Advanced Machine Learning

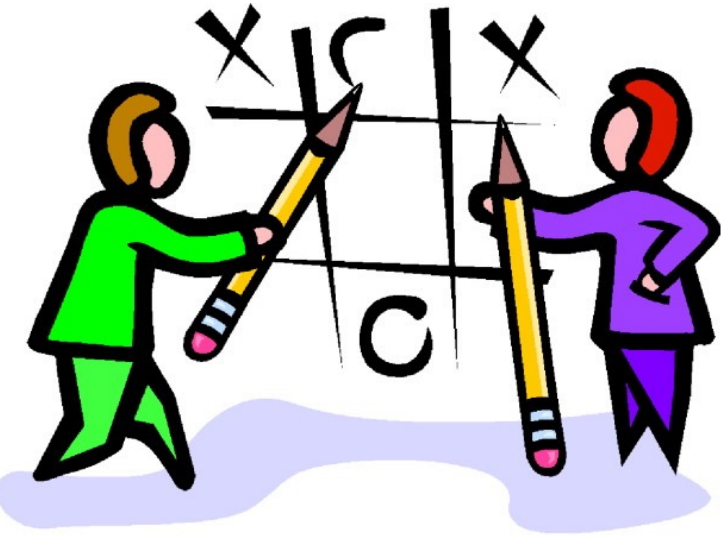
Learning and Games

MEHRYAR MOHRI MOHRI@

COURANT INSTITUTE & GOOGLE RESEARCH

Outline

- Normal form games
- Nash equilibrium
- von Neumann’s minimax theorem
- Correlated equilibrium
- Swap regret



