Correlated Equilibria

Aaron Roth

University of Pennsylvania

February 18 2021

Overview

- We've seen simple dynamics that converge to *Nash* equilibria in a variety of games.

Overview

- We've seen simple dynamics that converge to *Nash* equilibria in a variety of games.
- If we are sufficiently clever, might we find that this is so in all games?

Overview

- We've seen simple dynamics that converge to *Nash* equilibria in a variety of games.
- If we are sufficiently clever, might we find that this is so in all games?
- Unlikely... Finding Nash equilibria are as hard as finding general fixed points in the worst case.

Overview

- We've seen simple dynamics that converge to *Nash* equilibria in a variety of games.
- If we are sufficiently clever, might we find that this is so in all games?
- Unlikely... Finding Nash equilibria are as hard as finding general fixed points in the worst case.
- But maybe there is some richer family of equilibria we can shoot for...

Overview

- We've seen simple dynamics that converge to *Nash* equilibria in a variety of games.
- If we are sufficiently clever, might we find that this is so in all games?
- Unlikely... Finding Nash equilibria are as hard as finding general fixed points in the worst case.
- But maybe there is some richer family of equilibria we can shoot for...
- Analogous to our earlier relaxation from *Pure* to *Mixed* equilibria.

Traffic Lights

Consider the following two player traffic light game that will be familiar to those of you who can drive:

	STOP	GO
STOP	(0,0)	(0,1)
GO	(1,0)	(-100,-100)

Traffic Lights

Consider the following two player traffic light game that will be familiar to those of you who can drive:

	STOP	GO
STOP	(0,0)	(0,1)
GO	(1,0)	(-100,-100)

Two pure strategy Nash Equilibria: (GO,STOP), and (STOP,GO).

Traffic Lights

Consider the following two player traffic light game that will be familiar to those of you who can drive:

	STOP	GO
STOP	(0,0)	(0,1)
GO	(1,0)	(-100,-100)

Two pure strategy Nash Equilibria: (GO,STOP), and (STOP,GO).

But one player never gets any utility...

Traffic Lights

	STOP	GO
STOP	(0,0)	(0,1)
GO	(1,0)	(-100,-100)

There is also a mixed strategy Nash equilibrium:

Traffic Lights

	STOP	GO
STOP	(0,0)	(0,1)
GO	(1,0)	(-100,-100)

There is also a mixed strategy Nash equilibrium:

1. Suppose player 1 plays $(p, 1-p)$.

Traffic Lights

	STOP	GO
STOP	(0,0)	(0,1)
GO	(1,0)	(-100,-100)

There is also a mixed strategy Nash equilibrium:

1. Suppose player 1 plays $(p, 1-p)$.
2. If the equilibrium is to be fully mixed, player 2 must be indifferent between his two actions – i.e.:

$$0 = p - 100(1-p) \Leftrightarrow 101p = 100 \Leftrightarrow p = 100/101$$

Traffic Lights

	STOP	GO
STOP	(0,0)	(0,1)
GO	(1,0)	(-100,-100)

There is also a mixed strategy Nash equilibrium:

1. Suppose player 1 plays $(p, 1-p)$.
2. If the equilibrium is to be fully mixed, player 2 must be indifferent between his two actions – i.e.:

$$0 = p - 100(1-p) \Leftrightarrow 101p = 100 \Leftrightarrow p = 100/101$$

3. So both players play STOP with probability $p = 100/101$, and play GO with probability $(1-p) = 1/101$.

Traffic Lights

	STOP	GO
STOP	(0,0)	(0,1)
GO	(1,0)	(-100,-100)

There is also a mixed strategy Nash equilibrium:

1. Suppose player 1 plays $(p, 1-p)$.
2. If the equilibrium is to be fully mixed, player 2 must be indifferent between his two actions – i.e.:

$$0 = p - 100(1-p) \Leftrightarrow 101p = 100 \Leftrightarrow p = 100/101$$

3. So both players play STOP with probability $p = 100/101$, and play GO with probability $(1-p) = 1/101$.
4. This is even worse! Now both players get payoff 0 in expectation (rather than just one of them), and risk a horrific negative utility.

