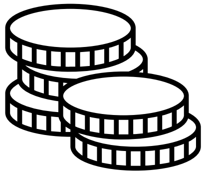
Games and Multiplicative Weight Updates

Yury Makarychev
TTIC and the University of Chicago

Two-player zero-sum games

Two players A and B play the following game. Each of the players has a set of possible moves (strategies) $\mathcal{A}$ and $\mathcal{B}$. A and B choose: $a \in \mathcal{A}, b \in \mathcal{B}$.



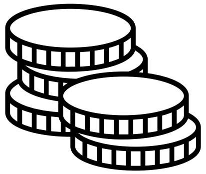
A

payoff function
$f(a, b)$

B

Two-player zero-sum games

Two players A and B play the following game. Each of the players has a set of possible moves (strategies) A and B. A and B choose: $a \in A, b \in B$.



A

payoff function

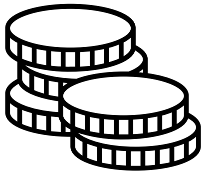
$f(\cdot\,,\cdot)$

B

Is it better to move first or second?

Two-player zero-sum games

Two players A and B play the following game. Each of the players has a set of possible moves (strategies) A and B. A and B choose: $a \in A, b \in B$.



A

payoff function

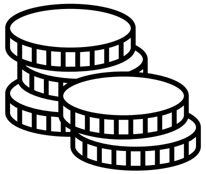
$f(\ ,\ )$

B

Is it better to move first or second?

Two-player zero-sum games

Two players A and B play the following game. Each of the players has a set of possible moves (strategies) A and B. A and B choose: $a \in A, b \in B$.



A

payoff function
$f(a, b)$

B

Is it better to move first or second?

Two-player zero-sum games

A

payoff function
$f(a, b)$

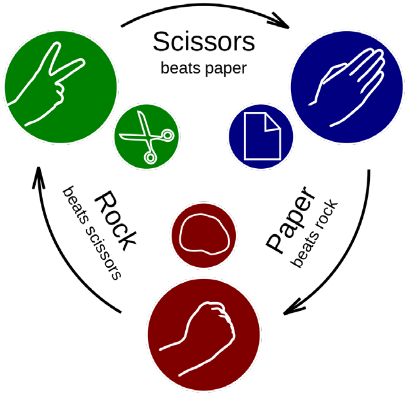
B

Is it better to move first or second?

$$\max_{a \in A} \min_{b \in B} f(a, b) \quad \text{vs} \quad \min_{b \in B} \max_{a \in A} f(a, b)$$

Thoughts?

Rock – Paper – Scissors



B \ A	R	P	S
R	0	1	−1
P	−1	0	1
S	1	−1	0

How do people play RPS in real life?

Randomization

A mixed strategy is probability distribution over pure strategies.

The payout for a mixed strategies α and β is

$$\mathbb{E}_{\substack{a \sim \alpha \\ b \sim \beta}} [f(a, b)]$$

Von Neumann Minimax Theorem

$$\max_{\alpha} \min_{\beta} \mathbb{E}[f(a, b)] = \min_{\beta} \max_{\alpha} \mathbb{E}[f(a, b)]$$

! The order of moves doesn't matter.

Randomization

A mixed strategy is probability distribution over pure strategies.

The payout for a mixed strategies α and β is

$$\mathbb{E}_{\substack{a\sim\alpha \\ b\sim\beta}}[f(a,b)]$$

Von Neumann Minimax Theorem

$$\max_{\alpha}\min_{\beta}\mathbb{E}[f(a,b)] = \min_{\beta}\max_{\alpha}\mathbb{E}[f(a,b)]$$

! The order of moves doesn't matter.

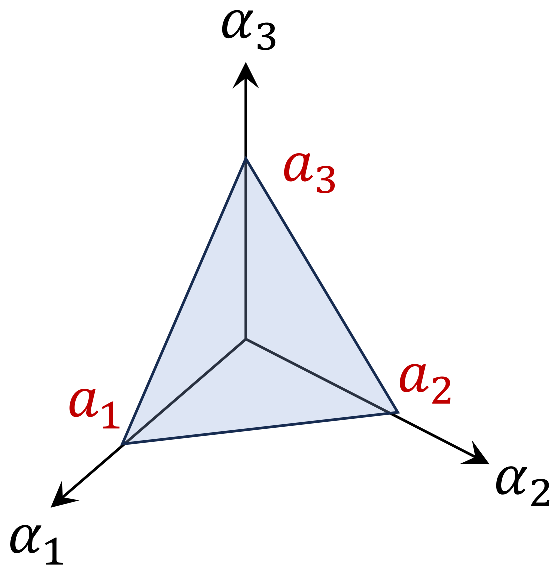
Second player

Q: Does the second player need to randomize?

