Stochastic Uncoupled Dynamics and Nash Equilibrium

Daniel Balle

ETH Zürich

May 2017

Introduction

Definition (Uncoupledness)

A dynamic process in a game is called **uncoupled** if the strategy of each player does not depend on the utility/payoff function of other players.

Introduction

Definition (Uncoupledness)

A dynamic process in a game is called **uncoupled** if the strategy of each player does not depend on the utility/payoff function of other players.

Hart and Mas-Colell '03

There are no deterministic uncoupled stationary dynamics that guarantee almost sure convergence to pure Nash equilibria in all games where such equilibria exist.

Introduction

The Bad News

	α	β	γ
α	$1, 0$	$0, 1$	$1, 0$
β	$0, 1$	$1, 0$	$1, 0$
γ	$0, 1$	$0, 1$	$1, 1$

Figure: A simple two-player game U

Observation

In each action combination a at least one of the two players is best-replying.

Introduction

The Bad News

	α	β	γ
α	$1,0$	$0,1$	$1,0$
β	$0,1$	$1,0$	$1,0$
γ	$0,1$	$0,1$	$1,1$

Figure: A simple two-player game U

Observation

In each action combination a at least one of the two players is best-replying.

Lemma

Under uncoupled dynamics f, if player i is best-replying in state a he will play the same move again.

Introduction

The Bad News

	α	β	γ
α	$\mathbf{1,2}$	$0,1$	$1,0$
β	$0,1$	$1,0$	$1,0$
γ	$0,1$	$0,1$	$\mathbf{1,0}$

Figure: Another two-player game U'

Proof 1/2.

- Pick some a where player 1 is best-replying
- Create a new game $U' = (u^1, \bar{u}^2)$ such that a becomes the unique Nash equilibria

Introduction

The Bad News

	α	β	γ
α	$1, \mathbf{2}$	$0, 1$	$1, 0$
β	$0, 1$	$1, 0$	$1, 0$
γ	$0, 1$	$0, 1$	$1, \mathbf{0}$

Figure: Another two-player game U'

Proof 2/2.

- The dynamics f converge and thus neither player will move away from a
- Yet by uncoupledness $f^1(U) = f^1(U')$

□

Introduction

The Bad News

	α	β	γ
α	$1, 0$	$0, 1$	$1, 0$
β	$0, 1$	$1, 0$	$1, 0$
γ	$0, 1$	$0, 1$	$1, 1$

Figure: The initial two-player game U

Observation 2

In any action combination a in which only player i plays γ, player i is not best-replying and thus player j is.

Introduction

The Bad News

	α	β	γ
α	$1, 0$	$0, 1$	$1, 0$
β	$0, 1$	$1, 0$	$1, 0$
γ	$0, 1$	$0, 1$	$1, 1$

Figure: The initial two-player game U

Observation 2

In any action combination a in which only player i plays γ, player i is not best-replying and thus player j is.

Conclusion.

It follows that the state (γ, γ) can never be reached when starting from any other state. $\square$

Introduction

But ...

What if players could **remember** previous plays?
What if they had **memories**?

Recall

Influence from the Past

Definition

A strategy has R-**recall** if only the last R action combinations matter, i.e. f^i is of the form $f^i(u^i; a(t-R), ..., a(t-1))$.

Recall

Influence from the Past

Definition

A strategy has R-**recall** if only the last R action combinations matter, i.e. f^i is of the form $f^i(u^i; a(t-R), ..., a(t-1))$.

Theorem

There exist uncoupled, ***2****-recall, stationary strategy mappings that guarantee almost sure convergence to pure Nash equilibria in every game where such equilibria exist.*

Proof

Definition (State)

A state is identified as the play of the two previous periods $(a', a) := (a(t-1), a(t)) \in A \times A$.

Proof

Definition (State)

A state is identified as the play of the two previous periods $(a', a) := (a(t-1), a(t)) \in A \times A$.

Strategy f^i

- if $a' = a$ and a^i is a best reply of player i to a^{-i} then player i plays the same action a^i;
- otherwise player i picks an action $\bar{a}^i$ uniformly at random from A^i

Proof

	$a' = a$	$a' \neq a$
$a \in$ PNE	$\mathbf{S}_1$	$\mathbf{S}_2$
$a \notin$ PNE	$\mathbf{S}_4$	$\mathbf{S}_3$

Figure: State space S partition

Proof

	$a' = a$	$a' \neq a$
$a \in$ PNE	$\mathbf{S}_1$	$\mathbf{S}_2$
$a \notin$ PNE	$\mathbf{S}_4$	$\mathbf{S}_3$

Figure: State space S partition

Observation

Each state in S_1 is absorbing.

