Value of persistent information

Marcin Pęski, Juuso Toikka

University of Toronto, MIT

October 31, 2017

Introduction

- Informational advantage in repeated interactions
- Trade-off: use information now or save for later.
- Applications: insider trading, arms race, bargaining.
- Original motivation for repeated games with incomplete information (Aumann-Maschler).
- But, persistence of information seems important!
- How exactly?

- Zero-sum stochastic game:
  - payoffs $g(a, b, s)$,
  - actions a (maximizer) and b (minimizer),
  - state s with Markov transitions $P : S \to \Delta S$,
  - maximizer (player 1) observes the state,
  - minimizer (player 2) observes player 1's actions, but not the state,
  - initial beliefs.
- Value $v^{\delta}(\pi; g, P)$, where $\delta < 1$,

$$v(\pi, g, P) = \lim_{\delta \to 1} v^{\delta}(\pi; g, P).$$

The limit value does not depend on π if P is ergodic.

Introduction
Example (Renault, 2006)

- Two states s_1, s_2
  - states stay the same with prob. ρ and change with probability $1-\rho$,
  - the larger ρ, the more persistent is the state,
- Maximizer chooses U or D and the payoffs are

s_1	L	R
U	1	0
D	0	0

,

s_2	L	R
U	0	0
D	0	1

,

- Value is notoriously difficult to compute (Hörner at al, 2010).
- Monotonicity of value?

Definition

Operator Q is *better for maximizer* than P (i.e., $Q \succeq P$) if $v(g, Q) \geq v(g, P)$ for each game g.

- Goal: Characterize relation $P \preceq Q$.
- Idea:
  - persistence is bad for maximizer,
  - the above relation should capture some notion of persistence.

- Stochastic games
  - vs. repeated games with incomplete information (i.e., Aumann-Maschler)
  - Stochastic zero-sum games with Markovian private information Renault 06, Neyman 08, Hörner-Rosenborg-Solan-Vieille 10
- Comparison of information literature: (Blackwell 1953, Mertens-Gossner 01, Peski 08).
  - intuition: more information (in the Blackwell sense) is better for the minimizer,
  - here: more information means that P is more persistent.
  - however, it is difficult to separate the information and the payoff effects of transitions.
- Applications:
  - zero-sum stochastic games: value is monotonic in (Hörner at al, 2010),
  - individual rationality constraint in repeated games with Markov types (Athey-Bagwell 08, Escobar-Toikka 13, Hörner-Takahashi-Vieille 15),
  - one long-run vs. many short run players (zero-sum).

1. Introduction
2. Notations and definitions
3. Value of stochastic game
4. Comparison of operators (characterization of order $\preceq$)
5. Characterizations and corollaries
6. Extensions

Notations and Definitions

Beliefs

- $p, q \in \Delta S$ - space of (minimizer's) beliefs,
  - prior beliefs in period t: beliefs *before* the actions are chosen (and information revealed),
  - posterior beliefs in period t: beliefs *after* the actions are chosen,
- $\mu, \nu \in \Delta^2 S = \Delta(\Delta S)$ - distributions over beliefs,

- $P : \Delta S \to \Delta S$ Markov operator,
- p are posteriors today $\rightrightarrows$ Pp are prior beliefs tomorrow,
- μ is a distribution of posteriors today $\rightrightarrows$ $P\mu$ is a distribution of priors tomorrow, where

$$(Pp)(s) = \sum_{s'} p(s) P(s|s'),$$

$$(P\mu)(A) = \mu\{q : Pq \in A\} = \mu\left(P^{-1}A\right).$$

- *No persistence*: D_π - i.i.d. draws from distribution $\pi \in \Delta S$,
- **Persistent information** : P is aperiodic and irreducible,
  - $P^n \pi \to \pi_P$, where π_P is unique stationary distribution,
  - value $v(g, P)$ does not depend on the initial distribution.
- *Permanent information*: $P = I$,
  - repeated game with incomplete information,
- *Alternating case*: $|S| = 2$ and $A = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$,
  - $(Ap)_s = 1 - p_s$.

