Basics of Online Convex Optimization

Yanze Song

November 30, 2023

1 Intro to OCO

1.1 Problem Formulation

Online Convex Optimization (*OCO*) can be considered as a repetitive game between the player/algorithm and the environment/adversary. Denote the convex decision space by $\mathcal{C}$, the convex loss function by $f : \mathcal{C} \to \mathbb{R}$, and the time horizon by T. For $t = 1, 2, ..., T$:

- Algorithm: Picks a decision vector $\boldsymbol{w}_t \in \mathcal{C}$,
- Adversary: Picks and sends a convex loss f_t with **full information** to the algorithm,
- Algorithm: Suffer, learn, and adapt from the loss.

1.2 Performance Metric

We use *regret*, the difference between the loss generated by the algorithm and the loss by the best **fixed** decision in hindsight, to quantify the performance of the algorithm:

$$\text{Regret}_T \coloneqq \sum_{t=1}^{T} f_t(\boldsymbol{w}_t) - \min_{\boldsymbol{u}\in\mathcal{C}} \sum_{t=1}^{T} f_t(\boldsymbol{u}),$$

and it can also be defined w.r.t. some fixed decision $\boldsymbol{u} \in \mathcal{C}$:

$$\text{Regret}_T(\boldsymbol{u}) \coloneqq \sum_{t=1}^{T} f_t(\boldsymbol{w}_t) - \sum_{t=1}^{T} f_t(\boldsymbol{u}).$$

This measures how much we "regret" for not picking the best fixed decision in hindsight. Since we're comparing with a fixed decision, this is more precisely called *static regret*, and it can indeed be negative. Then two natural questions may arise:

- *Why do we use the best fixed decision instead of the offline optimal as the benchmark?* The best fixed decision already has all information ahead of time, while the algorithm only receives one piece at a time, and hence, competing with it is already highly non-trivial.
- *What if the best fixed decision doesn't perform well on the losses either?* Then we will switch to other stronger regret measures, e.g., comparing with the offline optimal, *dynamic regret*:

$$\text{D-Regret}_T \coloneqq \sum_{t=1}^{T} f_t(\boldsymbol{w}_t) - \min_{\boldsymbol{u}_t\in\mathcal{C}} \sum_{t=1}^{T} f_t(\boldsymbol{u}_t).$$

Another reason for studying static regret is that the algorithms for those stronger regret measures are often designed by extending the ideas from static regret, so it's the foundation.

1.3 Goal

Due to the online setting, it's unrealistic to achieve zero regret, so we reasonably allow some mistakes. If we can achieve *sub-linear regret*, i.e.,

$$\text{Regret}_T \in o(T),$$

then $\lim_{T\to\infty} \text{Regret}_T/T = 0$, which means that on average the algorithm is performing as well as the best fixed decision. In this case, the algorithm/learner is called "no-regret", indicating that we don't regret for making those decisions.

1.4 Feedback Model

We assume full information feedback for general OCO, i.e., f_t will be fully observed when revealed.

1.5 Adversary Model

2 OCO & OLO

Theorem 2.0.1. *The regret of any algorithm on an OLO problem is no better than if it is on the corresponding OCO problem. In other words, for any algorithm $\mathcal{A}$,*

$$\operatorname*{Regret}_{\mathcal{C}'}(\mathcal{A}) \leq \operatorname*{Regret}_{\mathcal{L}}(\mathcal{A}),$$

provided that the classes of convex and linear functions, $\mathcal{C}'$ and $\mathcal{L}$, are equivalent.

Proof. Let $\boldsymbol{w}_t \in \mathcal{C}$ be our t^{th} decision, and $\boldsymbol{z}_t \in \partial f_t(\boldsymbol{w}_t)$. Then for all $\boldsymbol{w} \in \mathcal{C}$, by convexity we have

$$f_t(\boldsymbol{w}_t) + \langle \boldsymbol{z}_t, \boldsymbol{w} - \boldsymbol{w}_t \rangle \leq f_t(\boldsymbol{w}) \Rightarrow f_t(\boldsymbol{w}_t) - f_t(\boldsymbol{w}) \leq \langle \boldsymbol{z}_t, \boldsymbol{w}_t \rangle - \langle \boldsymbol{z}_t, \boldsymbol{w} \rangle,$$

and if we sum over t, we have

$$\sum_{t=1}^{T} f_t(\boldsymbol{w}_t) - f_t(\boldsymbol{w}) \leq \sum_{t=1}^{T} \langle \boldsymbol{z}_t, \boldsymbol{w}_t \rangle - \langle \boldsymbol{z}_t, \boldsymbol{w} \rangle.$$

Let $\boldsymbol{w}^*$ maximize the LHS. Substituting $\boldsymbol{w}^*$ into above, and by the maximality of the RHS, we have

$$\max_{\boldsymbol{w} \in \mathcal{C}} \sum_{t=1}^{T} f_t(\boldsymbol{w}_t) - f_t(\boldsymbol{w}) \leq \max_{\boldsymbol{w} \in \mathcal{C}} \sum_{t=1}^{T} \langle \boldsymbol{z}_t, \boldsymbol{w}_t \rangle - \langle \boldsymbol{z}_t, \boldsymbol{w} \rangle,$$

which completes the proof. □

Remark 2.0.1. The theorem implies that linear loss is the worst-case instance in OCO, so it's sufficient to consider the linear loss only. In the end, for any OCO problem, our decisions will exactly be the ones generated by solving the corrsponding OLO problem (we don't solve OCO directly).

