Scribe: Minghan Wu, Yin Lei
May 16, 2024
Editor/TA: Marialena Sfyraki

Lecture 14: Online Convex Optimization (Continued)

1 Example when FTL Fails

Let the decision space be $z = [-1,1] \subseteq \mathbb{R}$. Let $\ell_t(z) = c_t z \in \mathbb{R}$ where $c_t = \begin{cases} -0.5 & t = 1 \\ 1 & t \text{ is even} \\ -1 & t \text{ is odd} \end{cases}$.

Analysis

1st Round: Let $z_1 = \theta \in [-1,1]$, $c_1 = -0.5$. Then:

$$\ell_1(z_1) = c_1 z_1 = -0.5 \cdot \theta$$

2nd Round:

$$\begin{aligned} z_2 &= argmin_{z\in[-1,1]} \ell_1(z) \\ &= argmin_{z\in[-1,1]}(-0.5z) \\ &= 1 \end{aligned}$$

$$\ell_2(z_2) = c_2 z_2 = 1 \cdot 1 = 1$$

3rd Round:

$$\begin{aligned} z_3 &= argmin_{z\in[-1,1]}[\ell_1(z) + \ell_2(z)] \\ &= argmin_{z\in[-1,1]}(-0.5 + 1)z \\ &= -1 \end{aligned}$$

$$\ell_3(z_3) = c_3 z_3 = (-1) \cdot (-1) = 1$$

4th Round:

$$\begin{aligned} z_4 &= argmin_{z\in[-1,1]}[\ell_1(z) + \ell_2(z) + \ell_3(z)] \\ &= argmin_{z\in[-1,1]}(-0.5 + 1 - 1)z \\ &= 1 \end{aligned}$$

$$\ell_4(z_4) = c_4 z_4 = 1 \cdot 1 = 1$$

To summarize, the the loss of FTL is

$$\sum_{t=1}^{T} c_t z_t = -0.5 z_1 + (T-1)\cdot 1 = \Omega(T).$$

On the other hand, in hindsight, if we fix committing $z_* = 0$, then

$$Regret_T(z_*) = \underbrace{\sum_{t=1}^{T} \ell_t(z_t)}_{=\Omega(T)} - \underbrace{\sum_{t=1}^{T} \ell_t(z_*)}_{=0} = \sum_{t=1}^{T} \ell_t(z_t) = \Omega(T).$$

2 FTRL

To fix the problem, we need to introduce a regularizer.

2.1 Algorithm: Follow the Regularized Leader (FTRL)

At round t, play

$$z_t = \arg\min_{z\in Z} \sum_{s=1}^{t-1} \ell_s(z) + R(z),$$

where $R(z) : Z \to \mathbb{R}$ is strongly convex.

Remark: FTRL on $\ell_1, \dots, \ell_T$ is equivalent to FTL on $\ell_0 = R, \ell_1, \dots, \ell_T$.
For FTRL we have

$$\tilde{z}_{T+1} = \arg\min_{z} \sum_{t=1}^{T} \ell_t(z) + R(z).$$

For FTL we have

$$z_{T+1} = \arg\min_{z} \sum_{t=0}^{T} \ell_t(z),$$

where we set $\ell_0(z) = R(z)$.

2.2 OGD and FTRL

Algorithm: Online Gradient Descent (OGD)

Protocol/Setting

Init $z_1 = 0$
for $t = 1, 2, \ldots$ **do**
  $z_{t+1} = z_t - \eta c_t$, where $c_t \in \partial \ell_t(z_t)$.
end for

Consider

$$\ell_t(z) = \langle z, c_t \rangle \quad \text{and} \quad R(z) = \frac{1}{2\eta}\|z\|_2^2$$

and

$$z_t = \arg\min_{z \in Z} \underbrace{\sum_{s=1}^{t-1} \ell_s(z) + R(z)}_{:=\Phi(z)}$$

Then, we have

$$\begin{aligned}
\nabla \Phi(z_t) = 0 &\Leftrightarrow \sum_{s=1}^{t-1} c_s + \frac{1}{2\eta} 2z = 0 \\
&\Leftrightarrow z_t = -\eta \sum_{s=1}^{t-1} c_s \\
&\Leftrightarrow z_t = z_{t-1} - \eta c_{t-1} \\
&\Leftrightarrow z_t = z_{t-1} - \eta \nabla \ell_{t-1}(z_{t-1}) \ , \quad \text{since } \ell_t(z_t) = c_t.
\end{aligned}$$

Observe that this is OGD.

