Introduction

September 9, 2022 8:21 PM

Optimization: single decision-maker optimizes a single objective function
Game theory: multiple agents

Game

- Setup: a number of players/agents, $I = \{1, \dots, N\}$.
- Each player $i \in I$ has a number of actions, $u_i \in \Omega_i$ (acton set), $u = (u_1, \dots, u_N)$ is the players' action profile.
- Each player has individual payoff U_i or cost function J_i.
- Each player takes an action to maximize its payoff or minimize the loss
  - Each player's success in making decisions depends on the decisions of the others.

Forms

- Tree
- Normal (matrix)

Features

- Competitive
  - Non-cooperative (exists competition)
    - Coordination (what's good for one is good for all)
    - Constant-sum (zero-sum or opposing interest)
    - Games of conflicting interests
- Repetition
  - One shot: interact for only a single round
  - Repeated games: each time the same game
  - Dynamic games: characterized by a state, game changes when players interact repeatedly
- Knowledge information
  - Costs of other players
  - Own cost/payoff matrix/function, actions and costs of other players

Solution

- A set of rules to decide how to play the game
- Player is rational if he makes choices that optimizes his expected utility
- Minimax solution
  - Minimizes the player's maximum (worst) expected cost
  - Security strategy
- Best response
  - Play the strategy that gives the lowest cost given your opponents' strategies
  - If each player plays a BR to the strategy of all others, we get Nash equilibrium
  - No regret

Classical game theory

- Equilibrium analysis based on Nash equilibrium
- Alternative justification
  - As the limit point of a repeated play in which less than fully rational players myopically update their behavior

Learning

- Adaptive: best response, fictitious play
- Evolutionary dynamics
  - Selection of strategies according to performance against the aggregate and random

mutations

- Bayesian learning

For 2×2 matrix $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$, $A^{-1} = \frac{1}{\det A}\begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$.

2-payer games

September 13, 2022 5:56 PM

2-player zero-sum games (finite)

- Players $I = \{1,2\}$.
- Player 1 P_1.
  - Decision u_1.
  - $M_1 = \{1, \dots, j, \dots, m\}$.
    - m is the number of possible actions.
    - j is the jth element of P_1.
    - Alternatively, use unit vector, $e_j = \begin{pmatrix} 0 \\ \dots \\ 1 \\ \dots \\ 0 \end{pmatrix}$, 1 at jth position.
- Player 2 P_2.
  - Decision u_2.
  - $M_2 = \{1, \dots, k, \dots, n\}$.
    - n is the number of possible actions.
    - k is the jth element of P_2.
    - Alternatively, use unit vector, $e_k = \begin{pmatrix} 0 \\ \dots \\ 1 \\ \dots \\ 0 \end{pmatrix}$, 1 at kth position.
- Cost function
  - For player 1, $J_1(u_1, u_2)$, e.g. $J_1\left(e_j, e_k\right)$.
  - For player 2, $J_2(u_1, u_2)$.

Def: Game G is a zero sum game if $J_1(u_1, u_2) + J_2(u_1, u_2) = 0$, $\forall u_1, u_2$, or equivalently, $J_2(u_1, u_2) = -J_1(u_1, u_2)$.

Objectives

- P_1 minimizes J_1.
- P_2 minimizes $J_2 = -J_1$ or equivalently maximizes J_1.
- Let $J = J_1$ and use a single cost function, P_1 wants to min it, P_2 wants to max it.
- $J\left(e_j, e_k\right) = a_{jk}$ (scalar) is the cost when P_1 selects jth action and P_2 selects kth action.
- Cost matrix: $A = \begin{pmatrix} \dots & \dots & \dots \\ \dots & a_{jk} & \dots \\ \dots & \dots & \dots \end{pmatrix}$.
  - $J\left(e_j, e_k\right) = e_j^T A e_k = a_{jk}$.
  - P_1 select rows, P_2 select cols.

Security strategy

- Def: P_1 minimizes its worst cost, i.e., in each row, he computes $\max_k a_{jk}$ and picks the row j^* such that $\forall j \in M_1$, $\max_k a_{j^*k} \leq \max_k a_{jk}$. i.e. $j^* = \arg\min_j \max_k a_{jk}$ is the security strategy for P_1.
  - Associated cost is $J_u = J_{ceiling} = \min_j \max_k a_{jk}$.
- Similarly for P_2, in each col k, he finds $\min_j a_{jk}$ and selects col k^* such that $\forall k \in M_2$, $\min_j a_{jk^*} \geq \min_j a_{jk}$, i.e. $k^* = \arg\max_k \min_j a_{jk}$ is the security strategy for P_2.
  - Associated cost is $J_L = J_{floor} = \max_k \min_j a_{jk}$.
- When both P_1 and P_2 use their security strategy, we get a security solution $\left(j^*, k^*\right)$ and the corresponding outcome is $J_0 = a_{j^*k^*} = J\left(e_{j^*}, e_{k^*}\right)$.

- Note: $J_L \le J_0 \le J_u$ (not always equal).

Regret

- If neither P_1 nor P_2 regret their choice, then (j^*, k^*) is an equilibrium.
- A saddle point equilibrium (j^*, k^*) such that (max in row) $a_{j^*k} \le a_{j^*k^*} \le a_{jk^*}$ (min in col).
  - At this moment only, $J_L = J_u = J_0$.

Examples

- $A = \begin{pmatrix} 5 & 3 & -3 \\ 1 & 2 & 0 \\ 3 & 4 & 1 \end{pmatrix}$.
  - P_1 chooses $j^* = 2$, because $2 = \min\{5,2,4\}$ (min of max in each row), $J_u = 2$.
  - P_2 chooses $k^* = 2$, because $2 = \max\{1,2,-3\}$ (max of min in each col), $J_L = 2$.
  - Security strategy is $(j^*, k^*) = (2,2)$ and $J_0 = 2$, P_1 will not regret.
- $A = \begin{pmatrix} 4 & 0 & -1 \\ 0 & -1 & 3 \\ 1 & 2 & 1 \end{pmatrix}$.
  - P_1 chooses $j^* = 3$, because $2 = \min\{4,3,2\}$, $J_u = 2$.
  - P_2 chooses $k^* = 1$, because $0 = \max\{0,-1,-1\}$, $J_L = 0$.
  - Security strategy is $(j^*, k^*) = (3,1)$ and $J_L = 0 \le J_0 = 1 \le J_u = 2$
  - After knowing P_2's choice, P_1 regrets and may choose $j = 2$.

Matching-penny game

- $A = \begin{pmatrix} -1 & 1 \\ 1 & -1 \end{pmatrix}$.
- The security strategies are $(j^*, k^*) = (1,1), (2,2), (1,2), (2,1)$.
- For any of the 4 cases, one of P_1, P_2 regrets. There is no saddle point equilibrium in this case.

Mixed strategies

- Let P_1, P_2 randomize their choices, x_j be the probability that P_1 selects jth action, y_k be the probability that P_2 selects kth action, $x_j, y_k \in [0,1]$. Let $x = \begin{pmatrix} x_1 \\ \vdots \\ x_j \\ \vdots \\ x_n \end{pmatrix}$ be the probability vector for P_1, $y = \begin{pmatrix} y_1 \\ \vdots \\ y_j \\ \vdots \\ y_n \end{pmatrix}$ be the probability vector for P_2, $\sum_{j=1}^{m} x_j = \sum_{k=1}^{n} y_k = 1$.
  - x is the mixed strategy of P_1, y is the mixed strategy of P_2.
  - If $x_j = 1$, $x = e_j$ is the pure strategy.
- Simplex $x \in \Delta \subset \mathbb{R}^m$, represented by $\sum_{j=1}^{m} x_j = 1$.
- Since outcomes (costs) are no longer deterministic, expected value of cost is $E(J) = \bar{J}(x,y) = \sum_{k=1}^{n} \sum_{j=1}^{m} a_{jk} x_j y_k = x^T A y$.
  - If $x = e_j$, $y = e_k$, $\bar{J}(e_j, e_k) = e_j^T A e_k = a_{jk}$.
- Def: x^* is a mixed security strategy for P_1 if $x^* = \arg\min_{x \in \Delta_1} \max_{y \in \Delta_2} x^T A y$ with $\bar{J}_U$, y^* is a mixed security strategy for P_2 if $x^* = \arg\max_{y \in \Delta_2} \min_{x \in \Delta_1} x^T A y$ with $\bar{J}_L$.
- Def: (x^*, y^*) is a saddle point equilibrium in mixed strategy if $\forall x \in \Delta_1$, $y \in \Delta_2$, $(x^*)^T A y \le (x^*)^T A y^* \le x^T A y^*$.
  - If $x^* = e_{j^*}$, $y^* = e_{k^*}$, we recover pure strategy saddle point.
- Von Neumann theorem: in any 2 player zero sum finite game, $\bar{J}_L = \bar{J}_U = \bar{J}(x^*, y^*)$ and any such game has a saddle point equilibrium (no regret) in mixed strategy.

