Robust strategies in an evolving unknown environment

Jérôme Renault
TSE, Université Toulouse 1 & GDR Jeux

Sestri-Levante, september 2014

In a dynamic environment, a decision maker wants to minimize his cost against all possible realizations of the adversary. He is a small player and does not influence the environment.

General problem

Data: - S is the set of states of the environment. There is a given family $\mathscr{F}$ of possible stochastic process over S.
- J is the set of actions that the decision-maker can take at any period, and $g : S \times J \longrightarrow I\!R$ is a given cost function.
- There is a set of signals U and observation functions $l_t : S^t \longrightarrow \Delta(U)$ for each period t. Assume S, J, U finite.

Progress: One μ in $\mathscr{F}$ is the true process, and a sequence $(s_1, ..., s_t, ...)$ is generated according to μ. At the beginning of each period t, the DM observes a signal u_t generated with $l_t(s_1, ..., s_t)$, then chooses j_t and bears a cost $g(s_t, j_t)$.

Strategy for the DM: $\tau = (\tau_t)_{t \geq 1}$, with $\tau_t : (U \times J)^{t-1} \times U \longrightarrow \Delta(J)$. $\tau_t(u_1, j_1, ..., u_t)$ is the probability over J used by the DM at period t after observing $u_1, ..., u_t$ and having played $j_1, ..., j_{t-1}$. A process μ over S and a strategy τ jointly generate a process over $S \times U \times J$.

Dynamic Optimization Problem:
$v := \inf\{w \in I\!R, \exists \tau, \forall \mu \in \mathscr{F}, \limsup_T I\!E_{\mu,\tau}\left(\frac{1}{T}\sum_{t=1}^{T} g(s_t, j_t)\right) \leq w\}$.
Problem: Compute v ? Construct $(\varepsilon)-$ optimal strategies τ ?

General problem

Data: - S is the set of states of the environment. There is a given family $\mathscr{F}$ of possible stochastic process over S.
- J is the set of actions that the decision-maker can take at any period, and $g : S \times J \longrightarrow I\!R$ is a given cost function.
- There is a set of signals U and observation functions $l_t : S^t \longrightarrow \Delta(U)$ for each period t. Assume S, J, U finite.

Progress: One μ in $\mathscr{F}$ is the true process, and a sequence $(s_1, ..., s_t, ...)$ is generated according to μ. At the beginning of each period t, the DM observes a signal u_t generated with $l_t(s_1, ..., s_t)$, then chooses j_t and bears a cost $g(s_t, j_t)$.

Strategy for the DM: $\tau = (\tau_t)_{t \geq 1}$, with $\tau_t : (U \times J)^{t-1} \times U \longrightarrow \Delta(J)$. $\tau_t(u_1, j_1, ..., u_t)$ is the probability over J used by the DM at period t after observing $u_1,...,u_t$ and having played $j_1, ..., j_{t-1}$. A process μ over S and a strategy τ jointly generate a process over $S \times U \times J$.

Dynamic Optimization Problem:
$v := \inf\{w \in I\!R, \exists \tau, \forall \mu \in \mathscr{F}, \limsup_T I\!E_{\mu,\tau}\left(\frac{1}{T}\sum_{t=1}^{T} g(s_t, j_t)\right) \leq w\}$.
Problem: Compute v ? Construct $(\varepsilon)-$ optimal strategies τ ?

Remark: Many variants for the overall cost criterium.

Possible examples:

1) You want to invest on energy ressources, S contains the prices of gaz, oil, coal, wind energy... You are not influencing the evolution of these prices.

2) A state in S gives the distribution of fishes over the planet (Sestri Levante area) and the position of other fishing boats. The DM has to choose where to go fishing at each period.

3) S is the size of the ozone hole (or the temperature, or the size of the Amazon rainforest). You want to predict it: $g(s,j)$ is increasing in $|s-j|$.

4) You sell a few products online, and need to decide products and prices for your website. You are not influencing the global demand, nor the behavior of the major companies.

Remark: Many variants for the overall cost criterium.