2.3 Regret of FTRL

In the previous lecture, we proved the following lemma:

Lemma 1. *Let $z_0, z_1, z_2, \ldots,$ be the sequence of points generated by FTL. Then, for any benchmark $z_* \in Z$,*

$$\mathit{Regret}_{T+1}(z_*) = \sum_{t=0}^{T} \ell_t(z_t) - \sum_{t=0}^{T} \ell_t(z_*) \leq \sum_{t=0}^{T} \ell_t(z_t) - \sum_{t=0}^{T} \ell_t(z_{t+1}).$$

Lemma 2. *Let $z_1, z_2, \ldots,$ be the sequence of points generated by FTRL. Then, for any benchmark $z_* \in Z$,*

$$\begin{aligned} Regret_T(z_*) &= \sum_{t=0}^{T} \ell_t(z_t) - \sum_{t=0}^{T} \ell_t(z_*) \\ &\leq R(z_*) - R(z_1) + \sum_{t=1}^{T} \ell_t(z_t) - \sum_{t=1}^{T} \ell_t(z_{t+1}). \end{aligned}$$

Proof. From Lemma 1 we have that:

$$\sum_{t=0}^{T} \ell_t(z_t) - \sum_{t=0}^{T} \ell_t(z_*) \leq \sum_{t=0}^{T} \ell_t(z_t) - \sum_{t=0}^{T} \ell_t(z_{t+1})$$

Expanding the left-hand side:

$$\sum_{t=0}^{T} \ell_t(z_t) - \sum_{t=0}^{T} \ell_t(z_*) = \left(R(z_0) + \sum_{t=1}^{T} \ell_t(z_t) \right) - \left(R(z_*) + \sum_{t=1}^{T} \ell_t(z_*) \right).$$

Expanding the right-hand side:

$$\sum_{t=0}^{T} \ell_t(z_t) - \sum_{t=0}^{T} \ell_t(z_{t+1}) = \left(R(z_0) + \sum_{t=1}^{T} \ell_t(z_t) \right) - \left(R(z_1) + \sum_{t=1}^{T} \ell_t(z_{t+1}) \right).$$

Note that the two sides share the first term. Therefore, we have that

$$\begin{aligned} &\sum_{t=0}^{T} \ell_t(z_t) - \sum_{t=0}^{T} \ell_t(z_*) \leq \sum_{t=0}^{T} \ell_t(z_t) - \sum_{t=0}^{T} \ell_t(z_{t+1}) \\ \Leftrightarrow &\sum_{t=1}^{T} \ell_t(z_t) - \sum_{t=1}^{T} \ell_t(z_*) \leq R(z_*) - R(z_1) + \sum_{t=1}^{T} \ell_t(z_t) - \sum_{t=1}^{T} \ell_t(z_{t+1}). \end{aligned}$$

□

The following theorem gives an upper bound of $Regret_T(z_*)$ with respect to z_*.

Theorem 1. *Suppose $\ell_t(z) = \langle z, c_t \rangle$. Let $R(z) = \frac{1}{2\eta}\|z\|_2^2$. Then FTRL has*

$$Regret_T(z_*) \leq \frac{1}{2\eta}\|z_*\|_2^2 + \eta \sum_{t=1}^{T} \|c_t\|_2^2$$

for any z_*.

Proof. From lemma 2 we have,

$$\begin{aligned}
Regret_T(z_*) &\leq R(z_*) - R(z_1) + \sum_{t=1}^{T} \ell_t(z_t) - \sum_{t=1}^{T} \ell_t(z_{t+1}) \\
&= \frac{1}{2\eta}\|z_*\|_2^2 - \frac{1}{2\eta}\|z_1\|_2^2 + \sum_{t=1}^{T}\langle z_t, c_t\rangle - \sum_{t=1}^{T}\langle z_{t+1}, c_t\rangle \\
&\leq \frac{1}{2\eta}\|z_*\|_2^2 + \sum_{t=1}^{T}\langle z_t - z_{t+1}, c_t\rangle \\
&= \frac{1}{2\eta}\|z_*\|_2^2 + \eta\sum_{t=1}^{T}\|c_t\|_2^2,
\end{aligned}$$

where the last equality was obtained from the fact $z_{t+1} = z_t - \eta\nabla\ell_t(z_t) = z_t - \eta c_t$. □