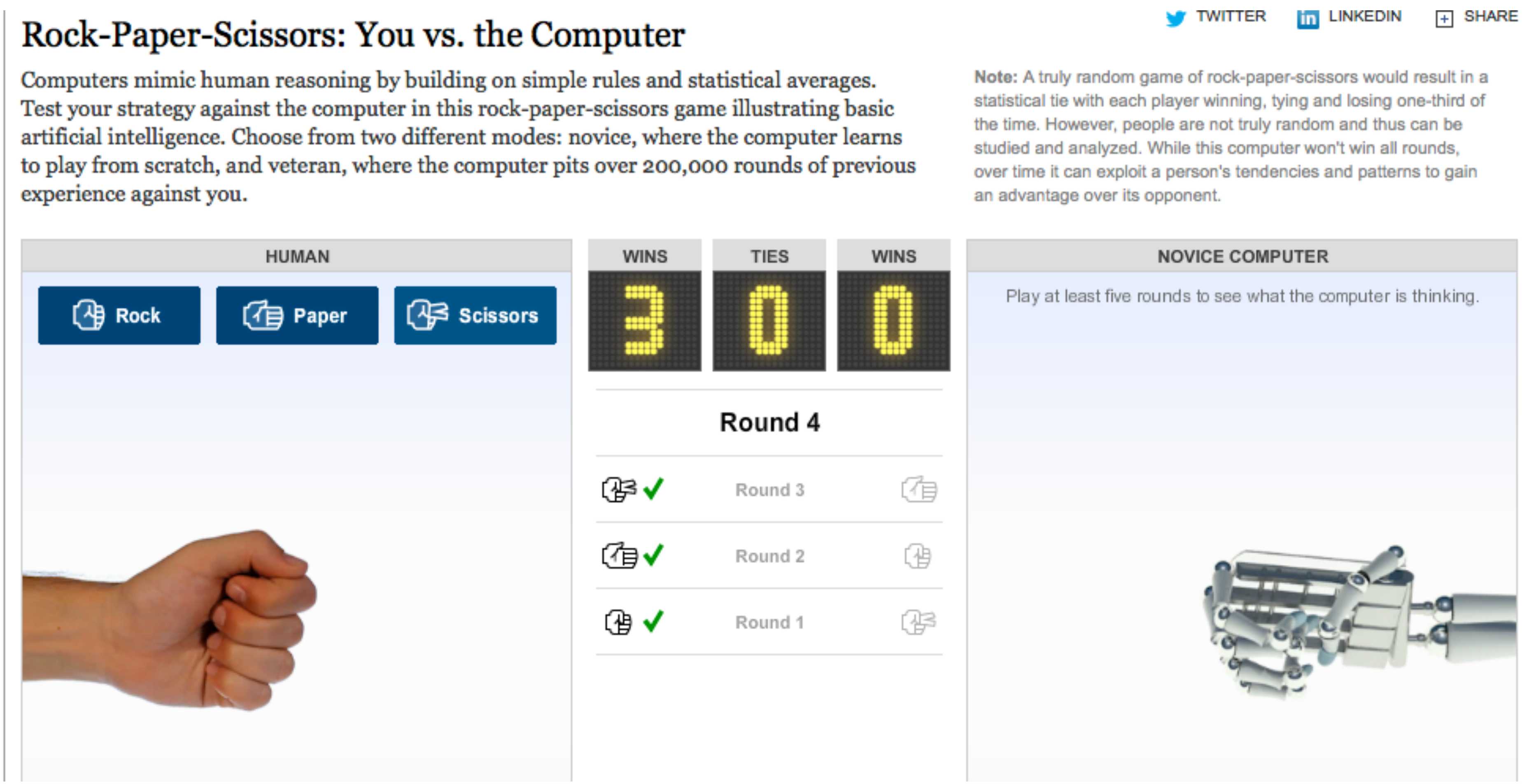
Normal Form Games: Example

- Rock-Paper-Scissors.

	R	P	S
R	0,0	-1,1	1,-1
P	1,-1	0,0	-1,1
S	-1,1	1,-1	0,0

Be Truly Random

- http://goo.gl/3sVFzN



Normal Form Games

- p players.
- For each player $k \in [1, p]$:
  - set of actions (or pure strategies) $\mathcal{A}_k$.
  - payoff function $u_k \colon \prod_{k=1}^{p} \mathcal{A}_k \to \mathbb{R}$.
- Goal of each player: maximize his payoff in a repeated game.

Prisoner's Dilemma

- Silence/Betrayal.
  - for each player, the best action is B, regardless of the other player's action.
  - but, with (B, B), both are worse off than (S, S).

	S	B
S	2,2	0,3
B	3,0	1,1

Matching Pennies

- Player A wins when pennies match, player B otherwise.
  - other versions: penalty kick.
  - no pure strategy Nash equilibrium.

	H	T
H	1,-1	-1,1
T	-1,1	1,-1

Battle of The Sexes

- Opera/Football.
  - two pure strategy Nash equilibria.

	O	F
O	3,2	0,0
F	0,0	2,3

Mixed Strategies

- Strategies:
  - pure strategies: elements of $\prod_{k=1}^{p} \mathcal{A}_k$.
  - mixed strategies: elements of $\prod_{k=1}^{p} \Delta_1(\mathcal{A}_k)$.
- Payoff: for each player $k \in [1, p]$, when players play mixed strategies $(\mathsf{p}_1, \ldots, \mathsf{p}_p)$,

$$\mathop{\mathrm{E}}_{a_j \sim \mathsf{p}_j}[u_k(\mathbf{a})] = \sum_{\mathbf{a}=(a_1,\ldots,a_p)} \mathsf{p}_1(a_1) \cdots \mathsf{p}_p(a_p) u_k(\mathbf{a}).$$

Nash Equilibrium

- Definition: a mixed strategy $(\mathsf{p}_1, \ldots, \mathsf{p}_p)$ is a (mixed) Nash equilibrium if for all $k \in [1, p]$ and $\mathsf{q}_k \in \Delta_1(\mathcal{A}_k)$,

$$u_k(\mathsf{q}_k, \mathsf{p}_{-k}) \leq u_k(\mathsf{p}_k, \mathsf{p}_{-k}).$$

  - if for all k, p_k is a pure strategy, then $(\mathsf{p}_1, \ldots, \mathsf{p}_p)$ is said to be a pure Nash equilibrium.

Nash Equilibrium: Examples

- Prisoner's dilemma: (B, B) is a pure Nash equilibrium. Dominant strategy: both better off playing B regardless of the other player's action.
- Matching Pennies: no pure Nash equilibrium; clear mixed Nash equilibrium: uniform probability for both.
- Battle of The Sexes:
  - pure Nash equilibria: both (O, O) and (F, F).
  - mixed Nash equilibria: ((2/3, 1/3), (1/3, 2/3)).
  - payoff of 2/3 for both in mixed case: less than payoffs in pure cases!

Nash's Theorem

- Theorem: any normal form game with a finite set of players and finite set of actions admits a (mixed) Nash equilibrium.

