# Mirror Descent

Lalit Jain

April 29, 2021

We assume that we have an action set $A$, a convex set $D$, a function $F : D \to \mathbb{R}$ with domain $D$ such that $A \subset D$. We further assume that $F$ is Legendre,

- $F$ is strictly convex and differentiable on $D^\circ$.
- The range of $F$ is $\mathbb{R}^d$
- $\lim_{x\to\partial D} \|\nabla F(x)\| = +\infty$

Legendre functions satisfy several very nice properties. Firstly $\arg\min_D f \subset D^\circ$. Secondly $\nabla f$ is an invertible map and we can compute it's inverse (almost) explicitly! More on this later.

**Definition 1.** *The Bregman divergence induced by $F$ is given by,*

$$B_F(y, x) = F(y) - F(x) - \langle \nabla F(x), y - x\rangle$$

By definition, the Bregman diveregence is just the error in the first order Taylor series approximation. Since $F$ is convex, the Bregman divergence is always positive. Also if $F$ is strongly convex with respect to a norm $\|\cdot\|$, by definition, $DF(x, y) \geq \frac{\eta}{2}\|x - y\|^2$. Finally we have the following "triangle inequality" for any three points $x, y, z \in A$

$$B_F(z, x) + B_F(x, y) - B_F(z, y) = \langle \nabla F(y) - \nabla F(x), z - x\rangle$$

We consider a game with an oblivious adversary. At each time we plan an arm $a_t \in A$ and observe a loss $\ell_t(a_t) = \langle a_t, g_t\rangle$

**Algorithm 1** Online Mirror Descent

**Input:** $F, D, A, T, \eta$

1: Initialize: $a_1 = \arg\min_{a\in A} F(a)$
2: **for** $t = 0, 1, 2, \cdots, T$ **do**
3: Player observes loss $\ell_t(a_t) = \langle g_t, a_t\rangle$
4: Update $y_t$

$$\nabla F(y_{t+1}) = \nabla F(a_t) - \eta g_t$$

5: Update: $a_{t+1} = \arg\min_{a\in A} B_F(a, y_{t+1})$
6: **end for**

The condition that $F$ is Legendre ensures that $y_{t+1}$ is indeed defined. We point out that it could be replaced with a condition that $A \subset D^\circ$ though this can often be too restrictive (as in Example 2 with the simplex below).

**Example 1:** Let $F(x) = \frac{1}{2}\|x\|_2^2$ on $D = \mathbb{R}^d$ and $A = K$ with $K$ convex, then $\nabla F(x) = x$ and we can compute the divergence,

$$B_F(x, y) = \frac{1}{2}\|x\|^2 - \frac{1}{2}\|y\|^2 - \langle y, x - y\rangle = \frac{1}{2}\|x - y\|_2^2.$$

So we see that we recover gradient descent $a_{t+1} = \Pi_K(a_t - \eta g_t)$. Next class we will see gradient descent as a "quadratic approximation" descent method.

**Example 2:** Take $X = \Delta_d$, $D = \mathbb{R}_+^d$, and $F(x) = \sum_{i=1}^d x_i \log(x_i) - x_i$. Then $F$ is indeed Legendre and $\nabla F(x) = \log(x_i)$ and the induced divergence is

$$\nabla B_F(x, y) = \sum_{i=1}^d x_i \log\left(\frac{x_i}{y_i}\right) - \sum_{i=1}^d x_i - y_i$$

**Algorithm 2** Online Stochastic Mirror Descent

**Input:** $F, D, X, T, \eta$

Initialize: $\bar{a}_1 = \arg\min_{a \in A} F(a)$
**for** $t = 0, 1, 2, \cdots, T$ **do**
  Player chooses a distribution $x_t \in \Delta_A$ with $E[x_t] = \bar{a}_t$
  Sample action $a_t \sim P_t$ and observe $\ell_t(a_t) = \langle a_t, g_t \rangle$
  Compute an estimator $\tilde{g}_t$ with $\mathbb{E}[\tilde{g}_t] = g_t$
  Update $y_t$

$$\nabla F(y_{t+1}) = \nabla F(\bar{a}_t) - \eta \tilde{g}_t$$

  Update: $\bar{a}_{t+1} = \arg\min_{a \in \text{conv}(A)} B_F(a, y_{t+1})$
**end for**

