Repeated Games with Stage Duration and Public Signals

Ivan Novikov

Université Paris-Dauphine

19/10/2023

Table of contents

Zero-sum stochastic games (1)

A zero-sum stochastic game is a 5-tuple (Ω, I, J, g, P), where:

- Ω is a non-empty set of states;
- I is a non-empty set of actions of player 1;
- J is a non-empty set of actions of player 2;
- $g : I \times J \times \Omega \to \mathbb{R}$ is a payoff function of player 1;
- $P : I \times J \times \Omega \to \Delta(\Omega)$ is a transition probability function.

We assume that I, J, Ω are finite.
$\Delta(\Omega) :=$ the set of probability measures on Ω.

Zero-sum stochastic games (2)

A stochastic game (Ω, I, J, g, P) proceeds in stages as follows. At each stage n:

1. The players observe the current state ω_n;
2. Players choose their mixed actions, $x_n \in \Delta(I)$ and $y_n \in \Delta(J)$;
3. Pure actions $i_n \in I$ and $j_n \in J$ are chosen according to $x_n \in \Delta(I)$ and $y_n \in \Delta(J)$;
4. Player 1 obtains a payoff $g_n = g(i_n, j_n, \omega_n)$, while player 2 obtains payoff $-g_n$;
5. The new state w_{n+1} is chosen according to the probability law $P(i_n, j_n, \omega_n)$.

The above description of the game is known to the players.

Strategies

- Strategies σ, τ of players consist in choosing at each stage a mixed action;
- The players can take into account the previous actions of players, as well as the current and previous states.

λ-discounted game Γ^λ

- What is the goal of the players?
- Total payoff: $E^{\omega}_{\sigma,\tau}\left(\lambda \sum_{i=1}^{\infty}(1-\lambda)^{i-1} g_i\right)$;
- Depends on $\lambda \in (0,1)$, initial state ω, and strategies of the players;
- Player 1 wants to maximize total payoff, while player 2 wants to minimize it;
- Value $v_\lambda : \Omega \to \mathbb{R}$:

$$v_\lambda(\omega) = \sup_\sigma \inf_\tau E^{\omega}_{\sigma,\tau}\left(\lambda \sum\nolimits_{i=1}^{\infty}(1-\lambda)^{i-1} g_i\right) =$$
$$= \inf_\tau \sup_\sigma E^{\omega}_{\sigma,\tau}\left(\lambda \sum\nolimits_{i=1}^{\infty}(1-\lambda)^{i-1} g_i\right).$$

Limit of λ-discounted game Γ^λ

- $v_\lambda(\omega) = \sup_\sigma \inf_\tau E^\omega_{\sigma,\tau} \left(\lambda \sum_{i=1}^\infty (1-\lambda)^{i-1} g_i\right)$;
- One can ask: what happens if players become more and more patient? I.e., players are willing to wait a lot to obtain a big payoff;
- Mathematically, it means that $\lambda \to 0$;
- Thus, one is interested in the uniform (in ω) limit $\lim_{\lambda\to 0} v_\lambda(\omega)$.

Table of contents

Kernel

- Kernel $Q : I \times J \times \Omega \to \Delta(\Omega)$

$$Q(i,j,\omega)(\omega') = \begin{cases} P(i,j,\omega)(\omega') & \text{if } \omega \neq \omega'; \\ P(i,j,\omega)(\omega') - 1 & \text{if } \omega = \omega'. \end{cases}$$

- Recall that $P(i,j,\omega)(\omega')$ is the probability that the next state is ω', if the current state is ω and players' actions are (i,j);
- Hence the closer kernel Q is to 0, the more probable it is that the next state coincides with the current one.

Stochastic games with stage duration

- Consider a family of stochastic games G_h, parametrized by $h \in (0, 1]$;
- h represents stage duration;
- Players now play at times $h, 2h, 3h, \ldots$, instead of playing at times $1, 2, 3, \ldots$;
- State space Ω and action spaces I and J of player 1 and player 2 are independent of h;
- Payoff function g_h of player 1 and kernel Q_h depend on h.