Proof. Here's a more general proof: for each round, if we can find $\hat{f}_t \in \hat{\mathcal{F}}$ (some function class) s.t.

- $\hat{f}_t(\boldsymbol{w}_t) = f_t(\boldsymbol{w}_t)$,
- $\hat{f}_t(\boldsymbol{w}) \leq f_t(\boldsymbol{w})$ for all $\boldsymbol{w} \in \mathcal{C}$,

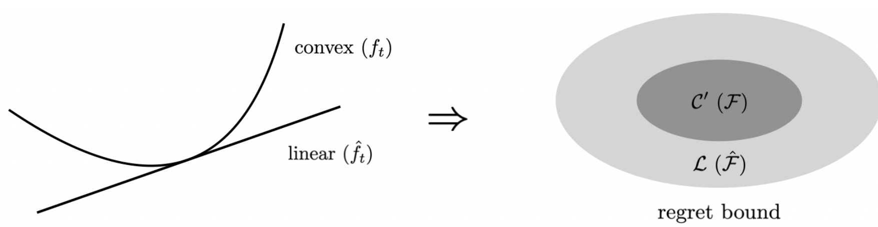
or slightly more strictly, $\hat{f}_t$ should be s.t. for all $\boldsymbol{w} \in \mathcal{C}$,

$$f_t \textit{ has an under-estimator } \hat{f}_t \in \hat{\mathcal{F}} \textit{ tangent at } \boldsymbol{w}.$$

Then we will have

$$\underbrace{\sum_{t=1}^{T} f_t(\boldsymbol{w}_t) - \min_{\boldsymbol{u} \in \mathcal{C}} \sum_{t=1}^{T} f_t(\boldsymbol{u})}_{\operatorname*{Regret}_{\mathcal{F}}} \leq \underbrace{\sum_{t=1}^{T} \hat{f}_t(\boldsymbol{w}_t) - \min_{\boldsymbol{u} \in \mathcal{C}} \sum_{t=1}^{T} \hat{f}_t(\boldsymbol{u})}_{\operatorname*{Regret}_{\hat{\mathcal{F}}}}.$$

In our case where $\mathcal{F} = \mathcal{C}'$, we choose $\hat{\mathcal{F}} = \mathcal{L}$ and $\hat{f}_t$ as the linearization of f_t at $\boldsymbol{w}_t$. The following graph may help visualize the relationship:



□

3 OCO Algorithms

3.1 Follow The Leader

A natural idea is to use the decision that minimizes the cumulative loss, which is our benchmark up to the previous round. This is called the *Follow-the-Leader* (*FTL*) algorithm:

$$\boldsymbol{w}_{t+1} := \operatorname*{argmin}_{\boldsymbol{w}\in\mathcal{C}} \sum_{i=1}^{t} f_i(\boldsymbol{w}).$$

Lemma 3.1.1 (*Regret Bound of FTL*)**.** *Let $\{\boldsymbol{w}_1, \boldsymbol{w}_2, ...\}$ be the decisions generated by FTL, then for all $\boldsymbol{u} \in \mathcal{C}$, we have*

$$\mathrm{Regret}_T(\boldsymbol{u}) \leq \sum_{t=1}^{T}(f_t(\boldsymbol{w}_t) - f_t(\boldsymbol{w}_{t+1})).$$

Proof. By definition of the regret, we can rewrite the above, and prove the following by induction:

$$\sum_{t=1}^{T}(f_t(\boldsymbol{w}_t) - f_t(\boldsymbol{u})) \leq \sum_{t=1}^{T}(f_t(\boldsymbol{w}_t) - f_t(\boldsymbol{w}_{t+1})) \Leftrightarrow \sum_{t=1}^{T} f_t(\boldsymbol{w}_{t+1}) \leq \sum_{t=1}^{T} f_t(\boldsymbol{u}),$$

