Solving two-person zero-sum sequential games via efficient computation of Nash equilibria: proof-of-concept on Kuhn's 3-card Poker

DOHMATOB Elvis [1,2]

[1]Parietal Team, INRIA
[2]Université Paris-Sud

Problem

Given a sequential two-person zero-sum game (e.g Texas Hold'em Poker), construct an optimal player.

Two-person zero sum Sequential games

- There are 3 players: nature (player 0), Alice (player 1), and Bob (player 2).
- Players take turns in making moves
- Sets of moves of any two distinct players don't overlap (i.e $A_p \cap A_q = \emptyset$ if $p \neq q$)

Two-person zero sum Sequential games

- There are 3 players: nature (player 0), Alice (player 1), and Bob (player 2).
- Players take turns in making moves
- Sets of moves of any two distinct players don't overlap (i.e $A_p \cap A_q = \emptyset$ if $p \neq q$)

Two-person zero sum Sequential games

- There are 3 players: nature (player 0), Alice (player 1), and Bob (player 2).
- Players take turns in making moves
- Sets of moves of any two distinct players don't overlap (i.e $A_p \cap A_q = \emptyset$ if $p \neq q$)

Strategy profiles, simplexes and complexes

Definition

- The *n-simplex*, denoted Δ_n, is the subset of $\mathbb{R}^{n+1}$ defined by

$$\Delta_n := \left\{ x \in \mathbb{R}^{n+1} \mid x_i \geq 0 \ \forall i, \text{ and } \sum_{i=1}^{n+1} x_i = 1 \right\} \tag{1}$$

- Each kink $\delta_i := (0, ..., 0, 1, 0, ..., 0)$ of Δ_n represents a *pure-strategy*. Any non-kink point x of Δ_n represents a *mixed-strategy*. Thus Δ_n is simply the *convex hall* of $(n+1)$-dimensional pure strategies δ_i.
- A *complex* is an abstract *polyhedron* whose kinks are themselves simplexes. Points on a complex correspond to *behavioural strategies*.

Strategy profiles, simplexes and complexes

Definition

- The *n-simplex*, denoted Δ_n, is the subset of $\mathbb{R}^{n+1}$ defined by

$$\Delta_n := \left\{ x \in \mathbb{R}^{n+1} \mid x_i \geq 0 \ \forall i, \text{ and } \sum_{i=1}^{n+1} x_i = 1 \right\} \quad (1)$$

- Each kink $\delta_i := (0, ..., 0, 1, 0, ..., 0)$ of Δ_n represents a *pure-strategy*. Any non-kink point x of Δ_n represents a *mixed-strategy*. Thus Δ_n is simply the *convex hall* of $(n+1)$-dimensional pure strategies δ_i.
- A *complex* is an abstract *polyhedron* whose kinks are themselves simplexes. Points on a complex correspond to *behavioural strategies*.

Strategy profiles, simplexes and complexes

Definition

- The *n-simplex*, denoted Δ_n, is the subset of $\mathbb{R}^{n+1}$ defined by

$$\Delta_n := \left\{ x \in \mathbb{R}^{n+1} \mid x_i \geq 0 \ \forall i, \text{ and } \sum_{i=1}^{n+1} x_i = 1 \right\} \tag{1}$$

- Each kink $\delta_i := (0, ..., 0, 1, 0, ..., 0)$ of Δ_n represents a *pure-strategy*. Any non-kink point x of Δ_n represents a *mixed-strategy*. Thus Δ_n is simply the *convex hall* of $(n+1)$-dimensional pure strategies δ_i.
- A *complex* is an abstract *polyhedron* whose kinks are themselves simplexes. Points on a complex correspond to *behavioural strategies*.

Simplexes and complexes (examples of)

NE minimax and best-response in sequential-games

Thanks to the sequence-form representation, we have:

- NE problem for Alice and Bob is the the *saddle-point problem*

$$\underset{x \in Q_1}{\text{maximize}}\ \underset{y \in Q_2}{\text{minimize}}\ y^T Ax, \tag{2}$$

  where:

$$\left.\begin{array}{l} Q_1 := \{x \in \mathbb{R}^n_+ \mid Ex = e\} \text{ is Alice's complex, and} \\ Q_2 := \{y \in \mathbb{R}^m_+ \mid Fy = f\} \text{ is Bob's complex.} \end{array}\right\} \tag{3}$$

  A saddle-point (x^*, y^*) corresponds to a NE for the game.

- Given a fixed behavioural strategy $y_0 \in Q_2$ for Bob, a *best-response* behavioural strategy x_0 for Alice [1] is a solution to the LCP

$$\underset{x \in Q_1}{\text{maximize}}\ y_0^T Ax \tag{4}$$

[1] Of course, there is an analogous concept for Bob.

