Empirical-evidence Equilibria in Stochastic Games

Nicolas Dudebout
Context

Multiagent problems
- stock market
- group of robots

Game-theoretic approach
- selfish agents
- different solution concepts
- predictive
- prescriptive
Empirical-evidence Equilibrium (EEE)

Motivation
Definition
Existence
Comparison
Characterization
Predictive Use
Prescriptive Use
Graphical convention

\[ x \rightarrow f \rightarrow y \]
\[ y^+ \sim f(x) \]

\[ x \rightarrow g \rightarrow y \]
\[ y^{t+1} \sim g(x^1, \ldots, x^t, y^1, \ldots, y^t) \]

\[ x \rightarrow h \rightarrow y \]
\[ y^+ \sim h(y, x) \]
Markov Decision Process (MDP)

\[
\max_{\sigma} \mathbb{E}_\sigma \left[ \sum_{t=0}^{\infty} \delta^t \cdot u(x^t, a^t) \right]
\]

\[
x^+ \sim f(x, a)
\]

\[
u(x, a), \delta
\]
Stochastic Game

\[ u_1(x_1, a_1, a_2), \delta_1 \]

\[ u_2(x_2, a_2, a_1), \delta_2 \]
Partially Observable Markov Decision Process (POMDP)

\[ u(x, a), \delta \]
\[ u_1(x_1, a_1, a_2), \delta_1 \quad \text{and} \quad u_2(x_2, a_2, a_1), \delta_2 \]

\[ u(x, a), \delta \]
Recap
• Multiagent problems
• Game-theoretic approach
• Nash equilibrium in stochastic game $\leftrightarrow$ unknown POMDPs

POMDP intractable

MDP solved
Stochastic Game

\[ u_1(x_1, a_1, a_2), \delta_1 \]

\[ u_2(x_2, a_2, a_1), \delta_2 \]
\[ u(x, a, s^+), \delta \]

\[
\begin{align*}
  w^+ &\sim n(w, x, a) \\
  s^+ &\sim \nu(w^+) 
\end{align*}
\]
Simple Consistency

$0100010001001010010110111010...$

$\mathbb{P}[0], \mathbb{P}[1]$

$\mu$  

$s$

$\mu[s] = \mathbb{P}[s]$

$u(x, a, s^+), \delta$

$\sigma$

$f$

$x$

$\alpha$

$\mu$

$s$
Two Systems

\[ u(x, a, s^+), \delta \]

Real System: \[ R \]

Mockup System: \[ M \]
Consistency

\[ \begin{align*}
\sigma & \xrightarrow{a} f \\
\text{Nature} & \xrightarrow{s} \\
\end{align*} \]

\[ \begin{align*}
\mu[s^+] &= \mathbb{P}[s^+] \\
&= \lim_{t \to \infty} \mathbb{P}[S^{t+1} = s^+] \\
&= \sum_{w^+, w, x, a} \nu(w^+) [s^+] \cdot \pi[w, x] \cdot \sigma(x)[a] \cdot n(w, x, a)[w^+] 
\end{align*} \]
Depth-1 Consistency

01010101010101010101010101...

\[ P[00], P[01], P[11], P[10] \iff P[0\mid0], P[1\mid0], P[0\mid1], P[1\mid1] \]
Two Systems

Real System: $\mathbf{R}$

Mockup System: $\mathbf{M}$

$\mathbb{P}[s^+ | z]$
Depth-k Consistency

\[
\begin{align*}
\mu(z)[s^+] &= \mathbb{P}[s^+ | z] \\
&= \lim_{t \to \infty} \mathbb{P}[S^{t+1} = s^+ | Z^t = z]
\end{align*}
\]
Recap
Start with one agent

Arbitrarily fix a model \( m^k \)

Split hard problem:

- Markov chain \( R \)
  \( \Rightarrow \) consistent predictor \( \mu \)

- MDP \( M \)
  \( \Rightarrow \) optimal strategy \( \sigma \)

EEEs are fixed points of:
\[ u_1(x_1, a_1, s_1^+), \delta_1 \]
Empirical-evidence Equilibrium

$m^{k_1}$ and $m^{k_2}$ fixed

$(\mu_1, \sigma_1, \mu_2, \sigma_2)$ is an $\varepsilon$ empirical-evidence equilibrium ($\varepsilon$ EEE) if:

- $\mu_1$ is consistent with $R$
- $\mu_2$ is consistent with $R$
- $\sigma_1$ is $\varepsilon$ optimal for $M_1$
- $\sigma_2$ is $\varepsilon$ optimal for $M_2$
EEE vs Nash

- optimization complexity fixed by agent not opponents
- always implementable
- each agent knows when at equilibrium
- less intrinsic to the problem
Existence of $\varepsilon$ EEEs

Theorem
For any $m^{k_1}$ and $m^{k_2}$, there exists an $\varepsilon$ EEE.