When $T = 1$, we have LHS $= f_1(\boldsymbol{w}_2) \leq f_1(\boldsymbol{u})$, because $w_2 := \operatorname{argmin}_{\boldsymbol{w}\in\mathcal{C}} f_1(\boldsymbol{w})$, and this proves the base case. Then we assume $\sum_{t=1}^{T} f_t(\boldsymbol{w}_{t+1}) \leq \sum_{t=1}^{T} f_t(\boldsymbol{u})$ for some $T \geq 1$, but we'll only use when $\boldsymbol{u} = \boldsymbol{w}_{T+2}$. We then have

$$\begin{aligned}
\sum_{t=1}^{T+1} f_t(\boldsymbol{w}_{t+1}) &= \sum_{t=1}^{T} f_t(\boldsymbol{w}_{t+1}) + f_{T+1}(\boldsymbol{w}_{T+2}) \\
&\leq \sum_{t=1}^{T} f_t(\boldsymbol{w}_{T+2}) + f_{T+1}(\boldsymbol{w}_{T+2}) \\
&\leq \sum_{t=1}^{T+1} f_t(\boldsymbol{w}_{T+2}),
\end{aligned}$$

and this completes the proof, since $\boldsymbol{w}_{T+2} := \operatorname{argmin}_{\boldsymbol{w}\in\mathcal{C}} \sum_{t=1}^{T+1} f_t(\boldsymbol{w})$. □

Remark 3.1.1. If the decisions generated by FTL are not stable, i.e., two consecutive decisions are far from each other, then $f_t(\boldsymbol{w}_t) - f_t(\boldsymbol{w}_{t+1})$ will likely be large, and hence the regret bound can be high (not necessarily tight) after summing up T terms.

Example 3.1.1 (*Failure of FTL*)**.** The regret of FTL is $\Omega(T)$, which can be achieved by a convex feasible set $[0, 1]$, a sequence of linear loss $f_t(w) = z_t w$, where

$$z_t = \begin{cases} -0.5, & \text{if } t = 1, \\ 1, & \text{elif } t \text{ is even}, \\ -1, & \text{elif } t \text{ is odd}, \end{cases}$$

and pick $u = 0$. We can observe the instability: decisions are alternating between the boundaries. This happens because linear function has no curvature (or has linearity), allowing that

- Upshape/Downshape can be switched easily, which destabilizes the decisions,
- Optima are always on the boundary, keeping a non-decreasing (fixed) distance between decisions.

Example 3.1.2 (*Success of FTL*)**.** Consider an online quadratic optimization problem with $\mathcal{C} = \mathbb{R}^n$ and $f_t(\boldsymbol{w}) = \frac{1}{2}\|\boldsymbol{w} - \boldsymbol{z}_t\|^2$. Then FTL enjoys $\mathcal{O}(L^2 \log T)$ regret bound, where $L = \max_t \|\boldsymbol{z}_t\|$. For each t, we have

$$\boldsymbol{w}_t = \operatorname*{argmin}_{\boldsymbol{w}} \sum_{i=1}^{t-1} \frac{1}{2}\|\boldsymbol{w} - \boldsymbol{z}_t\|^2 = \frac{1}{t-1}\sum_{i=1}^{t-1} \boldsymbol{z}_i,$$

and hence

$$\boldsymbol{w}_{t+1} = \frac{1}{t}\sum_{i=1}^{t} \boldsymbol{z}_i = (1-\frac{1}{t})\boldsymbol{w}_t + \frac{1}{t}\boldsymbol{z}_t.$$

Then we use the lemma,

$$\begin{aligned} f_t(\boldsymbol{w}_t) - f_t(\boldsymbol{w}_{t+1}) &= \frac{1}{2}\|\boldsymbol{w}_t - \boldsymbol{z}_t\|^2 - \frac{1}{2}\|\boldsymbol{w}_{t+1} - \boldsymbol{z}_t\|^2 \\ &= \frac{1}{2}(1-(1-\frac{1}{t})^2)\|\boldsymbol{w}_{t+1} - \boldsymbol{z}_t\|^2 \\ &= (\frac{1}{t} - \frac{1}{2t^2})^2\|\boldsymbol{w}_{t+1} - \boldsymbol{z}_t\|^2 \\ &\leq \frac{1}{t}\|\boldsymbol{w}_{t+1} - \boldsymbol{z}_t\|^2. \end{aligned}$$

Then for all t, $\|\boldsymbol{w}_t\| = \|\frac{1}{t-1}\sum_{i=1}^{t-1}\boldsymbol{z}_t\| \leq \frac{1}{t-1}\sum_{i=1}^{t-1}\|\boldsymbol{z}_t\| \leq \frac{1}{t-1}\sum_{i=1}^{t-1} L = L$, and hence for all i, t, $\|\boldsymbol{w}_i - \boldsymbol{z}_t\| \leq \|\boldsymbol{w}_i\| + \|\boldsymbol{z}_t\| \leq 2L$. Then after summing up we have

$$\sum_{t=1}^{T}(f_t(\boldsymbol{w}_t) - f_t(\boldsymbol{w}_{t+1})) \leq 4L^2\sum_{t=1}^{T}\frac{1}{t} \leq 4L^2(\log T + 1) \in \mathcal{O}(L^2\log T),$$

since $\sum_{t=1}^{T} 1/t \leq \log T + 1$ for all t, and by lemma we completes the proof.