NE minimax and best-response in sequential-games

Thanks to the sequence-form representation, we have:

- NE problem for Alice and Bob is the the *saddle-point problem*

$$\underset{x \in Q_1}{maximize}\ \underset{y \in Q_2}{minimize}\ y^T Ax, \tag{2}$$

  where:

$$\left.\begin{array}{l} Q_1 := \{x \in \mathbb{R}^n_+ \mid Ex = e\} \text{ is Alice's complex, and} \\ Q_2 := \{y \in \mathbb{R}^m_+ \mid Fy = f\} \text{ is Bob's complex.} \end{array}\right\} \tag{3}$$

  A saddle-point (x^*, y^*) corresponds to a NE for the game.

- Given a fixed behavioural strategy $y_0 \in Q_2$ for Bob, a *best-response* behavioural strategy x_0 for Alice [1] is a solution to the LCP

$$\underset{x \in Q_1}{maximize}\ y_0^T Ax \tag{4}$$

[1] Of course, there is an analogous concept for Bob.

Two-person zero sum Sequential games

- The game tree T (from Alice's perspective) is defined as follows. T has a set $V(T)$ of nodes (aka vertices) and a set $E(T)$ of edges.
- Each node v belongs to a single player, $p(v)$, called "the player to act at v"; $p(v) := 3$ if v is a leaf node. At node v, the player $p(v)$ has a set $C(v) \subset A_p$ of possible moves.
- For each player $p \in \{0, 1, 2\}$, the set of all nodes at which p acts is called the nodes of p, denoted $V_p(T)$.
- There are two kinds of nodes: the leaf nodes $L(T)$, at which the game must end and decision nodes $D(T)$, at which the player to act must make a move. For example in Poker, a leaf node is reached at countdown, or when a player folds, or when the player runs out of money.
- There is a special node root(T) defined by $root(T) := (/, p_0)$,
- In accordance with the rules of the game, every other node $y \in V(T) - \{root(T)\}$ is of the form $y = v.(c, p(v))$

Two-person zero sum Sequential games

- The game tree T (from Alice's perspective) is defined as follows. T has a set $V(T)$ of nodes (aka vertices) and a set $E(T)$ of edges.
- Each node v belongs to a single player, $p(v)$, called "the player to act at v"; $p(v) := 3$ if v is a leaf node. At node v, the player $p(v)$ has a set $C(v) \subset A_p$ of possible moves.
- For each player $p \in \{0, 1, 2\}$, the set of all nodes at which p acts is called the nodes of p, denoted $V_p(T)$.
- There are two kinds of nodes: the leaf nodes $L(T)$, at which the game must end and decision nodes $D(T)$, at which the player to act must make a move. For example in Poker, a leaf node is reached at countdown, or when a player folds, or when the player runs out of money.
- There is a special node root(T) defined by $root(T) := (/, p_0)$,
- In accordance with the rules of the game, every other node $y \in V(T) - \{root(T)\}$ is of the form $y = v.(c, p(v))$

Two-person zero sum Sequential games

- The game tree T (from Alice's perspective) is defined as follows. T has a set $V(T)$ of nodes (aka vertices) and a set $E(T)$ of edges.
- Each node v belongs to a single player, $p(v)$, called "the player to act at v"; $p(v) := 3$ if v is a leaf node. At node v, the player $p(v)$ has a set $C(v) \subset A_p$ of possible moves.
- For each player $p \in \{0, 1, 2\}$, the set of all nodes at which p acts is called the nodes of p, denoted $V_p(T)$.
- There are two kinds of nodes: the leaf nodes $L(T)$, at which the game must end and decision nodes $D(T)$, at which the player to act must make a move. For example in Poker, a leaf node is reached at countdown, or when a player folds, or when the player runs out of money.
- There is a special node root(T) defined by $root(T) := (/, p_0)$,
- In accordance with the rules of the game, every other node $y \in V(T) - \{root(T)\}$ is of the form $y = v.(c, p(v))$

Two-person zero sum Sequential games

- The game tree T (from Alice's perspective) is defined as follows. T has a set $V(T)$ of nodes (aka vertices) and a set $E(T)$ of edges.
- Each node v belongs to a single player, $p(v)$, called "the player to act at v"; $p(v) := 3$ if v is a leaf node. At node v, the player $p(v)$ has a set $C(v) \subset A_p$ of possible moves.
- For each player $p \in \{0, 1, 2\}$, the set of all nodes at which p acts is called the nodes of p, denoted $V_p(T)$.
- There are two kinds of nodes: the leaf nodes $L(T)$, at which the game must end and decision nodes $D(T)$, at which the player to act must make a move. For example in Poker, a leaf node is reached at countdown, or when a player folds, or when the player runs out of money.
- There is a special node root(T) defined by $root(T) := (/, p_0)$,
- In accordance with the rules of the game, every other node $y \in V(T) - \{root(T)\}$ is of the form $y = v.(c, p(v))$

