6.S890: Topics in Multiagent Learning

Lecture 3 – Prof. Daskalakis

Fall 2023

Game-Theoretic Formalism

- **Def:** A finite *n-player game* is described by:
  - a set of pure strategies/actions per player: S_p
  - a utility/payoff function per player: $u_p : \times_q S_q \to \mathbb{R}$
  - u_p: can be thought of as n-dimensional tensor
- **Def:** A *randomized/mixed strategy* for player p is any $x_p \in \Delta^{S_p}$
  - assigns probability $x_p(j)$ to each $j \in S_p$
  - i.e. Δ^{S_p} is the simplex whose vertices are identified with the elements of S_p
- **Def:** a player's *expected utility* is
  - $u_p(x_1, \ldots, x_n) = \sum_{s \in \times_q S_q} u_p(s)\, x_1(s_1) \cdot \ldots \cdot x_n(s_n) \equiv u_p \cdot (x_1 \otimes x_2 \otimes \cdots \otimes x_n)$
- A 2-player game can be described by a pair of matrices $(R\,, C)_{m \times n}$
  - rows $\overset{1-1}{\longleftrightarrow} S_1$; columns $\overset{1-1}{\longleftrightarrow} S_2$
  - player 1: "row player"; player 2: "column player"
  - mixed strategies $x \in \Delta^m$ for row player, $y \in \Delta^n$ for column player
  - expected utility of row player: $x^T R y$; expected utility of column player: $x^T C y$

Nash Equilibrium

- **Def:** a collection of mixed strategies $x_1, \dots, x_n$ is a *Nash equilibrium* iff

$$\forall i, x_i': \qquad u_i(x_i, x_{-i}) \geq u_i(x_i', x_{-i})$$

(recall that: if $x_1, \dots, x_n$ are player strategies, then x_{-i} denotes the strategies of all players except player i's)

- In 2-player games: (x, y) is Nash equilibrium iff

$$\forall x': x^T R y \geq x'^T R y$$
$$\forall y': x^T C y \geq x^T C y'$$

Nash's Theorem

[Nash 1950]: Every finite game (i.e. game with a finite number of players and a finite number of pure strategies per player) has a Nash equilibrium.

- **Proof (last time):** using Brouwer's fixed point theorem.

- **[Brouwer 1911]:** Every continuous function $f: D \to D$ from a convex compact set D to itself has a fixed point, i.e. some $x^* = f(x^*)$.

Two-player *Zero-Sum* games

Minimax Theorem [von Neumann'28]: Consider a two-player game zero-sum game $(R, C)_{m\times n}$ i.e. $R + C = 0$. Then $\min_{x\in\Delta^m} \max_{y\in\Delta^n} x^T C y = \max_{y\in\Delta^n} \min_{x\in\Delta^m} x^T C y$ (*)

Interpretation:

- (*) says: "If $\forall y, \exists x$ s.t. $x^T C y \leq v^*$ $\Rightarrow \exists x, \forall y$ s.t. $x^T C y \leq v^*$"
- If x^* is argmin of LHS, y^* argmax of RHS, v^* optimal value of (*), then (x^*, y^*) is a *Nash equilibrium*, i.e. if *min* and *max* adopt x^* and y^* then (i) *min* pays v^* to *max* and (ii) no player can improve by unilaterally deviating
- why? Because
  - under (x^*, y^*) *min* pays *max* at most v^* (since v^* optimum of LHS and x^* is argmin)
  - under (x^*, y^*) *max* receives from *min* at least v^* (since v^* optimum of RHS and y^* is argmax)
  - by the above two: under (x^*, y^*) min pays exactly v^* to max, hence (i) is proven
  - to prove (ii), suppose $\exists x$ that is a better response for min to y^*i.e. $x^T C y^* < {x^*}^T C y^* = v^*$
    - the existence of such x violates the fact that the optimum of RHS is v^* and y^* is an argmax for RHS
    - similarly the existence of a better response to x^* by max violates that the optimum of LHS is v^* and x^* is an argmin for the LHS

