Zero-Sum Repeated Games with Stage Duration and Public Signals

Ivan Novikov, Université Paris Dauphine-PSL

1. Repeated games with public signals

A *zero-sum repeated game with symmetric information, in which the signal depends only on the current state* is a 7-tuple $(A, \Omega, f, I, J, g, q)$, where:

- A is the finite set of signals;
- Ω is the finite set of states;
- $f : \Omega \to A$ is a deterministic signaling function on Ω;
- I (resp. J) is the finite set of actions of player 1 (resp. player 2);
- $g : I \times J \times \Omega \to \mathbb{R}$ is the payoff function of player 1;
- $q : I \times J \to \{$matrices $|\Omega| \times |\Omega|$ such that $q_{\omega_1\omega_2}(i,j) \geq 0$ for $\omega_1 \neq \omega_2$, $q_{\omega\omega}(i,j) \in [-1, 0]$, and $\sum_{\omega_2 \in \Omega} q_{\omega_1\omega_2}(i,j) = 0\}$ is the kernel function.

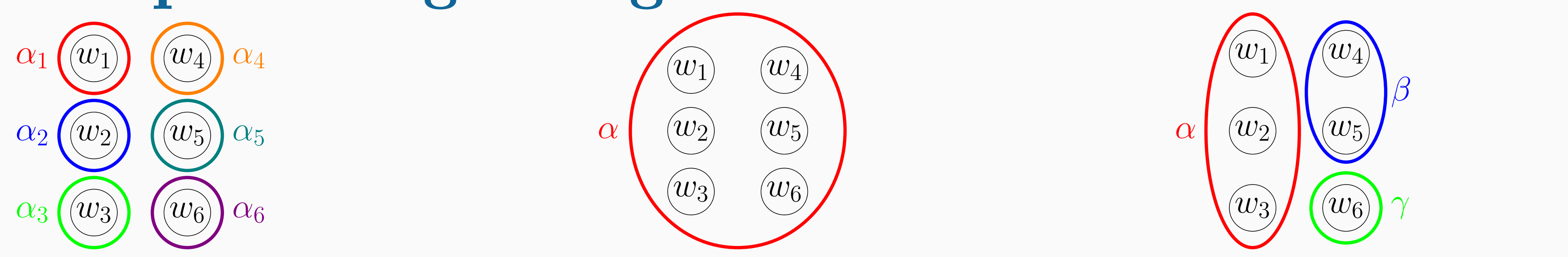
The game is played in discrete time. An initial state $\omega_1 \in \Omega$ is chosen according to given probability law p_0. At stage $n \in \mathbb{N}^*$:

1. The current state is $\omega_n \in \Omega$. Players do not observe it, but they observe the signal $\alpha_n = f(\omega_n)$;
2. Players choose their mixed actions, $x_n \in \Delta(I)$ and $y_n \in \Delta(J)$;
3. An action $i_n \in I$ (resp. $j_n \in J$) is chosen according to x_n (resp. y_n). Both actions are observed by both players;
4. The payoff is $g_n = g(i_n, j_n, \omega_n)$. The new state is ω_{n+1} with probability $\begin{cases} q_{\omega_n\omega_{n+1}}(i_n, j_n) & \text{, if } \omega_{n+1} \neq \omega_n; \\ 1 + q_{\omega_n\omega_{n+1}}(i_n, j_n) & \text{, if } \omega_{n+1} = \omega_n. \end{cases}$

A *strategy* consists of choosing x_n and y_n at each stage (taking into account past actions and received signals).

If A is a singleton, then the game is a *state-blind repeated game*. If $A = \Omega$ and $f = Id$, then the game is a stochastic game, i.e. players can observe the state.

