Asymptotic Perron Method for Stochastic Games and Control

Mihai Sîrbu, The University of Texas at Austin

Methods of Mathematical Finance
a conference in honor of Steve Shreve's 65th birthday
Carnegie Mellon University, Pittsburgh, June 1-5, 2015

The important things first

Happy Birthday Steve!

Outline

- Perron's method
- Previous work
  - One-player (control problems) and HJB's
  - Games and Isaacs equations
- Perron's method over asymptotic solutions
  - One-player (control problems) and HJB's
  - Games and Isaacs equations: some modeling and results
- Conclusions

Perron's method

Consider a PDE

$$\mathscr{L}v = 0 \quad + \text{"boundary conditions"}$$

(having the meaning of averaging, i.e. satisfying a maximum principle). For example, $\mathscr{L}v = \Delta v$.

1. Sub-solutions: $\mathscr{L}v \geq 0 \quad + \text{"boundary conditions"}$
2. Super-solutions: $\mathscr{L}v \leq 0 \quad + \text{"boundary conditions"}$.

Perron's method:

1. $v^- =$ supremum of sub-solutions
2. $v^+ =$ infimum of super-solutions

General principle: $v^- = v^+$ is a solution.
Remark: the original meaning of sub and super-solutions was in an averaging sense, without differentiability.

Perron and viscosity solutions

Ishii (87):

If we consider sub and super-solutions in the viscosity sense, then

$$v^+ = v^-$$

is a solution in the viscosity sense (a comparison principle is needed).

- if the PDE is the HJB(I) associated to a control problem (game) it is not clear that this viscosity solution is equal to the value function, unless one has a separate proof of V being a viscosity solution.
- using this route requires first a proof of the DPP (quite delicate/maybe avoidable).

Some previous modification of Perron for HJB's

(joint with E. Bayraktar, SICON 13).

The Control Problem (without technical conditions):

1. The state equation

$$\begin{cases} dX_t = b(t, X_t, u_t)dt + \sigma(t, X_t, u_t)dW_t, \\ X_s = x \in \mathbb{R}^d, \end{cases}$$

starting at initial time s at position x, and which is controlled by one player $u \in U$. The BM W is d'-dimensional.

2. A reward: $g(X_T^{s,x;u})$ at time T, $g : \mathbb{R}^d \to \mathbb{R}$.
3. The value function

$$V(t, x) \triangleq \sup_u \mathbb{E}[g(X_T^{s,x;u})].$$

The HJB and Stochastic semi-solutions

Formal HJB (for the value function)

$$\begin{cases} V_t + \sup_u L_t^u V = 0 \\ V(T,x) = g(x) \end{cases}$$

where $L_t^u V = b(t,x,u)V_x + \frac{1}{2} Tr(\sigma(t,x,u)\sigma^T(t,x,u)V_{xx})$.
Idea: Use semi-solutions in the stochastic sense of Stroock-Varadhan (adapted to the non-linear case).

Definition (Stochastic semi-solutions)

1. $w : [0, T] \times \mathbb{R}^d \to \mathbb{R}$ is a super-solution if, for each s, x and each control u the process $(w(t, X_t^{s,x;u}))_{s \le t \le T}$, is a super-martingale and $w(T, \cdot) \ge g$.
2. $w : [0, T] \times \mathbb{R}^d \to \mathbb{R}$ is a sub-solution if, for each s, x there exist a control u such that the process $(w(t, X_t^{s,x;u}))_{s \le t \le T}$ is a sub-martingale and $w(T, \cdot) \le g$.

Denote by $\mathscr{U}, \mathscr{L}$ the collections of super and sub-solutions in the stochastic sense as above.

Perron's method over stochastic solutions

Remark: by construction we have

1. if $w \in \mathscr{U}$ then $w \geq V$
2. if $w \in \mathscr{L}$ then $w \leq V$

Perron's method:

1. $v^+ \triangleq \inf_{w \in \mathscr{U}} w \geq V$,
2. $v^- \triangleq \sup_{w \in \mathscr{L}} w \leq V$,

so

$$v^- \leq V \leq v^+.$$

Theorem (Bayraktar, S.)

