Marginal values of Stochastic Games: How fragile is my game?

L. Attia[2]   **R. Saona**[1]   M. Oliu-Barton[2]

[1]ISTA

[2]CEREMADE, CNRS, Université Paris Dauphine, PSL Research Institute

Models are approximations

A model is an approximation of reality.
Conclusions should approximately hold on perturbed models.
Such an approximation is better quantified.

Value Strategies

Example, perturbed

$$M = \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix}$$

The optimal strategy is given by,

$$p^* = \left(\frac{1}{2}, \frac{1}{2} \right)^\top .$$

Therefore,

$$\mathrm{val} M = 0 \,.$$

Example, perturbed

Consider $\varepsilon > 0$.

$$M(\varepsilon) = \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix} + \begin{pmatrix} 1 & -3 \\ 0 & 2 \end{pmatrix} \varepsilon \,.$$

The optimal strategy is given by, for $\varepsilon < 1/2$,

$$p^*_\varepsilon = \left(\frac{1+\varepsilon}{2+3\varepsilon}, \frac{1+2\varepsilon}{2+3\varepsilon} \right)^\top \,.$$

Therefore,

$$\operatorname{val} M(\varepsilon) = \frac{\varepsilon^2}{2+3\varepsilon} \,.$$

Example, perturbed 2

Consider $\varepsilon > 0$.

$$M(\varepsilon) = \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix} + \begin{pmatrix} -1 & 3 \\ 0 & -2 \end{pmatrix} \varepsilon \,.$$

The optimal strategy is given by, for $\varepsilon < 2/3$,

$$p^*_\varepsilon = \left(\frac{1-\varepsilon}{2-3\varepsilon}, \frac{1-2\varepsilon}{2-3\varepsilon} \right)^\top \,.$$

Therefore,

$$\mathrm{val} M(\varepsilon) = \frac{\varepsilon^2}{2-3\varepsilon} \,.$$

Questions

Definition (Value-positivity problem)

Is the value function increasing?

Definition (Functional form problem)

What is the value function and some optimal strategy function?

Definition (Uniform value-positivity problem)

Can the max-player guarantee the unperturbed value in the perturbed game with a fixed strategy?

Question for Stochastic Games

Definition (Marginal value)

Consider a stochastic game Γ and a perturbation H.
The marginal value is

$$D_H \operatorname{val}(\Gamma) := \lim_{\varepsilon \to 0^+} \frac{\operatorname{val}(\Gamma + H\varepsilon) - \operatorname{val}(\Gamma)}{\varepsilon},$$

i.e., the right derivative at zero of $\varepsilon \mapsto \operatorname{val}(\Gamma + H\varepsilon)$.

To appear in Mathematics of Operations Research
https://doi.org/10.1287/moor.2023.0297

Preliminaries

Stochastic Games

Stcohastic games. A game $\Gamma = (K, k, I, J; g, q, \lambda)$, where

- K is a finite set of states,
- $k \in K$ is the initial state,
- I and J are the finite action sets respectively of Player 1 and 2,
- $g \colon K \times I \times J \to \mathbb{R}$ is the payoff function,
- $q \colon K \times I \times J \to \Delta(K)$ is the transition function, and
- $\lambda \in [0, 1]$ is the discount rate.

Payoff and Values

Payoff.

$$\gamma_\lambda(\sigma,\tau) := \mathbb{E}^k_{\sigma,\tau}\left(\sum_{m\geq 1}\lambda(1-\lambda)^{m-1}G_m\right)$$

$$\gamma_0(\sigma,\tau) := \mathbb{E}^k_{\sigma,\tau}\left(\liminf_{\lambda\to 0}\sum_{m\geq 1}\lambda(1-\lambda)^{m-1}G_m\right)$$

Value.

$$\mathrm{val}(\Gamma) := \sup_\sigma \inf_\tau \gamma_\lambda(\sigma,\tau)\,.$$

Perturbation

Perturbation.

