On the principle of smooth fit in optimal stopping problems

Amir Aliev
Moscow State University, Faculty of Mechanics and Mathematics, Department of Probability Theory, 119992, Moscow, Russia.

Keywords: optimal stopping, smooth fit, free-boundary problems
AMS: 60G40

Abstract

Given a function $G: \mathbb{R}^n \to \mathbb{R}$ and a Markov process $(X_t)_{t \geq 0}$ consider the optimal stopping problem $V(x) = \sup_{\tau} E_x G(X_\tau)$. It is well known that the optimal stopping time is given by $\tau^* = \inf\{t : X_t \in D\}$ where the set $D$ is given by $D := \{x : V(x) = G(x)\}$ and $V$ is the smallest superharmonic function that dominates $G$. This characterization could be reformulated as the free-boundary problem, where both $V$ and $D$ have to be determined. To get the explicit solution of the arising differential equations one has to impose additional boundary conditions on $V$. Often, this conditions follow from the principle of smooth fit that states $V' = G'$ on $\partial D$ (in one-dimensional case) under various assumptions on $G$ and $X$. This results are well-known and widely used, but only weakest of them could be extended directly to multi-dimensional case due to topological complications. Therefore, "smooth fit" is considered in topology weaker than standard topology of $\mathbb{R}^n$. In the article sufficient conditions are presented for this principle of smooth fit to hold in multi-dimensional case.

1. Introduction

The theory of optimal stopping the principle of smooth fit basically appears as "natural" condition that is "usually" satisfied, starting with the work of Mikhalevich (Mikh58). Kolmogorov once said that "diffusion doesn't like angles". As the candidate solution of the problem could be verified, this condition could be imposed without rigorous justification. Nevertheless, the theorems giving the sufficient conditions for this principle to hold are useful not only from theoretical point of view, but also to understand the cases when smooth fit fails to hold, and thus avoid unnecessary verifications.

In one-dimensional case for time-homogeneous diffusion process this theorems are well-known, starting with the work of Grigelionis and Shiryaev (GSh66) (see Shiryaev and Peskir (PSh06) for detailed reference). Generally speaking, they demand both "regular" behavior of the process and smoothness of the gain function. More precisely, they are given by

**Theorem 1.** Let functions $G$ and $S$ (scale function of $X$) be differentiable at $b$. Then function $V$ is also differentiable at $b$ and $V'(b) = G'(b)$. 

---

15th European Young Statisticians Meeting 15thEYSM
September 10-14, 2007 Castro Urdiales (Spain)
Theorem 2. Let the distribution of the increments of $X$ be independent from the starting point and let $b$ be regular for the set $D$ of stopping with respect to the process $X$. Let, moreover, $G \in C^1(b)$. Then $V \in D(b)$ and $V'(b) = G'(b)$.

The examples when smooth fit fails to hold due to lack of this properties are also well-known. The goal of this paper is to treat the multi-dimensional case, which is rather complicated because in the topology of $\mathbb{R}^n$, $n \geq 2$ the behavior of the process is harder to treat. Thus, proof of main result from one-dimensional case couldn’t be extended directly to the general case. In this paper we prove slightly weaker smooth fit principle in Theorem 1 introducing some artificial topology. in some cases this result could be strengthened, as we see in Theorem 2, by imposing different (and rather strict) conditions on the process.

2. Probability model

We consider $(X_t)_{t \geq 0}$ - a diffusion process in $\mathbb{R}^n$ with drift vector $a(x)$ and diffusion matrix $B(x)$ defined on the standard probability space $(\Omega, \mathcal{F}, \mathcal{F}_t, P)$, with $X_t$ adapted to the filtration $\mathcal{F}_t$. Starting with $G(x) : \mathbb{R}^n \to \mathbb{R}$ define $V(x) = \sup_\tau E_x G(X_\tau)$. Finding $V$ and optimal stopping time is the optimal stopping problem. Let functions $V$ and $G$ be equal in $b$ (optimal point).

3. Definitions

Recall, that in multi-dimensional case the scale function of $X$ is the solution to the differential equation

$$
(a(x), s'(x)) + \frac{1}{2} S p B(x) B^*(x) s''(x) = 0.
$$

Such a solution is not unique, and for our purposes the existence of $n$ linearly independent in $b$ solutions would suffice. This vector-solution would be further denoted by $S$. The main feature of the scale function we will use is that the process $S(X_t)$ is a martingale (see [3], ch. 3, §3).