Corollary 1. *If additionally to Theorem 1 we have* $\|z_*\|_2 \leq D$ *and* $\|c_t\|_2 \leq G$*, then by setting* $\eta = \frac{D}{G\sqrt{2T}}$*, we have*

$$Regret_T(z_*) \leq DG\sqrt{2T}$$

Proof. From Theorem 1, if $\|z_*\|_2 \leq D$ and $\|c_t\|_2 \leq G$, then

$$\begin{aligned}
Regret_T(z_*) &\leq \frac{1}{2\eta}\|z_*\|_2^2 + \sum_{t=1}^{T}\eta\|c_t\|_2^2 \\
&= \frac{1}{2\eta}D^2 + \eta T G^2
\end{aligned}$$

By setting $\eta = \frac{D}{G\sqrt{2T}}$, we have

$$\begin{aligned}
Regret_T(z_*) &\leq \frac{G\sqrt{2T}}{2D}D^2 + \frac{D}{G\sqrt{2T}}TG^2 \\
&= GD\sqrt{2T}
\end{aligned}$$

□

2.4 Online Linear Optimization (OLO)

Below is a statement of the Online Linear Optimization Setting:

Protocol/Setting

1: **for** $t = 1, 2, \ldots$ **do**
2: Commit a point z_t with its convex decision space $Z \subset R^d$.
3: Receive a loss function $\ell_t(z) = c_t^\top z_t$ and incurs a loss $\ell_t(z_t)$.
4: **end for**

We consider the problem of minimizing regret, that is

$$Regret_T(z_*) = \sum_{t=1}^{T} \langle c_t, z_t \rangle - \sum_{t=1}^{T} \langle c_t, z_* \rangle,$$

where $z_* \in Z$ is any fixed benchmark in Z.
Below is a statement of the Online Convex Optimization Setting:

Protocol/Setting

1: **for** $t = 1, 2, \dots$ **do**
2: Commit a point z_t with its convex decision space $Z \subset R^d$.
3: Receive a loss function $\ell_t(\cdot) : Z \to \mathbb{R}$ and incurs a loss $\ell_t(z_t)$.
4: **end for**

We consider the problem of minimizing regret, that is

$$Regret_T(z_*) = \sum_{t=1}^{T} \ell_t(z_t) - \sum_{t=1}^{T} \ell_t(z_*),$$

where $z_* \in Z$ is any fixed benchmark in Z.

One way to solve a Online Convex Optimization problem is to reduce it to an Online Linear Optimization problem. Given an algorithm, we can get an upper bound for OCO regret from the regret of the reduced OLO problem.

First, recall that g_x is a subgradient of $f : Z \to \mathbb{R}$ at $x \in Z$ if

$$f(y) \geq f(x) + \langle g_x, y - x \rangle \ , \ \forall y \in Z.$$

If we let

$$\begin{aligned} &f(\cdot) \leftarrow \ell_t(\cdot) \\ &g_x \leftarrow \nabla \ell_t(\cdot) := c_t \\ &y \leftarrow z_* \\ &x \leftarrow z_t, \end{aligned}$$

then, by the 1st-order characterization of convexity,

$$\underbrace{\ell_t(z_t) - \ell_t(z_*)}_{\text{per-round regret of OCO}} \quad \leq \quad \underbrace{\langle z_t - z_*, c_t \rangle}_{\text{per-round regret of OLO}} \quad ,$$

where c_t is a subgradient of $\ell_t(\cdot)$ at z_t.

Remark: If ℓ_t is differentiable, then $c_t = \nabla \ell_t(z_t)$ and we obtain

$$\underbrace{\ell_t(z_t) - \ell_t(z_*)}_{\text{per-round regret of OCO}} \leq \underbrace{\langle z_t - z_*, \nabla \ell_t(z_t) \rangle}_{\text{per-round regret of OLO}} .$$

Now, assume that loss functions $\{\ell_t(\cdot)\}_{t=1}^T$ are convex, and the decision space Z is convex. Consider we have an OLO algorithm $A(\cdot)$.

