

L08

Online Learning and a proof of Minimax Theorem

CS 280 Algorithmic Game Theory

Ioannis Panageas

Playing the experts game

Definition. For each day $t = 1 \dots T$, you have to choose between alternatives A, B (e.g., rain or not rain).

- Choose A or B according to some rule.
- One of the alternatives realizes.
- If you choose *correctly* you are *not penalized* otherwise you *lose one point*.
- Imagine that there are n *experts* who on each day t , recommend either A or B .

Playing the experts game

Definition. For each day $t = 1 \dots T$, you have to choose between alternatives A, B (e.g., rain or not rain).

- Choose A or B according to some rule.
- One of the alternatives realizes.
- If you choose *correctly* you are *not penalized* otherwise you *lose one point*.
- Imagine that there are n *experts* who on each day t , recommend either A or B .

Can you be *correct all the time*? What is the “*right*” objective?

Playing the experts game

Definition. For each day $t = 1 \dots T$, you have to choose between alternatives A, B (e.g., rain or not rain).

- Choose A or B according to some rule.
- One of the alternatives realizes.
- If you choose *correctly* you are *not penalized* otherwise you *lose one point*.
- Imagine that there are n *experts* who on each day t , recommend either A or B .

Can you be *correct all the time*? What is the “*right*” objective?

Perform close to best expert!

Playing the experts game

Algorithm (**Weighted Majority**). *We define the following algorithm:*

1. Initialize $w_i^0 = 1$ for all $i \in [n]$.
2. **For** $t=1 \dots T$ **do**
3. **If** $\sum_{i \text{ choose } A} w_i^{t-1} \geq \sum_{i \text{ choose } B} w_i^{t-1}$
4. **Choose** A , **otherwise** B .
5. **End If**
6. **For** expert i that made a mistake **do**
7. $w_i^t = (1 - \epsilon)w_i^{t-1}$.
8. **End For**
9. **For** expert i that did not make a mistake **do**
10. $w_i^t = w_i^{t-1}$.
11. **End For**
12. **End For**

Remarks:

- ϵ is the **stepsize** (to be chosen later).
- Performs almost as good as “**best**” expert (fewest mistakes)

We assume $0 < \epsilon \leq 1/2$.

Playing the experts game

Theorem (Weighted Majority). *Let M_T , M_T^B be the total number of mistakes the algorithm and best expert make until step T , respectively. It holds that*

$$M_T \leq 2(1 + \epsilon)M_T^B + \frac{2 \log n}{\epsilon}.$$

Proof. Let's define the **potential** function $\phi_t = \sum_i w_i^t$.

Playing the experts game

Theorem (Weighted Majority). *Let M_T , M_T^B be the total number of mistakes the algorithm and best expert make until step T , respectively. It holds that*

$$M_T \leq 2(1 + \epsilon)M_T^B + \frac{2 \log n}{\epsilon}.$$

Proof. Let's define the **potential** function $\phi_t = \sum_i w_i^t$.

- $\phi_0 = n$.
- $\phi_{t+1} \leq \phi_t$ (why?).

Playing the experts game

Theorem (Weighted Majority). Let M_T , M_T^B be the total number of mistakes the algorithm and best expert make until step T , respectively. It holds that

$$M_T \leq 2(1 + \epsilon)M_T^B + \frac{2 \log n}{\epsilon}.$$

Proof. Let's define the **potential** function $\phi_t = \sum_i w_i^t$.

- $\phi_0 = n$.
- $\phi_{t+1} \leq \phi_t$ (why?).

Observe that if we make a mistake at time $t + 1$, then the majority was wrong, that is at least $\frac{\phi_t}{2}$ will be multiplied by $(1 - \epsilon)$.

