On the Convergence to a Mean-Field Equilibrium

Josu Doncel
Polaris group
INRIA

joint work with Nicolas Gast, Bruno Gaujal

Atelier en Évaluation de Performances
Toulouse

16/03/2016

Nash Equilibrium

The set of decisions such that no player has benefit of unilateral deviation

Theorem (Nash)

There exists a solution for every non-cooperative game.

Nash Equilibrium

The set of decisions such that no player has benefit of unilateral deviation

Theorem (Nash)

There exists a solution for every non-cooperative game.

Theorem (Papadimitriou et al)

Computing a Nash equilibrium is PPAD-complete.
(Polynomial Parity Arguments on Directed graphs)

$\Rightarrow$ Huge number of players or strategies: not tractable

Nash Equilibrium

The set of decisions such that no player has benefit of unilateral deviation

Theorem (Nash)

There exists a solution for every non-cooperative game.

Theorem (Papadimitriou et al)

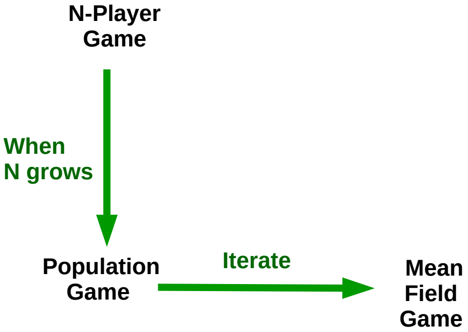
Computing a Nash equilibrium is PPAD-complete.
(Polynomial Parity Arguments on Directed graphs)

$\Rightarrow$ Huge number of players or strategies: not tractable

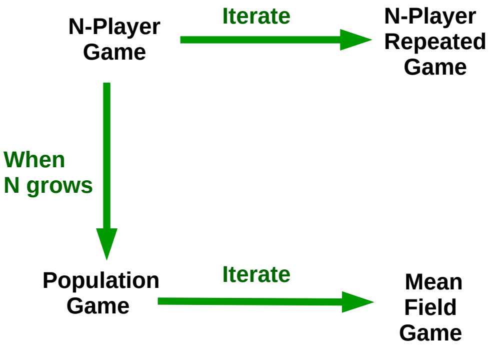
Mean-Field Games (Lions and Lasry)

Infinite number of rational objects in interaction.

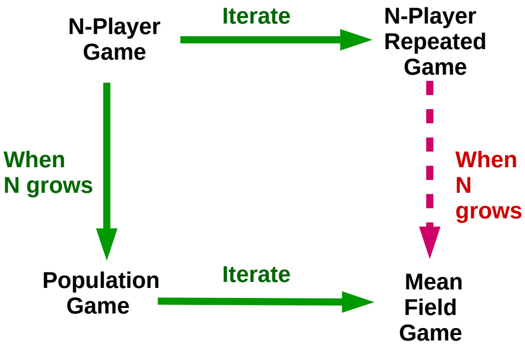
Our question



Our question



Our question



Outline

Outline

Discrete Time

Infinite Homogeneous Players

- State: $\mathcal{S} = \{1, \ldots, S\}$
- Actions: $\mathcal{A} = \{1, \ldots, A\}$

Infinite Homogeneous Players

- State: $\mathcal{S} = \{1, \ldots, S\}$
- Actions: $\mathcal{A} = \{1, \ldots, A\}$

- Population distribution: $\mathbf{m}^a(t)$, $a \in \mathcal{A}$

Discrete Time

Infinite Homogeneous Players

- State: $\mathcal{S} = \{1, \ldots, S\}$
- Actions: $\mathcal{A} = \{1, \ldots, A\}$

- Population distribution: $\mathbf{m}^a(t)$, $a \in \mathcal{A}$

Player 0

- $a^0 \in \mathcal{A}$
- Instan. cost: $c(a^0(t), \mathbf{m}^a(t))$

Discrete Time

Infinite Homogeneous Players

- State: $\mathcal{S} = \{1, \ldots, S\}$
- Actions: $\mathcal{A} = \{1, \ldots, A\}$