Under some technical conditions, v^+ is a USC viscosity sub-solution and v^- is a LSC viscosity super-solution.

Corollary

A comparison result for semi-continuous viscosity solutions implies

1. *V the unique continuous viscosity solution of the HJB*
2. *the (DPP) holds, i.e. $V(s,x) = sup_u \mathbb{E}[V(\tau, X_\tau^{s,x;u})]$ for all stopping times $\tau \geq s$.*

Some previous work on games

(S. SICON 14):

- model symmetric games over feedback strategies, but rebalanced at discrete stopping rules (elementary feedback strategies for both players)
- use a Perron construction over sub/super-martingales to treat the case of zero-sum symmetric games (where DPP is known to be particularty delicate)
- the definition of stochastic semi-solutions has to be changed in a non-trivial way to account for the (dynamic) strategic behavior
- upper and lower value functions are solutions of the Isaacs equations
- (versions) of the (DPP) for games are obtained

More about games to follow.

Some comments (on previous work)

- the proofs are rather elementary. The analytical part mimics the proof of Ishii and then we apply Itô to the (smooth) test functions.
- amounts to verification for non smooth solutions
- proving a DPP is particularly complicated for games (Fleming-Souganidis)

Perron's method over asymptotic semi-solutions

Goal: design a modification of Perron's method that works well with discretized Markov controls/strategies (in a similar elementary way).

Simple Markov strategies (in one player problems)

Fix $0 \leq s \leq T$. A path of the state equation is a $y \in C[s, T]$.

Definition (time grids, simple Markov strategies)

1. A time grid for $[s, T]$ is a finite sequence π of $s = t_0 < t_1 < \cdots < t_n = T$.
2. Fix a time grid π as above. A feedback strategy

$$\alpha : [s, T] \times C[s, T] \to U,$$

is called simple Markov over π if there exist some measurable functions $\alpha_k : \mathbb{R}^d \to U, k = 1, \ldots, n$ such that

$$\alpha(t, y(\cdot)) = \sum_{k=1}^{n} 1_{\{t_{k-1} < t \leq t_k\}} \alpha_k(y(t_{k-1})).$$

- $\mathscr{A}^M(s, \pi)$: simple Markov strategies over π
- $\mathscr{A}^M(s) \triangleq \bigcup_\pi \mathscr{A}^M(s, \pi)$ all simple Markov strategies

Use only simple Markov strategies

Define

$$V_\pi(s,x) \triangleq \sup_{\alpha \in \mathscr{A}^M(s,\pi)} \mathbb{E}[g(X_T^{s,x;\alpha,v})] \le V(s,x).$$

and the value function over all simple Markov strategies

$$V_M(s,x) \triangleq sup_{\alpha \in \mathscr{A}^M(s)} \mathbb{E}[g(X_T^{s,x;\alpha})] = sup_\pi V_\pi(s,x) \le V(s,x).$$

Natural question: can the value function be approximated by discretized Markov strategies, i.e. $V_M = V$?

Under some technical conditions, Krylov has proved this property using PDE arguments and regularity properties of the value function (obtaining the DPP also).
Some Perron-type construction may be simpler (work for games?)

- the previous method cannot show that. Why? Because, if $w \in \mathscr{L}$ then we do not know that $w \le V^M$
- if we use simple Markov strategies in the Def of $\mathscr{L}$ the Perron Construction does not work anymore.

Need for new concept of sub-solution of the HJB

Modify the definition of sub-solutions (only) such that

1. sub-solutions w stay below V_M
2. $w^- \triangleq \sup w \leq V_M$ is still a viscosity super-solution

Intuition: consider a (strict) smooth sub-solution of the HJB

$$\begin{cases} w_t + \sup_u L_t^u w > 0 \\ V(T,x) \leq g(x). \end{cases}$$

Start at time s at position x with the feedback control attaining the sup above (argmax $u^* = u^*(s,x)$) and hold it constant until T. The process

$$(w(t, X_t^{s,x;u^*}))s \leq t \leq T$$

is a sub-martingale until the first time τ when $w_t + L_t^{u^*} w \leq 0$. However,

$$\mathbb{P}(\tau \leq t) = o(t-s).$$

In between s and t we have a sub-martingale property plus an $o(t-s)$ correction.

