Asymptotic value of repeated games and iteration of non expensive maps

Sylvain Sorin

IMJ-PRG
Université P. et M. Curie - Paris 6

Atelier Italo-Français:
Optimisation et processus dynamiques en apprentissage et dans les problèmes inverses,
Sestri Levante, 8-12 Septembre 2014

Table of contents

Table of contents

1.1 General framework

A finite two person zero-sum stochastic game is defined by:
a state space Ω,
actions spaces I and J,
a transition probability Q from $\Omega \times I \times J$ to $\Delta(\Omega)$
a real payoff function g on $\Omega \times I \times J$.
Time is discrete: at stage n, given the past history including ω_n, the players choose (at random) i_n and j_n, the stage payoff is $g_n = g(\omega_n, i_n, j_n)$ and the law of the new state ω_{n+1} is $q(\omega_n, i_n, j_n)$.

The Shapley operator (1953) associated to this game gives, for any real function $f \in F$, from Ω to $\mathbb{R}$, the value of the one shot game with terminal payoff f:

$$\mathbf{T}(f)(\omega) = \mathtt{val}_{X\times Y}[g(\omega, i, j) + \sum_{\omega'} q(\omega, i, j)(\omega')f(\omega')]$$

where $X = \Delta(I), Y = \Delta(J)$.

1.1 General framework

A finite two person zero-sum stochastic game is defined by:
a state space Ω,
actions spaces I and J,
a transition probability Q from $\Omega \times I \times J$ to $\Delta(\Omega)$
a real payoff function g on $\Omega \times I \times J$.
Time is discrete: at stage n, given the past history including ω_n, the players choose (at random) i_n and j_n, the stage payoff is $g_n = g(\omega_n, i_n, j_n)$ and the law of the new state ω_{n+1} is $q(\omega_n, i_n, j_n)$.

The Shapley operator (1953) associated to this game gives, for any real function $f \in F$, from Ω to $\mathbb{R}$, the value of the one shot game with terminal payoff f:

$$\mathbf{T}(f)(\omega) = \mathtt{val}_{X \times Y}[g(\omega, i, j) + \sum_{\omega'} q(\omega, i, j)(\omega') f(\omega')]$$

where $X = \Delta(I), Y = \Delta(J)$.

1.1 General framework

A finite two person zero-sum stochastic game is defined by:
a state space Ω,
actions spaces I and J,
a transition probability Q from $\Omega \times I \times J$ to $\Delta(\Omega)$
a real payoff function g on $\Omega \times I \times J$.
Time is discrete: at stage n, given the past history including ω_n, the players choose (at random) i_n and j_n, the stage payoff is $g_n = g(\omega_n, i_n, j_n)$ and the law of the new state ω_{n+1} is $q(\omega_n, i_n, j_n)$.

The Shapley operator (1953) associated to this game gives, for any real function $f \in F$, from Ω to $\mathbb{R}$, the value of the one shot game with terminal payoff f:

$$\mathbf{T}(f)(\omega) = \mathtt{val}_{X \times Y}[g(\omega, i, j) + \sum_{\omega'} q(\omega, i, j)(\omega') f(\omega')]$$

where $X = \Delta(I), Y = \Delta(J)$.

1.2. Extensions

There are natural conditions for this operator to be well defined and

$\mathbf{T}$ is monotonic

$\mathbf{T}$ translates the constants.

A lot of similar properties hold for such operators from F to F, that are non expansive (for the uniform norm).

A typical example is a translation, hence without fixed point.

1.3. This operator extends to general repeated games (incomplete information, signals ...) by defining an adequate state space and transition probabilities.

What is crucial is to keep the stationarity structure defined by g and Q.

1.2. Extensions

There are natural conditions for this operator to be well defined and

$\mathbf{T}$ is monotonic

$\mathbf{T}$ translates the constants.

A lot of similar properties hold for such operators from F to F, that are non expansive (for the uniform norm).

A typical example is a translation, hence without fixed point.

1.3. This operator extends to general repeated games (incomplete information, signals ...) by defining an adequate state space and transition probabilities.

What is crucial is to keep the stationarity structure defined by g and Q.

1.2. Extensions

There are natural conditions for this operator to be well defined and

$\mathbf{T}$ is monotonic

$\mathbf{T}$ translates the constants.

A lot of similar properties hold for such operators from F to F, that are non expansive (for the uniform norm).

A typical example is a translation, hence without fixed point.

1.3. This operator extends to general repeated games (incomplete information, signals ...) by defining an adequate state space and transition probabilities.

What is crucial is to keep the stationarity structure defined by g and Q.