2. Examples of signalling on the states



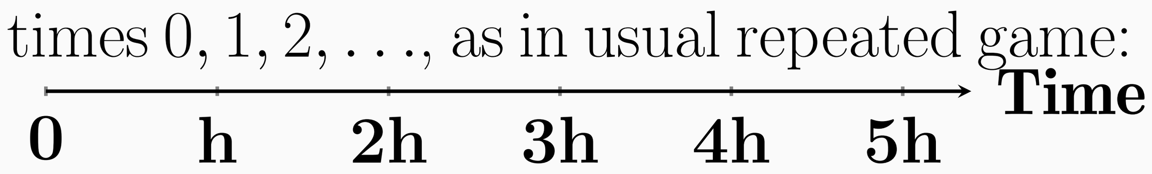
(a) Perfect observation of the state, i.e. there are 6 signals $\alpha_1, \ldots, \alpha_6$; and $f(w_i) := \alpha_i$.

(b) State-blind case. There is only one signal α, and $f(w_i) := \alpha$.

(c) Intermediate case. $f(w_1) = f(w_2) = f(w_3) = \alpha$, $f(w_4) = f(w_5) = \beta$, $f(w_6) = \gamma$

3. Stage duration

Fix a repeated game $G = (A, \Omega, f, I, J, g, q)$. *Game with stage duration* h is the repeated game $G_h = (A, \Omega, f, I, J, g^h, q^h)$, where $g^h = hg$ and $q^h = hq$, i.e. the kernel and the payoff function are proportional to stage duration h. We may think that the players act at times $0, h, 2h, \ldots$, and not at times $0, 1, 2, \ldots$, as in usual repeated game:



For each $\lambda > 0$ and small h, the λ-discounted payoff of the game G_h is (depending on the strategy profile (σ, τ), and initial probability on the states p)

$$G^p_{\sigma,\tau} = E^p_{\sigma,\tau}\left(\lambda \sum_{i=1}^{\infty} (1 - \lambda h)^{i-1} g^h_i\right).$$

The value $v_{h,\lambda} : \Delta(\Omega) \to \mathbb{R}$ is defined by $v_{h,\lambda}(p) = \sup_\sigma \inf_\tau G^p_{\sigma,\tau} = \inf_\tau \sup_\sigma G^p_{\sigma,\tau}$.

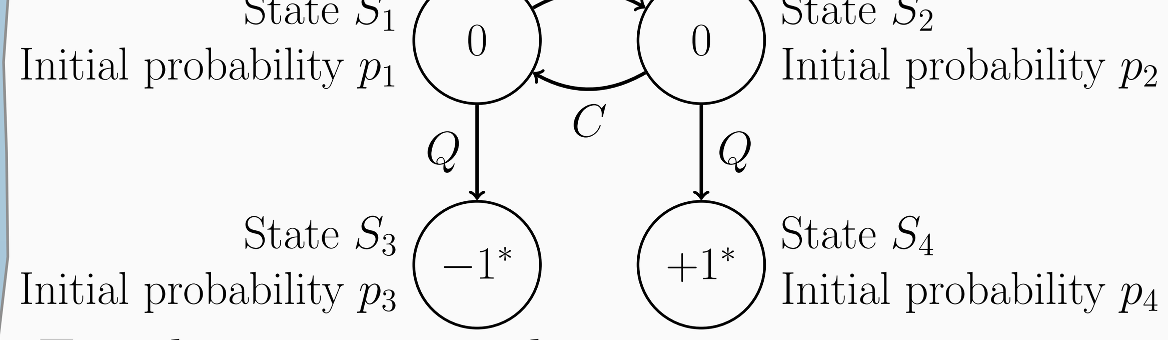
Stochastic games with stage duration were first introduced in [1].

Proposition 1 [1,2]: If players can observe the states (i.e. $f(\omega_i) = \alpha_i$), then:

1. $\lim_{h\to 0} v_{h,\lambda} = v_{1,\frac{\lambda}{1+\lambda}}$;
2. $\lim_{h\to 0} v_{h,\lambda}$ is a unique solution of
$\lambda v(\omega) = \mathtt{Val}_{I\times J}\left[g(i,j,\omega) + \sum_{\omega' \in \Omega} q_{\omega\omega'}(i,j) \cdot v(\omega')\right]$;
3. The limits $\lim_{\lambda\to 0} \lim_{h\to 0} v_{h,\lambda}$ and $\lim_{\lambda\to 0} v_{1,\lambda}$ exist or do not exist simultaneously. In the case of existence, they are equal to each other.

