**E1 245: Online Prediction & Learning** **Fall 2015**

Lecture 18 — October 6

*Lecturer: Aditya Gopalan* *Scribe: Mohammadi Zaki*

# 18.1 TWO MIRROR-DESCENT ALGORITHMS FOR ONLINE CONVEX OPTIMIZATION (O.C.O.)

Let $K$ be a convex set and let $R$ be a Legendre function over $K$.

### 18.1.1 "Lazy" Mirror Descent ("Follow the Regularized Leader FTRL")

**Algorithm:**

**Initialize:** $\tilde{w}_1 \in \mathbb{R}^d$, $w_1 = \Pi_{R,K}(\tilde{w}_1)$
**Repeat:** $\forall t = 2, 3, 4, \ldots,$

$$\begin{aligned} \nabla R(\tilde{w}_t) &= \nabla R(\tilde{w}_{t-1}) - z_{t-1} \\ \Rightarrow \tilde{w}_t &= \nabla R^*(\nabla R(\tilde{w}_{t-1}) - z_{t-1}) \\ w_t &= \Pi_{R,K}(\tilde{w}_t) \end{aligned}$$

See Fig. 18.1 for a schematic description.

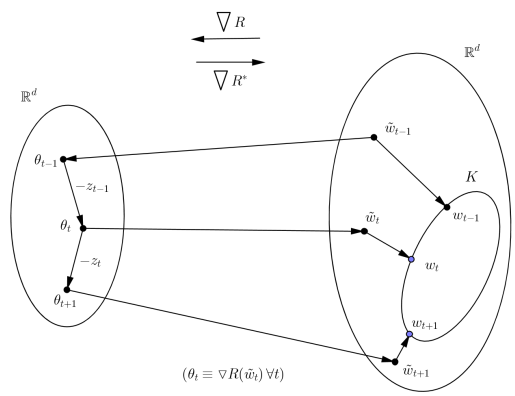



**Figure 18.1.** "LAZY" OMD

### 18.1.2 "Active" Mirror Descent or Proximal Point Algorithm

**Algorithm:**

**Initialize:** $w_1 \in K$
**Repeat:** $\forall t = 2, 3, 4, \ldots,$

$$\nabla R(y_t) = \nabla R(w_{t-1}) - z_{t-1}$$
$$w_t = \Pi_{R,K}(y_t).$$

See Fig. 18.2 for a schematic description.

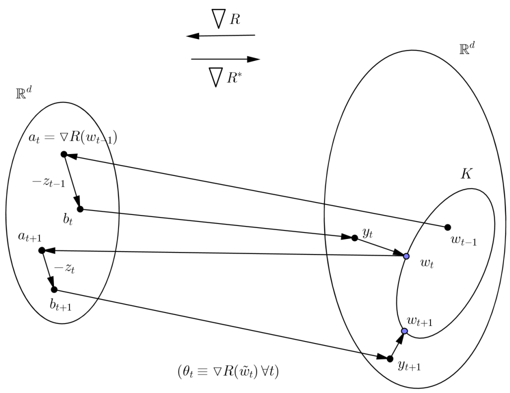



**Figure 18.2.** "ACTIVE" OMD

**Theorem 18.1 (Regret Bound for Active OMD).** *For the active OMD algorithm with R a legendre function,* $\forall u \in K$,

$$\sum_{t=1}^{T}[f_t(w_t) - f_t(u)] \leq D_R(u, w_1) - D_R(u, w_{T+1}) + \sum_{t=1}^{T} D_R(w_t, y_{t+1}).$$

[
**Aside:** For the LAZY version of OMD,

$$\sum_{t=1}^{T}[f_t(w_t) - f_t(u)] \leq D_R(u, w_1) - D_R(u, \tilde{w}_{T+1}) + \sum_{t=1}^{T} D_R(w_t, \tilde{w}_{t+1}).\Big]$$

