**10-716: Advanced Machine Learning** **Spring 2019**

# Lecture 24: April 25

*Lecturer: Pradeep Ravikumar* *Scribes: Quanyang Lu, Ben Lengerich*

**Note**: *LaTeX template courtesy of UC Berkeley EECS dept.*



## 24.1 Preliminaries

**Definition 24.1 (K-Person Normal Form Games)** *A (finite) $K$ -person game given in its strategic (or normal) form is defined as follows. Player $k, k = 1, \cdots, K$ has $N_k \in \mathbb{N}^+$ possible actions (or pure strategies) to choose from. We denote the $K$-tuple of all players' strategy by $\mathbf{i} = (i_1, \cdots, i_K) \in \bigotimes_{k=1}^K \{1, \cdots, N_k)$, then the* ***loss*** *suffered by player $k$ is $l^{(k)}(\mathbf{i})$, where $l^{(k)} : \bigotimes_{k=1}^K \{1, \cdots, N_k \to [0, 1]$ for each $k = 1, \cdots, K$ are given loss function for all players. Notice that although the loss function is player-specific, it depends all on players' actions.*

***Definition 24.2 (Mixed Strategy)*** *A mixed strategy for player $k$ is a probability distribution $\mathbf{p}^{(k)} = (p_1^{(k)}, \cdots, p_{N_k}^{(k)}$ over the set $\{1, \cdots, N_k\}$ of actions. For mixed strategy, player $k$ choose an action according to the distribution $\mathbf{p}^{(k)}$ Denote the action played by player $k$ by $I^{(k)}$. Thus $I^{(k)}$ is a random variable taking values in the set $\{1, \cdots, N_k\}$ and distributed according to $\mathbf{p}^{(k)}$. Let $\mathbf{I} = (I^{(1)}, \cdots, I^{(K)})$ denote the $K$-tuple of all actions played by all players. If the random variables $I^{(1)}, \cdots, I^{(K)}$ are independent of each other (i.e., the players randomize independently of each other), we denote their joint distribution by $\pi$. That is, $\pi$ is the joint distribution over the set $\bigotimes_{k=1}^K \{1, \cdots, N_k\}$ of all possible $K$-tuples of actions obtained by the product of the mixed strategies $\mathbf{p}^{(1)}$. The product distribution $\pi$ is called* ***mixed strategy profile****.*

$$\pi(\mathbf{i}) = \mathbb{P}[\mathbf{I} = \mathbf{i}] = p_{i_1}^{(1)} \times \cdots \times p_{i_K}^{(K)}$$

***Definition 24.3 (Expected Loss)*** *We define the* ***expected loss*** *of player $k$ by:*

$$\begin{aligned}\mathbb{E}_\pi l^{(k)} &= \sum_{\mathbf{i} \in \bigotimes_{k=1}^K \{1, \cdots, N_k\}} \pi(\mathbf{i}) l^{(k)}(\mathbf{i}) \\ &= \sum_{i_1=1}^{N_1} \cdots \sum_{i_K=1}^{N_K} p_{i_1}^{(1)} \cdots p_{i_K}^{(K)} l^{(k)}(i_1, \cdots, i_K)\end{aligned}$$

## 24.2 2 Person Zero-Sum Game

### 24.2.1 Definition

We now consider the definition of Nash Equilibrium. For k-player game, the joint strategy distribution is defined as,

$$\Pi = P^{(1)} \times P^{(2)} ... \times P^{(k)}$$

For any $j \in \{1...k\}$ and any mixed strategy $Q(j) \neq P(j)$ for each $j$, define a alternative strategy,

$$\Pi'_j = P^{(1)} \times ...Q^{(j)}... \times P^{(k)}$$

If $\forall j$ and $Q(j)$, we have,

$$\mathbb{E}_{I\sim\Pi} l^{(j)}(I) \leq \mathbb{E}_{I\sim\Pi'_j} l^{(j)}(I)$$

We call such $\Pi$ is Nash Equilibrium strategy.

**Theorem 24.4 (Nash's Existence Theorem)** *Any finite game has a Nash Equilibrium. Set of Nash Equilibrium strategies can be complex.*

### 24.2.2 Game setting

In a 2 person zero-sum game, K = 2, and $\forall I = (i_1, i_2)$, we have

$$l^{(1)(I)} = -l^{(2)}(I)$$

Define $l \equiv l^{(1)}$, Player 1's goal is to minimize $l$, and Player 2's goal is to maximize $l$.