Computing mixed-strategy security strategies and saddle point strategy

- Graphical method for $A_{2\times 2}$ and can be extended to $2 \times n$.
- In general, a mixed strategy $x = \begin{pmatrix} x_1 \\ \cdots \\ x_m \end{pmatrix} = \sum_{j=1}^{m} x_j e_j \in \Delta_1$ is a linear combination of pure strategies.
- $A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$.
  - For P_1, $V_1(x) = \min_{x\in\Delta_1} \max_{k\in\{1,2\}} x^T A e_k = \min_{x_1\in[0,1]} \max \begin{Bmatrix} (a_{11}-a_{21})x_1 + a_{21} \\ (a_{12}-a_{22})x_1 + a_{22} \end{Bmatrix}$.
    - $x_1^* = \arg\min_{x_1\in[0,1]} \max \begin{Bmatrix} (a_{11}-a_{21})x_1 + a_{21} \\ (a_{12}-a_{22})x_1 + a_{22} \end{Bmatrix}, x_2^* = 1 - x_1^*$.
  - For P_2, $V_2(y) = \max_{y\in\Delta_2} \min_{j\in\{1,2\}} e_j^T A y = \max_{y_1\in[0,1]} \min \begin{Bmatrix} (a_{11}-a_{12})y_1 + a_{12} \\ (a_{21}-a_{22})y_1 + a_{22} \end{Bmatrix}$.
    - $y_1^* = \arg\max_{y_1\in[0,1]} \min \begin{Bmatrix} (a_{11}-a_{12})y_1 + a_{12} \\ (a_{21}-a_{22})y_1 + a_{22} \end{Bmatrix}, y_2^* = 1 - y_1^*$.
- For the matching penny game
  - $x_1^* = x_2^* = y_1^* = y_2^* = \frac{1}{2}$.

Dominated strategies
- $A_{m\times n}$ cost matrix.
- For P_1, strategy j dominates r if $a_{jk} \le a_{rk}, \forall k \in M_2$ and $a_{jk} < a_{rk}$ for at least one $k \in M_2$.
- For P_2, strategy k dominates q if $a_{jk} \ge a_{jq}, \forall j \in M_1$ and $a_{jk} > a_{jq}$ for at least one $j \in M_1$.
- Prop: In a matrix game A, assume strategy $j_1, \dots, j_l$ are dominated, then P_1 has an optimal strategy $x_{j_1} = \cdots = x_{j_l} = 0$. Any optimal strategy after removing these from the game will be optimal for the original game
- e.g. $A = \begin{pmatrix} 2 & 1 & 4 \\ 0 & 2 & 1 \\ 1 & 5 & 3 \\ 4 & 3 & 2 \end{pmatrix}$ can be reduced to $A = \begin{pmatrix} 1 & 4 \\ 2 & 1 \end{pmatrix}$.
  - $x^* = \begin{pmatrix} x_1^* \\ x_2^* \\ 0 \\ 0 \end{pmatrix}, y^* = \begin{pmatrix} 0 \\ y_2^* \\ y_3^* \end{pmatrix}$.

2 player nonzero sum games
- Everything carries through from zero sum games except that we cannot use a single cost function (matrix).
- P_1 has cost J_1, P_2 has cost J_2, $J_1 + J_2 \ne 0$.

Prisoner's dilemma game
- $A = \begin{pmatrix} 5 & 0 \\ 15 & 1 \end{pmatrix}, B = \begin{pmatrix} 5 & 15 \\ 0 & 1 \end{pmatrix}$.
  - j for row, k for col.
- Security strategy: optimum (min cost) in worst case scenario.
  - $j^* = \arg\min_{j\in\{1,2\}} \max_{k\in\{1,2\}} e_j^T A e_k = \arg\min_{j\in\{1,2\}}\{5,15\} = 1$ (confess).
  - $k^* = \arg\min_{k\in\{1,2\}} \max_{j\in\{1,2\}} e_j^T B e_k = \arg\min_{j\in\{1,2\}}\{5,15\} = 1$ (confess).
  - $(j^*, k^*) = (1,1)$.
- No-regret
  - After P_1 knowing P_2 selects $k^* = 1$ (first col in A), P_1 will not regret as $5 < 15$.
  - After P_2 knowing P_1 selects $j^* = 1$ (first row in B), P_2 will not regret as $5 < 15$.
  - There is an equilibrium

No-regret equation for nonzero sum games
- With security strategy (j^*, k^*).

- For P_1, $e_{j^*}^T A e_{k^*} \le e_j^T A e_{k^*}, \forall j \in \{1, \dots, n\}$.
- For P_2, $e_{j^*}^T B e_{k^*} \le e_{j^*}^T B e_k, \forall k \in \{1, \dots, m\}$.

Chicken game

- $A = \begin{pmatrix} 0 & 1 \\ -1 & 10 \end{pmatrix}, B = \begin{pmatrix} 0 & -1 \\ 1 & 10 \end{pmatrix}$.
- $j^* = \arg\min_{j\in\{1,2\}} \max_{k\in\{1,2\}} e_j^T A e_k = \arg\min_{j\in\{1,2\}}\{1,10\} = 1$ (swerve).
- $k^* = \arg\min_{k\in\{1,2\}} \max_{j\in\{1,2\}} e_j^T A e_k = \arg\min_{j\in\{1,2\}}\{1,10\} = 1$ (swerve).
- After P_1 knows P_2 selects $k^* = 1$, P_1 will regret and selct $j = 2$ to get cost -1.

Extension to mixed strategies on 2PZS games

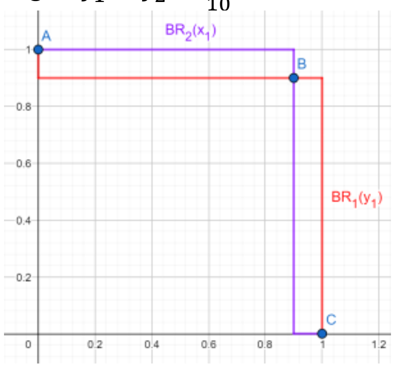
- P_1: minimum expected cost $\bar{J}_1(x, y) = x^T A y$ w.r.t. $x \in \Delta_1$.
- P_2: minimum expected cost $\bar{J}_2(x, y) = x^T B y$ w.r.t. $y \in \Delta_2$.
- Nash equilibrium: (x^*, y^*) is a no-regret (Nash) equilibrium if:
  - $(x^*)^T A y^* \le x^T A y^*$ for all $x \in \Delta_1$.
  - $(x^*)^T B y^* \le (x^*)^T B y$ for all $y \in \Delta_2$.
  - Note: for zero sum games, $B = -A$, so $(x^*)^T A y \le (x^*)^T A y^* \le x^T A y^*$.

Best Response strategy

- P_1's best response strategy $\xi \in \Delta_1$ is made up of a set of strategies obtained as follows
  - Given P_2 strategy $y \in \Delta_2$, $BR_1(y) = \{\xi \in \Delta_1 : \xi^T A y \le x^T A y, \forall x \in \Delta_1\}$.
  - $BR_1(y) : \Delta_2 \rightrightarrows \Delta_1$ is a set valued map.
  - $BR_1(y) = \arg\min_{x\in\Delta_1} x^T A y$ for given $y \in \Delta_2$.
- P_2's best response strategy $\eta \in \Delta_2$ is made up of a set of strategies obtained as follows
  - $BR_2(x) = \{\eta \in \Delta_2 : x^T B \eta \le x^T B y, \forall y \in \Delta_2\}$.
  - $BR_2(y) : \Delta_1 \rightrightarrows \Delta_2$ is a set valued map.
  - $BR_2(y) = \arg\min_{y\in\Delta_2} x^T A y$ for given $x \in \Delta_1$.
- For the NE definition
  - $P_1 : x^* \in BR_1(y^*)$, $P_2 : y^* \in BR_2(x^*)$ iff (x^*, y^*) is a Nash equilibrium.
  - NE lies at the intersection of their BR strategy maps.

Graphical computation of NE in 2×2 games

- $A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}, B = \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix}$.
- $x = \begin{pmatrix} x_1 \\ 1 - x_1 \end{pmatrix} \in \Delta_1, y = \begin{pmatrix} y_1 \\ 1 - y_1 \end{pmatrix} \in \Delta_2, x_1, y_1 \in [0,1]$.
- $\xi = \begin{pmatrix} \xi_1 \\ 1 - \xi_1 \end{pmatrix} \in \Delta_1, \eta = \begin{pmatrix} \eta_1 \\ 1 - \eta_1 \end{pmatrix} \in \Delta_2, \xi_1, \eta_1 \in [0,1]$.
- $BR_1(y) = \left\{ \begin{pmatrix} \xi_1 \\ 1 - \xi_1 \end{pmatrix} : (\xi_1, 1 - \xi_1) \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \begin{pmatrix} y_1 \\ 1 - y_1 \end{pmatrix} \le (x_1, 1 - x_1) \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \begin{pmatrix} y_1 \\ 1 - y_1 \end{pmatrix} \right\}$.
  - $= \arg\min_{x_1\in[0,1]} x_1(\tilde{a} y_1 - \tilde{c}_1)$ where $\begin{cases} \tilde{a} = a_{11} - a_{12} - a_{21} + a_{22} \\ \tilde{c}_1 = a_{22} - a_{12} \end{cases}$.
  - $= \begin{cases} 0, if\ \tilde{a} y_1 - \tilde{c}_1 > 0 \\ [0,1], if\ \tilde{a} y_1 - \tilde{c}_1 = 0. \\ 1, if\ \tilde{a} y_1 - \tilde{c}_1 < 0 \end{cases}$
- $BR_2(x) = \arg\min_{y_1\in[0,1]} y_1(\tilde{b} x_1 - \tilde{d}_2)$ where $\begin{cases} \tilde{b} = b_{11} - b_{12} - b_{21} + b_{22} \\ \tilde{d}_2 = b_{22} - b_{21} \end{cases}$.
  - $= \begin{cases} 0, if\ \tilde{b} x_1 - \tilde{d}_2 > 0 \\ [0,1], if\ \tilde{b} x_1 - \tilde{d}_2 = 0. \\ 1, if\ \tilde{b} x_1 - \tilde{d}_2 < 0 \end{cases}$
- Plotting them on the same graphs, the NE will be the intersections.

