Regret Minimization: A New Solution Concept

Work based on an article from Joseph Y. Halpern and Rafael Pass, Cornell University, 2009

Jean-Benoît Rossel

Game Theory and Algorithms

May 19th, 2011

ÉCOLE POLYTECHNIQUE
FÉDÉRALE DE LAUSANNE

Abstract

- Nash Equilibrium (NE): most commonly-used solution concept in game theory.
- Sometimes, NE gives unreasonable answers (Bertrand Duopoly, Centipede Game, Traveler's Dilemma).
- Why is it problematic?
- Presentation of a new solution concept, called Regret Minimization.
- Application of this concept to examples.

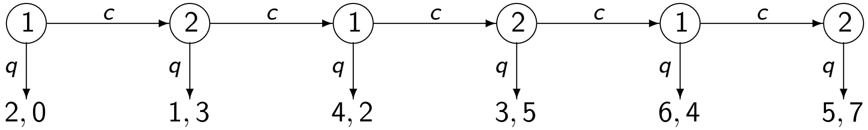
Bertrand Duopoly

- Two firms produce a same good.
- There is a demand of 100 units, and each firm can pick any price $p_1, p_2 \in \{\$0, \ldots, \$200\}$.
- Utility function for firm i ($i \in \{1, 2\}$) is

$$u_i(p_1, p_2) = \begin{cases} 100p_i & \text{if} \quad p_i < p_{3-i}, \\ 50p_i & \text{if} \quad p_i = p_{3-i}, \\ 0 & \text{if} \quad p_i > p_{3-i}. \end{cases}$$

- Only NE: $(0, 0)$ and $(1, 1)$.
- Generally, higher prices are chosen!

Centipede Game



Three classes of strategies for player 1: [1], [3] and [5] **(and also continue without stopping)**. Same for player 2 with [2], [4], [6].

[t] denotes the set of strategies where a player decides to stop at round t.

([1], [2]) is the only SPE of the game. But experimentally, people never quit the game at the first round.

Traveler's Dilemma [Kaushik Basu, 1994]

- Two travelers have identical luggage that is lost by an airline.
- The airline offers to recompense them.
- Each traveler may ask for any dollar amount between 2 and 100 (without cooperation).
- If they ask for the same amount, that is what they get.
- Otherwise, they both get the lower amount m, with a reward of p for whoever chooses m, and a penalty of p for whoever chooses the too high amount ($p \geq 2$).

  $\hookrightarrow$ Utility of player 1 is $u_1(x, y) = \min(x, y) + p \cdot \text{sgn}(y - x)$.
- Example with $p = 2$, $u_1(84, 92) = 86$, and $u_2(84, 92) = 82$.

Payoff matrix for $p = 2$

	2	3	4	$\cdots$	98	99	100
2	$(2,2)$	$(4,0)$	$(4,0)$		$(4,0)$	$(4,0)$	$(4,0)$
3	$(0,4)$	$(3,3)$	$(5,1)$	$\cdots$	$(5,1)$	$(5,1)$	$(5,1)$
4	$(0,4)$	$(1,5)$	$(4,4)$		$(6,2)$	$(6,2)$	$(6,2)$
$\vdots$		$\vdots$		$\ddots$		$\vdots$	
98	$(0,4)$	$(1,5)$	$(2,6)$		$(98,98)$	$(100,96)$	$(100,96)$
99	$(0,4)$	$(1,5)$	$(2,6)$	$\cdots$	$(96,100)$	$(99,99)$	$(101,97)$
100	$(0,4)$	$(1,5)$	$(2,6)$		$(96,100)$	$(97,101)$	$(100,100)$

Utility of player 1 is $u_1(x,y) = \min(x,y) + p \cdot \text{sgn}(y-x)$.

What would you ask for?

Nash Equilibrium for Traveler's Dilemma

- People usually are interested in high amounts... So player 1 could ask for 100, and think player 2 will do the same.
- But if he supposes that player 2 will ask for 100, player 1 would have better to ask for 99.
- By knowing this, player 2 should ask for 98.
- By knowing this, player 1 should ask for 97.
- ...
- After iterated deletion of weakly dominant strategies, $(2, 2)$ is the only NE of this game!
- But would a reasonable person choose 2? Even a game theorist?

