Schwarz inequality (for more details and background, see Orabona (2019, Sections 4.1.1 and 4.2.2) and Hazan et al. (2016, Section 2.1)):

(a) **Strong convexity:**

Definition 3. *Let $\alpha > 0$. A differentiable function $f : \mathcal{X} \to \mathbb{R}$ is α-strongly-convex with respect to $\|\cdot\|$ iff for all $x, x' \in \mathcal{X}$:*

$$f(x') \geq f(x) + \langle \nabla f(x), x' - x \rangle + \frac{\alpha}{2} \|x - x'\|^2 .$$

(b) **Dual norms:**

Definition 4. *The dual norm $\|\cdot\|_\star$ of $\|\cdot\|$ is given by $\|x\|_\star = \max_{x' : \|x'\| \leq 1} \langle x', x \rangle$.*

Example 5. Fix $p \geq 2$. The dual norm of the ℓ_p norm is the ℓ_q norm for $(1/p) + (1/q) = 1$. The dual norm of the ℓ_1 norm is the ℓ_∞ norm, which is given by $\|x\|_\infty = \max_{i \in [n]} |x_i|$.

Lemma 6 (Generalized Cauchy-Schwarz)**.** *Let $\|\cdot\|$ be a norm. Then $\langle x, x' \rangle \leq \|x\|_\star \|x\|$ for all $x, x' \in \mathbb{R}^n$.*

2.2 Intuition for FTRL via Leader-based Algorithms

Before stating its regret guarantee, we develop intuition for the FTRL family via the perspective of *leader-based* algorithms.

Be-the-Leader (BTL). Consider the OCO setting as described above with loss functions $\{f_t\}$. The *Be-the-Leader* (BTL) algorithm is an unimplementable, clairvoyant method for choosing iterates: at every step $t+1$, BTL outputs the action $x_t \in \mathcal{X}$ that minimizes the sum of the loss *through and including round* $t+1$:

$$x_{t+1} := \operatorname*{argmin}_{x \in \mathcal{X}} \sum_{k=1}^{t+1} f_k(x) \ . \tag{BTL}$$

The BTL algorithm is unimplementable given that x_{t+1} depends on ℓ_{t+1}, which violates the setup of the OCO model. However, this clairvoyant strategy always leads to *non-positive* regret:

Lemma 7. *Let $\{x_t\}$ be the iterates of* (BTL) *on loss functions $\{f_t\}$. Then* $\mathrm{Reg}_{BTL}(T) \leq 0$.

Follow-the-Leader (FTL). Given the strong regret guarantee of BTL, a natural idea for making the method implementable is to simply remove the dependence on the current round's loss function, and to instead choose the next action based on the sum of losses through the *previous round* (for which the learner *does* already have access to in the OCO setting). This yields the *Follow-the-Leader* (FTL) strategy with iterates given by:

$$x_{t+1} := \operatorname*{argmin}_{x \in \mathcal{X}} \sum_{k=1}^{t} f_k(x) \ . \tag{FTL}$$

Here, note the only key difference between (FTL) and (BTL) is the upper limit in the summation (through round $t+1$ for BTL, and through round t for FTL).

For FTL, we can derive the following, general regret guarantee:

Lemma 8. *Let $\{x_t\}$ be the iterates of* (FTL) *on any sequence of loss functions $\{f_t\}$. Then*

$$\mathrm{Reg}_{FTL}(T) \leq \sum_{t=1}^{T} \|\nabla f_t(x_t)\|_\star \|x_t - x_{t+1}\| \ .$$

Remark 9. We make several remarks on the FTL regret bound:

- **Dependence on path-length in FTL regret bound**:
  Suppose that the sequence of loss functions have uniformly bounded gradients. Then Lemma 8 reveals that the regret of FTL depends on the *stability* of the iterates $\{x_t\}$. In particular, the regret is bounded above by the *path-length* $\sum_{t=1}^{T} \|x_t - x_{t+1}\|$. How large can this path-length quantity grow? In general, linearly in T.
- **Example – Worst-case bound on path-length in experts setting**:
  To make the previous point clear, consider the experts setting where $\mathcal{X} = \Delta_n$ and $f_t(x) = \langle \ell_t, x \rangle$ for all t. Thus the update rule (FTL) reduces to $x_{t+1} := \mathrm{argmin}_{x \in \mathcal{X}} \sum_{k=1}^{t} \langle \ell_t, x \rangle$ for all $t \geq 1$. Given the linear nature of this objective, each x_t is a vertex of Δ_n (represented by a standard basis vector $e_i \in \mathbb{R}^n$). For every t such that $x_t \neq x_{t+1}$, it then follows that $\|x_t - x_{t+1}\| = 1$. Thus, in the worst case, supposing $x_t \neq x_{t+1}$ for all $t \geq 1$, then $\sum_{t=1}^{T} \|x_t - x_{t+1}\| \leq T$.
- **Example – Concrete $\Omega(T)$ regret lower bound for FTL in experts setting**:
  The intuition of the previous point directly leads to a concrete *lower bound* for FTL algorithm. In particular, consider the experts setting with $n = 2$, and suppose $x_1 = (p, 1-p)$ for $p > 0.5$. Suppose the adversary then selects $\ell_1 = (0, -0.9)$, and $\ell_t = (-1, 0)$ for even $t \geq 2$, and $\ell_t = (0, -1)$ for odd $t \geq 3$. Using the definition of (FTL) and $\mathrm{Reg}_{FTL}(T)$, one can compute for this example that $\mathrm{Reg}_{FTL}(T) \geq \Omega(T)$. Thus FTL is *not* a no-regret algorithm for online learning.

