CSE599s, Spring 2012, Online Learning Lecture 3 - 04/03/2012

The Online Optimization Game

Lecturer: Brendan McMahan *Scribe: Evan Herbst*

1 Online Optimization Game

Again, define the online optimization game:

- for $t = 1..T$ (eventually we'll have bounds that hold $\forall T$, so that we can let $T = \infty$)
  - player chooses a predictor $w_t \in \mathcal{W} \subseteq \mathbb{R}^n$
  - adversary reveals loss function $f_t : \mathcal{W} \to \mathbb{R}$
  - player pays $f_t(w_t)$

For real applications, think T maybe O(10^9), n maybe O(10^7).

The worst-case sum of losses is arbitrarily high, so instead of minimizing the sum of losses we'll minimize regret:

$$\text{Regret} = \sum_{t=1}^{T} f_t(w_t) - \min_{w \in \mathcal{W}} \sum_{t=1}^{T} f_t(w).$$

We also can't minimize regret wrt strategies that can change w at each iteration, because we'd still do arbitrarily badly. For now we'll write w as though it's constant over time, but later we will develop algorithms that have low regret against the set of strategies that switch w a known finite number of times (switching regret) and to the set of strategies in which $|w_t - w_{t-1}|$ is bounded (drifting regret).

2 Example Online Optimizers

An example algorithm: **Follow-The-Leader (FTL)**:

$$w_{t+1} = \operatorname*{argmin}_{w \in W} \sum_{s=1}^{t} f_s(w) \equiv \operatorname*{argmin}_{w} f_{1:t}(w)$$

Against linear functions, FTL can have the worst possible regret, O(T).

Another example algorithm (not realizable because it sees the future): **Be-The-Leader (BTL)**: as FTL, but play w_{t+1} on round t instead of playing w_t on round t.

Suppose $W \subset \mathbb{R}$ and the loss is required to be linear ($f_t(w) = g_t w$). Table for the same example we did for FTL last week:

t	w_t	g_t	**loss**(t)	$g_{1:t}$
1	-1	.5	-.5	.5
2	1	-1	-1	-.5
3	-1	1	-1	.5

Etc. Regret(BTL) will be about $-T$, not about T as it was for FTL. This is because now each g_t entry gets filled in *before* the w_t entry on the same line.

Theorem 1 (BTL Theorem)**.** *For arbitrary bounded f_t, Regret(BTL) ≤ 0. Equivalently, $\sum_{t=1}^{T} f_t(w_{t+1}) \leq \sum_{t=1}^{T} f_t(w_{T+1})$, which will be our IH.*

Proof. By induction.

Base case: $T = 1$. $f_1(w_2) \leq f_1(w_2)$.

For the induction step, suppose the IH holds for T. Then,

$$\begin{aligned}
\sum_{t=1}^{T+1} f_t(w_{t+1}) &= \sum_{t=1}^{T} f_t(w_{t+1}) + f_{T+1}(w_{T+2}) \\
&\leq \sum_{t=1}^{T} f_t(w_{T+1}) + f_{T+1}(w_{T+2}) && \text{IH} \\
&\leq \sum_{t=1}^{T} f_t(w_{T+2}) + f_{T+1}(w_{T+2}) && \text{def. } \mathrm{w_{T+1}} \\
&= \sum_{t=1}^{T} f_t(w_{T+2}).
\end{aligned}$$

□

Since the difference between FTL and BTL is whether we play w_t or w_{t+1}, let's try to bound the regret of FTL using that of BTL.

Theorem 2 (FTL Theorem)**.** *For $\forall w^* \in W$,*

$$\textit{Regret}(\textit{FTL vs } w^*) \leq \sum_{t=1}^{T} (f_t(w_t) - f_t(w_{t+1})),$$

again assuming the f_t are bounded.