$$H = (\tilde{g}, \tilde{q}, \tilde{\lambda}) \,,$$

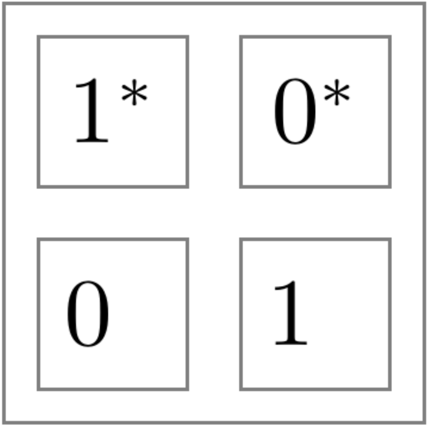
where

- $\tilde{g} \colon K \times I \times J \to \mathbb{R}$
- $\tilde{q} \colon K \times I \times J \to \mathbb{R}$
- $\tilde{\lambda} \in \mathbb{R}$

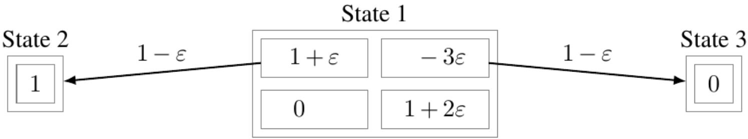
are such that $(\Gamma + H\varepsilon)$ is a stochastic game for small enough ε.

Note: No perturbation of available strategies.

Example



Example, perturbed



Previous results

Mills 1956

Theorem

Consider a matrix game M_0. For all perturbations M_1,

$$D_{M_1}\mathrm{val}(M_0) = \max_{p \in P(M_0)} \min_{q \in Q(M_0)} p^\top M_1 q \,.$$

In other words, defining $M(\varepsilon) = M_0 + M_1\varepsilon$,

$$D_{M_1}\mathrm{val}(M_0) = \mathrm{val}_{O^*(M_0)}(D\, M(0)) \,.$$

Kohlberg 1974

Theorem

There is a stochastic game such that

$$\mathrm{val}(\Gamma_\lambda) = \frac{1-\sqrt{\lambda}}{1-\lambda} = 1 - \sqrt{\lambda} + o(\sqrt{\lambda})\,.$$

Filar and Vrieze 1997

Theorem

Consider a stochastic game Γ *with* $\lambda > 0$*. For all perturbations* H*,*

$$|\mathrm{val}(\Gamma + H\varepsilon) - \mathrm{val}(\Gamma)| \leq \frac{\varepsilon}{\lambda} C(\Gamma, H)\,.$$

Solan 2003

Theorem

Consider a stochastic game Γ with $\lambda \geq 0$. For all perturbations H that neither perturb the discount factor ($\tilde{\lambda} = 0$) nor introduce new transitions,

$$|\mathrm{val}(\Gamma + H\varepsilon) - \mathrm{val}(\Gamma)| \leq \varepsilon\, C(\Gamma, H)\,.$$

Semi-algebraic theory

Theorem

Consider a stochastic game Γ *with* $\lambda \geq 0$. *For a perturbations* H, *where, if* $\lambda = 0$, *then* H *does not introduce new transitions, then*

$$\varepsilon \mapsto \operatorname{val}(\Gamma + H\varepsilon) \text{ is a Puiseux series.}$$

State of affairs

In many reasonable cases, the marginal value exists.

How can we compute it?

Oliu-Barton and Attia 2019

Theorem

Consider a stochastic game Γ *with* $\lambda > 0$*.*
Then, $\operatorname{val}(\Gamma)$ *is the unique solution of*

$$\operatorname{val}(\Delta^k - z\Delta^0) = 0\,,$$

where Δ^k *and* Δ^0 *are matrices constructed from* Γ *and* Δ^0 *is strictly positive.*

Results

Marginal discounted value

Theorem

Consider a stochastic game Γ *with* $\lambda > 0$ *and a perturbation* H*. Then,* $D_H \operatorname{val}(\Gamma)$ *is the unique* $z \in \mathbb{R}$ *satisfying*

$$\operatorname{val}_{O^*(\Gamma)} \left(D_H \Delta^k - \operatorname{val}(\Gamma) D_H \Delta^0 - z \Delta^0 \right) = 0 \,.$$