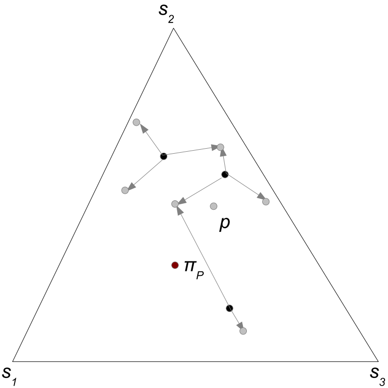
Definition

A mean preserving spread is a measurable $m : \Delta S \to \Delta^2 S$ such that

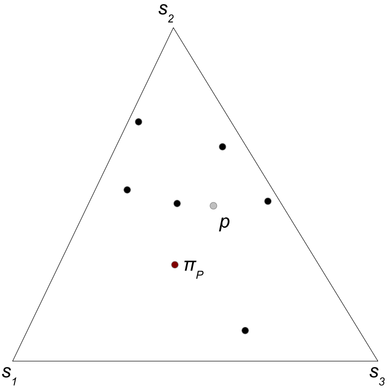
$$Em(.|q) = q \text{ for each } q$$

Notations and Definitions

Mean preserving spread



Distribution of priors



Distribution of posteriors

ν is a mean preserving spread of μ.

Definition

ν is a *mean preserving spread* of μ, if there exists a m.p.s. $m : \Delta S \to \Delta^2 S$ such that

$$\nu = \mu \star m,$$

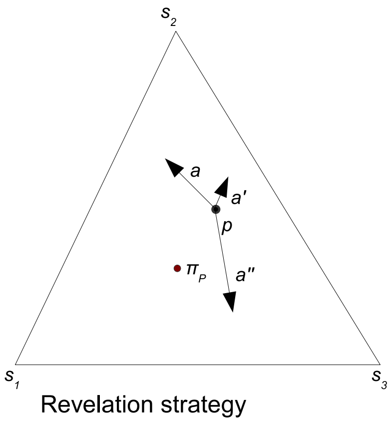
or

$$\nu(dp) = \int m(dp|q)\, d\mu(q).$$

- We write $\mu \leq^B \nu$.
- If $\mu \leq^B \nu$, then $P\mu \leq^B P\nu$.

Value of stochastic zero-sum game

Revelation of information



- Maximizer's Markov strategy $\alpha : \Delta S \times S \to \Delta A$
- Decompose the maximizer's behavior into
  - revelation strategy m,
  - actions that reveal information: $a : \operatorname{supp} m(p) \to A$.
- For each $m \in \Delta^2 S$, define

$$\hat{g}(m) = \max_{a:\operatorname{supp} m(p) \to A} \min_{\beta \in \Delta B} \int g(a(q), \beta, q)\, dm(q).$$

  - payoff in a zero-sum game, in which minimizer chooses β and the maximizer chooses actions that respect the revelation m.

Value of stochastic zero-sum game

Distributions over beliefs μ: stationarity

- Elements of equilibrium
  - $m : \Delta S \to \Delta^2 S$ (Markov) revelation strategy
  - μ stationary distribution over priors (i.e., beliefs at the beginning of the period),
  - the average payoff is equal to

$$\int \hat{g}\left(m\left(p\right)\right) d\mu\left(p\right).$$

- Stationary distribution μ over priors:
  - $\mu \star m$ is a distribution over posteriors, and
  - $P\left(\mu \star m\right)$ is a distribution over prior beliefs in the next period,
  - because μ is stationary:

$$P\left(\mu \star m\right) = \mu.$$

Value of stochastic zero-sum game

Theorem

For each g, each ergodic P, each stationary distribution π of P,

$$v(g, P) = \max_{\mu, m:\ P(\mu \star m) \leq^B \mu} \int \hat{g}(m(p))\, d\mu(p).$$

- when $\delta \to 1$, the value converges to the average revelation payoff over the stationary distribution,
- the second inequality can be replaced by equality
- proof: "stationarity" of the problem.

Value of stochastic zero-sum game

$$v(g, P) = \max_{\mu, m:\ P(\mu \star m) \leq^B \mu} \int \hat{g}(m(p))\, d\mu(p).$$

- maximization of a functional that depends on g, but not P
- over the set of (μ, m) that depends on P but not on g.

Comparison of operators

- Operator Q is *better for maximizer* than P (i.e., $Q \succeq P$) if $v(g, Q) \geq v(g, P)$ for each game g.

Theorem

Let P, Q be ergodic . The following are equivalent.
(a) $P \preceq Q$.
(b) for all μ, m,

$$P(\mu * m) \leq^B \mu \Longrightarrow Q(\mu * m) \leq^B \mu,$$

(c) for all ν,

$$P\nu \leq^B \nu \Longrightarrow Q\nu \leq^B P\nu.$$

Comparison of operators

$$P\nu \leq^B \nu \Longrightarrow Q\nu \leq^B P\nu.$$

- Fixed point-ish flavor.
- Here, ν is a distribution of posteriors (i.e., $\mu * m$),
  - $P\nu \leq^B \nu$ means that ν is "stabilizable",
  - $Q\nu \leq^B P\nu$ is exactly the condition for next period's priors to be more informative under P than under Q,
- For each "stabilizable" end-of-the-period information ν, the next-period information is worse under Q than under P,
  - "P leads to smaller loss of information"; "Q adds more noise"
  - that is, information is *more persistent* under P than under Q.

