CS 6840 – Algorithmic Game Theory (2 pages)	Spring 2012
Lecture 11 Scribe Notes	
Instructor: Éva Tardos	*Norris Xu*

1 Lecture 11 — Wednesday 15 February 2012 No-Regret Learning

1.1 No-Regret Learning

Suppose we have a game with cost, and players $1, \ldots, n$. For a strategy vector $s = (s_1, \ldots, s_n)$, we have $c_i(s)$ is the cost for player i if strategy vector s is played.

We play this game on days $1, \ldots, T$, with $s^t = (s_1^t, \ldots, s_n^t)$ being the strategy vector used on day t. The overall cost is therefore $\sum_{t=1}^T \sum_i c_i(s^t)$ and player i's cost is $\sum_{t=1}^T c_i(s^t)$.

We can model Best Response: each day, one person best responds, and the rest do the same thing as before. But best response is unrealistic (e.g. rock-paper-scissors). In reality, players learn based on others' strategies.

Definition (No-Regret)**.** A sequence of strategy vectors $s^1, \ldots, s^T$ is *no-regret* for player i if:

$$\sum_{t=1}^T c_i(s^t) \leq \min_x \underbrace{\sum_{t=1}^T c_i(x, s_{-i}^t)}_{\text{cost of strategy } x}$$

Definition (Vanishing Regret)**.** A sequence of strategy vectors $s^1, \ldots, s^T$ has *vanishing regret* for player i if, assuming that $0 \leq c_i(s) \leq 1$ for any strategy vector s:

$$\limsup_{T \to \infty} \frac{1}{T} \sum_{t=1}^T c_i(s^t) - \min_x \frac{1}{T} \sum_{t=1}^T c_i(x, s_{-i}^t) \leq 0$$

Theorem. If a cost game is (λ, μ)-smooth[1] and all players have no regret on a sequence $s^1, \ldots, s^T$ of plays, then:

$$\sum_{t=1}^T \sum_i c_i(s^t) \leq \underbrace{\frac{\lambda}{1-\mu}}_{\text{same bound as Price of Anarchy}} T \min_s \sum_i c_i(s)$$

[1]**Definition** ((λ, μ)-smooth). A cost game is (λ, μ)-*smooth* if for any strategy vectors s, s^*:

$$\sum_i c_i(s_i^*, s_{-i}) \leq \lambda \sum_i c_i(s^*) + \mu \sum_i c_i(s)$$

Proof. Let s^* be the min cost vector. Then for any player i, s_i^* is no-regret. Then:

$$\sum_{t=1}^{T} c_i(s^t) \le \min_x \sum_{t=1}^{T} c_i(x, s_{-i}^t) \le \sum_{t=1}^{T} c_i(s_i^*, s_{-i}^t)$$

$$\sum_{t=1}^{T} \sum_i c_i(s^t) \le \sum_{t=1}^{T} \sum_i c_i(s_i^*, s_{-i}^t)$$

$$\le \sum_{t=1}^{T} \left(\lambda \sum_i c_i(s^*) + \mu \sum_i c_i(s^t) \right) = \lambda T \sum_i c_i(s^*) + \mu \sum_{t=1}^{T} \sum_i c_i(s^t)$$

$$\sum_{t=1}^{T} \sum_i c_i(s^t) \le \frac{\lambda}{1-\mu} T \sum_i c_i(s^*)$$

□

Lemma. The sequence $s, s, \ldots, s$ is no-regret for all players if and only if s is a Nash equilibrium.

1.2 Randomized Strategy

For all players i and strategies x, let the probability distribution $p_i(x)$ be the probability of i playing x. Assuming that the players' choices are independent, the probability of playing the strategy vector $s = (s_1, \ldots, s_n)$ is then $p(s) = \prod_i p_i(s_i)$. The expected cost $\mathrm{E}_s[c_i(s)]$ for player i is $\sum_s p(s) c_i(s) = \sum_{s=(s_1,\ldots,s_n)} \prod_j p_j(s_j) c_i(s)$.

A set of probability distributions p_i form a Nash equilibrium if for any player i and strategy x, $\mathrm{E}_s[c_i(s)] \le \mathrm{E}_s[c_i(x, s_{-i})]$; equivalently, for any player i, $\mathrm{E}_s[c_i(s)] \le \min_x \mathrm{E}_s[c_i(x, s_{-i})]$.