- E.g. NE for chicken games.
    - $\begin{cases}\tilde{a} = 10 \\ \tilde{c_1} = 9\end{cases}, \begin{cases}\tilde{b} = 10 \\ \tilde{d_2} = 9\end{cases}.$
    - $BR_1(y_1) = \begin{cases} 0, if\ y_1 > \frac{9}{10} \\ [0,1], if\ y_1 = \frac{9}{10} \\ 1, if\ y_1 < \frac{9}{10} \end{cases}.$
    - $BR_2(x_1) = \begin{cases} 0, if\ x_1 > \frac{9}{10} \\ [0,1], if\ x_1 = \frac{9}{10} \\ 1, if\ x_1 < \frac{9}{10} \end{cases}.$
    - Three intersections
        - (0,1): $(x^*, y^*) = \left(\begin{pmatrix}0\\1\end{pmatrix}, \begin{pmatrix}1\\0\end{pmatrix}\right)$, P_1 doesn't swerve, P_2 swerves, no regret $J_1 = -1$, $J_2 = 1$.
        - (1,0): $(x^*, y^*) = \left(\begin{pmatrix}1\\0\end{pmatrix}, \begin{pmatrix}0\\1\end{pmatrix}\right)$, P_1 swerves, P_2 doesn't swerve, no regret $J_1 = 1$, $J_2 = -1$.
        - $\left(\frac{9}{10}, \frac{9}{10}\right)$: $(x^*, y^*) = \left(\begin{pmatrix}\frac{9}{10}\\\frac{1}{10}\end{pmatrix}, \begin{pmatrix}\frac{9}{10}\\\frac{1}{10}\end{pmatrix}\right)$, P_1, P_2 both swerve with 90% probability, no regret $J_1 = J_2 = \frac{1}{10}$.
        - 

N-player games

October 11, 2022 6:20 PM

N-player finite games $(A_1, A_2, \dots, A_N)$

- Setup: $I = \{1, \dots, N\}$ set of players $P_1, \dots, P_N$.
- P_i=player i, $i \in I$.
- Ω_i=finite action set with $|\Omega_i| = m_i$.
- Δ_i=mixed strategy set.
- $x_i = \begin{pmatrix} x_{i1} \\ \dots \\ x_{ij} \\ \dots \\ x_{im_i} \end{pmatrix} \in \mathbb{R}^{m_i}$ with x_{ij}=probability of choosing jth action in $\{1, \dots, m_i\}$.
    - $\sum_{j=1}^{m_i} x_{ij} = 1$, $x_{ij} \geq 0$.
- Let N-tuple of all mixed strategies be $x = (x_1, \dots, x_i, \dots, x_N) \in \Delta = \Delta_1 \times \cdots \Delta_N$
    - $x_{-i} = (x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_N)$ is mixed strategy of everyone except P_i.
    - Write $x = (x_i, x_{-i})$.
    - Note: when $N = 2$, $x = (x_1, x_2) = (x_1, x_{-1}) = (x_2, x_{-2})$ with $x_{-1} = x_2$, $x_1 = x_{-2}$.
- Expected cost of P_i given x_{-i}, $\bar{J}_i(x_i, x_{-i}) = \sum_{j=1}^{m_i} J_i(e_{ij}, x_{-i}) x_{ij}$ (linear in x_i).

Nash Equilibrium for N-player

- $x^* = (x_1^*, \dots, x_i^*, \dots, x_N^*) = (x_i^*, x_{-i}^*) \in \Delta$, such that $\forall i \in [m]$, $\bar{J}_i(x_i^*, x_{-i}^*) \leq \bar{J}_i(x_i, x_{-i}^*)$, $\forall x_i \in \Delta_i$.
- BR map of P_i: $BR_i(x_{-i}) = \{\xi_i \in \Delta_i : \bar{J}_i(\xi_i, x_{-i}) \leq \bar{J}_i(x_i, x_{-i}), \forall x_i \in \Delta_i\} = \arg\min_{x_i \in \Delta_i} \bar{J}_i(x_i, x_{-i})$.
- Note: $x_i^* \in BR_i(x_{-i}^*)$, $\forall i = 1, \dots, N$ or equivalently $x^* \in BR(x^*)$ where $BR(x^*) = \begin{pmatrix} BR_1(x_{-1}^*) \\ \dots \\ BR_N(x_{-N}^*) \end{pmatrix}$.
    - x^* is a fixed point of BR map.

Nash theorem:

- Any N-player finite game has at least 1 NE in mixed strategy
- Proving the existence of a fixed point of BR map
    - Apply Kakutani's theorem for $\Phi = BR$, $S = \Delta = \Delta_1 \times \cdots \times \Delta_N$.

Mathematics background:

- Graph of a function:
    - Point valued: $Graph(f) = \{(y, x) : y = f(x), x \in dom(f)\} = dom(f) \times Range(f)$.
    - Set valued: $Graph(f) = \{(y, x) : y \in f(x), x \in S\} \subset S^2$.
- A closed set contains all its limit points.
    - $\forall x$ such that $\exists \{x_n\} \in S$ such that $\lim_{n \to \infty} x_n = x$, then $x \in S$.
- A set is compact if its closed and bounded.
- A set is convex if $\forall x, y \in S, \alpha, \beta \in [0,1]$, $\alpha + \beta = 1$, then $\alpha x + \beta y \in S$.

2 fixed point theorems:

- Brower's fixed point theorem: Let S be a compact convex set in $\mathbb{R}^n$, $f: S \to S$ a continuous function, then $\exists x \in S$ such that $x = f(x)$.
- Kakutani's fixed point theorem: Let S be a compact convex set in $\mathbb{R}^n$ and $\Phi : S \rightrightarrows S$ (set valued) with the image of $x \in S$ denoted $\Phi(x) \subset S$ such that:
    - $\Phi(x)$ is nonempty and convex for any $x \in S$.
    - Φ has a closed graph for any $x \in S$.

- Then there exists at least one $x \in S$ such that $x \in \Phi(x)$.

Dominated strategies for N player non-cooperate game

- A strategy $z_i \in \Delta_i$ weakly dominates strategy $x_i \in \Delta_i$, if $J_i(z_i, x_{-i}) \le J_i(x_i, x_{-i})$, $\forall x_{-i} \in \Delta_{-i}$.
- A strategy x_i is undominated if no such z_i exists.
- If $J_i(z_i, x_{-i}) < J_i(x_i, x_{-i})$ $\forall x_{-i} \in \Delta_{-i}$, then z_i strictly dominates x_i.
- If we replace Δ_i and Δ_{-i} with pure strategy set,we get the same definition as before
  - $J_i(x, y) = x^T A y$.
  - When y is a pure strategy, it chooses a column of A.
  - When x is a pure strategy, it chooses a row.
- e.g. $A = \begin{pmatrix} 2 & 0 \\ 0 & 2 \\ \frac{3}{2} & \frac{3}{2} \end{pmatrix}$.
  - No row dominates another row.
  - However, for $x_1 = \left(\frac{1}{2}, \frac{1}{2}, 0\right)$, $J_1(x_1, y) = y_1 + y_2 = 1$ for all $y \in \Delta_y$.
  - For $x_2 = (0,0,1)$, $J_1(x_2, y) = \frac{3}{2}(y_1 + y_2) = \frac{3}{2}$.
  - x_1 strictly dominates x_2.

Support characterization of NE

- Def: for a mixed strategy $x_i \in \Delta_i$, we define its support or carrier as the set of pure strategies that are assigned positive probabilities.
  - $\text{supp}(x_i) = \{j \in M_i : x_{i,j} > 0\}$.
- e.g.
  - $x_i^1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$, $\text{supp}(x_i^1) = \{1\}$.
  - $x_i^2 = \begin{pmatrix} 0.5 \\ 0.5 \\ 0 \end{pmatrix}$, $\text{supp}(x_i^2) = \{1,2\}$.
  - $x_i^3 = \begin{pmatrix} 1/3 \\ 1/3 \\ 1/3 \end{pmatrix}$, $\text{supp}(x_i^3) = \{1,2,3\}$.
- e.g.
  - 2PZSG, where $A = \begin{pmatrix} 1 & 3 & 5 \\ 4 & 2 & 3 \\ 6 & 1 & 4 \end{pmatrix}$.
  - NE is $(x^*, y^*) = \left(\begin{pmatrix} 1/5 \\ 4/5 \\ 0 \end{pmatrix}, \begin{pmatrix} 2/5 \\ 0 \\ 3/5 \end{pmatrix}\right)$, $J_1(x^*, y^*) = \frac{17}{5}$.
  - $\text{supp}(x^*) = \{1,2\}$.
  - $\text{supp}(y^*) = \{1,3\}$.
  - $Ay^* = \begin{pmatrix} 17/5 \\ 17/5 \\ 24/5 \end{pmatrix}$, $Ax^* = \begin{pmatrix} 17/5 \\ 11/5 \\ 17/5 \end{pmatrix}$.
  - For player 1 with supported action j, we have $J_1(e_1, y^*) = J_1(x^*, y^*)$.
- Support characterization theorem: let $x^* = (x_i^*, x_{-i}^*) \in \Delta_x$. Then $x^* \in NE(G)$ is a mixed strategy NE if and only if $\forall i \in N, \forall j \in \text{supp}(x_i^*), J_i(e_{ij}, x_{-i}^*) = \min_{w_i \in \Delta_i} J_i(w_i, x_{-i}^*)$.
  - Proof ($\Leftarrow$): let $x_i^* \in \Delta_i$, $x^* \in \Delta_x$.
    - $\forall j \in \text{supp}(x_i^*), J_i(e_{ij}, x_{-i}^*) \le J_i(w_i, x_{-i}^*)$, $\forall w_i \in \Delta_i$, w_i is any point in the set of mixed strategies.
    - $J_i(x_i^*, x_{-i}^*) = \sum_{j \in \text{supp}(x_i^*)} J_i(e_{ij}, x_{-i}^*) x_{ij}^*$.