Remark 3.1.2. Consider the following questions:

1. *FTL fully trusts the cumulative losses from the previous rounds, while the new loss is adversarially chosen. However, from lemma we know that the regret will be small if the decisions by FTL are stable, despite the adversarial setting. How does FTL with stability compete the adversary without "chasing" the loss?*

First of all, we're competing with the fixed decisions, not the adversary. Therefore, we don't need to care about the loss as long as our decisions are closed to the best fixed decision. From the previous example, solving $\operatorname{argmin}_{\boldsymbol{w}\in\mathcal{C}}\sum_{i=1}^{t} f_t(\boldsymbol{w})$ are pretty good at chasing the best fixed decision in the quadratic case.

2. *From the lemma we know if the decisions by FTL are stable, then they stay quite closed to the best fixed decisions. Why intuitively?*

By the definition of FTL we have,

$$\{\boldsymbol{w}_2, \boldsymbol{w}_3, ...\} = \{\boldsymbol{u}_1^*, \boldsymbol{u}_2^*, ...\}.$$

In this case, if $\{\boldsymbol{w}_t\}_{t=1}^{T}$ are stable, then $\boldsymbol{w}_i \approx \boldsymbol{u}_t^*$ for all i, t. In particular, the best fixed decision throughout the whole time horizon $\boldsymbol{u}_T^*$ is closed to $\{\boldsymbol{w}_t\}_{t=1}^{T}$. Essentially we're saying that *provided that the decisions by FTL are stable, then* $\operatorname{argmin}_{\boldsymbol{w}\in\mathcal{C}}\sum_{i=1}^{t} f_t(\boldsymbol{w})$ *is very good at chasing the best fixed decision.*

3. *How do we achieve stability?*

FTL itself takes no precaution for stability, and we showed that linear loss can indeed ruin stability. Therefore, more constraints will be needed on the loss. Informally speaking, it's sufficient to require the loss to have some curvature s.t.

the whole shape doesn't change much after adding any new loss,

in which case the optimum $\{\boldsymbol{u}_1^*, \boldsymbol{u}_2^*, ...\} = \{\boldsymbol{w}_2, \boldsymbol{w}_3, ...\}$ don't change much, guaranteeing stability. By contrast, the shape of linear function can dramatically change after adding another one.

3.2 Convexity & Lipschitzness Review

3.2.1 Strong Convexity

Definition 3.2.1 (*Strongly Convex Functions*)**.** *A function $f : \mathcal{C} \to \mathbb{R}$ is σ-strongly convex over $\mathcal{C}$ w.r.t. some norm $\|\cdot\|$ if for all $\boldsymbol{w} \in \mathcal{C}$, we have*

$$\forall \boldsymbol{z} \in \partial f(\boldsymbol{w}), \forall \boldsymbol{u} \in \mathcal{C} : f(\boldsymbol{u}) \geq f(\boldsymbol{w}) + \langle \boldsymbol{z}, \boldsymbol{u} - \boldsymbol{z} \rangle + \frac{\sigma}{2}\|\boldsymbol{u} - \boldsymbol{w}\|^2.$$

Remark 3.2.1. The definition indicates the followings:

- f is strongly convex if for all $\boldsymbol{w} \in \mathcal{C}$, there's a quadratic under-estimator tangent at $\boldsymbol{w}$,
- If f is σ-strongly convex, then the distance between f and its linearization at $\boldsymbol{w}_0$ is at least $\frac{\sigma}{2}\|\boldsymbol{w} - \boldsymbol{w}_0\|^2$.

Lemma 3.2.1. *Let $f : \mathcal{C} \to \mathbb{R}$ be a σ-strongly convex over $\mathcal{C}$ w.r.t. some norm $\|\cdot\|$, and $\boldsymbol{w}^* = \operatorname{argmin}_{\boldsymbol{w} \in \mathcal{C}} f(\boldsymbol{w})$. Then for all $\boldsymbol{w} \in \mathcal{C}$, we have*

$$f(\boldsymbol{w}) - f(\boldsymbol{w}^*) \geq \frac{\sigma}{2}\|\boldsymbol{w} - \boldsymbol{w}^*\|^2.$$

Proof. For simplicity, we assume f is differentiable. By strong convexity,

$$f(\boldsymbol{w}) - \boldsymbol{w}^* \geq \langle \nabla f(\boldsymbol{w}^*), \boldsymbol{w} - \boldsymbol{w}^* \rangle + \frac{\sigma}{2}\|\boldsymbol{w} - \boldsymbol{w}^*\|^2.$$

By convexity, $\langle \nabla f(\boldsymbol{w}^*), f(\boldsymbol{w}) - \boldsymbol{w}^* \rangle \geq 0$, and this completes the proof. This lemma gives a bit more information than $f(\boldsymbol{w}) - f(\boldsymbol{w}^*) \geq 0$. $\square$

Lemma 3.2.2. *Let $f : \mathcal{C} \to \mathbb{R}$ be σ-strongly convex, and $g : \mathcal{C} \to \mathbb{R}$ be convex. Then $h = f + g$ is also σ-strongly convex.*

