Announcements

- Project ideas and guidelines posted.
- A project poster session will be held on last class.
- Tershia visiting.
- We will do a poll on the remaining topics to cover. See below for some potential ones.
  - Blackwell, Approachability ≈ Calibrated Forecast ≈ Correlated Equilibrium.
  - Finance: option pricing and its relation to Black-Scholes.
  - Bandit settings (also called limited feedback models).
  - Relationship between regret minimization and generalization error in statistical learning theory.

16.1 Online Convex Optimization: Problem Formulation

Given a convex decision set $X \subset \mathbb{R}^n$,

For $t = 1, \cdots, T$ |
  Learner chooses $x_t \in X$,
  Nature chooses $l_t : X \rightarrow \mathbb{R}$.
|}

$$\text{Regret}_T \triangleq \sum_{t=1}^{T} (l_t(x_t) - l_t(u)), \text{ where } u \text{ is arbitrary point in } X \text{ (could be } \arg\min_x \sum_{t=1}^{T} l_t(x)).$$

**Question:** Why compared to a fixed $u$?

An answer: there exist reductions that let you compete against $u$ that changes over time.

**Challenge** Find an algorithm and its bound that competes against $u_1, \ldots, u_T$.

Hint: the bound will depend on $\sum_{t=1}^{T-1} \|u_{t+1} - u_t\|$.

16.2 A Review of Follow The Leader (FTL) Algorithm

Follow The Leader:

$$x_t = \arg\min_x \sum_{s=1}^{t-1} l_s(x)$$
Regret bound:

\[ \text{Regret}_T(\text{FTL}) \leq \sum_{t=1}^{T} (l_t(x_t) - l_t(x_{t+1})) \]

If loss function takes the form of

\[ l_t(x) = \frac{1}{2} \| z_t - x \|^2 \]

(which can come from log loss of Gaussian density estimation), then

\[ l_t(x_t) - l_t(x_{t+1}) = O\left(\frac{1}{t}\right) \]

hence

\[ \text{Regret}_T(\text{FTL}) \leq O(\log T) \]

In general FTL is a bad algorithm, however it rocks in this special case. The intuition is that it works because of curvature of the objective function, which stems from the curvature of \( l_t \)'s.

### 16.3 Adding Curvature to FTL: Follow The Regularized Leader (FTRL)

Follow The Regularized Leader

\[ x_t = \arg\min_x \sum_{s=1}^{t-1} l_s(x) + \frac{1}{\eta} R(x) \]

where

- \( \eta \) – learning rate.
- \( R \) – the “regularizer”, which is a “curved” convex function. Several usual choices of \( R \):
  1. \( R(x) = \sum x_i \log x_i \). (Equivalent to EWA)
  2. \( R(x) = \frac{1}{2} \| x - x_1 \|^2 \), where \( x_1 \) is the initial point. (Similar to OGD)
  3. \( R(x) \) is log of the barrier function.\(^1\)

**Exercise:** Prove that if \( X = \Delta_n \), \( l_t(x) = l^t \cdot x \), then FTRL \( \Leftrightarrow \) EWA.

**Proposition 16.1.** If \( X \) is the ball of radius \( C \) and \( l_t(x) = \nabla_i x \), then FTRL with \( l^2 \)-norm regularizer

\[ x_t = \arg\min_{x \in X} \sum_{s=1}^{t-1} \nabla_s x + \frac{1}{\eta} \frac{\| x - x_1 \|^2}{2} \]

is equivalent to

1. **Compute the minimizer without constraining** \( x \in X \).
2. **Project the minimizer onto the boundary of** \( X \).

---

\(^1\) In 1-dimension, the barrier function is \( \log \frac{1}{x} + \log \frac{1}{1-x} \). This choice of \( R(x) \) works well in bandit setting. See [http://ie.technion.ac.il/~ehazan/papers/ieeeitbandit.pdf](http://ie.technion.ac.il/~ehazan/papers/ieeeitbandit.pdf) for further reading
Lecture 16: FTRL

Connection to Online Gradient Descent (OGD)

\[
\text{OGD: } x_{t+1} = \text{Proj}_X(x_t - \eta \nabla_t) \\
\text{FTRL: } x_{t+1} = \text{Proj}_X(x_1 - \sum_{s=1}^{t-1} \nabla_s)
\]

If you never “try” to leave \(X\), the two algorithms are essentially the same.

16.4 Regret Bound for FTRL

Observation 16.2. FTRL \(\Leftrightarrow\) FTL++ (FTL with one more function \(l_0(x) = \frac{1}{\eta} R(x)\)).

Note that

\[
\text{Regret}(\text{FTL}^{++}) = \text{Regret}(\text{FTRL}) + \frac{1}{\eta}(R(x_0) - R(u))
\]

Let \(x_0 = \arg\min_{x \in X} \frac{1}{\eta} R(x)\). As a consequence, \(x_1 = x_0\). Applying the regret bound of FTL, we have

\[
\text{Regret}(\text{FTL}^{++}) = \sum_{t=0}^{T} (l_t(x_t) - l_t(u)) \leq \sum_{t=0}^{T} (l_t(x_t) - l_t(x_{t+1})) = \sum_{t=1}^{T} (l_t(x_t) - l_t(x_{t+1})) \quad \text{(since } x_1 = x_0)\]

Hence,

\[
\text{Regret}(\text{FTRL}) = \text{Regret}(\text{FTL}^{++}) - \frac{1}{\eta}(R(x_0) - R(u)) \leq \frac{1}{\eta}(R(u) - R(x_1)) + \sum_{t=1}^{T} (l_t(x_t) - l_t(x_{t+1}))
\]
It seems that FTRL is worse than FTL, since besides \( \frac{1}{\eta} (R(u) - R(x_1)) \) the remaining part of the regret bound looks the same. However, the remaining part actually gets improved due to curvature.

\[
l_t(x_t) - l_t(x_{t+1}) \leq \nabla l_t(x_t) (x_t - x_{t+1}) \leq \|\nabla l_t\| \|x_t - x_{t+1}\|
\]

where \(\|x_t - x_{t+1}\|\) is roughly of order \(\eta\) for OGD. Therefore, the whole regret bound looks like

\[
\frac{a}{\eta} + bT\eta
\]

where \(a\) and \(b\) are bounds on \(\|\nabla l_t\|^2\) and \(\|x_1 - x^*\|^2\) respectively. It is easy to see that the bound is optimized when \(\eta = \Theta\left(\frac{1}{\sqrt{T}}\right)\), and the optimal bound is \(O(\sqrt{T})\).