---

for $t = 1, 2, \ldots$ **do**
    Commit a point $z_t \in Z \subset R^d$, recommended by $A(\cdot)$.
    Obtain $\nabla \ell_t(z_t) \in \mathbb{R}^d$ from the Nature.
    Feed the cost vector $c_t \leftarrow \nabla \ell_t(z_t) \in \mathbb{R}^d$ to $A(\cdot)$,
    i.e., $A(\cdot)$ receives a loss function $\tilde{\ell}_t(z) = c_t^\top z$ and incurs a loss $\tilde{\ell}_t(z_t)$.
end for

OLO's Regret is given by

$$\text{Regret}_T(z_*)^{\text{OLO}} = \sum_{t=1}^T \langle z_t, \nabla \ell_t(z_t) \rangle - \sum_{t=1}^T \langle z_*, \nabla \ell_t(z_t) \rangle.$$

Thus, the regret of our reduction approach is

$$\begin{aligned} \text{Regret}_\text{T}{}^{\text{OCO}}(z_*) &:= \sum_{t=1}^T \ell_t(z_t) - \ell_t(z_*) \\ &\leq \sum_{t=1}^T \langle z_t, \nabla \ell_t(z_t) \rangle - \sum_{t=1}^T \langle z_*, \nabla \ell_t(z_t) \rangle \\ &= \text{Regret}_T(z_*)^{\text{OLO}}. \end{aligned}$$

2.5 Regret of FTRL with Lipschitz Loss Functions

Lemma 3. *Let $R : Z \to \mathbb{R}$ be a μ-strongly convex function over Z the decision space with respect to a norm $\|\cdot\|$. If ℓ_t is G_t-Lipschitz with respect to $\|\cdot\|$, then*

$$\ell_t(z_t) - \ell_t(z_{t+1}) \leq G_t \|z_t - z_{t+1}\| \leq \frac{G_t^2}{\mu}$$

Proof. Define $F_t(z) = \sum_{s=1}^{t-1} \ell_s(z) + R(z)$. Then $z_t = \arg\min_{z \in Z} F_t(z)$.

By strong convexity,

$$F_t(z_{t+1}) \geq F_t(z_t) + \langle z_{t+1} - z_t, \nabla F_t(z_t) \rangle + \frac{\mu}{2} \|z_{t+1} - z_t\|^2$$

and

$$F_t(z_t) \geq F_t(z_{t+1}) + \langle z_t - z_{t+1}, \nabla F_t(z_{t+1}) \rangle + \frac{\mu}{2} \|z_t - z_{t+1}\|^2$$

.

Note that $\nabla F_t(z_t) = 0$ and $\nabla F_t(z_{t+1}) = 0$ from optimality.

Combine them together:

$$F_{t+1}(z_t) - F_t(z_t) \geq F_{t+1}(z_{t+1}) - F_t(z_{t+1}) + \mu \|z_t - z_{t+1}\|^2$$

The last inequality is equivalent to

$$\ell_t(z_t) - \ell_t(z_{t+1}) \geq \mu \|z_t - z_{t+1}\|^2$$

.

Finally, to prove the lemma, note that $\|\ell_t(z_t) - \ell_t(z_{t+1})\| \leq G_t \|z_t - z_{t+1}\|$ from Lipschitz condition, and hence $G_t \|z_t - z_{t+1}\| \geq \mu \|z_t - z_{t+1}\|^2$. Rearrange to get $G_t \|z_t - z_{t+1}\| \leq \frac{G_t^2}{\mu}$. □

Theorem 2. *Suppose each $l_t(\cdot)$ is convex and G_t-Lipschitz w.r.t. a norm $\|\cdot\|$. Then FTRL with a μ-strongly convex regularizer $R(\cdot)$ satisfies:*

$$Regret_T(z^*) \leq R(z^*) - R(z_1) + \frac{TG^2}{\mu}$$

where $G := max_t(G_t)$.

Proof.

$$\begin{aligned} Regret_T(z_*) &= \sum_{t=0}^{T} (\ell_t(z_t) - \ell_t(z_*)) \\ &\leq R(z_*) - R(z_1) + \sum_{t=1}^{T} (\ell_t(z_t) - \ell_t(z_{t+1})) \\ &\leq R(z_*) - R(z_1) + \sum_{t=1}^{T} \frac{G_t^2}{\mu} \\ &\leq R(z_*) - R(z_1) + \frac{TG^2}{\mu} \end{aligned}$$

The first inequality comes from lemma 2.