Proof. By definition, we have for $\boldsymbol{w}, \boldsymbol{u} \in \mathcal{C}$,

$$\begin{aligned} f(\boldsymbol{u}) &\geq f(\boldsymbol{w}) + \langle \boldsymbol{z}, \boldsymbol{u} - \boldsymbol{w} \rangle + \frac{\sigma}{2}\|\boldsymbol{u} - \boldsymbol{w}\|^2, \\ g(\boldsymbol{u}) &\geq g(\boldsymbol{w}) + \langle \boldsymbol{p}, \boldsymbol{u} - \boldsymbol{w} \rangle, \end{aligned}$$

and by adding we have

$$\underbrace{f(\boldsymbol{u}) + g(\boldsymbol{u})}_{h(\boldsymbol{u})} \geq \underbrace{f(\boldsymbol{w}) + g(\boldsymbol{w})}_{h(\boldsymbol{w})} + \langle \underbrace{\boldsymbol{z} + \boldsymbol{p}}_{\in \partial h(\boldsymbol{w})}, \boldsymbol{u} - \boldsymbol{w} \rangle + \frac{\sigma}{2}\|\boldsymbol{u} - \boldsymbol{w}\|^2,$$

so $h = f + g$ is σ-strongly convex. $\square$

3.2.2 Lipschitzness

Definition 3.2.2 (*Lipschitz Continuity*)**.** *A function $f : \mathcal{X} \to \mathbb{R}$ is L-Lipschitz continuous w.r.t. some norm $\|\cdot\|$ if for all $\boldsymbol{w}, \boldsymbol{u} \in \mathcal{X}$, we have*

$$|f(\boldsymbol{w}) - f(\boldsymbol{u})| \leq L\|\boldsymbol{w} - \boldsymbol{u}\|.$$

Lemma 3.2.3 (*Bounded Gradient*)**.** *Assume $f : \mathcal{C} \to \mathbb{R}$ is convex and L-Lipschitz w.r.t. some norm $\|\cdot\|$, then for all $\boldsymbol{w} \in \mathcal{C}$, we have*

$$\|\nabla f(\boldsymbol{w})\| \leq L.$$

Proof. For simplicity, we assume $\boldsymbol{w}$ is in the interior of $\mathcal{C}$, then

$$\|\nabla f(\boldsymbol{w})\| = \sqrt{\langle \nabla f(\boldsymbol{w}), \nabla f(\boldsymbol{w}) \rangle} = \sqrt{\langle \nabla f(\boldsymbol{w}), \frac{1}{\eta}(\boldsymbol{u} - \boldsymbol{w}) \rangle},$$

where $\boldsymbol{u} = \boldsymbol{w} + \eta \nabla f(\boldsymbol{w}) \in \mathcal{C}$ for some small $\eta > 0$, then by convexity and lipschitzness,

$$\sqrt{\langle \nabla f(\boldsymbol{w}), \frac{1}{\eta}(\boldsymbol{u} - \boldsymbol{w}) \rangle} \leq \sqrt{|f(\frac{1}{\eta}\boldsymbol{u}) - f(\frac{1}{\eta}\boldsymbol{w})|} \leq \sqrt{\frac{L}{\eta}\|\boldsymbol{u} - \boldsymbol{w}\|} = \sqrt{L\|\nabla f(\boldsymbol{w})\|},$$

that is, $\|\nabla f(\boldsymbol{w})\| \leq \sqrt{L\|\nabla f(\boldsymbol{w})\|} \Rightarrow \|\nabla f(\boldsymbol{w})\| \leq L$. $\square$

3.3 Follow the Regularized Leader

The regret bound for FTL can be high if the decisions are unstable. Instead of directly minimizing the previous t losses, we instead minimize a slightly different expression with an extra regularizer R, and this gives the *Follow-the-Regularized-Leader* (*FTRL*):

$$\boldsymbol{w}_{t+1} \coloneqq \operatorname*{argmin}_{\boldsymbol{w}\in\mathcal{C}} \sum_{i=1}^{t} f_i(\boldsymbol{w}) + R(\boldsymbol{w}).$$

Lemma 3.3.1 (*Regret Bound of FTRL*)**.** *Let $\{\boldsymbol{w}_1, \boldsymbol{w}_2, ...\}$ be the decisions generated by FTL, then for all $\boldsymbol{u} \in \mathcal{C}$, we have*

$$\operatorname{Regret}_T(\boldsymbol{u}) \leq R(\boldsymbol{u}) - \min_{\boldsymbol{w}\in\mathcal{C}} R(\boldsymbol{w}) + \sum_{t=1}^{T}(f_t(\boldsymbol{w}_t) - f_t(\boldsymbol{w}_{t+1})).$$