Proof

- Define function $\Phi\colon \prod_{k=1}^{p} \Delta_1(\mathcal{A}_k) \to \prod_{k=1}^{p} \Delta_1(\mathcal{A}_k)$ by

$$\Phi(\mathsf{p}_1, \ldots \mathsf{p}_p) = (\mathsf{p}'_1, \ldots \mathsf{p}'_p)$$

with $\forall k \in [1, p], j \in [1, n_k],\ p_k'^j = \frac{p_k^j + c_k^{j+}}{1 + \sum_{j=1}^{n_k} c_k^{j+}},$

where $c_k^j = u_k(\mathbf{e}_j, \mathsf{p}_{-k}) - u_k(\mathsf{p}_k, \mathsf{p}_{-k}),\ c_k^{j+} = \max(0, c_k^j).$

- Φ is a continuous function mapping from a non-empty compact convex set to itself, thus, by Brouwer's fixed-point theorem, there exists $(\mathsf{p}_1, \ldots \mathsf{p}_p)$ such that

$$\Phi(\mathsf{p}_1, \ldots \mathsf{p}_p) = (\mathsf{p}_1, \ldots \mathsf{p}_p).$$

Proof

- Observe that for any $k \in [1, p]$,

$$\sum_{j=1}^{n_k} p_k^j c_k^j = \sum_{j=1}^{n_k} p_k^j u_k(\mathbf{e}_j, \mathsf{p}_{-k}) - u_k(\mathsf{p}_k, \mathsf{p}_{-k}) = 0.$$

- Thus, for any $k \in [1, p]$, there exists at least one j such that $c_k^j \leq 0$ with $p_k^j > 0$. For that j, $c_k^{j+} = 0$ and

$$\begin{aligned} p_k^j = \frac{p_k^j}{1 + \sum_{j=1}^{n_k} c_k^{j+}} &\Rightarrow 1 + \sum_{j=1}^{n_k} c_k^{j+} = 1 \\ &\Rightarrow c_k^{j+} = 0, \forall j \\ &\Rightarrow u_k(\mathbf{e}_j, \mathsf{p}_{-k}) \leq u_k(\mathsf{p}_k, \mathsf{p}_{-k}), \forall j \\ &\Rightarrow u_k(\mathsf{q}_k, \mathsf{p}_{-k}) \leq u_k(\mathsf{p}_k, \mathsf{p}_{-k}), \forall \mathsf{q}_k. \end{aligned}$$

Nash Equilibrium: Problems

- Different equilibria:
  - not clear which one will be selected.
  - different payoffs.
- Circular definition.
- Finding any Nash equilibrium is a PPAD-complete (polynomial parity argument on directed graphs) problem (Daskalakis et al., 2009).
- Not a natural model of rationality if computationally hard.

Zero-Sum Games: Order of Play

- If row player plays p then column player plays q solution of
$$\min_{\mathsf{q} \in \Delta_1(\mathcal{A}_2)} \mathop{\mathrm{E}}_{\substack{a_1 \sim \mathsf{p} \\ a_2 \sim \mathsf{q}}} [u_1(\mathbf{a})].$$
- Thus, if row player starts, he plays p to maximize that quantity and the payoff is
$$\max_{\mathsf{p} \in \Delta_1(\mathcal{A}_1)} \min_{\mathsf{q} \in \Delta_1(\mathcal{A}_2)} \mathop{\mathrm{E}}_{\substack{a_1 \sim \mathsf{p} \\ a_2 \sim \mathsf{q}}} [u_1(\mathbf{a})].$$
- Similarly, if column player plays first, the expected payoff is
$$\min_{\mathsf{q} \in \Delta_1(\mathcal{A}_2)} \max_{\mathsf{p} \in \Delta_1(\mathcal{A}_1)} \mathop{\mathrm{E}}_{\substack{a_1 \sim \mathsf{p} \\ a_2 \sim \mathsf{q}}} [u_1(\mathbf{a})].$$

von Neumann's Theorem

(von Neumann, 1928)

- Theorem (von Neumann's minimax theorem): for any two-player zero-sum game with finite action sets,

$$\max_{\mathbf{p} \in \Delta_1(\mathcal{A}_1)} \min_{\mathbf{q} \in \Delta_1(\mathcal{A}_2)} \mathop{\mathrm{E}}_{\substack{a_1 \sim \mathbf{p} \\ a_2 \sim \mathbf{q}}} [u_1(\mathbf{a})] = \min_{\mathbf{q} \in \Delta_1(\mathcal{A}_2)} \max_{\mathbf{p} \in \Delta_1(\mathcal{A}_1)} \mathop{\mathrm{E}}_{\substack{a_1 \sim \mathbf{p} \\ a_2 \sim \mathbf{q}}} [u_1(\mathbf{a})].$$

  - common value called value of the game.
  - mixed Nash equilibria coincide with maximizing and minimizing pairs and they all have the same payoff.