- $\leq \sum_{j \in \mathrm{supp}(x_i^*)} J_i(w_i, x_{-i}^*) x_{ij}^*$.
- $= J_i(w_i, x_{-i}^*) \sum_{j \in \mathrm{supp}(x_i^*)} x_{ij}^* = J_i(w_i, x_{-i}^*)$.
- So x^* is an NE.

○ Proof (⇒): let $x^* = (x_i^*, x_{-i}^*)$ be an NE.
- $J_i(x_i^*, x_{-i}^*) \leq J_i(w_i, x_{-i}^*), \forall w_i \in \Delta_i$.
- $J_i(x_i^*, x_{-i}^*) \leq J_i(e_{ij}, x_{-i}^*), \forall j \in M_i$.
- Let $J_i(e_{ij}, x_{-i}^*) = J_i(x_i^*, x_{-i}^*) + \epsilon_{ij}$, $\epsilon_{ij} \geq 0$.
- $J_i(x_i^*, x_{-i}^*) = \sum_{j \in \mathrm{supp}(x_i^*)} J_i(e_{ij}, x_{-i}^*) x_{ij}^*$.
- $= \sum_j \left(J_i(x_i^*, x_{-i}^*) + \epsilon_{ij}\right) x_{ij}^*$.
- $= J_i(x_i^*, x_{-i}^*) + \sum_j \epsilon_{ij} x_{ij}^*$.
- So $\sum_j \epsilon_{ij} x_{ij}^* = 0$.
- And thus, $\forall j \in \mathrm{supp}(x_i^*), J_i(e_{ij}, x_{-i}^*) = J_i(x_i^*, x_{-i}^*) = \min_{w_i \in \Delta_i} J_i(w_i, x_{-i}^*)$.

Repeated games

October 18, 2022 6:12 PM

N-player game: P_i chooses its mixed strategy $x_i^k \in \Delta_i$ at iteration k of the play

- let e_i^k be the action of P_i at time k.
- $P\left(e_i^k = e_j\right) = x_{ij}^k$.
- $E\left(e_i^k\right) = \sum_{j=1}^{m_i} e_j x_{ij}^k = x_i^*$.

Let x_{-i}^k be the mixed strategy of the others at iteration k, expected cost of P_i at iteration k is $\bar{J}_i\left(x_i^k, x_{-i}^k\right)$.

Goal: use an iterative process to update their own strategy such that in the long run, $x^k = \left(x_i^k, x_{-i}^k\right)$ converges to a NE, $\lim_{k\to\infty} x^k = x^*$, where $x^* = \left(x_i^*, x_{-i}^*\right)$.

Recall: $\forall i \in I, \overline{J_i}\left(x_i^*, x_{-i}^*\right) \leq \overline{J_i}\left(x_i, x_{-i}^*\right)\ \forall x_i \in \Delta_i \Leftrightarrow \forall i \in I, x_i^* \in BR_i\left(x_{-i}^*\right) = \arg\min_{x_i \in \Delta_i} J_i\left(x_i, x_{-i}^*\right)$

$\Leftrightarrow x^* \in BR(x^*) = \begin{pmatrix} \vdots \\ BR_i\left(x_{-i}^*\right) \\ \vdots \end{pmatrix}$.

Note: in general, each P_i will have

- Some info ω_i: its own cost J_i, x_{-i}^k, $x_{i'}^k$ for some $i' \in I$, action $e_{i'}^k$, $i' \in I$.
- Internal state $Z_i^k \in \mathbb{R}^{q_i}$: update based on the ω_i^k and the play. They will map this internal state into a strategy to use
  - $x_i^k = \sigma_i\left(z_i^k\right)$ where $\sigma_i : \mathbb{R}^{q_i} \to \Delta_i$.
  - $z_i^{k+1} = z_i^k + \alpha_k f\left(z_i^k, \omega_i^k\right)$.
- Iterative process of P_i: $\Sigma_i = \begin{cases} z_i^k \\ x_i^k \end{cases}, \Sigma_{-i} = \begin{cases} z_{-i}^k \\ x_{-i}^k \end{cases}$.
  - $\Sigma = \begin{cases} z^k \\ x^k \end{cases}$.
- It gives a feedback, interconnected discrete time dynamical system

BR-play

- $\omega_i^k \to J_i, x_{-i}^k$, at next iteration $(k+1)$, P_i sets $x_i^{k+1} \in BR_i\left(x_{-i}^k\right)$.
- In the limit, $x_i^{k+1} = x_i^k = \overline{x_i}$, $x_{-i}^{k+1} = x_{-i}^k = \overline{x_{-i}}$, so $\overline{x_i} \in BR_i\left(\overline{x_{-i}}\right)$, $\forall i \in I$.
  - For 2 players, $\overline{x_1} \in BR_1\left(\overline{x_2}\right)$, $\overline{x_2} \in BR_2\left(\overline{x_1}\right)$.
- $\bar{x} = \left(\overline{x_i}, \overline{x_{-i}}\right)$ is a NE.

Smooth (perturbed) BR-play

- $\widetilde{BR} = \arg\min_{x_i \in \Delta_i} \overline{J_i}(x_i, x_{-i}) - \epsilon v_i\left(x_i\right)$, for small $\epsilon > 0$, v_i strictly convex in x_i.
  - Example v_i: softmax, softmin.
- Algorithm: $x_i^{k+1} = \widetilde{BR}_i(x_{-i}^k)$ since $\widetilde{BR}_i$ is not set-valued with perturbation.
- Limit point: $\overline{x_i} = \widetilde{BR}_i\left(\overline{x_{-i}}\right)$, for all i.
  - $\overline{x_i}$ is a perturbed NE (logit equilibrium/Nash distribution).

Relaxed BR-play

- $x_i^{k+1} = \alpha_k \widetilde{BR}_i\left(x_{-i}^k\right) + \left(1 - \alpha_k\right) x_i^k$ with $0 < \alpha_k < 1$.
- Limit point: $\overline{x_i} = \widetilde{BR}_i\left(\overline{x_{-i}}\right)$, for all i.

Fictitious play

- $\omega_i^k \to J_i, e_{i'}^k, \forall i' \in I$.
  - $e_{i'}^k$ is the action used by player i' at previous play k.

- Idea: use empirical frequency of an action as approximation of probability of using that action.
  - For player i, for $i' \in I$, $\widehat{x_{i'}^k} = \frac{1}{k}\sum_{k'=0}^{k-1} e_{i'}^{k'}$ is the estimation of mixed strategy of $P_{i'}$.
- For all $i' \in I$, $i' \neq i$, we need to build $\widehat{x_{-i}^k}$ as the state Z_i^k.
- Set $x_i^k = \widetilde{BR}_i\left(\widehat{x_{-i}^k}\right)$, i.e. best response to a fictitious strategy.
- To find Z_i^{k+1} iteratively:
  - $\widehat{x_i^{k+1}} = \frac{1}{k+1}\sum_{k'=0}^{k} e_{i'}^{k'} = \frac{1}{k+1}\left(\frac{k}{k}\sum_{k'=0}^{k-1} e_{i'}^{k'} + e_{i'}^k\right) = \frac{1}{k+1}\left(e_{i'}^k - \widehat{x_{i'}^k}\right) + \widehat{x_{i'}^k}$.
  - $Z_i^{k+1} = Z_i^k + \frac{1}{k+1}\left(e_{-i'}^k - Z_{-i}^k\right)$.

Tutorial content

- Definition:
  - Play same game multiple times
  - Each agent tries to improve their cost/update action
- Perspectives
  - Design an algorithm to perform well in certain games
  - Create a model of how players perform and analyze outcome
  - e.g. population game, predator/prey dynamics
- Agents
  - Goal: minimize cost in the game
  - Learn a strategy to minimize cost
  - If we get to a point where all agents stop updating their strategies, because they cannot improve their cost, then we are at an NE.
  - $J_i\left(x_i^k, x_{-i}^k\right) \le J_i\left(y_i, x_{-i}^k\right), \forall y_i$.
- How do agents update their strategies
  - w_i observations from the environment.
    - Cost, other players' actions.
    - Affect the potential algorithm
  - z_i internal state.
  - x_i strategy.
- Player i's process
  - $z_i^{k+1} = z_i^k + \gamma_k f_i\left(z_i^k, w_i^k\right) = \widetilde{f}_i\left(z_i^k, w_i^k\right)$.
  - $x_i^{k+1} = \sigma\left(z_i^{k+1}\right)$ is the action.
  - γ_k is the learning rate.
- Best response dynamics
  - $w_i^k = J_i, x_i^k, x_{-i}^k$.
  - $z_i^{k+1} = \widetilde{f}_i\left(z_i^k, w_i^k\right) = x_{-i}^k$ (other players' BR).
  - $x_i^{k+1} = BR_i\left(z_i^{k+1}\right) = BR_i\left(x_{-i}^k\right)$.
  - For rock-paper-scissors: NE= $\left(\frac{1}{3}, \frac{1}{3}, \frac{1}{3}\right)$ (completely randomly).
- Fictitious play (finite action game)
  - $w_i^k = e_{-i}^k$ (realized pure strategy at iteration k).
  - $z_i^{k+1} = \frac{k}{k+1} z_i^k + \frac{1}{k+1} e_{-i}^k$
    - measure the frequency of pure strategy
    - Estimate of the mixed strategies of other players
  - $x_i^{k+1} = BR_i\left(z_i^{k+1}\right)$.
- Reinforcement learning (fictitious play)
  - $w_i^k = u_i\left(x_i^k, x_{-i}^k\right) = -J\left(x_i^k, x_{-i}^k\right)$ (assuming $u_i \ge 0$).
  - $z_i^{k+1} = z_i^k + \pi_i^k e_i^k$.
  - $\left[x_i^{k+1}\right]_m = \frac{\left[z_i^{k+1}\right]_m}{\sum_p \left[z_i^{k+1}\right]_p}$ where $[\cdot]_m$ is the mth element of the vector.