4. An example

Consider a 1-player state-blind game with 2 actions, C and Q.



For this game, we have

$$\lim_{h\to 0} v_{h,\lambda}(p_1,p_2,p_3,p_4) = p_4 - p_3 + \frac{1}{1+\lambda}\max\{0, p_2 - p_1\};$$
$$v_{1,\lambda}(p_1,p_2,p_3,p_4) = p_4 - p_3 + p_2(1-\lambda) + p_1(1-\lambda)^2.$$

Note all of the assertions of Proposition 1 do not hold here.

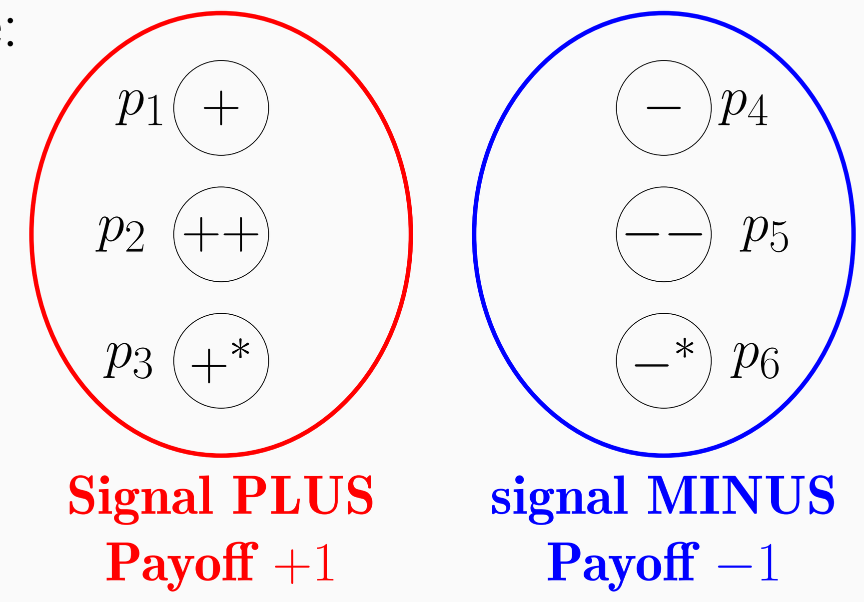
5. First result

If X is finite, $\zeta \in \Delta(X)$, and μ is a $|X| \times |X|$ matrix, denote $\zeta * \mu(x) := \sum_{x' \in X} \zeta(x') \cdot \mu_{x'x}$.

Theorem 1: If $(A, \Omega, f, I, J, g, q)$ is a state-blind repeated game, then uniform limit $\lim_{h\to 0} v_{h,\lambda}(p)$ exists and is a unique viscosity solution of
$\lambda v(p) = \mathtt{val}_{I\times J}[g(i,j,p) + \langle (p * q(i,j))(\cdot), \nabla v(p)\rangle]$.

6. Second result

Consider a game from [4], similar to a game from [5]. It has 6 states, $+, ++, +^*, -, --, -^*$. The states $+^*$ and $-^*$ are absorbing. Payoff and signalling structure:



The actions of player 1 are T, M, B, and the actions of player 2 are L, M, R, Q. The transition probability matrices:

	L	M	R	Q
T	$-$	$--$	$--$	$--$
M	$--$	$-$	$-$	$-$
B	$++$	$++$	$++$	$++$

Figure 2: State $-$

	L	M	R	Q
T	$+$	$++$	$++$	$--$
M	$++$	$+$	$++$	$--$
B	$++$	$++$	$+$	$--$

Figure 3: State $+$

	L	M	R	Q
T	$\frac{1}{2}\{-\} + \frac{1}{2}\{--\}$	$--$	$--$	$--$
M	$--$	$\frac{1}{2}\{-\} + \frac{1}{2}\{--\}$	$\frac{1}{2}\{-\} + \frac{1}{2}\{--\}$	$\frac{1}{2}\{-\} + \frac{1}{2}\{--\}$
B	$-^*$	$-^*$	$-^*$	$-^*$