Stochastic games with stage duration

- Payoff $g_h = hg$;
- Kernel $Q_h = hQ$;
- $h = 1$: "Usual" stochastic game;
- When h small, g_h is close to zero (players receive almost nothing each turn), and Q_h is close to zero (the next state with a high probability will be the same).

Comparison (1)

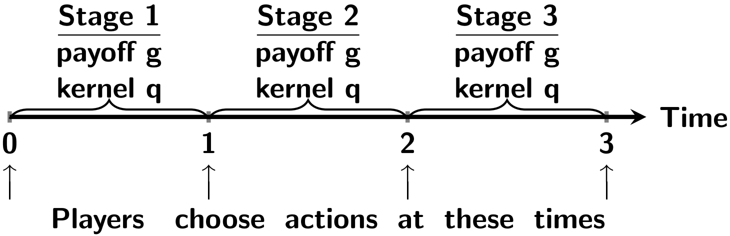


Figure: "Usual" stochastic game: duration of each stage is 1

Comparison (2)

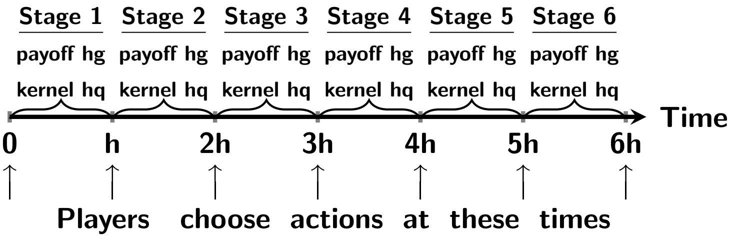


Figure: Stochastic game with stage duration h: stage payoff and kernel are proportional to h

Papers about games with stage duration

- "Stochastic games with short-stage duration" by Abraham Neyman;
- "Operator approach to values of stochastic games with varying stage duration" by Sylvain Sorin and Guillaume Vigeral.

Discounted games with stage duration

- For a game with stage duration h, the total payoff is (depending on the discount factor λ, initial state ω, and strategies σ, τ of players)
$$E^{\omega}_{\sigma,\tau}\left(\lambda \sum_{k=1}^{\infty}(1-\lambda h)^{k-1} g_h^k\right);$$
- Why such a choice? Easy explanation:
- The total payoff is λ-discounted game with stage duration 1 is $E^{\omega}_{\sigma,\tau}\left(\lambda \sum_{k=1}^{\infty}(1-\lambda)^{k-1} g^k\right)$. The total payoff of λ-discounted game with stage duration h is $E^{\omega}_{\sigma,\tau}\left(\sum_{k=1}^{\infty}\lambda h(1-\lambda h)^{k-1} g^k\right)$;
- So, it may be seen as a game with discount factor λh. I.e., the discount factor is proportional to h, just as the payoff g and the kernel Q.

Real meaning behind the total payoff of the game with stage duration h (1)

- Total payoff: $E^{\omega}_{\sigma,\tau}\left(\lambda \sum_{k=1}^{\infty}(1-\lambda h)^{k-1} g_h^k\right)$;
- When h is small, the total payoff of the λ-discounted stochastic game with stage duration h is close to the total payoff of the analogous λ-discounted continuous-time game;
- In a continuous-time game, players can choose actions at any time, and at each time t they receive instantaneous payoff g_t. The total payoff is (depending on the discount factor λ) $\int_0^{\infty} \lambda e^{-\lambda t} g_t dt$. The received payoff during a period of time $[(n-1)h, nh]$ is $\int_{(n-1)h}^{nh} \lambda e^{-\lambda t} g_t dt \underset{h\to 0}{\approx} \lambda h g_{(n-1)h}(1-\lambda h)^{n-1}$.