Define,

$$N_1 = N, N_2 = M, (i, j) = (i_1, i_2)$$

Then the mixed strategy for Player 1 and Player 2 is,

$$P = (P_1, ..., P_n)$$

$$Q = (Q_1, ..., Q_m)$$

$$\Pi = (P \times Q)$$

### 24.2.3 Von Neumman's Minimax Theorem

In the above setting, assuming $\Pi = (P \times Q)$ is the Nash Equilibrium mixed strategy,

$$\sum_{i=1}^{M}\sum_{j=1}^{N} P_i Q'_j l(i,j) \leq \sum_{i=1}^{M}\sum_{j=1}^{N} P_i Q_j l(i,j) \leq \sum_{i=1}^{M}\sum_{j=1}^{N} P'_i Q_j l(i,j)$$

Define,

$$l(P, Q) = \sum_{i=1}^{M}\sum_{j=1}^{N} P_i Q_j l(i,j)$$

we obtain,

$$\max_{Q'} l(P, Q') \leq l(P, Q) \leq \min_{P'} l(P', Q)$$

$$\max_{Q'} l(P, Q') \leq \min_{P'} l(P', Q)$$

By adding $\min_P$ and $\max_Q$ on left and right side, respectively,

$$\min_{P} \max_{Q'} l(P, Q') \leq \max_{Q} \min_{P'} l(P', Q)$$

Also, from the fact that,

$$l(P', Q) \geq \min_{P'} l(P', Q)$$

take $\max_Q$ on both side, we obtain,

$$\max_Q l(P', Q) \geq \max_Q \min_{P'} l(P', Q)$$

Because the above inequality holds for any P',

$$\min_{P'} \max_Q l(P', Q) \geq \max_Q \min_{P'} l(P', Q)$$

Therefore,

$$\min_{P'} \max_Q l(P', Q) = \max_Q \min_{P'} l(P', Q)$$

For any Nash Equilibrium strategy $\Pi = P \times Q$, the value of the game is defined as,

$$V = l(P, Q) = \min_{P'} \max_Q l(P', Q) = \max_Q \min_{P'} l(P', Q)$$

**Theorem 24.5** *Any mixed strategy (P, Q) s.t. $l(P, Q) = V$ is a Nash Equilibrium strategy.*

## 24.3 Generalization of Von Neumman Theorem

Now we will generalize the Von Neumman theorem to see min-max max-min equivalence for a broad variety of functions.

**Theorem 24.6 (7.1 in BL)** *Let $f : \mathcal{X} \times \mathcal{Y} \to \mathbb{R}$ be a bounded real-valued function, where $\mathcal{X}$ and $\mathcal{Y}$ are convex sets and $\mathcal{X}$ is compact.*

*Suppose that $f(\cdot, y)$ is convex and continuous for each fixed $y \in \mathcal{Y}$ and $f(x, \cdot)$ is concave for each fixed $x \in \mathcal{X}$. Then*

$$\inf_{x\in\mathcal{X}} \sup_{y\in\mathcal{Y}} f(x, y) = \sup_{y\in\mathcal{Y}} \inf_{x\in\mathcal{X}} f(x, y) \tag{24.1}$$

**Proof:** First note that for any function $f$

$$\inf_{x\in\mathcal{X}} \sup_{y\in\mathcal{Y}} f(x, y) \geq \sup_{y\in\mathcal{Y}} \inf_{x\in\mathcal{X}} f(x, y) \tag{24.2}$$

because $\forall x' \in \mathcal{X}$, $\sup_y f(x', y) \geq \sup_y \inf_x f(x, y)$.

Now let us consider the reverse inequality. Wlog, we assume that $f(x, y) \in [0, 1]\ \forall\ (x, y) \in \mathcal{X} \times \mathcal{Y}$. Our strategy is to define a covering of $\mathcal{X}$ and examine the worst-case difference between the expressions.

Fix $\epsilon > 0$, $n > 0$. $\mathcal{X}$ is compact, so there is a set $\{x^{(1)}, \ldots, x^{(N)}\} \subset \mathcal{X}$ such that $\forall x \in \mathcal{X}$, $d(x^{(i)}, x) \leq \epsilon$. Fix $y_0$. For $t = 1, \ldots, n$, let

$$x_t = \frac{\sum_{i=1}^N x^{(i)} e^{-\eta \sum_{s=0}^{t-1} f(x^{(i)}, y_s)}}{\sum_{i=1}^N e^{-\eta \sum_{s=0}^{t-1} f(x^{(i)}, y_s)}} \tag{24.3}$$

where $\eta = \sqrt{8lnN/n}$ and $y_t$ satisfies $f(x_t, y_t) \geq \sup_{y\in\mathcal{Y}} f(x_t, y) - 1/n$. Then, because $f(\cdot, y_t)$ is convex, we have that

$$\frac{1}{n}\sum_{t=1}^{n} f(x_t, y_t) \leq \min_{i=1,\ldots,N} \frac{1}{n}\sum_{t=1}^{n} f(x^{(i)}, y_t) + \sqrt{\frac{\ln N}{2n}} \tag{24.4}$$