**Proof:** Let us look at a single term of L.H.S,

$$\begin{aligned}
f_t(w_t) - f_t(u) &\leq \langle \nabla f_t(w_t), w_t - u \rangle && \text{because f is convex.} \\
&= \langle \nabla R(w_t) - \nabla R(y_{t+1}), w_t - u \rangle \\
&= \langle w_t - u, \nabla R(w_t) - \nabla R(y_{t+1}) \rangle \\
&= D_R(u, w_t) - D_R(u, y_{t+1}) + D_R(w_t, y_{t+1}) && \text{By 3-point inequality} \\
&\leq D_R(u, w_t) - D_R(u, w_{t+1}) + D_R(w_t, y_{t+1}) \\
&[\text{"}\mathit{Generalized\,Pythagoras}\text{"} \\
&\forall u \in K, \forall w : \\
&D_R(u, w) \leq D_R(u, \Pi_{R,K}(w)) + D_R(\Pi_{R,K}(w), w)]
\end{aligned}$$

Adding over t = 1,2,..., T,

$$f_t(w_t) - f_t(u) \leq D_R(u, w_1) - D_R(u, w_{T+1}) + \sum_{t=1}^{T} D_R(w_t, y_{t+1}).$$

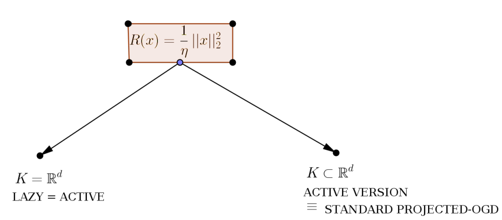
$\square$

**Note:** This can be specialized to give concrete regret bounds, e.g., [1, Zinkevich, '03].
Consider active OMD with $R(x) = \frac{1}{2\eta}||x||_2^2$.

$$\begin{aligned}
y_t &= w_{t-1} - \eta \nabla f_{t-1}(w_{t-1}) \\
w_t &= \Pi_{R,K}(y_t) && [\equiv POGD.] \\
Regret_T(POGD) &\leq \frac{1}{2\eta} D^2 + \frac{\eta}{2} T G^2. \\
&where, \\
D &= \max_{x,y \in K} ||x - y||_2^2 && (\text{"diameter"}). \\
G &= \sup_{t \leq T, x \in K} ||\nabla f_t(x)||_2^2. && (\text{"Largest possible gradient"})
\end{aligned}$$

### 18.1.3 Examples:

1. $R(x) = \frac{1}{2\eta}||x||_2^2$.



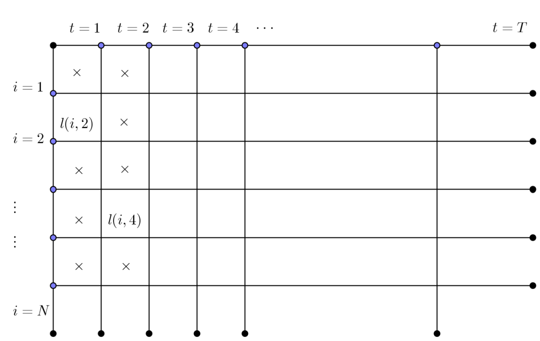
## 18.2 Online Learning with limited/partial information-BANDITS

Recall: The problem of competing with the set of $N$ actions ($\equiv$ experts problem or O.C.O. with linear losses).
At each time $t = 1, 2, \ldots,$

1. Environment sets a loss $l(i,t), \forall i = 1, 2, \ldots, N$.
2. Learner/algorithm plays action $I_t \in [N]$.
3. Algorithm suffers loss $l(I_t, t)$.
4. Algorithm gets to "see" only $l(I_t)$ and NOT $(l(i,t))_{i=1}^{N}$.

### 18.2.1 Assumptions on $\{l(i,t)\}_{i,t}$

We will assume "Oblivious adversary" ,i.e., $l(i,t)$ is deterministic but arbitrarily chosen at $t = 0$. ($l(i,t)$ chosen is "independent" of the algorithm chosen).




**Notes:**

1. $Regret(T) = \sum_{t=1}^{T} l(I_t, t) - \min_{i\in[N]} \sum_{t=1}^{T} l(i,t)$.

2. Randomization is absolutely necessary here to avoid linear regret as for any deterministic algorithm the adversary can set worst loss for each round forcing the algorithm to incur regret of $\mathcal{O}(T)$.

# Bibliography

[1] Martin Zinkevich, ''Online convex programming and generalized infinitesimal gradient ascent,'' in *Proc. of the Twenty International Conference on Machine Learning* (ICML'03), *pp.* 928-936, AAAI Press, 2003.