Table of contents

2.1. Repeated games and recursive formula

Repeated games are played in stages and the parameter of a stage corresponds to the number of repetitions. There is no intrinsic duration or weigth to a stage.

The n-stage game corresponds to the evaluation $g_1 + ... + g_n$. Its value V_n satisfies:

$$V_n = \mathbf{T}(V_{n-1})$$

(which extends the basic dynamic programming equation).

Its normalized value v_n is given by:

$$v_n = \frac{T^n(0)}{n}$$

2.1. Repeated games and recursive formula

Repeated games are played in stages and the parameter of a stage corresponds to the number of repetitions. There is no intrinsic duration or weigth to a stage.

The n-stage game corresponds to the evaluation $g_1 + ... + g_n$. Its value V_n satisfies:

$$V_n = \mathbf{T}(V_{n-1})$$

(which extends the basic dynamic programming equation).

Its normalized value v_n is given by:

$$v_n = \frac{T^n(0)}{n}$$

Similarly for the λ discounted game the (un-normalized) value V_λ satisfies:

$$V_\lambda(\omega) = \mathtt{val}_{X\times Y}[g(\omega,i,j) + (1-\lambda)\sum_{\omega'} q(\omega,i,j)(\omega')V_\lambda(\omega')]$$

by stationarity, hence

$$V_\lambda = \mathbf{T}((1-\lambda)V_\lambda)$$

and the normalized value is $v_\lambda = \lambda V_\lambda$.

It is thus natural to introduce the operator $\Phi(\varepsilon,f) = \varepsilon\mathbf{T}(\frac{(1-\varepsilon)}{\varepsilon}f)$ and one has

$$v_n = \Phi(\frac{1}{n}, v_{n-1})$$

$$v_\lambda = \Phi(\lambda, v_\lambda)$$

Similarly for the λ discounted game the (un-normalized) value V_λ satisfies:

$$V_\lambda(\omega) = \mathtt{val}_{X\times Y}[g(\omega,i,j) + (1-\lambda)\sum_{\omega'} q(\omega,i,j)(\omega')V_\lambda(\omega')]$$

by stationarity, hence

$$V_\lambda = \mathbf{T}((1-\lambda)V_\lambda)$$

and the normalized value is $v_\lambda = \lambda V_\lambda$.

It is thus natural to introduce the operator $\Phi(\varepsilon,f) = \varepsilon\mathbf{T}(\frac{(1-\varepsilon)}{\varepsilon}f)$ and one has

$$v_n = \Phi(\frac{1}{n}, v_{n-1})$$

$$v_\lambda = \Phi(\lambda, v_\lambda)$$

More generally a (normalized) evaluation $\Theta = \{\theta_n\}_{n\geq 1}$ ($\theta_n \geq 0, \sum_{n\geq 1} \theta_n = 1$) specifies a game G_Θ with payoff $\sum \theta_n g_n$ and its value v_Θ satisfies the following **recursive formula**

$$v_\Theta(\omega) = \mathtt{val}_{X\times Y}[\theta_1 g(\omega,i,j) + (1-\theta_1)\sum_{\omega'} q(\omega,i,j)(\omega')v_{\Theta(1)}(\omega')]$$

where $\Theta(1)$ is the normalization of Θ after stage 1.

The **asymptotic analysis** is the study of the sequences of values as $\mathsf{E}(\Theta) \to \infty$.
One can describe these games as having " vanishing stage weight ".

More generally a (normalized) evaluation $\Theta = \{\theta_n\}_{n\geq 1}$ ($\theta_n \geq 0, \sum_{n\geq 1} \theta_n = 1$) specifies a game G_Θ with payoff $\sum \theta_n g_n$ and its value v_Θ satisfies the following **recursive formula**

$$v_\Theta(\omega) = \mathtt{val}_{X\times Y}[\theta_1 g(\omega,i,j) + (1-\theta_1)\sum_{\omega'} q(\omega,i,j)(\omega')v_{\Theta(1)}(\omega')]$$

where $\Theta(1)$ is the normalization of Θ after stage 1.

The **asymptotic analysis** is the study of the sequences of values as $\mathsf{E}(\Theta) \to \infty$.
One can describe these games as having " vanishing stage weight ".