A plays mixed strategy α.

To minimize f, can B play a pure strategy b?
Or does B have to play a mixed strategy β?

Geometric view



Let $a_1, \dots, a_m$ be pure strategies of A.

A mixed strategy is $\alpha = \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_m \end{pmatrix} \in \Delta_m$, where Δ_m is the unit simplex

$$f(\alpha, \beta) \equiv \mathbb{E} f(a, b) = \sum_{i,j} \alpha_i \beta_j f(a_i, b_j) = \alpha^T F \beta$$

where $F_{ij} = f(a_i, b_j)$

Geometric view

Mixed strategies: Δ_m and Δ_n, which are convex and compact sets.

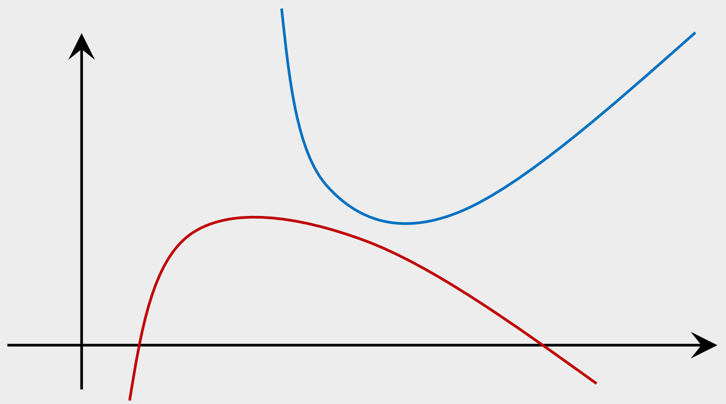
The objective $f(\alpha, \beta) = \alpha^T F \beta$ is linear in each argument.

Von Neumann Minimax Theorem (general)

Assume that $A \subseteq \mathbb{R}^m$ and $B \subseteq \mathbb{R}^m$ are convex and compact.

- f is continuous
- $a \mapsto f(a, b)$ is concave for every b
- $b \mapsto f(a, b)$ is convex for every a

Then $\max_{a \in A} \min_{b \in B} f(a, b) = \min_{b \in B} \max_{a \in A} f(a, b)$



Finding optimal strategies using LP

Player A:

LP variables: $\alpha_1, \dots, \alpha_m$

maximize p

s.t.

$$(\beta^T F^T)\alpha \geq p \quad \text{for every } \beta \in \Delta_n$$

$$\sum_{i=1}^{m} \alpha_i = 1$$

$$\alpha_i \geq 0$$

There are infinitely many β.
Can we consider only some of them?

Finding optimal strategies using LP

Player A:

LP variables: $\alpha_1, \dots, \alpha_m$

maximize p

s.t.

$$F^T \alpha \geq \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix} \cdot p$$

$$\sum_{i=1}^{m} \alpha_i = 1$$

$$\alpha_i \geq 0$$

Finding optimal strategies using LP

Player A:

LP variables: $\alpha_1, \dots, \alpha_m$

maximize p

s.t.

$$\begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix} \cdot p - F^T \alpha \leq 0$$

$$\sum_{i=1}^{m} \alpha_i = 1$$

$$\alpha_i \geq 0$$

Player B:

LP variables: $\beta_1, \dots, \beta_m$

minimize c

s.t.

$$\begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix} \cdot c - F\beta \geq 0$$

$$\sum_{i=1}^{m} \beta_j = 1$$

$$\beta_j \geq 0$$

Multiplicative weights update

- Suppose that $\alpha \in \Delta_n$
- B is a convex and closed set
- f is linear in α

$$f(\alpha, b) = \sum_i \alpha_i f(a_i, b)$$

- f is concave in b

Idea

- Start with some $\alpha \in \Delta_n$
- Find the best response $b \in B$

How should we augment our strategy α in response to b?

- Increase the probability of those a_i that are good responses to b
- Decrease the probability of those a_i that are bad responses to b

- Repeat

We will assume that $|f(\alpha, b)| \leq 1$ for now.