Real meaning behind the total payoff of the game with stage duration h (2)

- Total payoff $E^{\omega}_{\sigma,\tau}\left(\lambda \sum_{k=1}^{\infty}(1-\lambda h)^{k-1} g_h^k\right)$;
- Thus the total payoff of a continuous-time game is close to the total payoff of analogous discrete-time game;
- One can prove strictly that the values of λ-discounted games with stage duration h tend to the value of analogous continuous-time λ-discounted game when $h \to 0$.

Discounted games with stage duration (main properties)

- We denote by $v_{h,\lambda}$ the value of the game with total payoff $E^{\omega}_{\sigma,\tau}\left(\lambda\sum_{k=1}^{\infty}(1-\lambda h)^{k-1}g^k_h\right)$;
- Main question: What happens with $v_{h,\lambda}$ when $h\to 0$?

Proposition (A. Neyman)

$\lim_{h\to 0} v_{h,\lambda}$ *exists and is a unique solution of a functional equation.*

Proposition (S. Sorin, G. Vigeral)

$\lim_{\lambda\to 0}\lim_{h\to 0} v_{h,\lambda}$ *exists if and only if* $\lim_{\lambda\to 0} v_{1,\lambda}$ *exists.*

- $\lim_{\lambda\to 0} v_{1,\lambda}$ should be considered as the limit value of the discrete-time stochastic game, whereas $\lim_{\lambda\to 0}\lim_{h\to 0} v_{h,\lambda}$ should be considered as the limit value of analogous continuous-time game.

Table of contents

Repeated Games with Public Signals (1)

- Now players cannot perfectly obseserve the current state;
- Players know the initial probability distribution on the states and some information about the current state.

Repeated Games with Public Signals (2)

A *zero-sum repeated game with symmetric information,* is a 7-tuple $(A, \Omega, f, I, J, g, P)$, where:

- A is a non-empty set of signals;
- Ω is a non-empty set of states;
- $f : \Omega \to A$ is a partition of Ω;
- I is a non-empty set of actions of player 1;
- J is a non-empty set of actions of player 2;
- $g : I \times J \times \Omega \to \mathbb{R}$ is stage payoff function of player 1;
- $P : I \times J \times \Omega \to \Delta(\Omega)$ is the transition probability function.

We assume that I, J, Ω, A are finite.

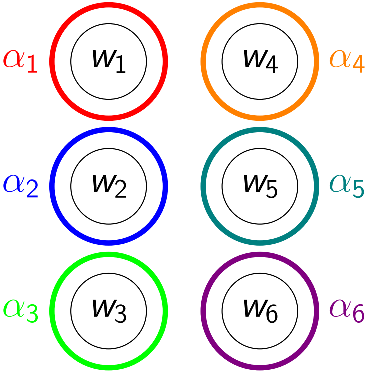
Repeated Games with Public Signals (3)

The game $(A, \Omega, f, I, J, g, P)$ proceeds in stages as follows. At each stage n:

1. The current state is ω_n. Players do not observe it, but they observe the signal $\alpha_n = f(\omega_n) \in A$ and the actions of each other at the previous stage;
2. Players choose their mixed actions, $x_n \in \Delta(I)$ and $y_n \in \Delta(J)$;
3. Pure actions $i_n \in I$ and $j_n \in J$ are chosen according to $x_n \in \Delta(I)$ and $y_n \in \Delta(J)$;
4. Player 1 obtains a payoff $g_n = g(i_n, j_n, \omega_n)$, while player 2 obtains payoff $-g_n$;
5. The new state w_{n+1} is chosen according to the probability law $P(i_n, j_n, \omega_n)$. The new signal is $\alpha_{n+1} = f(\omega_{n+1})$.