Traffic Lights

The four possible action profiles have roughly the following probabilities under this equilibrium:

	STOP	GO
STOP	98%	$<$1%
GO	$<$1%	$\approx$ 0.01%

Traffic Lights

The four possible action profiles have roughly the following probabilities under this equilibrium:

	STOP	GO
STOP	98%	<1%
GO	<1%	$\approx$ 0.01%

A better outcome would be the following, which is fair, has social welfare 1, and doesn't risk death:

	STOP	GO
STOP	0%	50%
GO	50%	0%

Traffic Lights

The four possible action profiles have roughly the following probabilities under this equilibrium:

	STOP	GO
STOP	98%	<1%
GO	<1%	$\approx$ 0.01%

A better outcome would be the following, which is fair, has social welfare 1, and doesn't risk death:

	STOP	GO
STOP	0%	50%
GO	50%	0%

Is there a Nash equilibrium that achieves this?

Traffic Lights

The four possible action profiles have roughly the following probabilities under this equilibrium:

	STOP	GO
STOP	98%	$<$1%
GO	$<$1%	$\approx$ 0.01%

A better outcome would be the following, which is fair, has social welfare 1, and doesn't risk death:

	STOP	GO
STOP	0%	50%
GO	50%	0%

Is there a Nash equilibrium that achieves this?

Worse: *there is no set of mixed strategies that creates this distribution over action profiles.*

Traffic Lights

	STOP	GO
STOP	0%	50%
GO	50%	0%

1. The reason is not that this play is not rational... (it is!)

Traffic Lights

	STOP	GO
STOP	0%	50%
GO	50%	0%

1. The reason is not that this play is not rational... (it is!)
2. The problem is that mixed strategies (as defined) requires that players randomize independently — *without coordination*.

Traffic Lights

	STOP	GO
STOP	0%	50%
GO	50%	0%

1. The reason is not that this play is not rational... (it is!)
2. The problem is that mixed strategies (as defined) requires that players randomize independently — *without coordination*.
3. In the traffic light game, players correlate! They use a traffic light.

Traffic Lights

	STOP	GO
STOP	0%	50%
GO	50%	0%

1. The reason is not that this play is not rational... (it is!)
2. The problem is that mixed strategies (as defined) requires that players randomize independently — *without coordination*.
3. In the traffic light game, players correlate! They use a traffic light.
4. Obeying traffic lights is not just a matter of obedience... Following the suggestion of the traffic light is a best response!

Traffic Lights

	STOP	GO
STOP	0%	50%
GO	50%	0%

1. The reason is not that this play is not rational... (it is!)
2. The problem is that mixed strategies (as defined) requires that players randomize independently — *without coordination*.
3. In the traffic light game, players correlate! They use a traffic light.
4. Obeying traffic lights is not just a matter of obedience... Following the suggestion of the traffic light is a best response!
5. We can generalize this...

Correlated Equilibrium

Definition

A *correlated equilibrium* is a distribution $\mathcal{D}$ over action profiles A such that for every player i, and every action a_i^*:

$$\mathrm{E}_{a\sim\mathcal{D}}[u_i(a)] \geq \mathrm{E}_{a\sim\mathcal{D}}[u_i(a_i^*, a_{-i})|a_i]$$

Correlated Equilibrium

Definition

A *correlated equilibrium* is a distribution $\mathcal{D}$ over action profiles A such that for every player i, and every action a_i^*:

$$\mathrm{E}_{a \sim \mathcal{D}}[u_i(a)] \geq \mathrm{E}_{a \sim \mathcal{D}}[u_i(a_i^*, a_{-i})|a_i]$$

In words:

> a distribution over action profiles a such that after a profile a is drawn, playing a_i is a best response for player i conditioned on seeing a_i, given that everyone else will play according to a.

Correlated Equilibrium

Definition

A *correlated equilibrium* is a distribution $\mathcal{D}$ over action profiles A such that for every player i, and every action a_i^*:

$$\mathrm{E}_{a\sim\mathcal{D}}[u_i(a)] \geq \mathrm{E}_{a\sim\mathcal{D}}[u_i(a_i^*, a_{-i})|a_i]$$

In words:

> a distribution over action profiles a such that after a profile a is drawn, playing a_i is a best response for player i conditioned on seeing a_i, given that everyone else will play according to a.