Moreover, define the topology $\mathcal{W}$ in $\mathbb{R}^n$ as the sets $W$ such that for any $x \in W$ and a sequence of contours $\gamma_n$, tending to the point $x$

$$
\lim_{\gamma_n \to x} \frac{E_x \| X_{\tau(\gamma_n)} - x \| I\{X_{\tau(\gamma_n)} \not\in W\}}{E_x \| X_{\tau(\gamma_n)} - x \|} = 0 \quad (1)
$$

(here and further $\| \cdot \|$ designated the standard norm of $\mathbb{R}^n$). It could be easily shown that this is indeed a topology, and that it contains all sets open in the standard topology of $\mathbb{R}^n$. When $n = 1$ and rather weak conditions imposed on the process, $\mathcal{W}$ coincides with the standard topology of $\mathbb{R}^n$, but in the multi-dimensional case it is richer.
4. Main theorem

**Theorem 1.** Let functions $G$ and $S$ be differentiable in $b$, and Jacobian determinant of $S$ be non-zero. Then function $V - G$ is infinitesimal in $W$ concerning $||x - b||$ as $x \to b$.

The general theory of optimal stopping states that the function $V$ is superharmonic. Thus the inequality $G(b) = V(b) \geq E_b V(X_\tau)$ holds for any stopping time $\tau$ and, extracting $E_b G(X_\tau)$ from both sides we arrive to

$$E_b (G(b) - G(X_\tau)) \geq E_b (V(X_\tau) - G(X_\tau)) \geq 0 \quad (2)$$

Fix $\varepsilon > 0$ and define $B := \{x : V(x) - G(x) < \varepsilon ||b - x||\}$. Notice, that for the diffusion processes function $V$ is continuous (see [1], ch. 3, 2.1). Hence function $V(x) - G(x)$ is also continuous, and for any $x$ not equal to $b$ there exists a neighborhood $U(x)$ such that $U(x) \subset B$. It follows that (1) holds for any $x \in B$, except for $b$.

Notice, that if $S$ is differentiable in $b$, then for any stopping time $\tau$ we get (with the condition of uniform integrability of $S(X_{t\wedge \tau})$) the following property $0 = E_b (S(X_\tau) - S(b)) = E_b J \times (X_\tau - b) + \alpha_1 (X_\tau - b)$, where $J$ - Jacobi matrix of $S$ in $b$, and $\alpha_1 (x) \sim o(||x||)$ as $x \to 0$. Together with $\det S(b)$ we conclude that $E_b (X_\tau - b) = -J^{-1} E_b \alpha_1 (X_\tau - b)$. The left-hand side of (2) could be rewritten (with the residual $\alpha_2 \sim o(||x||)$) as $E_b (G(b) - G(X_\tau)) = E_b (\nabla G \times (X_\tau - b) + \alpha_2 (X_\tau - b)) = -\nabla G J^{-1} E_b \alpha_1 (X_\tau - b) + E_b \alpha_2 (X_\tau - b)$. Thus, the left-hand side of (2) could be expressed (denoting the sum of infinitesimals $a$) as $E_b \alpha(X_\tau - b)$. Consider the sequence of contours tending to $b$, i.e. $\gamma_n \to b : \forall \delta \exists N \forall n > N \gamma_n \in U_\delta (b)$ (here $U_\delta (b)$ - $\delta$-neighborhood of $b$). Fix $\beta > 0$. Then there exist $\delta$ and $N$ such that $0 < ||t|| < \delta \Rightarrow |\alpha(t)| < \beta ||t||$ and $\forall n > N \gamma_n \in U_\delta (b)$. Hence, for all $n > N$ we have

$$\psi(\gamma_n) = \frac{E_b (G(b) - G(X_\tau(\gamma_n)))}{E_b ||X_\tau(\gamma_n) - b||} \leq \frac{E_b |\alpha(X_\tau(\gamma_n) - b)|}{E_b ||X_\tau(\gamma_n) - b||} < \beta.$$ 

It follows, that $\lim_{\gamma_n \to b} \psi(\gamma_n) = 0$, and with (2) in mind we get

$$\lim_{\gamma_n \to b} \frac{E_b (V(X_\tau(\gamma_n)) - G(X_\tau(\gamma_n)))}{E_b ||X_\tau(\gamma_n) - b||} = 0.$$ 