FTL with Regularization. The points of Remark 9 reveal the key failure point of FTL: the greedy nature of the update rule is highly sensitive to variations in the losses, which can lead to a lack of stability in the iterates and thus linear regret. The FTRL algorithm attempts to counter this lack of stability via *regularization*. Recalling the update rule of FTRL as stated in Algorithm 2, we have

$$x_{t+1} = \operatorname*{argmin}_{x \in \mathcal{X}} \left\{ \sum_{k=1}^{t} f_k(x) + \frac{R(x)}{\eta} \right\} . \tag{FTRL}$$

The objective in the update (FTRL) is exactly that in (FTL) with the addition of the regularization term. As mentioned in Remark 2, the amount of regularization is controlled by the stepsize parameter η. For larger stepsize settings, the amount of regularization decreases, and the update of (FTRL) approaches that of (FTL). In the experts setting, the regularization leads to choosing distributions $x_t \in \Delta_n$ that are more *uniform* or have greater *entropy* (as opposed to the point mass distributions that are chosen by FTL).

Main regret guarantee for FTRL. We will prove the following guarantees for FTRL:

Theorem 10. *Let $\{x_t\}$ be the iterates of FTRL (Algorithm 2) with strictly convex regularizer $R : \mathcal{X} \to \mathbb{R}$ and stepsize $\eta > 0$ on any sequence of convex and differentiable loss functions $\{f_t\}$. Then for any $T \geq 1$, the following bounds hold:*

(1) Suppose for $M > 0$ that $|R(x) - R(x')| \leq M$ for all $x, x' \in \mathcal{X}$. Then:

$$\mathrm{Reg}_{FTRL}(T) \;\leq\; \sum_{t=1}^{T} \|\nabla f_t(x_t)\|_\star \|x_t - x_{t+1}\| + \frac{M}{\eta} .$$

(2) Moreover, assume R is 1-strongly-convex with respect to $\|\cdot\|$, then for all $t \geq 1$:

$$\|x_t - x_{t+1}\| \;\leq\; \eta \cdot \|\nabla f_t(x_t)\|_\star .$$

(3) If in addition there exists $L > 0$ such that $\|\nabla f_t(x)\|_\star \leq L$ for all $t \geq 1$ and $x \in \mathcal{X}$, then setting $\eta := \frac{\sqrt{M}}{L\sqrt{T}}$:

$$\mathrm{Reg}_{FTRL}(T) \;\leq\; 2L\sqrt{MT} .$$

Remark 11. We make the following remarks on Theorem 10:

- **On the general regret bound of part (1)**:
  Part (1) of the theorem gives a general regret bound similar to that of Lemma 8 for FTL. In particular, this bound depends on the stability of the iterates (as well as the norm of the gradients of the losses) over all rounds. This bound also now depends on the *diameter* of the regularizer M.
- **Stable iterates via strongly-convex regularizer**:
  Part (2) of the theorem develops the crucial bound on the stability of the iterates $\|x_t - x_{t+1}\| \leq O(\eta)$ under strong convexity of the regularizer.
- **Using a time-invariant stepsize**:
  For simplicity, and similar to the analysis of Online Gradient Descent from Lecture 01, we stated and analyzed here the FTRL algorithm with time-invariant stepsize η. To obtain the final, sublinear regret bound scaling like $O(\sqrt{T})$ in Part (3) of the theorem, we set $\eta = O(1/\sqrt{T})$. For an analysis using time-varying stepsizes scaling like $\eta_t = 1/\sqrt{T})$, see Orabona (2019, Section 7.2).

2.3 Proof of FTRL Regret via BTL/FTL Coupling and Stability Bound

We now develop the proof of Theorem 10. To start, we prove the regret bounds for BTL and FTL which were stated above in Lemma 7 and Lemma 8. Together, these lead to the proof of Part (1) of the main theorem, which we restate as Lemma 12. We then restate and prove Part (2) of the main theorem, which establishes the stability of FTRL iterates under strong convexity of the regularizer, in Lemma 13. The stepsize-instantiated bound in Part (3) of the main theorem is finally restated and proved in Lemma 14.