Two-player *Zero-Sum* games

Minimax Theorem [von Neumann'28]: Consider a two-player game zero-sum game $(R, C)_{m\times n}$ i.e. $R + C = 0$. Then $\min_{x\in\Delta^m} \max_{y\in\Delta^n} x^T C y = \max_{y\in\Delta^n} \min_{x\in\Delta^m} x^T C y$ (*)

Interpretation:

- (*) says: "If $\forall y, \exists x$ s.t. $x^T C y \le v^* \Rightarrow \exists x, \forall y$ s.t. $x^T C y \le v^*$"
- If x^* is argmin of LHS, y^* argmax of RHS, v^* optimal value of (*), then (x^*, y^*) is a *Nash equilibrium*, i.e. if *min* and *max* adopt x^* and y^* then (i) *min* pays v^* to *max* and (ii) no player can improve by unilaterally deviating
- thus von Neumann's theorem establishes the *existence of a Nash equilibrium in two-player zero-sum games*

von Neumann: "As far as I can see, there could be no theory of games … without that theorem …
I thought there was nothing worth publishing until the Minimax Theorem was proved"

Connection to mathematical programming:

- **[von Neumann-Dantzig'47, Adler'13, Brooks-Reny'21]:** minimax eq computation ⇔ Linear Programming
- Generalizes to convex-concave objectives w/ general convex compact constraint sets
    - In this case, equivalence to convex programming

Proof of von Neumann's Minimax Theorem

Minimax Theorem [von Neumann'28]: Consider a two-player game zero-sum game $(R, C)_{m\times n}$ i.e. $R + C = 0$. Then $\min_{x\in\Delta^m} \max_{y\in\Delta^n} x^T C y = \max_{y\in\Delta^n} \min_{x\in\Delta^m} x^T C y$ (*)

- Here we'll do a proof using Strong Linear Programming duality
- **Proof:**

LHS:

$$LP1: \quad \min z$$
$$x^T C e_j \leq z, \forall j$$
$$x \in \Delta^m$$

RHS:

$$LP2: \quad \max w$$
$$e_i^T C y \geq w, \forall i$$
$$y \in \Delta^n$$

LP1 and LP2 are duals!

Strong LP duality: LP1=LP2

von Neumann and Dantzig

[picture from Game Theory Alive, by Anna Karlin and Yuval Peres]

- On October 3, 1947, I visited him (von Neumann) for the first time at the Institute for Advanced Study at Princeton.
- I remember trying to describe to von Neumann, as I would to an ordinary mortal, the Air Force problem. I began with the formulation of the linear programming model in terms of activities and items, etc.
- Von Neumann did something which I believe was uncharacteristic of him. “Get to the point,” he said impatiently. Having at times a somewhat low kindling-point, I said to myself “O.K., if he wants a quicky, then that’s what he will get.”
- In under one minute I slapped the geometric and algebraic version of the problem on the blackboard. Von Neumann stood up and said “Oh that!” Then for the next hour and a half, he proceeded to give me a lecture on the mathematical theory of linear programs.
- At one point seeing me sitting there with my eyes popping and my mouth open (after I had searched the literature and found nothing), von Neumann said: “I don’t want you to think I am pulling all this out of my sleeve at the spur of the moment like a magician. I have just recently completed a book with Oskar Morgenstern on the theory of games. What I am doing is conjecturing that the two problems are equivalent. The theory that I am outlining for your problem is an analogue to the one we have developed for games.” Thus I learned about Farkas’ Lemma, and about duality for the first time.

Nash Equilibrium Existence:
two-player zero-sum games

	1/3	1/3	1/3
1/3	0,0	-1,1	1,-1
1/3	1,-1	0,0	-1 , 1
1/3	-1,1	1 , -1	0,0

[von Neumann ’28:]
In two-player zero-sum games, it always exists.
[original proof used fixed point arguments]

[Danzig ’47]

⇔

LP duality

[Adler ’13]
[Brooks-Reny’21]

No-regret Learning
cf next week lectures

Nash Equilibrium Existence: general games

[John Nash ’50]: A Nash equilibrium exists in every finite game.