Possible examples:
1) You want to invest on energy ressources, S contains the prices of gaz, oil, coal, wind energy... You are not influencing the evolution of these prices.
2) A state in S gives the distribution of fishes over the planet (Sestri Levante area) and the position of other fishing boats. The DM has to choose where to go fishing at each period.
3) S is the size of the ozone hole (or the temperature, or the size of the Amazon rainforest). You want to predict it: $g(s,j)$ is increasing in $|s-j|$.
4) You sell a few products online, and need to decide products and prices for your website. You are not influencing the global demand, nor the behavior of the major companies.

Trivial cases

1) In each period, the signal u_t determines the state s_t: Choose for each t an action j_t minimizing $g(s_t, j)$.

2) There is a single process μ^* in $\mathscr{F}$: Choose at each period t an action j_t minimizing $I\!E_{\mu^*}\left(g(s_t, j)\middle| u_1, ... u_t\right)$.

3) $\mathscr{F}$ is the set of all processes on S, and $l_t(s_1, ..., s_t)$ does not depend on s_t: Play in each period an optimal strategy of the column player in the matrix game with payoffs $(g(s,t))_{s,t}$.
The value is
$v = \min_{y \in \Delta(J)} \max_{s \in S} \sum_j y_j g(s,j) = \max_{x \in \Delta(S)} \min_{j \in J} \sum_s x_s g(s,j)$.

	j_1	j_2
s_1	3	0
s_2	0	2

Here, play $\frac{2}{5} j_1 + \frac{3}{5} j_2$ at each period. Value $v = \frac{6}{5}$.

4) $\mathscr{F}$ is the set of all processes on S, and $l_t(s_1,...,s_t)$ only depends on s_t. Play in each period t a mixed optimal strategy of player 2 in the finite game where: first, player 1 chooses s in S then u is selected w.r.t. $l_t(s)$ and observed by player 2, finally player 2 chooses j in J and has cost $g(s,j)$.

The optimal cost for period t is:

$$v_t = \min_{y:U\to\Delta(J)} \max_{s\in S} \sum_{u,j} l_t(u|s)y(j|u)g(s,j) = \max_{x\in\Delta(S)} \min_{y:U\to J} \sum_{s,u} x_s l_t(u|s)g(s,y(u)),$$

and the overall value is: $v = \limsup_T \frac{1}{T}\sum_{t=1}^T v_t$.

What it is *not* about:

	j_1	j_2	j_3
s_1	1	1	1
s_2	1	1	0
s_3	1	0	1

Suppose $\mathscr{F}$ contains 3 deterministic processes: $s_1s_1s_1...$, $s_2s_2s_2...$ and $s_3s_3s_3....$ And that the signal reveals the previous state.
$v=1$, and any τ is optimal

A particular case: Controlled Markov Chain

We assume it from now on:
1) $\mathscr{F}$ is the set of possible processes of a controlled Markov chain. Assume $S = K \times I$, where K is a set of physical parameters, and I is a set of actions for the adversary. An initial probability p_0 on K, and a transition $q : K \times I \longrightarrow \Delta(K)$ are given and known to the DM. The adversary observes the parameters and can choose any sequence $i_1,...,i_t,...$ depending on the parameters. A strategy of the adversary is thus a sequence $\sigma = (\sigma_t)_{t\geq 1}$, with $\sigma_t : (K \times I)^{t-1} \times K \longrightarrow \Delta(I)$. Every σ induces a process μ_σ over $(K \times I)^\infty$, and $\mathscr{F}$ is the set of all μ_σ.
2) The DM only observes at the beginning of period t the previous action i_{t-1} (but not the current action nor the parameters).

We obtain a dynamic game where:
The adversary (player 1, maximizer) observes the parameters but not the actions of the DM.
The DM (player 2, minimizer) observes the actions of player 1 but not the parameters.

A particular case: Controlled Markov Chain

We assume it from now on:
1) $\mathscr{F}$ is the set of possible processes of a controlled Markov chain. Assume $S = K \times I$, where K is a set of physical parameters, and I is a set of actions for the adversary. An initial probability p_0 on K, and a transition $q : K \times I \longrightarrow \Delta(K)$ are given and known to the DM. The adversary observes the parameters and can choose any sequence $i_1,...,i_t,...$ depending on the parameters. A strategy of the adversary is thus a sequence $\sigma = (\sigma_t)_{t \geq 1}$, with $\sigma_t : (K \times I)^{t-1} \times K \longrightarrow \Delta(I)$. Every σ induces a process μ_σ over $(K \times I)^\infty$, and $\mathscr{F}$ is the set of all μ_σ.
2) The DM only observes at the beginning of period t the previous action i_{t-1} (but not the current action nor the parameters).