Asymptotic sub-solutions, precise definition

Take the previous observation and make it (more) dynamic.

Definition (Asymptotic Stochastic Sub-Solutions)

$w : [0, T] \times \mathbb{R}^d \to \mathbb{R}$ is called an asymptotic sub-solution if it is (bounded), continuous and satisfies $w(T, \cdot) \leq g(\cdot)$.
There exists a gauge function $\varphi = \varphi_w : (0, \infty) \to (0, \infty)$, depending on w such that

1. $\lim_{\varepsilon \searrow 0} \varphi(\varepsilon) = 0$,
2. for each s, and for each time $s \leq r \leq T$, there exists a measurable function $\alpha : \mathbb{R}^d \to U$ such that, for each $\xi \in \mathscr{F}_r$, if we start at the time r at the (random) condition ξ and keep the feedback control $\alpha(\xi)$ constant on $[r, T]$ we have

$$w(r, \xi) \leq \mathbb{E}[w(t, X_t^{s,\xi;\alpha(\xi)}) | \mathscr{F}_r] + (t - r)\varphi(t - r) \ a.s.$$

Denote by $\mathscr{L}_a$ the set of asymptotic sub-solutions.

Important property of asymptotic sub-solutions

Lemma

Any $w \in \mathscr{L}_a$ satisfies $w \leq V_M \leq V$.

Proof:

1. Fix ε, then fix δ such that $\varphi(\delta) \leq \varepsilon$. Choose $\|\pi\| \leq \delta$
2. construct, recursively, going from time t_{k-1} to time t_k, some measurable $\alpha_k : \mathbb{R}^d \to U$ satisfying the Definition 5.
3. for $\alpha(t, y(\cdot)) = \sum_{k=1}^{n} 1_{\{t_{k-1} < t \leq t_k\}} \alpha_k(y(t_{k-1}))$ we have

$$w(t_{k-1}, X_{t_{k-1}}^{s,x;\alpha}) \leq \mathbb{E}[w(t_k, X_{t_k}^{s,x;\alpha}) | \mathscr{F}_{t_{k-1}}] + \varepsilon(t_k - t_{k-1})$$

Telescoping sum: $w(s,x) \leq \mathbb{E}[w(T, X_T^{s,x;\alpha})] + \varepsilon(T - s)$.

Summary: if $|\pi| \leq \delta(\varepsilon)$ there exists $\alpha \in \mathscr{A}^M(s,\pi)$ such that

$$w(s,x) \leq \mathbb{E}[g(X_T^{s,x;\alpha})] + \varepsilon \times (T - s) \leq V_\pi(s,x) + \varepsilon(T - s).$$

Letting $\varepsilon \searrow 0$ we obtain the conclusion.

Perron method over asymptotic sub-solutions

Define

$$w^- \triangleq \sup_{w \in \mathscr{L}_a} w \leq V_M \leq V.$$

Theorem (Perron over asymptotic sub-solutions, HJB)

The function w^- is an LSC viscosity super-solution of the HJB.

- the proof is (again) based on the analytic construction of Ishii
- negating the viscosity super-solution property, the test function is a strict classic sub-solution (locally)
- apply Itô to the test function, together with a very similar observation we made for strict classic sub-solution, i.e. sub-martingale property plus an $o(t-r)$ correction.

Corollary

A comparison result (which holds under some technical assumptions) implies that $V_M = V$ and, actually, $V_\pi \nearrow V$ as $\|\pi\| \searrow 0$ uniformly on compacts.

Overview of the method

- adding a correction to the sub-martingale property (over feedback controls), we have a (rather elementary) tool to approach Dynamic Programing over simple Markov strategies for control problems.
- we can apply it to games, where the Dynamic Programming arguments are more difficult

Zero-sum differential games

1. The state equation

$$\begin{cases} dX_t = b(t, X_t, u_t, v_t)dt + \sigma(t, X_t, u_t, v_t)dW_t, \\ X_s = x \in \mathbb{R}^d, \end{cases}$$

starting at initial time s at position x, and which is controlled by both players.