- Proof: (b) <-> (c) easy,
- Proof: (b) -> (a) immediate from the characterization of value.

- Proof: not (b) -> not (a)
  - suppose that $P(\mu_0 * m_0) \leq^B \mu_0$ and $Q(\mu_0 * m_0) \nleq^B \mu_0$.
  - Blackwell: there exists a concave function $f : \Delta S \to R$ st.

  $$\mu_0[f] - Q(\mu_0 * m_0)[f] > 0.$$

  Use f to construct g and $\hat{g}$ st.

  $$\int \hat{g}(m(p))\, d\mu(p) = \mu[f] - Q(\mu * m)[f].$$

  It follows that

  $$\upsilon(g, P) \geq \int \hat{g}(m_0(p))\, d\mu_0(p) > 0.$$

  - Because f is concave,

  $$\forall_{(\mu,m)} \text{ st. } _{Q(\mu * m) \leq^B \mu} \mu[f] - Q(\mu * m)[f] < 0,$$

  Hence, $\upsilon(g, Q) < 0$.

- W.l.o.g. there is a finite set L of functions $l : S \to R$.

$$f(p) = \min_{l \in L} \sum p(s)\, l(s),$$

- Let $A = B = L$, and for each $a, b \in L$,

$$g(a, b, s) = b(s) - \sum_{s'} Q(s'|s)\, a(s').$$

- We show that

$$\int \left(\min_{\beta \in \Delta B} \int \left(\max_a g(a, \beta, q) \right) dm(q|p) \right) d\mu(p) = \mu[f] - Q(\mu * m)[f].$$

- We have

$$\max_{\alpha} \sum_{s} q(s)\, g(a, \beta, s)$$
$$= \sum_{s} \beta(s)\, q(s) - \min_{\alpha} \sum_{s} q(s) \sum_{s'} Q\left(s' | s\right) a(s)$$
$$= \sum_{s} \beta(s)\, q(s) - f(Qq),$$

and

$$\min_{\beta \in \Delta B} \left(\int g^{*}(\beta, q)\, dm(|qp) \right)$$
$$= \left(\min_{\beta \in \Delta B} \sum_{s} \beta(s)\, p(s) \right) - \left(\int f(Qq)\, dm(q|p) \right)$$
$$= f(p) - \left(\int f(Qq)\, dm(q|p) \right).$$

Partial characterizations

Order properties

Corollary

Let P, Q, Q' be ergodic.

- *If $P \preceq Q$ or $Q \preceq P$, then $\pi_P = \pi_Q$.*
- *If $P \preceq Q$ and $Q \preceq P$, then $P = Q$.*
- *If $P \preceq Q$ and $P \preceq Q'$, then $P \preceq \lambda Q + (1 - \lambda) Q'$.*

Partial characterizations

Simple (but not complete) characterization

- For each $\alpha \in \mathcal{A} := \{(\alpha_1, \alpha_2, ..., \alpha_\infty) : \alpha_i \geq 0, \sum \alpha_i = 1\}$, let

$$P^\alpha := \sum_{k=1}^{\infty} \alpha_k P^k$$

.

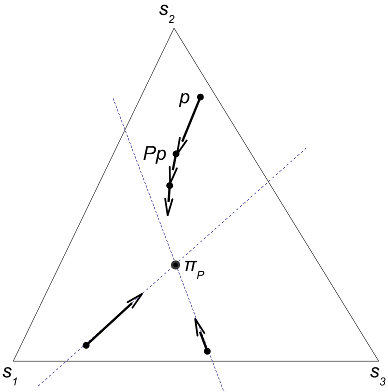
Theorem

For each ergodic P, Q:

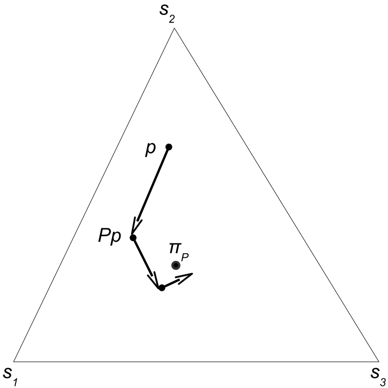
1. *Sufficient condition: If $Q = P^\alpha$ for some $\alpha \in \mathcal{A}$, then $P \preceq Q$.*
2. *Necessary condition: If $P \preceq Q$, then, for each p, there exists $\alpha_p \in \mathcal{A}$ such that $Qp = P^{\alpha_p} p$.*
3. *If P has purely real eigenvalues, then the necessary and the sufficient conditions are equivalent.*