We obtain a dynamic game where:
The adversary (player 1, maximizer) observes the parameters but not the actions of the DM.
The DM (player 2, minimizer) observes the actions of player 1 but not the parameters.

Very particular case: Constant parameter

Here, we assume moreover that the parameter k is selected according to p_0 and remains fixed throughout.

Example 1: $p_0 = (1/2, 1/2)$

$$\begin{array}{c} \\ i_1 \\ i_2 \end{array} \overset{\begin{array}{ccc} j_1 & j_2 & j_3 \end{array}}{\underset{k_1}{\begin{pmatrix} 4 & 0 & 2 \\ 4 & 0 & -2 \end{pmatrix}}} \qquad \begin{array}{c} \\ i_1 \\ i_2 \end{array} \overset{\begin{array}{ccc} j_1 & j_2 & j_3 \end{array}}{\underset{k_2}{\begin{pmatrix} 0 & 4 & -2 \\ 0 & 4 & 2 \end{pmatrix}}}$$

Prop: There exists a strategy τ of the DM such that for all σ and all k:

$$\limsup_T I\!E_{\sigma,\tau}\left(\frac{1}{T}\sum_{t=1}^{T} g(s_t, j_t)\Big| k\right) \leq 1.$$

(Blackwell approachability strategy, Kohlberg) So $v \leq 1$.

Very particular case: Constant parameter

Here, we assume moreover that the parameter k is selected according to p_0 and remains fixed throughout.
Example 1: $p_0 = (1/2, 1/2)$

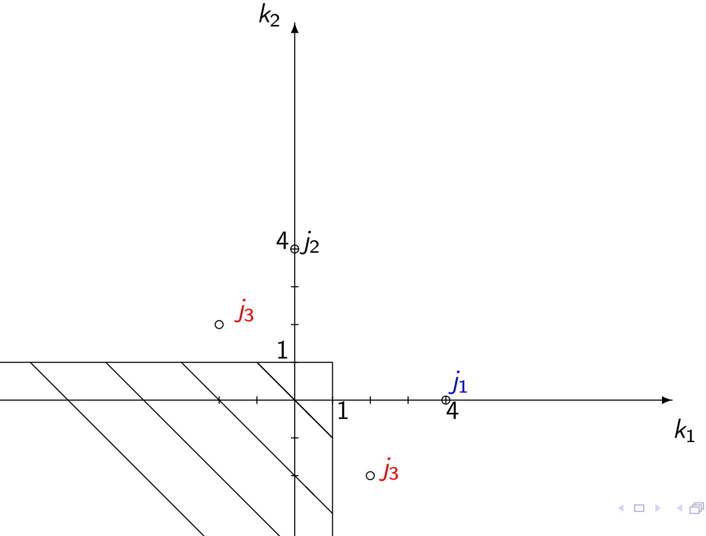
$$\begin{array}{cc} & \begin{array}{ccc} j_1 & j_2 & j_3 \end{array} \\ \begin{array}{c} i_1 \\ i_2 \end{array} & \begin{pmatrix} 4 & 0 & 2 \\ 4 & 0 & -2 \end{pmatrix} \\ & k_1 \end{array} \qquad \begin{array}{cc} & \begin{array}{ccc} j_1 & j_2 & j_3 \end{array} \\ \begin{array}{c} i_1 \\ i_2 \end{array} & \begin{pmatrix} 0 & 4 & -2 \\ 0 & 4 & 2 \end{pmatrix} \\ & k_2 \end{array}$$

Prop: There exists a strategy τ of the DM such that for all σ and all k:

