Basics of Online Convex Optimization

Yanze Song

November 16, 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 at each round.

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 $\mathcal{A}$,*

$$\underset{C}{\mathrm{Regret}}(\mathcal{A}) \leq \underset{L}{\mathrm{Regret}}(\mathcal{A}),$$

provided that the classes of convex and linear functions, C and L, are equivalent.

Proof. Let $f_t : \mathcal{C} \to \mathbb{R}$ be a convex function, $\boldsymbol{w}_t \in \mathcal{C}$ be our t^{th} decision, and $\boldsymbol{g}_t \in \partial f_t(\boldsymbol{w}_t)$. Then for all $\boldsymbol{w} \in \mathcal{C}$, by convexity we have

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

and if we sum over t, we have

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

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

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

which completes the proof. $\square$

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 $\boldsymbol{w} \in \mathcal{C}$,

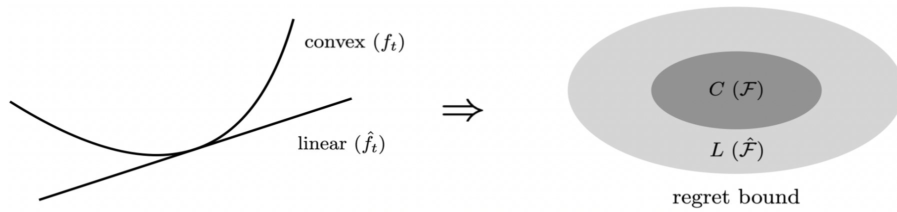
or slightly more strictly, $\hat{f}_t$ should be s.t. for all $\boldsymbol{w} \in \mathcal{C}$,

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

Then we will have

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

In our case where $\mathcal{F} = C$, we choose $\hat{\mathcal{F}} = L$ 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 total loss of all previous rounds. 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*

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

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

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

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

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

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.

Consider the following question:

> *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 with full power. Therefore, we don't need to chase the loss as long as our decisions are closed to the best fixed decision. How do we achieve stability? FTL itself takes no precaution to guarantee stability, and we showed that linear loss can indeed ruin stability. Therefore, more constraints will be needed on the loss. It's sufficient to require the loss defined 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

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. □

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. □

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 $\boldsymbol{w}, \boldsymbol{u} \in X$, 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*

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

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 $\mathrm{Regret}_T(\boldsymbol{u})$, 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

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

- Penalty: Penalize for not directly minimizing the cumulative loss,
- Stability: Since we're still using FTL-based algorithm.

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

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

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. We'll first prove $\|\boldsymbol{w}_t - \boldsymbol{w}_{t+1}\| \leq L_t/\sigma$, and equivalently,

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

Therefore, it's sufficient to prove $\sigma\|\boldsymbol{w}_t - \boldsymbol{w}_{t+1}\|^2 \leq f_t(\boldsymbol{w}_t) - f_t(\boldsymbol{w}_{t+1})$ by lipschitzness. 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 $\boldsymbol{w}_t = \operatorname{argmin}_{\boldsymbol{w}\in\mathcal{C}} F_{t-1}(\boldsymbol{w})$ and $\boldsymbol{w}_{t+1} = \operatorname{argmin}_{\boldsymbol{w}\in\mathcal{C}} F_t(\boldsymbol{w})$.

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 by summing up we finish the proof for the intermediate result. Then

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

and this completes the whole proof. □

Remark 3.3.2. Under the online context, we dedicate to achieve robust results with minimal

3.4 Online Gradient Descent

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 are L_t-Lipschitz for all t, and $B = \max_{\boldsymbol{w}} \|\boldsymbol{w}\|$, then OGD with $\eta = \frac{B}{L\sqrt{2T}}$ where $L = \max_t L_t$ enjoys the following regret bound:*

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

Proof. For general FTRL we have

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

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

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

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). $\square$

3.5 Online Mirror Descent

3.6 Online Newton Step

4 Experts Problem

5 Lower Bound of OCO

6 Beyond General Convexity Assumptions

7 References