The second inequality comes from lemma 3.

□

3 Application: Prediction with Experts' Advice

3.1 Prediction with Experts' Advice

Suppose now that the learner's decision space is the probability simplex:

$$\Delta_n := \{z \in \mathbb{R}^n : \sum_{i=1}^{n} z[i] = 1, z[i] \geq 0\}$$

We note that Δ_n in this case defines a discrete probability distribution. In this case, OLO at each time step has the following interpretation:

At each time t, a player plays $z_t \in \Delta_n$, and receives a cost vector c_t. We have that

$$\begin{aligned} \ell_t(z_t) &= \langle z_t, c_t \rangle \\ &= \sum_{i=1}^{n} z_t[i] c_t[i] \\ &= \sum_{i=1}^{n} P(i_t = i) c_t[i] \\ &= \mathbb{E}_{i_t \sim z_t}[c_t[i_t]]. \end{aligned}$$

The loss function $\ell_t(z_t)$ gives us the expected cost at each time step t.

Now, consider a more concrete example. Suppose again that we have

$$Z = \Delta_n := \left\{ z \in \mathbb{R}^n : \sum_{i=1}^{n} z[i] = 1, z[i] \geq 0 \right\}$$

Then, using the FTRL approach we have:

At round t, play

$$z_t = \arg\min_{z \in Z} \sum_{s=1}^{t-1} \ell_s(z) + R(z),$$

where $R(z)$ is μ-strongly convex. Suppose additionally that each $\ell_t(\cdot)$ is convex and G_t-Lipschitz w.r.t. a norm $\|\cdot\|$, that is

$$|\ell_t(x) - \ell_t(y)| \leq G_t \|x - y\|.$$

Then, recall Theorem 2 in section 2.5:

Theorem 2: *Suppose each $\ell_t(\cdot)$ is convex and G_t-Lipschitz w.r.t. a norm $\|\cdot\|$. Then, FTRL with a μ-strongly convex regularizer $R(\cdot)$ satisfies:*

$$Regret_T(z^*) \leq R(z^*) - R(z_1) + \frac{TG^2}{\mu},$$

where $G := max_t(G_t)$.

Below we give a general example of application of this theorem.

Example 1: Suppose each $\ell_t(\cdot)$ satisfies the above conditions, and $R(\cdot) \leftarrow \frac{\mu}{2}\|\cdot\|^2$. Then we have,

$$\begin{aligned}
\text{Regret}_T(z^*) &\leq \frac{\mu}{2}\|z^*\|^2 + \frac{TG^2}{\mu} \\
&\leq \frac{\mu D^2}{2} + \frac{TG^2}{\mu} \quad \text{(Suppose the decision space is bounded by } D) \\
&= GD\sqrt{T} + GD\sqrt{T} \quad (\text{Set } \mu \text{ to be } \frac{G\sqrt{T}}{D}) \\
&= O(GD\sqrt{T})
\end{aligned}$$

This result shows us that the regret is bounded sublinearly w.r.t. T. We now turn to a specific problem of prediction with expert advice.

Example 2: Suppose

$$Z = \Delta_n := \left\{ z \in \mathbb{R}^n : \sum_{i=1}^{n} z[i] = 1, z[i] \geq 0 \right\}$$

and consider FTLR:

At round t, play

$$z_t = \arg\min_{z \in Z} \sum_{s=1}^{t-1} \ell_s(z) + \eta\phi(z),$$

where $\phi(z) : Z \to \mathbb{R}$ is strongly convex. Let $\phi(z) := \sum_{i=1}^{d} z_i \log z_i$ be the negative entropy, which is 1-strongly convex w.r.t. $\|\cdot\|_1$ (this was proved in HW1).