- Population distribution: $\mathbf{m}^a(t)$, $a \in \mathcal{A}$

Player 0

- $a^0 \in \mathcal{A}$
- Instan. cost: $c(a^0(t), \mathbf{m}^a(t))$

$$V(a^0, a) = \mathbb{E}\left[(1-\delta)\sum_{t=0}^{\infty} \delta^t \cdot c(a^0(t), \mathbf{m}^a(t))\right]$$

Mean-Field Equilibrium

Best Response to a

Set of strategies

$$BR(a) = \arg\min_{a^0} V(a^0, a)$$

Mean-Field Equilibrium

Best Response to a

Set of strategies

$$BR(a) = \arg\min_{a^0} V(a^0, a)$$

Mean-Field Equilibrium

$$a^{MFE} \in BR(a^{MFE})$$

$\Rightarrow$ Fixed-point problem

N-Player Repeated Game

- Player 0 $\Rightarrow a^0$
- Others $\Rightarrow a$

N-Player Repeated Game

- Player 0 $\Rightarrow a^0$
- Others $\Rightarrow a$

- $\mathbf{M}^a(t)$: population distribution at time t
- Instan. cost: $c(a^0, \mathbf{M}^a(t))$

$$V^N(a^0, a) = \mathbb{E}\left[(1-\delta)\sum_{t=0}^{\infty} \delta^t \cdot c(a^0, \mathbf{M}^a(t))\right]$$

N-Player Repeated Game

- Player 0 $\Rightarrow a^0$
- Others $\Rightarrow a$

- $\mathbf{M}^a(t)$: population distribution at time t
- Instan. cost: $c(a^0, \mathbf{M}^a(t))$

$$V^N(a^0, a) = \mathbb{E}\left[(1-\delta)\sum_{t=0}^{\infty} \delta^t \cdot c(a^0, \mathbf{M}^a(t))\right]$$

N-player game equilibrium

For all a'

$$V^N(a, a) \leq V^N(a', a)$$

Outline

MFE existence

Continuity assumptions on $\mathbf{m}$

- $P_{ija}(\mathbf{m})$
- $c(a^0, \mathbf{m})$

MFE existence

Continuity assumptions on $\mathbf{m}$

- $P_{ija}(\mathbf{m})$
- $c(a^0, \mathbf{m})$

Theorem

Any discrete time mean-field game with discounted cost that satisfies the continuity assumptions has a mean-field equilibrium.

Best-response has a fixed-point: Kakutani

Convergence

Every repeated game Nash equilibria converge to a MFE?

Convergence

Every repeated game Nash equilibria converge to a MFE?

Player Strategies of N-player repeated game

- Local: depend on state and time $\Rightarrow a(i, t)$
- Markov: depend on $\mathbf{m}$, state and time $\Rightarrow a(i, \mathbf{m}(t))$

Convergence

Every repeated game Nash equilibria converge to a MFE?

Player Strategies of N-player repeated game

- Local: depend on state and time $\Rightarrow a(i,t)$
- Markov: depend on $\mathbf{m}$, state and time $\Rightarrow a(i,\mathbf{m}(t))$

Theorem

Every local equilibrium converges to a mean-field equilibrium.

Proof: $V^N(\pi',\pi) \to V(\pi',\pi)$ when $N \to \infty$

(H. Tembine, J.-Y. L. Boudec, R. El-Azouzi, and E. Altman. Mean-field asymptotics of markov decision evolutionary games and teams. GameNets' 09.)

Prisoner's dilemma

$\mathcal{S} = \mathcal{A} = \{C, D\}$

	C	D
C	1,1	3,0
D	0,3	2,2

Prisoner's dilemma

$\mathcal{S} = \mathcal{A} = \{C, D\}$

	C	D
C	1,1	3,0
D	0,3	2,2

N Players: mean-field version

$\Rightarrow m_C + m_D = 1$

$$c(i, m) = \begin{cases} m_C + 3m_D & \text{if } i = C \\ 2m_D & \text{if } i = D \end{cases}$$

Prisoner's dilemma

$\mathcal{S} = \mathcal{A} = \{C, D\}$

	C	D
C	1,1	3,0
D	0,3	2,2

N Players: mean-field version
$\Rightarrow m_C + m_D = 1$

$$c(i, m) = \begin{cases} m_C + 3m_D & \text{if } i = C \\ 2m_D & \text{if } i = D \end{cases}$$

Static game equilibrium $\Rightarrow$ Always D

- MFE?
- Repeated Game NE?