Proof. Running FTRL on $\{f_1, f_2, ...\}$ is equivalent to running FTL on $\{f_0, f_1, ...\}$, where $f_0 = R$. Therefore by lemma, we have for all $\boldsymbol{u} \in \mathcal{C}$,

$$\sum_{t=0}^{T}(f_t(\boldsymbol{w}_t) - f_t(\boldsymbol{u})) \leq \sum_{t=0}^{T}(f_t(\boldsymbol{w}_t) - f_t(\boldsymbol{w}_{t+1})),$$

$$f_0(\boldsymbol{w}_0) - f_0(\boldsymbol{u}) + \sum_{t=1}^{T}(f_t(\boldsymbol{w}_t) - f_t(\boldsymbol{u})) \leq f_0(\boldsymbol{w}_0) - f_0(\boldsymbol{w}_1) + \sum_{t=1}^{T}(f_t(\boldsymbol{w}_t) - f_t(\boldsymbol{w}_{t+1})).$$

Then cancelling $f_0(\boldsymbol{w}_0)$ out, replacing f_0 by R, $\sum_{t=1}^{T}(f_t(\boldsymbol{w}_t) - f_t(\boldsymbol{u}))$ by $\operatorname{Regret}_T(\boldsymbol{u})$, and $\boldsymbol{w}_1 \coloneqq \operatorname{argmin}_{\boldsymbol{w}\in\mathcal{C}} R(\boldsymbol{w})$ complete the proof. ☐

Remark 3.3.1. To understand the regret bound

$$\operatorname{Regret}_T \leq \underbrace{\max_{\boldsymbol{w}\in\mathcal{C}} R(\boldsymbol{w}) - \min_{\boldsymbol{w}\in\mathcal{C}} R(\boldsymbol{w})}_{\text{penalty}} + \underbrace{\sum_{t=1}^{T}(f_t(\boldsymbol{w}_t) - f_t(\boldsymbol{w}_{t+1}))}_{\text{stability}},$$

- Penalty: Penalize for not directly minimizing the cumulative loss, which we know is good at chasing the best fixed decision,
- Stability: Since we're still using FTL-based algorithm, stability still matters.

Lemma 3.3.2 (*Regret Bound of FTRL*)**.** *Assume f_t is L_t-Lipschitz for all t, and R is a σ-strongly-convex function. Then FTRL enjoys the following regret bound:*

$$\operatorname{Regret}_T \leq B_R + \frac{TL^2}{\sigma},$$

where $B_R = \max_{\boldsymbol{w}\in\mathcal{C}} R(\boldsymbol{w}) - \min_{\boldsymbol{w}\in\mathcal{C}} R(\boldsymbol{w})$ and $T = \max_t\{L_t\}$.

Proof. The B_R part is trivial. We'll first prove $\|\boldsymbol{w}_t - \boldsymbol{w}_{t+1}\| \leq L_t/\sigma$ for future use, and equivalently,

$$\|\boldsymbol{w}_t - \boldsymbol{w}_{t+1}\| \leq \frac{L_t}{\sigma} \Leftrightarrow \sigma\|\boldsymbol{w}_t - \boldsymbol{w}_{t+1}\|^2 \leq L_t\|\boldsymbol{w}_t - \boldsymbol{w}_{t+1}\|.$$

By lipschitzness, it's sufficient to prove $\sigma\|\boldsymbol{w}_t - \boldsymbol{w}_{t+1}\|^2 \leq f_t(\boldsymbol{w}_t) - f_t(\boldsymbol{w}_{t+1})$. Let $F_t(\boldsymbol{w}) = \operatorname{argmin}_{\boldsymbol{w}\in\mathcal{C}} \sum_{i=1}^{t} f_i(\boldsymbol{w}) + R(\boldsymbol{w})$, then we have

$$\boldsymbol{w}_t = \operatorname*{argmin}_{\boldsymbol{w}\in\mathcal{C}} F_{t-1}(\boldsymbol{w}),\ \boldsymbol{w}_{t+1} = \operatorname*{argmin}_{\boldsymbol{w}\in\mathcal{C}} F_t(\boldsymbol{w}),$$

and hence by strong convexity,

$$F_{t-1}(\boldsymbol{w}_{t+1}) - F_{t-1}(\boldsymbol{w}_t) \geq \frac{\sigma}{2}\|\boldsymbol{w}_t - \boldsymbol{w}_{t+1}\|^2,\ F_t(\boldsymbol{w}_t) - F_t(\boldsymbol{w}_{t+1}) \geq \frac{\sigma}{2}\|\boldsymbol{w}_t - \boldsymbol{w}_{t+1}\|^2,$$

and after summing up we finish the proof for the intermediate result. Then

$$\sum_{t=1}^{T}(f_t(\boldsymbol{w}_t) - f_t(\boldsymbol{w}_{t+1})) \le \sum_{t=1}^{T} L_t \|\boldsymbol{w}_t - \boldsymbol{w}_{t+1}\| \le \sum_{t=1}^{T} \frac{L_t^2}{\sigma} \le \frac{TL^2}{\sigma},$$

and this completes the whole proof. □

Remark 3.3.2. To achieve a sub-linear regret, we need to additionally bound B_R, or essentially bound the range of $\boldsymbol{w}$ (we'll see this in the next section). Now we analyze how the assumptions potentially guarantee stability and lead to sub-linear regret intuitively:

- Bounded Domain: Restricting the range of $\boldsymbol{w}$ to fully bound the penalty term, and partially guaranteeing stability,
- Lipschitzness of f_t (bounded ∇f_t): Restricting the range of $\boldsymbol{w}$ isn't enough, because it's possible that the gradient is so large that taking a small step from $\boldsymbol{w}_t$ to $\boldsymbol{w}_{t+1}$ leads to large $f(\boldsymbol{w}_t) - f(\boldsymbol{w}_{t+1})$.