Hence, if we make a mistake then $\phi_{t+1} \leq (1 - \epsilon)\frac{\phi_t}{2} + \frac{\phi_t}{2} = (1 - \frac{\epsilon}{2})\phi_t$

Playing the experts game

Theorem (Weighted Majority). Let M_T , M_T^B be the total number of mistakes the algorithm and best expert make until step T , respectively. It holds that

$$M_T \leq 2(1 + \epsilon)M_T^B + \frac{2 \log n}{\epsilon}.$$

Proof. Let's That is $\phi_{t+1} \leq (1 - \frac{\epsilon}{2})\phi_t$ when we do a mistake, otherwise just $\phi_{t+1} \leq \phi_t$. Since we have M_T mistakes, then

- ϕ_0

- ϕ_t

$$\phi_T \leq \left(1 - \frac{\epsilon}{2}\right)^{M_T} \phi_0.$$

Observe that if ϕ_t is at least $\frac{\phi_0}{2}$ will be multiplied by $(1 - \epsilon)$.

Hence, if we make a mistake then $\phi_{t+1} \leq (1 - \epsilon)\frac{\phi_t}{2} + \frac{\phi_t}{2} = (1 - \frac{\epsilon}{2})\phi_t$

Playing the experts game

Proof cont. Moreover, assuming the best expert (say i^*) did M_T^B mistakes, we have

$$\phi_T \geq w_{i^*}^T = (1 - \epsilon)^{M_T^B}.$$

Playing the experts game

Proof cont. Moreover, assuming the best expert (say i^*) did M_T^B mistakes, we have

$$\phi_T \geq w_{i^*}^T = (1 - \epsilon)^{M_T^B}.$$

We conclude that

$$(1 - \epsilon)^{M_T^B} \leq \left(1 - \frac{\epsilon}{2}\right)^{M_T} n.$$

By taking the log, $M_T^B \log(1 - \epsilon) \leq \log(1 - \epsilon/2)M_T + \log n$.

Playing the experts game

Proof cont. Moreover, assuming the best expert (say i^*) did M_T^B mistakes, we have

$$\phi_T \geq w_{i^*}^T = (1 - \epsilon)^{M_T^B}.$$

We conclude that

$$(1 - \epsilon)^{M_T^B} \leq \left(1 - \frac{\epsilon}{2}\right)^{M_T} n.$$

By taking the log, $M_T^B \log(1 - \epsilon) \leq \log(1 - \epsilon/2)M_T + \log n$.

Since $-x - x^2 \leq \log(1 - x) \leq -x$, $M_T^B(-\epsilon - \epsilon^2) \leq -M_T\epsilon/2 + \log n$.

The general setting

Definition. *At each time step $t = 1 \dots T$.*

- *Player* chooses $x_t \in \Delta_n$.
- *Adversary* chooses $u_t \in [-1, 1]^n$.
- *Player* gets payoff $x_t^\top u_t$ and observes u_t .

The general setting

Definition. At each time step $t = 1 \dots T$.

- *Player* chooses $x_t \in \Delta_n$.
- *Adversary* chooses $u_t \in [-1, 1]^n$.
- *Player* gets payoff $x_t^\top u_t$ and observes u_t .

Player's goal is to minimize the (time average) **Regret**, that is:

$$\begin{aligned} & \frac{1}{T} \left[\max_{x \in \Delta_n} \sum_{t=1}^T x^\top u_t - \sum_{t=1}^T x_t^\top u_t \right] . \\ &= \frac{1}{T} \left[\max_{i^* \in [n]} \sum_{t=1}^T u_{t,i^*} - \sum_{t=1}^T x_t^\top u_t \right] . \end{aligned}$$

No-regret if $\limsup \text{Regret} \leq 0$ as $T \rightarrow \infty$.

Multiplicative Weights Update

Algorithm (MWU). *We define the following algorithm:*

1. Initialize $p_i^1 = \frac{1}{n}$ for all $i \in [n]$.
2. **For** $t=1 \dots T$ **do**
3. **For** each i that gives payoff $u_{t,i}$ **do**
4. $p_i^{t+1} = p_i^t \frac{1 + \epsilon u_{t,i}}{Z^t}$.
5. **End For**
6. **End For**

Remarks:

- ϵ is the **stepsize** (to be chosen later).
- Performs almost as good as “**best**” expert (fewest mistakes).
- The algorithm is also called **Multiplicative Weights Update!**
- $Z^t = \sum_i p_i^t (1 + \epsilon u_{t,i})$ is **renormalization** constant.

We assume $0 < \epsilon \leq 1/2$.