Recall that $\ell_t(z) = \langle z, c_t \rangle$. It is easy to show that $l_t(\cdot)$ is convex. To show $l_t(\cdot)$ is G_t-Lipschitz, we note that

$$\begin{aligned} l_t(x) - l_t(y) &= \langle x, c_t \rangle - \langle y, c_t \rangle \\ &= \langle x - y, c_t \rangle \\ &\leq \|c_t\|_\infty \|x - y\|_1 \quad \text{(Holder's inequality)} \end{aligned}$$

Assume additionally that the cost is bounded, that is $\forall i \in [d], c_t[i] \in [0, 1]$, we have

$$\|c_t\|_\infty = \max_i (c_t[i]) \leq 1.$$

This result implies that

$$l_t(x) - l_t(y) \leq \|x - y\|_1.$$

Thus, each $l_t(\cdot)$ is G_t-Lipschitz w.r.t. L1-norm. On the other hand, since we know that $R(z)$ (the negative entropy) is 1-strongly convex w.r.t. L1-norm, all conditions of theorem 2.5 are satisfied. Recall again that the theorem gives us:

$$\begin{aligned} \text{Regret}_T(z^*) &\leq R(z^*) - R(z_1) + \frac{TG^2}{\eta} \\ &= \eta\phi(z^*) - \eta\phi(z_1) + \frac{TG^2}{\eta} \end{aligned}$$

To further proceed, we need the following lemma:

Lemma: Let $\phi(z) = \sum_{i=1}^n z_i \log z_i$ be negative entropy. Then,

$$\max_{z \in \Delta_n} \phi(z) - \min_{z \in \Delta_n} \phi(z) = \log n.$$

With this lemma, we can get:

$$\begin{aligned} \text{Regret}_T(z^*) &= \eta\phi(z^*) - \eta\phi(z_1) + \frac{TG^2}{\eta} \\ &\leq \eta \log n + \frac{TG^2}{\eta} \\ &= O\left(G\sqrt{\log(n)T}\right) \quad (\text{Setting } \eta = \sqrt{\frac{TG^2}{\log n}}) \end{aligned}$$

Remark: This result tells us that the regret depends logarithmically on the number of experts/items (i.e., $\sqrt{\log n}$), which is in contrast to the case of using the squared of the l_1 norm as the regularizer (i.e., $\sqrt{n}$).

Update: The update of z at each t is:

$$z_t = \arg\min_{z \in \Delta_n} \sum_{s=1}^{t-1} \langle z, c_s \rangle + \eta \sum_{i=1}^{n} z_i \log z_i$$

The update step has a closed form solution, for each $i \in [n]$,

$$z_t[i] = \frac{\exp(-\eta \sum_{i=1}^{t-1} c_s[i])}{\sum_{j=1}^{n} \exp(-\eta \sum_{i=1}^{t-1} c_s[j])}$$

This is known as Multiplicative Weight update. The weights are updated multiplicatively and iteratively according to the feedback of how well an expert performed: reducing it in case of poor performance (large loss), and increasing it otherwise.

3.2 An Intuitive Interpretation

Suppose that we are investing on stocks and receive advice from three different experts. $c_t[i]$ gives us the observed cost of choosing the advice of expert i at time t. At the end of turn t, $z_t[i]$ indicates the current weight we give to the advice of expert i (how much confidence we currently have in that expert).
Based on the previous analysis, $l_t(z_t) = \mathbb{E}_{i_t \sim z_t}[c_t[i_t]]$ gives us the expected cost based on the current weight at time t. The cumulative regret corresponds to the total extra cost of distributing the capital as we did, in comparison to the best/benchmark path in hindsight.
Moreover, the update

$$z_t[i] = \frac{\exp(-\eta \sum_{i=1}^{t-1} c_s[i])}{\sum_{j=1}^{n} \exp(-\eta \sum_{i=1}^{t-1} c_s[j])}$$

indicates that if the cumulative cost of trusting expert i is high, then the confidence we have in it should become low, and vice versa.

4 Online Mirror Descent

Recall first the Mirror Descent:

Mirror Descent Algorithm: Suppose $f(z)$ is convex but not necessarily differentiable. $g_t \in \partial f(z_t)$ is the sub-gradient of $f(\cdot)$ at z_t.

```
for t = 1, 2, ... do
    z_{t+1} = arg min_{z∈C} ⟨g_t, z − z_t⟩ + (1/η) D^φ_{z_t}(z).
end for
```

In discussing Mirror Descent, we proved the following theorem:

Theorem 3. *Choose a generating function $\phi(z)$ that is 1-strongly convex w.r.t. some norm $\|\cdot\|$. Then, Mirror Descent has*

$$\sum_{t=1}^{T} f(z^t) - f(z^*) \leq \frac{1}{\eta} D^{\phi}_{z_1}(z_*) + \sum_{t=1}^{T} \frac{\eta}{2} \|g_t\|_*^2,$$

where $g_t \in \partial f(z_t)$ is the sub-gradient of $f(\cdot)$ at z_t and $D^{\phi}_{z}(z_*)$ is the initial Bregman Divergence.

