Global and Stochastic Analysis
Vol. 10 No. 3 (December, 2023)

ON OPTIMAL SOLUTIONS OF STOCHASTIC DIFFERENTIAL INCLUSIONS WITH BACKWARD MEAN DERIVATIVES

YURI GLIKLIKH, ALINA GRANKINA

ABSTRACT. For a stochastic differential inclusion given in terms of backward mean derivatives, we prove the existence of optimal solution minimizing a certain cost criterion.

Introduction

The notion of mean derivatives (forward, backward, symmetric and antisymmetric) was introduced by E. Nelson in [1, 2, 3]. In [4] (see also [5] where all preliminaries about mean derivatives are given) an additional mean derivative, called quadratic, was introduced so that form some Nelson's mean derivative and the quadratic one it became in principle possible to find a stochastic process having those derivatives.

In this paper we investigate stochastic differential inclusion given in terms of backward mean derivatives. The case of stochastic differential inclusion with forward mean derivatives was investigated in [6]. We prove the existence of optimal solution minimizing a certain cost criterion. The preliminaries about set-valued mappings can de found in [7].

Some remarks on notations. In this paper we deal with equations and inclusions in the linear space $\mathbb{R}^n$, for which we always use coordinate presentation of vectors and linear operators. Vectors in $\mathbb{R}^n$ are considered as columns. If X is such a vector, the transposed row vector is denoted by X^*. Linear operators from $\mathbb{R}^n$ to $\mathbb{R}^n$ are represented as $n \times n$ matrices, the symbol $*$ means transposition of a matrix (pass to the matrix of conjugate operator). The space of $n \times n$ matrices is denoted by $L(\mathbb{R}^n, \mathbb{R}^n)$.

By $\mathrm{S}(n)$ we denote the linear space of symmetric $n \times n$ matrices that is a subspace in $L(\mathbb{R}^n, \mathbb{R}^n)$. The symbol $\mathrm{S}_+(n)$ denotes the set of positive definite symmetric $n \times n$ matrices that is a convex open set in $\mathrm{S}(n)$. Its closure, i.e., the set of positive semi-definite symmetric $n \times n$ matrices, is denoted by $\bar{\mathrm{S}}_+(n)$.

Everywhere below for a set B in $\mathbb{R}^n$ or in $L(\mathbb{R}^n, \mathbb{R}^n)$ we use the norm introduced by usual formula $\|B\| = \sup\limits_{y \in B} \|y\|$.

For the sake of simplicity we consider equations, their solutions and other objects on a finite time interval $t \in [0, T]$.

Date: Date of Submission March 11, 2023 ; Date of Acceptance June 7, 2023 , Communicated by Igor V. Pavlov.

2010 *Mathematics Subject Classification.* Primary Primary 60H10; Secondary 60H30.

Key words and phrases. Mean derivatives; stochastic differential inclusions; optimal solution.

1. Mean derivatives

In this section we briefly describe preliminary facts about mean derivatives. See details in [1, 2, 3, 5].

Consider a stochastic process $\xi(t)$ in $\mathbb{R}^n$, $t \in [0,T]$, given on a certain probability space $(\Omega, \mathcal{F}, \mathsf{P})$ and such that $\xi(t)$ is an L_1 random element for all t. It is known that such a process determines 3 families of σ-subalgebras of the σ-algebra $\mathcal{F}$:

(i) "the past" $\mathcal{P}^\xi_t$ generated by preimages of Borel sets from $\mathbb{R}^n$ under all mappings $\xi(s) : \Omega \to \mathbb{R}^n$ for $0 \le s \le t$;

(ii) "the future" $\mathcal{F}^\xi_t$ generated by preimages of Borel sets from $\mathbb{R}^n$ under all mappings $\xi(s) : \Omega \to \mathbb{R}^n$ for $t \le s \le T$;

(iii) "the present" ("now") $\mathcal{N}^\xi_t$ generated by preimages of Borel sets from $\mathbb{R}^n$ under the mapping $\xi(t) : \Omega \to \mathbb{R}^n$.

All the above families we suppose to be complete, i.e., containing all sets of probability zero.