So,

$$\begin{aligned}
&\inf_{x\in\mathcal{X}} \sup_{y\in\mathcal{Y}} f(x, y) \\
&\qquad \leq \sup_{y\in\mathcal{Y}} f\big(\frac{1}{n}\sum_{t=1}^{n} x_t, y\big) \\
&\qquad \leq \sup_{y\in\mathcal{Y}} \frac{1}{n}\sum_{t=1}^{n} f(x_t, y) \\
&\qquad \leq \frac{1}{n}\sum_{t=1}^{n} \sup_{y\in\mathcal{Y}} f(x_t, y) \\
&\qquad \leq \frac{1}{n}\sum_{t=1}^{n} f(x_t, y_t) + \frac{1}{n} \\
&\qquad \leq \min_{i=1,\ldots,N} \frac{1}{n}\sum_{t=1}^{n} f(X^{(i)}, y_t) + \sqrt{\frac{\ln N}{2n}} + \frac{1}{n} \\
&\qquad \leq \min_{i=1,\ldots,N} f\big(x^{(i)}, \frac{1}{n}\sum_{t=1}^{n} y_t\big) + \sqrt{\frac{\ln N}{2n}} + \frac{1}{n} \\
&\qquad \leq \sup_{y\in\mathcal{Y}} \min_{i=1,\ldots,N} f\big(x^{(i)}, y\big) + \sqrt{\frac{\ln N}{2n}} + \frac{1}{n}
\end{aligned}$$

Letting $n \to \infty$, we have that

$$\inf_{x\in\mathcal{X}} \sup_{y\in\mathcal{Y}} f(x, y) \leq \sup_{y\in\mathcal{Y}} \min_{i=1,\ldots,N} f(x^{(i)}, y). \tag{24.5}$$

Letting $\epsilon \to 0$, we have that

$$\inf_{x\in\mathcal{X}} \sup_{y\in\mathcal{Y}} f(x, y) \leq \sup_{y\in\mathcal{Y}} \inf_{x\in\mathcal{X}} f(x, y) \tag{24.6}$$

which completes our proof. ■

This theorem implies von Neumann's minimax theorem for two-player zero-sum games because the loss function is bounded, linear, and the simplex of all mixed strategies is a compact set.

## 24.4 Repeated Two-Player Zero-Sum Games

We investigate repeated games starting from two-player zero-sum games. At each round t, based on the past plays of both players, the row player chooses an action $I_t \in \{1, \cdots, N\}$, according to the mixed strategy $p_t = (p_{1,t}, \cdots, p_{N,t})$ and the column player chooses an action $J_t = \{1, \cdots, M\}$ according to the mixed strategy $q_t = (q_{1,t}, \cdots, q_{M,t})$. The distributions $p_t$ and $q_t$ may depend on the past plays of both. The row players loss at time $t$ is $l(I_t, J_t)$ and the column players loss is $(I_t, J_t)$. At each time instant, after making the play, the row player observes the losses $l(i, J_t)$ he would have suffered had he played strategy $i, i = 1, \cdots, N$

We consider the view point of row player, he/she wants to minimize the difference between the his/her cumulative loss and the cumulative loss of the best constant strategy.

$$\sum_{t=1}^{n} l(I_t, J_t) - \min_{i=1,\cdots,N} \sum_{t=1}^{n} l(i, J_t)$$

A strategy such as the exponentially weighted average mixed strategy:

$$p_{i,t} = \frac{\exp\left(-\eta \sum_{s=1}^{t-1} l(i, J_s)\right)}{\sum_{k=1}^{N} \exp\left(-\eta \sum_{s=1}^{t-1} l(k, J_s)\right)}, i = 1, \cdots, N, \eta > 0$$

Could make the above difference grow sub-linearly almost surely. More precisely, according to Hannan-Consistency:

$$\limsup_{n\to\infty} \left( \sum_{t=1}^{n} l(I_t, J_t) - \min_{i=1,\cdots,N} \sum_{t=1}^{n} l(i, J_t) \right) \leq 0, a.s.$$

We now show the following remarkable fact: if the row player plays according to any Hannan-Consistent strategy, then his average loss cannot be much larger than the value of the game, regardless of the opponent's strategy.