Algorithm

- Start with some $\alpha^{(1)} = \left(\frac{1}{n}, \dots, \frac{1}{n}\right)^T$
- For $t = 1$ to T
  - Find the best response $b^{(t)} \in B$ to $\alpha^{(t)}$
  - $\alpha_i^{(t+1)} = \left(1 + \varepsilon f\left(a_i, b^{(t)}\right)\right) \alpha_i^{(t)}$
  - normalize: $\alpha^{(t+1)} = \alpha^{(t+1)} / \left\|\alpha^{(t+1)}\right\|_1$
- Return

$$\alpha_{ALG} = \sum_{t=1}^{T} \frac{\alpha^{(t)}}{T} \quad \text{and} \quad b_{\mathrm{ALG}} = \sum_{t=1}^{T} \frac{b^{(t)}}{T}$$

$$\alpha_i^{(t+1)} = \left(1 + \varepsilon f\left(a_i, b^{(t)}\right)\right) \alpha_i^{(t)}$$

Analysis

Define potential

$$w_i^{(1)} = \frac{1}{n} \quad \text{for all } i \in \{1, \dots, m\}$$

$$w_i^{(t+1)} = \left(1 + \varepsilon f\left(a_i, b^{(t)}\right)\right) w_i^{(t)}$$

$$W^{(t)} = \sum_i w_i^{(t)}$$

! The update formula for w_i is the same as for α_i but we don't normalize w_i.

$$\Rightarrow \alpha_i^{(t)} = w_i^{(t)} / W^{(t)}$$

$$w_i^{(t+1)} = \left(1 + \varepsilon f\left(a_i, b^{(t)}\right)\right) w_i^{(t)}$$

Analysis

$$W^{(1)} = w_1^{(1)} + \cdots + w_n^{(1)} = \frac{1}{n} + \cdots + \frac{1}{n} = 1$$

$$W^{(t+1)} = \sum_i \left(1 + \varepsilon f\left(a_i, b^{(t)}\right)\right) w_i^{(t)} = \sum_i w_i^{(t)} + \sum_i \varepsilon f\left(a_i, b^{(t)}\right) w_i^{(t)}$$

$$w_i^{(t+1)} = \left(1 + \varepsilon f\left(a_i, b^{(t)}\right)\right) w_i^{(t)}$$

Analysis

$$W^{(1)} = w_1^{(1)} + \cdots + w_n^{(1)} = \frac{1}{n} + \cdots + \frac{1}{n} = 1$$

$$W^{(t+1)} = \sum_i w_i^{(t)} + \sum_i \varepsilon f\left(a_i, b^{(t)}\right) w_i^{(t)}$$

$$w_i^{(t+1)} = \left(1 + \varepsilon f\left(a_i, b^{(t)}\right)\right) w_i^{(t)}$$

Analysis

$$W^{(1)} = w_1^{(1)} + \cdots + w_n^{(1)} = \frac{1}{n} + \cdots + \frac{1}{n} = 1$$

$$W^{(t+1)} = \sum_i w_i^{(t)} + \sum_i \varepsilon f\left(a_i, b^{(t)}\right) w_i^{(t)} = W^{(t)} + \varepsilon \sum_i f\left(a_i, b^{(t)}\right) W^{(t)} \alpha_i^{(t)}$$

$$= \left(1 + \varepsilon \textstyle\sum_i f\left(a_i, b^{(t)}\right) \alpha_i^{(t)}\right) W^{(t)}$$

$$= \left(1 + \varepsilon\, f\left(\alpha^{(t)}, b^{(t)}\right)\right) W^{(t)}$$

Analysis

$$W^{(1)} = w_1^{(1)} + \dots + w_n^{(1)} = \frac{1}{n} + \dots + \frac{1}{n} = 1$$

$$W^{(t+1)} = \left(1 + \varepsilon\, f\left(\alpha^{(t)}, b^{(t)}\right)\right) W^{(t)}$$

$$W^{(T+1)} = \prod_{t=1}^{T} \left(1 + \varepsilon\, f\left(\alpha^{(t)}, b^{(t)}\right)\right)$$

Use: $e^{x - O(x^2)} \leq 1 + x \leq e^x$ if $x \in \left(-\frac{1}{2}, \frac{1}{2}\right)$

$$e^{x-O(x^2)} \leq 1 + x$$

Analysis

$$W^{(1)} = w_1^{(1)} + \cdots + w_n^{(1)} = \frac{1}{n} + \cdots + \frac{1}{n} = 1$$

$$W^{(t+1)} = \left(1 + \varepsilon\, f\left(\alpha^{(t)}, b^{(t)}\right)\right) W^{(t)}$$

$$\boxed{W^{(T+1)} = \prod_{t=1}^{T} \left(1 + \varepsilon\, f\left(\alpha^{(t)}, b^{(t)}\right)\right) \leq \exp\left(\sum_{t=1}^{T} \varepsilon\, f\left(\alpha^{(t)}, b^{(t)}\right)\right)}$$

