An accretive operator approach for stochastic games with ergodic payoff

Antoine Hochart

INRIA and CMAP Ecole polytechnique

Journées SMAI-MODE 2016
Toulouse – 25 mars 2016

Zero-sum stochastic games – model

Definition:

- **Finite state space**: $\mathbf{S} = \{1, \dots, n\}$.
- **Action spaces**: A_i (Player MIN) and B_i (Player MAX) in state i.
- **Transition payment** from Player MIN to Player MAX: $\mathbf{r_i^{ab}} \in \mathbb{R}$ with $i \in S,\ a \in A_i,\ b \in B_i$.
- **Transition probability**: $\mathbf{P_i^{ab}} = (P_{ij}^{ab})_{1 \leqslant j \leqslant n} \in \Delta(S)$.

Play:

$$\begin{array}{ccccccccc} i_0 & \longrightarrow & a_0, b_0 & \xrightarrow{\ P_{i_0}^{a_0 b_0}\ } & i_1 & \longrightarrow & a_1, b_1 & \xrightarrow{\ P_{i_1}^{a_1 b_1}\ } & i_2 \longrightarrow \cdots \\ & & \downarrow & & & & \downarrow & & \\ & & r_{i_0}^{a_0 b_0} & & & & r_{i_1}^{a_1 b_1} & & \end{array}$$

Zero-sum stochastic games – value

Value in finite horizon:

- **Play**: $(i_\ell, a_\ell, b_\ell)_{\ell\in\mathbb{N}}$ with $i_\ell \in S,\ a_\ell \in A_{i_\ell},\ b_\ell \in B_{i_\ell}$.
- **Payoff** of the k-stage game with initial state i, strategy σ for Player MIN and strategy τ for Player MAX:

$$J_i^k(\sigma,\tau) = \mathbb{E}_{i,\sigma,\tau}\left[\sum_{\ell=0}^{k-1} r_{i_\ell}^{a_\ell b_\ell}\right] \quad .$$

- **Value** of the k-stage game with initial state i:

$$v_i^k = \inf_\sigma \sup_\tau J_i^k(\sigma,\tau) = \sup_\tau \inf_\sigma J_i^k(\sigma,\tau) \quad .$$

Assumption

The value v_i^k exists for all $i \in S$ and $k \in \mathbb{N}$.

Zero-sum stochastic games – ergodic payoff

Mean payoff vector:

$$\chi := \lim_{k \to \infty} \frac{v^k}{k} \quad .$$

Existence of the mean payoff:

- recursive games (*Everett*, '57);
- absorbing games (*Kohlberg*, '74);
- finite stochastic games (*Bewley & Kohlberg*, '76);
- games with incomplete information (*Aumann & Maschler*, '95);
- Markov chain games with incomplete information (*Renault*, '06).

Counterexamples:

- stochastic game (*Vigeral*, '13);
- zero-sum repeated game with symmetric information (*Ziliotto*, '13).

Question

When is the mean payoff independent of the initial state?

Operator approach

Shapley operator $T : \mathbb{R}^S \to \mathbb{R}^S$,

$$T_i(x) = \inf_{a \in A_i} \sup_{b \in B_i} \left(r_i^{ab} + \sum_{j=1}^{n} P_{ij}^{ab} x_j \right) \left(= \sup_{b \in B_i} \inf_{a \in A_i} \left(r_i^{ab} + \sum_{j=1}^{n} P_{ij}^{ab} x_j \right) \right) \quad , \quad i \in S \ .$$

Dynamic programming principle: $\quad v_i^0 = 0 \ , \quad v_i^{k+1} = T_i(v^k) \quad .$

$$\implies \quad \chi := \lim_{k \to \infty} \frac{v^k}{k} = \lim_{k \to \infty} \frac{T^k(0)}{k}$$

Constant-mean-payoff problem related to the solvability of the **ergodic equation**:

$$T(u) = \lambda e + u \ , \quad \lambda \in \mathbb{R}, \ u \in \mathbb{R}^S \ . \qquad \text{(Erg)}$$

- If the ergodic equation is solvable, then the **ergodic constant** λ gives the mean payoff for every initial state: $\chi_i = \lambda, \ \forall i \in S$. $\quad$ (u **bias vector**.)
- When T is **polyhedral** (finite action spaces), the ergodic equation is solvable iff the mean payoff vector χ is constant.

Solvability of the ergodic equation

Definition (ergodic game)

We say that a zero-sum repeated game with Shapley operator T is **ergodic** when the ergodic equation is solvable for all **perturbed operators** $g + T$ with $g \in \mathbb{R}^S$ (i.e., perturbed games in which the payments in state i is increased by g_i).