For example: Conditioned on seeing STOP, you know your opponent will GO, so STOP is a best response. Conditioned on seeing GO, you know your opponent will STOP, so GO is a best response.

Hierarchies

1. Observe: Nash Equilibria are also Correlated Equilibria — they just correspond to uncorrelated distributions. (a_i contains no information about a_{-i}).

Hierarchies

1. Observe: Nash Equilibria are also Correlated Equilibria — they just correspond to uncorrelated distributions. (a_i contains no information about a_{-i}).
2. But Correlated Equilibria are a strictly larger/richer set.

Hierarchies

1. Observe: Nash Equilibria are also Correlated Equilibria — they just correspond to uncorrelated distributions. (a_i contains no information about a_{-i}).
2. But Correlated Equilibria are a strictly larger/richer set.
3. We can define still larger sets!

Definition

A *coarse correlated equilibrium* is a distribution $\mathcal{D}$ over action profiles A such that for every player i, and every action a_i^*:

$$\mathrm{E}_{a\sim\mathcal{D}}[u_i(a)] \geq \mathrm{E}_{a\sim\mathcal{D}}[u_i(a_i^*, a_{-i})]$$

Hierarchies

1. Observe: Nash Equilibria are also Correlated Equilibria — they just correspond to uncorrelated distributions. (a_i contains no information about a_{-i}).
2. But Correlated Equilibria are a strictly larger/richer set.
3. We can define still larger sets!

Definition

A *coarse correlated equilibrium* is a distribution $\mathcal{D}$ over action profiles A such that for every player i, and every action a_i^*:

$$\mathrm{E}_{a \sim \mathcal{D}}[u_i(a)] \geq \mathrm{E}_{a \sim \mathcal{D}}[u_i(a_i^*, a_{-i})]$$

4. The difference: the suggestion just has to be a best response on average, not *conditioned* on having seen it.

Hierarchies

1. Observe: Nash Equilibria are also Correlated Equilibria — they just correspond to uncorrelated distributions. (a_i contains no information about a_{-i}).
2. But Correlated Equilibria are a strictly larger/richer set.
3. We can define still larger sets!

Definition

A *coarse correlated equilibrium* is a distribution $\mathcal{D}$ over action profiles A such that for every player i, and every action a_i^*:

$$\mathrm{E}_{a\sim\mathcal{D}}[u_i(a)] \geq \mathrm{E}_{a\sim\mathcal{D}}[u_i(a_i^*, a_{-i})]$$

4. The difference: the suggestion just has to be a best response on average, not *conditioned* on having seen it.
5. Whether it is sensible depends on whether you have to commit to following the correlating device up front, or have the option of deviating after seeing the suggestion.

Hierarchies

CCE can occasionally suggest obviously bad actions. CE cannot. Consider:

	A	B	C
A	(1,1)	(-1,-1)	(0,0)
B	(-1,-1)	(1,1)	(0,0)
C	(0,0)	(0,0)	(-1.1,-1.1)

	A	B	C
A	$1/3$		
B		$1/3$	
C			$1/3$

Hierarchies

CCE can occasionally suggest obviously bad actions. CE cannot. Consider:

	A	B	C
A	(1,1)	(-1,-1)	(0,0)
B	(-1,-1)	(1,1)	(0,0)
C	(0,0)	(0,0)	(-1.1,-1.1)

	A	B	C
A	$1/3$		
B		$1/3$	
C			$1/3$

The payoff for each player for playing according to this distribution is:

$$(1/3) \cdot 1 + (1/3) \cdot 1 - (1/3) \cdot 1.1 = 0.3$$

Hierarchies

	A	B	C
A	(1,1)	(-1,-1)	(0,0)
B	(-1,-1)	(1,1)	(0,0)
C	(0,0)	(0,0)	(-1.1,-1.1)

	A	B	C
A	1/3		
B		1/3	
C			1/3

the payoff a player would get by playing the fixed action A or B while his opponent randomized would be:

$$(1/3) \cdot 1 - (1/3) \cdot 1 + (1/3) \cdot 0 = 0$$

and the payoff for playing C would be strictly less than zero.