Mean-Field Equilibrium

$$c(i,m) = \begin{cases} m_C + 3m_D & \text{if } i = C \\ 2m_D & \text{if } i = D \end{cases}$$

$a^D = BR(a)$

The best response to any action a is always D.

$$V(i,m) = (1-\delta)\sum_{t=0}^{\infty} \delta^t \left(x_C(t)\cdot(m_C(t)+3m_D(t)) + x_D(t)\cdot 2m_D(t)\right)$$

Mean-Field Equilibrium

$$c(i,m) = \begin{cases} m_C + 3m_D & \text{if } i = C \\ 2m_D & \text{if } i = D \end{cases}$$

$a^D = BR(a)$

The best response to any action a is always D.

$$\begin{aligned} V(i,m) &= (1-\delta)\sum_{t=0}^{\infty} \delta^t \left(x_C(t)\cdot(m_C(t)+3m_D(t)) + x_D(t)\cdot 2m_D(t)\right) \\ &= (1-\delta)\sum_{t=0}^{\infty} \delta^t (x_C(t) + 2m_D(t)) \end{aligned}$$

Mean-Field Equilibrium

$$c(i,m) = \begin{cases} m_C + 3m_D & \text{if } i = C \\ 2m_D & \text{if } i = D \end{cases}$$

$a^D = BR(a)$

The best response to any action a is always D.

$$\begin{aligned} V(i,m) &= (1-\delta)\sum_{t=0}^{\infty} \delta^t \left(x_C(t)\cdot(m_C(t)+3m_D(t)) + x_D(t)\cdot 2m_D(t)\right) \\ &= (1-\delta)\sum_{t=0}^{\infty} \delta^t (x_C(t) + 2m_D(t)) \end{aligned}$$

Minimum when $x_C(t) = 0$

Mean-Field Equilibrium

$$c(i,m) = \begin{cases} m_C + 3m_D & \text{if } i = C \\ 2m_D & \text{if } i = D \end{cases}$$

$a^D = BR(a)$

The best response to any action a is always D.

$$\begin{aligned} V(i,m) &= (1-\delta)\sum_{t=0}^{\infty} \delta^t \left(x_C(t)\cdot(m_C(t)+3m_D(t)) + x_D(t)\cdot 2m_D(t)\right) \\ &= (1-\delta)\sum_{t=0}^{\infty} \delta^t (x_C(t) + 2m_D(t)) \end{aligned}$$

Minimum when $x_C(t) = 0 \Rightarrow$ BR(a): Always D

Mean-Field Equilibrium

$$c(i,m) = \begin{cases} m_C + 3m_D & \text{if } i = C \\ 2m_D & \text{if } i = D \end{cases}$$

$a^D = BR(a)$

The best response to any action *a* is always D.

$$\begin{aligned} V(i,m) &= (1-\delta)\sum_{t=0}^{\infty} \delta^t \left(x_C(t)\cdot(m_C(t)+3m_D(t)) + x_D(t)\cdot 2m_D(t)\right) \\ &= (1-\delta)\sum_{t=0}^{\infty} \delta^t (x_C(t) + 2m_D(t)) \end{aligned}$$

Minimum when $x_C(t) = 0 \Rightarrow$ BR(a): Always D

MFE: a^D

$$a^D \in BR(a^D)$$

Repeated Game Nash Equilibrium

$$c(i, m^a) = \begin{cases} m_C + 3m_D & \text{if } i = C \\ 2m_D & \text{if } i = D \end{cases}$$

Action a^*

- First k rounds, play D
- From k, play C if all play C
  Otherwise, play D for ever