Recall that the value of the game characterized by the loss matrix $l$ is defined as:

$$V = \max_{\mathbf{p}} \min_{\mathbf{q}} \bar{l}(\mathbf{p}, \mathbf{q}) = \max_{\mathbf{p}} \min_{\mathbf{q}} \sum_{i=1}^{N} \sum_{j=1}^{M} p_i q_j l(i, j)$$

We overload the notation, and also write:

$$\bar{l}(\mathbf{p}, j) = \sum_{i=1}^{N} p_i l(i, j), \quad \bar{l}(i, \mathbf{q}) = \sum_{j=1}^{M} q_j l(i, j)$$

**Theorem 24.7** *Assume that in a two-person zero-sum game the row player plays according to a Hannan-Consistent strategy, then:*

$$\limsup_{n\to\infty} \frac{1}{n} \sum_{t=1}^{n} l(I_t, J_t) \leq V$$

**Proof:** Per Hannan Consistency, we only need to prove that

$$\min_{i=1,\cdots,N} l(i, J_t) \leq V$$

Since the minimization of a linear function over a simplex is the same as the minimization over the simplex's corners only, we have:

$$\min_{i=1,\cdots,N} \frac{1}{n} \sum_{t=1}^{n} l(i, J_t) = \min_{\mathbf{p}\in\Delta} \frac{1}{n} \sum_{t=1}^{n} \bar{l}(\mathbf{p}, J_t)$$

where $\frac{1}{n}\sum_{t=1}^{n} l(i, J_t)$ are the values that function $\frac{1}{n}\sum_{t=1}^{n} \bar{l}(\mathbf{p}, J_t)$ at each simplex corner.
Let $\hat{q}_{j,n} = \sum_{t=1}^{n}$ be the empirical probability of the column player's action being $j$:

$$\begin{aligned}
\min_{\mathbf{p}} \frac{1}{n} \sum_{t=1}^{n} \bar{l}(\mathbf{p}, J_t) &= \min_{\mathbf{p}} \sum_{j=1}^{M} \hat{q}_{j,n} \bar{l}(\mathbf{p}, j) \\
&= \min_{\mathbf{p}} \bar{l}(\mathbf{p}, \hat{\mathbf{q}}_n), \text{ where } \hat{\mathbf{q}}_n = (\hat{q}_{1,n}, \cdots, \hat{q}_{M,n} \\
&\leq \max_{\mathbf{q}} \min_{\mathbf{p}} \bar{l}(\mathbf{p}, \mathbf{q}) = V
\end{aligned}$$

■

**Corollary 24.8** *If both the row player and the column player follows some Hannan-Consistent strategy (e.g. exponentially weighted average mixed strategy). Then*

$$\lim_{n\to\infty} \frac{1}{n}\sum_{t=1}^{n} l(I_t, J_t) = V \quad a.s.$$

**Proof:** By theorem 24.7

$$\limsup_{n\to\infty} \frac{1}{n}\sum_{t=1}^{n} l(I_t, J_t) \le V, \quad a.s. \tag{1}$$

Then we apply the same theorem to the column player:

$$\liminf_{n\to\infty} \frac{1}{n}\sum_{t=1}^{n} l(I_t, J_t) \ge \min_{\mathbf{p}} \max_{\mathbf{q}} \bar{l}(\mathbf{p}, \mathbf{q}) = V \quad a.s. \tag{2}$$

Combine (1) and (2) gives us:

$$\lim_{n\to\infty} \frac{1}{n}\sum_{t=1}^{n} l(I_t, J_t) = V \quad a.s.$$

■

**Corollary 24.9** *If both players follow some Hannan-Consistent strategy, then the product distribution $\hat{\mathbf{p}}_n \times \hat{\mathbf{q}}_n$ formed by the empirical distribution of play*

$$\hat{p}_{i,n} = \frac{1}{n}\sum_{t=1}^{n} \mathbb{1}[I_t = i], \quad \hat{q}_{j,n} = \frac{1}{n}\sum_{t=1}^{n} \mathbb{1}[I_t = j]$$

*of the two players converges a.s. to the set of Nash Equilibria $\pi = \mathbf{p} \times \mathbf{q}$ of the game.*

**Proof:** By 24.9, we know that

$$\lim_{n\to\infty} \bar{l}(\hat{\mathbf{p}}_n, \hat{\mathbf{q}}_n) = V \quad a.s.$$

By 24.5, the empirical distribution of play will also be Nash Equilibrium strategy a.s. ■