Deep influence in Economics, enabling other existence results.

Proof non-constructive (uses Brouwer’s fixed point theorem)

No simpler proof has been discovered

[Daskalakis-Goldberg-Papadimitriou’06]: no simpler proof exists

i.e. **Nash Equilibrium** ⇔ **Brouwer’s Fixed Point Theorem**

cf lectures week after next

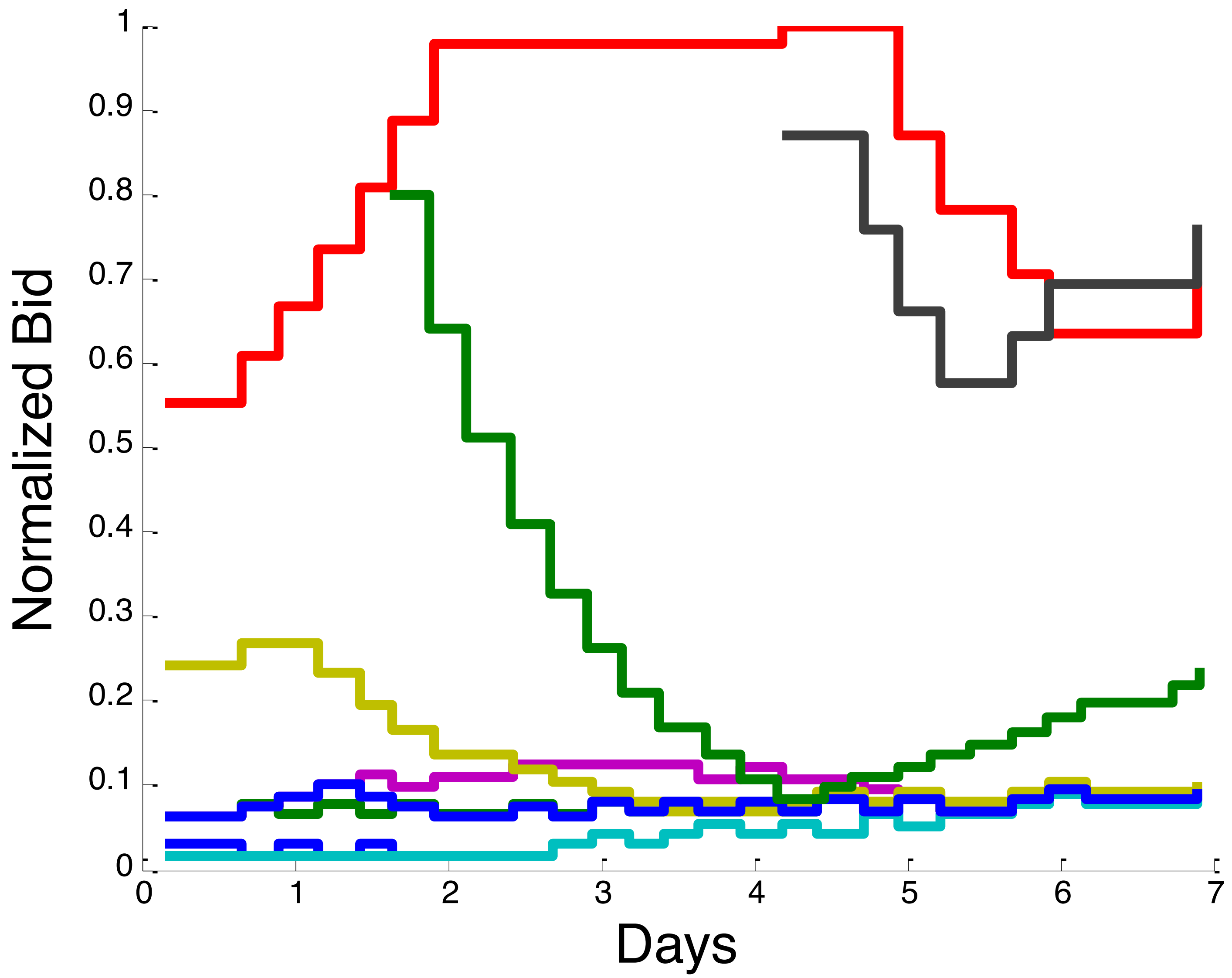
Beyond Nash Equilibrium?