Marginal undiscounted value

Theorem

Consider a stochastic game Γ *with* $\lambda = 0$
and a (undiscounted) perturbation $H = (\tilde{g}, \tilde{q}, \tilde{\lambda} = 0)$.
Asume that $\varepsilon \mapsto \operatorname{val}(\Gamma + H\varepsilon)$ *is continuous at zero.*
Let p *be a polynomial such that, for all* $\varepsilon > 0$ *small enough,*

$$p(\varepsilon, \operatorname{val}(\Gamma + H\varepsilon)) = 0$$

and such that $\partial_2 p(0, \operatorname{val}(\Gamma)) \neq 0$.
Then,

$$D_H \operatorname{val}(\Gamma) = -\frac{\partial_1 p(0, \operatorname{val}(\Gamma))}{\partial_2 p(0, \operatorname{val}(\Gamma))} .$$

Sketch proof: Marginal discounted value

Sketch proof.

For every ε, define the matrix game $M(\varepsilon) := \Delta_\varepsilon^k - \mathrm{val}(\Gamma + H\varepsilon)\Delta_\varepsilon^0$.
By Oliu-Barton and Attia, for all ε, we have $\mathrm{val}(M(\varepsilon)) = 0$.
Differentiating, by Mills, we have

$$\begin{aligned} D\,\mathrm{val}(M)(0) &= \mathrm{val}_{O^* M(0)}(D\,M(0)) \\ &= \mathrm{val}_{O^* M(0)}(D_H\,\Delta^k - \mathrm{val}(\Gamma)\,D_H\,\Delta^0 - D_H\,\mathrm{val}(\Gamma)\,\Delta^0) \\ &= 0\,. \end{aligned}$$

□

Half-true: $O^* M(0) = O^*(\Gamma)$ is not proven in full generality.
Instead, take optimal strategies and Taylor approximations.

Sketch proof: Marginal undiscounted value

Sketch proof.

Consider the polynomial p such that $p(\varepsilon, \mathrm{val}(\Gamma + H\varepsilon)) = 0$ and $\partial_2 p(0, \mathrm{val}(\Gamma)) \neq 0$.
Differentiating, we obtain

$$D\, p(\cdot, \mathrm{val}(\Gamma + H \cdot))(0) = \partial_1 p(0, \mathrm{val}(\Gamma)) + \partial_2 p(0, \mathrm{val}(\Gamma))\, D_H \mathrm{val}(\Gamma) = 0 .$$

Reordering

$$D_H \mathrm{val}(\Gamma) = -\frac{\partial_1 p(0, \mathrm{val}(\Gamma))}{\partial_2 p(0, \mathrm{val}(\Gamma))} .$$

□

Great, but where does p come from?

Matrices Δ^k and Δ^0, explained

Consider the perturbed Big Match.
Fix a pure stationary strategy $(i, j) = (\text{Top}, \text{Left})$.
The induced Markov Chain has payoffs

$$g(i,j) = (1, 1+\varepsilon, 0)^\top$$

and transition matrix

$$Q(i,j) = \begin{pmatrix} 1 & 0 & 0 \\ 1-\varepsilon & \varepsilon & 0 \\ 1 & 0 & 1 \end{pmatrix}$$

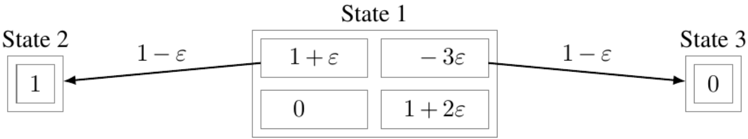
Then,

$$\Delta^0_\varepsilon(i,j) = \det(Id - (1-\lambda)Q) = \lambda^2(1 - \varepsilon(1-\lambda))\,.$$

Also,

$$\Delta^k_\varepsilon(i,j) = \lambda^2(1+\varepsilon)\,.$$

Example, perturbed



Oliu-Barton and Attia 2019

Theorem

Consider a stochastic game Γ *with* $\lambda > 0$.
Then, $\mathrm{val}(\Gamma)$ *is the unique solution of*

$$\mathrm{val}(\Delta^k - z\Delta^0) = 0\,,$$

where Δ^k *and* Δ^0 *are matrices constructed from* Γ *and* Δ^0 *is strictly positive.*