For Online Mirror Descent, the algorithm is modified to be as follows:

Online Mirror Descent Algorithm Suppose $\ell_t(z)$ is convex but not necessarily differentiable. $g_t \in \partial \ell_t(z_t)$ is the sub-gradient of $\ell_t(\cdot)$ at z_t.

```
for t = 1, 2, ... do
    z_{t+1} = arg min_{z∈C} ⟨g_t, z − z_t⟩ + (1/η) D^φ_{z_t}(z)
end for
```

As the two algorithms are essentially identical, OMD admits a similar theorem:

Theorem 4. *Choose a generating function $\phi(z)$ that is 1-strongly convex w.r.t. some norm $\|\cdot\|$. Then, Mirror Descent has*

$$Regret_T(z^*) = \sum_{t=1}^{T} \ell(z^t) - \ell(z^*) \leq \frac{1}{\eta} D^{\phi}_{z_1}(z_*) + \sum_{t=1}^{T} \frac{\eta}{2} \|g_t\|_*^2,$$

where $g_t \in \partial f(z_t)$ is the sub-gradient of $f(\cdot)$ at z_t and $D^{\phi}_{z}(z_*)$ is the initial Bregman Divergence.

The two proofs are similar, so we only prove the latter one here.

Proof. Since $\ell_t(\cdot)$ is convex, by the first-order condition of convexity we have:

$$\begin{aligned} f(x_*) &\geq l_t(z_t) + \langle \nabla l_t(z_t), z_* - z_t \rangle & (1)\\ \Leftrightarrow l(z_t) - l(z_*) &\leq \langle g_t, z_t - z_* \rangle \quad \text{(by rearrangement)} & (2)\\ &= \langle g_t, x_t - x_{t+1} + x_{t+1} - x_* \rangle & (3)\\ &= \langle g_t, x_t - x_{t+1} \rangle + \langle g_t, x_{t+1} - x_* \rangle & (4) \end{aligned}$$

Since $z_{t+1} = \arg\min_{z\in C} \langle g_t, z - z_t \rangle + \frac{1}{\eta} D^{\phi}_{z_t}(z)$, by the optimality condition we have:

$$\left\langle g_t + \frac{1}{\eta}(\nabla\phi(z_{t+1}) - \nabla\phi(z_t)), z - z_{t+1} \right\rangle \geq 0, \forall z \in C$$

Set $z \leftarrow z^*$ gives us:

$$\begin{aligned} &\langle g_t + \frac{1}{\eta}(\nabla\phi(z_{t+1}) - \nabla\phi(z_t)), z_* - z_{t+1} \rangle \geq 0\\ \Leftrightarrow &\langle g_t, z_* - z_{t+1} \rangle + \frac{1}{\eta}\langle \nabla\phi(z_{t+1}) - \nabla\phi(z_t), z_* - z_{t+1} \rangle \geq 0\\ \Leftrightarrow &\langle g_t, z_{t+1} - z_* \rangle \leq \frac{1}{\eta}\langle \nabla\phi(z_{t+1}) - \nabla\phi(z_t), z_* - z_{t+1} \rangle \end{aligned}$$

Plug this result back into (5) gives us:

$$f(z_t) - f(z_*) \leq \frac{1}{\eta}\langle \nabla\phi(z_{t+1}) - \nabla\phi(z_t), z_* - z_{t+1} \rangle + \langle g_t, z_t - z_{t+1} \rangle \tag{5}$$

Now, to further proceed, we need the following lemma:

Lemma: (Three Point Equality) $\forall x, y, z \in C$,

$$\langle \nabla\phi(x) - \nabla\phi(y), z - x \rangle = D^{\phi}_y(z) - D^{\phi}_x(z) - D^{\phi}_y(x)$$

Applying this result to (5), we get:

$$\ell_t(z_t) - \ell_t(z_*) \leq \frac{1}{\eta}\left(D^{\phi}_{z_t}(z_*) - D^{\phi}_{z_{t+1}}(z_*) - D^{\phi}_{z_t}(z_{t+1})\right) + \langle g_t, z_t - z_{t+1} \rangle.$$