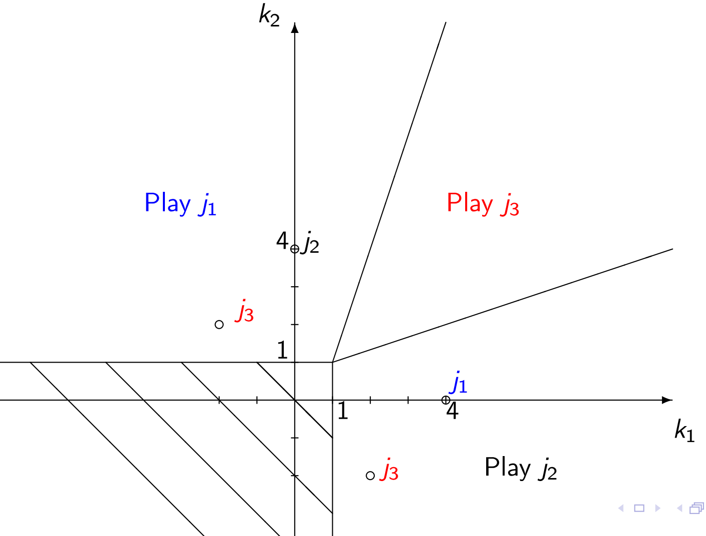
$$\limsup_T I\!E_{\sigma,\tau}\left(\frac{1}{T}\sum_{t=1}^{T} g(s_t, j_t) \Big| k\right) \leq 1.$$

(Blackwell approachability strategy, Kohlberg) So $v \leq 1$.

$$\begin{pmatrix} 4 & 0 & 2 \\ 4 & 0 & -2 \end{pmatrix} \qquad \begin{pmatrix} 0 & 4 & -2 \\ 0 & 4 & 2 \end{pmatrix}$$



$$\begin{pmatrix} 4 & 0 & 2 \\ 4 & 0 & -2 \end{pmatrix} \qquad \begin{pmatrix} 0 & 4 & -2 \\ 0 & 4 & 2 \end{pmatrix}$$



Prop (Blackwell): Let C be a closed convex set of $I\!R^K$, and $(g_t)_t$ a bounded sequence in $I\!R^K$. Write $\bar{g}_T = \frac{1}{T}\sum_{t=1}^{T} g_t$, and assume that for each T,

$$< \bar{g}_T - p_C(\bar{g}_T), g_{T+1} - p_C(\bar{g}_T) > \leq 0.$$

Then

$$d(\bar{g}_T, C) \xrightarrow[T\to\infty]{} 0.$$

Can the DM have a cost lower than 1 ? No, so $v = 1$, and the above τ is optimal.

For 2 reasons. Define $r(p) = \mathrm{Val}(\sum_k p^k G^k)$ for each p in $\Delta(K)$.

Prop (Blackwell): Let C be a closed convex set of $I\!R^K$, and $(g_t)_t$ a bounded sequence in $I\!R^K$. Write $\bar{g}_T = \frac{1}{T}\sum_{t=1}^{T} g_t$, and assume that for each T,

$$< \bar{g}_T - p_C(\bar{g}_T), g_{T+1} - p_C(\bar{g}_T) > \leq 0.$$

Then

$$d(\bar{g}_T, C) \xrightarrow[T\to\infty]{} 0.$$

Can the DM have a cost lower than 1 ? No, so $v = 1$, and the above τ is optimal.

For 2 reasons. Define $r(p) = \mathrm{Val}(\sum_k p^k G^k)$ for each p in $\Delta(K)$.

Prop (Blackwell): Let C be a closed convex set of $I\!R^K$, and $(g_t)_t$ a bounded sequence in $I\!R^K$. Write $\bar{g}_T = \frac{1}{T}\sum_{t=1}^{T} g_t$, and assume that for each T,

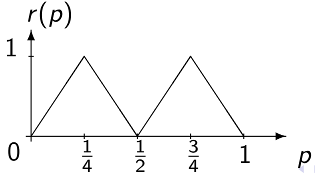
$$< \bar{g}_T - p_C(\bar{g}_T), g_{T+1} - p_C(\bar{g}_T) > \leq 0.$$

Then

$$d(\bar{g}_T, C) \xrightarrow[T \to \infty]{} 0.$$

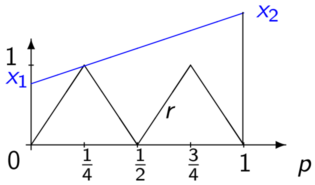
Can the DM have a cost lower than 1 ? No, so $v = 1$, and the above τ is optimal.

For 2 reasons. Define $r(p) = \mathrm{Val}(\sum_k p^k G^k)$ for each p in $\Delta(K)$.

