# Alternatives to Nash equilibrium

Lecture 6

Tomáš Kroupa

# Outline

1. Correlated equilibrium
2. Regret matching
3. Stackelberg equilibrium

# Correlated equilibrium

# Probabilistic interpretation of NE

- Assume that players follow Nash equilibrium $(p_1, \ldots, p_n)$
- Every player $i$ samples a pure strategy $s_i \in S_i$ based on $p_i$ *independently* of the other players
- This means that the probability of $\mathbf{s} = (s_1, \ldots, s_n) \in \mathbf{S}$ is

$$p(\mathbf{s}) = p_1(s_1) \cdots p_n(s_n)$$

*We may allow players to follow other random signals.*

# Correlation of pure strategies

A *correlation mechanism* is a probability distribution $p$ over $\mathbf{S}$.

**The extensive-form game** $\Gamma(p)$ proceeds as follows:

1. A strategy profile (*signal*) $\mathbf{s}$ is sampled from $p$
2. Each player $i$ learns about $s_i$ but not about $\mathbf{s}_{-i}$
3. Each player $i$ picks $s'_i \in S_i$, so the payoff is $u_i(s'_1, \ldots, s'_n)$

Strategies in $\Gamma(p)$ are maps $\sigma_i \colon S_i \to S_i$. A player $i$ adopting the signalled strategy $s_i$ is using the strategy $\sigma^*_i(s_i) = s_i$.

# Correlated equilibrium

A *correlated equilibrium* in a normal-form game is a correlation mechanism $p$ such that $(\sigma_1^*, \ldots, \sigma_n^*)$ is a Nash equilibrium in the extensive-form game $\Gamma(p)$.

- Does every game have a correlated equilibrium $p$?
- How to compute such $p$?

# Correlated equilibrium, equivalently

A correlation mechanism $p$ is a *correlated equilibrium* if, and only if, for each player $i$ and every $s_i, s_i' \in S_i$ with $s_i \neq s_i'$,

$$\sum_{\mathbf{s}_{-i} \in \mathbf{S}_{-i}} p(s_i, \mathbf{s}_{-i}) u_i(s_i', \mathbf{s}_{-i}) \leq \sum_{\mathbf{s}_{-i} \in \mathbf{S}_{-i}} p(s_i, \mathbf{s}_{-i}) u_i(s_i, \mathbf{s}_{-i}).$$

This means that the set of all CE $p$ is a convex polytope.

# Example: The game of Chicken

$$\begin{bmatrix} 6,6 & 2,7 \\ 7,2 & 0,0 \end{bmatrix}$$

The set of correlated equilibria is given by

$$\begin{aligned} 7p(1,1) &\leq 6p(1,1) + 2p(1,2) \\ 6p(2,1) + 2p(2,2) &\leq 7p(2,1) \\ 7p(1,1) &\leq 6p(1,1) + 2p(2,1) \\ 6p(1,2) + 2p(2,2) &\leq 7p(1,2) \end{aligned}$$

# Properties of correlated equilibria

- In any game, every NE $(p_1, \ldots, p_n)$ induces a CE given by

$$p(\mathbf{s}) = p_1(s_1) \cdots p_n(s_n), \qquad \mathbf{s} = (s_1, \ldots, s_n) \in \mathbf{S}$$

- A single CE can be found by solving the linear program where the objective is to maximize the *social welfare*

$$\sum_{i \in N} \sum_{\mathbf{s} \in \mathbf{S}} p(\mathbf{s}) u_i(\mathbf{s})$$

or some other criterion

# Regret matching

# Motivation - learning in games

- *Best response dynamics* converges only to pure equilibria
- *Fictitious play* is slow and may fail to converge

We seek a simple adaptive procedure for playing a game:

- Players observe the history of past plays
- Not only best response actions may be played!
- The probability of strategy is proportional to its *regret*

# Regret

Each player $i$ plays a pure strategy $s_i^t$ in iteration $t$. We define the following regrets of player $i$ in iteration $t$ for strategy $s_i$:

- *Instantenous regret* $r_i^t(s_i) = u_i(s_i, \mathbf{s}_{-i}^t) - u_i(\mathbf{s}^t)$
- *Expected regret*

$$R_i^t(s_i) = \frac{1}{t} \sum_{\tau=1}^{t} r_i^\tau(s_i)$$

- *Positive regret* $R_i^t(s_i)_+ = \max\{R_i^t(s_i), 0\}$

# Regret matching

1. Pick mixed strategies $p_1^t, \ldots, p_n^t$ arbitrarily when $t = 1$
2. For each $i \in N$, sample $s_i^t$ from $p_i^t$:
   i. If $\sum_{s_i' \in S_i} R_i^t(s_i')_+ > 0$, then

   $$p_i^{t+1}(s_i) = \frac{R_i^t(s_i)_+}{\sum_{s_i' \in S_i} R_i^t(s_i')_+}, \quad s_i \in S_i.$$

   ii. Otherwise $p_i^{t+1}(s_i) = \frac{1}{|S_i|}$, for all $s_i \in S_i$.
3. Set $t \leftarrow t + 1$ and go to 2.