For the sake of convenience we denote by E^ξ_t the conditional expectation $E(\cdot|\mathcal{N}^\xi_t)$ with respect to the "present" $\mathcal{N}^\xi_t$ for $\xi(t)$.

Following [1, 2, 3], introduce the following notions of forward and backward mean derivatives.

Definition 1.1. (i) The forward mean derivative $D\xi(t)$ of $\xi(t)$ at the time instant t is an L_1 random element of the form

$$D\xi(t) = \lim_{\triangle t \to +0} E^\xi_t\Big(\frac{\xi(t+\triangle t) - \xi(t)}{\triangle t}\Big), \tag{1.1}$$

where the limit is supposed to exist in $L_1(\Omega, \mathcal{F}, \mathsf{P})$ and $\triangle t \to +0$ means that $\triangle t$ tends to 0 and $\triangle t > 0$.

(ii) The backward mean derivative $D_*\xi(t)$ of $\xi(t)$ at t is the L_1-random element

$$D_*\xi(t) = \lim_{\Delta t \to +0} E^\xi_t\Big(\frac{\xi(t) - \xi(t-\Delta t)}{\Delta t}\Big) \tag{1.2}$$

where (as well as in (i)) the limit is assumed to exist in $L^1(\Omega, \mathcal{F}, \mathsf{P})$ and $\Delta t \to +0$ means that $\Delta t \to 0$ and $\Delta t > 0$.

Remark 1.2. If $\xi(t)$ is a Markov process then evidently E^ξ_t can be replaced by $E(\cdot|\mathcal{P}^\xi_t)$ in (1.1) and by $E(\cdot|\mathcal{F}^\xi_t)$ in (1.2). In initial Nelson's works there were two versions of definition of mean derivatives: as in our Definition 1.1 and with conditional expectations with respect to "past" and "future" as above that coincide for Markov processes. We shall not suppose $\xi(t)$ to be a Markov process and give the definition with conditional expectation with respect to "present" taking into account the physical principle of locality: the derivative should be determined by the present state of the system, not by its past or future.

Following [4] (see also [5]) we introduce the differential operator D_2 that differentiates an L_1 random process $\xi(t)$, $t \in [0,T]$ according to the rule

$$D_2\xi(t) = \lim_{\triangle t \to +0} E^\xi_t\Big(\frac{(\xi(t+\triangle t) - \xi(t))(\xi(t+\triangle t) - \xi(t))^*}{\triangle t}\Big), \tag{1.3}$$

where $(\xi(t+\triangle t)-\xi(t))$ is considered as a column vector (vector in $\mathbb{R}^n$), $(\xi(t+\triangle t)-\xi(t))^*$ is a row vector (transposed, or conjugate vector) and the limit is supposed to exists in $L_1(\Omega,\mathcal{F},\mathsf{P})$. We emphasize that the matrix product of a column on the left and a row on the right is a matrix. It is shown that $D_2\xi(t)$ takes values in $\bar{S}_+(n)$, the set of symmetric semi-positive definite matrices. We call D_2 the quadratic mean derivative.

Remark 1.3. From the properties of conditional expectation (see, e.g., [8]) it follows that there exist Borel mappings $a(t,x)$, $a_*(t,x)$ and $\alpha(t,x)$ from $R\times\mathbb{R}^n$ to $\mathbb{R}^n$ and to $\bar{\mathrm{S}}_+$, respectively, such that $D\xi(t)=a(t,\xi(t))$, $D_*\xi(t)=a_*(t,\xi(t))$ and $D_2\xi(t)=\alpha(t,\xi(t))$. Following [8] we call $a(t,x)$, $a_*(t,x)$ and $\alpha(t,x)$ the regressions.

Let Borel measurable mappings $a(t,x)$ and $\alpha(t,x)$ from $[0,T]\times\mathbb{R}^n$ to $\mathbb{R}^n$ and to $\bar{\mathrm{S}}_+(n)$, respectively, be given. We call the system of the form

$$\begin{cases} D\xi(t)=a(t,\xi(t)),\\ D_2\xi(t)=\alpha(t,\xi(t)),\end{cases} \tag{1.4}$$

a first order differential equation with forward mean derivatives.

