Prediction, Learning and Games - Chapter 7
Prediction and Playing Games

Walid Krichene

November 4, 2013

The setting

- K players
- player k: N_k actions

The setting

- K players
- player k: N_k actions
- joint action vector $i \in [N_1] \times \cdots \times [N_K]$
- loss vector $\ell(i) \in [0, 1]$

The setting

- K players
- player k: N_k actions
- joint action vector $i \in [N_1] \times \cdots \times [N_K]$
- loss vector $\ell(i) \in [0, 1]$
- players randomize: mixed strategy profile $\pi \in \Delta^{[N_1]} \times \cdots \times \Delta^{[N_K]}$

$$\pi(i) = p^{(1)}(i_1) \dots p^{(K)}(i_K)$$

- joint random action $I \in [N_1] \times [N_K]$, $I \sim \pi$
- expected loss $\mathrm{E}[\ell^{(k)}] = \sum_i \pi(i)\ell^{(k)}(i)$
- by abuse of notation $\ell^{(k)}(\pi)$ (can think of it as a linear extension of ℓ from $[N_1] \times \cdots \times [N_K]$ to $\Delta^{[N_1]} \times \cdots \times \Delta^{[N_K]}$)

Nash equilibrium

Nash equilibrium

p is a Nash eq. if for all k and all p'

$$\ell^{(k)}(p^{(k)}, p^{(-k)}) \leq \ell^{(k)}(p', p^{(-k)})$$

- no one has an incentive to unilaterally deviate.
- set of Nash equilibria: $\mathcal{N}$

Correlated equilibria

Correlated equilibria (Aumann)

- Probability distribution P over $[N_1] \times \cdots \times [N_K]$, which is not necessarily a product distribution

Correlated equilibria

Correlated equilibria (Aumann)

- Probability distribution P over $[N_1] \times \cdots \times [N_K]$, which is not necessarily a product distribution
- Some coordinator gives advice $I \sim P$
- If player unilaterally deviates, worse-off in expectation.

Correlated equilibria

Correlated equilibria (Aumann)

- Probability distribution P over $[N_1] \times \cdots \times [N_K]$, which is not necessarily a product distribution
- Some coordinator gives advice $I \sim P$
- If player unilaterally deviates, worse-off in expectation.

Equivalent to "switching j and j' is a bad idea". For all j, j'

$$\sum_{i : i^{\mathbf{k}} = j} P(i)[\ell^{(k)}(j, i^{(-k)}) - \ell^{(k)}(j', i^{(-k)})] \leq 0$$

Link to internal regret.

Correlated equilibria

Correlated equilibria (Aumann)

- Probability distribution P over $[N_1] \times \cdots \times [N_K]$, which is not necessarily a product distribution
- Some coordinator gives advice $I \sim P$
- If player unilaterally deviates, worse-off in expectation.

Equivalent to "switching j and j' is a bad idea". For all j, j'

$$\sum_{i : i^{\mathbf{k}} = j} P(i) [\ell^{(k)}(j, i^{(-k)}) - \ell^{(k)}(j', i^{(-k)})] \leq 0$$

Link to internal regret.

Structure of correlated equilibria

- Think of P as a vector in $\Delta^{[N_1] \times \cdots \times [N_K]}$
- Linear inequalities in $P \Rightarrow$ closed convex polyhedron
- Convex combination of Nash eq $\Rightarrow$ correlated eq. (but not $\Leftarrow$)

Repeated games

- k maintains $p_t^{(k)}$, draws action $I_t^k \sim p_t^{(k)}$
- observes all players' actions $I_t = (I_t^{(1)}, \ldots, I_t^{(K)})$

Repeated games

- k maintains $p_t^{(k)}$, draws action $I_t^k \sim p_t^{(k)}$
- observes all players' actions $I_t = (I_t^{(1)}, \dots, I_t^{(K)})$
- Uncoupled play: player only knows his loss function.

Repeated games

- k maintains $p_t^{(k)}$, draws action $I_t^k \sim p_t^{(k)}$
- observes all players' actions $I_t = (I_t^{(1)}, \dots, I_t^{(K)})$
- Uncoupled play: player only knows his loss function. In fact, only his regret.

Repeated games

- k maintains $p_t^{(k)}$, draws action $I_t^k \sim p_t^{(k)}$
- observes all players' actions $I_t = (I_t^{(1)}, \dots, I_t^{(K)})$
- Uncoupled play: player only knows his loss function. In fact, only his regret.