Partial characterizations

Operator $P : \Delta S \to \Delta S$



Real eigenvalues



Complex eigenvalues

Action of operator P

- For general operators, the sufficient is not necessary.
  - we do not know whether the necessary condition is sufficient,
  - but the necessary condition is not really easier than our full characterization.
- Persistence of information:
  - $P \preceq P^n$,
  - for each $\alpha \in [0, 1]$, if π is P-invariant:

$$P \preceq \alpha P + (1 - \alpha) D_\pi .$$

Corollary

$P \preceq D_\pi$ for each P-invariant π.

- D_π is the best operator.

- There is no worst operator (unless $|S| = 2$).
- In particular, permanent case ($P = I$) is not the worst.

Corollary

If $P \neq \alpha I + (1-\alpha) D_\pi$ for some $\alpha \in (0,1)$ and $\pi \in \Delta S$, then, there exists game g such that

$$v(\pi, g, I) > v(g, P).$$

.

Example

(s_1, s_2, s_3)	L	R
U	-2,0,3	0,-2,3
D	-1,1,0	1,-1,0

- Suppose $P = I$.
- Minimizer play L if $p_1 \geq p_2$ and R if $p_2 \geq p_1$.
- Maximizer plays U if $s = s_3$ and to play D otherwise.
- Minimizer only ever learns $\{s_1, s_2\}$ or $\{s_3\}$.
- If $\pi = (\frac{1}{3}, \frac{1}{3}, \frac{1}{3})$, then the value of the game is $\frac{2}{3} \cdot 0 + \frac{1}{3} \cdot 3 = 1$.
- The argument also applies for each $P = \alpha I + (1 - \alpha) D_\pi$.

Partial characterizations

No worst operator:Example

Example

(s_1, s_2, s_3)	L	R
U	-2,0,3	0,-2,3
D	-1,1,0	1,-1,0

- Suppose that

$$P = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \frac{1}{2} & \frac{1}{2} \\ 0 & \frac{1}{2} & \frac{1}{2} \end{bmatrix},$$

- $\pi = (\frac{1}{3}, \frac{1}{3}, \frac{1}{3})$ is P-invariant.
- If maximizer reveals s_3, then, the next period belief is that s_2 is much more likely than s_1 - and better payoff for the minimizer for s_2.
- If maximizer does not reveal s_3, she does not benefit from payoff 3 at this state.

Partial characterizations

$|S| = 2$

- Suppose that $|S| = 2$, and ergodic distr. is $\pi = \left(\frac{1}{2}, \frac{1}{2}\right)$.
- Then, each operator is $P(\rho) = \begin{bmatrix} \rho & 1-\rho \\ 1-\rho & \rho \end{bmatrix}$,
  - larger $\rho > \frac{1}{2}$ means more persistence,
  - smaller $\rho < \frac{1}{2}$ means more alternating

Partial characterizations

$|S| = 2$

Corollary

If $\frac{1}{2} \leq \xi \leq \rho$, then

$$D_\pi = M\left(\frac{1}{2}\right) \succeq M(\xi) \succeq M(\rho).$$

If $\rho < \frac{1}{2}$ and $\frac{1}{2} + \frac{1}{2}(2\rho - 1)^2 \geq \xi \geq \rho \geq 0$, then

$$D_\pi = M\left(\frac{1}{2}\right) \succeq M(\xi) \succeq M(\rho) \succeq M(0) = A.$$

- monotonicity of value in $\rho > \frac{1}{2}$ in Renualt's example (for any game).

Partial characterizations

$|S| = 2$

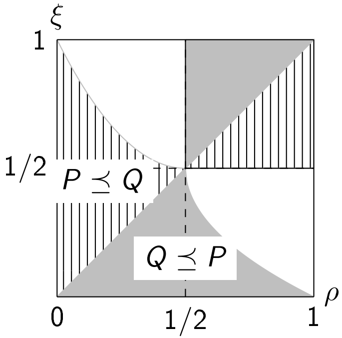


Figure: $P = M(\rho)$ and $Q = M(\xi)$.