Repeated Game Nash Equilibrium

$$c(i, m^a) = \begin{cases} m_C + 3m_D & \text{if } i = C \\ 2m_D & \text{if } i = D \end{cases}$$

Action a^*

- First k rounds, play D
- From k, play C if all play C
  Otherwise, play D for ever

Nash Equilibrium

No player gets benefit from unilateral deviation

Repeated Game Nash Equilibrium

$$c(i, m^a) = \begin{cases} m_C + 3m_D & \text{if } i = C \\ 2m_D & \text{if } i = D \end{cases}$$

Action a^*

- First k rounds, play D
- From k, play C if all play C
  Otherwise, play D for ever

Nash Equilibrium

No player gets benefit from unilateral deviation

- First k rounds, cost increases

Repeated Game Nash Equilibrium

$$c(i, m^a) = \begin{cases} m_C + 3m_D & \text{if } i = C \\ 2m_D & \text{if } i = D \end{cases}$$

Action a^*

- First k rounds, play D
- From k, play C if all play C
  Otherwise, play D for ever

Nash Equilibrium

No player gets benefit from unilateral deviation

- First k rounds, cost increases
- From k, if play D, an immediate advantage $\Rightarrow$ punished until the end

What happened?

Theorem (Folk Theorem)

Let G be a symmetric matricial game, and let V^* be the cost under the strategy that repeats the Nash equilibrium of the static game G. Then any feasible cost V smaller than V^* is the cost of an equilibrium of the repeated game if the discount factor is large enough.

Reward and Punishment Principle

What happened?

Theorem (Folk Theorem)

Let G be a symmetric matricial game, and let V^* be the cost under the strategy that repeats the Nash equilibrium of the static game G. Then any feasible cost V smaller than V^* is the cost of an equilibrium of the repeated game if the discount factor is large enough.

Reward and Punishment Principle

$\Rightarrow$ None of these equilibria convergence to a mean-field equilibrium

Why?

What happened?

Theorem (Folk Theorem)

Let G be a symmetric matricial game, and let V^* be the cost under the strategy that repeats the Nash equilibrium of the static game G. Then any feasible cost V smaller than V^* is the cost of an equilibrium of the repeated game if the discount factor is large enough.

Reward and Punishment Principle

$\Rightarrow$ None of these equilibria convergence to a mean-field equilibrium

Why?

MFG assumption: Individuals action does not influence the mass

- The mass can not punish

What happened?

Theorem (Folk Theorem)

Let G be a symmetric matricial game, and let V^* be the cost under the strategy that repeats the Nash equilibrium of the static game G. Then any feasible cost V smaller than V^* is the cost of an equilibrium of the repeated game if the discount factor is large enough.

Reward and Punishment Principle

$\Rightarrow$ None of these equilibria convergence to a mean-field equilibrium

Why?

MFG assumption: Individuals action does not influence the mass

- The mass can not punish

The Folk Theorem do not scale

Outline

Continuous Time and Discounted Case

$Q_{ija}(\mathbf{m})$: transition rate matrix
Assump: $Q_{ija}(\mathbf{m})$ continuous in $\mathbf{m}$

Theorem

Under the continuity assumptions, a MFE exists.

Continuous Time and Discounted Case

$Q_{ija}(\mathbf{m})$: transition rate matrix
Assump: $Q_{ija}(\mathbf{m})$ continuous in $\mathbf{m}$

Theorem

Under the continuity assumptions, a MFE exists.

Asynchronous N-player repeated game

- $\mathcal{T}_N = \{0, 1/N, 2/N, \dots\}$
- $R(t) \Rightarrow$ choose strategy

Continuous Time and Discounted Case

$Q_{ija}(\mathbf{m})$: transition rate matrix
Assump: $Q_{ija}(\mathbf{m})$ continuous in $\mathbf{m}$

Theorem

Under the continuity assumptions, a MFE exists.