Regret-minimizing procedure

Does the repeated play converge to $\mathcal{N}$ (or other equilibrium concept) in some sense?

Summary of results

- Minimizing regret:
  - average marginal distributions $\bar{p}^{(1)} \times \dots \times \bar{p}^{(K)}$ converge to $\mathcal{N}$
  - average joint distribution $\overline{p^{(1)} \times p^{(K)}}$ converges to the Hannan set
- Minimizing internal regret:
  - average joint distribution $\overline{p^{(1)} \times p^{(K)}}$ converges to correlated equilibria

Regret-minimization Vs. fictitious play

On iteration t, play best response to cumulative losses

$$\ell_t^{(k)} : \text{ vector of losses } \ell_t^{(k)}(i^{(k)}) = \ell^{(k)}(i^{(k)}, I_t^{(-k)})$$

$$L_t^{(k)} = \sum_{\tau=1}^{t} \ell_\tau^{(k)}$$

Regret-minimization Vs. fictitious play

On iteration t, play best response to cumulative losses

$$\ell_t^{(k)} : \text{ vector of losses } \ell_t^{(k)}(i^{(k)}) = \ell^{(k)}(i^{(k)}, I_t^{(-k)})$$

$$L_t^{(k)} = \sum_{\tau=1}^{t} \ell_\tau^{(k)}$$

Fictitious play

$$i_{t+1}^{(k)} \in \arg\min_{i \in [N_k]} L_t^{(k)}(i)$$

equivalent to

$$\min_{p \in \Delta^{[N_k]}} \left\langle p, L_t^{(k)} \right\rangle$$

Regret-minimization Vs. fictitious play

On iteration t, play best response to cumulative losses

$$\ell_t^{(k)} : \text{ vector of losses } \ell_t^{(k)}(i^{(k)}) = \ell^{(k)}(i^{(k)}, I_t^{(-k)})$$

$$L_t^{(k)} = \sum_{\tau=1}^{t} \ell_\tau^{(k)}$$

Fictitious play

$$i_{t+1}^{(k)} \in \arg\min_{i \in [N_k]} L_t^{(k)}(i)$$

equivalent to

$$\min_{p \in \Delta^{[N_k]}} \left\langle p, L_t^{(k)} \right\rangle$$

- Only converges in very special cases (e.g. 2 players, 2 actions each)
- Not Hannan-consistent (does not minimize average expected regret)

Regret-minimization Vs. fictitious play

Fictitious play

$$p_{t+1}^{(k)} \in \arg\min_{p \in \Delta^{[N_k]}} \left\langle p, L_t^{(k)} \right\rangle$$

Regret-minimization

Some algorithms can be written as

$$p_{t+1}^{(k)} \in \arg\min_{p \in \Delta^{[N_k]}} \left\langle p, L_t^{(k)} \right\rangle + \frac{1}{\gamma} R(p)$$

A Minimax theorem

Minimax Theorem via Regret minimization

- $f(x, y)$
- $\mathcal{X}$ convex compact, $f(\cdot, y)$ is convex continuous
- $\mathcal{Y}$ convex compact, $f(x, \cdot)$ is concave continuous.

A Minimax theorem

Minimax Theorem via Regret minimization

- $f(x, y)$
- $\mathcal{X}$ convex compact, $f(\cdot, y)$ is convex continuous
- $\mathcal{Y}$ convex compact, $f(x, \cdot)$ is concave continuous.

Then

$$\inf_x \sup_y f(x, y) = \sup_y \inf_x f(x, y)$$

A Minimax theorem

Minimax Theorem via Regret minimization

- $f(x, y)$
- $\mathcal{X}$ convex compact, $f(\cdot, y)$ is convex continuous
- $\mathcal{Y}$ convex compact, $f(x, \cdot)$ is concave continuous.

Then

$$\inf_x \sup_y f(x, y) = \sup_y \inf_x f(x, y)$$

$\geq$: OK. When x plays first, worse than playing second.

A Minimax theorem

proof of $\leq$:
Idea: sequential play (x_t), (y_t), then

$$\inf_x \sup_y f(x,y) \leq \frac{1}{t} \sum_{\tau \leq t} f(x_\tau, y_\tau) \leq \sup_y \inf_x f(x,y) + O(\frac{1}{\sqrt{t}})$$

A Minimax theorem

Let $\epsilon > 0$.
Find a finite cover of $\mathcal{X}$ by balls $B(a^{(i)}, \epsilon)$. Use $a^{(i)}$ as the actions.