Regret bound for BTL. We start by proving the non-positive regret guarantee of BTL from Lemma 7 (restated here):

Lemma 7. *Let $\{x_t\}$ be the iterates of (BTL) on loss functions $\{f_t\}$. Then* $\mathrm{Reg}_{BTL}(T) \leq 0$.

Proof. We prove the claim by induction on t:

- **Base case (t=1)**: Observe by definition of (BTL) that $f_1(x_1) = \min_{x\in\mathcal{X}} f_1(x)$. Then

$$\mathrm{Reg}_{BTL}(1) = f_1(x_1) - \min_{x\in\mathcal{X}} f_1(x_1) = 0\,.$$

- **Inductive case (t>1)**: Assume the claim holds for all $t \geq 1$. Observe by definition of (BTL) that $\min_{x\in\mathcal{X}} \sum_{k=1}^{t+1} f_k(x) = \sum_{k=1}^{t+1} f_k(x_{t+1})$. Then

$$\begin{aligned}
\mathrm{Reg}_{BTL}(t+1) &= \sum_{k=1}^{t+1} f_k(x_k) - \min_{x\in\mathcal{X}} \sum_{k=1}^{t+1} f_k(x) \\
&= \sum_{k=1}^{t+1} f_k(x_k) - \sum_{k=1}^{t+1} f_k(x_{t+1}) \\
&= \sum_{k=1}^{t} f_k(x_k) - \sum_{k=1}^{t} f_k(x_{t+1}) + \big(f_{t+1}(x_{t+1}) - f_{t+1}(x_{t+1})\big) \\
&\leq \sum_{k=1}^{t} f_k(x_k) - \min_{x\in\mathcal{X}} \sum_{k=1}^{t} f_k(x) \;=\; \mathrm{Reg}_{BTL}(t) \;\leq\; 0\,.
\end{aligned}$$

Here, the final inequality comes from the inductive hypothesis. □

Regret bound for FTL. Using the BTL regret bound, we derive the bound for the FTL algorithm from Lemma 8 (restated here):

Lemma 8. *Let $\{x_t\}$ be the iterates of (FTL) on any sequence of loss functions $\{f_t\}$. Then*

$$\mathrm{Reg}_{FTL}(T) \leq \sum_{t=1}^{T} \|\nabla f_t(x_t)\|_\star \|x_t - x_{t+1}\|\,.$$

Proof. Fix the set of loss functions $\{f_t\}$ The proof strategy is to consider the iterates of BTL on $\{f_t\}$, and to relate these iterates to those of FTL:

(i) **Couple the BTL and FTL iterates**:
Let $\{\tilde{x}_t\}$ denote the iterates of BTL on $\{f_t\}$. By definition of (BTL) and (FTL), this means for all $t \geq 1$ that

$$\tilde{x}_t = \operatorname*{argmin}_{x\in\mathcal{X}} \sum_{k=1}^{t} f_k(x) \quad \text{and} \quad x_{t+1} = \operatorname*{argmin}_{x\in\mathcal{X}} \sum_{k=1}^{t} f_k(x)\,.$$

Thus $\tilde{x}_t = x_{t+1}$ for all $t \geq 1$.

(ii) **Couple the FTL and BTL regret terms**:
Using the definitions of $\mathrm{Reg}_{FTL}(T)$ and $\mathrm{Reg}_{BTL}(T)$, we can then further write

$$\begin{aligned}\mathrm{Reg}_{FTL}(T) - \mathrm{Reg}_{BTL}(T) &= \sum_{t=1}^{T} f_t(x_t) - \min_{x\in\mathcal{X}} \sum_{t=1}^{T} f_t(x) - \Big(\sum_{t=1}^{T} f_t(\tilde{x}_t) - \min_{x\in\mathcal{X}} \sum_{t=1}^{T} f_t(x) \Big) \\ &= \sum_{t=1}^{T} f_t(x_t) - f_t(x_{t+1}) \,,\end{aligned}$$

where we use the relationship $\tilde{x}_t = x_{t+1}$. Rearranging terms, we have

$$\mathrm{Reg}_{FTL}(T) \;=\; \sum_{t=1}^{T} f_t(x_t) - f_t(x_{t+1}) + \mathrm{Reg}_{BTL}(T) \;\le\; \sum_{t=1}^{T} f_t(x_t) - f_t(x_{t+1}) \,, \tag{3}$$

where the inequality comes from the fact that $\mathrm{Reg}_{BTL}(T) \le 0$ (Lemma 7).