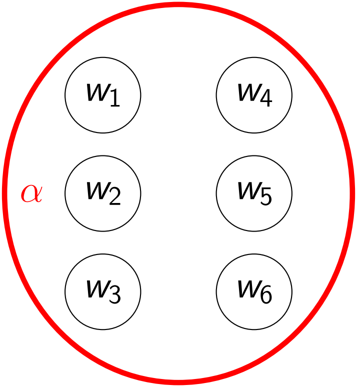
The above description of the game is known to the players.

Examples of the partition function f (1)



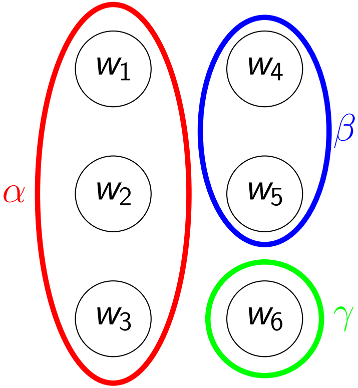
The perfect observation of the state, i.e. there are 6 signals $\alpha_1, \dots, \alpha_6$; and $f(w_i) := \alpha_i$.

Examples of the partition function f (2)



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

Examples of the partition function f (3)



Neither the perfect observation of the state nor the state-blind case. There are 3 signals, and $f(w_1) = f(w_2) = f(w_3) = \alpha$, $f(w_4) = f(w_5) = \beta$, $f(w_6) = \gamma$.

Stage duration

- We still can consider games with stage duration h in this new setting;
- Payoff $g_h = hg$;
- Kernel $Q_h = hQ$;
- State space Ω, signal set A, partition function f, and action spaces I and J of player 1 and player 2 are independent of h;
- The total payoff is still $E^{\omega}_{\sigma,\tau}\left(\lambda \sum_{k=1}^{\infty}(1-\lambda h)^{k-1} g_h^k\right)$;
- $v_{h,\lambda}$ is the value of the game with such a total payoff.

First result

Theorem

In the state-blind case, $\lim_{h\to 0} v_{h,\lambda}$ *exists and is a unique viscosity solution of a partial differential equation.*

- The main tool in the proof is the Shapley equation;
- Open question: can we say that $\lim_{h\to 0} v_{h,\lambda}$ in a general case of games with public signals; if yes, is it a unique viscosity solution of some PDE?

Second result

Theorem

There is a game in which uniform limit $\lim_{\lambda \to 0} \lim_{h \to 0} v_{h,\lambda}$ *exists, but pointwise limit* $\lim_{\lambda \to 0} v_{1,\lambda}$ *does not exist.*

Open question: can we say that

1. For any fixed $h \in (0,1]$, the limit $\lim\limits_{\lambda \to 0} v_{h,\lambda}$ does not exist?
2. We have $\left| \limsup\limits_{\lambda \to 0} v_{h,\lambda}(p) - \liminf\limits_{\lambda \to 0} v_{h,\lambda}(p) \right| \to 0$ as $h \to 0$, uniformly in p?

Second result (proof)

Theorem

There is a game in which uniform limit $\lim_{\lambda\to 0} \lim_{h\to 0} v_{h,\lambda}$ *exists, but pointwise limit* $\lim_{\lambda\to 0} v_{1,\lambda}$ *does not exist.*

- $\lim_{\lambda\to 0} v_{1,\lambda}$ does not exist: the game in this case is equivalent to a game from "Hidden stochastic games and limit equilibrium payoffs" (Jérôme Renault and Bruno Ziliotto), which is known to not have the limit value;
- $\lim_{\lambda\to 0} \lim_{h\to 0} v_{h,\lambda}$ exists: use the first result.

Generalization: varying stage duration

- Now we allow different stage durations for different stages;
- There is a sequence $\{h_i\}_{i\in\mathbb{N}}$;
- Players act in times $h_1, h_1 + h_2, h_1 + h_2 + h_3, \ldots$;
- i-th stage payoff is $h_i g$ and i-th stage kernel is $h_i Q$;
- The analogues of the above theorems hold in this more general model.

This is all.

Thank you!