Proposition (zero-player case)

A finite Markov chain (zero-player game) with transition matrix $P \in \mathbb{R}^{n \times n}$ is ergodic iff for every $g \in \mathbb{R}^n$, there exist $\lambda \in \mathbb{R}$ and $u \in \mathbb{R}^n$ such that $g + Pu = \lambda e + u$.

THEOREM 1

A stochastic game with **finite state space** is ergodic iff all the slice spaces are bounded in the Hilbert's seminorm.

Definition

- **slice space**: $\mathcal{S}_\alpha^\beta(T) := \{x \in \mathbb{R}^S \mid \alpha e + x \leqslant T(x) \leqslant \beta e + x\}$
- **Hilbert's seminorm**: $\|x\|_{\mathsf{H}} := \max_i x_i - \min_i x_i$

Ergodicity conditions for zero-sum repeated games

Sufficient condition:

Theorem (corollary of Gaubert & Gunawardena, TAMS '04)

If all the slice spaces $\mathcal{S}_\alpha^\beta(T)$ are bounded in the Hilbert's seminorm, then
$\forall g \in \mathbb{R}^n, \;\; \exists(\lambda, u) \in \mathbb{R} \times \mathbb{R}^n, \;\; g + T(u) = \lambda e + u.$

Algorithmic aspects:

Corollary (Akian, Gaubert & H., CDC '15)

Ergodicity of a game can be checked in time $O(2^{|S|}\text{poly}(|A|, |B|))$ by a deterministic Turing machine with oracles $\Omega^\pm$.

- Oracles $\Omega^\pm$ based on $\lim_{\rho \to \pm\infty} T_i(\rho e_J)$.
- **coNP-hard** problem, but fixed-parameter tractable.

From ergodic equation to fixed-point problem

Let $T : \mathbb{R}^S \to \mathbb{R}^S$ be a Shapley operator.

- T is order-preserving: $x \leqslant y \implies T(x) \leqslant T(y)$.
- T is additively homogeneous: $T(x + \lambda e) = T(x) + \lambda e \; , \quad \forall \lambda \in \mathbb{R}$.
- T is **nonexpansive w.r.t. the Hilbert's seminorm**:

$$\|T(x) - T(y)\|_{\mathsf{H}} \leqslant \|x - y\|_{\mathsf{H}} \; .$$

Let $\mathbb{TR}^S := \mathbb{R}^S/\mathbb{R}e$ be the **"additive projective space"** (the set of $x + \mathbb{R}e$).

- $(\mathbb{TR}^S, \|\cdot\|_H)$ is a finite-dimensional normed space.
- T can be quotiented into a map $[T] : \mathbb{TR}^S \to \mathbb{TR}^S$ nonexpansive w.r.t. $\|\cdot\|_H$.

Lemma

T has a (unique up to an additive constant) bias vector, iff $[T]$ has a (unique) fixed point.

Accretive operators

Let $(\mathcal{X}, \|\cdot\|)$ be a finite-dimensional normed space, and $(\mathcal{X}^*, \|\cdot\|^*)$ its dual.

Definition

- The (normalized) **duality mapping** on $\mathcal{X}$ is the set-valued map

$$J : \mathcal{X} \rightrightarrows \mathcal{X}^*, \quad x \mapsto \{x^* \in \mathcal{X}^* \mid \|x^*\|^* = \|x\|, \ \langle x, x^* \rangle = \|x\|^2\} \ .$$

- A set-valued map $A : \mathcal{X} \rightrightarrows \mathcal{X}$ is **accretive** if

$$\forall x, y \in \mathcal{X}, u \in A(x), v \in A(y), \quad \exists x^* \in J(x-y), \quad \langle u - v, x^* \rangle \geqslant 0 \ .$$

- A is **m-accretive** if it is accretive and $\mathrm{rg}(\mathrm{Id} + \lambda A) = \mathcal{X}$ for some $\lambda > 0$.
- A is **coaccretive** if A^{-1} is accretive.

Lemma

If $T : \mathcal{X} \to \mathcal{X}$ is nonexpansive, then $A := \mathrm{Id} - T$ is m-accretive.

Local boundedness of coaccretive maps

Proposition

Let $\mathcal{X}$ be a finite-dimensional normed space.
Let $A : \mathcal{X} \rightrightarrows \mathcal{X}$ be a **coaccretive** map.
Then, A is locally bounded at each point in the interior of $\operatorname{dom}(A) := \{x \in \mathcal{X} \mid A(x) \neq \emptyset\}$.

Theorem (Fitzpatrick, Hess, Kato, '72, see also Browder, '68)

Let $\mathcal{X}$ be a Banach space such that $\mathcal{X}'$ is uniformly convex.
Let $A : \mathcal{X} \rightrightarrows \mathcal{X}$ be an **accretive** map.
For every point in the interior of its domain, A is locally bounded.

