Available online at www.sciencedirect.com

ScienceDirect

IFAC Papers Online
CONFERENCE PAPER ARCHIVE

IFAC-PapersOnLine 49-18 (2016) 310–315

Necessary Optimality Conditions for Weak Local Minima in Stochastic Control ⋆

Helene Frankowska * **Haisen Zhang** ** **Xu Zhang** ***

* *CNRS, UPMC Univ Paris 06, 75252 Paris, France (e-mail: helene.frankowska@imj-prg.fr).*
** *School of Mathematics and Statistics, Southwest University, Chongqing 400715, China (e-mail: haisenzhang@yeah.net)*
*** *School of Mathematics, Sichuan University, Chengdu 610064, China (e-mail: zhang_xu@scu.edu.cn)*

Abstract: This paper is devoted to the first and second order necessary optimality conditions for stochastic optimal control problems of the Bolza type. The control system is governed by a stochastic differential equation whose drift and diffusion terms are control dependent and the set of controls may be nonconvex. The optimal controls under consideration are those providing the weak local minima. The derived first order necessary condition involves one adjoint equation and a pointwise variational inequality. They are very similar to the known necessary conditions of the deterministic optimal control theory. The second order necessary condition is stated in the integral form and involves two adjoint equations. For sufficiently regular singular weak local minimizers this second order integral inequality implies a pointwise condition. To obtain these results we use the classical variational approach reinforced by the set-valued analysis and the Malliavin calculus. The proofs of our results being quite long and technical will appear elsewhere.

Keywords: Stochastic optimal control, maximum principle, Malliavin calculus, variational equation, adjoint equation, second order necessary conditions.

1. INTRODUCTION

Let $T > 0$, $(\Omega, \mathcal{F}, \mathbb{F}, P)$ be a complete filtered probability space (satisfying the usual assumptions) on which a 1-dimensional standard Wiener process $W(\cdot)$ is defined such that $\mathbb{F} = \{\mathcal{F}_t\}_{0\le t\le T}$ is the natural filtration generated by $W(\cdot)$ (augmented by all the P-null sets).

This paper is devoted to the following stochastic optimal control problem of the Bolza type:

$$\text{Minimize } \ J(u(\cdot))$$

over process-control pairs of the stochastic control system

$$\begin{cases} dx(t) = b(t, x(t), u(t))dt + \sigma(t, x(t), u(t))dW(t) \\ x(0) = x_0, \end{cases} \tag{1}$$

with the cost functional $J(\cdot)$ being given by

$$J(u(\cdot)) = \mathbb{E}\Big(\int_0^T f(t, x(t), u(t))dt + g(x(T))\Big). \tag{2}$$

Here, $m, n \in \mathbb{N}$, $x_0 \in \mathbb{R}^n$, $U \subset \mathbb{R}^m$, b, $\sigma : [0,T] \times \mathbb{R}^n \times U \times \Omega \to \mathbb{R}^n$, $f : [0,T] \times \mathbb{R}^n \times U \times \Omega \to \mathbb{R}$ and $g : \mathbb{R}^n \times \Omega \to \mathbb{R}$ are given; $u(\cdot)$ is a control taking values in U, while $x(\cdot)$ is the corresponding solution to (1). As usual, when the context is clear, we omit the ω $(\in \Omega)$ argument in the functions under consideration.

Denote by $\mathcal{B}(X)$ the Borel σ-field on a metric space X, and by $\mathcal{U}_{ad}$ the set of $\mathcal{B}([0,T]) \otimes \mathcal{F}$-measurable and $\mathbb{F}$-adapted stochastic processes with values in U such that $\|u\|_2 := \big(\mathbb{E}\int_0^T |u(t)|^2 dt\big)^{1/2} < \infty$. Every $u \in \mathcal{U}_{ad}$ is called an admissible control, the corresponding stochastic process x described by (1) is called an admissible state, and (x, u) is called an admissible pair. A control $\bar{u} \in \mathcal{U}_{ad}$ is optimal if

$$J(\bar{u}) = \inf_{u\in\mathcal{U}_{ad}} J(u). \tag{3}$$

It is called a weak local minimizer if there exists a $\delta > 0$ such that $J(u) \ge J(\bar{u})$ for all $u \in \mathcal{U}_{ad}$ satisfying $\|u - \bar{u}\|_2 < \delta$. We investigate here some first and second order necessary optimality conditions satisfied by the weak local minimizers.

Earlier results concerning the first order optimality conditions can be found in Bensoussan (1981), Bismut (1978), Haussmann (1976), Kushner (1972) and the references contained therein. The classical needle variation approach of the deterministic control theory applies also in the stochastic case. However, different from the deterministic setting, for stochastic optimal control problems when this approach is used, the cost functional needs to be expanded up to the *second order*. This leads to *two adjoint equations* in the *first order necessary optimality conditions*. A stochastic maximum principle for this general case can be found in Peng (1990). Furthermore, to derive second order

⋆ The first author is partially supported by the Gaspard Monge Program for Optimisation(PGMO); the second author is partially supported by NSF of China under grants 11401404 and 11471231 and the fundamental research funds for the central universities under grants SWU114074 and XDJK2015C142; the third author is partially supported by NSF of China under grant 11231007, and by the Chang Jiang Scholars Program from Chinese Education Ministry.