Proof

- Playing second is never worse:

$$\max_{\mathsf{p}\in\Delta_1(\mathcal{A}_1)} \min_{\mathsf{q}\in\Delta_1(\mathcal{A}_2)} \mathop{\mathrm{E}}_{\substack{a_1\sim\mathsf{p}\\ a_2\sim\mathsf{q}}}[u_1(\mathbf{a})] \leq \min_{\mathsf{q}\in\Delta_1(\mathcal{A}_2)} \max_{\mathsf{p}\in\Delta_1(\mathcal{A}_1)} \mathop{\mathrm{E}}_{\substack{a_1\sim\mathsf{p}\\ a_2\sim\mathsf{q}}}[u_1(\mathbf{a})].$$

  - straightforward:

$$\forall\mathsf{p}\in\Delta_1(\mathcal{A}_1), \forall\mathsf{q}\in\Delta_1(\mathcal{A}_2), \quad \min_{\mathsf{q}\in\Delta_1(\mathcal{A}_2)} \mathop{\mathrm{E}}_{\substack{a_1\sim\mathsf{p}\\ a_2\sim\mathsf{q}}}[u_1(\mathbf{a})] \leq \mathop{\mathrm{E}}_{\substack{a_1\sim\mathsf{p}\\ a_2\sim\mathsf{q}}}[u_1(\mathbf{a})]$$

$$\Rightarrow \quad \forall\mathsf{q}\in\Delta_1(\mathcal{A}_2), \quad \max_{\mathsf{p}\in\Delta_1(\mathcal{A}_1)} \min_{\mathsf{q}\in\Delta_1(\mathcal{A}_2)} \mathop{\mathrm{E}}_{\substack{a_1\sim\mathsf{p}\\ a_2\sim\mathsf{q}}}[u_1(\mathbf{a})] \leq \max_{\mathsf{p}\in\Delta_1(\mathcal{A}_1)} \mathop{\mathrm{E}}_{\substack{a_1\sim\mathsf{p}\\ a_2\sim\mathsf{q}}}[u_1(\mathbf{a})]$$

$$\Rightarrow \quad \max_{\mathsf{p}\in\Delta_1(\mathcal{A}_1)} \min_{\mathsf{q}\in\Delta_1(\mathcal{A}_2)} \mathop{\mathrm{E}}_{\substack{a_1\sim\mathsf{p}\\ a_2\sim\mathsf{q}}}[u_1(\mathbf{a})] \leq \min_{\mathsf{q}\in\Delta_1(\mathcal{A}_2)} \max_{\mathsf{p}\in\Delta_1(\mathcal{A}_1)} \mathop{\mathrm{E}}_{\substack{a_1\sim\mathsf{p}\\ a_2\sim\mathsf{q}}}[u_1(\mathbf{a})].$$

Proof

- Set-up: at reach round,
  - column player selects q_t using RWM.
  - row player selects $\mathsf{p}_t = \max_{\mathsf{p} \in \Delta_1(\mathcal{A}_1)} \mathsf{p}^\top \mathbf{U} \mathsf{q}_t$.
- Thus, letting $T \to +\infty$ in the following completes the proof:

$$
\begin{aligned}
\min_{\mathsf{q} \in \Delta_1(\mathcal{A}_2)} \max_{\mathsf{p} \in \Delta_1(\mathcal{A}_1)} \mathop{\mathrm{E}}_{\substack{a_1 \sim \mathsf{p} \\ a_2 \sim \mathsf{q}}} [u_1(\mathbf{a})] &= \min_{\mathsf{q} \in \Delta_1(\mathcal{A}_2)} \max_{\mathsf{p} \in \Delta_1(\mathcal{A}_1)} \mathsf{p}^\top \mathbf{U} \mathsf{q} \\
&\leq \max_{\mathsf{p} \in \Delta_1(\mathcal{A}_1)} \mathsf{p}^\top \mathbf{U} \left[\frac{1}{T} \sum_{t=1}^{T} \mathsf{q}_t \right] = \max_{\mathsf{p} \in \Delta_1(\mathcal{A}_1)} \frac{1}{T} \sum_{t=1}^{T} \mathsf{p}^\top \mathbf{U} \mathsf{q}_t \\
&\leq \frac{1}{T} \sum_{t=1}^{T} \max_{\mathsf{p} \in \Delta_1(\mathcal{A}_1)} \mathsf{p}^\top \mathbf{U} \mathsf{q}_t = \frac{1}{T} \sum_{t=1}^{T} \mathsf{p}_t^\top \mathbf{U} \mathsf{q}_t = \min_{\mathsf{q}} \frac{1}{T} \sum_{t=1}^{T} \mathsf{p}_t^\top \mathbf{U} \mathsf{q} + \frac{R_T}{T} \\
&= \min_{\mathsf{q}} \left[\frac{1}{T} \sum_{t=1}^{T} \mathsf{p}_t^\top \right] \mathbf{U} \mathsf{q} + \frac{R_T}{T} \leq \max_{\mathsf{p}} \min_{\mathsf{q}} \mathsf{p}^\top \mathbf{U} \mathsf{q} + \frac{R_T}{T}.
\end{aligned}
$$

