Game Theory

Lecturer: Brendan McMahan *Scribe: Saghar Hosseini*

Game theory is a study of strategic decision making where a set of rational players are playing against each other. Let's consider a zero-sum two-player game where each player's gain or loss is balanced by the loss or gain of the other player. Player I chooses her action from an action set, i.e., $i \in \{1, 2, \ldots, m\}$ and player II chooses his action $j \in \{1, 2, \ldots, n\}$. The game's payoff matrix is denoted by M and is a representation of loss or gain of players. For example player I pays M_{ij} to player II.

Min-Max Theorem

Based on Min-Max Theorem we have

$$\min_{p \in \Delta(m)} \max_{q \in \Delta(n)} p^T M q = \max_{q \in \Delta(n)} \min_{p \in \Delta(m)} p^T M q,$$

where player II has the privilege of playing second and see what player I has chosen. Also, note that since $p \in \Delta(m)$ and $q \in \Delta(n)$ the objective function is equivalent to the expected value of M_{ij} where i and j are drawn from the probability distributions p and q respectively.

Let's consider the worst case where the player plays against an adaptive all knowing adversary which tries to maximize the regret. The number of rounds T is known and fixed.

$$\min_{w_1} \max_{g_1} \min_{w_2} \max_{g_2} \ldots \min_{w_T} \max_{g_T} [\sum_{t=1}^{T} g_t.w_t - \min_{u \in W} g_{1:T}.u] = V_T \in \mathbb{R},$$

where $W = \{w | \, \|w\|_2 \le B\}$, $w_t \in \mathbb{R}^d$, $g_t \in \tilde{G}$, and $\tilde{G} = \{g | \, \|g\|_2 \le G\}$ which is a convex set. The cost of the best fixed comparator can be expressed as

$$\min_{\|u\|_2 \le B} g_{1:T}.u = -B \max_{\|u\|_2 \le 1} g_{1:T}.u = -B \|g_{1:T}\|_* = -B \|g_{1:T}\|_2 \,,$$

Min-Max Adversary

The adversary follows the following strategy:

$$\|g_t\| = G, \; g_t.w_t = 0, \; g_t.g_{1:t-1} = 0,$$

which implies $\sum_{t=1}^{T} g_t.w_t = 0$ and subsequently

$$V_T = -\min_{u \in W} g_{1:T}.u = B \|g_{1:T}\|_2$$

In order to find a the regret bound we need the following lemmas.
Lemma 1: there exist $x, y \in \mathbb{R}$ such that $x.y = 0$, then

$$\|x + y\| = \sqrt{\|x\|^2 + \|y\|^2}.$$

Proof. We have

$$\|x + y\|^2 = (x + y).(x + y) = x^2 + 2x.y + y^2 = \|x\|^2 + \|y\|^2 \,,$$

and the statement of the lemma follows. □

Based on Lemma 1 we can provide a bound on $\|g_{1:t}\|$ in the following lemma.
Lemma 2: for any $t \in \{1, 2, \dots\}$ we have $\|g_{1:t}\| = G\sqrt{t}$.

Proof. The proof by induction is used. We know that $\|g_1\| = G$. Suppose that $\|g_{1:t-1}\| = G\sqrt{t-1}$, thus based on lemma 1 we have

$$\|g_{1:t}\| = \|g_{1:t-1} + g_t\| = \sqrt{G^2(t-1) + G^2} = G\sqrt{t}.$$

□

Therefore, the adversary gets at least $V_T = BG\sqrt{T}$. Note that the regret for Online Gradient Descent (OGD)is bounded as

$$\forall u, \ \ \text{Regret}(u) \leq \frac{\|u\|^2}{2\eta} + \frac{\eta}{2}\sum_{t=1}^{T} g_t^2,$$

where with $\eta = \frac{B}{G\sqrt{T}}$ the regret bound is $BG\sqrt{T}$. Therefore, the player has two choices:
1) OGD with fixed learning rate $\eta = \frac{B}{G\sqrt{T}}$.
2) OGD with adaptive learning rate

$$\eta_t = \frac{B}{\sqrt{\|g_{1:t}\|^2 + G^2(T-t)}}.$$

Note that

$$w_{t+1} = -\eta_t g_{1:t} \Rightarrow \|w_{t+1}\| = \eta_t \|g_{1:t}\| \leq \frac{B}{\sqrt{\|g_{1:t}\|^2}} \|g_{1:t}\| \ \Rightarrow \|w_{t+1}\| \leq B,$$

which implies that the projected OGD is equivalent to OGD against a min-max adversary and the best strategy is to use OGD.
In addition, since

$$\forall u, \ \ \text{Regret}(u) \leq \frac{\|u\|^2}{2\eta} + \frac{\eta}{2}\sum_{t=1}^{T} g_t^2,$$

we have

$$\text{loss} \leq \min_{u \in W}(g_{1:T}.u + \frac{\|u\|^2}{2\eta}) + \frac{\eta}{2}\sum_{t=1}^{T} g_t^2 = -\frac{\eta}{2}(g_{1:T}^2 - \sum_{t=1}^{T} g_t^2),$$

and the following theorem provides the exact loss for OGD.

Theorem 1. *The loss of OGD algorithm is*

$$\text{loss} = -\frac{\eta}{2}(g_{1:T}^2 - \sum_{t=1}^{T} g_t^2).$$

Proof. We know that

$$\text{loss} = \sum_{t=1}^{T} g_t.w_t,$$

and based on the update rule in OGD we have $w_t = -\eta g_{1:t-1}$ and subsequently

$$\text{loss} = \sum_{t=1}^{T} g_t.(-\eta g_{1:t-1}) = -\eta \sum_{t=1}^{T} g_t.g_{1:t-1}.$$

Moreover, since $\sum_{t=1}^{T} g_t.g_{1:t-1} = \frac{1}{2}(g_{1:T}^2 - \sum_{t=1}^{T} g_t^2)$, the statment of the theorem follows. □

We can show that the loss in the above theorem satisfies the regret bound for OGD. Based on the definition of regret for a comparator u we have

$$\text{Regret} = \text{loss} - g_{1:T}.u = -\frac{\eta}{2}(g_{1:T}^2 - \sum_{t=1}^{T} g_t^2) - g_{1:T}.u,$$

Thus,

$$\text{Regret} \leq \frac{\eta}{2}\sum_{t=1}^{T} g_t^2 + \max_{z\in\mathbb{R}^d}(-\frac{\eta}{2}z^2 - z.u) = \frac{\|u\|^2}{2\eta} + \frac{\eta}{2}\sum_{t=1}^{T} g_t^2.$$

Generally, any algorithm for online linear algorithm results in

$$\text{loss} \leq -\psi(g_{1:T}) \quad \forall g_1, g_2, \ldots, g_T$$

if and only if

$$\text{Regret}(u) \leq \psi^*(u) \quad \forall u \in \mathbb{R}^d,$$

where the convex conjugate of $\psi(u)$ is defined as

$$\psi^*(u) = \max_{g\in\mathbb{R}^d} g.u - \psi(u)$$