2405-8963 © 2016, IFAC (International Federation of Automatic Control) Hosting by Elsevier Ltd. All rights reserved.
Peer review under responsibility of International Federation of Automatic Control.
10.1016/j.ifacol.2016.10.183

necessary optimality conditions, the cost functional needs to be expanded up to the forth order, leading, in turn, to four adjoint equations, see Zhang and Zhang (2015b). For this reason, unlike in the deterministic case, much stronger smoothness assumptions (with respect to the state variable x) on b, σ, f, g have to be imposed.

In this work we show that under less regularity assumptions and using the direct variational approach, the first (resp. second) order necessary conditions involving just one adjoint equation (resp. two adjoint equations) can be obtained in a way similar to the deterministic one.

To illustrate our approach let us assume first that U is convex. Then the usual convex variations allow to construct control perturbations and only one adjoint equation is needed to describe the first order necessary condition (see Bensoussan (1981)) and two adjoint equations are needed to establish the second order necessary condition (see Zhang and Zhang (2015a)). Such variations allow to avoid difficulties brought by the perturbations with respect to the measure. When the control region is nonconvex, the usual convex variations cannot be used, since there may exist a control $u \in \mathcal{U}_{ad}$ such that $\bar{u} + \varepsilon(u - \bar{u}) \notin \mathcal{U}_{ad}$ for some small $\varepsilon > 0$. Nevertheless, if the perturbation direction v is chosen so that for any $\varepsilon > 0$ one can find a $v^\varepsilon \in \mathcal{U}_{ad}$ converging to v (in a suitable sense) and such that $\bar{u}+\varepsilon v^\varepsilon \in \mathcal{U}_{ad}$, then the variational approach can be applied (we call this approach the *classical variational analysis approach*). Indeed, this method has been used extensively in optimization and optimal control theory in the deterministic setting. Using this approach, in Hoehener (2012) and Frankowska and Osmolovskii (2015), some second order integral type necessary conditions for deterministic optimal controls were established. Furthermore, it was shown in Frankowska and Tonon (2013) and Frankowska and Hoehener (2015) that these integral necessary conditions imply pointwise ones. In the stochastic setting, the first and second order integral type necessary conditions for stochastic optimal controls with convex control constraints were derived in Bonnans and Silva (2012).

In this paper, we shall use the classical variational analysis approach to establish the first and second order necessary optimality conditions for stochastic optimal controls in the general setting, that is, when the control region is allowed to be nonconvex and the diffusion term contains the control variable. Compared to the existing results obtained by the needle variations, cf. Peng (1990) and Zhang and Zhang (2015b), the main advantage of the classical variational analysis approach is due to less regularity requirements and to fewer adjoint equations needed to state these conditions.

The outline of the paper is as follows. In Section 2, we recall some notions and notations. In Section 3, we state the first order necessary conditions for stochastic optimal controls, while, in Section 4 we establish the corresponding second order necessary conditions.

2. PRELIMINARIES

Denote by $\langle\cdot,\cdot\rangle$ and $|\cdot|$ respectively the inner product and norm in $\mathbb{R}^n$ and by $C_b^\infty(\mathbb{R}^n;\mathbb{R}^m)$ the set of C^∞-smooth functions from $\mathbb{R}^n$ to $\mathbb{R}^m$ with bounded derivatives. Let $\mathbb{R}^{n\times m}$ be the space of all $n \times m$-real matrices. For any $A \in \mathbb{R}^{n\times m}$, denote by $A^\top$ its transpose. Also, write $\mathbf{S}^n := \{A \in \mathbb{R}^{n\times n} \mid A^\top = A\}$.

Let $\varphi : [0,T]\times\mathbb{R}^n\times U\times\Omega \to \mathbb{R}^d$ ($d \in \mathbb{N}$) be a given function. For a.e. $(t,\omega) \in [0,T]\times\Omega$, we denote by $\varphi_x(t,x,u,\omega)$, $\varphi_u(t,x,u,\omega)$ the first order partial derivatives of φ with respect to x and u at (t,x,u,ω), by $\varphi_{(x,u)^2}(t,x,u,\omega)$ the Hessian of φ with respect to (x,u) at (t,x,u,ω), and by $\varphi_{xx}(t,x,u,\omega)$, $\varphi_{xu}(t,x,u,\omega)$, $\varphi_{uu}(t,x,u,\omega)$ the second order partial derivatives of φ with respect to x and u at (t,x,u,ω).

