6.S979 | Lecture 8 | Oct 1, 2019

Online Convex Optimization

- Unattainability of $\frac{1}{T}\sum_t \min f_t(\cdot)$ target
  - imagine $\mathcal{X} = [0,1]$
  - nature chooses $f_t(x) = \begin{cases} x & \text{wpr } 1/2 \\ 1-x & \text{wpr } 1/2 \end{cases}$
  - no matter what x_t's algorithm chooses:

$$\mathbb{E}\, f_t(x_t) = \frac{1}{2} \leadsto \mathbb{E}\left[\sum_{t=1}^{T} f_t(x_t)\right] = \frac{1}{2}T$$

  - however, $\mathbb{E}\left[\sum_{t=1}^{T} \min f_t(\cdot)\right] = 0$ $\left(\text{also whp } \sum_{t=1}^{T} f_t(x_t) \geq 0.49T\right)$

- Follow-the-Leader is not no-regret

$$f_t(x) = \begin{cases} -x\left(1+\frac{t+1}{T}\right) & , \; t\text{: even} \\ x\left(1+\frac{t}{T}\right) & , \; t\text{: odd} \end{cases} \qquad \mathcal{X} = [0,1]$$

$$x \cdot \left[\left(1+\frac{1}{T}\right) - \left(1+\frac{2}{T}\right) + \left(1+\frac{3}{T}\right) + \dots + \left(1+\frac{t-1}{T}\right)\right]$$

then $\forall$ t: even: $\arg\min \sum_{\tau<t} f_t(x) = \arg\min(x) = 0$

but loss $= 0$

t: odd: $\arg\min \sum_{\tau<T} f_t(x) = \arg\min(-x) = 1$

but loss $= 1 + \frac{t}{T}$

total loss of FTL $\overset{\sum_{t\ \mathrm{odd}}(1+\frac{t}{T})}{=} O(T)$

however, $\sum_t f_t(0) = 0$

- Follow-the-Regularized Leader w/ ℓ_2^2-regularizer
  - Suppose $R(x) = \frac{1}{2}\|x\|_2^2$
  - suppose f_t's are linear, w/ gradients g_t
  - $\mathcal{X} = \mathbb{R}^n$
  - minimize $\underbrace{\sum_{\tau<t} f_\tau(x) + \frac{1}{\eta}\frac{1}{2}\|x\|_2^2}_{g(x)}$ ⇝ minimized where $\nabla g(x) = 0$

$$\nabla g(x) = \sum_{\tau<t} g_t + \frac{1}{\eta} x$$

$$\nabla g(x) = 0 \iff x = -\eta \cdot \sum_{\tau<t} g_\tau$$

so $x_t = -\eta \cdot \sum_{\tau<t} g_\tau \quad \forall t$

hence: $x_t = x_{t-1} - \eta \cdot g_{t-1} \quad \forall\, t = 2, \ldots, T$

assuming initial condition $x_1 = 0$

in general FTRL w/ ℓ_2^2 regularizer on function sequene $h_t(x) = f_t(x_t) + \nabla f_t(x_t) \cdot (x - x_t)$, $t = 1, \ldots, T$ same as gradient descent on fn' sequene $f_t(x)$, $t = 1, \ldots, T$ + regret bounds obtained by pretending h_t carry over to f_t

∘ FTRL & Min-Max

$\min_{x \in X} \max_{y \in Y} f(x,y)$ ← convex-concave

- suppose both min & max player run FTRL

$$x_t = \arg\min \sum_{\tau<t} f(\cdot, y_\tau) + \frac{1}{\eta} R(\cdot)$$

$$y_t = \arg\min -\sum_{\tau<t} f(x_\tau, \cdot) + \frac{1}{\eta} R(\cdot)$$

- x-player no-regret property:

$$\frac{1}{T}\sum_{t=1}^{T} f(x_t, y_t) \le \min_x \frac{1}{T}\sum_{t=1}^{T} f(x, y_t) + o(1)$$

$$\xrightarrow{\text{concavity}} \le \min_x f\left(x, \frac{1}{T}\sum_t y_t\right) + o(1) \qquad (*)$$

$$\le \min_x \max_y f(x,y) + o(1)$$

- y-player no-regret property:

$$\frac{1}{T}\sum_{t=1}^{T} f(x_t, y_t) \ge \max_y \frac{1}{T}\sum_{t=1}^{T} f(x_t, y) - o(1)$$

$$\xrightarrow{\text{convexity}} \ge \max_y f\left(\frac{1}{T}\sum x_t, y\right) - o(1) \qquad (**)$$

$$\ge \underbrace{\max_y \min_x f(x,y)}_{= \min_x \max_y f(x,y)} - o(1)$$

- so:

$$\min_x \max_y f(x,y) - o(1) \le \frac{1}{T}\sum_{t=1}^{T} f(x_t,y_t) \le \min_x \max_y f(x,y) + o(1)$$

- moreover, let $\bar{x}_T = \frac{1}{T}\sum_{t=1}^{T} x_t$, $\bar{y}_T = \frac{1}{T}\sum_{t=1}^{T} y_t$

$$f(\bar{x}_T,\bar{y}_T) \le \max_y f(\bar{x}_T, y) \underset{(**)}{\overset{(*)}{\le}} \min_x f(x,\bar{y}_T) + o(1)$$

$$f(\bar{x}_T,\bar{y}_T) \ge \min_x f(x,\bar{y}_T) \underset{(**)}{\overset{(*)}{\ge}} \max_y f(\bar{x}_T, y) - o(1)$$

$\rightsquigarrow$ $(\bar{x}_T, \bar{y}_T)$ approximate min-max equilibrium