Then the update step is $\log y_{t+1} = \log a_t - \eta g_t$ so, $y_{t+1} = a_t \exp(-\eta g_t)$. It remains to show that $B_F(x, y) = \|x\|_1$. With this we have

$$a_{t+1,i} = \frac{\exp(-\eta \sum_{s=1}^t g_{s,i})}{\sum_{j=1}^d \exp(-\eta \sum_{s=1}^t g_{s,j})}$$

**Theorem 1.** *Assume that $F$ is Legendre on $D$ and that $A \subset D$ is a closed convex set with $D \cap A \neq \emptyset$. Then for any $x \in A$*

$$\sum_{t=1}^T \ell_t(x_t) - \sum_{t=1}^T \ell_t(x) \leq \frac{B_F(x, x_1)}{\eta} + \frac{\eta}{2} \sum_{t=1}^T \psi_t(g_t)$$

*where*

$$\psi_t(g_t) = \min\{\|g\|^2_{(\nabla^2 F(z_t))^{-1}}, \|g\|^2_{(\nabla^2 F(z_t'))^{-1}}\}$$

*where $z_t$ is a point on the line between $x_t$ and $x_{t+1}$ and $z_t'$ is a point on the line between $x_t$ and $y_{t+1}$. Similarly if $\mathbb{E}[\tilde{g}_t | \mathcal{F}_t] = g_t$ for all $t \geq 1$ we have*

$$\mathbb{E}[\sum_{s=1}^t \ell_t(x_t)] - \sum_{s=1}^t \ell_t(x) \leq \frac{B_F(x, x_1)}{\eta} + \frac{\eta}{2} \sum_{t=1}^T \mathbb{E}\psi_t(\tilde{g}_t)$$

*Proof.* We leave the majority of the proof for next time. We will quickly show that the first implies the second. Note that,

$$\begin{aligned}
\mathbb{E}[\ell_t(a_t) - \ell_t(x)] &= \mathbb{E}[\langle a_t - x, g_t \rangle] \\
&= \mathbb{E}[\langle \bar{a}_t - x, g_t \rangle] \\
&= \mathbb{E}[\mathbb{E}[\langle \bar{a}_t - x, g_t \rangle | \bar{a}_t]] \\
&= \mathbb{E}[\mathbb{E}[\langle \bar{a}_t - x, \tilde{g}_t \rangle | \bar{a}_t]] \\
&= \mathbb{E}[\langle \bar{a}_t - x, \tilde{g}_t \rangle]
\end{aligned}$$

Now we apply the first part of the theorem and take expectations on both sides to arrive at the final result. □

Let's now apply this theorem to analyze the regret of previous algorithms.
**Example 1 continued.** Return to the context of Example 1 in the full information setting. Assume $\text{diam}(K) = D$ and $\max_{t \geq 1} \|g_t\|^2 \leq L$. In this case $\nabla^2 F = I$ so

$$R_t \leq \frac{\|x - x_1\|_2^2}{\eta} + \frac{\eta}{2} \sum_{s=1}^t \|g_t\|_2^2 \leq \frac{D}{\eta} + \frac{\eta T L}{2} \leq \sqrt{LDT}$$

using $\eta = \frac{\sqrt{D}}{\sqrt{LT}}$ which matches the result from gradient descent.
**Example 2 continued.** We compute the Hessian,

$$\nabla^2 F(x) = \text{diag}(\{\frac{1}{x_i}\}_{i=1}^d)$$

From the above calculation when $\|g_t\|_\infty \leq 1$, $\|g_t\|^2_{(\nabla^2 F(z_t))^{-1}} = \sum_{i=1}^d z_{t,i} g_{t,i}^2 \leq 1$

$$R_t \leq \frac{\log(d)}{\eta} + \frac{\eta T}{2} \leq O(\sqrt{T\log(d)})$$

Now we apply OSMD to the setting where $A = \{e_1, \cdots, e_d\}$ with the entropic regularizer. This just reduces to the multi-armed bandit setting. We set $P_t = \bar{a}_t$ and $\tilde{g}_t = \frac{\mathbf{1}\{a_t = e_i\}}{\bar{a}_{t,i}} g_{t,i}$. Then,