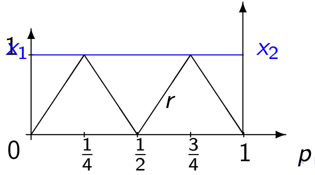


Reason 1) (Kohlberg) Prop: Let x be a vector in $I\!R^K$.

$$\exists \tau \, \forall \sigma \, \limsup_T I\!E_{\sigma,\tau}\left(\frac{1}{T}\sum_{t=1}^{T} g(s_t, j_t)\Big| k\right) \leq x_k \Longleftrightarrow \forall p \in \Delta(K) < x, p > \geq r(p)$$



Best choice, since the initial proba is 1/2: $x_1 = x_2 = 1$.



Reason 2 (Aumann Maschler): Dual approach. There exists a strategy σ of the adversary such that for all τ and all T:

$$\mathbb{E}_{\sigma,\tau}\left(\frac{1}{T}\sum_{t=1}^{T} g(s_t, j_t)\right) \geq 1.$$

$$\begin{array}{c} \\ i_1 \\ i_2 \\ \\ \end{array} \begin{array}{c} \begin{array}{ccc} j_1 & j_2 & j_3 \end{array} \\ \left(\begin{array}{ccc} 4 & 0 & 2 \\ 4 & 0 & -2 \end{array}\right) \\ k_1 \end{array} \qquad \begin{array}{c} \\ i_1 \\ i_2 \\ \\ \end{array} \begin{array}{c} \begin{array}{ccc} j_1 & j_2 & j_3 \end{array} \\ \left(\begin{array}{ccc} 0 & 4 & -2 \\ 0 & 4 & 2 \end{array}\right) \\ k_2 \end{array}$$

Optimal strategy for player 1: if k_1 (resp. k_2): with proba 3/4 (resp.1/4) play always i_1 and with proba 1/4 (resp.3/4) play always i_2.
Given i_1, the belief of player 2 on the state is $3/4k_1 + 1/4k_2$, and the corresponding cost matrix is $\left(\begin{array}{ccc} 3 & 1 & 1 \\ 3 & 1 & -1 \end{array}\right)$, and i_1 gives a cost at least 1. Similarly given i_2, the belief of player 2 on the state is $1/4k_1 + 3/4k_2$, and the corresponding cost matrix is $\left(\begin{array}{ccc} 1 & 3 & -1 \\ 1 & 3 & 1 \end{array}\right)$, and i_2 gives a cost at least 1.

Reason 2 (Aumann Maschler): Dual approach. There exists a strategy σ of the adversary such that for all τ and all T:

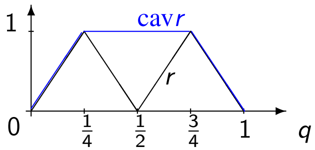
$$\mathbb{E}_{\sigma,\tau}\left(\frac{1}{T}\sum_{t=1}^{T} g(s_t,j_t)\right) \geq 1.$$

$$\underset{k_1}{\begin{array}{c c} & \begin{array}{ccc} j_1 & j_2 & j_3 \end{array} \\ \begin{array}{c} i_1 \\ i_2 \end{array} & \left(\begin{array}{ccc} 4 & 0 & 2 \\ 4 & 0 & -2 \end{array}\right) \end{array}} \qquad \underset{k_2}{\begin{array}{c c} & \begin{array}{ccc} j_1 & j_2 & j_3 \end{array} \\ \begin{array}{c} i_1 \\ i_2 \end{array} & \left(\begin{array}{ccc} 0 & 4 & -2 \\ 0 & 4 & 2 \end{array}\right) \end{array}}$$

Optimal strategy for player 1: if k_1 (resp. k_2): with proba 3/4 (resp.1/4) play always i_1 and with proba 1/4 (resp.3/4) play always i_2.
Given i_1, the belief of player 2 on the state is $3/4k_1 + 1/4k_2$, and the corresponding cost matrix is $\left(\begin{array}{ccc} 3 & 1 & 1 \\ 3 & 1 & -1 \end{array}\right)$, and i_1 gives a cost at least 1. Similarly given i_2, the belief of player 2 on the state is $1/4k_1 + 3/4k_2$, and the corresponding cost matrix is $\left(\begin{array}{ccc} 1 & 3 & -1 \\ 1 & 3 & 1 \end{array}\right)$, and i_2 gives a cost at least 1.