Consider other equilibrium concepts
e.g. correlated equilibrium

Consider outcomes of dynamical behavior
e.g. no-regret learning

Data-Set from Microsoft's Bing

“*Econometrics for Learning Agents*” **[Nekipelov, Syrgkanis, Tardos'15]** Stationarity of behavior inconsistent with data-sets



Menu

- **Refresher: Nash & von Neumann**
- Correlated Equilibrium

Menu

- **Refresher: Nash & von Neumann**
- **Correlated Equilibrium**

Traffic Lights

Consider the following two-player traffic light game:

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

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

- but under these one player never gets any utility...

One mixed strategy equilibrium: each player plays GO w/ prob $\frac{1}{101}$ STOP w/ prob $\frac{100}{101}$

- results in an accident w/ probability $\approx 0.01\%$
- no player goes w/ probability 98%

Traffic Lights

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%

No Nash equilibrium attains the above probabilities
Worse still: *no pair of mixed strategies can attain this distribution over action profiles*

Traffic lights do!
by correlating players’ behavior

Obeying traffic lights is not just a matter of obedience...
following the suggestion of the traffic light is a best response!

Correlated Equilibrium [Aumann'74]

Def: A *correlated equilibrium* is a joint distribution $D(s_1, \dots, s_n)$ over pure strategy profiles such that for every player i, every pair of pure strategies s_i and s_i' s.t. s_i is sampled with non-zero probability by D for player i:

$$\mathbb{E}_{s_{-i}\sim D(\cdot|s_i)}[u_i(s_i, s_{-i})] \geq \mathbb{E}_{s_{-i}\sim D(\cdot|s_i)}[u_i(s_i', s_{-i})]$$

In words: "A distribution D over pure strategy profiles such that after a profile s is drawn from D, playing s_i is optimal for player i conditioned on seeing s_i but not seeing s_{-i} and assuming everyone else will play according to s_{-i}."

For example: Conditioned on seeing STOP, you know your opponent will see GO, so STOP is a best response for you assuming opponent will follow the GO recommendation. Conditioned on seeing GO, you know your opponent will see STOP, so GO is a best response assuming opponent will follow the STOP recommendation.

Hierarchies

1. Observe: Nash Equilibria are also Correlated Equilibria — they just correspond to product measures (wherein s_i contains no information about s_{-i}).
2. But Correlated Equilibria are a larger/richer set.
3. We can define still larger sets!

Def: A *coarse correlated equilibrium* is a distribution $D(s_1, \dots, s_n)$ over pure strategy profiles such that for every player i and every pure strategy s_i':

$$\mathbb{E}_{s\sim D}[u_i(s)] \geq \mathbb{E}_{s\sim D}[u_i(s_i', s_{-i})]$$

4. The difference: the recommendation just has to be optimal 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" D up front, or have the option of deviating after seeing its suggestion.

Hierarchies

Consider game:

	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)

and joint distribution D:

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

Expected payoff of, say the row player, from committing to follow recommendation of D assuming that the column player does also is > 0

Expected payoff of row from playing fixed strategy *A* or *B* assuming that the column player follows the recommendation of D is: $(1/3) \cdot 1 - (1/3) \cdot 1 + (1/3) \cdot 0 = 0$

Expected payoff of row from playing fixed strategy *C* assuming that the column player follows the recommendation of D is*:* 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 CCE *but not* a CE.