Infinite action games

October 25, 2022 6:29 PM

N player infinite action (continuous kernel) game

- $I = \{1,2,\dots,N\}$.
- Each P_i has continum action set Ω_i.
  - Ω_i is convex, non-empty and compact set in $\mathbb{R}^{n_i}$.
  - e.g. $\Omega_i = [a,b]$.
  - $\Omega = \Omega_1 \times \cdots \times \Omega_N$.
- Cost: $J_i : \Omega \to \mathbb{R}$.
  - Jointly continuous in its arguments
- Let $u_i \in \Omega_i$ be one action of P_i, the action profile is $u = (u_i, u_{-i})$.

NE in CK game

- $u^* = (u_i^*, u_{-i}^*)$ is a NE of game if $J_i(u_i^*, u_{-i}^*) \le J_i(u_i, u_{-i}^*)$, $\forall u_i \in \Omega_i$, $\forall i$.
- $BR_i(u_{-i}) = \arg\min_{u_i \in \Omega_i} J_i(u_i, u_{-i})$.
  - u^* is a NE if $u_i^* \in BR_i(u_{-i}^*)$, $\forall i \in I$ (intersection of all BRs).

For 2 player zero sum game, we still want to have:

- P_1, $\min_{u_1} \max_{u_2} J$.
- P_2, $\max_{u_2} \min_{u_1} J$.
- However, we may be able to solve them separately using partial gradients.

Note: we can also have mixed-strategy, but we don't need it in general
Existence of a NE is guaranteed under relatively mild assumption

DFG theorem:

- Consider a CK game where $\Omega_i \in \mathbb{R}^{n_i}$ is non-empty, convex, compact. J_i is jointly continuous on its argument and convex in u_i. Then the game admits at least one NE in pure strategies
- Possible relaxation of Ω_i: convex and closed if J_i is assumed to be radially unbounded.
  - i.e. as $\|u_i\| \to \infty$, $J_i(u_i, u_{-i}) \to \infty$ for all given u_{-i}.

Optimization

- $f : \Omega \to \mathbb{R}$ continuous.
- f is convex if $\forall u, v \in \Omega$, $\alpha \in [0,1]$, $f(\alpha u + (1-\alpha)v) \le \alpha f(u) + (1-\alpha) f(v)$.
- When f is C^1, f is convex if $f(v) - f(u) \ge \nabla f^T(u)(v-u)$, $\forall u, v \in \Omega$.
  - $\nabla f(u)$ is monotone, $(\nabla f(v) - \nabla f(u))^T (v-u) \ge 0$.
- When f is C^2, f is convex if $\nabla^2 f(u) \ge 0$, $\forall u \in \Omega$.
- Normal cone:
  - $N_\Omega(u^*) = \{v : v^T(u-u^*) \le 0, \forall u \in \Omega\}$.
  - If $u^* \in int(\Omega)$, $\nabla f(u^*) = 0$.
  - If $u^* \in \partial\Omega$, $\nabla f(u^*)$ should point into Ω.
- Minimization of a convex function
  - $u^* = \arg\min_{u \in \Omega} f(u)$.
  - If f is C^1, u^* is a minimizer if $\nabla f(u^*)^T(u-u^*) \ge 0$, $\forall u \in \Omega \Leftrightarrow -\nabla f(u^*) \in N_\Omega(u^*)$.
- Euclidean projection.
  - $T_\Omega(y) = \arg\min_{x \in \Omega} \|x-y\|^2$ is the Euclidean projection of y to a set Ω.
  - If Ω is convex, $T_\Omega(y)$ is unique: $y - T_\Omega(y)$ $\perp$tangent plane to Ω.
  - u^* is a min of f if $u^* = T_\Omega(u^* - \alpha \nabla f(u^*))$, $\forall \alpha > 0$.
    - $T_\Omega(u^* + \alpha N_\Omega(u^*)) = u^*$.

Partial gradient

- Assume J_i is C^1 and convex in u_i, define $\nabla_{u_i} J_i = \frac{\partial J_i}{\partial u_i}(u_i, u_{-i})$.
- Stacked partial gradient (pseudo gradient) of the game: $F(u) = \begin{pmatrix} \nabla_{u_1} J_1 \\ \vdots \\ \nabla_{u_N} J_N \end{pmatrix}$.
- Not a true gradient unless $J_1 = J_2 = \cdots = J_N$(potential games)
- Then $BR_i(u_i) = \arg\min_{u_i \in \Omega_i} J_i(u_i, u_{-i})$ is equivalent to the following:
  - $\left(\nabla_{u_i} J_i(u^*)\right)^T (u_i - u_i^*) \geq 0, \forall i \in I$.
  - $\sum_{i=1}^N \left(\nabla_{u_i} J_i(u^*)\right)^T (u_i - u_i^*) \geq 0 \Leftrightarrow F(u^*)^T (u - u^*) \geq 0, \forall u \in \Omega$.
  - $-F(u^*) \in N_\Omega(u^*)$.
  - $u^* \in T_\Omega(u^* - \alpha F(u^*)), \forall \alpha > 0$.
- This is the characterization of the NE.

Note: if $u^* \in int(\Omega)$, $F(u^*) = 0$ u^* is an inner NE.
If $\Omega = \mathbb{R}$, $f(u^*) = 0$ always.

BR-play

- At iteration k, given other players' actions u_{-i}^k, then at next iteration $k+1$, P_i can play a BR_i to these actions: $u_i^{k+1} = BR_i(u_{-i}^k)$.
- It is equivalent to solving a minimization at each iteration
- Variants: $u_i^{k+1} = (1 - \alpha_i) u_i^k + \alpha_i BR(u_{-i}^k)$ with $\alpha_i \in (0,1), \forall i \in I$.

Projected gradient play/better-response play

- $u_i^{k+1} = T_{\Omega_i}\left(u_i^k - \alpha \nabla_{u_i} J_i(u_i^k, u_{-i}^k)\right)$.
- Cheaper computationally, since only gradient is calculated
- If $\Omega_i = \mathbb{R}$, then $u_i^{k+1} = u_i^k - \alpha \nabla_{u_i} J_i(u_i^k, u_{-i}^k), \forall i \in I$.

Dynamics

- $\dot{x} = Ax$ is asymptotically stable iff eigenvalues of A are in the open left half plane, $Re(\lambda) < 0$.
- $x^{k+1} = Ax^k$ is asymptotically stable iff $|\lambda| < 1$.

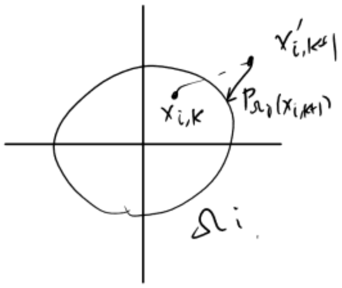
Example (2 player quadratic game)

- $P_1: J_1(u_1, u_2) = 2u_1^2 - 2u_1 - u_1 u_2$.
  - $\nabla_{u_1}^2 J_1 = 4$, J_1 is convex w.r.t. u_1.
- $P_2: J_2(u_1, u_2) = u_2^2 - \frac{1}{2} u_2 - u_1 u_2$.
  - $\nabla_{u_2}^2 J_2 = 2$, J_2 is convex w.r.t. u_2.
- $\Omega_1 = \Omega_2 = \mathbb{R}$.
- J_1, J_2 are radially unbounded, so there is a NE.
- Gradient play
  - $\nabla_{u_1} J_1 = 4u_1 - 2 - u_2, \nabla_{u_2} J_2 = 2u_2 - \frac{1}{2} - u_1$.
  - $F(u) = \begin{pmatrix} 4u_1 - 2 - u_2 \\ 2u_2 - \frac{1}{2} - u_1 \end{pmatrix}$.
  - Note: for a NE, $u^* \in int(\Omega)$, $F(u^*) = 0 \Rightarrow \begin{cases} 4u_1^* - 2 - u_2^* = 0 \\ 2u_2^* - \frac{1}{2} - u_1^* = 0 \end{cases}$.
  - $\begin{cases} u_1^{k+1} = u_1^k - \alpha(4u_1^k - 2 - u_2^k) \\ u_2^{k+1} = u_2^k - \alpha\left(2u_2^k - \frac{1}{2} - u_1^k\right) \end{cases}, \alpha \in (0,1)$.
  - Consider the limit point, $\begin{cases} \overline{u_1} = \overline{u_1} - \alpha(4\overline{u_1} - 2 - \overline{u_2}) \\ \overline{u_2} = \overline{u_2} - \alpha\left(2\overline{u_2} - \frac{1}{2} - \overline{u_1}\right) \end{cases} \Rightarrow \begin{cases} \overline{u_1} = u_1^* \\ \overline{u_2} = u_2^* \end{cases}$.
  - Check the convergence

- Let $\begin{cases}\widetilde{u_1^k} = \overline{u_1} - u_1^* \\ \widetilde{u_2^k} = \overline{u_2} - u_2^*\end{cases}$, then $\begin{cases}\widetilde{u_1^{k+1}} = (1-4\alpha)\widetilde{u_1^k} + \alpha\widetilde{u_2^k} \\ \widetilde{u_2^{k+1}} = \alpha\widetilde{u_1^k} + (1-2\alpha)\widetilde{u_2^k}\end{cases}$.
- $\widetilde{u^{k+1}} = \begin{pmatrix}1-4\alpha & \alpha \\ \alpha & 1-2\alpha\end{pmatrix}\widetilde{u^k}$.
- $A = \begin{pmatrix}1-4\alpha & \alpha \\ \alpha & 1-2\alpha\end{pmatrix}$, $|\lambda_A| \le 1$ gives condition on α.