(iii) **Use convexity of losses to simplify**:
By convexity of f_t and applying Lemma 6 we have for all t that

$$f_t(x_t) - f_t(x_{t+1}) \;\le\; \langle \nabla f_t(x_t), x_t - x_{t+1} \rangle \;\le\; \|\nabla f_t(x_t)\|_\star \|x_t - x_{t+1}\| \,. \tag{4}$$

Substituting (4) into (3), we conclude

$$\mathrm{Reg}_{FTL}(T) \;\le\; \sum_{t=1}^{T} \|\nabla f_t(x_t)\|_\star \|x_t - x_{t+1}\| \,. \qquad \square$$

Regret bound for FTRL via FTL. Using the regret bound for FTL of Lemma 8, we can further derive a general regret bound for FTRL. This yields claim (1) of Theorem 10, which we restate and prove below as the following lemma:

Lemma 12. *Let $\{x_t\}$ be the iterates of FTRL (Algorithm 2) with regularizer $R : \mathcal{X} \to \mathbb{R}$ and stepsize $\eta > 0$ on any sequence of convex and differentiable loss functions $\{f_t\}$. Assume for $M > 0$ that $|R(x) - R(x')| \le M$ for all $x, x' \in \mathcal{X}$. Then*

$$\mathrm{Reg}_{FTRL}(T) \;\le\; \sum_{t=1}^{T} \|\nabla f_t(x_t)\|_\star \|x_t - x_{t+1}\| + \frac{M}{\eta} \,.$$

Proof. The proof strategy is to consider the output of running FTL on the same sequence of loss functions $\{f_t\}$ encountered by FTRL, but offset by one day and with the regularizer R as the first day's loss function. We break up the proof into the following steps:

(i) **Define sequence of coupled loss functions.**

First, we define the functions $\hat{f}_t : \mathcal{X} \to \mathbb{R}$ by

$$\begin{cases} \hat{f}_1(x) := R(x)/\eta \\ \hat{f}_{t+1}(x) := f_t(x) \text{ for } t \ge 1. \end{cases}$$

Now let $\{\hat{x}_t\}$ be the iterates of (FTL) run on the sequence $\{\hat{f}_t\}$, with $\hat{x}_1 = x_1$. Then by definition of (FTL) and (FTRL), observe for all $t \ge 1$ that

$$\begin{aligned}\hat{x}_{t+1} = \operatorname*{argmin}_{x\in\mathcal{X}} \sum_{t=1}^{t} \hat{f}_t(x) &= \operatorname*{argmin}_{x\in\mathcal{X}} \frac{R(x)}{\eta} + \sum_{t=2}^{t} \hat{f}_t(x) \\ &= \operatorname*{argmin}_{x\in\mathcal{X}} \frac{R(x)}{\eta} + \sum_{t=1}^{t-1} \hat{f}_{t+1}(x) \\ &= \operatorname*{argmin}_{x\in\mathcal{X}} \frac{R(x)}{\eta} + \sum_{t=1}^{t-1} f_t(x) \;=\; x_t \,.\end{aligned}$$

Thus we have the key property that $\hat{x}_{t+1} = x_t$ for all $t \geq 1$.

(ii) **Bound the regret of FTL on the coupled loss sequence.**

Fixing $x \in \mathcal{X}$, it follows from the regret bound for FTL in Lemma 8 that we can write and further simplify

$$\begin{aligned}
\sum_{t=1}^{T+1} \hat{f}_t(x) - \hat{f}_t(\hat{x}_t) \;\leq\; \mathrm{Reg}_{FTL}(T+1) \;&\leq\; \sum_{t=1}^{T+1} \|\nabla \hat{f}_t(\hat{x}_t)\|_\star \|\hat{x}_t - \hat{x}_{t+1}\| &(5)\\
&= \sum_{t=2}^{T+1} \|\nabla \hat{f}_t(\hat{x}_t)\|_\star \|\hat{x}_t - \hat{x}_{t+1}\| &(6)\\
&= \sum_{t=1}^{T} \|\nabla \hat{f}_{t+1}(\hat{x}_{t+1})\|_\star \|\hat{x}_{t+1} - \hat{x}_{t+2}\| &(7)\\
&= \sum_{t=1}^{T} \|\nabla f_t(x_t)\|_\star \|x_t - x_{t+1}\| &(8)
\end{aligned}$$

Here, in expression (6) we use the fact that $\hat{x}_1 = \hat{x}_2 = x_1$, in expression (7) we re-index the sum, and in expression (8) we use the facts that $\hat{f}_{t+1} = f_t$ and $\hat{x}_{t+1} = x_t$ for all $t \geq 1$.

(iii) **Relate FTL regret on coupled losses to FTRL regret on true losses.**

On the other hand, we also have for fixed $x \in \mathcal{X}$ that

$$\begin{aligned}
\sum_{t=1}^{T} \hat{f}_t(\hat{x}_t) - \hat{f}_t(x) \;&=\; \frac{R(\hat{x}_1)}{\eta} + \sum_{t=2}^{T} \hat{f}_t(\hat{x}_t) - \left(\frac{R(x)}{\eta} + \sum_{t=2}^{T} \hat{f}_t(x) \right) &(9)\\
&= \sum_{t=1}^{T} \hat{f}_{t+1}(\hat{x}_{t+1}) - \hat{f}_{t+1}(x) + \frac{R(\hat{x}_1) - R(x)}{\eta} &(10)\\
&= \sum_{t=1}^{T} f_t(x_t) - f_t(x) + \frac{R(x_1) - R(x)}{\eta} &(11)
\end{aligned}$$

where we again use the relationships $\hat{f}_{t+1} = f_t$, $\hat{x}_{t+1} = x_t$, and $\hat{x}_1 = x_1$ as above.