Hierarchies

Solution Concept Recap:

$$DSE \subset^{*} PSNE \subset MSNE \subset CE \subset CCE$$

Key

DSE : dominant strategy equilibrium
PSNE : pure Nash equilibrium
MSNE : (mixed) Nash equilibrium
CE : correlated equilibrium
CCE : coarse correlated equilibrium
$\subset$: strict inclusion for some games inclusion for all games
$\subset^{*}$: inclusion for games w/ no identical-payoff strategies; otherwise $DSE \subset MSNE$

1. Starting at MSNE, we have guaranteed existence.
2. **Claim:** Starting at CE, we have computational tractability.
   - First, can write CE as a linear program - *today*
   - Second, no-regret learning converges to CCE in general (and with extra work can define no-regret learner that converge to CE) – *next week*

Menu

- **Refresher: Nash & von Neumann**
- **Correlated Equilibrium**

Menu

- **Refresher: Nash & von Neumann**
- **Correlated Equilibrium**
  - **computation**

Linear Programming Formulation

Def: A *correlated equilibrium* is a joint distribution $D(s_1, \ldots, s_n)$ over pure strategy profiles such that for every player i, every pair of pure strategies s_i and s_i' s.t. s_i is sampled with non-zero probability by D for player i:

$$\mathbb{E}_{s_{-i} \sim D(\cdot|s_i)}[u_i(s_i, s_{-i})] \geq \mathbb{E}_{s_{-i} \sim D(\cdot|s_i)}[u_i(s_i', s_{-i})]$$

max 0

s.t. $\forall i, \forall s_i, s_i'$: $\sum_{s_{-i}} D(s_i, s_{-i}) u_i(s_i, s_{-i}) \geq \sum_{s_{-i}} D(s_i, s_{-i}) u_i({s_i}', s_{-i})$

$\forall s$: $D(s) \geq 0$

$\sum_s D(s) = 1$

N.B. first constraint same as

$\forall i, \forall s_i, s_i'$: $D(s_i) \sum_{s_{-i}} D(s_{-i}|s_i) u_i(s_i, s_{-i}) \geq D(s_i) \sum_{s_{-i}} D(s_{-i}|s_i) u_i(s_i', s_{-i})$

$\Leftrightarrow \forall i, \forall s_i \; s.t. D(s_i) > 0, \forall s_i'$: $\sum_{s_{-i}} D(s_{-i}|s_i) u_i(s_i, s_{-i}) \geq \sum_{s_{-i}} D(s_{-i}|s_i) u_i(s_i', s_{-i})$

Linear Programming Formulation

Def: A *correlated equilibrium* is a joint distribution $D(s_1, \dots, s_n)$ over pure strategy profiles such that for every player i, every pair of pure strategies s_i and s_i' s.t. s_i is sampled with non-zero probability by D for player i:

$$\mathbb{E}_{s_{-i} \sim D(\cdot|s_i)}[u_i(s_i, s_{-i})] \geq \mathbb{E}_{s_{-i} \sim D(\cdot|s_i)}[u_i(s_i', s_{-i})]$$

max any linear function of D

s.t. $\forall i, \forall s_i, s_i'$: $\sum_{s_{-i}} D(s_i, s_{-i}) u_i(s_i, s_{-i}) \geq \sum_{s_{-i}} D(s_i, s_{-i}) u_i(s_i{}', s_{-i})$

$\forall s$: $D(s) \geq 0$

$\sum_s D(s) = 1$

N.B. first constraint same as

$\forall i, \forall s_i, s_i'$: $D(s_i) \sum_{s_{-i}} D(s_{-i}|s_i) u_i(s_i, s_{-i}) \geq D(s_i) \sum_{s_{-i}} D(s_{-i}|s_i) u_i(s_i', s_{-i})$

$\Leftrightarrow \forall i, \forall s_i \ s.t. D(s_i) > 0, \forall s_i'$: $\sum_{s_{-i}} D(s_{-i}|s_i) u_i(s_i, s_{-i}) \geq \sum_{s_{-i}} D(s_{-i}|s_i) u_i(s_i', s_{-i})$