Projected gradient play

- Algorithm to compute NE
- Each agent does a projected gradient descent
- $x_{i,k+1} = P_{\Omega_i}\left(x_{i,k} - \alpha_k \nabla_i J_i(x_i, x_{-i})\right)$.
- If agent i's action $x_i \in \Omega_i$.



- For all agents, $x_{k+1} = P_\Omega\left(x_k - \alpha_k F(x_k)\right)$, where $F(x) = \begin{pmatrix}\vdots \\ \nabla J_i(x_i, x_{-i}) \\ \vdots\end{pmatrix}$.
  - $P_\Omega(x) = \arg\min_{\omega\in\Omega}\|x-\omega\|^2$.
  - $\min_{\omega\in\Omega} f(x) = \frac{1}{2}\|\omega - x\|^2$, with $\nabla f(\omega) = \omega - x$.
  - For $x^* = \arg\min f(x)$, $\nabla f(x^*)^T(y - x^*) \ge 0, \forall y \in \Omega$.
  - Namely $\nabla f(\omega^*)^T(y-\omega^*) \ge 0, \forall y \in \Omega$.
  - $(\omega^* - x)^T(y - \omega^*) \ge 0, \forall y \in \Omega$.
  - $\left(P_\Omega(x) - x\right)^T\left(y - P_\Omega(x)\right) \ge 0, \forall y \in \Omega$.
  - $\left(P_\Omega(x) - x\right)^T\left(P_\Omega(y) - P_\Omega(x)\right) \ge 0$, since $P_\Omega(y) = y$.
  - $\left(P_\Omega(y) - y\right)^T\left(P_\Omega(y) - P_\Omega(x)\right) \ge 0$.
  - $(x-y)^T\left(P_\Omega(y) - P_\Omega(x)\right) \ge \|P_\Omega(x) - P_\Omega(y)\|^2$.
  - $\|x - y\| \ge \|P_\Omega(x) - P_\Omega(y)\|$ (no expansive).
- Assumption
  - $F(x)$ is strongly monotone, $\left(F(x) - F(y)\right)^T(x-y) \ge \mu\|x-y\|^2$, $\mu > 0$.
  - $F(x)$ is Lipschitz continuous, $\|F(x) - F(y)\| \le L\|x-y\|$, $L \ge 0$.
  - So $\mu \le \frac{|F(x)-F(y)|}{|x-y|} \le L$.
- Convergence (As $k \to \infty$, what happenes to the distance between x_{k+1} and x^*):
  - $\|x_{k+1} - x^*\|^2 = \left\|P_\Omega\left(x_k - \alpha_k F(x_k)\right) - P_\Omega(x^* - \alpha_k F(x^*))\right\|^2$.
  - $\le \|x_k - \alpha_k F(x_k) - x^* + \alpha_k F(x^*)\|^2$.
  - $= \left\|(x_k - x^*) - \alpha_k\left(F(x_k) - F(x^*)\right)\right\|^2$.
  - $= \|x_k - x^*\|^2 + \alpha_k^2\|F(x_k) - F(x^*)\|^2 - 2\alpha_k(x_k - x^*)^T\left(F(x_k) - F(x^*)\right)$.
  - $\le \|x_k - x^*\|^2 + {\alpha_k}^2 L^2\|x_k - x^*\|^2 - 2\alpha_k\mu\|x_k - x^*\|^2$.
  - $= (1 + \alpha_k^2 L^2 - 2\alpha_k\mu)\|x_k - x^*\|^2$.
  - Assume $\alpha_k = \alpha$.
  - $\|x_{k+1} - x^*\|^2 \le (1 + \alpha^2 L^2 - 2\alpha\mu)^{k+1}\|x_0 - x^*\|^2$.
  - If we want $\|x_{k+1} - x^*\|^2 \to 0$ as $k \to \infty$, we need $|1 + \alpha^2 L^2 - 2\alpha\mu| < 1$, $\alpha \in \left(0, \frac{2\mu}{L^2}\right)$.

- Then the algorithm converges.
- When $\alpha = \frac{\mu}{L^2}$, minimizes $1 + \alpha^2 L^2 - 2\alpha\mu$, fastest convergence.

- If $F(x)$ is merely monotone, $\mu = 0$, we might not get convergence.
  - $F(x) = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$, $\Omega = \mathbb{R}^2$.
  - $x_{k+1} = x_k - \alpha \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} x_k = \begin{pmatrix} 1 & -\alpha \\ \alpha & 1 \end{pmatrix} x_k$.
  - Eigenvalues are $\{1 \pm \alpha_i\}$, $\|1 + \alpha_i\| \geq 1$, doesn't converge.
  - But it may converge in $\Omega = \left\{x : \left\|x - \begin{pmatrix} 0.5 \\ 0 \end{pmatrix}\right\| \leq 0.5\right\}$.

Continuous time nonlinear systems

November 15, 2022 7:08 PM

Let $\dot{x} = f(x)$,

- $x \in \mathbb{R}^n$ is the state vector.
- f is Lipschitz continuous (differentiable).
- Then $\dot{x} = f(x)$ has unique solution $x(t) \in \mathbb{R}^n$ for any given initial condition $x(0) = x_0$.
- Trajectory: $\phi(t, x_0) = x(t)$.
  - $f(x)$ is tangent to the trajectory.

Equilibrium

- An equilibrium of $\dot{x} = f(x)$ is a set $\{x_{eq} : f(x_{eq}) = 0\}$.
- If $x_0 = x_{eq}$, then $x(t) = x_{eq}$.
- If $x(0) \neq x_{eq}$.
  - Stable: x_{eq} is a stable equilibrium if when starting sufficiently close from x_{eq}, the trajectory stays arbitrarily close to x_{eq}
    - $\forall \epsilon > 0, \exists \delta > 0$ s.t. $\forall x(0) \in B_\delta(x_{eq}), x(t) \in B_\epsilon(x_{eq})$.
  - Asymptotically stable: if in addition, $x(t) \to x_{eq}$ as $t \to \infty$.
  - Unstable: if x_{eq} is not stable.
    - $\exists \epsilon > 0, \forall \delta > 0, x(0) \in B_\delta(x_{eq})$, exists T finite s.t. $x(t) \notin B_\epsilon(x_{eq}), \forall t \geq T$.

Linearization method for testing stability

- Compute the Jacobian matrix $A_L = Df(x)\big|_{x=x_{eq}} = \frac{\partial f}{\partial x}\Big|_{x=x_{eq}} \in \mathbb{R}^{n\times n}$.
- Let $z = x - x_{eq}$,
- $\dot{z} = \dot{x} = f(x) \approx f(x_{eq}) + \frac{\partial f}{\partial x}(x_{eq})(x - x_{eq}) + \cdots = f(x_{eq}) + A_L z$.
- So the linearization of $\dot{x} = f(x)$ around x_{eq} is $\dot{z} = A_L z$.
  - Continuous time linear system
  - Equilibrium point is at $z = 0$.
- Stability of $\dot{z} = A_L z$ based on $eig(A_L)$
  - If $\forall i$, $Re(\lambda_i) < 0$, then $z = 0$ is stable.
  - If $\exists i$, $Re(\lambda_i) > 0$, then $z = 0$ is unstable.
  - If $Re(\lambda_i) = 0$, then can be stable (oscillating) or unstable.
- Hartmon-Grobman theorem: if all $eig(A_L)$ have $Re(\lambda_i) \neq 0$, then any stability/unstability of $z = 0$ for $\dot{z} = A_L z$ is equivalent to any stability/unstability of x_{eq} for $\dot{x} = f(x)$.
  - Note: if $Re(\lambda_i) = 0$, we cannot say anything based on linearization.

e.g. $\dot{x} = ax^3$, $a \neq 0$, $a \in \mathbb{R}$ a parameter.

- $f(x) = ax^3 = 0$ gives $x_{eq} = 0$.
- Let $z = x$.
- $A_L = 3ax^2\big|_{x=0} = 0$, so $\dot{z} = 0$, $z(t) = z(0) =$const.
- Linearization fails
- Consider $V(x) = \frac{1}{2}x^2$, $V(x) \geq 0, \forall x, V(0) = 0$.
- $\dot{V} = \frac{dV}{dx}\frac{dx}{dt} = ax^4$.
  - If $a < 0$, $\dot{V}(x) < 0\ \forall x$, $\dot{V}(0) = 0$, V is strictly decreasing in time along trajectory of $\dot{x} = f(x)$ towards 0. Hence, $x(t) \to 0$ as $t \to \infty$, $x = 0$ is stable.
  - If $a > 0$, $\dot{V}(x) > 0\ \forall x$. V is strictly increasing in time, so $x = 0$ is unstable.

Lyapunov Theorem

- Let $V : \mathbb{R}^n \to \mathbb{R}$ be C^1 (continuously differentiable), such that the following holds, then x_{eq} is an asymptotically stable equilibrium for $\dot{x} = f(x)$.