2.2 Results
specific for :
n-stage or discounted games
stochastic or incomplete information games
Shapley, Aumann-Maschler, Mertens-Zamir, Bewley, Kohlberg (70-90)

operator approach Rosenberg- Sorin (2000)
comparison principle $\Phi(\varepsilon, f) \leq f$ for ε small enough,
derived game $\frac{\Phi(\varepsilon, f) - \Phi(0, f)}{\varepsilon}$
variational approach, Laraki (2002)

2.2 Results
specific for :
n-stage or discounted games
stochastic or incomplete information games
Shapley, Aumann-Maschler, Mertens-Zamir, Bewley, Kohlberg (70-90)

operator approach Rosenberg- Sorin (2000)
comparison principle $\Phi(\varepsilon, f) \leq f$ for ε small enough,
derived game $\frac{\Phi(\varepsilon, f) - \Phi(0, f)}{\varepsilon}$
variational approach, Laraki (2002)

2.3 Chernoff's formula

$$V_n - V_{n-1} = (T - Id)(V_{n-1})$$

Consider

$$\dot{U}(t) = -(Id - T)(U(t))$$

then

$$\|U(t) - \mathbf{T}^n(U(0))\| \leq \|U'(0)\|\sqrt{t + (n-t)^2}$$

In particular with $U(0) = 0$ and $t = n$

$$\|\frac{U(n)}{n} - v_n\| \leq \frac{\|\mathbf{T}(0)\|}{\sqrt{n}}$$

It is thus natural to consider $u(t) = \frac{U(t)}{t}$ which satisfies an equation of the form

$$\dot{x}(t) = \Phi(\varepsilon(t), x(t)) - x(t) \qquad (I)$$

which is no longer autonomous.
(Vigeral, 2009)
a) If $\varepsilon(t) = \lambda$, then $\|x(t) - v_\lambda\| \to 0$
b) If $\varepsilon(t) \sim \frac{1}{t}$, then $\|x(n) - v_n\| \to 0$
c) If $\frac{\varepsilon'(t)}{\varepsilon^2(t)} \to 0$ then $\|x(t) - v_{\varepsilon(t)}\| \to 0$
Hence $\lim v_n$ and $\lim v_\lambda$ mimick solutions of similar perturbed evolution equations. Moreover:
d) Let $\bar{x}$ solution associated to $\bar{\varepsilon}$. Then $\|u(t) - \bar{u}(t)\| \to 0$ as soon as
i) $\varepsilon(t) \sim \bar{\varepsilon}(t)$ as $t \to \infty$ or
ii) $|\varepsilon - \bar{\varepsilon}| \in L^1$
Stability results in the spirit of almost orbits (Peypouquet (2007), Alvarez and Peypouquet (2009)).

2.5 Neyman (2003)

If v_λ is of bounded variation in the sense that for any sequence λ_i decreasing to 0

$$\sum_i \|v_{\lambda_{i+1}} - v_{\lambda_i}\| < \infty \qquad (1)$$

then $\lim_{n\to\infty} v_n = \lim_{\lambda\to 0} v_\lambda$.

(application: finite stochastic games, semi algebraic case ...)

2.6 Random duration

Neyman and Sorin (2009)

An uncertain duration process a random tree with finite expected length where the nodes at distance n correspond to the information at stage n.

The “random iterate" $\mathbf{T}^{\Theta}$ is well defined.

$v_{\Theta} = \frac{\mathbf{T}^{\Theta}(0)}{E(\theta)}$ has the same properties than above, as $E(\theta)$ goes to ∞.

2.7. The game on $[0,1]$.

One can then introduce a game Γ_Θ on $[0,1]$ where t represents the fraction of the total duration. Stage n in G_Θ corresponds to the interval $[t_{n-1}, t_n[$ in Γ_Θ with $t_0 = 0$ and $t_n = \sum_{1\leq m\leq n} \theta_m$. To a game G_Θ is thus associated a sequence of values $\{w_\Theta(t_n, .)\}_{n\geq 0}$ (for the un-normalized game starting at t_n) hence a function $w_\theta(t,.)$ on $[0,1]$ by linear interpolation.

$$w_\Theta(0,\omega) = \mathtt{val}_{X\times Y}[\theta_1 g(\omega,i,j) + \sum_{\omega'} q(\omega,i,j)(\omega') w_\Theta(t_1,\omega')]$$

Cardaliaguet, Laraki, Sorin (2012)

The family of values has accumulation points.

Any limit point is a viscosity solution of some HJB equation.

Comparison principle implies uniqueness, hence convergence.