Some problems with Nash Equilibrium

- NE requires mutual knowledge of rationality.
- NE implicitly assumes that the players know what strategy the other players are using.
- Why should a player assume that the other players will choose their part of NE?

  $\hookrightarrow$ Unreasonable, especially in one-shot game!
- What should we do if there is more than one NE?
- $\hookrightarrow$ introduction of a new solution concept, Regret Minimization.
- Goal: capture the intuition that a player wants to do well no matter what the other players do.

Regret Minimization

A few notations:

- $\mathcal{S}_i$ = set of strategies for player i,
- $\mathcal{S}_{-i} = \prod_{j \in \{1,\dots,n\} \setminus \{i\}} \mathcal{S}_j$ = set of strategies for all players except i,
- $u_i : \mathcal{S}_i \times \mathcal{S}_{-i} \to \mathbb{R}$ = payoff function for player i.

We can define now, for $a_i \in \mathcal{S}_i$,

- $regret(a_i | a_{-i}) = \max_{a \in \mathcal{S}_i} u_i(a, a_{-i}) - u_i(a_i, a_{-i}), \forall\, a_{-i} \in \mathcal{S}_{-i}$,
- $regret(a_i) = \max_{a_{-i} \in \mathcal{S}_{-i}} regret(a_i | a_{-i})$.
- *Minimax regret rule*: choose the action a that has the smallest regret.

  $\hookrightarrow$ minimize the "I wish I had chosen a' instead of a" feeling.

Application to Traveler's Dilemma

Step 1: We suppose that $p \in \{2, \ldots, 49\}$ and that players only use pure strategies $a_1, a_2 \in \{2, \ldots, 100\}$.

- If $a_1 > a_2$, player 1 gets $a_2 - p$, he could have won $a_2 - 1 + p$ (or a_2 if $a_2 = 2$).

$$\hookrightarrow regret(a_1|a_2) = \begin{cases} 2p - 1 & \text{if } a_2 \neq 2, \\ p & \text{if } a_2 = 2. \end{cases}$$

- If $a_1 = a_2$, player 1 gets a_2, he could have won $a_2 - 1 + p$ (or a_2 if $a_2 = 2$).

$$\hookrightarrow regret(a_1|a_2) = \begin{cases} p - 1 & \text{if } a_2 \neq 2, \\ 0 & \text{if } a_2 = 2. \end{cases}$$

- If $a_1 < a_2$, player 1 gets $a_1 + p$, he could have won $a_2 - 1 + p$.

$$\hookrightarrow regret(a_1|a_2) = a_2 - a_1 - 1.$$

Application to Traveler's Dilemma

Hence $regret(a_1) = \max\{2p - 1, 99 - a_1\}$, and all strategies between $100 - 2p$ and 100 minimize regret.

Step 2: Iteration of this process: we suppose that $p \in \{2, \ldots, 49\}$ and that players only use pure strategies a_1 and a_2 between $100 - 2p$ and 100. We can see that

- $regret(100 - 2p) = 2p - 1$,
- $regret(100 - 2p + 1) = 2p - 2$,
- $regret(a_1) = 2p - 1$ for all $a_1 > 100 - 2p + 1$.

Conclusion: $100 - 2p + 1$ is the strategy that *survives iterated regret minimization*.

Formal definition: Deletion Operator

- Let $\mathcal{S} = \mathcal{S}_1 \times \cdots \times \mathcal{S}_n$ be a set of pure or mixed strategies.
- Let $\mathcal{D} : \mathcal{S} \to \mathcal{S}$ such that $\mathcal{D}(\mathcal{S}) \subseteq \mathcal{S}$.

  $\hookrightarrow$ strategy profiles in $\mathcal{S} \setminus \mathcal{D}(\mathcal{S})$ are deleted.

