EECS598: Prediction and Learning: It's Only a Game Fall 2013

Lecture 15: Online Convex Optimization: Part III

Prof. Jacob Abernethy *Scribe: Murray (Yanzhe) Lei*

Announcements

- CRLT in-class evaluation on Wednesday lecture.
- Project ideas coming soon.
- Schedule individual meeting with Prof. Jacob Abernethy by the end of next week to talk about final project.

1 Online Convex Optimization

1.1 General Framework of Online Convex Optimization (OCO)

Assume we have a *decision space* $X \subset \mathbb{R}^n$, which is convex, closed and compact.

For $t = 1, ..., T$

- Player plays some $x_t \in X$.
- Nature reveals some $l_t : X \to \mathbb{R}$ which is convex.
- Player suffers loss of $l_t(x_t)$.

Our goal is to minimize the regret regards the best static decision in hindsight:

$$Regret_T := \sum_{t=1}^{T} l_t(x_t) - \min_{x \in X} \sum_{t=1}^{T} l_t(x) \tag{1.1}$$

1.2 Online Gradient Descent (OGD) Approach

Assuming $\{l_t(\cdot)\}_{t=1}^T$ are differentiable. Starting with some arbitrary $x_1 \in X$.

For $t = 1, ..., T$

- $x_{t+1} \leftarrow \Pi_X(x_t - \eta \bigtriangledown l_t(x_t))$, where $\Pi_X(\cdot)$ is the projection function.

The performance of *OGD* is described as follows.

Theorem 1.1. *If there exists some positive constant G, D such that*

$$\| \bigtriangledown l_t(x)\|_2 \le G, \forall t, x \in X \tag{1.2}$$

$$\|X\|_2 ;= \max_{x,y \in X} \|x - y\|_2 \le D \tag{1.3}$$

Then $Regret_T(OGD) \le \mathcal{O}(DG\sqrt{T})$.

Observation: The bound may not be optimal! Here are two examples.

- $X = \Delta_n$, $l_t(x) = \underline{l}^t \cdot x$, where $0 \le l_i^t \le 1$ for $i = 1, ..., n$.

  This is the expert problem in action setting. We know that EWA will give us $\mathcal{O}(\sqrt{log(n)T})$ bound on regret. But Theorem 1.1 only gives us $\mathcal{O}(\sqrt{nT})$ because $G = \sqrt{n}, D = \sqrt{2}$.

- $X = \Delta_n$, $l_t(x) = -log(b^t \cdot x)$, where $b^t \in (0, \infty)^n$.

  This is the universal portfolio selection problem. We know that UCRP gives us $\mathcal{O}(nlog(T))$ bound on regret.. In this case we cannot even apply OGD because $\nabla l_t(x) = -\frac{b^t}{b^t \cdot x}$ is unbounded.

2 Follow the Regularized Leader (FTRL)

Recall that the Follow the Leader Algorithm works pretty bad in some adversarial setting. However, after closer look into the algorithm, it actually works really well if the loss function is "regularized"! This motivate the FTRL algorithm.

2.1 Follow the Leader Algorithm

Follow the leader (FTL):

$$x_t = arg\min_{x \in X} \sum_{s=1}^{t-1} l_s(x).$$

What if we can observe the loss function 1 period ahead of time?

Be the leader (BTL):

$$\hat{x}_t = arg\min_{x \in X} \sum_{s=1}^{t} l_s(x).$$

Claim 1: $Regret_T(BTL) \le 0$.

Proof: By induction and definition of $\hat{x}_t$: for all t

$$\begin{aligned} Regret_t(BTL) &= \sum_{s=1}^{t} [l_s(\hat{x}_s) - l_s(\hat{x}_t)] = \sum_{s=1}^{t-1} [l_s(\hat{x}_s) - l_s(\hat{x}_t)] \\ &\le \sum_{s=1}^{t} [l_s(\hat{x}_s) - l_s(\hat{x}_{t-1})] = Regret_{t-1}(BTL) \le ... \le Regret_0(BTL) = 0. \square \end{aligned}$$

This tells us that BTL performs very well despite its fictitiousness. Then how far is FTL from BTL?

Claim 2: $Regret_T(FTL) \leq \sum_{t=1}^{T}[l_t(x_t) - l_t(x_{t+1})]$.