Fact: (Fenchel Young Inequality)

$$\langle u, v \rangle \leq \frac{\eta}{2}\|u\|^2 + \frac{1}{2\eta}\|v\|_*^2$$

Apply F-Y inequality to the second term of the RHS of the previous inequality, and we get:

$$\ell_t(z_t) - \ell_t(z_*) \leq \frac{1}{\eta}\left(D^{\phi}_{z_t}(z_*) - D^{\phi}_{z_{t+1}}(z_*) - D^{\phi}_{z_t}(z_{t+1})\right) + \frac{\eta}{2}\|g_t\|_*^2 + \frac{1}{2\eta}\|z_t - z_{t+1}\|^2.$$

Finally, by 1-strong convexity, we can derive that

$$D^{\phi}_{z_{t+1}}(z_t) \geq \frac{1}{2}\|z_t - z_{t+1}\|^2.$$

Plugging in this result, we get

$$\ell_t(z_t) - \ell_t(z_*) \leq \frac{1}{\eta}\left(D^{\phi}_{z_t}(z_*) - D^{\phi}_{z_{t+1}}(z_*)\right) + \frac{\eta}{2}\|g_t\|_*^2,$$

for any feasible t.
Taking the sum of these inequalities over all $t = 1, 2, \ldots, T$ gives us

$$\sum_{t=1}^{T} \ell_t(z_t) - \ell_t(z_*) \leq \frac{1}{\eta} D^{\phi}_{z_1}(z_*) + \sum_{t=1}^{T} \frac{\eta}{2}\|g_t\|_*^2.$$

□

5 Optimistic Mirror Descent

Optimistic Mirror Descent is an algorithm modified from online mirror descent. The algorithm is as follows:

Optimistic Mirror Descent: Suppose $\ell_t(z)$ is convex but not necessarily differentiable. $g_t \in \partial \ell_t(z_t)$ is the sub-gradient of $\ell_t(\cdot)$ at z_t. Assume there is a good guess m_t of g_t.

---

for $t = 1, 2, \ldots$ **do**
  $z_{t-\frac{1}{2}} = \arg\min_{z\in C}\langle g_{t-1}, z\rangle + \frac{1}{\eta} D^{\phi}_{z_{t-\frac{3}{2}}}(z)$.
  $z_t = \arg\min_{z\in C}\langle m_t, z\rangle + \frac{1}{\eta} D^{\phi}_{z_{t-\frac{1}{2}}}(z)$.
end for

The algorithm aims to make use of the estimation m_t and attempts to incorporate the result of online mirror descent and the initial estimation to achieve a more optimal outcome. It admits the following bound:

$$\text{Regret}_T(z^*) \leq \frac{1}{\eta} D^{\phi}_{z_1}(z_*) + \sum_{t=1}^{T} \frac{\eta}{2}\|g_t - m_t\|_*^2$$

From the bound, if the guess m_t is close to g_t, then the regret can be potentially better than $O(\sqrt{T})$ of those non-optimistic variants, e.g., Online Mirror Descent.

Moreover, even if the guess m_t is a poor estimate of g_t, it can still have $O(\sqrt{T})$ regret as those non-optimistic variants. Specifically, if the size of both m_t and g_t is bounded by G, then the second term of the regret bound above can be upper-bounded as

$$\|g_t - m_t\|_*^2 \leq (\|g_t\|_*^2 + \|m_t\|_*^2) \leq 4G^2$$

Thus, the regret bound is still $O(\sqrt{T})$ with an appropriately chosen η.

Bibliographic notes

The following materials are recommended for further studies: [Hazan (2019)], [Orabona (2020)], and [Erven (2021)].

References

[Hazan (2019)] Elad Hazan. Introduction to Online Convex Optimization. 2019.

[Orabona (2020)] Francesco Orabona. A Modern Introduction to Online Learning. 2020

[Erven (2021)] Tim van Erven. Why FTRL Is Better than Online Mirror Descent. 2021. `https://urldefense.com/v3/__https://www.timvanerven.nl/blog/ftrl-vs-omd/__;!!Mih3wA!FOOHQhY1QowoVUZqZrzfFSygPCXb1hunKwmggdon9T_Ouv1kMGC_NvMgcsetWLmdtbk91ZUwNf4LMIU$`