Multiplicative Weights Update

Theorem (MWU). *It holds that*

$$\frac{1}{T} \sum_t u_t^\top p^t \geq \max_{x \in \Delta_n} \frac{1}{T} \sum_t x^\top u_t - \frac{\log n}{\epsilon T} - \epsilon.$$

Proof. Let's define the **potential** function $\phi_t = \sum_i w_i^t$ where $w_i^t = \prod_{s=1}^{t-1} (1 + \epsilon u_{s,i})$.

Multiplicative Weights Update

Theorem (MWU). *It holds that*

$$\frac{1}{T} \sum_t u_t^\top p^t \geq \max_{x \in \Delta_n} \frac{1}{T} \sum_t x^\top u_t - \frac{\log n}{\epsilon T} - \epsilon.$$

Proof. Let's define the **potential** function $\phi_t = \sum_i w_i^t$ where $w_i^t = \prod_{s=1}^{t-1} (1 + \epsilon u_{s,i})$.

Let the best strategy be i^* , we have

$$\phi_{T+1} \geq w_{i^*}^{T+1} \geq e^{\epsilon \sum_{s=1}^T u_{s,i^*} - \epsilon^2 \sum_{s=1}^T u_{s,i^*}^2}.$$

$$\text{Now } \phi_{t+1} = \sum w_i^{t+1} = \sum w_i^t (1 + \epsilon u_{t,i})$$

Multiplicative Weights Update

Theorem (MWU). *It holds that*

$$\frac{1}{T} \sum_t u_t^\top p^t \geq \max_{x \in \Delta_n} \frac{1}{T} \sum_t x^\top u_t - \frac{\log n}{\epsilon T} - \epsilon.$$

Proof. Let's define the **potential** function $\phi_t = \sum_i w_i^t$ where $w_i^t = \prod_{s=1}^{t-1} (1 + \epsilon u_{s,i})$.

Let the best strategy be i^* , we have

$$\phi_{T+1} \geq w_{i^*}^{T+1} \geq e^{\epsilon \sum_{s=1}^T u_{s,i^*} - \epsilon^2 \sum_{s=1}^T u_{s,i^*}^2}.$$

$$\begin{aligned} \text{Now } \phi_{t+1} &= \sum w_i^{t+1} = \sum w_i^t (1 + \epsilon u_{t,i}) \\ &= \sum \phi_t p_i^t (1 + \epsilon u_{t,i}) \end{aligned}$$

Multiplicative Weights Update

Theorem (MWU). *It holds that*

$$\frac{1}{T} \sum_t u_t^\top p^t \geq \max_{x \in \Delta_n} \frac{1}{T} \sum_t x^\top u_t - \frac{\log n}{\epsilon T} - \epsilon.$$

Proof. Let's define the **potential** function $\phi_t = \sum_i w_i^t$ where $w_i^t = \prod_{s=1}^{t-1} (1 + \epsilon u_{s,i})$.

Let the best strategy be i^* , we have

$$\phi_{T+1} \geq w_{i^*}^{T+1} \geq e^{\epsilon \sum_{s=1}^T u_{s,i^*} - \epsilon^2 \sum_{s=1}^T u_{s,i^*}^2}.$$

$$\begin{aligned} \text{Now } \phi_{t+1} &= \sum w_i^{t+1} = \sum w_i^t (1 + \epsilon u_{t,i}) \\ &= \sum \phi_t p_i^t (1 + \epsilon u_{t,i}) \\ &= \phi_t \sum p_i^t (1 + \epsilon u_{t,i}) \end{aligned}$$

Multiplicative Weights Update

Proof cont. Therefore

$$\phi_{t+1} = \phi_t \left(1 + \epsilon \sum_i p_i^t u_{t,i} \right)$$

Multiplicative Weights Update

Proof cont. Therefore

$$\begin{aligned}\phi_{t+1} &= \phi_t \left(1 + \epsilon \sum_i p_i^t u_{t,i} \right) \\ &\leq \phi_t e^{\epsilon \sum_i p_i^t u_{t,i}} = \phi_t e^{\epsilon u_t^\top p^t}\end{aligned}$$