Remark 3.3.3. Consider the following questions:

1. *What motivates us to add a regularizer?*

Previously, we mentioned that FTL is good at solving curved loss. Here, without further assumptions on the input loss, we cheat to add an irrelevant regularizer to bring the curvature to the problem to have stability (this is our ultimate goal), which mimics the case where the losses are curved.

2. *Now that we have stability, why do we achieve $\mathcal{O}(\sqrt{T})$ only but no further to $\mathcal{O}(\log T)$, ignoring that OCO has the $\Omega(\sqrt{T})$ lower bound?*

Because the mimic is imperfect: although we have stability, we're now staying a little further from the best fixed decision for not directly solving $\operatorname{argmin}_{\boldsymbol{w}\in\mathcal{C}} \sum_{i=1}^{t} f_t(\boldsymbol{w})$, which is known good at chasing the best fixed decision given stability. This distance essentially leads to the additional penalty term of the regret bound.

3. *Why strong convexity?*

Informally speaking, strong convexity satisfies the requirement: *the whole shape doesn't change much after adding any new loss.* Besides, we know that *if f is σ-strongly convex and g is convex, then $f+g$ is σ-strongly convex*, i.e., the curvature never decays.

4. *Can we choose any strongly convex regularizer?*

No, because there's a trade-off here: if it's too strong, then the stability term will be small, while the penalty term will be large (small range of $\boldsymbol{w}$ will lead to large B_R); if it's too weak, the penalty term will be small, but it has less effect in stability. Besides, we should also choose the one that fully exploits the geometry of the problem.

3.4 Online Gradient Descent

3.4.1 Unconstrained Domain

The *Online Gradient Descent* algorithm is an instance of FTRL. Consider an OLO with $f_t(\boldsymbol{w}) = \langle \boldsymbol{z}_t, \boldsymbol{w}_t \rangle$ and $\mathcal{C} = \mathbb{R}^n$, we're to instantiate FTRL with $R(\boldsymbol{w}) = \frac{1}{2\eta}||\boldsymbol{w}||_2^2$ for some $\eta > 0$, so

$$\boldsymbol{w}_{t+1} \coloneqq \operatorname*{argmin}_{\boldsymbol{w}} \sum_{i=1}^{t} \langle \boldsymbol{z}_t, \boldsymbol{w} \rangle + \frac{1}{2\eta}||\boldsymbol{w}||_2^2.$$

We then have $\boldsymbol{w}_{t+1} = -\eta \sum_{i=1}^{t} \boldsymbol{z}_i = -\eta \sum_{i=1}^{t-1} \boldsymbol{z}_i - \eta \boldsymbol{z}_t = \boldsymbol{w}_t - \eta \boldsymbol{z}_t$, where $\boldsymbol{z}_t = \nabla f_t(\boldsymbol{w})$ for all $\boldsymbol{w}$, so in particular we have $\boldsymbol{z}_t = \nabla f_t(\boldsymbol{w}_t)$. Rewriting this yields OGD:

$$\boldsymbol{w}_{t+1} \coloneqq \boldsymbol{w}_t - \eta \nabla f_t(\boldsymbol{w}_t).$$

Lemma 3.4.1 (*Regret Bound of OGD*)**.** *Assume f_t is L_t-Lipschitz for all t, and $B = \max_{\boldsymbol{w}} \|\boldsymbol{w}\|_2$, then OGD with $\eta = \frac{B}{L\sqrt{2T}}$ where $L = \max_t\{L_t\}$ enjoys the following regret bound:*

$$\text{Regret}_T \leq BL\sqrt{2T} \in \mathcal{O}(BL\sqrt{T}).$$

Proof. For general FTRL we have

$$\text{Regret}_T \leq B_R + \frac{TL^2}{\sigma} = \max_{\boldsymbol{w}} \frac{1}{2\eta}\|\boldsymbol{w}\|_2^2 - \min_{\boldsymbol{w}} \frac{1}{2\eta}\|\boldsymbol{w}\|_2^2 + \frac{TL^2}{\sigma}$$

where $\sigma = 1/\eta$, since $1/2\eta\|\boldsymbol{w}\|_2^2$ is $1/\eta$-strongly convex. Then

$$\text{Regret}_T \leq \frac{1}{2\eta}B^2 + \eta T L^2,$$

and plugging in η completes the proof (η is chosen s.t. the whole expression is minimized, and this happens when the two terms are equal). □