Proof

	$a' = a$	$a' \neq a$
$a \in$ PNE	$\mathbf{S}_1$	$\mathbf{S}_2$
$a \notin$ PNE	$\mathbf{S}_4$	$\mathbf{S}_3$

Figure: State space S partition

Observation

Each state in S_1 is absorbing.

Strategy f^i.

- if $a' = a$ and a^i is a best reply of player i to a^{-i} then player i plays the same action a^i;

Proof

	$a' = a$	$a' \neq a$
$a \in$ PNE	$\mathbf{S}_1$	$\mathbf{S}_2$
$a \notin$ PNE	$\mathbf{S}_4$	$\mathbf{S}_3$

Figure: State space S partition

Lemma

For all states $s \in S_2 \cup S_3 \cup S_4$ there is a strictly positive probability $p > 0$ to reach a state $s' \in S_1$ in finitely many periods.

Proof

	$a' = a$	$a' \neq a$
$a \in$ PNE	$\mathbf{S}_1$	$\mathbf{S}_2$
$a \notin$ PNE	$\mathbf{S}_4$	$\mathbf{S}_3$

Figure: State space S partition

Lemma

For all states $s \in S_2 \cup S_3 \cup S_4$ there is a strictly positive probability $p > 0$ to reach a state $s' \in S_1$ in finitely many periods.

S_2.

All players randomly pick a^i again $\Rightarrow (a, a) \in S_1$

Proof

	$a' = a$	$a' \neq a$
$a \in$ PNE	$\mathbf{S}_1$	$\mathbf{S}_2$
$a \notin$ PNE	$\mathbf{S}_4$	$\mathbf{S}_3$

Figure: State space S partition

Lemma

For all states $s \in S_2 \cup S_3 \cup S_4$ there is a strictly positive probability $p > 0$ to reach a state $s' \in S_1$ in finitely many periods.

S_3.

All players randomly pick some $\bar{a}^i$ s.t. $\bar{a} \in$ PNE $\Rightarrow (a, \bar{a}) \in S_2$

Proof

	$a' = a$	$a' \neq a$
$a \in$ PNE	$\mathbf{S}_1$	$\mathbf{S}_2$
$a \notin$ PNE	$\mathbf{S}_4$	$\mathbf{S}_3$

Figure: State space S partition

Lemma

For all states $s \in S_2 \cup S_3 \cup S_4$ there is a strictly positive probability $p > 0$ to reach a state $s' \in S_1$ in finitely many periods.

S_4.

Some player randomly picks $\bar{a}^i \Rightarrow (a, \bar{a}) \in S_2 \cup S_3$

Proof

	$a' = a$	$a' \neq a$
$a \in$ PNE	$\mathbf{S}_1$	$\mathbf{S}_2$
$a \notin$ PNE	$\mathbf{S}_4$	$\mathbf{S}_3$

Figure: State space S partition

Conclusion.

f induces an absorbing Markov Chain over S. □

Can We Go Further?

What about **Mixed** Nash Equilibria?

Mixed Equilibira

Theorem

For every small enough $\epsilon > 0$, there are no uncoupled, ***finite recall****, stationary strategy mappings f that guarantee in every game, the almost sure convergence of the behavior probabilities to Nash ϵ-equilibria.*

Proof

	α	β
α	$1, 0$	$0, 1$
β	$0, 1$	$1, 0$

(a) $U = (u^1, u^2)$

	α	β
α	$1, 1$	$0, 0$
β	$0, 1$	$1, 0$

(b) $U' = (u^1, \bar{u}^2)$

Proof

	α	β
α	$1, 0$	$0, 1$
β	$0, 1$	$1, 0$

(a) $U = (u^1, u^2)$

	α	β
α	$1, 1$	$0, 0$
β	$0, 1$	$1, 0$

(b) $U' = (u^1, \bar{u}^2)$

Observation

- Unique equilibria in U is $\underline{x}$ with $\underline{x}^i = (0.5, 0.5)$
- Unique equlibiria in U' is $\underline{a} = (\alpha, \alpha)$

Proof

	α	β
α	$1, 0$	$0, 1$
β	$0, 1$	$1, 0$

(a) $U = (u^1, u^2)$

	α	β
α	$1, 1$	$0, 0$
β	$0, 1$	$1, 0$

(b) $U' = (u^1, \bar{u}^2)$

Suppose f exists.