Extensions

- Main results (i.e. value and comparison) extend to
  - non-ergodic operators,
  - public signal,
  - imperfect monitoring.
- Partial characterization (sufficient condition) for finite discount factor.

- (Discounted) distribution over prior beliefs if no information is ever revealed

$$P^{\infty,\delta}\pi = \sum_{k=0}^{\infty} (1-\delta)\,\delta^k \mathrm{Dirac}_{P^k\pi}.$$

- Limit $P^{\infty}\pi = \lim_{\delta\to 1} P^{\infty,\delta}\pi$ always exists:
  - if P is ergodic, then $P^{\infty}\pi = \mathrm{Dirac}_{\pi^*}$ (does not depend on π),
  - if $P = I$, then $P^{\infty}\pi = \mathrm{Dirac}_{\pi}$,
  - if $P = A$, then $P^{\infty}\pi = \frac{1}{2}\mathrm{Dirac}_{\pi} + \frac{1}{2}\mathrm{Dirac}_{-\pi}$.
- In equilibrium, stationary distribution μ must "respect" the initial information of the minimizer,
  - it must be that

$$P^{\infty}\pi \leq^{B} \mu.$$

Theorem

*(**Value of the stochastic zero-sum game**) For each g, each P, each stationary distribution π of P,*

$$v(\pi, g, P) = \max_{\mu, m : P^{\infty}\pi \leq^{B} \mu \text{ and } P(\mu \star m) \leq^{B} \mu} \int \hat{g}(m(p))\, d\mu(p).$$

$$v(\pi, g, P) = \max_{\mu, m : P^{\infty}\pi \leq^{B} \mu \text{ and } P(\mu \star m) \leq^{B} \mu} \int \hat{g}(m(p))\, d\mu(p).$$

- ergodic P:

$$v(g, P) = \max_{\mu, m \text{ st. } P(\mu \star m) \leq^{B} \mu} \int \hat{g}(m(p))\, d\mu(p),$$

- repeated game with incomplete information,

$$\begin{aligned} v(\pi, g, I) &= \max_{\mu : E\mu = \pi} \int \hat{g}(\mathrm{Dirac}_p)\, d\mu(p), \\ &= (\mathrm{cav}\hat{g})(\pi) \end{aligned}$$

- repeated game with incomplete information and alternating state

$$v(\pi, g, A) = \left(\mathrm{cav}\left(\frac{1}{2}\hat{g} + \frac{1}{2}\hat{g}^{-}\right)\right)(\pi).$$

- Public signal observed before actions (beginning of the period) $F : S \to \Delta Z$,
  - $n^F : \Delta S \to \Delta^2 S$.
- **Value**: For every game g and ergodic P,

$$v(g, P, F) = \max_{\mu, m : P(\mu * m) * n^F \leq^B \mu} \int \hat{g}(m(p)) d\mu(p).$$

- **Comparison**:$(P, F) \preceq_{Pub} (Q, G)$.
  - for every $(\mu, m) \in \Delta^2 S \times \mathcal{M}$ such that $P(\mu * m) * n^F \leq^B \mu$, we have $Q(\mu * m) * n^G \leq^B \mu$.
  - for every $\nu \in \Delta^2 S$ such that $P\nu * n^F \leq^B \nu$, we have $Q\nu * n^G \leq^B P\nu * n^F$.

- Monitoring: $F_a \in \Delta Z$,
  - signal z (and not action a) is observed.
- **Value**: the same, if we replace $\hat{g}$ by

$$\hat{g}_F(\nu) = \min_{\beta \in \Delta B} \max_{\alpha \text{ st. } m^{\alpha,p,F} \leq^B \nu} \sum_s p(s)\, g(\alpha(s), \beta, p).$$

- **Comparison**: $P \preceq_{Im} Q$ if for each game g and each imperfect monitoring F,

$$v(\pi; g, F, P) \leq v(\pi; g, F, Q).$$

  - the Comparison Theorem holds *verbatim*.

Theorem

For any zero-sum game, any ergodic P, any zero-sum g, any discount factor, any $\alpha \in (0,1)$, if π is invariant dist. of P, then

$$v^{\delta}(\pi; g, P) \leq v^{\delta}(\pi; g, \alpha P + (1-\alpha) D_{\pi}).$$

- $P \prec_{\delta,\pi} \alpha P + (1-\alpha) D_{\pi}$.

Conclusions

- We analyze stochastic games with incomplete information.
  - formula for the value,
  - comparison of operators with respect to the value of the game
- More persistence (in some sense) is good for the minimizer,
- The main result is interpretable, and easy to use in proofs, but not in calculations.