Combining expressions (11) and (8) and rearranging, we have for any fixed $x \in \mathcal{X}$ that

$$\sum_{t=1}^{T} f_t(x_t) - f_t(x) \;\leq\; \sum_{t=1}^{T} \|\nabla f_t(x_t)\|_\star \|x_t - x_{t+1}\| + \frac{R(x) - R(x_1)}{\eta}\,. \tag{12}$$

Minimizing both sides of (12) over all $x \in \mathcal{X}$ and recalling by definition of M that $\min_{x \in \mathcal{X}} |R(x) - R(x_1)| \leq M$, we find

$$\begin{aligned}
\mathrm{Reg}_{FTRL}(T) \;&=\; \sum_{t=1}^{T} f_t(x_t) - \min_{x \in \mathcal{X}} \sum_{t=1}^{T} f_t(x)\\
&\leq\; \sum_{t=1}^{T} \|\nabla f_t(x_t)\|_\star \|x_t - x_{t+1}\| + \frac{M}{\eta}\,.
\end{aligned}$$

□

Stability of FTRL Iterates via Strongly-Convex Regularizer. Assuming strong-convexity of the regularizer R, we derive a bound on the *stability* of the FTRL iterates with respect to the stepsize η. This yields claim (ii) of the main Theorem 10, which we restate and prove below in the following lemma:

Lemma 13. *Let $\{x_t\}$ be the iterates of FTRL (Algorithm 2) with regularizer $R : \mathcal{X} \to \mathbb{R}$ and stepsize $\eta > 0$ on any sequence of convex and differentiable loss functions $\{f_t\}$. Suppose further that R is 1-strongly-convex. Then for all $t \geq 1$:*

$$\|x_{t+1} - x_t\| \leq \eta \cdot \|\nabla f_t\|_\star\,.$$

Proof. Fix $t \geq 1$. To prove the statement, we define the functions $F : \mathcal{X} \to \mathbb{R}$ and $G : \mathcal{X} \to \mathbb{R}$ that are minimized by the actions x_{t+1} and x_t, respectively. Specifically, for $x \in \mathcal{X}$, let

$$\begin{aligned} F(x) &= \sum_{k=1}^{t} f_k(x) + \frac{R(x)}{\eta} \\ G(x) &= \sum_{k=1}^{t-1} f_k(x) + \frac{R(x)}{\eta} \ . \end{aligned} \tag{13}$$

The following properties of F and G hold:

(a) For any fixed $x \in \mathcal{X}$:

$$F(x) - G(x) = \sum_{k=1}^{t} f_k(x) - \sum_{k=1}^{t-1} f_k(x) = f_t(x) \ .$$

(b) F and G are $(1/\eta)$-strongly-convex.

This follows from the assumption that the regularizer R is 1-strongly-convex, and the fact that all f_t are convex.

(c) $x_{t+1} = \operatorname{argmin}_{x\in\mathcal{X}} F(x)$ and $x_t = \operatorname{argmin}_{x\in\mathcal{X}} G(x)$.

This follows from the definition of F and G and the update rule of $FTRL$. Note in particular that strong-convexity of F and G imply the uniqueness of their argmin sets.

Moreover, this optimality condition also implies $\nabla F(x_{t+1}) = \nabla G(x_t) = 0 \in \mathbb{R}^n$.

Combining Observations (b) and (c) with the strong-convexity property of Definition 3, we then have

$$F(x_t) \geq F(x_{t+1}) + \frac{1}{2\eta}\|x_{t+1} - x_t\|^2 \tag{14}$$

$$\text{and} \quad G(x_{t+1}) \geq G(x_t) + \frac{1}{2\eta}\|x_{t+1} - x_t\|^2 \ . \tag{15}$$

Adding expressions (14) and (15) and simplifying yields

$$\begin{aligned} \frac{1}{\eta}\|x_{t+1} - x_t\|^2 &\leq F(x_t) - G(x_t) - (F(x_{t+1}) - G(x_{t+1})) && (16) \\ &= f_t(x_t) - f_t(x_{t+1}) && (17) \\ &\leq \langle \nabla f_t(x_t), x_t - x_{t+1} \rangle && (18) \\ &\leq \|\nabla f_t(x_t)\|_\star \cdot \|x_t - x_{t+1}\| \ . && (19) \end{aligned}$$

Here, (17) follows from Property (a), (18) follows from the first-order convexity property, and (19) follows from Lemma 6.