Linear Programming Formulation

Def: A *correlated equilibrium* is a joint distribution $D(s_1, \dots, s_n)$ over pure strategy profiles such that for every player i, every pair of pure strategies s_i and s_i' s.t. s_i is sampled with non-zero probability by D for player i:

$$\mathbb{E}_{s_{-i} \sim D(\cdot|s_i)}[u_i(s_i, s_{-i})] \geq \mathbb{E}_{s_{-i} \sim D(\cdot|s_i)}[u_i(s_i', s_{-i})]$$

max $\sum_s D(s) \sum_i u_i(s)$ (social welfare)

s.t. $\forall i, \forall s_i, s_i'$: $\sum_{s_{-i}} D(s_i, s_{-i}) u_i(s_i, s_{-i}) \geq \sum_{s_{-i}} D(s_i, s_{-i}) u_i(s_i', s_{-i})$

$\forall s$: $D(s) \geq 0$

$\sum_s D(s) = 1$

N.B. first constraint same as

$\forall i, \forall s_i, s_i'$: $D(s_i) \sum_{s_{-i}} D(s_{-i}|s_i) u_i(s_i, s_{-i}) \geq D(s_i) \sum_{s_{-i}} D(s_{-i}|s_i) u_i(s_i', s_{-i})$

$\Leftrightarrow \forall i, \forall s_i \text{ s.t. } D(s_i) > 0, \forall s_i'$: $\sum_{s_{-i}} D(s_{-i}|s_i) u_i(s_i, s_{-i}) \geq \sum_{s_{-i}} D(s_{-i}|s_i) u_i(s_i', s_{-i})$

Linear Programming Formulation

Def: A *correlated equilibrium* is a joint distribution $D(s_1, \dots, s_n)$ over pure strategy profiles such that for every player i, every pair of pure strategies s_i and s_i' s.t. s_i is sampled with non-zero probability by D for player i:

$$\mathbb{E}_{s_{-i}\sim D(\cdot|s_i)}[u_i(s_i, s_{-i})] \geq \mathbb{E}_{s_{-i}\sim D(\cdot|s_i)}[u_i(s_i', s_{-i})]$$

max $\sum_s D(s) u_{i^*}(s)$ (welfare of a particular player i^*)

s.t. $\forall i, \forall s_i, s_i'$: $\sum_{s_{-i}} D(s_i, s_{-i}) u_i(s_i, s_{-i}) \geq \sum_{s_{-i}} D(s_i, s_{-i}) u_i({s_i}', s_{-i})$

$\forall s$: $D(s) \geq 0$

$\sum_s D(s) = 1$

N.B. first constraint same as

$$\forall i, \forall s_i, s_i': \quad D(s_i) \sum_{s_{-i}} D(s_{-i}|s_i) u_i(s_i, s_{-i}) \geq D(s_i) \sum_{s_{-i}} D(s_{-i}|s_i) u_i(s_i', s_{-i})$$

$$\Leftrightarrow \forall i, \forall s_i \; s.t. D(s_i) > 0, \forall s_i': \quad \sum_{s_{-i}} D(s_{-i}|s_i) u_i(s_i, s_{-i}) \geq \sum_{s_{-i}} D(s_{-i}|s_i) u_i(s_i', s_{-i})$$

N.B.

- LP has as many variables as there are pure strategy profiles
  - n players, s strategies each: s^n
  - so to solve this LP need time $\mathrm{poly}(s^n)$
  - Technically, this is polynomial in the description of the game, as every utility tensor has as many entries as there are variables in the LP
- Can we do better?
  - Yes, if we are OK with approximate CE/CCE computation
    - using no-regret learning (next week!) can find CCE/CE in time roughly: $\frac{\mathrm{poly}(s,n)}{\epsilon^2}$
  - Yes, if the game has more structure,
    - e.g. graphical game (every player's utility depends on d other players, where $d << n$)
    - use "Ellipsoid Against Hope" algorithm of [Papadimitriou'05]