Proof

- Let $p^* \in \underset{p \in \Delta_1(\mathcal{A}_1)}{\text{argmax}} \min_{q \in \Delta_1(\mathcal{A}_2)} u_1(p, q)$ and

$$q^* \in \underset{q \in \Delta_1(\mathcal{A}_2)}{\text{argmin}} \max_{p \in \Delta_1(\mathcal{A}_1)} u_1(p, q).$$

  - p^* and q^* exist by the continuity of u_1 and the compactness of the simplices.
  - By definition of p^* and q^* and the minmax theorem:

$$v = \min_q u_1(p^*, q) \leq u_1(p^*, q^*) \leq \max_p u_1(p, q^*) = v.$$

  - Thus, (p^*, q^*) is a Nash equilibrium.

Proof

- Conversely, assume that (p^*, q^*) is a Nash equilibrium. Then,

$$u_1(p^*, q^*) = \max_p u_1(p, q^*) \geq \min_q \max_p u_1(p, q) = v$$

$$u_1(p^*, q^*) = \min_q u_1(p^*, q) \leq \max_p \min_q u_1(p, q) = v.$$

  - This implies equalities and

$$u_1(p^*, q^*) = \max_p \min_q u_1(p, q) = \min_q \max_p u_1(p, q).$$

Notes

- Unique value: all Nash equilibria have the same payoff (less problematic than general case).
- Potentially several equilibria but no need to cooperate.
- Computationally efficient: convergence in $O\left(\sqrt{\frac{\log N}{T}}\right)$.
- Plausible explanation of how an equilibrium is reached — note that both players can play RWM.
- In general non-zero-sum games regret minimization does not lead to an equilibrium.

Yao's Lemma

(Yao, 1977)

- Theorem: for any two-player zero-sum game with finite action sets,

$$\max_{\mathsf{p} \in \Delta_1(\mathcal{A}_1)} \min_{a_2 \in \mathcal{A}_2} \mathop{\mathrm{E}}_{a_1 \sim \mathsf{p}} [u_1(\mathbf{a})] = \min_{\mathsf{q} \in \Delta_1(\mathcal{A}_2)} \max_{a_1 \in \mathcal{A}_1} \mathop{\mathrm{E}}_{a_2 \sim \mathsf{q}} [u_1(\mathbf{a})].$$

  - consequence: for any distribution D over the inputs, the cost of a randomized algorithm is lower bounded by the minimum D-average cost of a deterministic algorithm.
  - to determine a lower bound for the cost of a randomized algorithm, it suffices to inspect the complexity of deterministic algorithms with randomized inputs.

General Finite Games

- Regret notion not relevant: (external) regret minimization may not lead to a Nash equilibrium.
- Notion of equilibrium: several issues related to Nash equilibria.

  → new notion of equilibrium, new notion of regret.

Correlated Equilibrium - Tale

- There is an authority or a correlation mechanism device.
- The authority defines a probability distribution p over the p-tuple of the players' actions.
- The authority draws $(a_1, \ldots, a_p) \sim$ p and reveals to each player k only his action a_k.
- The authority is a correlated equilibrium if player k has no incentive to deviate from the action recommended: the utility of any other action is lower than a_k, conditioned on the fact that he was told a_k, assuming that other players follow the recommendation they received.

Correlated Equilibrium

(Aumann, 1974)

- Definition: consider a normal form game with $p < +\infty$ players and finite action sets $\mathcal{A}_k$, $k \in [1, p]$. Then, a probability distribution p over $\prod_{k=1}^{p} \mathcal{A}_k$ is a correlated equilibrium if for all $k \in [1, p]$, for all $a_k \in \mathcal{A}_k$ with positive probability and all $a'_k \in \mathcal{A}_k$,

$$\operatorname*{E}_{\mathbf{a} \sim \mathbf{p}}[u_k(a_k, a_{-k}) \mid a_k] \geq \operatorname*{E}_{\mathbf{a} \sim \mathbf{p}}[u_k(a'_k, a_{-k}) \mid a_k].$$

Notes

- Think of the joint distribution as a correlation device.
- The set of all correlated equilibria is a convex set (it is a polyhedron): defined by a system of linear inequalities, including the simplex constraints. Solution via solving an LP problem.
- The set of Nash equilibria in general is not convex. It is defined by the intersection of the polyhedron of correlated equilibria and the constraints

$$\mathsf{p}(\mathbf{a}) = \mathsf{p}_1(a_1) \times \cdots \times \mathsf{p}_p(a_p).$$

Traffic Lights

- Stop/Go.