Obviously, $E_b (V(X_\tau(\gamma_n)) - G(X_\tau(\gamma_n))) = E_b (V(X_\tau(\gamma_n)) - G(X_\tau(\gamma_n)))I\{X_\tau(\gamma_n) \in B\} + E_b (V(X_\tau(\gamma_n)) - G(X_\tau(\gamma_n)))I\{X_\tau(\gamma_n) \notin B\} \geq \varepsilon E_b ||X_\tau(\gamma_n) - b|| I\{X_\tau(\gamma_n) \notin B\}$. Thus

$$\lim_{\gamma_n \to b} \frac{E_b ||X_\tau(\gamma_n) - b|| I\{X_\tau(\gamma_n) \notin B\}}{E_b ||X_\tau(\gamma_n) - b||} = 0.$$
It proves that (1) holds also in $b$, and the set $B$ is open in the topology $\mathcal{W}$. It is valid for any $\varepsilon > 0$, hence $V(x) - G(x) = o(\|x - b\|)$, $x \to b$ in the topology $\mathcal{W}$.

5. Additional theorem

**Theorem 2.** Let the distribution of the increments of $X$ be independent from the starting point and let $b$ be regular for the set $D$ of stopping with respect to the process $X$. Let, moreover, function $G$ be differentiable in any direction and let the derivatives be bounded in the neighborhood of the border $\partial D$. Then $V$ is differentiable in any direction in $b$ and its derivatives coincide with the corresponding derivatives of $G$.

The proof repeats the steps of the similar theorem in one-dimensional case (see [1], ch. 4, 2.1, th. 2). Fix $\varepsilon > 0$ and a direction $\vec{m}$ with $||\vec{m}|| = 1$. From $V(b) = G(b)$ and $V \geq G$ we get immediately the inequality

$$\varepsilon^{-1} (V(b + \varepsilon \vec{m}) - V(b)) \geq \varepsilon^{-1} (G(b + \varepsilon \vec{m}) - G(b)),$$

and passing to the limit $\varepsilon \downarrow 0$ we get

$$\lim_{\varepsilon \to 0} \varepsilon^{-1} (V(b + \varepsilon \vec{m}) - V(b)) \geq G'_{\vec{m}}(b). \quad (3)$$

Let $\tau^\varepsilon := \inf\{t : X_t \in D\}$ (optimal stopping moment) and $\sigma^\varepsilon = \inf\{t : b + \varepsilon \vec{m} + X_t \in D\}$. From $V(b + \varepsilon \vec{m}) = E_{b+\varepsilon \vec{m}} G(X_{\tau^\varepsilon}) = E_0 G(b + \varepsilon \vec{m} + X_{\sigma^\varepsilon})$ and $V(b) \geq E_b V(X_{\sigma^\varepsilon}) \geq E_0 G(b + X_{\sigma^\varepsilon})$ we get

$$\varepsilon^{-1} (V(b + \varepsilon \vec{m}) - V(b)) \leq \varepsilon^{-1} (E_0 [G(b + \varepsilon \vec{m} + X_{\sigma^\varepsilon}) - G(b + X_{\sigma^\varepsilon})]).$$

It follows from Lagrange theorem and the existence of derivative $\frac{\partial G}{\partial \vec{m}}$ we get

$$\varepsilon^{-1} (E_0 [G(b + \varepsilon \vec{m} + X_{\sigma^\varepsilon}) - G(b + X_{\sigma^\varepsilon})]) = E_0 G'_{\vec{m}}(b + \theta \varepsilon \vec{m} + X_{\sigma^\varepsilon}), \quad \theta \in [0; 1].$$

If $|G'_{\vec{m}}(b + \theta \varepsilon \vec{m} + X_{\sigma^\varepsilon})| \leq Z$, where $Z$ is an integrable random variable, then we can pass to the limit $\varepsilon \downarrow 0$ according to Lebesgue dominated convergence theorem. In our case, passing to the limit is justified as the derivatives are bounded in the neighborhood of the border. From the regularity condition we get $\sigma^\varepsilon \to 0$, which gives us (as $\varepsilon \downarrow 0$)

$$\lim_{\varepsilon \to 0} \varepsilon^{-1} (V(b + \varepsilon \vec{m}) - V(b)) \leq \lim_{\varepsilon \to 0} E_0 G'_{\vec{m}}(b + \theta \varepsilon \vec{m} + X_{\sigma^\varepsilon}) = G'_{\vec{m}}(b).$$

Together with (3) it proves the theorem.

**Acknowledgements:** The author is grateful to professor Shiryaev for the formulation of the problem.
6. Bibliography