    - V is positive definite at x_{eq}, i.e. $V(x) > 0, \forall x \neq x_{eq}$ and $V(x_{eq}) = 0$.
    - $\dot{V}$ is negative definite at x_{eq}, i.e. $\dot{V}(x) < 0, \forall x \neq x_{eq}$ and $\dot{V}(x_{eq}) = 0$.
- If $\dot{V}$ is negative semidefinite, i.e. $\dot{V}(x) \leq 0\ \forall x \neq x_{eq}$ at x_{eq}, then x_{eq} is a stable equilirbium for $\dot{x} = f(x)$.
- If $\dot{V}(x)$ is positive definite, i.e. $\dot{V}(x) > 0\ \forall x \neq x_{eq}$, then x_{eq} is unstable.
- Potential choices for V
    - $V(x) = \frac{1}{2}\left\|x - x_{eq}\right\|^2$.
    - $V(x) = \frac{1}{2}\left(x - x_{eq}\right)^T M\left(x - x_{eq}\right)$ with $M \geq 0$.

Example: continuous time gradient play

- $\dot{x} = -f(x)$.
    - $f(x) = \begin{pmatrix} \vdots \\ \frac{\partial J_i}{\partial x_i}(x_i, x_{-i}) \\ \vdots \end{pmatrix}$ is the partial gradient.
    - $x^* = 0$ is equivalent to $f(x^*) = 0$.
- Consider $J_1 = J_2 = \cdots = J_N = P$, i.e. $f(x)$ is the true gradient.
    - $f(x) = \nabla P(x)$ is a potential game.
- Assume $P(x)$ is strictly convex, let $V(x) = P(x) - P(x^*)$.
    - For x^*, $\nabla P(x^*) = 0$, $P(x^*)$ is the minimum.
    - $V(x) = P(x) - P(x^*) > 0, \forall x \neq x^*$, V is positive definite at x^*.
    - $\dot{V}(x) = \left(\nabla V(x)\right)^T\left(-\nabla P(x)\right) = -\nabla P(x)^T \nabla P(x) = -\|\nabla P(x)\|^2 \leq 0$, $\dot{V}$ is negative semidefinite, $\dot{V}(x^*) = -\|\nabla P(x^*)\|^2 = 0$.
- By Lyapunov theorem, x^* is a stable equilibrium.
- $\dot{V}(x) = 0 \Leftrightarrow \|\nabla P(x)\| = 0 \Leftrightarrow \nabla P(x) = 0 \Leftrightarrow x = x^*$.
- Here x^* is asymptotically stable.

Population games (Evolutionary games)

November 2, 2022 8:29 AM

A large population of agents with a finite number of strategies, $M = \{1, \dots, j, \dots, m\}$.
Let x_j =fraction (population) of agents that use strategy $j \in M$ (frequency)

- $\forall j \in M, x_j \geq 0, \sum_{j=1}^{m} x_j = 1$.
- Let $x = \begin{pmatrix} x_1 \\ \vdots \\ x_m \end{pmatrix}$, $x \in \Delta$ (simplex).
- This is similar to mixed strategy in finite N-player games, but here it is the distribution of strategies in the population of agents, call it population state.
- Note: $x = \begin{pmatrix} 1 \\ \vdots \\ 0 \end{pmatrix}$ means all agents use strategy 1.

Canonical example of population games

- Agents are paired randomly to play a (symmetric) matrix game
  - i.e. if an agent uses jth strategy (jth strategist) is paired with a kth strategist, then the payoff/cost is $A = [a_{jk}]$.
- Assume: the success of a strategy depends on how many are using it
- Let $J_j(x)$ denote the expected cost of a jth strategist when in population state x.
  - $J_j(x) = \sum_{k=1}^{m} x_k a_{jk} = (Ax)_{jth\ row}$.
  - $J(x) = \begin{pmatrix} J_1(x) \\ \vdots \\ J_m(x) \end{pmatrix} = Ax$ is the vector of expected cost of all strategies when in population state x.
  - $J(x)$ is linear in x.

Nash equilibrium

- NE state is $x^* \in \Delta$ s.t. $(x^*)J(x^*) \leq y^T J(x^*) = \sum_j y_j J_j(x^*)$, $\forall y \in \Delta$.
  - Equivalently, $(x^*)Ax^* \leq y^T Ax^*$, $\forall y \in \Delta$.
- Recall, for 2 player finite game (A, B), $NE = (x^*, y^*)$.
  - For symmetric game, $B = A^T$, (A, A^T).
  - If we have $y^* = x^*$, the game is symmetric (x^*, x^*).

Evolutionary stable state (ESS)

- A population state $x^* \in \Delta$ is an ESS if
  - ESS-1: $(x^*)J(x^*) \leq y^T J(x^*)$, $\forall y \in \Delta$ (NE).
  - ESS-2: if $(x^*)J(x^*) = y^T J(x^*)$, then $(x^*)J(y) < y^T J(y)$ (refinement).
    - y is an alternative best response to x^*.
    - $y^T J(y)$ is the average cost of the population when in state y.
- Note: ESS-1 and ESS-2 are equivalent to $\exists 0 < \epsilon_y < 1$, $\forall 0 < \epsilon < \epsilon_y$, $(x^*)^T J(w) \leq y^T J(w)$, where $w = (1-\epsilon)x^* + \epsilon y = x^* + \epsilon(y - x^*)$, $\forall y \in x$, i.e. x^* is robust to invasion (perturbation).
- Note: if x^* is a strict NE, $(x^*)^T J(x^*) < y^T J(x^*)$, then x^* is ESS.

Revision protocols and mean dynamics ($\dot{x}$)

- Assume agents can revise the strategies they use and switch to another one i.e. from jth strategy to kth strategy
- Assume:
  - Agents have an internal clock and revision instances follow Poisson distribution with rate R

- i.e. over $[0, t]$, the mean number of revisions is Rt.
  - At a particular revision instance an agent switches from jth to kth strategy with a conditional switch rate p_{jk}, with condition $\max_x \sum_{j,k} p_{jk}(x) \leq R$ (the probability of switching is proportional to $p_{jk}(x)$).
  - Revision of all agents are independent.
- Look at rate of change in fraction (frequency) that uses jth strategy $\dot{x_j} = \frac{dx_j}{dt}$.
  - $\forall j \in M$, $\dot{x}$ is the mean dynamics.
- Consider dt interval of time, mean number of revisions is Rdt, and the mean number of revisions for jth strategist is $x_j Rdt$.
  - Expected number of switches from jth to kth strategy is $\frac{p_{jk}}{R} x_j Rdt = p_{jk} x_j dt$
  - Expected number of switches from kth to jth is $p_{jk} x_k dt$.
  - Expected change in x_j over dt, denoted dx_j is $dx_j = \sum_{k=1}^{m} p_{kj} x_k dt - \sum_{k=1}^{m} p_{jk} x_j dt$.
  - $\dot{x_j} = \sum_{k=1}^{m} x_k p_{kj} - x_j \sum_{k=1}^{m} p_{jk}$, $\forall j \in [m]$.
  - $\dot{x} = f(x) = \begin{pmatrix} \dot{x_1} \\ \vdots \\ \dot{x_m} \end{pmatrix}$ is the mean dynamics is a set of nolinear ODEs.
- Different revision protocols $\Rightarrow$ different p_{jk} $\Rightarrow$ different mean dynamics with different properties and stable points

E.g. pairwise imitation protocol

- $p_{jk} = x_k \left[J_j(x) - J_k(x)\right]_+$ where $[v]_+ = \begin{cases} v, v \geq 0 \\ 0, v < 0 \end{cases}$.
  - If $J_k < J_j$, switch to k, proportional to the difference.
  - If more people are using k (x_k is high), higher chance of switching.
- Mean dynamics
  - $\dot{x_j} = \sum_{k=1}^{m} x_k x_j \left[J_k(x) - J_j(x)\right]_+ - x_j \sum_{k=1}^{m} x_k \left[J_j(x) - J_k(x)\right]_+$,
  - $= x_j \sum_{k=1}^{m} x_k \left(J_k(x) - J_j(x)\right)$,
  - $= x_j \left(\sum_{k=1}^{m} x_k J_k(x) - J_j(x) \sum_{k=1}^{m} x_k\right)$,
  - $= x_j \left(x^T J(x) - J_j(x)\right)$, $\forall j \in [m]$. (replicator dynamics/RD)

Replicator Dynamics (RD)

- The mean dynamic for pairwise-imitation protocol in a population game
- $\dot{x_j} = x_j\left[x^T J(x) - J_j(x)\right]$, $\forall j \in [m]$.
- If $J(x) = Ax$ (a symetric matrix game $A = A^T$), $\dot{x_j} = x_j\left[x^T Ax - (Ax)_{jth\ row}\right]$, $\dot{x} = f(x)$ is nonlinear.
- For $A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$, $x = (x_1, x_2)^T$, $x_1 + x_2 = 1$.
  - $\dot{x_1} = -(a_1 x_1 - a_2 x_2) x_1 x_2$. (combine $a_{11} x_1^2 + a_{21} x_1 x_2$ and $a_{22} x_2^2 + a_{12} x_1 x_2$)
  - $\dot{x_2} = (a_1 x_1 - a_2 x_2) x_1 x_2$. (since $\dot{x_2} = -\dot{x_1}$)
  - Where $a_1 = a_{11} - a_{21}$, $a_2 = a_{22} - a_{12}$.
  - Note: $\dot{x_2} = -\dot{x_1}$, $x_2 = 1 - x_1$.

e.g. $A = \begin{pmatrix} 5 & 0 \\ 15 & 1 \end{pmatrix}$, $x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$, $\Delta = \{x \in \mathbb{R}^2 : x_1 + x_2 = 1, x_1 \geq 0, x_2 \geq 0\} = \{x \in \mathbb{R}^2 : x_{1,2} \geq 0, 1^T x - 1 = 0\}$.