$$\begin{aligned}
\mathbb{E}[\|\tilde{g}\|^2_{(\nabla^2 F(z_t'))^{-1}}] &= \mathbb{E}[\sum_{i=1}^d z'_{t,i} \frac{\mathbf{1}\{a_t = e_i\}}{\bar{a}_{t,i}^2} g_{t,i}^2] \\
&\leq \mathbb{E}[\sum_{i=1}^d \bar{a}_{t,i} \frac{\mathbf{1}\{a_t = e_i\}}{\bar{a}_{t,i}^2} g_{t,i}^2] \\
&\leq \mathbb{E}[\sum_{i=1}^d \frac{\mathbf{1}\{a_t = e_i\}}{\bar{a}_{t,i}} g_{t,i}^2] \\
&\leq \sum_{i=1}^d g_{t,i}^2 \leq d \qquad \text{(assume an oblivious adversary)}
\end{aligned}$$

where the first to second line used $y_{t+1} = \bar{a}_t \exp(-\eta g_t) \leq \bar{a}_t$ so since $\|g\|_\infty \leq 1$ we have that $z'_{t,i} \leq \bar{a}_{t,i}$ for all $1 \leq i \leq d$. We can now compute the final regret to be $O(\sqrt{Td\log(d)})$. The $\log(d)$ can be removed with a more careful choice of the potential function.

**Rewriting OMD:**

Here is another way to write the mirror descent update:

$$\begin{aligned}
a_{t+1} &= \arg\min_{a\in A} B_F(a, y_{t+1}) \\
&= \arg\min_{a\in A} F(a) - \nabla F(y_{t+1})^\top a \\
&= \arg\min_{a\in A} F(a) - (\nabla F(a_t) - \eta g)^\top a \\
&= \arg\min_{a\in A} F(a) - F(a_t) - \nabla F(a_t)^\top (a - a_t) + \eta g^\top a \\
&= \arg\min_{a\in A} \eta g_t^\top a + B_F(a, a_t)
\end{aligned}$$

Finally we remark that due to the assumption that $F$ is Legendre and strongly convex with domain $D$ then

$$y_{t+1} = \arg\min_{a\in D} \langle \eta g_t, a\rangle + B_F(a, a_t).$$

Indeed,

$$\begin{aligned}
\arg\min_{a\in D} \langle \eta g_t, a\rangle + B_F(a, a_t) &= \arg\min_{a\in D} \langle \eta g_t, a\rangle + F(a) - \nabla F(a_t)^\top a \\
&= \arg\min_{a\in D} \langle \eta g_t - \nabla F(a_t), a\rangle + F(a)
\end{aligned}$$

This last function is also Legendre so it's minima occurs in the interior of $D$ and so by the first order optimality conditions,

$$\eta g_t - \nabla F(a_t) = \nabla F(a)$$

so we see that $y_{t+1}$ is indeed the minimum.

**Proof of Theorem 1** We need two facts. Firstly the "law of cosines"

$$B_F(z, x) + B_F(x, y) - B_F(z, y) = \langle \nabla F(y) - \nabla F(x), z - x\rangle.$$

Secondly, for $a \in A$ and $y \in D$, let $y' = \arg\min_{a\in A} B_F(a, y)$, then

$$B_F(a, y') + B_F(y', y) \geq B_F(a, y)$$

Now,

$$\begin{aligned}\ell_t(a_t) - \ell_t(a) &\leq \nabla g_t^\top (a_t - a) \\ &= \frac{1}{\eta}(\nabla F(y_{t+1}) - \nabla F(a_t))^\top (a - a_t) \\ &= \frac{1}{\eta}(B_F(a, a_t) - B_F(a, y_{t+1}) + B_F(a_t, y_{t+1})) \\ &\leq \frac{1}{\eta}(B_F(a, a_t) - B_F(a, a_{t+1}) - B_F(a_{t+1}, y_{t+1}) + B_F(a_t, y_{t+1}))\end{aligned}$$

where the first inequality is from convexity, the second is from the definition of the step, the third is the law of cosines, the last is from projections. Thus summing overall iterations we see

$$R_t \leq \frac{B(a, a_1)}{\eta} + \sum_{t=1}^{T} B_F(a_t, y_{t+1}) - B_F(a_{t+1}, y_{t+1})$$