Simplifying further, we find

$$\frac{1}{\eta}\|x_t - x_{t+1}\| \leq \|\nabla f_t(x_t)\|_\star \cdot \|x_t - x_{t+1}\| \qquad \Longrightarrow \qquad \|x_t - x_{t+1}\| \leq \eta \cdot \|\nabla f_t(x_t)\|_\star \ ,$$

which proves the claim. □

Instantiation of FTRL regret bound. We now prove Claim (3) of Theorem 10, which gives a quantitative regret bound for FTRL under the assumption of a strongly-convex regularizer and uniformly bounded gradients of the loss functions:

Lemma 14. *Let $\{x_t\}$ be the iterates of FTRL (Algorithm 2) with regularizer $R : \mathcal{X} \to \mathbb{R}$ and stepsize $\eta > 0$ on any sequence of convex and differentiable loss functions $\{f_t\}$. Suppose that R is 1-strongly-convex, that $|R(x) - R(x')| \le M$ for all $x, x' \in \mathcal{X}$, and that $\|\nabla f_t(x)\| \le L$ for all $x \in \mathcal{X}$ and $t \ge 1$. Then setting $\eta := \frac{\sqrt{M}}{L\sqrt{T}}$, we have*

$$\mathrm{Reg}_{FTRL}(T) \le 2L\sqrt{MT} .$$

Proof. Combining Lemmas 12 and 13, we have for any fixed $\eta > 0$

$$\mathrm{Reg}_{FTRL}(T) \;\le\; \sum_{t=1}^{T} \eta \|\nabla f_t(x_t)\|^2 + \frac{M}{\eta} \;\le\; \eta L^2 T + \frac{M}{\eta}$$

where in the final inequality we apply the bounded gradients assumption. Plugging in the stepsize setting $\eta = \frac{\sqrt{M}}{L\sqrt{T}}$ then yields the stated bound. □

2.4 Regret bound for MWU via FTRL

We now show that the MWU algorithm for the experts setting (Algorithm 1) is an instantiation of FTRL using *entropic* regularization.

(Negative) Entropy Regularizer. The entropy function $H : \Delta_n \to \mathbb{R}$ is given by $H(x) = -\sum_{i=1}^{n} x_i \log x_i$. This function is 1-strongly *concave* with respect to the ℓ_1 norm $\|\cdot\|_1$. We will consider the use of the regularizer $R(x) = -H(x)$ (i.e., the negative entropy function). Given the strong-concavity of H, the negative entropy regularizer $R(x) = \sum_{i=1}^{n} x_i \log x_i$. is 1-strongly-convex with respect to the ℓ_1 norm (see Hazan et al. (2016, Section 5.4.2) for more background on the entropy function).

Instantiated FTRL Update Rule. Using the negative entropy regularizer, the update rule in (FTRL) becomes:

$$x_{t+1} := \operatorname*{argmin}_{x \in \Delta_n} \left\{ \langle x, \sum_{k=1}^{t} \ell_k \rangle + \frac{1}{\eta} \cdot \left(\sum_{i=1}^{n} x_i \log x_i \right) \right\} \tag{MWU}$$

Using properties of the entropy function, it can be shown that the objective in (MWU) has a closed-form solution via the *softmax* function. In particular, we have the following proposition:

Proposition 15. *Let $x_{t+1} \in \Delta_n$ be defined as in (MWU). Then for all $i \in [n]$,*

$$x_{t+1}(i) \;=\; \frac{\exp\left(-\eta \cdot \sum_{k=1}^{t} \ell_t(i) \right)}{\sum_{j=1}^{n} \exp\left(-\eta \cdot \sum_{k=1}^{t} \ell_t(j) \right)} \, . \tag{20}$$

Moreover, using properties of the exponential function, we also have an equivalence between the softmax expression for x_t from Proposition 15 and the multiplicative-update-based expression for x_t from the definition of Algorithm 1:

Proposition 16. *Fix a sequence of loss vectors $\{\ell_t\}$. Let $\{p_t\}$ be the iterates of Algorithm 1 initialized from $p_1 = (1/n, \ldots, 1/n) \in \Delta_n$, and let $\{x_t\}$ be the iterates as given in expression (20) of Proposition 15 intialized from $x_1 = p_1$. Then $x_t = p_t$ for all $t \ge 1$.*

The proofs of Proposition 15 and Proposition 16 are left as exercises.