Proof.

- $\varepsilon$ and Gibbs distribution $\Rightarrow \mu_i \mapsto \sigma_i$ is a function
- $\mu \mapsto \mu$ is a continuous function
- set of predictors is compact and convex
- Brouwer’s fixed point theorem
Existence of EEEs

Theorem
For any $m^{k_1}$ and $m^{k_2}$, there exists a EEE.

Proof.
- $\mu \mapsto \mu$ is a closed-graph correspondence
- set of predictors is compact and convex
- Kakutani’s fixed point theorem
Theorem
Exogenous EEEs in perfect-monitoring repeated games yield correlated equilibria of the underlying one-shot game.

Repeated game:
Stochastic game without a state

Correlated equilibrium:
Nash equilibrium with common source of randomness
Recap
- multiagent EEE identical to single agent
- each agent arbitrarily picks a model \( m^k \)
- EEEs always exist
- EEEs induce correlated equilibria in repeated games
Asset Management Example

State  holdings  \( x_i \in [0, M] \)
Action sell one, hold, or buy one  \( a_i \in \{-1, 0, 1\} \)
Signal price  \( p \in \{\text{Low}, \text{High}\} \)
Dynamic  \( x_i^+ = x_i + a_i \)
Stage cost  \( p \cdot a_i \)
Nature market trend  \( w \in \{\text{Bull}, \text{Bear}\} \)
Model depth 0
Iterative Process

Update Rule \[ \mu_i^{r+1} = (1 - \alpha^r) \mu_i^r + \alpha^r (\tilde{\mu}_i - \mu_i^r) \]
Theoretical Predictor

Update Rule

\[ \mu_i^{r+1} = (1 - \alpha)\mu_i^r + \alpha(\tilde{\mu}_i - \mu_i^r) \]
\[ \mathbb{P}[\text{High}] \]

\[ \mu_1^r \quad \mu_2^r \]

Round \( r \)
Empirical Predictor

\[
\begin{align*}
\mu_i^{r+1} &= (1 - \alpha^r) \mu_i^r + \alpha^r (\tilde{\mu}_i^T - \mu_i^r) \\
\alpha^r &\text{ non-summable, square-summable}
\end{align*}
\]
Hawk-dove Game

Repeated game

<table>
<thead>
<tr>
<th></th>
<th>h</th>
<th>d</th>
</tr>
</thead>
<tbody>
<tr>
<td>H</td>
<td>-1, -1</td>
<td>6, 0</td>
</tr>
<tr>
<td>D</td>
<td>0, 6</td>
<td>3, 3</td>
</tr>
</tbody>
</table>

Nash equilibria (H, d) and (D, h)

Want correlated equilibrium alternating between the two
Hawk-dove Game

Depth-2 models

Strategies:
\[ \sigma_1(d, h) = 0.999 \text{H} + 0.001 \text{D} \]
\[ \sigma_1(h, d) = 0.999 \text{D} + 0.001 \text{H} \]
\[ \sigma_1(h, h) = 0.5 \text{H} + 0.5 \text{D} \]
\[ \sigma_1(d, d) = 0.5 \text{H} + 0.5 \text{D} \]

Associated predictors:
\[ \mu_1(d, h) = 0.996 d + 0.004 h \]
\[ \mu_1(h, d) = 0.996 h + 0.004 d \]
\[ \mu_1(h, h) = 0.5 h + 0.5 d \]
\[ \mu_1(d, d) = 0.5 h + 0.5 d \]

Strategy approximately optimal as \( \delta \) close enough to one

Generalizes to any convex combination of pure Nash equilibria
Recap
Predictive given models and adaptation rule a EEE emerges
Prescriptive implement desired outcome as a EEE
Extensions

• $n$ agents
• endogenous models $z^+ \sim m(z, x, a, s)$
• notions of consistency: approximate, weak, and eventual
• convergence of empirical iterative process when theoretical one converges
Empirical-evidence Equilibrium (EEE)

- **Motivation**: intractable problem
- **Definition**: split into Markov chain and consistent MDPs
- **Existence**: fixed-point theorems
- **Comparison**: lower computational requirements
- **Characterization**: correlated equilibrium in repeated game
- **Predictive Use**: model to understand stock price
- **Prescriptive Use**: desired outcome encoded as EEE