2.7. The game on $[0,1]$.

One can then introduce a game Γ_Θ on $[0,1]$ where t represents the fraction of the total duration. Stage n in G_Θ corresponds to the interval $[t_{n-1}, t_n[$ in Γ_Θ with $t_0 = 0$ and $t_n = \sum_{1 \leq m \leq n} \theta_m$. To a game G_Θ is thus associated a sequence of values $\{w_\Theta(t_n, .)\}_{n \geq 0}$ (for the un-normalized game starting at t_n) hence a function $w_\theta(t, .)$ on $[0,1]$ by linear interpolation.

$$w_\Theta(0,\omega) = \mathtt{val}_{X\times Y}[\theta_1 g(\omega,i,j) + \sum_{\omega'} q(\omega,i,j)(\omega') w_\Theta(t_1,\omega')]$$

Cardaliaguet, Laraki, Sorin (2012)

The family of values has accumulation points.

Any limit point is a viscosity solution of some HJB equation.

Comparison principle implies uniqueness, hence convergence.

2.8 Limit game

2.9 Counterexamples

Vigeral 2013
Ziliotto 2013

oscillations
Sorin and Vigeral 2013

Table of contents

3.1. Presentation

An alternative approach to the repeated game framework while modelizing multistage interactions is to consider a time homogeneous state process Z_t defined on $\mathbb{R}^+ = [0, +\infty)$ with values in Ω, an evaluation given by a continuous decreasing probability density $k(t)$ on $\mathbb{R}^+$ and a discrete time game induced by a partition $\Pi = \{t_0 = 0, t_1, ..., t_n, ...\}$ of $\mathbb{R}^+$.

References include Fleming, Zachrisson, Tanaka, Wakuta, Prieto-Rumeau and Hernandez-Lerma, Neyman ...

The time interval $L_n = [t_{n-1}, t_n[$ (which corresponds to stage n) has duration $\delta_n = t_n - t_{n-1}$.

The law of Z_t on L_n is determined by $Z_{t_{n-1}}$ and the choices i_n, j_n at time t_{n-1}, that last for stage n (hence $(i_t, j_t) = (i_n, j_n)$ for $t \in L_n$).

One considers the asymptotics of the value $v_{\Pi,k}$ as the mesh $\bar{\delta} = \sup \delta_n$ of the partition vanishes.

One can describe these games as having " vanishing stage duration".

3.1. Presentation

An alternative approach to the repeated game framework while modelizing multistage interactions is to consider a time homogeneous state process Z_t defined on $\mathbb{R}^+ = [0, +\infty)$ with values in Ω, an evaluation given by a continuous decreasing probability density $k(t)$ on $\mathbb{R}^+$ and a discrete time game induced by a partition $\Pi = \{t_0 = 0, t_1, ..., t_n, ...\}$ of $\mathbb{R}^+$.

References include Fleming, Zachrisson, Tanaka, Wakuta, Prieto-Rumeau and Hernandez-Lerma, Neyman ...

The time interval $L_n = [t_{n-1}, t_n[$ (which corresponds to stage n) has duration $\delta_n = t_n - t_{n-1}$.

The law of Z_t on L_n is determined by $Z_{t_{n-1}}$ and the choices i_n, j_n at time t_{n-1}, that last for stage n (hence $(i_t, j_t) = (i_n, j_n)$ for $t \in L_n$).

One considers the asymptotics of the value $v_{\Pi,k}$ as the mesh $\bar{\delta} = \sup \delta_n$ of the partition vanishes.

One can describe these games as having " vanishing stage duration".

3.2. Recursive equation

Consider the case where the trajectory of Z_t follows a continuous Markov process specified by a transition rate $\mathbf{q} \in \mathscr{M}$: $\mathbf{q}$ is a real map on $I \times J \times \Omega \times \Omega$ with $\mathbf{q}(i,j)[\omega, \omega'] \geq 0$ if $\omega' \neq \omega$ and $\sum_{\omega' \in \Omega} \mathbf{q}(i,j)[\omega, \omega'] = 0$.

Proposition

The value $v_{\Pi,k}(t,z)$ satisfies the following recursive equation:

$$\begin{aligned} v_{\Pi,k}(t_{n-1}, Z_{t_{n-1}}) &= \mathtt{val}_{X \times Y} \mathsf{E}_{z,x,y}[\int_{t_{n-1}}^{t_n} \mathbf{g}(Z_s, x, y) k(s) ds + v_{\Pi,k}(t_n, Z_{t_n})] \\ &= \mathtt{val}_{X \times Y} [\mathsf{E}_{z,x,y}(\int_{t_{n-1}}^{t_n} \mathbf{g}(Z_s, x, y) k(s) ds) \\ &\quad + \mathsf{P}^{\delta_n}(x,y)[Z_{t_{n-1}}, .] \circ v_{\Pi,k}(t_n, .)] \end{aligned}$$

Proposition

The family of values $\{v_{\Pi,k}\}_{\Pi}$ has at least an accumulation point as $\bar{\delta}$ goes to 0.