Two-person zero sum Sequential games

- The game tree T (from Alice's perspective) is defined as follows. T has a set $V(T)$ of nodes (aka vertices) and a set $E(T)$ of edges.
- Each node v belongs to a single player, $p(v)$, called "the player to act at v"; $p(v) := 3$ if v is a leaf node. At node v, the player $p(v)$ has a set $C(v) \subset A_p$ of possible moves.
- For each player $p \in \{0, 1, 2\}$, the set of all nodes at which p acts is called the nodes of p, denoted $V_p(T)$.
- There are two kinds of nodes: the leaf nodes $L(T)$, at which the game must end and decision nodes $D(T)$, at which the player to act must make a move. For example in Poker, a leaf node is reached at countdown, or when a player folds, or when the player runs out of money.
- There is a special node root(T) defined by $root(T) := (/, p_0)$,
- In accordance with the rules of the game, every other node $y \in V(T) - \{root(T)\}$ is of the form $y = v.(c, p(v))$

Two-person zero sum Sequential games

- The game tree T (from Alice's perspective) is defined as follows. T has a set $V(T)$ of nodes (aka vertices) and a set $E(T)$ of edges.
- Each node v belongs to a single player, $p(v)$, called "the player to act at v"; $p(v) := 3$ if v is a leaf node. At node v, the player $p(v)$ has a set $C(v) \subset A_p$ of possible moves.
- For each player $p \in \{0, 1, 2\}$, the set of all nodes at which p acts is called the nodes of p, denoted $V_p(T)$.
- There are two kinds of nodes: the leaf nodes $L(T)$, at which the game must end and decision nodes $D(T)$, at which the player to act must make a move. For example in Poker, a leaf node is reached at countdown, or when a player folds, or when the player runs out of money.
- There is a special node root(T) defined by $root(T) := (/, p_0)$,
- In accordance with the rules of the game, every other node $y \in V(T) - \{root(T)\}$ is of the form $y = v.(c, p(v))$

Sequences and payoff matrix

- Let S_p be the sequences of moves for player p. Then any $\sigma \in S_p$ is either the empty sequence set or can be written as $\sigma_{h_p} c$ where h_p is an information set of p and c is a choice at h_p.
- Precisely,

$$S_p = \{\emptyset\} \cup \{\sigma_h c \mid h \in H_p, c \in C_h\} \tag{5}$$

- The payoff matrix $A = (a_{\sigma,\tau})$ by

$$a_{\sigma,\tau} := \sum_{\text{leaf } t:\ \sigma_1(t)=\sigma, \sigma_2(t)=\tau} \beta_0(\sigma_0(t))a(t) \tag{6}$$

Sequences and payoff matrix

- Let S_p be the sequences of moves for player p. Then any $\sigma \in S_p$ is either the empty sequence set or can be written as $\sigma_{h_p} c$ where h_p is an information set of p and c is a choice at h_p.
- Precisely,

$$S_p = \{\emptyset\} \cup \{\sigma_h c \mid h \in H_p, c \in C_h\} \tag{5}$$

- The payoff matrix $A = (a_{\sigma,\tau})$ by

$$a_{\sigma,\tau} := \sum_{\text{leaf } t:\ \sigma_1(t)=\sigma, \sigma_2(t)=\tau} \beta_0(\sigma_0(t)) a(t) \tag{6}$$

Sequences and payoff matrix

- Let S_p be the sequences of moves for player p. Then any $\sigma \in S_p$ is either the empty sequence set or can be written as $\sigma_{h_p} c$ where h_p is an information set of p and c is a choice at h_p.
- Precisely,

$$S_p = \{\emptyset\} \cup \{\sigma_h c \mid h \in H_p, c \in C_h\} \tag{5}$$

- The payoff matrix $A = (a_{\sigma,\tau})$ by

$$a_{\sigma,\tau} := \sum_{\text{leaf } t:\ \sigma_1(t)=\sigma, \sigma_2(t)=\tau} \beta_0(\sigma_0(t))a(t) \tag{6}$$

NE computation Kuhn's Poker

$$E = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 1 \\ -1 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ -1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & -1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & -1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 & 1 & 1 & 0 \end{pmatrix},$$

$$e = (1, 0, 0, 0, 0, 0, 0)$$

Non-Linear Ledoit-Wolf: Why? How?

- (p/n large) or (Σ eigenvalues close to one another) $\Rightarrow$ linear shrinkage OK.
  (p/n small) or (Σ eigenvalues dispersed) $\Rightarrow$ non-linear shrinkage better.
- Marčenko-Pastur equation: relationship between $\boldsymbol{S}$ and $\boldsymbol{\Sigma}$ eigenvalues under large-dimensional asymptotics.