2. A penalty/reward: (if a genuine game) the second player (v) pays to the first player (u) the amount $g(X_T^{s,x;u,v})$ at time T.

Formal zero-sum game

$$\sup_u \inf_v \mathbb{E}[g(X_T^{s,x;u,v})], \quad \inf_v \sup_u \mathbb{E}[g(X_T^{s,x;u,v})].$$

Inputs of the game

- the coefficients of the state equation b, σ
- the sets where the two players can take action: $u \in U$, $v \in V$
- for each initial time s, a (fixed) probability space $(\Omega, \mathscr{F}, \mathbb{P})$ and a Brownian motion W with respect to the filtration $(\mathscr{F}_t^s)_{s \leq s \leq T}$. Allow the filtration to be larger than the ones generated by W.

Standing assumptions

- g is continuous and bounded,
- b and σ are continuous on their whole corresponding domains and (locally) uniformly Lipschitz in x and have linear growth.
- U, V are compact

Objective

Look at

- modeling of such games
- apply the Asymptotic Perron tool to the dynamic programming analysis:
  - relate the value functions to Dynamic Programming Equation(s)
  - (more important) study existence of value for the game: $\sup\inf = \inf\sup$

Why?

- Unlike one player (control) problem, it is unclear what to use for u, v. There is no widely accepted notion of strategy/control, so no "canonical" formulation anyway (Cardaliaguet lecture notes).
- the dynamic programming analysis (i.e. relation to Isaacs equation) is significantly more complicated (see Fleming-Souganidis, for an Elliott-Kalton formulation of the game)

Some literature on games

Very selective list of works

- Isaacs: deterministic case over (heuristically) feedback strategies
- Krasovskii-Subbotin: engineering-type literature, very interesting modeling over feedback strategies and counter-strategies
- Elliott-Kalton: deterministic case, use so called strategies for the stronger player, (open loop) controls for weaker player (no symmetric formulation)
- Fleming-Souganidis: use Elliott-Kalton strategies in stochastic games. Prove the value functions are viscosity solutions of DPE (Bellman-Isaacs equations). No symmetric formulation.
- large interesting literature on games studied using BSDE's: Hamadene and Lepeltier, El Karoui-Hamadene, Buckdahn-Li (more others)

Literature cont'd

1. Most of the mathematical literature uses an Elliott-Kalton formulation. Value functions do not compare by definition, but after complicated analysis. Actually, the value functions belong to different games, depending on one player or another having an informational advantage.

2. Ekaterinburg school of games (mainly Krasovskii-Subbotin): use (discretized) feedback strategies (mostly in) deterministic games. Recently, symmetric formulation of games where both players use strategies based only on the knowledge of the past of the state have been re-considered:

- Cardaliaguet-Rainer '08 (strong formulation, path-wise feedback strategies with delay)
- Pham -Zhang '12 (path-wise feedback strategies, discretized in time and space, called "feedback controls")

The model(s) we use for games

We continue the line of modeling with feedback strategies and counter-strategies in Krasovskii-Subbotin and recently in Fleming-Hernandez-Hernandez.

The lower Isaacs equation

Formal equation:

$$\begin{cases} V_t + \sup_u \inf_v L_t^{u,v} V = 0 \\ V(T, x) = g(x) \end{cases}$$

where

$L_t^{u,v} V = b(t, x; u, v) V_x + \frac{1}{2} Tr(\sigma(t, x; u, v) \sigma^T(t, x; u, v) V_{xx}).$

- analytic representation of a game where v has an informational advantage over u
- would like to model this situation using feedback strategies

Feedback strategies

(Krasovskii-Subbotin, Cardaliaguet-Rainer, Pham-Zhang)
The player using such strategy

- observes the state only
- does not observe the other player's actions
- do not observe the noise

Therefore, if $C([s, T])$ is the path-space for the state, we can consider

$$\alpha : (s, T] \times C[s, T] \to U$$

OR

$$\beta : (s, T] \times C[s, T] \to V$$

adapted to the filtration $\mathbb{B} = (\mathscr{B}_t)_{s \leq t \leq T}$ defined by

$$\mathscr{B}_t^s \triangleq \sigma(y(u), s \leq u \leq t), \quad 0 \leq t \leq T.$$

As for one player, we denote by y the paths of the state equation.