Hierarchies

	A	B	C
A	(1,1)	(-1,-1)	(0,0)
B	(-1,-1)	(1,1)	(0,0)
C	(0,0)	(0,0)	(-1.1,-1.1)

	A	B	C
A	1/3		
B		1/3	
C			1/3

the payoff a player would get by playing the fixed action A or B while his opponent randomized would be:

$$(1/3) \cdot 1 - (1/3) \cdot 1 + (1/3) \cdot 0 = 0$$

and the payoff for playing C would be strictly less than zero.

Hence this is a CCE *even though* conditioned on being told to play C, it is not a best response. This means that the given distribution is a coarse correlated equilibrium, *but not* a CE.

Hierarchies

Solution Concept Recap

$$DSE \subset PSNE \subset MSNE \subset CE \subset CCE$$

Hierarchies

Solution Concept Recap

$$DSE \subset PSNE \subset MSNE \subset CE \subset CCE$$

1. Starting at MSNE, we have guaranteed existence.

Hierarchies

Solution Concept Recap

$$DSE \subset PSNE \subset MSNE \subset CE \subset CCE$$

1. Starting at MSNE, we have guaranteed existence.
2. Want to show: Starting at CE, we have computational tractability.

Characterization in Terms of Regret

Definition

For a *strategy modification rule* $F_i : A_i \to A_i$ and an action profile $a \in A$:

$$\mathrm{Regret}_i(a, F_i) = u_i(F_i(a_i), a_{-i}) - u_i(a)$$

i.e. it is how much player i regrets not applying F_i to change his action.

We say that F_i is a constant strategy modification rule if $F_i(a_i) = F_i(a_i')$ for all $a_i, a_i' \in A_i$.

Characterization in Terms of Regret

Definition

For a *strategy modification rule* $F_i : A_i \to A_i$ and an action profile $a \in A$:

$$\mathrm{Regret}_i(a, F_i) = u_i(F_i(a_i), a_{-i}) - u_i(a)$$

i.e. it is how much player i regrets not applying F_i to change his action.

We say that F_i is a constant strategy modification rule if $F_i(a_i) = F_i(a_i')$ for all $a_i, a_i' \in A_i$.

We can give an alternative characterization of CCE:

Definition

A distribution $\mathcal{D}$ is a coarse correlated equilibrium if for every player i and for every *constant* strategy modification rule F_i:

$$\mathrm{E}_{a \sim \mathcal{D}}[\mathrm{Regret}_i(a, F_i)] \leq 0$$

Characterization in Terms of Regret

1. An immediate consequence of this definition is that if $a^1, \ldots, a^T$ are a sequence of actions with $\Delta(T)$ regret, then $\bar{a} = \frac{1}{T}\sum_{t=1}^{T} a^t$ forms a $\Delta(T)$-approximate coarse correlated equilibrium.

Characterization in Terms of Regret

1. An immediate consequence of this definition is that if $a^1, \ldots, a^T$ are a sequence of actions with $\Delta(T)$ regret, then $\bar{a} = \frac{1}{T}\sum_{t=1}^{T} a^t$ forms a $\Delta(T)$-approximate coarse correlated equilibrium.
2. In particular, if everyone plays an (arbitrary) game with the PW algorithm, after T steps they will have generated a sequence of plays that corresponds to a $\Delta(T) = 2\sqrt{\log k/T}$-approximate CCE

Characterization in Terms of Regret

1. An immediate consequence of this definition is that if $a^1, \ldots, a^T$ are a sequence of actions with $\Delta(T)$ regret, then $\bar{a} = \frac{1}{T}\sum_{t=1}^{T} a^t$ forms a $\Delta(T)$-approximate coarse correlated equilibrium.
2. In particular, if everyone plays an (arbitrary) game with the PW algorithm, after T steps they will have generated a sequence of plays that corresponds to a $\Delta(T) = 2\sqrt{\log k/T}$-approximate CCE
3. Can we approach computing CE in the same way? First step: characterize CE in terms of regret.