Generalization to any sets K, I, J and cost matrices. Well known (Aumann Maschler 1966):

$$v = \mathrm{cav}\, r(p_0) = \inf\{w(p_0), w \text{ concave}, w \geq r\}.$$



Generalization to signals for the DM.
Remark: same value if the adversary is restricted to play *mixtures* of *stationary* strategies.

General case of controlled Markov chain

The state may move according to a Markov chain controlled by the adversary.

Example 2: $p_0 = (1/2, 1/2)$ is the initial distribution.

$$\begin{array}{cc} & \begin{array}{cc} j_1 & j_2 \end{array} \\ \begin{array}{c} i_1 \\ i_2 \end{array} & \left(\begin{array}{cc} 1_{k_1} & 0_{k_1} \\ 0_{p_0} & 0_{p_0} \end{array} \right) \\ & \text{state } k_1 \end{array} \qquad \begin{array}{cc} & \begin{array}{cc} j_1 & j_2 \end{array} \\ \begin{array}{c} i_1 \\ i_2 \end{array} & \left(\begin{array}{cc} 0_{p_0} & 0_{p_0} \\ 0_{p_0} & 1_{p_0} \end{array} \right) \\ & \text{state } k_2 \end{array}$$

For example, if the state is k_1 and (i_2, j_1) is played, then the cost is 0, and the next state is selected according to p_0.

State variable: Belief of player 2 on the current state in K. Write $X = \Delta(K)$.
If $p \in \Delta(K)$ and $a \in \Delta(I)^K$, define $q(p,a)$ in $\Delta(X)$ as the law of the belief of the DM over the state of the next period if his current belief was p and player 1 has played a^k if the state was k.

Theorem: (R., Venel 2012)

$$v = \inf\{w(p_0), w : \Delta(X) \to [0,1] \text{ affine } C^0 \ s.t.$$

$$(1)\ \forall p \in X, w(p) \geq \sup_{a \in \Delta(I)^K} w(q(p,a))$$

$$(2)\ \forall (u,y) \in RR, w(u) \geq y\}.$$

where

$$RR = \{(u,y) \in \Delta(X) \times [0,1], \text{there exists } a : X \to \Delta(I)^K \text{ measurable } s.t.$$

$$\int_{p \in X} q(p,a(p))du(p) = u \text{ and } \int_{p \in X} \min_{j \in J} g(p,a(p),j)du(p) = y \qquad \}.$$

State variable: Belief of player 2 on the current state in K. Write $X = \Delta(K)$.
If $p \in \Delta(K)$ and $a \in \Delta(I)^K$, define $q(p,a)$ in $\Delta(X)$ as the law of the belief of the DM over the state of the next period if his current belief was p and player 1 has played a^k if the state was k.

Theorem: (R., Venel 2012)

$$v = \inf\{w(p_0), w : \Delta(X) \to [0,1] \text{ affine } C^0 \; s.t.$$

$$(1)\ \forall p \in X, w(p) \geq \sup_{a \in \Delta(I)^K} w(q(p,a))$$

$$(2)\ \forall (u,y) \in RR, w(u) \geq y\}.$$

where

$$RR = \{(u,y) \in \Delta(X) \times [0,1], \text{there exists } a : X \to \Delta(I)^K \text{ measurable } s.t.$$

$$\int_{p \in X} q(p,a(p))du(p) = u \text{ and } \int_{p \in X} \min_{j \in J} g(p,a(p),j)du(p) = y \qquad \}.$$

Back to example 2 : p_0 is the initial distribution.

$$\begin{array}{cc} & \begin{array}{cc} j_1 & j_2 \end{array} \\ \begin{array}{c} i_1 \\ i_2 \end{array} & \begin{pmatrix} 1_{\circlearrowleft} & 0_{\circlearrowleft} \\ 0_{p_0} & 0_{p_0} \end{pmatrix} \\ & \text{state } k_1 \end{array} \qquad \begin{array}{cc} & \begin{array}{cc} j_1 & j_2 \end{array} \\ \begin{array}{c} i_1 \\ i_2 \end{array} & \begin{pmatrix} 0_{p_0} & 0_{p_0} \\ 0_{p_0} & 1_{p_0} \end{pmatrix} \\ & \text{state } k_2 \end{array}$$

Here, $v = 1/3$.

Optimal strategy for player 2: play j_2 if the last action of player 1 was i_1, and $2/3 j_1 + 1/3 j_2$ if it was i_2.