For any $\alpha,\beta \in [1,+\infty)$ and $t \in [0,T]$, let $L^\beta_{\mathcal{F}_t}(\Omega;\mathbb{R}^n)$ be the space of $\mathbb{R}^n$-valued, $\mathcal{F}_t$ measurable random variables ξ such that $\mathbb{E}\,|\xi|^\beta < +\infty$; $L^\beta([0,T]\times\Omega;\mathbb{R}^n)$ be the space of $\mathbb{R}^n$-valued, $\mathcal{B}([0,T])\otimes\mathcal{F}$-measurable processes φ such that $\|\varphi\|_\beta := \big[\mathbb{E}\int_0^T|\varphi(t,\omega)|^\beta dt\big]^{\frac{1}{\beta}} < +\infty$; $L^\beta_{\mathbb{F}}(\Omega;L^\alpha(0,T;\mathbb{R}^n))$ be the space of $\mathbb{R}^n$-valued, $\mathcal{B}([0,T])\otimes\mathcal{F}$-measurable, $\mathbb{F}$-adapted processes φ such that $\|\varphi\|_{\alpha,\beta} := \big[\mathbb{E}\,\big(\int_0^T|\varphi(t,\omega)|^\alpha dt\big)^{\frac{\beta}{\alpha}}\big]^{\frac{1}{\beta}} < +\infty$; $L^\beta_{\mathbb{F}}(\Omega;C([0,T];\mathbb{R}^n))$ be the space of $\mathbb{R}^n$-valued, $\mathcal{B}([0,T])\otimes\mathcal{F}$-measurable, and $\mathbb{F}$-adapted continuous processes φ such that $\|\varphi\|_{\infty,\beta} := \big[\mathbb{E}\,\big(\sup_{t\in[0,T]}|\varphi(t,\omega)|^\beta\big)\big]^{\frac{1}{\beta}} < +\infty$; $L^\infty([0,T]\times\Omega;\mathbb{R}^n)$ be the space of $\mathbb{R}^n$-valued, $\mathcal{B}([0,T])\otimes\mathcal{F}$-measurable processes φ such that $\|\varphi\|_\infty := \mathrm{ess\ sup}_{(t,\omega)\in[0,T]\times\Omega}|\varphi(t,\omega)| < +\infty$ and $L^\beta(0,T;L^\beta_{\mathbb{F}}([0,T]\times\Omega;\mathbb{R}^n))$ be the space of $\mathbb{R}^n$-valued, $\mathcal{B}([0,T])\otimes\mathcal{B}([0,T])\otimes\mathcal{F}$ measurable functions φ such that $\|\varphi\|_\beta := \big[\mathbb{E}\int_0^T\int_0^T|\varphi(s,t,\omega)|^\beta ds dt\big]^{\frac{1}{\beta}} < +\infty$ and for any $t \in [0,T]$, the process $\varphi(\cdot,t,\cdot)$ is $\mathbb{F}$-adapted.

For a separable Banach space X with norm $\|\cdot\|_X$, $B(x,r)$ denotes the open ball centered at $x \in X$ of radius $r > 0$. For any set $K \subset X$, ∂K is its boundary. The distance between a point $x \in X$ and K is defined by $dist\,(x,K) := \inf_{y\in K}\|y - x\|_X$.

The adjacent cone $T^b_K(x)$ to K at $x \in K$ is given by

$$T^b_K(x) := \Big\{v \in X \;\Big|\; \lim_{\varepsilon\to 0^+}\frac{dist\,(x+\varepsilon v, K)}{\varepsilon} = 0\Big\}.$$

If in the above $\lim_{\varepsilon\to 0^+}$ is replaced by $\liminf_{\varepsilon\to 0^+}$, then we obtain a larger cone, the so called *contingent cone* $T_K(x)$ to K at x.

Definition 1. For any $x \in K$ and $v \in T^b_K(x)$, the second order adjacent subset to K at (x,v) is defined by

$$T^{b(2)}_K(x,v) := \Big\{h \in X \;\Big|\; \lim_{\varepsilon\to 0^+}\frac{dist\,(x+\varepsilon v+\varepsilon^2 h, K)}{\varepsilon^2} = 0\Big\}.$$

It is well known that $T^b_K(x)$ is a nonempty closed cone. However, when $x \in \partial K$ and $v \in T^b_K(x)$, the set $T^{b(2)}_K(x,v)$, in general, may not be a cone and may be empty.

Let $(\Xi,\mathscr{G})$ be a measurable space, and $F : \Xi \rightsquigarrow 2^X$ be a set-valued map. F is called measurable if for any $A \in \mathcal{B}(X)$, $F^{-1}(A) := \{\xi \in \Xi \mid F(\xi)\cap A \neq \emptyset\} \in \mathscr{G}$.

The following result is a special case of Theorem 8.5.1 from Aubin and Frankowska (1990).

Lemma 2. Suppose $(\Xi,\mathscr{G},\mu)$ is a complete σ-finite measure space, $p \geq 1$ and K is a closed set in X. Define