Now,

$$\begin{aligned}B_F(a_t, y_{t+1}) - B_F(a_{t+1}, y_{t+1}) &= F(a_t) - F(y_{t+1}) - \nabla F(y_{t+1})^\top (a_t - y_{t+1}) \\ &\quad - F(a_{t+1}) + F(y_{t+1}) + \nabla F(y_{t+1})^\top (a_{t+1} - y_{t+1}) \\ &= F(a_t) - F(a_{t+1}) - \nabla F(y_{t+1})^\top (a_t - a_{t+1}) \\ &= F(a_t) - F(a_{t+1}) - (\nabla F(a_t) - \eta g_t)^\top (a_t - a_{t+1}) \\ &= \langle \eta g_t, a_t - a_{t+1} \rangle - B_F(a_{t+1}, a_t)\end{aligned}$$

This expression on it's own is pretty useful for regret minimization. However, we still need to get to the theorem statement. By Taylor's theorem,

$$B_F(a_{t+1}, a_t) = \frac{1}{2}(a_{t+1} - a_t)^\top \nabla^2 F(z_t)(a_{t+1} - a_t)$$

where $z_t$ is some point on the line $x_t + t(x_{t+1} - x_t), 0 \leq t \leq 1$. Thus using the fact that $\langle x, y \rangle \leq \frac{\|x\|_2^2 + \|y\|_2^2}{2}$ (the Fenchel-Young inequality)

$$\langle \eta g_t, a_t - a_{t+1} \rangle - B_F(a_{t+1}, a_t) \leq \frac{\eta^2}{2}\|g_t\|^2_{(\nabla^2 F(z_t))^{-1}} + \frac{1}{2}\|a_t - a_{t+1}\|^2_{(\nabla^2 F(z_t))^{-1}} - B_F(a_{t+1}, a_t) = \frac{\eta^2}{2}\|g_t\|^2_{(\nabla^2 F(z_t))^{-1}}$$

For the second term in the minimum, note that by definition,

$$\begin{aligned}\langle \eta g_t, a_t - a_{t+1} \rangle - B_F(a_{t+1}, a_t) &\leq \max_{a \in D} \langle \eta g_t, a_t - a \rangle - B_F(a, a_t) \\ &\leq -(\min_{a \in D} \langle \eta g_t, a - a_t \rangle + B_F(a, a_t))) \\ &\leq \langle \eta g_t, a_t - y_{t+1} \rangle - B_F(y_{t+1}, a_t)\end{aligned}$$

and now proceed as before.

**Remark:** We finally mention one other expression for this term that is useful in it's own right (and can give a slightly less informative but perhaps more algebraic proof of the guarantees of EXP3). Recall that by definition,

$$F^*(y) = \sup_{x \in D} \langle y, x \rangle - F(y)$$

is the Fenchel conjugate of $F$. The Fenchel conjugate satisfies two extremely useful properties:

- $\nabla F^{-1} = \nabla F^*$
- $B_F(x, y) = B_{F^*}(\nabla F(y), \nabla F(x))$

The second fact is a consequence of the first. Thus,

$$B_F(a_t, y_{t+1}) - B_F(a_{t+1}, y_{t+1}) \leq B_{F^*}(\nabla F(y_{t+1}), \nabla F(a_t)).$$

**Follow the Regularized Leader:** Now consider the setting where $A = D$ and $F$ is Legendre. Then

$$\nabla F(a_{t+1}) = \nabla F(a_t) - \eta g_t = \nabla F(a_1) - \eta \sum_{s=1}^{t} y_s = \eta \sum_{s=1}^{t} g_s$$

This algorithm has a nice interpretation. Instead of mapping back to the primal space, we do our gradient steps in the dual space and map back at the end. Follow the Regularized Leader chooses

$$a_{t+1} = \arg\min_{a \in A} \sum_{s=1}^{t} \langle a, \eta g_t \rangle + F(a)$$

which leads to the same update! In general if $A \neq D$ then the projection onto $A$ can lead to a setting where the two algorithms have different updates. In the case of exponential weights, the two algorithms agree.
The material in this section is cobbled together from:

- Bandit Algorithms, Lattimore and Szepesvari
- Regret Analysis of Stochastic and Nonstochastic Multi-armed, Bubeck and Cesa-Bianchi
- Introduction to Online Convex Optimization, Hazan
- A Modern Introduction to Online Learning, Orabona