Non-Linear Ledoit-Wolf: deeper into the "how?"

- H_n empirical d.f. of sample ($\boldsymbol{S}_n$) eigenvalues $\boldsymbol{\tau}_n$.
  $H_n(\boldsymbol{\tau}) \equiv \frac{1}{p}\sum_{i=1}^{p} \mathbb{1}_{[\tau_i, +\infty[}(\tau)$
- F_n e.d.f. of population ($\boldsymbol{\Sigma}_n$) eigenvalues $\boldsymbol{\lambda}_n$.
  $\boldsymbol{Y}_n = \boldsymbol{X}_n\Sigma_n^{1/2}$, $\boldsymbol{X}_n \sim \mathcal{N}(\boldsymbol{0}, \mathrm{Id})$
- F_n (resp. H_n) converges almost surely to F (resp. H).
- Stieltjes transform of a non-decreasing function G:
  $\forall z \in \mathbb{C}^+, m_G(z) \equiv \int_{-\infty}^{+\infty} \frac{1}{u-z} dG(u)$
- Marčenko-Pastur equation:
  $\forall z \in \mathbb{C}^+, m_F(z) = \int_{-\infty}^{+\infty} \frac{1}{\tau[1-\frac{p}{n}-\frac{p}{n}zm_F(z)]-z} dH(\tau)$

Non-Linear Ledoit-Wolf: Implementation...

$$Q_{n,p} : [0, \infty[^p \to [0, \infty[^p$$
$$\boldsymbol{t} \equiv (t_1, \ldots, t_p)^\mathsf{T} \mapsto Q_{n,p}(\boldsymbol{t}) \equiv (q^1_{n,p}(t), \ldots, q^p_{n,p}(t))^\mathsf{T}$$

$$\hat{\boldsymbol{\tau}}_n = \underset{\boldsymbol{t} \in [0,\infty[^p}{\operatorname{argmin}} \frac{1}{p} \sum_{i=1}^p [q^i_{n,p}(\boldsymbol{t}) - \lambda_{n,i}]^2$$

$$\breve{m}^{\hat{\tau}_n}_{n,p}(\lambda_i) \simeq \frac{n-p}{p\lambda_i} - \frac{n}{p} \frac{1}{\hat{\boldsymbol{\tau}}_{n,i}\lambda_i}$$

$\hat{d}^*_i = \frac{\lambda_i}{|1 - \frac{p}{n} - \frac{p}{n}|^2 \breve{m}^{\hat{\tau}_n}_{n,p}(\lambda_i)}$; Covariance: $S^*_n \boldsymbol{U}_n \hat{\boldsymbol{D}}^*_n \boldsymbol{U}^\mathsf{T}_n$, with $\boldsymbol{S}_n = \boldsymbol{U}_n \boldsymbol{D}_n \boldsymbol{U}^\mathsf{T}_n$

About the $Q_{n,p}$ function

$$Q_{n,p} : [0, \infty[^p \to [0, \infty[^p$$
$$\boldsymbol{t} \equiv (t_1, \ldots, t_p)^\mathsf{T} \mapsto Q_{n,p}(\boldsymbol{t}) \equiv (q^1_{n,p}(t), \ldots, q^p_{n,p}(t))^\mathsf{T}$$

$$\forall i = 1, \ldots, p \quad q^i_{n,p}(t) \equiv p \int_{(i-1)/p}^{i/p} (F^{\boldsymbol{t}}_{n,p})^{-1}(u) du$$

$$(F^{\boldsymbol{t}}_{n,p})^{-1}(u) \equiv \sup\{x \in \mathbb{R} : F^{\boldsymbol{t}}_{n,p}(x) \leq u\}$$

$$F^{\boldsymbol{t}}_{n,p}(x) \equiv \max\left\{1 - \frac{n}{p}, \frac{1}{p} \sum_{i=1}^{p} \mathbb{1}_{\{t_i = 0\}}\right\} \text{ if } x = 0,$$
$$F^{\boldsymbol{t}}_{n,p}(x) \equiv \lim_{\eta \to 0^+} \frac{1}{\pi} \int_{-\infty}^{+\infty} \operatorname{Im}\left[m^{\boldsymbol{t}}_{n,p}(\zeta + i\,\eta)\right] d\zeta \text{ otherwise}$$

$$m \equiv m^{\boldsymbol{t}}_{n,p}(z) \text{ sol. in } \left\{m \in \mathbb{C} : -\frac{n-p}{nz} + \frac{p}{n}\right\} \text{ of}$$
$$m = \frac{1}{p} \sum_{i=1}^{p} \frac{1}{t_i(1 - \frac{p}{n} - \frac{p}{n} z m) - z}$$