Fig. 4: State $--$

	L	M	R	Q
T	$\frac{1}{3}\{++\} + \frac{2}{3}\{+\}$	$++$	$++$	$+^*$
M	$++$	$\frac{1}{3}\{++\} + \frac{2}{3}\{+\}$	$++$	$+^*$
B	$++$	$++$	$\frac{1}{3}\{++\} + \frac{2}{3}\{+\}$	$+^*$

Fig. 5: State $++$

Theorem 2: Denote $p = (p_1, p_2, p_3, p_4, p_5, p_6)$. For the above game,

1. $\lim_{\lambda\to 0} v_{1,\lambda}(p)$ does not exist [4,5];
2. Uniform limit $\lim_{\lambda\to 0} \lim_{h\to 0} v_{h,\lambda}(p)$ exists, and it equals

$$\begin{cases} \frac{7}{11}p_1 - p_2 - p_3 + \frac{5}{11}p_4 - p_5 + p_6 & \text{, if } \frac{p_2}{p_1+p_2} \leq \frac{2}{3} \text{ and } \frac{p_5}{p_4+p_5} \leq \frac{3}{4}; \\ \frac{7}{11}p_1 - p_2 - p_3 + \frac{7}{11}p_4 + \frac{7}{11}p_5 + p_6 & \text{, if } \frac{p_2}{p_1+p_2} \leq \frac{2}{3} \text{ and } \frac{p_5}{p_4+p_5} > \frac{3}{4}; \\ -\frac{5}{11}p_1 - \frac{5}{11}p_2 - p_3 + \frac{5}{11}p_4 - p_5 + p_6 & \text{, if } \frac{p_2}{p_1+p_2} > \frac{2}{3} \text{ and } \frac{p_5}{p_4+p_5} \leq \frac{3}{4}; \\ -\frac{5}{11}p_1 - \frac{5}{11}p_2 - p_3 + \frac{7}{11}p_4 + \frac{7}{11}p_5 + p_6 & \text{, if } \frac{p_2}{p_1+p_2} > \frac{2}{3} \text{ and } \frac{p_5}{p_4+p_5} > \frac{3}{4}. \end{cases}$$

7. Final remarks

1. In case when players observe the state, the value of a continuous-time zero-sum Markov game (CTZSMG) coincides with the value of analogous stochastic game when its stage duration tends to 0. In the state-blind case, the same thing holds, if instead of a CTZSMG we consider its discretization, in which players can act only at times $0, h, 2h, \ldots$, where h tends to 0. This implies that the general case with public signals may also be connected with CTZSMGs. The discretization of CTZSMGs is studied in [3].
2. Instead of assuming that each stage has duration h, one may assume that n-th stage duration is h_n, so that n-th stage payoff and kernel are $h_n g$ and $h_n q$ respectively. We are interested in the behavior of the value when $\sup\{h_n\} \to 0$. One can prove that Theorems 1 and 2 still hold in this more general framework.

References

[1] A. Neyman. Stochastic games with short-stage duration. Dynamic Games and Applications, 3(2):236–278, 2013.

[2] S. Sorin and G. Vigeral. Operator approach to values of stochastic games with varying stage duration. International Journal of Game Theory, 45(1):389–410, 2016.

[3] S. Sorin. Limit value of dynamic zero-sum games with vanishing stage duration. Mathematics of Operations Research, 43(1):51–63, 2017.

[4] J. Renault and B. Ziliotto. Hidden stochastic games and limit equilibrium payoffs. Games and Economic Behavior, 124:122–139, 2020.

[5] B. Ziliotto. Zero-sum repeated games: Counterexamples to the existence of the asymptotic value and the conjecture maxmin = $\lim v_n$. The Annals of Probability, 44(2):1107 – 1133, 2016.

LaTeX TikZposter