The set of strategy profiles that *survive iterated deletion with respect to $\mathcal{D}$ and $\mathcal{S}$* is

$$\mathcal{D}^{\infty}(\mathcal{S}) = \bigcap_{k \geq 0} \mathcal{D}^k(\mathcal{S}),$$

where $\mathcal{D}^0(\mathcal{S}) = \mathcal{S}$, and $\mathcal{D}^{k+1}(\mathcal{S}) = \mathcal{D}\left(\mathcal{D}^k(\mathcal{S})\right)$.

Regret Minimization as a Deletion Operator

We define now

- $\mathcal{RM}_i(\mathcal{S})$ = set of strategies in $\mathcal{S}_i$ that minimize regret with respect to $\mathcal{S}_{-i}$.
- $\mathcal{RM}(\mathcal{S}) = \mathcal{RM}_1(\mathcal{S}) \times \cdots \times \mathcal{RM}_n(\mathcal{S})$.

In a strategy profile that survives iterated regret minimization, a player is not making a best response to the strategies used by the other players since, intuitively, he does not know what these strategies are. He chooses a strategy that ensures that he does reasonably well compared to the best he could have done, no matter what the other players do.

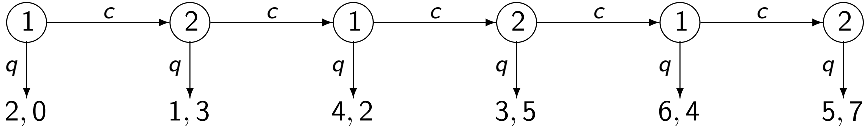
Bertrand Duopoly

Recall: Utility function for firm i ($i \in \{1,2\}$) is $u_i(p_1,p_2) = \begin{cases} 100p_i & \text{if} \quad p_i < p_{3-i}, \\ 50p_i & \text{if} \quad p_i = p_{3-i}, \\ 0 & \text{if} \quad p_i > p_{3-i}, \end{cases}$ and $p_i \in \{0,\dots,200\}$. Only NE are $(0,0)$ and $(1,1)$.

Suppose firm 1 chooses p_1 and firm 2 chooses p_2. We have $regret(0) = 199 \cdot 100$. Suppose now that $p_1 \geq 1$.

- If $p_2 > p_1$, $regret(p_1|p_2) = (p_2 - 1 - p_1)100$ (worst: $p_2 = 200$).
- If $p_1 = p_2$, $regret(p_1|p_2) = (p_2/2 - 1)100$ if $p_2 > 1$, and 0 if $p_2 = 1$.
- If $p_2 < p_1$, $regret(p_1|p_2) = (p_2 - 1)100$ (worst: $p_2 = p_1 - 1$).
  $\hookrightarrow$ $regret(p_1) = \max\left((199 - p_1)100, (p_1 - 2)100\right)$.
  $\hookrightarrow$ Firm 1 minimizes regret by choosing 100 or 101 (same for firm 2).
- Second round: 100 is the unique strategy that survives iterated regret minimization.

Centipede Game



Recall: $[t]$ denotes the set of strategies where a player decides to stop at round t.

For player 1, we have $regret([1]||[2]) = 0$, $regret([1]||[4]) = 2$, and $regret([1]||[6]) = 4$. Hence $regret([1]) = 4$.

Similarly, $regret([3]) = \max\{1, 0, 2\}$, $regret([5]) = \max\{1, 1, 0\}$. Hence, trying to continue until step 5 minimizes regret for him.
Continue without stopping gives also a regret of 1.

Computation of regret with respect to mixed strategy

Proposition. *Let $\mathcal{S}_i$ and $\mathcal{S}_{-i}$ be sets of pure strategies, and $\Delta(\mathcal{S})$ be the set of probability distributions on $\mathcal{S}$. Suppose that player i wants to use mixed strategy σ_i. Then*

$$regret(\sigma_i) = \max_{a_{-i} \in \mathcal{S}_{-i}} regret(\sigma_i | a_{-i}).$$

Proof. By definition, we have

$$regret(\sigma_i) = \max_{\sigma_{-i} \in \Delta(\mathcal{S}_{-i})} regret(\sigma_i | \sigma_{-i}).$$