Asynchronous N-player repeated game

- $\mathcal{T}_N = \{0, 1/N, 2/N, \dots\}$
- $R(t) \Rightarrow$ choose strategy

Theorem

In continuous time, under the continuity assumptions and if $\mathbb{E}(|R(t)|^2) < \infty$, every local strategy converge to a MFE

Markov strategies

Continuous Time and Discounted Case

$Q_{ija}(\mathbf{m})$: transition rate matrix
Assump: $Q_{ija}(\mathbf{m})$ continuous in $\mathbf{m}$

Theorem

Under the continuity assumptions, a MFE exists.

Asynchronous N-player repeated game

- $\mathcal{T}_N = \{0, 1/N, 2/N, \dots\}$
- $R(t) \Rightarrow$ choose strategy

Theorem

In continuous time, under the continuity assumptions and if $\mathbb{E}(|R(t)|^2) < \infty$, every local strategy converge to a MFE

Markov strategies do not converge
$\Rightarrow$ Prisoner's dilemma

MFE $\Rightarrow a^D$

$$c(i, m^a) = \begin{cases} m_C + 3m_D & \text{if } i = C \\ 2m_D & \text{if } i = D \end{cases}$$

Continuous Time and Discounted Case

$Q_{ija}(\mathbf{m})$: transition rate matrix
Assump: $Q_{ija}(\mathbf{m})$ continuous in $\mathbf{m}$

Theorem

Under the continuity assumptions, a MFE exists.

Asynchronous N-player repeated game

- $\mathcal{T}_N = \{0, 1/N, 2/N, \dots\}$
- $R(t) \Rightarrow$ choose strategy

Theorem

In continuous time, under the continuity assumptions and if $\mathbb{E}(|R(t)|^2) < \infty$, every local strategy converge to a MFE

Markov strategies do not converge
$\Rightarrow$ Prisoner's dilemma

MFE $\Rightarrow a^D$

$$c(i, m^a) = \begin{cases} m_C + 3m_D & \text{if } i = C \\ 2m_D & \text{if } i = D \end{cases}$$

$$\pi(m) = \begin{cases} C & \text{if } m_C = 1 \\ D & \text{if } m_C < 1 \end{cases}$$

Continuous Time and Finite Horizon

Theorem

Under the continuity assumptions, a MFE exists.

Previous example $\Rightarrow$ MFE=NE (Always D)

Continuous Time and Finite Horizon

Theorem

Under the continuity assumptions, a MFE exists.

Previous example $\Rightarrow$ MFE=NE (Always D)

Prisoner's dilemma $\Rightarrow$ P: punish

	C	D	P
C	1,1	3,0	3,0
D	0,3	2,2	4,0
P	0,3	0,4	3,3

Static Nash Equilibria: D and P

Continuous Time and Finite Horizon

Theorem

Under the continuity assumptions, a MFE exists.

Previous example $\Rightarrow$ MFE=NE (Always D)

Prisoner's dilemma $\Rightarrow$ P: punish

	C	D	P
C	1,1	3,0	3,0
D	0,3	2,2	4,0
P	0,3	0,4	3,3

Static Nash Equilibria: D and P

Repeated Game Nash Equilibrium:

- if $t < 1$, play C if $m_C = 1$, play P otherwise
- if $t \geq 1$, play D if $m_P = 0$, play P otherwise.

Outline

Conclusions

Convergence:

- Local strategy $\rightarrow$ MFE
- Markov strategies do not

Discrete and continuous time

Conclusions

Convergence:

- Local strategy $\rightarrow$ MFE
- Markov strategies do not

Discrete and continuous time

MFE existence:

- Discrete and continuous time
- Discounted and finite horizon

Future work

Convergence:

- Best class of actions $\Rightarrow$ convergence

Future work

Convergence:

- Best class of actions $\Rightarrow$ convergence

MFE:

- Uniqueness conditions

Future work

Convergence:

- Best class of actions $\Rightarrow$ convergence

MFE:

- Uniqueness conditions

Full version available:
Josu Doncel, Nicolas Gast, Bruno Gaujal. Mean-Field Games with Explicit Interactions. `https://hal.inria.fr/hal-01277098`