- x plays first
- Applies regret-minimizing procedure
- action set $\{a^{(i)}\}_i$
- losses $f(a^{(i)}, y_t)$

A Minimax theorem

Let $\epsilon > 0$.
Find a finite cover of $\mathcal{X}$ by balls $B(a^{(i)}, \epsilon)$. Use $a^{(i)}$ as the actions.

- x plays first
- Applies regret-minimizing procedure
- action set $\{a^{(i)}\}_i$
- losses $f(a^{(i)}, y_t)$

- y plays second
- plays best response
- $y_t \in \arg\max_{y \in \mathcal{Y}} f(x_t, y)$

A Minimax theorem

Let $\epsilon > 0$.
Find a finite cover of $\mathcal{X}$ by balls $B(a^{(i)}, \epsilon)$. Use $a^{(i)}$ as the actions.

- x plays first
- Applies regret-minimizing procedure
- action set $\{a^{(i)}\}_i$
- losses $f(a^{(i)}, y_t)$

- y plays second
- plays best response
- $y_t \in \arg\max_{y \in \mathcal{Y}} f(x_t, y)$

$$\bar{x}_t = \frac{1}{t} \sum_{\tau=1}^{t} x_\tau \qquad \bar{y}_t = \frac{1}{t} \sum_{\tau=1}^{t} y_\tau$$

A Minimax theorem

$$
\begin{aligned}
\inf_x \sup_y f(x, y) &\le \sup_y f(\bar{x}_t, y) \\
&\le \frac{1}{t} \sum_{\tau \le t} f(x_\tau, y) && \text{by convexity} \\
&\le \frac{1}{t} \sum_{\tau \le t} f(x_\tau, y_\tau) && y \text{ plays optimally} \\
&\le \min_i \frac{1}{t} \sum_{\tau \le t} f(a^{(i)}, y_\tau) + \sqrt{\frac{\ln N}{2t}} && \text{by the regret bound} \\
&\le \min_i f(a^{(i)}, \bar{y}_t) \sqrt{\frac{\ln N}{2t}} && \text{by concavity} \\
&\le \sup_y \min_i f(a^{(i)}, y) + \sqrt{\frac{\ln N}{2t}}
\end{aligned}
$$

Then $t \to \infty$ and $\epsilon \to 0$

Zero-sum two player game

$$\ell^{(row)}(i,j) = -\ell^{(column)}(i,j)$$

- Matrix of losses: M such that $M_{i,j} = \ell^{(row)}(i,j)$
- Joint strategy: (p, q)
- Loss of row player: $\sum_{i,j} p_i M_{i,j} q_j = \langle p, Mq \rangle$

Zero-sum two player game

$$\ell^{(row)}(i,j) = -\ell^{(column)}(i,j)$$

- Matrix of losses: M such that $M_{i,j} = \ell^{(row)}(i,j)$
- Joint strategy: (p, q)
- Loss of row player: $\sum_{i,j} p_i M_{i,j} q_j = \langle p, Mq \rangle$

Nash equilibrium

(p, q) is a Nash eq. iff for all p', q'

$$\langle p, Mq' \rangle \leq \langle p, Mq \rangle \leq \langle p', Mq \rangle$$

Zero-sum two player game

$$\ell^{(row)}(i,j) = -\ell^{(column)}(i,j)$$

- Matrix of losses: M such that $M_{i,j} = \ell^{(row)}(i,j)$
- Joint strategy: (p,q)
- Loss of row player: $\sum_{i,j} p_i M_{i,j} q_j = \langle p, Mq \rangle$

Nash equilibrium

(p,q) is a Nash eq. iff for all p', q'

$$\langle p, Mq' \rangle \leq \langle p, Mq \rangle \leq \langle p', Mq \rangle$$

In fact, by Von Neumann's minimax theorem,

$$\min_{p'} \max_{q'} \langle p', Mq' \rangle = \max_{q'} \min_{p'} \langle p', Mq' \rangle$$

Call it V, the value of the game.