Characterization in Terms of Regret

Definition

A distribution $\mathcal{D}$ is a correlated equilibrium if for all players i and for *all* strategy modification rules F_i:

$$\mathrm{E}_{a\sim\mathcal{D}}[\mathrm{Regret}_i(a, F_i)] \leq 0$$

To see this:

Characterization in Terms of Regret

Definition

A distribution $\mathcal{D}$ is a correlated equilibrium if for all players i and for *all* strategy modification rules F_i:

$$\mathrm{E}_{a \sim \mathcal{D}}[\mathrm{Regret}_i(a, F_i)] \leq 0$$

To see this:

1. Note that a strategy modification rule F_i lets player i consider different deviations for each suggested action a_i.

Characterization in Terms of Regret

Definition

A distribution $\mathcal{D}$ is a correlated equilibrium if for all players i and for *all* strategy modification rules F_i:

$$\mathrm{E}_{a \sim \mathcal{D}}[\mathrm{Regret}_i(a, F_i)] \leq 0$$

To see this:

1. Note that a strategy modification rule F_i lets player i consider different deviations for each suggested action a_i.
2. So if there are no beneficial deviations of this sort, player i must be playing a best response even conditioned on seeing his suggestion.

Characterization in Terms of Regret

Definition

A distribution $\mathcal{D}$ is a correlated equilibrium if for all players i and for *all* strategy modification rules F_i:

$$\mathrm{E}_{a \sim \mathcal{D}}[\mathrm{Regret}_i(a, F_i)] \leq 0$$

To see this:

1. Note that a strategy modification rule F_i lets player i consider different deviations for each suggested action a_i.
2. So if there are no beneficial deviations of this sort, player i must be playing a best response even conditioned on seeing his suggestion.
3. Are there learning algorithms that efficiently converge to correlated equilibrium?

Characterization in Terms of Regret

Definition

A distribution $\mathcal{D}$ is a correlated equilibrium if for all players i and for *all* strategy modification rules F_i:

$$\mathrm{E}_{a\sim\mathcal{D}}[\mathrm{Regret}_i(a, F_i)] \leq 0$$

To see this:

1. Note that a strategy modification rule F_i lets player i consider different deviations for each suggested action a_i.
2. So if there are no beneficial deviations of this sort, player i must be playing a best response even conditioned on seeing his suggestion.
3. Are there learning algorithms that efficiently converge to correlated equilibrium?
4. Look for learning algorithms with stronger regret guarantees...

Stronger Regret Guarantees

1. We want an algorithm for learning in the experts setting that can promise...

Stronger Regret Guarantees

1. We want an algorithm for learning in the experts setting that can promise...
2. Given any k experts and an arbitrary sequence of losses $\ell^1, \ldots, \ell^T$, the algorithm chooses a sequence of experts $a_1, \ldots, a_t$ such that:

$$\frac{1}{T}\sum_{t=1}^{T} \ell_{a^t} \leq \frac{1}{T}\sum_{t=1}^{T} \ell_{F(a^t)} + \Delta(T)$$

for all strategy modification rules F and for $\Delta(T) = o(1)$.

Stronger Regret Guarantees

1. We want an algorithm for learning in the experts setting that can promise...
2. Given any k experts and an arbitrary sequence of losses $\ell^1, \ldots, \ell^T$, the algorithm chooses a sequence of experts $a_1, \ldots, a_t$ such that:

$$\frac{1}{T}\sum_{t=1}^{T} \ell_{a^t} \leq \frac{1}{T}\sum_{t=1}^{T} \ell_{F(a^t)} + \Delta(T)$$

   for all strategy modification rules F and for $\Delta(T) = o(1)$.
3. "No Swap Regret"

Stronger Regret Guarantees

1. We want an algorithm for learning in the experts setting that can promise...
2. Given any k experts and an arbitrary sequence of losses $\ell^1, \ldots, \ell^T$, the algorithm chooses a sequence of experts $a_1, \ldots, a_t$ such that:

   $$\frac{1}{T}\sum_{t=1}^{T} \ell_{a^t} \leq \frac{1}{T}\sum_{t=1}^{T} \ell_{F(a^t)} + \Delta(T)$$

   for all strategy modification rules F and for $\Delta(T) = o(1)$.
3. "No Swap Regret"
4. We'll see how to do this! (Next lecture).

Thanks!

See you next class — stay healthy, and wear a mask!