Regret Bound for MWU via Theorem 10. Propositions 15 and 16 establish that MWU is the instantiation of FTRL using the negative entropy regularizer. Thus, by applying the guarantees of FTRL from Theorem 10, we have the following regret bound for MWU:

Theorem 17. *Let $\{x_t\}$ be the iterates of MWU on an instance of the experts setting with loss functions $f_t(x) = \langle \ell_t, x \rangle$ for all $t \geq 1$. Suppose that $\ell_t \in [-1, 1]^n$ for all $t \geq 1$. Then setting $\eta = \frac{\sqrt{\log n}}{\sqrt{T}}$:*

$$\mathrm{Reg}_{MWU}(T) \leq 2\sqrt{\log n \cdot T} \,.$$

Proof. As stated above, the negative entropy regularizer $R(x) = \sum_{i=1}^n x_i \log x_i$ is 1-strongly convex with respect to the ℓ_1 norm $\|\cdot\|_1$. Moreover, $-\log n \leq R(x) \leq 0$ for all $x \in \mathcal{X}$. Additionally, recall by Definition 4 and Example 4 that the ℓ_∞ norm $\|\cdot\|_\infty$ is the dual norm of $\|\cdot\|_1$. Together, the following properties then hold:

(i) For all $t \geq 1$, we have
$$\|\nabla f_t(x_t)\|_\star = \|\ell_t\|_\infty \leq 1 \,.$$

(ii) For all $x, x' \in \mathcal{X}$,
$$|R(x) - R(x')| \leq |\max_{x \in \mathcal{X}} R(x) - \min_{x \in \mathcal{X}} R(x)| \leq \log n \,.$$

Then applying Parts (1) and (2) of Theorem 10 with $L = 1$ and $M = \log n$, we find

$$\mathrm{Reg}_{MWU}(T) \leq \eta T + \frac{\log n}{\eta} \,.$$

Finally, optimizing the bound over η (which amounts to the stated stepsize setting of $\eta := \frac{\sqrt{\log n}}{\sqrt{T}}$ yields the final regret bound $\mathrm{Reg}_{MWU}(T) \leq 2\sqrt{\log n \cdot T}$. □

3 Lower Bounds for Online Learning

In this section, we develop *lower bounds* for online learning. In particular, we will prove regret bounds for the *experts* setting scaling like $\Omega(\sqrt{T})$, and these bounds will hold *information-theoretically* (i.e., independent of any algorithm). This implies that the regret upper bound established in Theorem 10 are optimal in their dependence on T.

Lower Bound Construction for Experts Setting. We present now the components of a lower bound construction for the experts setting based on (Arora et al., 2012, Theorem 4.1). Consider the following sequence of loss vectors. For all $t \geq 1$:

$$\ell_t(1) = \frac{1}{2} \quad \text{and} \quad \ell_t(i) = \begin{cases} 1 & \text{w.p. } 1/2 \\ 0 & \text{w.p. } 1/2 \end{cases} \quad \text{for all } i \geq 2 \,. \tag{21}$$

For convenience, we define the following pieces of additional notation: first, for $i \in [n]$, let $X_i = \sum_{t=1}^T \ell_t(i)$. Further define $\Phi := \min_{i \in [n]} X_i$.

Given this construction, we have the following proposition:

Proposition 18. *let $\{\ell_t\}$ be the sequence of loss vectors defined in (21). Then for any sequence $\{x_t\}$ where each $x_t \in \Delta_n$, the following statements hold:*

(i) $\mathbf{E}\left[\sum_{t=1}^T \langle x_t, \ell_t \rangle\right] = \frac{T}{2}$*, where the expectation is over the randomness of the construction.*

(ii) $0 \leq \Phi \leq \frac{T}{2}$ *with probability 1 over the randomness of the construction.*

Proof. To prove part (i), observe by definition of the loss vectors ℓ_t that, for any $t \geq 1$,

$$\mathbf{E}\left[\langle x_t, \ell_t \rangle\right] = \mathbf{E}\left[\sum_{i=1}^{n} x_t(i) \cdot \ell_t(i)\right] = \sum_{i=1}^{n} x_t(i)\, \mathbf{E}\left[\ell_t(i)\right] = \frac{1}{2}\,.$$

Here, the equalities follow from the linearity of expectation, the fact that $\mathbf{E}[\ell_t(i)] = \frac{1}{2}$ for all $i \in [n]$ by construction, and the fact that $x_t \in \Delta_n$ and thus $\sum_{i=1}^{n} x_t(i) = 1$. Summing over all $t \in [T]$ and again using the linearity of expectation, we find $\mathbf{E}\left[\sum_{t=1}^{T} \langle x_t, \ell_t \rangle\right] = \frac{T}{2}$.

For the part (ii), notice by definition that $X_1 = \frac{T}{2}$ with probability 1 by construction. Moreover, $X_i \geq 0$ for all i given that each $\ell_t(i) \in \{0, 1\}$. Thus $0 \leq \Phi = \min_{i \in [n]} X_i \leq \frac{T}{2}$ with probability 1. □

We also restate the following bound on the tails of the random variable Φ.

