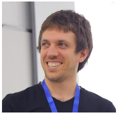
Marginal Values of Stochastic Games: Directional Derivative of the Value

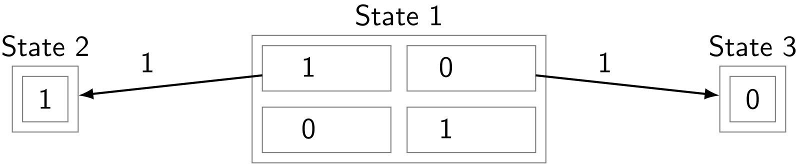




L. Attia[2] **R. Saona**[1] M. Oliu-Barton[2]

[1]Institute of Science and Technology Austria (ISTA)
→ London School of Economics (LSE)

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

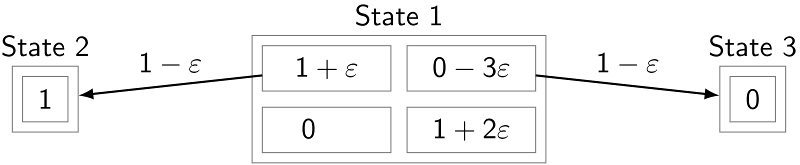
16th Viennese Conference on Optimal Control and Dynamic



The big Match

For all $\lambda \in \{0\} \cup (0,1]$, the (normalized) discounted value is

$$\mathrm{val}(\Gamma_\lambda) := \sup_\sigma \inf_\tau \mathbb{E}_{s_1}^{\sigma,\tau}\left[\lambda \sum_{m \geq 0} (1-\lambda)^m G_m\right] = 1/2\,.$$



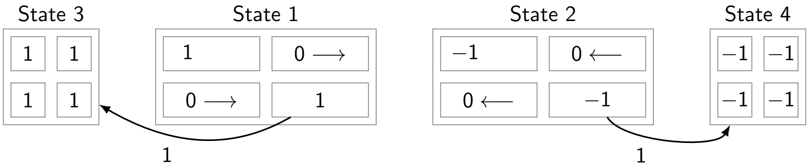
A perturbed big Match

The marginal value is

$$D_H \operatorname{val}(\Gamma_\lambda) := \lim_{\varepsilon \to 0^+} \frac{\operatorname{val}(\Gamma_\lambda(\varepsilon)) - \operatorname{val}(\Gamma_\lambda(0))}{\varepsilon} = \frac{1-\lambda}{2}\,.$$

So,

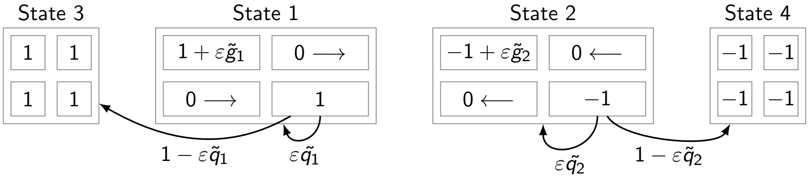
$$\operatorname{val}(\Gamma_\lambda(\varepsilon)) \approx \frac{1}{2} + \varepsilon \left(\frac{1-\lambda}{2} \right)\,.$$



The Kohlberg game

The undiscounted value is

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



The perturbed Kohlberg game

The marginal undiscounted value is

$$D_H \operatorname{val}(\Gamma_0) = \frac{\tilde{g}_1 + \tilde{g}_2 - \tilde{q}_1 + \tilde{q}_2}{4} .$$

So,

$$\operatorname{val}(\Gamma_\lambda(\varepsilon)) \approx 0 + \varepsilon \left(\frac{\tilde{g}_1 + \tilde{g}_2 - \tilde{q}_1 + \tilde{q}_2}{4} \right) .$$

Directional Derivative 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 $\varepsilon \mapsto \operatorname{val}(\Gamma + H\varepsilon)$.