Definition 1.4. We say that (1.4) has a solution on $[0,T]$ with initial condition $\xi(0)=x_0$, if there exist a probability space $(\Omega,\mathcal{F},\mathsf{P})$ and a process $\xi(t)$ given on $(\Omega,\mathcal{F},\mathsf{P})$ and taking values in $\mathbb{R}^n$ such that P-a.s. and for almost all t (1.4) is satisfied.

Several existence of solution theorems for (1.4) can be found in [4].

Definition 1.5. The smooth function $\varphi: X\to\mathbb{R}$ sending the topological space X to $\mathbb{R}$ is called proper if the preimage of every relatively compact set in $\mathbb{R}$ is relatively compact in X.

Denote by $\mathcal{L}$ the generator of Markov process generated by equation (1.4).

Theorem 1.6. *Let on $\mathbb{R}^n$ there exist a smooth proper positive function $\varphi:\mathbb{R}^n\to\mathbb{R}$ such that $\mathcal{L}\varphi<C$ for all $t\in[0,+\infty)$ and $x\in\mathbb{R}^n$ where $C>0$ is a certain real constant. Then the flow generated by equation (1.4) is complete, i.e. all solutions of (1.4) with deterministic initial values exist for $t\in[0,+\infty)$.*

Theorem 1.6 is a reformulation of [9, Theorem IX. 6A].

2. Auxiliary facts about inclusions with forward mean derivatives and upper semi-continuous right-hand sides

Let $\mathbf{a}(t,x)$ and $\boldsymbol{\alpha}(t,x)$ be set-valued mappings from $[0,T]\times\mathbb{R}^n$ to $\mathbb{R}^n$ and to $\bar{\mathrm{S}}_+(n)$, respectively. The system of the form

$$\begin{cases} D\xi(t)\in\mathbf{a}(t,\xi(t)),\\ D_2\xi(t)\in\boldsymbol{\alpha}(t,\xi(t)).\end{cases} \tag{2.1}$$

is called a first order differential inclusion with forward mean derivatives.

Definition 2.1. We say that (2.1) has a solution on $[0,T]$ with initial condition $\xi(0)=x_0$, if there exist a probability space $(\Omega,\mathcal{F},\mathsf{P})$ and a process $\xi(t)$ given on $(\Omega,\mathcal{F},\mathsf{P})$ and taking values in $\mathbb{R}^n$ such that P-a.s. and for almost all t (2.1) is satisfied.

Note that for simplicity here we consider only deterministic initial conditions, i.e., ξ_0 in Definition 2.1 is a point in $\mathbb{R}^n$.

Recall that for a mapping $F : X \to Y$ of a metric space X to a metric space Y its graph is the set of pairs $\{(x, F(x)) \mid x \in X\}$ in $X \times Y$. Note that for a set-valued F the value $F(x)$ is a set in Y.

For considering upper semicontinuous mean forward differential inclusions we need to recall the following

Definition 2.2. Let X and Y be metric spaces. For given $\varepsilon > 0$ a continuous single-valued mapping $f_\varepsilon : X \to Y$ is called an ε-approximation of the set-valued mapping $F : X \to Y$, if the graph of f belongs to ε-neighbourhood of the graph of F.

It is known (see, e.g., [7]), that for upper semicontinuous set-valued mappings with convex closed images in normed linear spaces the ε-approximations exist for each $\varepsilon > 0$.

Denote by Ω the Banach space $C^0([0,T], \mathbb{R}^n)$ of continuous curves in $\mathbb{R}^n$ given on $[0,T]$, with usual uniform norm. Introduce in Ω the σ-algebra $\mathcal{F}$ generated by cylinder sets. Everywhere below we use this notation. Recall that $\mathcal{F}$ is the Borel σ-algebra in Ω. Note that the elementary event in Ω is a curve that we denote by $x(\cdot)$. Its value at $t \in [0,T]$ is denoted by $x(t)$.