	S	G
S	4,4	1,5
G	5,1	0,0

  - Pure Nash equilibria: (S, G), (G, S). Mixed Nash equilibrium: ((1/2, 1/2), (1/2, 1/2)).
  - Correlated equilibria:

0	1/2
1/2	0

1/3	1/3
1/3	0

Internal Regret

- Definition: internal regret, $C_{a,b}$ functions $f\colon \mathcal{A} \to \mathcal{A}$ leaving all actions unchanged but a which is switched to b.

$$R_T = \sum_{t=1}^{T} \mathop{\mathrm{E}}_{a_t \sim \mathsf{p}_t} [l(a_t)] - \min_{f \in C_{a,b}} \sum_{t=1}^{T} \mathop{\mathrm{E}}_{a_t \sim \mathsf{p}_t} [l(f(a_t))].$$

- Definition: swap regret, C family of all functions $f\colon \mathcal{A} \to \mathcal{A}$.

$$R_T = \sum_{t=1}^{T} \mathop{\mathrm{E}}_{a_t \sim \mathsf{p}_t} [l(a_t)] - \min_{f \in C} \sum_{t=1}^{T} \mathop{\mathrm{E}}_{a_t \sim \mathsf{p}_t} [l(f(a_t))].$$

Swap Regret and Correlated Eq.

- Theorem: consider a finite normal form game played repeatedly. Assume that each player follows a swap regret minimizing strategy. Then, the empirical distribution of all plays converges to a correlated equilibrium.

Alternative Proof

- Let C be the convex set of correlated equilibria. If the sequence of empirical dist. $(\widehat{p}_t)_{t \in \mathbb{N}}$ does not converge to C, it admits a subsequence in $C_\eta = \{\mathsf{p}\colon d(\mathsf{p}, C) \geq \eta\}$ (compact set), thus, it admits a subsequence converging to $\widehat{p} \notin C$.

- Thus, there exist $\epsilon > 0$, $k \in [1, p]$, and $a_k, a'_k \in \mathcal{A}_k$ such that

$$\sum_{a_{-k} \in \mathcal{A}_{-k}} \widehat{p}(\mathbf{a}) [u_k(a'_k, a_{-k}) - u_k(a_k, a_{-k})] = \epsilon.$$

- Therefore, for t sufficiently large,

$$\sum_{a_{-k} \in \mathcal{A}_{-k}} \widehat{p}_{\tau(t)}(\mathbf{a}) [u_k(a'_k, a_{-k}) - u_k(a_k, a_{-k})] \geq \frac{\epsilon}{2}.$$

Alternative Proof

- Since $\widehat{p}_{\tau(t)}(\mathbf{a}) = \frac{1}{\tau(t)} \sum_{s=1}^{\tau(t)} 1_{\mathbf{a}_{-k,\tau(s)}=\mathbf{a}_{-k}} 1_{a_{k,\tau(s)}=a_k}$,

$$\frac{1}{\tau(t)} \sum_{s=1}^{\tau(t)} \Big[u_k(a'_k, \mathbf{a}_{-k,\tau(s)}) - u_k(a_k, \mathbf{a}_{-k,\tau(s)}) \Big] 1_{a_{k,\tau(s)}=a_k} \geq \frac{\epsilon}{2}.$$

- Thus, the internal regret of player k for switching a_k to a'_k is lower bounded by $\frac{\epsilon}{2}$ at time $\tau(t)$ and later, which implies that the player is not following a swap regret minimization strategy.

Proof

- Define the instantanous regret of player k at time t as

$$\widehat{r}_{k,t,j,j'} = 1_{a_{k,t}=j}[l_k(j, a_{-k,t}) - l_k(j', a_{-k,t})],$$

and $$r_{k,t,j,j'} = p_{k,t,j}[l_k(j, a_{-k,t}) - l_k(j', a_{-k,t})].$$

- Then, $\mathrm{E}[\widehat{r}_{k,t,j,j'} | \text{past} \wedge \text{other players' actions}] = r_{k,t,j,j'}$.
- Thus, for any (j, j'), $(r_{k,t,j,j'} - \widehat{r}_{k,t,j,j'})$ is a bounded martingale difference. By Azuma's inequality and the Borell-Cantelli lemma, for all k and (j, j'),

$$\limsup_{T \to +\infty} \frac{1}{T} \sum_{t=1}^{T} \widehat{r}_{k,t,j,j'} - r_{k,t,j,j'} = 0 \ \text{(a.s.)}.$$