Value of the game

$$(p,q) \text{ is a Nash eq. } \Leftrightarrow \langle p, Mq \rangle = V$$

Repeated zero-sum two-player game

- Row player plays first
- Given a sequence of column plays, $j_1, j_2, \ldots$, score performance of row sequence $i_1, i_2, \ldots$ by evaluating regret

$$\max_i \sum_{t=1}^{T} \ell^{(row)}(i_t, j_j) - \sum_{t=1}^{T} \ell^{(row)}(i, j_t)$$

Repeated zero-sum two-player game

- Row player plays first
- Given a sequence of column plays, $j_1, j_2, \ldots$, score performance of row sequence $i_1, i_2, \ldots$ by evaluating regret

$$\max_i \sum_{t=1}^{T} \ell^{(row)}(i_t, j_j) - \sum_{t=1}^{T} \ell^{(row)}(i, j_t)$$

- Can be extended to time varying: $\ell_t^{(row)}$

Repeated zero-sum two-player game

- Row player plays first
- Given a sequence of column plays, $j_1, j_2, \dots,$ score performance of row sequence $i_1, i_2, \dots$ by evaluating regret

$$\max_i \sum_{t=1}^{T} \ell^{(row)}(i_t, j_j) - \sum_{t=1}^{T} \ell^{(row)}(i, j_t)$$

- Can be extended to time varying: $\ell_t^{(row)}$

Note

Implicit assumption: oblivious opponent.
This is not the loss Row would have suffered had he played a constant strategy.

Repeated zero-sum two-player game

Convergence of average losses

If row strategies are Hannan-consistent, then

$$\limsup_{T} \frac{1}{T} \sum_{t=1}^{T} \ell^{(row)}(i_t, j_t) \leq V$$

Repeated zero-sum two-player game

Convergence of average losses

If row strategies are Hannan-consistent, then

$$\limsup_{T} \frac{1}{T} \sum_{t=1}^{T} \ell^{(row)}(i_t, j_t) \leq V$$

- Note: only upper bound: can do much better if column player is not adversarial

Repeated zero-sum two-player game

Convergence of average losses

If row strategies are Hannan-consistent, then

$$\limsup_{T} \frac{1}{T} \sum_{t=1}^{T} \ell^{(row)}(i_t, j_t) \leq V$$

- Note: only upper bound: can do much better if column player is not adversarial
- proof: suffices to show

$$\limsup_{T} \min_{i} \frac{1}{T} \sum_{t=1}^{T} \ell^{(row)}(i, j_t) \leq V$$

(then use regret bound)

Repeated zero-sum two-player game

proof:

$$\min_i \frac{1}{T}\sum_{t=1}^{T} \ell^{(row)}(i, j_t) = \min_p \frac{1}{T}\sum_{t=1}^{T} \langle p, Me_{j_t}\rangle$$
$$= \min_p \left\langle p, M\frac{1}{T}\sum_{t=1}^{T} e_{j_t} \right\rangle$$
$$\leq \min_p \max_q \langle p, Mq\rangle$$
$$= V$$

Repeated zero-sum two-player game

If both players minimize regret, average distributions converge to $\mathcal{N}$

- $\bar{p}_t = \frac{1}{t} \sum_{\tau \leq t} p_\tau$
- $\bar{q}_t = \frac{1}{t} \sum_{\tau \leq t} q_\tau$

$$\bar{p}_t \times \bar{q}_t \to \mathcal{N}$$

Repeated zero-sum two-player game

If both players minimize regret, average distributions converge to $\mathcal{N}$

- $\bar{p}_t = \frac{1}{t} \sum_{\tau \leq t} p_\tau$
- $\bar{q}_t = \frac{1}{t} \sum_{\tau \leq t} q_\tau$

$$\bar{p}_t \times \bar{q}_t \to \mathcal{N}$$

- Does not prove convergence of actual probability sequences (p_t, q_t)
- Typical example: sequences (p_t, q_t) oscillate, but averages $(\bar{p}_t, \bar{q}_t)$ converge to $\mathcal{N}$

Repeated zero-sum two-player game

If both players minimize regret, average distributions converge to $\mathcal{N}$

- $\bar{p}_t = \frac{1}{t} \sum_{\tau \le t} p_\tau$
- $\bar{q}_t = \frac{1}{t} \sum_{\tau \le t} q_\tau$

$$\bar{p}_t \times \bar{q}_t \to \mathcal{N}$$

- Does not prove convergence of actual probability sequences (p_t, q_t)
- Typical example: sequences (p_t, q_t) oscillate, but averages $(\bar{p}_t, \bar{q}_t)$ converge to $\mathcal{N}$