Feedback counter-strategies for v

(following Krasovskii-Subbotin)

- player u can only see the state
- the player v observes the state, and, in addition, the control u (in real time)

Intuition: the advantage of the player v actually comes **only** from the local observation of u_t. A counter-strategy for v depends on

1. the whole past of the state process X up to the present time t
2. (only) the current value of the adverse control u_t.

Definition (Feedback Counter-Strategies)

Fix a time s. A feedback counter-strategy for the player v is a mapping

$$\gamma : [s, T] \times C[s, T] \times U \to V,$$

which is measurable with respect to $\mathscr{P}^s \otimes \mathscr{U}^b / \mathscr{V}$. The action of the player v is $v_t = \gamma(t, X_\cdot, u_t)$.

More ways to think about the game

where v has an advantage over player u:

1. the lower value of a symmetric game over feedback strategies i.e. u announces the strategy to v (Pham-Zhang, S.)
2. the value function of a robust control problem where u is an intelligent maximizer (using feedback strategies) and v is a possible worst case scenario modeling Knigthian uncertainty (see S.), or
3. the genuine value of a sup-inf/inf-sup non symmetric game over feedback strategies vs. feedback counter-strategies (Krasovskii-Subbotin, Fleming -Hernandez -Hernandez).

Strategies/counter-strategies and open-loop controls

Well-posedness of the state eq with feedback strategies or counter-strategies is important, but we disregard it here.

Denote by $\mathscr{U}(s)$ and $\mathscr{V}(s)$ the set of **open-loop controls** for the u-player and the v-player, respectively. Precisely,

$$\mathscr{V}(s) \triangleq \{v : [s, T] \times \Omega \to V | \text{ predictable } \text{ w.r.t } \mathbb{F}^s = (\mathscr{F}_t^s)_{s \le t \le T}\},$$

and a similar definition is made for $\mathscr{U}(s)$.

Notation/symbols

1. α, β for the feedback strategies of players u and v,
2. u, v for the open loop controls,
3. γ for the feedback counter-strategy of the second player v.

Value functions

1. lower value function for the symmetric game in between two feedback players:

$$V^{-}(s,x) \triangleq \sup_{\alpha \in \mathscr{A}(s)} \left(\inf_{\beta \in \mathscr{B}(s)} \mathbb{E}[g(X_T^{s,x;\alpha,\beta})] \right).$$

2. the value of a robust control problem where the intelligent player u uses feedback strategies and the open-loop controls v parametrize worst case scenarios/Knightian uncertainty:

$$V_{rob}^{-}(s,x) \triangleq \sup_{\alpha \in \mathscr{A}(s)} \left(\inf_{v \in \mathscr{V}(s)} \mathbb{E}[g(X_T^{s,x;\alpha,v})] \right).$$

3. Lower and upper values of a non-symmetric game over strategies α vs counter-strategies γ

$$W^{-}(s,x) \triangleq \sup_{\alpha \in \mathscr{A}(s)} \left(\inf_{\gamma \in \mathscr{C}(s)} \mathbb{E}[g(X_T^{s,x;\alpha,\gamma})] \right) \leq$$

$$\inf_{\gamma \in \mathscr{C}(s)} \left(\sup_{\alpha \in \mathscr{A}(s)} \mathbb{E}[g(X_T^{s,x;\alpha,\gamma})] \right) \triangleq W^{+}(s,x).$$

Value functions cont'd

For mathematical reasons, we define yet another value function

$$V^{+}_{rob}(s,x) \triangleq \inf_{\gamma \in \mathscr{C}(s)} \left(\sup_{u \in \mathscr{U}(s)} \mathbb{E}[g(X_T^{s,x;u,\gamma})] \right) \geq W^{+}(s,x). \qquad (1)$$

We attach to V^{+}_{rob} the meaning of some robust optimization problem, but this is not natural, since the intelligent optimizer v can see in real time the "worst case scenario".

By simple observation

$$V^{-}_{rob} \leq W^{-} = V^{-} \leq W^{+} \leq V^{+}_{rob}.$$

Simple Markov strategies/counter-strategies definition

Recall we have defined Markov strategies already for one-player.

