* HW 6 to be released soon

Online Convex Optimization / Online Learning

Problem Model (Online Convex Opt)

1) We are given a convex body $K \subseteq \mathbb{R}^n$
2) T rounds:
   (a) Alg plays $x_t \in K$
   (b) Convex cost function $f_t : \mathbb{R}^n \to \mathbb{R}$ is revealed

E.g. $f_t(x) = (a_t x - b_t)^2$

Goal: $\min \sum_{t=0}^{T-1} f_t(x_t)$

Benchmark: $\min_{x^* \in K} \sum_{t=0}^{T-1} f_t(x^*) \neq \sum_{t=0}^{T-1} \min_{y_t^* \in K} f_t(y_t^*)$

Thm: Online gradient descent guarantees

$x_{t+1} = \prod_K \left[x_t - \eta \nabla f_t(x_t)\right]$

$$\underbrace{\frac{1}{T}\sum_{t=0}^{T-1}\left[f_t(x_t) - f_t(x^*)\right]}_{\text{Average regret}} \leq \frac{DG}{\sqrt{T}}$$

where $D = \text{Diameter}(K)$

& $G = \max_t \|\nabla f_t(x)\|$ for $x \in K$

Applications

(a) Experts Problem:

- n experts
- Each round $t \in \{1, \dots, T\}$
  (1) Alg chooses an expert/action $a_t \in \{1, \dots, n\}$ random choice is allowed
  (2) Cost of each expert $c_t(i) \in [-1, 1]$ revealed
  & Alg gets $c_t(a_t)$

& Alg gets $c_t(a_t)$

Objective: Min total Alg cost $\sum_{t=1}^{T} c_t(a_t)$

compared to best fixed expert $\min_i \sum_t c_t(i)$

<u>Remark</u>: This problem is useful in Forecasting (weather, stocks), Spam detection, LP solving

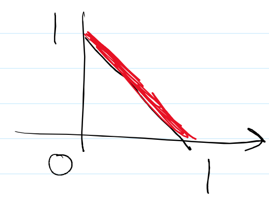
Thm: OGD implies a random alg with expected avg regret

$$\frac{1}{T}\left[\mathbb{E}\left[\sum_t c_t(a_t)\right] - \sum_t c_t(i)\right] \le \frac{2\sqrt{n}}{\sqrt{T}}$$

Pf: Let x_t denote Alg's t-th distrib over n experts

Let $K = \{ x \mid \sum_i x(i) = 1,\ x(i) \ge 0 \}$

$\Rightarrow$ diameter$(K) \le 2$



$f_t(x) = \sum_i c_t(i) \cdot x(i)$ ← linear fn

$\rightarrow \|\nabla f_t\| = \|c_t\| \le \sqrt{n}$

Hence, OGD implies

$$\frac{1}{T}\sum_t \underbrace{f_t(x_t)}_{= \sum_i x_t(i)\cdot c_t(i) = \mathbb{E}[\text{Alg's cost at step } t]} - \min_{x^* \in K} \underbrace{\sum_t f_t(x^*)}_{= \sum_i x^*(i) \sum_t c_t(i) = \min_i \sum_t c_t(i)} \le \frac{\overset{\le 2}{D}\ \overset{\le\sqrt{n}}{G}}{\sqrt{T}}$$

∎

Remark: There is a different alg, called Multiplicative Weights, with average regret $\sqrt{\log n}$ ... for Experts problem

Remark: There is a different alg, called Multiplicative Weights, with average regret $\frac{\sqrt{\log n}}{\sqrt{T}}$ for Experts problem

(b) Minimax Theorem

2-Player Zero Sum Game:

1) Row & Col players
2) Each have n actions
3) Matrix $M_{n \times n}$ where $M(i,j)$ is Row's reward & Column's cost

	Rock	Paper	Sciss
Rock	0	-1	1
Paper	1	0	-1
Sciss	-1	1	0

Strategy S: A distribution over n actions, i.e., $\sum_i s_i = 1$

Minimax value of 0-sum game

Thm: (von Neumann)

$$\underbrace{\min_{S'_{col}} \max_{S'_{row}}}_{\text{Go 2nd}} \mathbb{E}_{\substack{i \sim S'_{row} \\ j \sim S'_{col}}} \left[M(i,j) \right] \quad (A) \quad = \underbrace{\max_{S_{row}}}_{\text{Go 1st}} \min_{S_{col}} \mathbb{E}_{\substack{i \sim S_{row} \\ j \sim S_{col}}} \left[M(i,j) \right] \quad (B)$$

Pf: (A) $\geq$ (B): Easy $\because$ row player can ignore S'_{col}

(A) $\leq$ (B): Think of finding S_{row} in K = {Simplex over n actions}

$K = \{ x \mid \sum_i x(i) = 1, \; x(i) \geq 0 \}$

- Rather than finding S_{row} in one step, gradually improve it (using OGD).

for $t = 1$ to T

(1) Row player plays $x_t \in K$

(2) Column player selects a_t to $\min_{a_t \in [n]} \mathbb{E}_{i \sim x_t} \left[M(i, a_t) \right]$

Online Grad Desc

$$\underbrace{\frac{1}{T} \sum_t M(x_t, a_t)}_{\leq (B) \text{ since row}} \geq \underbrace{\frac{1}{T} \max_i \sum_t M(i, a_t)}_{\geq (A) \text{ since col player}} - \epsilon$$

$\leq$ B since row can play best x_t

$\geq$ A since col player could be unif over a_t

$\blacksquare$