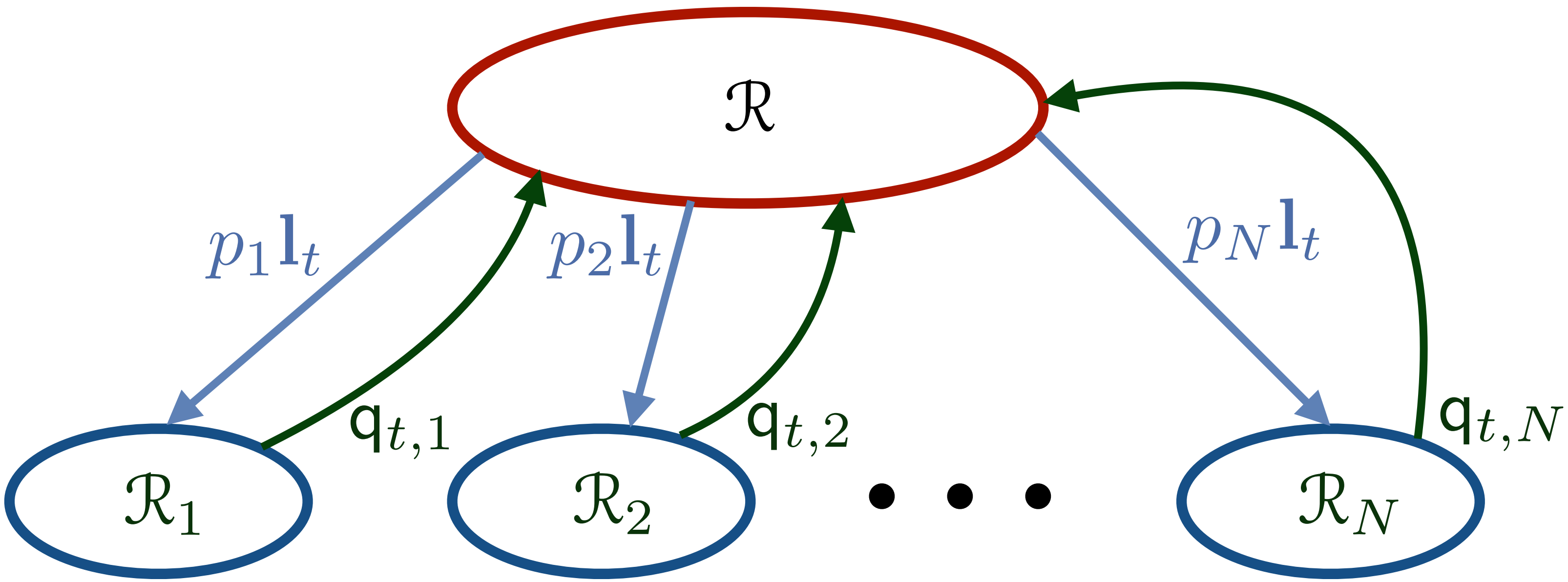
Therefore,

$$\forall k, \limsup_{T \to +\infty} \max_{j,j'} \frac{1}{T} \sum_{t=1}^{T} \widehat{r}_{k,t,j,j'} \leq 0 \ \text{(a.s.)}.$$

Swap Regret Algorithm

(Blum and Mansour, 2007)

- Theorem: there exists an algorithm with $O(\sqrt{NT \log N})$ swap regret.



$\mathcal{R}_i$s external regret minimization algorithms

Proof

- Define for all $t \in [1, T]$ the stochastic matrix

$$\mathbf{Q}_t = (q_{t,i,j})_{(i,j) \in [1,N]^2} = \begin{bmatrix} \mathsf{q}_{t,1}^\top \\ \vdots \\ \mathsf{q}_{t,N}^\top \end{bmatrix}.$$

- Since $\mathbf{Q}_t$ is stochastic, it admits a stationary distribution p_t :

$$\mathsf{p}_t^\top = \mathsf{p}_t^\top \mathbf{Q}_t \Leftrightarrow \forall j \in [1, N], p_{t,j} = \sum_{i=1}^{N} p_{t,i}\, q_{t,i,j}$$

- Thus,

$$\sum_{t=1}^{T} \mathsf{p}_t \cdot \mathbf{l}_t = \sum_{t=1}^{T} \sum_{j=1}^{N} p_{t,j} l_{t,j} = \sum_{t=1}^{T} \sum_{j=1}^{N} \sum_{i=1}^{N} p_{t,i} q_{t,i,j} l_{t,j} = \sum_{i=1}^{N} \sum_{t=1}^{T} \mathsf{q}_{t,i} \cdot (p_{t,i} \mathbf{l}_t) \leq \sum_{i=1}^{N} \min_j \sum_{t=1}^{T} p_{t,i} l_{t,j} + R_{T,i}.$$