Use: $e^{x-O(x^2)} \leq 1 + x \leq e^x$ if $x \in \left(-\frac{1}{2}, \frac{1}{2}\right)$

Analysis

$$W^{(T+1)} = \sum w_i^{(T+1)} \leq \exp\left(\sum_{t=1}^{T} \varepsilon\, f\left(\alpha^{(t)}, b^{(t)}\right)\right)$$

Also,

$$w_i^{(T+1)} = \frac{1}{n} \cdot \prod_{t=1}^{T} \left(1 + \varepsilon f\left(a_i, b^{(t)}\right)\right) \geq \frac{1}{n} \exp\left(\sum_{t=1}^{T} \varepsilon\, f\left(a_i, b^{(t)}\right) - O\left(\varepsilon^2 T\right)\right)$$

Analysis

$$\frac{1}{n}\exp\left(\sum_{t=1}^{T}\varepsilon\, f\left(a_i, b^{(t)}\right) - O\left(\varepsilon^2 T\right)\right) \leq \exp\left(\sum_{t=1}^{T}\varepsilon\, f\left(\alpha^{(t)}, b^{(t)}\right)\right)$$

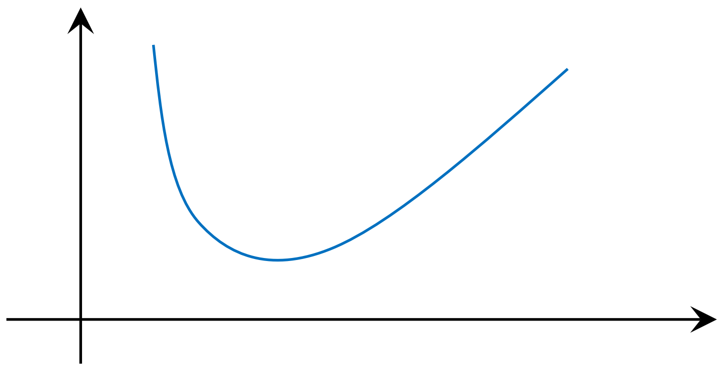
Analysis

$$\sum_{t=1}^{T} \varepsilon\, f\left(a_i, b^{(t)}\right) - O\left(\varepsilon^2 T\right) - \log n \leq \sum_{t=1}^{T} \varepsilon\, f\left(\alpha^{(t)}, b^{(t)}\right)$$

Analysis

$$\frac{1}{T}\sum_{t=1}^{T} f\left(a_i, b^{(t)}\right) \leq \frac{1}{T}\sum_{t=1}^{T} f\left(\alpha^{(t)}, b^{(t)}\right) + O\left(\varepsilon + \frac{\log n}{\varepsilon T}\right)$$

Analysis



$$f(a_i, b_{ALG}) \leq \frac{1}{T}\sum_{t=1}^{T} f\left(a_i, b^{(t)}\right) \leq \frac{1}{T}\sum_{t=1}^{T} f\left(\alpha^{(t)}, b^{(t)}\right) + \delta$$

where $\delta = O\left(\varepsilon + \frac{\log n}{\varepsilon T}\right)$

Analysis

$$f(a_i, b_{ALG}) \leq \frac{1}{T}\sum_{t=1}^{T} f\left(\alpha^{(t)}, b^{(t)}\right) + \delta \quad \text{where } \delta = O\left(\varepsilon + \frac{\log n}{\varepsilon T}\right)$$

- Proving near optimality of b_{ALG}

$$f\left(\alpha^{(t)}, b^{(t)}\right) \leq val$$

Thus,

$$f(a_i, b_{ALG}) \leq val + \delta$$

b_{ALG} is nearly optimal for pure strategies and thus for mixed strategies as well.

Analysis

$$f(a_i, b_{ALG}) \leq \frac{1}{T}\sum_{t=1}^{T} f\left(\alpha^{(t)}, b^{(t)}\right) + \delta \quad \text{where } \delta = O\left(\varepsilon + \frac{\log n}{\varepsilon T}\right)$$

- Proving near optimality of α_{ALG}. Consider a strategy $b \in B$.

$$f\left(\alpha^{(t)}, b^{(t)}\right) \leq f\left(\alpha^{(t)}, b\right)$$

Thus,

$$f(a_i, b_{ALG}) \leq \frac{1}{T}\sum_{t=1}^{T} f\left(\alpha^{(t)}, b\right) + \delta = f(\alpha_{ALG}, b) + \delta$$

Analysis

$$f(a_i, b_{ALG}) \leq f(\alpha_{ALG}, b) + \delta$$

Average with weights α_i^*:

$$val \leq f(\alpha^*, b_{ALG}) \leq f(\alpha_{ALG}, b) + \delta$$

Summary

- $f(a_i, b_{ALG}) \leq val + \delta$
- $f(\alpha_{ALG}, b) \geq val - \delta$

$$\delta = O\left(\varepsilon + \frac{\log n}{T\varepsilon}\right)$$

Q: What T should we choose to get an $O(\varepsilon)$ additive approximation?