Remark 3.4.1. For each round of decision, OGD fully trusts the previous loss only (like FTL fully trusts the cumulative loss), but still enjoys sub-linear regret although the new loss is adversarially chosen. This is due to the assumption:

$$f_t \text{ is } L_t\text{-Lipschitz} \Rightarrow \|\nabla f_t(\boldsymbol{w})\|_2 \leq L_t$$

for all $\boldsymbol{w} \in \mathcal{C}$, that is, the gradient doesn't change much throughout the horizon, which makes $\nabla f_t(\boldsymbol{w})$ a reasonable estimator of ∇f_{t+1}. By this, we notice the core difference between OGD and FTL-based algorithm:

- OGD: "Approximates" ∇f_{t+1} by ∇f_t, so it tries to compete with the adversary directly, but only achieves sub-linear regret when compares with the best fixed decision,
- FTRL: Takes no action in competing with the adversary, and only aims at staying closed to the best fixed decision.

This observation is counter-intuitive, since OGD is fully derived from FTRL.

3.4.2 Constrained Domain

If $\mathcal{C}$ is constrained, we need to additionally perform a projection $\boldsymbol{w}_t = \Pi_{\mathcal{C}}(\boldsymbol{u}_t) = \operatorname{argmin}_{\boldsymbol{w} \in \mathcal{C}} \|\boldsymbol{w} - \boldsymbol{u}_t\|_2$. There're the *lazy* and *agile* version of the projected OGD respectively:

$$\boldsymbol{u}_{t+1} = \boldsymbol{u}_t - \eta \nabla f_t(\boldsymbol{w}_t),\ \boldsymbol{w}_{t+1} = \Pi_{\mathcal{C}}(\boldsymbol{u}_t),$$
$$\boldsymbol{u}_{t+1} = \boldsymbol{w}_t - \eta \nabla f_t(\boldsymbol{w}_t),\ \boldsymbol{w}_{t+1} = \Pi_{\mathcal{C}}(\boldsymbol{u}_t).$$

3.5 Online Mirror Descent

3.5.1 Bregman Divergence Review

Definition 3.5.1 (*Bregman Divergence*)**.** *Let* $\phi : \mathcal{C} \to \mathbb{R}$ *be a strictly convex and differentiable function. The Bregman divergence of* ϕ*,* $D_\phi : \mathcal{C} \times \mathcal{C} \to \mathbb{R}$ *is a distance measure of two points on* ϕ*, where*

$$D_\phi(\boldsymbol{w} \| \boldsymbol{u}) = \phi(\boldsymbol{w}) - \underbrace{(\phi(\boldsymbol{u}) + \langle \nabla \phi(\boldsymbol{u}), \boldsymbol{w} - \boldsymbol{u} \rangle)}_{L_{\boldsymbol{u}}(\boldsymbol{w})}.$$

Remark 3.5.1. Euclidean Distance $\frac{1}{2}\|\boldsymbol{w} - \boldsymbol{u}\|^2$ is an instance of Bregman distance, where $\phi(\boldsymbol{w}) = \frac{1}{2}\|\boldsymbol{w}\|^2$.

3.5.2 Online Mirror Descent

Recall that the update method of OGD is $\boldsymbol{w}_{t+1} = \boldsymbol{w}_t - \eta \nabla f_t(\boldsymbol{w}_t)$, which is equivalent to the following equation:

$$\boldsymbol{w}_{t+1} = \operatorname*{argmin}_{\boldsymbol{w} \in \mathcal{C}} f_t(\boldsymbol{w}_t) + \langle \nabla f_t(\boldsymbol{w}_t), \boldsymbol{w} - \boldsymbol{w}_t \rangle + \underbrace{\frac{1}{2\eta}\|\boldsymbol{w} - \boldsymbol{w}_t\|^2}_{\text{Euclidean Distance}}.$$

If we replace Euclidean distance by the more general Bregman divergence $D_\phi(\boldsymbol{w}\|\boldsymbol{w}_t)$:

$$\boldsymbol{w}_{t+1} = \operatorname*{argmin}_{\boldsymbol{w} \in \mathcal{C}} f_t(\boldsymbol{w}_t) + \langle \nabla f_t(\boldsymbol{w}_t), \boldsymbol{w} - \boldsymbol{w}_t \rangle + \frac{1}{\eta} D_\phi(\boldsymbol{w}\|\boldsymbol{w}_t),$$

we yield the following formula, which is the update method for *Online Mirror Descent* (*OMD*)

$$\boldsymbol{w}_{t+1} = (\nabla \phi)^{-1}(\nabla \phi(\boldsymbol{w}_t) - \eta \nabla f_t(\boldsymbol{w}_t)),$$

where $\nabla \phi$ is known as the mirror map.

3.6 Online Newton Step

4 Experts Problem

5 Lower Bound of OCO

6 References