Existence stability of fixed points

Theorem (surjectivity condition for accretive operators)

Let $A : \mathcal{X} \rightrightarrows \mathcal{X}$ be an accretive map.

$$\mathrm{rg}(A) := \bigcup_{x \in \mathcal{X}} A(x) = \mathcal{X} \quad \Longrightarrow \quad \mathcal{S}_\gamma := \{x \in \mathcal{X} \mid \mathrm{dist}(A(x), 0) \leqslant \gamma\} \text{ bounded } \forall \gamma \geqslant 0 \ .$$

Theorem (Kirk & Schöneberg, '80)

Let $A : \mathcal{X} \rightrightarrows \mathcal{X}$ be an m-accretive map.

$$\mathcal{S}_\gamma := \{x \in \mathcal{X} \mid \mathrm{dist}(A(x), 0) \leqslant \gamma\} \text{ bounded } \forall \gamma \geqslant 0 \quad \Longrightarrow \quad \mathrm{rg}(A) = \mathcal{X} \ .$$

Corollary

Let $T : \mathcal{X} \to \mathcal{X}$ be a nonexpansive operator. T.F.A.E. :

1. for every vector $g \in \mathcal{X}$, the operator $g + T$ has at least one fixed point;
2. for every $\gamma > 0$, the set $\mathcal{S}_\gamma(T) = \{x \in \mathcal{X} \mid \|x - T(x)\| \leqslant \gamma\}$ is bounded.

proof.

Let $A = \mathrm{Id} - T$*:* $\quad \mathrm{rg}(A) = \mathcal{X} \iff \forall g \in \mathcal{X},\ \exists x \in \mathcal{X},\ g + T(x) = x \ .$

Uniqueness of the bias vector (up to an additive constant)

THEOREM 2

If a stochastic game with **finite state space** is ergodic, then the bias vector is unique (up to an additive constant) for a generic perturbation vector $g \in \mathbb{R}^S$ of the Shapley operator.

- In the **policy iteration algorithm**, uniqueness is important to avoid cycling.
- In **one-player games** (optimal control problems) with continuous time (Aubry-Mather theory, weak-KAM theory) the result is already known (Rifford).

Generic uniqueness of fixed points

Let $T : \mathcal{X} \to \mathcal{X}$ be a nonexpansive operator, and $A = \mathrm{Id} - T$.
Let $\mathrm{FP} : \mathcal{X} \rightrightarrows \mathcal{X},\ \ g \mapsto \{x \in \mathcal{X} \mid g + T(x) = x\}$ be the **fixed point map**.
Note that $\mathrm{FP} = A^{-1}$: $\quad x \in \mathrm{FP}(g) \quad \Longleftrightarrow \quad g = (\mathrm{Id} - T)(x) = A(x)$.

Assume that $\mathrm{dom}(\mathrm{FP}) = \mathcal{X}$ (i.e., that $g + T$ has a fixed point for every $g \in \mathcal{X}$).

Theorem

- FP has compact values and is upper semicontinuous.
- FP is continuous at point $g \in \mathcal{X} \quad \Longleftrightarrow \quad \mathrm{FP}(g)$ is a singleton.

Theorem (generic continuity, see Aubin and Frankowska, '09)

Let $F : \mathcal{Y} \rightrightarrows \mathcal{Z}$, with $\mathcal{Y}, \mathcal{Z}$ complete metric spaces and $\mathcal{Z}$ separable. If F is u.s.c., then it is continuous on a residual of $\mathcal{Y}$ (countable intersection of dense open subsets).

Corollary

The fixed point of $g + T$ is unique for every g in a residual of $\mathcal{X}$.

Summary of main result

Theorem

Let $T : \mathbb{R}^S \to \mathbb{R}^S$ be the Shapley operator of a stochastic game with finite state space. T.F.A.E. :

1. the ergodic equation $g + T(u) = \lambda e + u$ has a solution for all $g \in \mathbb{R}^S$;
2. all the subsets $\mathcal{S}_\gamma(T) := \{x \in \mathbb{R}^S \mid \|x - T(x)\|_{\mathsf{H}} \leqslant \gamma\}$ are bounded in the Hilbert's seminorm;
3. all the slice spaces $\mathcal{S}_\alpha^\beta(T) := \{x \in \mathbb{R}^S \mid \alpha e \leqslant T(x) - x \leqslant \beta e\}$ are bounded in the Hilbert's seminorm.

Moreover, if one of the properties holds, then the set of perturbation vectors $g \in \mathbb{R}^S$ for which $g + T$ has a unique bias (up to an additive constant) is a residual of $\mathbb{R}^S$.

Perspectives

- Can we extend these results to the case of games with infinite state space \differential games (Hamilton-Jacobi-Isaacs PDE)?
- Can we describe the set of bias vectors when the ergodic equation is solvable?

Thank you