Proof: By definition of x_t and the fact that $\hat{x}_t = x_{t+1}$:

$$\begin{aligned} Regret_T(FTL) &= \sum_{t=1}^{T}[l_t(x_t) - l_t(x_{T+1})] = \sum_{t=1}^{T}[l_t(x_{t+1}) - l_t(x_{T+1})] + \sum_{t=1}^{T}[l_t(x_t) - l_t(x_{t+1})] \\ &= \sum_{t=1}^{T}[l_t(\hat{x}_t) - l_t(x_{T+1})] + \sum_{t=1}^{T}[l_t(x_t) - l_t(x_{t+1})] = 0 + \sum_{t=1}^{T}[l_t(x_t) - l_t(x_{t+1})]. \square \end{aligned}$$

Now we are ready to introduce a situation where FTL works very well.

2.2 Online Density Estimation

Observe $Z_1, Z_2, \ldots Z_T \in \mathbb{R}^n$. Before seeing Z_t, we want to predict the mean μ_t. We assume the underlying model is normal distribution:

$$Z_t \sim N(\mu, I)$$

and the predictor pay log loss of the conditional probability:

$$-log(\mathbb{P}(Z_t|\mu_t)) = -log(C \cdot exp\{-\frac{\|Z_t - \mu_t\|^2}{2}\}) = \frac{\|Z_t - \mu_t\|^2}{2} + C'$$

What is the Maximum Likelihood Estimator (MLE)?

$$\begin{aligned} \mu^* &= \underset{\mu}{argmax} \prod_{t=1}^{T} \mathbb{P}(Z_t|\mu) = \underset{\mu}{argmin} \sum_{t=1}^{T} -log(\mathbb{P}(Z_t|\mu)) \\ &= \underset{\mu}{argmin} \sum_{t=1}^{T} \frac{\|Z_t - \mu_t\|^2}{2} = \frac{1}{T}\sum_{t=1}^{T} Z_t \end{aligned}$$

So, if we apply FTL by using MLE, what would be the regret?

$$FTL: \qquad \mu_t = \frac{1}{t}\sum_{s=1}^{t-1} Z_s$$

By Claim 2 in section 2.1

$$Regret_T(FTL) \leq \sum_{t=1}^{T}[l_t(x_t) - l_t(x_{t+1})] = \frac{1}{2}\sum_{t=1}^{T}[\|Z_t - \mu_t\|^2 - \|Z_t - \mu_{t+1}\|^2]$$

Notice that $\mu_{t+1} = \frac{t-1}{t}\mu_t + \frac{Z_t}{t}$, $\|Z_t - \mu_{t+1}\| = \|Z_t - \frac{Z_t}{t} - \frac{t-1}{t}\mu_t\| = (1-\frac{1}{t})\|Z_t - \mu_t\|$, we have

$$Regret_T(FTL) \leq= \frac{1}{2}\sum_{t=1}^{T}(1-(1-\frac{1}{t})^2)\|Z_t - \mu_t\|^2 \leq \sum_{t=1}^{T}\frac{\|Z_t - \mu_t\|^2}{t}$$

If we assume that $\|Z_t\| \leq D$, then $Regret_T(FTL) \leq \sum_{t=1}^{T}\frac{D^2}{t} \sim \mathcal{O}(D^2 log(T))$!

Question: Why FTL works pretty well here?

Answer: Because l_t's are curved, thus x_t's don't move so much across time.

Claim: If l_t's are α-strongly convex, i.e. for all $x \in X$ and $\delta > 0$

$$l_t(x+\delta) \geq l_t(x) + \nabla l_t(x) \cdot \delta + \alpha\frac{\|\delta\|^2}{2}$$

then FTL has regret bounded by $\mathcal{O}(\frac{1}{\alpha}log(T))$.

Thus we need to modify FTL when l_t's are not curved!

2.3 Follow the Regularized Leader (FTRL)

Choose a "regularizer" $R : X \to \mathbb{R}$

For $t = 1, ..., T$, predicts

$$x_t = arg\min_{x\in X}\{\sum_{s=1}^{t-1} l_s(x) + \frac{R(x)}{\eta}\}.$$

Observation: Assume linear loss function $l_t(x) = \underline{l}^t \cdot x$ and $X = \Delta_n$. For $R(x) = \sum_{i=1}^{n} x_i log(x_i)$ (negative entropy function), FTRL gives us

$$x_t = arg\min_{x\in\Delta_n}\{\sum_{s=1}^{t-1}\underline{l}^t \cdot x + \frac{\sum_{i=1}^{n} x_i log(x_i)}{\eta}\}, \forall t$$

whose close form solution is

$$x_i^t = (e^{-\eta\sum_{s=1}^{t-1} l_i^s})/(\sum_{i=1}^{n} e^{-\eta\sum_{s=1}^{t-1} l_i^s}), \forall i, t$$

This is exactly the distribution of EWA! So EWA is a special case of FTRL.