Available in Mathematics of Operations Research
DOI: 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.

$$\gamma_\lambda(\sigma,\tau) := \begin{cases} \mathbb{E}^k_{\sigma,\tau}\left(\lambda \sum_{m\geq 0}(1-\lambda)^m G_m\right) & \lambda > 0 \\ \liminf_{\lambda\to 0^+} \mathbb{E}^k_{\sigma,\tau}\left(\lambda \sum_{m\geq 0}(1-\lambda)^m G_m\right) & \lambda = 0 \end{cases}$$

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.

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,

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

Solan 2003

Theorem

Consider a stochastic game Γ *with* $\lambda \geq 0$*.*
For a perturbation H *where*

the discount factor is not perturbed, $\tilde{\lambda} = 0$*,*
there are no new transitions, $\operatorname{supp}(\tilde{q}) \subseteq \operatorname{supp}(q)$*,*

we have

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

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,*

we have

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

State of affairs

In many reasonable cases, the marginal value exists.

How to compute the marginal value?

Oliu-Barton and Attia 2019

Theorem

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

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

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

Results

Marginal DISCOUNTED value

Theorem

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

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

Sketch proof: Marginal discounted value

Sketch proof.

For every ε, define the matrix game

$$M(\varepsilon) := \Delta_\varepsilon^k - \operatorname{val}(\Gamma + H\varepsilon)\Delta_\varepsilon^0 \,.$$

By Oliu-Barton and Attia, for all ε, we have $\operatorname{val}(M(\varepsilon)) = 0$.
Differentiating, by Mills, we have

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

□

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

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 $\varepsilon \mapsto \mathrm{val}(\Gamma + H\varepsilon)$ *is continuous at zero.*
Let p be a polynomial such that, for all $\varepsilon > 0$ *small enough,*

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

Then,

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

Sketch proof: Marginal undiscounted value

Sketch proof.

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

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

Reordering,

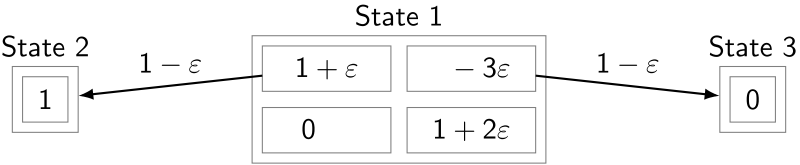
$$D_H \mathrm{val}(\Gamma) \cdot \partial_2 p(0, \mathrm{val}(\Gamma)) = -\partial_1 p(0, \mathrm{val}(\Gamma)) .$$

□

But where does p come from?

Matrices Δ^k and Δ^0, explained

Consider the perturbed Big Match.



A perturbed big Match

Fix a pure stationary strategy $(i, j) = (\text{Top, Left})$.
We compute $\Delta^0_\varepsilon(i, j)$ and $\Delta^k_\varepsilon(i, j)$.

Matrices Δ^k and Δ^0, explained

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,

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

Oliu-Barton and Attia 2019

Theorem

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

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

where Δ^k and Δ^0 are matrices constructed from Γ_λ 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*

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

Note that it may be that $\partial_2 p(0, \mathrm{val}(\Gamma)) = 0$.

Limit and marginal value

We know that

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

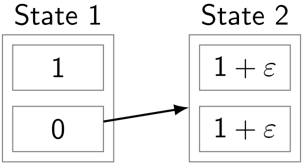
Does this occur with the marignal values?

No, there are examples where

$$\lim_{\lambda \to 0^+} D_H \quad \mathrm{val}(\Gamma_\lambda) \quad \neq \quad D_H \lim_{\lambda \to 0^+} \quad \mathrm{val}(\Gamma_\lambda)\,.$$

Limit and marginal value, example

A perturbed stochastic game where
$\lim_{\lambda \to 0} D_H \text{val}(\Gamma_\lambda) \neq D_H \lim_{\lambda \to 0} \text{val}(\Gamma_\lambda)$.



Thank you!