- RD: $\dot{x} = f(x) = \begin{pmatrix} \dot{x_1} \\ \dot{x_2} \end{pmatrix} = \begin{pmatrix} x_1 x_2 (10 x_1 + x_2) \\ -x_1 x_2 (10 x_1 + x_2) \end{pmatrix}$.
- Eq points: $x_1 = 0$ or $x_2 = 0$ or $10 x_1 + x_2 = 0$.
  - $x_{eq}^1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$, $x_{eq}^2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$.
- Tangent set to Δ: $T_\Delta = \{y \in \mathbb{R}^2 : 1^T y = 0\}$.
  - Same for higher dimensions. $\Delta = \{x \in \mathbb{R}^3 : x_{1,2,3} \geq 0, 1^T x - 1 = 0\}$, $T_\Delta = \{y \in$

- Since $1^T f(x) = 0$, $f(x) \in T_\Delta$.
  - Thus, $\frac{d}{dt}(x_1 + x_2) = 0$, $x_1 + x_2 = const$, $\forall t$.
- Since $x(0) \in \Delta$, $x_1(0) + x_2(0) = 1$, then $x_1(t) + x_2(t) = 1$, $x(t) \in \Delta$, $\forall t$.
  - Δ is invariant in time under RD.
- Use linearization method to check stability of $x_{eq}^1 = e_1$, $x_{eq}^2 = e_2$.
  - $A_L = \frac{\partial f}{\partial x} = \begin{pmatrix} \frac{\partial f_1}{\partial x_1} & \frac{\partial f_1}{\partial x_2} \\ \frac{\partial f_2}{\partial x_1} & \frac{\partial f_2}{\partial x_2} \end{pmatrix} = \begin{pmatrix} 20x_1x_2 + x_2^2 & 2x_1x_2 + 10x_1^2 \\ -20x_1x_2 - x_2^2 & -2x_1x_2 - 10x_1^2 \end{pmatrix}$.
  - We only need to consider the eigenvectors in the tangent set T_Δ, as that's the only direction we can move.
  - $A_L^1 = \frac{\partial f}{\partial x}\Big|_{x=e_1} = \begin{pmatrix} 0 & 10 \\ 0 & -10 \end{pmatrix}$.
    - $\dot{z} = A_L^1 z$ with $z = x - x_{eq}^1 \in T_\Delta$ ($1^T z = 1^T x - 1^T x_{eq}^1 = 1 - 1 = 0$).
    - $\lambda = \{0, -10\}$, $v_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \notin T_\Delta$, $v_2 = \begin{pmatrix} 1 \\ -1 \end{pmatrix} \in T_\Delta$, x_{eq}^1 is stable since -10 is in the $\{x : Re(x) < 0\}$.
  - $A_L^2 = \frac{\partial f}{\partial x}\Big|_{x=e_2} = \begin{pmatrix} 1 & 0 \\ -1 & 0 \end{pmatrix}$.
    - $\dot{z} = A_L^2 z$ with $z = x - x_{eq}^2 \in T_\Delta$.
    - $\lambda = \{0,1\}$, $v_1 = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \notin T_\Delta$, $v_2 = \begin{pmatrix} 1 \\ -1 \end{pmatrix} \in T_\Delta$, x_{eq}^2 is unstable since 1 is in the $\{x : Re(x) > 0\}$.
- Recall for a PD game NE=(e_1, e_1) (confess, confess), $x_{NE}^* = e_1$.
  - RD finds the NE, no matter from what initial condition.

General properties of the RD

- Δ is invariant under the RD, $\forall x(0) \in \Delta, x(t) \in \Delta, \forall t \in \mathbb{R}$.
- $\{e_1, \dots, e_j, \dots, e_m\}$ (vertices in $\mathbb{R}^m$) are equilibria points of RD.
  - If x^* is NE state ($(x - x^*)^T J(x^*) \geq 0$, $\forall x \in \Delta$), then x^* is an eq point of RD.
  - Note: it is not true that any eq point of RD is an NE (x^*).
- If x_{eq} is asymptotically stable eq point of RD, then x_{eq} is NE (proof by contradiction).
  - Note: not true that any NE is an asymptotically stable eq point of RD.
  - However, there are special classes of population games that the reverse is also true
    - Potential population games $J(x) = \nabla P(x)$, true gradients of potential function ($A = A^T$ for matrix game).
    - Strictly stable games:
      - $(x - y)^T (J(x) - J(y)) > 0$, $\forall x \neq y \in \Delta$.
      - If matrix games, $(x - y)^T (A + A^T)(x - y) > 0$, $\forall x \neq y$ with $A + A^T > 0$ on T_Δ.
- If x^* is ESS, then it is an asymptotically stable eq of RD.
  - For the PD example above, $x^* = e_1$ is ESS.
  - Note both 4 and reverse of 3 in strictly stable games can be shown using Lyapunov method.

Lyapunov function candidates for RD

- Quadratic form: $V(x) = (x - x_{eq})^T P (x - x_{eq})$.
- $V(x)$ is PD at x_{eq}, so P is a PD matrix.
- $\dot{V}(x) = \nabla V(x)^T f(x) < 0$, $\forall x \neq x_{eq}$.
  - $\dot{V}(x) = 2(x - x_{eq})^T P f(x)$, doesn't quite work for RD.
- An appropriate Lyapunov function is the relative entropy: $V(x) = \sum_{j \in \text{supp}\, x^*} x_j^* \ln\left(\frac{x_j^*}{x_j}\right)$.
  - $V(x)$ is PD, $\forall x \neq x^*$, using Jenson's inequality.

- For RD, $\dot{V}(x) = \nabla V(x)^T f(x) = \sum_{j \in \operatorname{supp} x^*} \frac{\partial V}{\partial x_j} \frac{dx_j}{dt} = \sum_j -x_j^* \frac{1}{x_j} x_j \left(x^T J(x) - J_j(x) \right)$,
  $= -\sum_j x_j^* \left(x^T J(x) \right) + \sum_j x_j^* J_j(x) = -x^T J(x) + (x^*)^T J(x) = (x^* - x)^T J(x)$.
- For NE, $(x^* - x)^T J(x^*) \leq 0$.
- $\dot{V} = (x^* - x)^T J(x) = (x^* - x)^T \left(J(x) + J(x^*) - J(x^*) \right) = (x^* - x)^T J(x^*) + (x^* - x)^T \left(J(x) - J(x^*) \right)$.
- For strictly stable games, $(x^* - x)^T \left(J(x) - J(x^*) \right) < 0$, thus $\dot{V} < 0$.

Rock paper scissor game

- $x^* = \left(\frac{1}{3}, \frac{1}{3}, \frac{1}{3} \right)$ no matter what formation.
- Standard: $A = \begin{pmatrix} 0 & 1 & -1 \\ -1 & 0 & 1 \\ 1 & -1 & 0 \end{pmatrix}$.
  - $A + A^T = 0$, not a strictly stable game, but a stable game.
- Modified: $= \begin{pmatrix} 0 & 0 & -1 \\ -1 & 0 & 0 \\ 0 & -1 & 0 \end{pmatrix}$.
  - $A + A^T > 0$, strictly stable, RD converges to x^*.

Learning in games

November 29, 2022 6:14 PM

2 types of algo for N-player repeated finite action game

- BR-play and variants
  - $x_i(k+1) = \widetilde{BR}_i\big(x_{-i}(k)\big)$.
  - $x_{-i}(k)$ is mixed strategies of all others
  - $\widetilde{BR}_i$ softmax of some cost functions J_i.
- Fictitious-play and variants
  - z_i: internal variable of P_i that is estimate of x_{-i}.
    - Updated by simply averaging history: $z_i^{k+1} = \frac{k}{k+1} z_i^k + \frac{1}{k+1} e_{-i}^k$.
  - $x_i(k) = \widetilde{BR}_i\big(z_i(k)\big)$.
  - Still require a cost function (structure info)

Relax the info that players have

- $\pi_i^k = -J_i\left(e_i^k, e_{-i}^k\right) \in \mathbb{R}$, received or realized payoff at iteration k of game.
- How to update internal variable z_i^k using this info?
- How to map internal variable into a mixed strategy x_i^k?

Erev-Roth algo (payoff based)

- z_i^k: a vector with #components = #actions of P_i=m_i.
- z_{ij}^k: a score variable for action j, $j \in [m_i]$.
- $e_i^k = e_j$ when P_i uses jth action.
- At iteration k:
  - If $e_i^k = e_j$, $z_{ij}^{k+1} = z_{ij}^k + \pi_i^k$.
  - Else, $z_{ij}^{k+1} = z_{ij}^k$.
  - For $j = 1, \dots, m_i$, $z_i^{k+1} = z_i^k + \pi_i^k e_i^k$.
  - Adds a corresponding payoff if using a strategy.
- To map z_i^k to $x_i^k \in \Delta_i$ ($x_{ij}^k \geq 0, \sum_j x_{ij}^k = 1$).
  - $x_{ij}^k = \frac{z_{ij}^k}{\sum_{j=1}^{m_i} z_{ij}^k}$ (assuming $z_{ij}^k \geq 0$).
  - $x_i^k = \frac{z_i^k}{\sum_{j=1}^{m_i} z_{ij}^k}$.
- The algorithm is two simple functions
  - $z_i^{k+1} = z_i^k + \pi_i^k e_i^k$.
  - $x_i^k = \frac{z_i^k}{\sum_{j=1}^{m_i} z_{ij}^k}$.
- The behavior of the stochastic algo can be analyzed based on its mean dynamics (deterministic CT set of ODEs) that have a function similar to RD.
- Similar convergence results can be obtained.