# Convergence to correlated equilibria

- Let $\mathbf{s}^t = (s_1^t, \ldots, s_n^t)$ be the strategy profile played according to $p_i^t$ at iteration $t$
- The empirical distribution of such strategy profiles is

$$q^t(\mathbf{s}) = \frac{|\{\tau = 1, \ldots, t \mid \mathbf{s}^\tau = \mathbf{s}\}|}{t}, \qquad \mathbf{s} \in \mathbf{S}$$

- The sequence of empirical distributions $q^1, q^2, \ldots$ converges to the set of correlated equilibria almost surely

# Stackelberg equilibrium

# Two-player Stackelberg game

Player 1 (*leader*) and player 2 (*follower*) interact as follows:

1. The leader *publicly* commits to a mixed strategy $p_1 \in \Delta_1$
2. The follower then selects a pure strategy $s_2 \in \mathbf{BR}_2(p_1)$

### The main problem

The leader wants to maximize $U_1(p_1, s_2)$, which depends on unknown $s_2 \in \mathbf{BR}_2(p_1)$. We need a *tie-breaking rule*.

# Tie-breaking

1. The set $\mathbf{BR}_2(p_1)$ contains only one element (no problem!)
2. The set $\mathbf{BR}_2(p_1)$ contains more than one element:
   a. $U_1(p_1, s_2) = U_1(p_1, t_2)$ for all $s_2, t_2 \in \mathbf{BR}_2(p_1)$
   b. The choice of best response is based on the application
   c. The follower breaks ties *in favor* of the leader
   d. The follower breaks ties *to the disadvantage* of the leader

# Strong Stackelberg equilibrium

The follower picks the best response $s_2$ *in favor* of the leader:

$$\max_{p_1 \in \Delta_1} \max_{s_2 \in \mathbf{BR}_2(p_1)} U_1(p_1, s_2)$$

*Strong SE* is a pair $(p_1^*, s_2^*)$ satisfying

$$\max_{s_2 \in \mathbf{BR}_2(p_1^*)} U_1(p_1^*, s_2) = \max_{p_1 \in \Delta_1} \max_{s_2 \in \mathbf{BR}_2(p_1)} U_1(p_1, s_2)$$

$$U_2(p_1^*, s_2^*) = \max_{s_2 \in S_2} U_2(p_1^*, s_2)$$

# Computation of strong SE

The optimal strategy of leader $p_1^*$ can be computed by LP since

$$\max_{p_1 \in \Delta_1} \max_{s_2 \in \mathbf{BR}_2(p_1)} U_1(p_1, s_2) = \max_{s_2 \in S_2} \max_{\substack{p_1 \in \Delta_1 \\ s_2 \in \mathbf{BR}_2(p_1)}} U_1(p_1, s_2)$$

- For each $s_2 \in S_2$ maximize $U_1(p_1, s_2)$ s.t.

$$\begin{aligned} U_2(p_1, s_2) \geq U_2(p_1, t_2) & \qquad \forall t_2 \in S_2 \\ p_1 \in \Delta_1 & \end{aligned}$$

- $p_1^*$ is the optimal solution of an LP with the maximal value

# Strong SE: Example

$$\begin{bmatrix} 2,1 & 4,0 \\ 1,0 & 3,1 \end{bmatrix} \qquad \mathbf{BR}_2(p_1) = \begin{cases} 2 & 0 \le p_1 < 0.5, \\ \{1,2\} & p_1 = 0.5, \\ 1 & 0.5 < p_1 \le 1, \end{cases}$$

$$\max_{s_2 \in \mathbf{BR}_2(p_1)} U_1(p_1, s_2) = \begin{cases} p_1 + 3 & 0 \le p_1 \le 0.5, \\ p_1 + 1 & 0.5 < p_1 \le 1. \end{cases}$$

This gives $p_1^* = 0.5$ (payoff 3.5) and $s_2 \in \{1,2\}$ (payoff 0.5).

# Weak Stackelberg equilibrium

The follower picks $s_2$ *to the disadvantage* of the leader:

$$\max_{p_1 \in \Delta_1} \min_{s_2 \in \mathbf{BR}_2(p_1)} U_1(p_1, s_2)$$

*Weak SE* is a pair $(p_1^*, s_2^*)$ satisfying

$$\min_{s_2 \in \mathbf{BR}_2(p_1^*)} U_1(p_1^*, s_2) = \max_{p_1 \in \Delta_1} \min_{s_2 \in \mathbf{BR}_2(p_1)} U_1(p_1, s_2)$$
$$U_2(p_1^*, s_2^*) = \max_{s_2 \in S_2} U_2(p_1^*, s_2)$$

# Weak SE: Example

$$\begin{bmatrix} 2,1 & 4,0 \\ 1,0 & 3,1 \end{bmatrix} \qquad \mathbf{BR}_2(p_1) = \begin{cases} 2 & 0 \le p_1 < 0.5, \\ \{1,2\} & p_1 = 0.5, \\ 1 & 0.5 < p_1 \le 1, \end{cases}$$

$$\min_{s_2 \in \mathbf{BR}_2(p_1)} U_1(p_1, s_2) = \begin{cases} p_1 + 3 & 0 \le p_1 < 0.5, \\ p_1 + 1 & 0.5 \le p_1 \le 1. \end{cases}$$

- The last function doesn't have maximum on $[0,1]$
- This means that the weak SE doesn't exist

# Zero-sum Stackelberg games

- By the zero-sum assumption, for all $s_2, t_2 \in \mathbf{BR}_2(p_1)$,

$$U_1(p_1, s_2) = U_1(p_1, t_2) = \min_{r_2 \in S_2} U_1(p_1, r_2)$$

- This implies that the leader solves the problem

$$\max_{p_1 \in \Delta_1} \min_{r_2 \in S_2} U_1(p_1, r_2)$$

  whose optimal solution is the maxmin strategy