- f assigns $x^i(\alpha) > 0$ in both U and U'

$\square$

Proof

	α	β
α	$1, 0$	$0, 1$
β	$0, 1$	$1, 0$

(a) $U = (u^1, u^2)$

	α	β
α	$1, 1$	$0, 0$
β	$0, 1$	$1, 0$

(b) $U' = (u^1, \bar{u}^2)$

Suppose f exists.

- f assigns $x^i(\alpha) > 0$ in both U and U'
- We eventually reach the state $s = (\underline{a}, ..., \underline{a})$ in both games

$\square$

Proof

	α	β
α	$1, 0$	$0, 1$
β	$0, 1$	$1, 0$

(a) $U = (u^1, u^2)$

	α	β
α	$1, 1$	$0, 0$
β	$0, 1$	$1, 0$

(b) $U' = (u^1, \bar{u}^2)$

Suppose f exists.

- f assigns $x^i(\alpha) > 0$ in both U and U'
- We eventually reach the state $s = (\underline{a}, ..., \underline{a})$ in both games
- $f^1(s)$ should then be close to the unique Nash Equilibrium

$\square$

Proof

	α	β
α	$1, 0$	$0, 1$
β	$0, 1$	$1, 0$

(a) $U = (u^1, u^2)$

	α	β
α	$1, 1$	$0, 0$
β	$0, 1$	$1, 0$

(b) $U' = (u^1, \bar{u}^2)$

Suppose f exists.

- f assigns $x^i(\alpha) > 0$ in both U and U'
- We eventually reach the state $s = (\underline{a}, ..., \underline{a})$ in both games
- $f^1(s)$ should then be close to the unique Nash Equilibrium
- Contradicts uncoupledness since $f^1(U; s) = f^1(U'; s)$

$\square$

But...

What about if players had arbitrary **memories**?

Memory

No Continuity Restriction

Definition

A player's strategy f^i has finite R-memory if it can be implemented by a finite-state automaton in $|A|^R$ states.

Memory

No Continuity Restriction

Definition

A player's strategy f^i has finite R-memory if it can be implemented by a finite-state automaton in $|A|^R$ states.

Theorem

For every $\epsilon > 0$ there exists an uncoupled, R-memory, stationary strategy mapping that guarantees the almost sure convergence of the behavior probabilities to a Nash ϵ-equilibrium $\underline{x}$.

Proof

Buckle up

Construction 1/2

- Find some K such that there is a Nash ϵ-Equilibrium $y = (y^1, ..., y^N)$ with all probabilities being multiples of $1/K$.

Proof

Buckle up

Construction 1/2

- Find some K such that there is a Nash ϵ-Equilibrium $y = (y^1, ..., y^N)$ with all probabilities being multiples of $1/K$.
- Define f to have $R = 2K + 1$ memory.

Proof

Buckle up

Construction 1/2

- Find some K such that there is a Nash ϵ-Equilibrium $y = (y^1, ..., y^N)$ with all probabilities being multiples of $1/K$.
- Define f to have $R = 2K + 1$ memory.

Definition

- The state $\tilde{s}^i$ of player i are his R memories $(a_0, a_1, ...a_{2K})$

Proof

Buckle up

Construction 1/2

- Find some K such that there is a Nash ϵ-Equilibrium $y = (y^1, ..., y^N)$ with all probabilities being multiples of $1/K$.
- Define f to have $R = 2K + 1$ memory.

Definition

- The state $\tilde{s}^i$ of player i are his R memories $(a_0, a_1, ...a_{2K})$
- s^i denotes his $2K$ most recent memories $(a_1, ..., a_{2K})$

Proof

Buckle up

Construction 1/2

- Find some K such that there is a Nash ϵ-Equilibrium $y = (y^1, ..., y^N)$ with all probabilities being multiples of $1/K$.
- Define f to have $R = 2K + 1$ memory.