Moreover,

$$\begin{aligned} regret(\sigma_i | \sigma_{-i}) &= \max_{\sigma_i' \in \Delta(\mathcal{S}_i)} U(\sigma_i', \sigma_{-i}) - U(\sigma_i, \sigma_{-i}) \\ &= U(\sigma_i^*, \sigma_{-i}) - U(\sigma_i, \sigma_{-i}) \\ &= \sum_{a_{-i} \in \mathcal{S}_{-i}} \left(\left(U(\sigma_i^*, a_{-i}) - U(\sigma_i, a_{-i}) \right) \sigma_{-i}(a_{-i}) \right). \end{aligned}$$

End of the proof

Hence

$$\begin{aligned} regret(\sigma_i|\sigma_{-i}) &\leq \sum_{a_{-i}\in\mathcal{S}_{-i}} regret(\sigma_i|a_{-i})\sigma_{-i}(a_{-i}) \\ &\leq \left(\max_{a_{-i}\in\mathcal{S}_{-i}} regret(\sigma_i|a_{-i})\right) \sum_{a_{-i}\in\mathcal{S}_{-i}} \sigma_{-i}(a_{-i}) \\ &= \max_{a_{-i}\in\mathcal{S}_{-i}} regret(\sigma_i|a_{-i}). \end{aligned}$$

To summarize, we get

$$\begin{aligned} regret(\sigma_i) &= \max_{\sigma_{-i}\in\Delta(\mathcal{S}_{-i})} regret(\sigma_i|\sigma_{-i}) \\ &\leq \max_{a_{-i}\in\mathcal{S}_{-i}} regret(\sigma_i|a_{-i}), \end{aligned}$$

so the inequality is actually an equality. This completes the proof. □

An example with mixed strategies

Consider Traveler's Dilemma with $p = 2$. We saw that strategy 97 survives iterated regret minimization with pure strategies, and that $regret(97) = 3$.

Consider now the mixed strategy σ that puts probability $1/2$ on 100, $1/4$ on 99, $1/8$ on 98, ..., and finally $1/2^{98}$ on both 3 and 2. We want to compute $regret(\sigma)$, but by last proposition, it is sufficient to compute $regret(\sigma|a)$ for each pure strategy of the other player.

- Best response to a is $a - 1$.
- Payoff with $a - 1$ is $a + 1$.
- We need to compute $U(\sigma, a)$.

Computation of *regret*(σ)

$$
\begin{aligned}
U(\sigma, a) &= (a-2)\cdot\mathbb{P}\{\sigma > a\} + a\cdot\mathbb{P}\{\sigma = a\} + \\
&\quad + \textstyle\sum_{k=2}^{a-1}(k+2)\cdot\mathbb{P}\{\sigma = k\} \\
&= (a-2)\textstyle\sum_{k=a+1}^{100}(1/2)^{101-k} + a(1/2)^{101-a} + \\
&\quad + \textstyle\sum_{k=3}^{a-1}(k+2)(1/2)^{101-k} + 4(1/2)^{98} \\
&= (a-2)(1-2^{a-100}) + a(1/2)^{101-a} + 4(1/2)^{98} + \\
&\quad a\cdot 2^{a-101} - 3\cdot 2^{-98} \\
&= a - 2 - a\cdot 2^{a-100} + 2^{a-99} + a\cdot 2^{a-100} + 2^{-98} \\
&= a - 2 + 2^{a-99} + 2^{-98}.
\end{aligned}
$$

We will have then

$$
\begin{aligned}
regret(\sigma|a) &= (a+1) - (a - 2 + 2^{a-99} + 2^{-98}) \\
&= 3 - 2^{a-99} - 2^{-98} < 3 = regret(97).
\end{aligned}
$$

Comments

- σ has non-zero probability on all actions,
- $\mathit{regret}(\sigma) < \mathit{regret}(97) = 3$.

  $\hookrightarrow$ we can do better by putting some weight even on actions that do not minimize regret!
- Optimal mixed strategy is really hard to compute (Halpern and Pass suppose that there is a unique mixed strategy which minimizes regret).
- Regret Minimization is a nice concept, seems to give plausible results, but it can be hard to deal with.

Thank you for your attention!