Origin of polynomial for marginal undiscounted value

Theorem

Consider a stochastic game Γ *with* $\lambda = 0$ *and a (undiscounted) perturbation* $H = (\tilde{g}, \tilde{q}, \tilde{\lambda} = 0)$*. Then, there is an explicit finite set of candidate polynomials including a polynomial* p *such that*

$$p(\varepsilon, \mathrm{val}(\Gamma + H\varepsilon)) = 0$$

but not necessarily $\partial_2 p(0, \mathrm{val}(\Gamma)) \neq 0$*.*

Limit and marginal value

We know that

$$\lim_{\lambda \to 0} \operatorname{val}(\Gamma_\lambda) = \operatorname{val}(\Gamma_0)\,.$$

Does this occur with the marignal values?
No,

$$\lim_{\lambda \to 0} D_H \operatorname{val}(\Gamma_\lambda) \neq D_H \operatorname{val}(\Gamma_0)\,.$$

Wrapping up

Marginal discounted value

Theorem

Consider a stochastic game Γ *with* $\lambda > 0$ *and a perturbation* H*. Then,* $D_H \operatorname{val}(\Gamma)$ *is the unique* $z \in \mathbb{R}$ *satisfying*

$$\operatorname{val}_{O^*(\Gamma)} \left(D_H \, \Delta^k - \operatorname{val}(\Gamma) \, D_H \, \Delta^0 - z \, \Delta^0 \right) = 0 \, .$$

Marginal undiscounted value

Theorem

Consider a stochastic game Γ *with* $\lambda = 0$
and a (undiscounted) perturbation $H = (\tilde{g}, \tilde{q}, \tilde{\lambda} = 0)$.
Asume that $\varepsilon \mapsto \operatorname{val}(\Gamma + H\varepsilon)$ *is continuous at zero.*
Let p be a polynomial such that, for all $\varepsilon > 0$ *small enough,*

$$p(\varepsilon, \operatorname{val}(\Gamma + H\varepsilon)) = 0$$

and such that $\partial_2 p(0, \operatorname{val}(\Gamma)) \neq 0$.
Then,

$$D_H \operatorname{val}(\Gamma) = -\frac{\partial_1 p(0, \operatorname{val}(\Gamma))}{\partial_2 p(0, \operatorname{val}(\Gamma))}\,.$$

Question for Stochastic Games

Definition (Marginal value)

Consider a stochastic game Γ and a perturbation H.
The marginal value is

$$D_H \operatorname{val}(\Gamma) := \lim_{\varepsilon \to 0^+} \frac{\operatorname{val}(\Gamma + H\varepsilon) - \operatorname{val}(\Gamma)}{\varepsilon},$$

i.e., the right derivative at zero of $\varepsilon \mapsto \operatorname{val}(\Gamma + H\varepsilon)$.

To appear in Mathematics of Operations Research
https://doi.org/10.1287/moor.2023.0297

Questions

Definition (Value-positivity problem)

Is the value function increasing?

Definition (Functional form problem)

What is the value function and some optimal strategy function?

Definition (Uniform value-positivity problem)

Can the max-player guarantee the unperturbed value in the perturbed game with a fixed strategy?

Extras

Perturbing the discount factor in Stochastic Games

Consider a Stochastic Game Γ and its parametrized polynomial matrix game $M_z := \Delta^k - z\Delta^0$.

Lemma (Value-positivity)

For all $z \in \mathbb{R}$, if M_z is value-positive, then, for all small λ,

$$\mathrm{val}_\lambda \Gamma \geq z\,.$$

Lemma (Uniform value-positivity)

For all $z \in \mathbb{R}$, if M_z is uniform value-positive, then there exists a fixed strategy $p \in (\Delta[m])^n$ such that, for all λ sufficiently small,

$$\mathrm{val}_\lambda(\Gamma; p) \geq z\,.$$