Definition

- The state $\tilde{s}^i$ of player i are his R memories $(a_0, a_1, ...a_{2K})$
- s^i denotes his $2K$ most recent memories $(a_1, ..., a_{2K})$
- Here all players will have the same memories $\tilde{s}$!

Proof

Definition

We denote by z^i the frequency distribution of actions from player i over his K most recent memories $(a_{K+1}, ..., a_{2K})$.

Proof

Definition

We denote by z^i the frequency distribution of actions from player i over his K most recent memories $(a_{K+1}, ..., a_{2K})$.

Example

For the $K = 5$ memories $((\alpha,\alpha),(\alpha,\beta),(\beta,\gamma),(\gamma,\gamma),(\gamma,\gamma))$:

$$z^1(\alpha) = 2/5, \quad z^1(\beta) = 1/5, \quad z^1(\gamma) = 2/5$$
$$z^2(\alpha) = 1/5, \quad z^2(\beta) = 1/5, \quad z^2(\gamma) = 3/5$$

Strategy f^i

Strategy f^i

if $\tilde{s} = (a_0, s)$ is not K-periodic (Mode I)

Strategy f^i

if $\tilde{s} = (a_0, s)$ is not K-periodic (Mode I)

- if s is K-periodic and

Strategy f^i

if $\tilde{s} = (a_0, s)$ is not K-periodic (Mode I)

- if s is K-periodic and
  - z^i is a ϵ-best reply to z^{-i}, **he continues his periodic play**

Strategy f^i

if $\tilde{s} = (a_0, s)$ is not K-periodic (Mode I)

- if s is K-periodic and
  - z^i is a ϵ-best reply to z^{-i}, he continues his periodic play
  - otherwise, **he breaks his periodic play**

Strategy f^i

if $\tilde{s} = (a_0, s)$ is not K-periodic (Mode I)

- if s is K-periodic and
  - z^i is a ϵ-best reply to z^{-i}, he continues his periodic play
  - otherwise, he breaks his periodic play
- if s is not K-periodic player i picks a random action $\bar{a}^i$

Strategy f^i

if $\tilde{s} = (a_0, s)$ is not K-periodic (Mode I)

- if s is K-periodic and
  - z^i is a ϵ-best reply to z^{-i}, he continues his periodic play
  - otherwise, he breaks his periodic play
- if s is not K-periodic player i picks a random action $\bar{a}^i$

The new state of f^i becomes $\tilde{s}' = (a_1, ..., a_{2K}, \bar{a})$.

Strategy f^i

if $\tilde{s} = (a_0, s)$ is not K-periodic (Mode I)

- if s is K-periodic and
  - z^i is a ϵ-best reply to z^{-i}, he continues his periodic play
  - otherwise, he breaks his periodic play
- if s is not K-periodic player i picks a random action $\bar{a}^i$

The new state of f^i becomes $\tilde{s}' = (a_1, ..., a_{2K}, \bar{a})$.

if $\tilde{s} = (a_0, s)$ is K-periodic (Mode II)

Strategy f^i

if $\tilde{s} = (a_0, s)$ is not K-periodic (Mode I)

- if s is K-periodic and
  - z^i is a ϵ-best reply to z^{-i}, he continues his periodic play
  - otherwise, he breaks his periodic play
- if s is not K-periodic player i picks a random action $\bar{a}^i$

The new state of f^i becomes $\tilde{s}' = (a_1, ..., a_{2K}, \bar{a})$.

if $\tilde{s} = (a_0, s)$ is K-periodic (Mode II)

- player i plays the mixed action z^i

Strategy f^i

if $\tilde{s} = (a_0, s)$ is not K-periodic (Mode I)

- if s is K-periodic and
  - z^i is a ϵ-best reply to z^{-i}, he continues his periodic play
  - otherwise, he breaks his periodic play
- if s is not K-periodic player i picks a random action $\bar{a}^i$

The new state of f^i becomes $\tilde{s}' = (a_1, ..., a_{2K}, \bar{a})$.

if $\tilde{s} = (a_0, s)$ is K-periodic (Mode II)

- player i plays the mixed action z^i

The state of f^i remains the same $\tilde{s}' = \tilde{s}$.

Proof

Observation

Once in Mode II each player will play the same mixed action z^i forever.

Proof

When do we reach Mode II?

Proof

When do we reach Mode II?