Lemma 19 (Lemma 4.2 of Arora et al. (2012))**.** *Let X_i and Φ be defined as above. Then for $\alpha := 0.25\sqrt{T \log(n-1)}$,*

$$\Pr\left[\Phi \leq \frac{T}{2} - \alpha\right] \;\geq\; 0.05\,.$$

With these preliminaries in hand, we now state the following guarantee:

Theorem 20. *Fix $n \geq 2$, and consider the random loss vectors defined in (21). Then for any sequence of distributions $\{x_t\}$, there exists a realization of the loss vectors $\{\ell_t\}$ such that*

$$\sum_{t=1}^{T} \langle x_t, \ell_t \rangle - \min_{i \in [n]} \sum_{t=1}^{T} \ell_t(i) \;\geq\; 0.0125\sqrt{T \log(n-1)}\,.$$

Proof. Our goal is to derive a lower bound on the *expected* regret given by the (random) sequence loss vectors $\{\ell_t\}$ on any fixed sequence of distributions $\{x_t\}$. Recalling the definitions of $X_i = \sum_{t=1}^{T} \ell_t(i)$ and $\Phi := \min_{i \in [n]} X_i$, this expected regret is given by

$$\mathbf{E}\left[\sum_{t=1}^{T} \langle x_t, \ell_t \rangle - \min_{i \in [n]} \sum_{t=1}^{T} \ell_t(i)\right] \;=\; \mathbf{E}\left[\sum_{t=1}^{T} \langle x_t, \ell_t \rangle\right] - \mathbf{E}\left[\Phi\right]\,. \tag{22}$$

Now by part (i) of Proposition 18, we have $\mathbf{E}[\sum_{t=1}^{T} \langle x_t, \ell_t \rangle] = T/2$ for any sequence $\{x_t\}$. Thus, to derive a lower bound on (22), it suffices to derive an upper bound on $\mathbf{E}[\Phi]$.

For this, recalling from Proposition 18 that $0 \leq \Phi \leq \frac{T}{2}$ it follows by definition of expectation that

$$\mathbf{E}[\Phi] = \sum_{0 \leq \phi \leq \frac{T}{2}} \phi \cdot \Pr[\Phi = \phi] \;=\; \sum_{0 \leq \phi \leq \frac{T}{2} - \alpha} \phi \cdot \Pr[\Phi = \phi] \;+\; \sum_{\frac{T}{2} - \alpha < \phi \leq \frac{T}{2}} \phi \cdot \Pr[\Phi = \phi] \tag{23}$$

For the first term of (23), we have

$$\sum_{0 \leq \phi \leq \frac{T}{2} - \alpha} \phi \cdot \Pr[\Phi = \phi] \;\leq\; \left(\frac{T}{2} - \alpha\right) \cdot \Pr\left[\Phi \leq \frac{T}{2} - \alpha\right]. \tag{24}$$

For the second term of (23), we similarly have

$$\sum_{\frac{T}{2} - \alpha < \phi \leq \frac{T}{2}} \phi \cdot \Pr[\Phi = \phi] \;\leq\; \frac{T}{2} \cdot \left(1 - \Pr\left[\Phi \leq \frac{T}{2} - \alpha\right]\right). \tag{25}$$

Substituting expressions (24) and (26) back into (23), we find

$$\mathbf{E}[\Phi] \;\le\; \frac{T}{2} - \alpha \cdot \Pr\left[\Phi \le \frac{T}{2} - \alpha\right] \;\le\; \frac{T}{2} - 0.05\alpha\,, \tag{26}$$

where the last inequality comes from applying the tail bound of Lemma 19.

Combining this upper bound on $\mathbf{E}[\Phi]$ with the fact that $\mathbf{E}[\sum_{t=1}^{T}\langle x_t, \ell_t\rangle] = T/2$, we conclude

$$\begin{aligned}\mathbf{E}\left[\sum_{t=1}^{T}\langle x_t, \ell_t\rangle - \min_{i\in[n]}\sum_{t=1}^{T}\ell_t(i)\right] \;&\le\; \frac{T}{2} - \left(\frac{T}{2} - 0.05\alpha\right)\\ &=\; 0.0125\sqrt{T\log(n-1)}\,,\end{aligned}$$

where in the final equality we apply the setting of $\alpha = 0.25\sqrt{T\log(n-1)}$.

As the bound holds in expectation over the randomness of the loss vectors $\{\ell_t\}$, it follows that there exists a realization of this sequence such that $\sum_{t=1}^{T}\langle x_t, \ell_t\rangle - \Phi \ge 0.0125\sqrt{T\log(n-1)}$ deterministically, which proves the claim. $\square$

References

Sanjeev Arora, Elad Hazan, and Satyen Kale. The multiplicative weights update method: a meta-algorithm and applications. *Theory of computing*, 8(1):121–164, 2012.

Yoav Freund and Robert E Schapire. A decision-theoretic generalization of on-line learning and an application to boosting. *Journal of computer and system sciences*, 55(1):119–139, 1997.

Elad Hazan et al. Introduction to online convex optimization. *Foundations and Trends® in Optimization*, 2(3-4):157–325, 2016.

Francesco Orabona. A modern introduction to online learning. *arXiv preprint arXiv:1912.13213*, 2019.