Proof

- Thus, for any $f\colon \mathcal{A} \to \mathcal{A}$,

$$\sum_{t=1}^{T} \mathsf{p}_t \cdot \mathbf{l}_t \leq \sum_{i=1}^{N} \sum_{t=1}^{T} p_{t,i} l_{t,f(i)} + R_{T,i}.$$

- For RWM, $R_{T,i} = O(\sqrt{L_{\min,i} \log N})$. Thus, by Jensen's inequality,

$$\sum_{i=1}^{N} R_{T,i} = N \frac{1}{N} \sum_{i=1}^{N} R_{T,i}$$

$$\leq O\left(N \sqrt{\frac{1}{N} \sum_{i=1}^{N} L_{\min,i} \log N}\right) \qquad \text{(Jensen's ineq.)}$$

$$\leq O\left(N \sqrt{\frac{1}{N} T \log N}\right) = O\left(\sqrt{NT \log N}\right). \qquad \left(\sum_{i=1}^{N} L_{\min,i} = \sum_{t=1}^{T} \sum_{i=1}^{N} p_{t,i} l_{t,j_i^*}\right)$$

Notes

- Surprising result:
  - no explicit joint distribution in the game!
  - correlation induced by the empirical sequence of plays by the players.
- Game matrix:
  - no need to know the full matrix (which could be huge with a lot of players).
  - only need to know the loss or payoff for actions taken.

Coarse Correlated Equilibrium

- Definition: consider a normal form game with $p < +\infty$ players and finite action sets $\mathcal{A}_k$, $k \in [1, p]$. Then, a probability distribution p over $\prod_{k=1}^{p} \mathcal{A}_k$ is a coarse correlated equilibrium if for all $k \in [1, p]$, for all $a_k \in \mathcal{A}_k$ and all $a'_k \in \mathcal{A}_k$,

$$\mathop{\mathrm{E}}_{\mathbf{a} \sim \mathbf{p}}[u_k(a_k, a_{-k})] \geq \mathop{\mathrm{E}}_{\mathbf{a} \sim \mathbf{p}}[u_k(a'_k, a_{-k})].$$

Notes

- Any correlated equilibrium is a coarse correlated equilibrium. Difference: realization a_k not known to player.
- Comparison with mixed Nash equilibria: (general) joint distribution vs. product distributions.
- Relationship with external regret, and external regret minimizers.

Conclusion

- Zero-sum finite games:
  - external regret minimization algorithms (e.g., RWM).
  - Nash equilibrium, value of the game reached.
- General finite games:
  - internal/swap regret minimization algorithms.
  - correlated equilibrium, can be learned.
- Questions:
  - Nash equilibria.
  - extensions: e.g., time selection functions (Blum and Mansour, 2007), conditional correlated equilibrium (MM and Yang, 2014).

References

- Robert Aumann. Subjectivity and correlation in randomized strategies. Journal of Mathematical Economics, 1:67–96, 1974.
- Avrim Blum and Yishay Mansour. From external to internal regret. Journal of Machine Learning Research, 8:1307–1324, 2007.
- Nicolò Cesa-Bianchi and Gábor Lugosi. *Prediction, learning, and games*. Cambridge University Press, 2006.
- Constantinos Daskalakis, Paul W. Goldberg, Christos H. Papadimitriou. The complexity of computing a Nash equilibrium. Commun. ACM 52(2): 89-97 (2009).
- Dean P. Foster and Rakesh V. Vohra. Calibrated learning and correlated equilibrium. Games and Economic Behavior, 21(12):40 – 55, 1997.
- Yoav Freund and Robert Schapire. Large margin classification using the perceptron algorithm. In *Proceedings of COLT 1998*. ACM Press, 1998.

References

- Ehud Lehrer. Correlated equilibria in two-player repeated games with nonobservable actions. Mathematics of Operations Research, 17(1), 1992.
- Nick Littlestone. From On-Line to Batch Learning. COLT 1989: 269-284.
- Nick Littlestone, Manfred K. Warmuth: The Weighted Majority Algorithm. *FOCS 1989*: 256-261.
- Mehryar Mohri and Scott Yang. Conditional swap regret and conditional correlated equilibrium. In NIPS. 2014.
- Noam Nisan, Tim Roughgarden, Éva Tardos, and Vijay V. Vazirani. Algorithmic Game Theory. Cambridge University Press, New York, NY, USA, 2007.
- John von Neumann. Zur Theorie der Gesellschaftsspiele. Mathematische Annalen, 100(1):295–320, 1928.

References

- Gilles Stoltz and Gábor Lugosi. Learning correlated equilibria in games with compact sets of strategies. Games and Economic Behavior, 59(1), 2007.
- Andrew Yao. Probabilistic computations: Toward a unified measure of complexity, in FOCS, pp. 222–227, 1977.