Mode I ($\tilde{s}$ not K-periodic)

- if s is K-periodic and
  - z^i is a ϵ-best reply to z^{-i} he continues his periodic play
  - otherwise he breaks his periodic play
- if s is not K-periodic player i picks a random action $\bar{a}^i$

Proof

When do we reach Mode II?

- For $\tilde{s} = (a_0, s)$ to be K-periodic, first s has to be

Mode I ($\tilde{s}$ not K-periodic)

- if s is K-periodic and
  - z^i is a ϵ-best reply to z^{-i} he continues his periodic play
  - otherwise he breaks his periodic play
- if s is not K-periodic player i picks a random action $\bar{a}^i$

Proof

When do we reach Mode II?

- For $\tilde{s} = (a_0, s)$ to be K-periodic, first s has to be

Mode I ($\tilde{s}$ not K-periodic)

- if s is K-periodic and
  - z^i is a ϵ-best reply to z^{-i} he continues his periodic play
  - otherwise he breaks his periodic play
- if s is not K-periodic player i **picks a random action** $\bar{a}^i$

Proof

When do we reach Mode II?

- For $\tilde{s} = (a_0, s)$ to be K-periodic, first s has to be
- Since $s' = (a_1, ..., a_{2K}, \bar{a})$ we get K-periodicity only if $\bar{a} = a_1$

Mode I ($\tilde{s}$ not K-periodic)

- if s is K-periodic and
  - z^i is a ϵ-best reply to z^{-i} he **continues his periodic play**
  - otherwise he breaks his periodic play
- if s is not K-periodic player i picks a random action $\bar{a}^i$

Proof

When do we reach Mode II?

- For $\tilde{s} = (a_0, s)$ to be K-periodic, first s has to be
- Since $s' = (a_1, ..., a_{2K}, \bar{a})$ we get K-periodicity only if $\bar{a} = a_1$

Mode I ($\tilde{s}$ not K-periodic)

- if s is K-periodic and
  - z^i **is a** ϵ**-best reply to** z^{-i} he continues his periodic play
  - otherwise he breaks his periodic play
- if s is not K-periodic player i picks a random action $\bar{a}^i$

Proof

When do we reach Mode II?

- For $\tilde{s} = (a_0, s)$ to be K-periodic, first s has to be
- Since $s' = (a_1, ..., a_{2K}, \bar{a})$ we get K-periodicity only if $\bar{a} = a_1$
- We enter Mode II when z is an ϵ-Equilibria!

Mode I ($\tilde{s}$ not K-periodic)

- if s is K-periodic and
  - z^i is a ϵ-best reply to z^{-i} he continues his periodic play
  - otherwise he breaks his periodic play
- if s is not K-periodic player i picks a random action $\bar{a}^i$

Proof

When do we reach Mode II?

- For $\tilde{s} = (a_0, s)$ to be K-periodic, first s has to be
- Since $s' = (a_1, ..., a_{2K}, \bar{a})$ we get K-periodicity only if $\bar{a} = a_1$
- We enter Mode II when z is an ϵ-Equilibria! □

Mode I ($\tilde{s}$ not K-periodic)

- if s is K-periodic and
  - z^i is a ϵ-best reply to z^{-i} he continues his periodic play
  - otherwise he breaks his periodic play
- if s is not K-periodic player i picks a random action $\bar{a}^i$

Memory is pretty cool

Memories

Memory is almost identical to Recall as players will always remember the R most recent plays. But once they reach a Nash equilibria this becomes the only memory they hold on to.

Summary

	Recall			Memory
	1	2	R	R
Pure				
Mixed				

Figure: Convergence of uncoupled dynamics

Summary

	Recall			Memory
	1	2	R	R
Pure	✗			
Mixed				

Figure: Convergence of uncoupled dynamics

Summary

	Recall			Memory
	1	2	R	R
Pure	✗	✓	✓	
Mixed				

Figure: Convergence of uncoupled dynamics

Summary

	Recall			Memory
	1	2	R	R
Pure	✗	✓	✓	
Mixed	✗	✗	✗	

Figure: Convergence of uncoupled dynamics

Summary

	Recall			Memory
	1	2	R	R
Pure	✗	✓	✓	✓
Mixed	✗	✗	✗	✓

Figure: Convergence of uncoupled dynamics