Definition (simple counter-strategies)

Fix $0 \leq s \leq T$. Fix a time grid π as above.
A counter-strategy $\gamma \in \mathscr{C}(s)$ is called a simple Markov counter-strategy **over** π if there exist some functions $\eta_k : \mathbb{R}^d \times U \to V, k = 1, \ldots, n$ measurable, such that

$$\gamma(t, y(\cdot), u) = \sum_{k=1}^{n} 1_{\{t_{k-1} < t \leq t_k\}} \eta_k(y(t_{k-1}), u).$$

Notation:

- $\mathscr{C}^M(s, \pi)$ is set of simple Markov counter strategies over π
- $\mathscr{C}^M(s) \triangleq \bigcup_\pi \mathscr{C}^M(s, \pi)$ is the set of all simple Markov counter-strategies.

State equation and value functions with Markov strategies

The state equation is well posed if one player is using simple Markov strategies or counter-strategies and the opposing player is using open-loop controls.

If we use Markov strategies or counter-strategies in the LHS, RHS they are well defined.

$$V_M^-(s,x) \triangleq \sup_{\alpha \in \mathscr{A}^M(s)} \left(\inf_{v \in \mathscr{V}(s)} \mathbb{E}[g(X_T^{s,x;\alpha,v})] \right) \leq V_{rob}^-(s,x)$$

as well as

$$V_M^+(s,x) \triangleq \inf_{\gamma \in \mathscr{C}^M(s)} \left(\sup_{u \in \mathscr{U}(s)} \mathbb{E}[g(X_T^{s,x;u,\gamma})] \right) \geq V_{rob}^+(s,x).$$

Main result/games

Theorem (S.)

Under the standing assumptions

- *all value functions are equal:* $V_M^- = V_M^+ =$ *the uviscosity solution of the lower Isaacs equation.*
- *the game* α *vs* γ *has* ε*-saddle points over simple Markov* α*'s and* γ*'s.*

(More) precisely: $\forall\ N,\ \varepsilon > 0,\ \exists\ \delta(N,\varepsilon) > 0$ *such that* $\forall s, \forall\ |\pi| \leq \delta,\ \exists\ \hat{\alpha} \in \mathscr{A}^M(s,\pi), \hat{\gamma} \in \mathscr{C}^M(s,\pi)$ *for which*

$$0 \leq W^-(s,x) - \inf_{v \in \mathscr{V}(s)} \mathbb{E}[g(X_T^{s,x;\hat{\alpha},v})] \leq \varepsilon$$

and

$$0 \leq \sup_{u \in \mathscr{U}(s)} \mathbb{E}[g(X_T^{s,x;u,\hat{\gamma}})] - W^+(s,x) \leq\ \varepsilon.$$

Conclusions

- dynamic programming arguments (the proof of the DPP) are clean only for discrete-time problems (Blackwell, Berstekas and Shreve)
- the (direct) proof of the DPP is more delicate for continuous-time problems
- Perron's method and its modifications appear to be useful in the dynamic programming analysis of continuous-time optimization problems, allowing for a verification without smoothness
- the asymptotic modification of Perron (adding a correction to martingales) is the last step in the program. It provides the strongest conclusion expected in a non-smooth case: cannot get more without smoothness of the solutions of the HJB(I)
- the case of games is of particular interest, since usual dynamic programming arguments run into even more difficulties there (Fleming-Souganidis)

For games

There are three possible interpretations of the game that lead to the lower Isaacs equation

1. lower value of a a symmetric over (restricted) feed-back strategies
2. a control problem of model uncertainty
3. the true value of a game over strategies vs. counter-strategies

Using Perron method over asymptotic solutions, we connect (in a unified way) all three problems.

More important: find an (approximate) saddle point over Markov strategies and counter-strategies.

Remarks:

- it is important how we define both the game and the semi-solutions (asymptotic correction)
- we have flexibility in defining semi-solutions (to make the proofs work, but keep comparison with the value function obvious)
- filtrations do not matter (just as if solutions to the Isaacs equation were smooth).