Multiplicative Weights Update

Proof cont. Therefore

$$\begin{aligned}\phi_{t+1} &= \phi_t \left(1 + \epsilon \sum_i p_i^t u_{t,i} \right) \\ &\leq \phi_t e^{\epsilon \sum_i p_i^t u_{t,i}} = \phi_t e^{\epsilon u_t^\top p^t}\end{aligned}$$

Telescopic product gives

$$\phi_{T+1} \leq \phi_1 e^{\epsilon \sum_t u_t^\top p^t} = n e^{\epsilon \sum_t u_t^\top p^t}.$$

Multiplicative Weights Update

Proof cont. Therefore

$$\begin{aligned}\phi_{t+1} &= \phi_t \left(1 + \epsilon \sum_i p_i^t u_{t,i} \right) \\ &\leq \phi_t e^{\epsilon \sum_i p_i^t u_{t,i}} = \phi_t e^{\epsilon u_t^\top p^t}\end{aligned}$$

Telescopic product gives

$$\phi_{T+1} \leq \phi_1 e^{\epsilon \sum_t u_t^\top p^t} = n e^{\epsilon \sum_t u_t^\top p^t}.$$

Therefore $e^{\epsilon \sum_{s=1}^T u_{s,i^*}} - \epsilon^2 \sum_{s=1}^T u_{s,i^*}^2 \leq n e^{\epsilon \sum_t u_t^\top p^t}$, or equivalently

$$\epsilon OPT - \epsilon^2 T \leq \epsilon OPT - \epsilon^2 \sum_{s=1}^T u_{s,i^*}^2 \leq \log n + \epsilon \sum_t u_t^\top p^t.$$

Multiplicative Weights Update

Proof cont. Therefore

$$\text{Set } \varepsilon \rightarrow \sqrt{\frac{\ln n}{T}} \text{ and we get regret}$$
$$2 \sqrt{\frac{\ln n}{T}} \text{ (No-regret!)}$$

For $n \geq 2$ and $T \geq 4 \log n$.

Telescopic product gives

$$\phi_{T+1} \leq \phi_1 e^{\epsilon \sum_t u_t^\top p^t} = n e^{\epsilon \sum_t u_t^\top p^t}.$$

Therefore $e^{\epsilon \sum_{s=1}^T u_{s,i^*}} - \epsilon^2 \sum_{s=1}^T u_{s,i^*}^2 \leq n e^{\epsilon \sum_t u_t^\top p^t}$, or equivalently

$$\epsilon OPT - \epsilon^2 T \leq \epsilon OPT - \epsilon^2 \sum_{s=1}^T u_{s,i^*}^2 \leq \log n + \epsilon \sum_t u_t^\top p^t.$$

No-regret learning and CCE

Let $s^1, \dots, s^T \in S$ be the joint action profiles played in T rounds. Their **empirical joint distribution** is

$$\chi_T(s) = \frac{1}{T} |\{t \in [T] : s^t = s\}|.$$

Equivalently, choose one past round uniformly at random.

If each player's average external regret is at most $\varepsilon \geq 0$, then for every i and fixed action $b' \in S_i$,

$$\begin{aligned} \mathbb{E}_{s \sim \chi_T} [u_i(b', s_{-i}) - u_i(s)] \\ = \frac{1}{T} \sum_{t=1}^T [u_i(b', s_{-i}^t) - u_i(s^t)] \leq \varepsilon. \end{aligned}$$

Hence χ_T is an **ε -coarse correlated equilibrium**. No-regret learning gives $\varepsilon \rightarrow 0$ as $T \rightarrow \infty$.

Minimax Theorem

Theorem (**Minimax** by John von Neumann). *Let A a matrix of size $n \times m$.*

$$\min_{x \in \Delta_n} \max_{y \in \Delta_m} x^\top A y = \max_{y \in \Delta_m} \min_{x \in \Delta_n} x^\top A y$$

Remarks

- The above is the **value** of the game.
- Note that It is always true (min-max inequality):

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

Minimax Theorem

Theorem (**Minimax** by John von Neumann). *Let A a matrix of size $n \times m$.*

$$\min_{x \in \Delta_n} \max_{y \in \Delta_m} x^\top A y = \max_{y \in \Delta_m} \min_{x \in \Delta_n} x^\top A y$$

Remarks

- The above is the **value** of the game.
- Note that It is always true (min-max inequality):