Proposition

Any accumulation point W of the family of values $\{v_{\Pi,k}\}_\Pi$ is a viscosity solution of

$$0 = \frac{d}{dt}v(t,z) + \mathtt{val}_{X\times Y}\{\mathbf{g}(z,x,y)k(t) + \mathbf{q}(x,y)[z,.]\circ v(t,\cdot)\}. \qquad (2)$$

Proposition

Equation (2) has a unique viscosity solution, hence the family of values converges as the mesh of Π vanishes.

3.3. Related differential game

The recursive equation is similar to the one induced by the discretization of a deterministic differential game (Fleming) $\mathscr{G}$ defined as follows:

the state space is $\Delta(\Omega)$

the action spaces are $\mathbf{X} = X^{\Omega}$ and $\mathbf{Y} = Y^{\Omega}$

the dynamics is

$$\dot{\zeta}_t(z) = \sum_{\omega \in \Omega} \mathbf{q}(\mathbf{x}(\omega), \mathbf{y}(\omega))[\omega, z]\zeta_t(\omega)$$

of the form

$$\dot{\zeta}_t = f(\zeta_t, \mathbf{x}, \mathbf{y})$$

3.4. Euler scheme

Define, for $\mu \in [0,1]$:

$$\mathbf{T}_\mu = Id + \mu(\mathbf{T} - Id).$$

Then

$$\|U_t(x) - \mathbf{T}_\mu^n(x)\| \leq \|x - \mathbf{T}x\| \sqrt{t\mu + (n\mu - t)^2}.$$

This corresponds to 2 approximations:

i) Comparison of the asymptotic behavior of U_t to iterations of $\{\mathbf{T}_\mu\}$ with size $\mu \leq 1$ and total length t ($n\mu = t$),

$$\|U_t(0) - \mathbf{T}_\mu^n(0)\| \leq \|\mathbf{T}(0)\|\sqrt{t},$$

hence the asymptotic behaviors of $f_n(0)/n$ and v_n are the same.

ii) Comparison on a compact interval $[0,L]$ between U_t and the linear interpolation of $\mathbf{T}_{L/n}^m, m = 0, ..., n$, for $t \in [0,L]$:

$$\|U_t(0) - \mathbf{T}_{L/n}^{nt/L}(0)\| \leq K \frac{L}{\sqrt{n}}.$$

and vanishes as the mesh $1/n$ goes to zero.

More generally for two Eulerian schemes satisfying:

$$x_{k+1} = \mathbf{T}_{\lambda_{k+1}} x_k,$$

$$\hat{x}_{\ell+1} = \mathbf{T}_{\hat{\lambda}_{\ell+1}} \hat{x}_\ell,$$

Vigeral (2009)

$$\|\hat{x}_\ell - x_k\| \le \|\hat{x}_0 - z\| + \|x_0 - z\| + \|z - \mathbf{T}z\| \sqrt{(\sigma_k - \hat{\sigma}_\ell)^2 + \tau_k + \hat{\tau}_\ell}, \quad \forall z \in X,$$

with $x_0 = x$, $\sigma_k = \sum_{i=1}^k \lambda_i$, $\tau_k = \sum_{i=1}^k \lambda_i^2$.
This leads to:

$$\|U_t(0) - \prod_{i=1}^{k} \mathbf{T}_{\lambda_i}(0)\| \le \|\mathbf{T}(0)\| \sqrt{t \max_i \lambda_i},$$

whenever $\sigma_k = t$.

Let G be a stochastic game with transition $P = Id + Q$ on the finite set Ω and G^h the game with stage duration h and transition $Ph = Id + hQ$.

Proposition

If $\mathbf{T}$ is the Shapley operator of G, $\mathbf{T}_h$ is the Shapley operator of the game G^h.

Proposition

$$\|V_n^h - f_{nh}(0)\| \leq Lh\sqrt{n}.$$

Then:

Proposition

The value $\overline{V}_t$ of the continuous time game of length t exists:

$$\|V_{\Pi,t} - \overline{V}_t\| \to 0, \text{ as the mesh of } \Pi \to 0$$

and satisfies:

$$\overline{V}_t = U_t(0).$$

Comments

Compact action spaces
Symmetric information
Similar tools: differential games, repeated games, vanishing stage duration

Double limit

Comments

Compact action spaces
Symmetric information
Similar tools: differential games, repeated games, vanishing stage duration

Double limit

Comments

Compact action spaces
Symmetric information
Similar tools: differential games, repeated games, vanishing stage duration

Double limit

Comments

Compact action spaces
Symmetric information
Similar tools: differential games, repeated games, vanishing stage duration

Double limit