- Also does not prove convergence of average joint distribution, $\overline{p \times q}$

$$\overline{p \times q}_t(i,j) = \frac{1}{T} \sum_{\tau \le t} p_t(i) q_t(j)$$

Note: average marginals Vs. average joint

They are not the same

$$\bar{p}_t \times \bar{q}_t(i,j) = \frac{1}{T^2} \sum_{t=1}^{T} \sum_{t'=1}^{T} p_t(i) q_{t'}(j)$$

$$\overline{p \times q}_t(i,j) = \frac{1}{T} \sum_{t=1}^{T} p_t(i) q_t(j)$$

Correlated equilibria and internal regret

Back to general K player game.
Let $\bar{P}$ be the average joint distribution

$$\bar{P}_t(i) = \frac{1}{t} \sum_{\tau \leq t} p^{(1)}(i_1) \dots p^{(K)}(i_K)$$

Correlated equilibria and internal regret

Back to general K player game.
Let $\bar{P}$ be the average joint distribution

$$\bar{P}_t(i) = \frac{1}{t} \sum_{\tau \leq t} p^{(1)}(i_1) \dots p^{(K)}(i_K)$$

If every player minimizes regret, $\bar{P}_t$ converges to the Hannan set

$$\mathcal{H} = \{P : \forall k,\ \forall j,\ \sum_i P(i)\ell^{(k)}(i^{(k)}, i^{(-k)}) \leq \sum_i P(i)\ell^{(k)}(j, i^{(-k)})\}$$

Correlated equilibria and internal regret

Back to general K player game.
Let $\bar{P}$ be the average joint distribution

$$\bar{P}_t(i) = \frac{1}{t} \sum_{\tau \leq t} p^{(1)}(i_1) \dots p^{(K)}(i_K)$$

If every player minimizes regret, $\bar{P}_t$ converges to the Hannan set

$$\mathcal{H} = \{P : \forall k, \ \forall j, \ \sum_i P(i) \ell^{(k)}(i^{(k)}, i^{(-k)}) \leq \sum_i P(i) \ell^{(k)}(j, i^{(-k)})\}$$

$\mathcal{H}$ contains correlated equilibria, but proper superset in most cases.

$$\mathcal{N} \subset \mathcal{C} \subset \mathcal{H}$$

Correlated equilibria and internal regret

Convergence of average joint to $\mathcal{C}$

If players minimize internal regret, then $\bar{P}_t \to \mathcal{C}$

Recall the definition of internal regret

Internal regret

$$R^{(k)}_{j,j',T} = \sum_{t=1}^{T} p^{(k)}_{j,t} (\ell(j, I^{(-k)}) - \ell(j', I^{(-k)}))$$

$$R_T = \max_{j,j'} R_{j,j',T}$$

Correlated equilibria and internal regret

Convergence of average joint to $\mathcal{C}$

If players minimize internal regret, then $\bar{P}_t \to \mathcal{C}$

Recall the definition of internal regret

Internal regret

$$R^{(k)}_{j,j',T} = \sum_{t=1}^{T} p^{(k)}_{j,t}(\ell(j, I^{(-k)}) - \ell(j', I^{(-k)}))$$

$$R_T = \max_{j,j'} R_{j,j',T}$$

Also recall that we can turn any regret-minimizing procedure over N actions, to internal-regret minimizing procedure, by taking the set of pairs (size N^2)

Summary of results

$$\mathcal{N} \subset \mathcal{C} \subset \mathcal{H}$$

- Minimizing regret:
  - $(\bar{p}_t^{(1)} \times \cdots \times \bar{p}_t^{(K)}) \to \mathcal{N}$
  - $(p_t^{(1)} \times \cdots \times p_t^{(K)}) \to \mathcal{N}$?
  - $\bar{P}_t \to \mathcal{H}$
- Minimizing internal regret:
  - $\bar{P}_t \to \mathcal{C}$, updates $O(N^2)$

Summary of results

$$\mathcal{N} \subset \mathcal{C} \subset \mathcal{H}$$

- Minimizing regret:
  - $(\bar{p}_t^{(1)} \times \cdots \times \bar{p}_t^{(K)}) \to \mathcal{N}$
  - $(p_t^{(1)} \times \cdots \times p_t^{(K)}) \to \mathcal{N}$?
  - $\bar{P}_t \to \mathcal{H}$
- Minimizing internal regret:
  - $\bar{P}_t \to \mathcal{C}$, updates $O(N^2)$
- Next
  - (Unknown game: exploration)
  - Blackwell approachability theorem
  - Potential-based approachability

Next week

- Finish Chapter 7 (approachability)
- Application to the routing game