Optimal strategy of player 1: Play i_2 in state k_2, and play i_1 with proba p^2/p^1 if player 2's belief on the current state is (p_1, p_2). Generates the invariant measure $2/3\ \delta_{(1/2,1/2)} + 1/3\ \delta_{(1,0)}$.

But computing the value and optimal strategies is difficult and unknown in general. Even when the chain is uncontrolled

$K = \{a, b\}$, $p_0 = (1/2, 1/2)$,

$$\begin{array}{cc} & j_1 \quad\quad j_2 \\ \begin{array}{c} i_1 \\ i_2 \end{array} & \left(\begin{array}{cc} 1_{(\alpha,1-\alpha)} & 0_{(\alpha,1-\alpha)} \\ 0_{(\alpha,1-\alpha)} & 0_{(\alpha,1-\alpha)} \end{array} \right) \end{array} \qquad \begin{array}{cc} & j_1 \quad\quad j_2 \\ \begin{array}{c} i_1 \\ i_2 \end{array} & \left(\begin{array}{cc} 0_{(1-\alpha,\alpha)} & 0_{(1-\alpha,\alpha)} \\ 0_{(1-\alpha,\alpha)} & 1_{(1-\alpha,\alpha)} \end{array} \right) \end{array}$$

state k_1 — state k_2

The state evolves according to the Markov chain with transition $M = \begin{pmatrix} \alpha & 1-\alpha \\ 1-\alpha & \alpha \end{pmatrix}$.

If $\alpha = 1$, the value is 1/4 (Aumann Maschler).

If $\alpha \in [1/2, 2/3]$, the value is $\frac{\alpha}{4\alpha - 1}$ (Hörner *et al.* 2010, Marino 2005 for $\alpha = 2/3$).

Recent advances for $\alpha \in [2/3, .73]$ (Bressaud & Quas 2014): $\frac{1}{v} = u_0 + u_0 u_1 + u_0 u_1 u_2 + \ldots$, where (u_n) is defined by $u_0 = 1$ and $u_{n+1} = max\{\psi(u_n), 1 - \psi(u_n)\}$ with $\psi(u) = 3\alpha - 1 - \frac{2\alpha - 1}{u}$.

What is the value for $\alpha = 0.9$?

But computing the value and optimal strategies is difficult and unknown in general. Even when the chain is uncontrolled
$K = \{a, b\}$, $p_0 = (1/2, 1/2)$,

$$\begin{array}{cc} & j_1 \quad\quad j_2 \\ \begin{array}{c} i_1 \\ i_2 \end{array} & \left(\begin{array}{cc} 1_{(\alpha,1-\alpha)} & 0_{(\alpha,1-\alpha)} \\ 0_{(\alpha,1-\alpha)} & 0_{(\alpha,1-\alpha)} \end{array} \right) \\ & \text{state } k_1 \end{array} \qquad \begin{array}{cc} & j_1 \quad\quad j_2 \\ \begin{array}{c} i_1 \\ i_2 \end{array} & \left(\begin{array}{cc} 0_{(1-\alpha,\alpha)} & 0_{(1-\alpha,\alpha)} \\ 0_{(1-\alpha,\alpha)} & 1_{(1-\alpha,\alpha)} \end{array} \right) \\ & \text{state } k_2 \end{array}$$

The state evolves according to the Markov chain with transition
$M = \left(\begin{array}{cc} \alpha & 1-\alpha \\ 1-\alpha & \alpha \end{array} \right)$.
If $\alpha = 1$, the value is $1/4$ (Aumann Maschler).
If $\alpha \in [1/2, 2/3]$, the value is $\frac{\alpha}{4\alpha-1}$ (Hörner *et al.* 2010, Marino 2005 for $\alpha = 2/3$).

Recent advances for $\alpha \in [2/3, .73]$ (Bressaud & Quas 2014):
$\frac{1}{v} = u_0 + u_0 u_1 + u_0 u_1 u_2 + \ldots$, where (u_n) is defined by $u_0 = 1$ and $u_{n+1} = max\{\psi(u_n), 1 - \psi(u_n)\}$ with $\psi(u) = 3\alpha - 1 - \frac{2\alpha-1}{u}$.

What is the value for $\alpha = 0.9$?

Grazie per l'attenzione !