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

Define $g(z) \triangleq \inf_{w \in W} f(z, w)$.

$$\forall w, \forall z, g(z) \leq f(z, w)$$

$$\implies \forall w, \sup_z g(z) \leq \sup_z f(z, w)$$

$$\implies \sup_z g(z) \leq \inf_w \sup_z f(z, w)$$

$$\implies \sup_z \inf_w f(z, w) \leq \inf_w \sup_z f(z, w)$$

Minimax Theorem

Theorem (**Minimax** by John von Neumann). *Let A a matrix of size $n \times m$.*

$$\min_{x \in \Delta_n} \max_{y \in \Delta_m} x^\top A y = \max_{y \in \Delta_m} \min_{x \in \Delta_n} x^\top A y$$

Proof. Let's use **no-regret learning** for both "players"!

After scaling, assume A has entries in $[-1, 1]$.

The row player minimizes and the column player maximizes.

Their reward vectors are $-Ay_t$ and $A^\top x_t$.

Minimax Theorem

Theorem (**Minimax** by John von Neumann). *Let A a matrix of size $n \times m$.*

$$\min_{x \in \Delta_n} \max_{y \in \Delta_m} x^\top A y = \max_{y \in \Delta_m} \min_{x \in \Delta_n} x^\top A y$$

Proof. Let's use **no-regret learning** for both "players"!

Let $x_1, \dots, x_T$ and $y_1, \dots, y_T$ be the MWU iterates. We define $\hat{x} = \frac{1}{T} \sum_{i=1}^T x_i$ and $\hat{y} = \frac{1}{T} \sum_{i=1}^T y_i$, with $T \geq \frac{4 \log(\max\{2, n, m\})}{\eta^2}$ and $\epsilon = \eta/2$.

Choose any x , then from the **no-regret** property for x we get that

Minimax Theorem

Theorem (**Minimax** by John von Neumann). *Let A a matrix of size $n \times m$.*

$$\min_{x \in \Delta_n} \max_{y \in \Delta_m} x^\top A y = \max_{y \in \Delta_m} \min_{x \in \Delta_n} x^\top A y$$

Proof. Let's use **no-regret learning** for both "players"!

Let $x_1, \dots, x_T$ and $y_1, \dots, y_T$ be the MWU iterates. We define $\hat{x} = \frac{1}{T} \sum_{i=1}^T x_i$ and $\hat{y} = \frac{1}{T} \sum_{i=1}^T y_i$, with $T \geq \frac{4 \log(\max\{2, n, m\})}{\eta^2}$ and $\epsilon = \eta/2$.

Choose any x , then from the **no-regret** property for x we get that

$$\frac{1}{T} \sum_t x_t^\top A y_t \leq \frac{1}{T} \sum_t x^\top A y_t + \eta$$

Minimax Theorem

Theorem (**Minimax** by John von Neumann). *Let A a matrix of size $n \times m$.*

$$\min_{x \in \Delta_n} \max_{y \in \Delta_m} x^\top A y = \max_{y \in \Delta_m} \min_{x \in \Delta_n} x^\top A y$$

Proof. Let's use **no-regret learning** for both "players"!

Let $x_1, \dots, x_T$ and $y_1, \dots, y_T$ be the MWU iterates. We define $\hat{x} = \frac{1}{T} \sum_{i=1}^T x_i$ and $\hat{y} = \frac{1}{T} \sum_{i=1}^T y_i$, with $T \geq \frac{4 \log(\max\{2, n, m\})}{\eta^2}$ and $\epsilon = \eta/2$.

Choose any x , then from the **no-regret** property for x we get that

$$\begin{aligned} \frac{1}{T} \sum_t x_t^\top A y_t &\leq \frac{1}{T} \sum_t x^\top A y_t + \eta \\ &= x^\top A \left(\frac{\sum_t y_t}{T} \right) + \eta. \end{aligned}$$

Minimax Theorem

Proof cont.

Choose any y , then from the **no-regret** property for y we get that

$$\begin{aligned}\frac{1}{T} \sum_t x_t^\top A y_t &\geq \frac{1}{T} \sum_t x_t^\top A y - \eta \\ &= \left(\frac{\sum x_t}{T} \right)^\top A y - \eta.\end{aligned}$$

Minimax Theorem

Proof cont.