Summary

- $f(a_i, b_{ALG}) \leq val + \delta$
- $f(\alpha_{ALG}, b) \geq val - \delta$

$$\delta = O\left(\varepsilon + \frac{\log n}{T\varepsilon}\right)$$

Q: What T should we choose to get an $O(\varepsilon)$ additive approximation?

A: $T = c\varepsilon^{-2} \log n$

Running time: $O(T(M + nP))$, where M is the time for computing the optimal response $b^{(t)}$ and P is the time for computing $f(a_i, b^{(t)})$.

Arbitrary Width ρ

Q: What if $\max|f(a, b)| > 1$? Can we use the algorithm as is?

A: No! In particular, this step may be problematic:

$$\alpha_i^{(t+1)} = \left(1 + \varepsilon f\left(a_i, b^{(t)}\right)\right) \alpha_i^{(t)}$$

Arbitrary Width ρ

Q: What if $\max|f(a, b)| > 1$? Can we use the algorithm as is?

A: No! In particular, this step may be problematic:

$$\alpha_i^{(t+1)} = \left(1 + \varepsilon f\left(a_i, b^{(t)}\right)\right) \alpha_i^{(t)}$$

Solution: rescale f. Assume $f \in [-\rho, \rho]$. ρ is the *width* of the game.

Apply our algorithm to $f' = \frac{f}{\rho} \in [-1,1]$.

Algorithm for arbitrary ρ

- Start with some $\alpha^{(1)} = \left(\frac{1}{n}, \dots, \frac{1}{n}\right)^T$
- For $t = 1$ to T
  - Find the best response $b^{(t)} \in B$ to $\alpha^{(t)}$

$$\alpha_i^{(t+1)} = \left(1 + \frac{\varepsilon f\left(a_i, b^{(t)}\right)}{\rho}\right)\alpha_i^{(t)}$$

  - normalize: $\alpha^{(t+1)} = \alpha^{(t+1)} / \left\|\alpha^{(t+1)}\right\|_1$
- Return $\alpha_{ALG} = \sum_{t=1}^{T} \frac{\alpha^{(t)}}{T}$ and $b_{\mathrm{ALG}} = \sum_{t=1}^{T} \frac{b^{(t)}}{T}$

Algorithm

The algorithm will find an additive $O(\varepsilon)$ approximation for f'

$\Rightarrow$ an additive $O(\varepsilon\rho)$ approximation for f.

Need $\varepsilon' = \varepsilon/\rho$ to get an $O(\varepsilon)$ approximation. Then, the number of iterations

$$T = O\left({\varepsilon'}^{-2} \log n\right) = O(\varepsilon^{-2}\rho^2 \log n)$$

Can we improve the dependence on ρ? If $f \in [0, \rho]$ or $f \in [-O(1), \rho]$

! Get $(1 \pm \varepsilon)val \pm \varepsilon$ guarantee when $T = c\varepsilon^{-2}\rho \log n$.

Revised analysis

Instead of using

$$e^{x-O(x^2)} \leq 1 + x \leq e^x$$

Use

$$e^{(1-\varepsilon)x} \leq 1 + x \leq e^x$$

for $x \in (0, \varepsilon)$.

Exercise: finish the analysis.

Oracles

- The procedure that computes the optimal response $b^{(t)}$ is called an oracle.
- We can use an approximate procedure that computes $b^{(t)}$ s.t.

$$f\left(\alpha^{(t)}, b^{(t)}\right) \leq (1+\varepsilon) \min_{b \in B} f\left(a^{(t)}, b\right) + \varepsilon$$

We will get nearly optimal strategies α_{ALG} and b_{ALG} with a $1 + O(\varepsilon)$ multiplicative and $O(\varepsilon)$ additive error.

Oracles

- Assume that $val \leq 1$ but we only want to find a strategy b_{ALG} s.t.

$$f(\alpha, b_{ALG}) \leq 1 + \varepsilon$$

(even if $val \ll 1$)

Then it is sufficient to use an oracle that finds a response $b^{(t)}$ with

\

$$f\left(\alpha^{(t)}, b^{(t)}\right) \leq 1 + \varepsilon$$

The width of the oracle is $\rho = \max_{i,b} |f(a_i, b)|$, where the maximum is over all possible responses b provided by the oracle.