Proof.

$$\begin{aligned}
\text{Regret(FTL)} &= \sum_{t=1}^{T} f_t(w_t) - \sum_{t=1}^{T} f_t(w^*) \\
&\leq \sum_{t=1}^{T} f_t(w_t) - \sum_{t=1}^{T} f_t(w_{T+1}) && \text{def. } w_{T+1} \\
&= \underbrace{\sum_{t=1}^{T} f_t(w_{t+1}) - \sum_{t=1}^{T} f_t(w_{T+1})}_{=\text{ regret(BTL) } \leq 0} + \sum_{t=1}^{T} (f_t(w_t) - f_t(w_{t+1})) \\
&\leq \sum_{t=1}^{T} (f_t(w_t) - f_t(w_{t+1})).
\end{aligned}$$

□

3 Transformations in Online Optimization

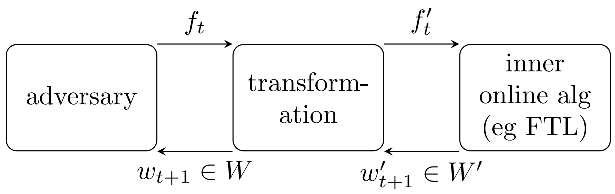

Figure 1: transformations in online optimization.

Sometimes we can transform the adversary's loss and parameter space into something that it's easier to prove things about. See fig. 1 for an overview. Information flows from upper left to lower left by the arrows.

An algorithm using a transformation: **Follow-The-Regularized-Leader (FTRL)**, which runs FTL on a regularized loss function. The udpate is defined by

$$w_{t+1} = \underset{w}{\operatorname{argmin}}(f_{1:t}(w) + r(w))$$

where the regularization function $r : \mathbb{R}^n \to \mathbb{R}$ satisfies $r(w) \geq 0$, and typically also $r(0) = 0$.

We view this as running FTL together with the following transformation: Given functions $f_t(w)$ chosen by the adversary, we let $f'_1(w) = g_t \cdot w + r(w)$ (that is, we add a regularization component to the first function we see), and take $f'_t = f_t$ for $t > 1$.

Again, we can prove a strong result that holds for arbitrary (potentially even non-convex f_t):
Theorem 3 (FTRL theorem)**.** *The FTRL algorithm has*

$$\mathit{Regret}(FTRL) \leq \sum_{t=1}^{T}(f_t(w_t) - f_t(w_{t+1})) + r(w^*).$$

Proof.

$$\sum_t f'_t(w_t) - \sum_t f'_t(w^*) \leq \sum_t f'_t(w_t) - f'_t(w_{t+1}) \quad \text{Applying the FTL theorem to } f'$$

$$\iff \underbrace{\sum_t f_t(w_t) - \sum_t f_t(w^*)}_{regret(FTRL)} + r(w_1) - r(w^*) \leq \sum_t (f_t(w_t) - f_t(w_{t+1})) + r(w_1) - r(w_2)$$

$$\iff \text{Regret(FTRL)} \leq \sum_t (f_t(w_t) - f_t(w_{t+1})) + r(w^*) - \underbrace{r(w_2)}_{\leq 0}$$

□

FTRL for linear f_t:

- $f_t(w) = g_t \cdot w$
- $|g_t| \leq G$ (bounded loss)
- $w_{t+1} = \operatorname{argmin}_w \left(\sum_t g_t \cdot w + \frac{\sigma}{2}|w|^2\right), \sigma \in \mathbb{R}^+$ (quadratic regularizer)

Then,

$$\begin{aligned}
\text{Regret(FTRL)} &\le \sum_{t=1}^{T}(f_t(w_t) - f_t(w_{t+1})) + r(w^*) && \text{FTRL theorem} \\
&= \sum_t g_t \cdot (w_t - w_{t+1}) + r(w^*) \\
&\le \sum_t G|w_t - w_{t+1}| + \frac{\sigma}{2}|w^*|^2 && \text{Cauchy-Schwarz} \\
&\le \frac{T}{\sigma}G^2 + \frac{\sigma}{2}|w^*|^2.
\end{aligned}$$

This is just gradient descent with a fixed learning rate $\frac{1}{\sigma}$, because

$$w_t - w_{t+1} = \frac{-g_{1:t-1} + g_{1:t}}{\sigma} = \frac{g_t}{\sigma}.$$

To choose σ optimally you need to know G and T. But in practice it doesn't matter that you don't know T. In general, you can always apply the "doubling trick" (not explained in class). Or even better, you can analyze a version of the algorithm that changes the amount of regularization adaptively, which we may analyze later.

Theorem 4. *$\forall w^*$ with $|w^*| \le R$, against linear loss functions $f_t(w) = g_t \cdot w$ with $|g_t| \le G$, FTRL with $r(w) = \frac{\sigma}{2}|w|^2$, where $\sigma = \frac{G\sqrt{2T}}{R}$, has regret $\le GR\sqrt{2T}$.*

Proof. Plug this value for σ into the regret bound above. □