Choose any y , then from the **no-regret** property for y we get that

$$\begin{aligned}\frac{1}{T} \sum_t x_t^\top A y_t &\geq \frac{1}{T} \sum_t x_t^\top A y - \eta \\ &= \left(\frac{\sum x_t}{T} \right)^\top A y - \eta.\end{aligned}$$

We conclude that for all x, y we have

$$\left(\frac{\sum x_t}{T} \right)^\top A y - 2\eta \leq x^\top A \left(\frac{\sum_t y_t}{T} \right).$$

Minimax Theorem

Proof cont.

Choose any y , then from the **no-regret** property for y we get that

$$\begin{aligned}\frac{1}{T} \sum_t x_t^\top A y_t &\geq \frac{1}{T} \sum_t x_t^\top A y - \eta \\ &= \left(\frac{\sum x_t}{T} \right)^\top A y - \eta.\end{aligned}$$

We conclude that for all x, y we have

$$\max_y \left(\frac{\sum x_t}{T} \right)^\top A y - 2\eta \leq \min_x x^\top A \left(\frac{\sum_t y_t}{T} \right).$$

Finally we get $\max_y \min_x x^\top A y \geq \min_x x^\top A \left(\frac{\sum_y y_t}{T} \right)$

Minimax Theorem

Proof cont.

Choose any y , then from the **no-regret** property for y we get that

$$\begin{aligned}\frac{1}{T} \sum_t x_t^\top A y_t &\geq \frac{1}{T} \sum_t x_t^\top A y - \eta \\ &= \left(\frac{\sum x_t}{T} \right)^\top A y - \eta.\end{aligned}$$

We conclude that for all x, y we have

$$\max_y \left(\frac{\sum x_t}{T} \right)^\top A y - 2\eta \leq \min_x x^\top A \left(\frac{\sum y_t}{T} \right).$$

$$\begin{aligned}\text{Finally we get } \max_y \min_x x^\top A y &\geq \min_x x^\top A \left(\frac{\sum y_t}{T} \right) \\ &\geq \max_y \left(\frac{\sum x_t}{T} \right)^\top A y - 2\eta\end{aligned}$$

Minimax Theorem

Proof cont.

Choose any y , then from the **no-regret** property for y we get that

$$\begin{aligned}\frac{1}{T} \sum_t x_t^\top A y_t &\geq \frac{1}{T} \sum_t x_t^\top A y - \eta \\ &= \left(\frac{\sum x_t}{T} \right)^\top A y - \eta.\end{aligned}$$

We conclude that for all x, y we have

$$\max_y \left(\frac{\sum x_t}{T} \right)^\top A y - 2\eta \leq \min_x x^\top A \left(\frac{\sum y_t}{T} \right).$$

$$\begin{aligned}\text{Finally we get } \max_y \min_x x^\top A y &\geq \min_x x^\top A \left(\frac{\sum y_t}{T} \right) \\ &\geq \max_y \left(\frac{\sum x_t}{T} \right)^\top A y - 2\eta \\ &\geq \min_x \max_y x^\top A y - 2\eta\end{aligned}$$

Minimax Theorem

Proof cont.

Choose an

Set $\eta \rightarrow 0$ and we are done!

$$\begin{aligned}\frac{1}{T} \sum_t x_t^\top A y_t &\geq \frac{1}{T} \sum_t x_t^\top A y - \eta \\ &= \left(\frac{\sum x_t}{T} \right)^\top A y - \eta.\end{aligned}$$

We conclude that for all x, y we have

$$\max_y \left(\frac{\sum x_t}{T} \right)^\top A y - 2\eta \leq \min_x x^\top A \left(\frac{\sum_t y_t}{T} \right).$$

$$\begin{aligned}\text{Finally we get } \max_y \min_x x^\top A y &\geq \min_x x^\top A \left(\frac{\sum_y y_t}{T} \right) \\ &\geq \max_y \left(\frac{\sum_t x_t}{T} \right)^\top A y - 2\eta \\ &\geq \min_x \max_y x^\top A y - 2\eta\end{aligned}$$