It is a well-known fact that every stochastic process η with continuous sample paths in $\mathbb{R}^n$, given on a certain probability space $(\tilde{\Omega}, \tilde{\mathcal{F}}, \mathsf{P})$ for $t \in [0,T]$, is a measurable mapping from $(\tilde{\Omega}, \tilde{\mathcal{F}})$ to $(\Omega, \mathcal{F})$. Thus it determines a measure μ_η on $(\Omega, \mathcal{F})$ by the standard formula $\mu_\eta(A) = \mathsf{P}(\eta^{-1}(A))$ for every $A \in \mathcal{F}$.

There is a standard process $c(t, x(\cdot))$ in $\mathbb{R}^n$ given on $(\Omega, \mathcal{F})$. It is the so-called "coordinate process" defined by the formula $c(t, x(\cdot)) = x(t)$. The coordinate process on the probability space $(\Omega, \mathcal{F}, \mu_\eta)$ is the standard description of the process $\eta(t)$ on this probability space. See details, e.g., in [10, 5].

We shall look for solutions of (2.1) with continuous sample paths and mainly the solution will be described as a coordinate process on Ω where the corresponding measure will be constructed.

Definition 2.3. The perfect solution of (2.1) is a stochastic process with continuous sample paths such that it is a solution in the sense of Definition 2.1 and the measure corresponding to it on the space of continuous curves, is a weak limit of measures generated by solutions of a sequence of diffusion-type Itô equations with continuous coefficients.

Remark 2.4. Note that perfect solutions, approximated by solutions of diffusion type equation, naturally arise in applications. But there is an open question whether any solution is perfect or non-perfect solutions also may exist.

Lemma 2.5. *Let $\alpha(t,x)$ be a jointly continuous (measurable, smooth) mapping from $[0,T] \times \mathbb{R}^n$ to $\mathrm{S}_+(n)$. Then there exists a jointly continuous (measurable, smooth, respectively) mapping $A(t,x)$ from $[0,T] \times \mathbb{R}^n$ to $L(\mathbb{R}^n, \mathbb{R}^n)$ such that for all $t \in R$, $x \in \mathbb{R}^n$ the equality $A(t,x)A^*(t,x) = \alpha(t,x)$ holds.*

The proof is available in [4, Lemma 2.2].

Theorem 2.6 ([6])**.** *Specify an arbitrary initial value $\xi_0 \in \mathbb{R}^n$. Let $\mathbf{a}(t,x)$ be an upper semicontinuous set-valued mapping with closed convex images from $[0,T] \times \mathbb{R}^n$ to $\mathbb{R}^n$ and let it satisfy the estimate*

$$\|\mathbf{a}(t,x)\|^2 < K(1+\|x\|^2) \tag{2.2}$$

for some $K > 0$.

Let $\boldsymbol{\alpha}(t,x)$ be an upper semicontinuous set-valued mapping with closed convex images from $[0,T] \times \mathbb{R}^n$ to $\bar{S}_+(n)$ such that for each $\alpha(t,x) \in \boldsymbol{\alpha}(t,x)$ the estimate

$$|tr\alpha(t,x)| < K(1+\|x\|^2) \tag{2.3}$$

takes place for some $K > 0$.

Then for every sequence $\varepsilon_i \to 0$, $\varepsilon_i > 0$, each pair of sequence $a_i(t,x)$ and $\alpha_i(t,x)$ of ε_i-approximations of $\mathbf{a}(t,x)$ and $\boldsymbol{\alpha}(t,x)$, respectively, generates a perfect solution of (2.1) *with initial condition ξ_0.*

3. Differential inclusions with backward mean derivatives

The system

$$\begin{cases} D_*\xi(t) = a(t,\xi(t)) \\ D_2\xi(t) = \alpha(t,\xi(t)) \end{cases} \tag{3.1}$$

is called a first order differential equation with backward mean derivatives.

Notice that we do not introduce the notion of backward analog of operator D_2 since, applying the properties of Itô integral, one can easily prove that for a diffusion process $\xi(t)$ the result of application of that analog coincides with $D_2\xi(t)$ (for the case of diffusion processes this follows from the results of [2, 3]).

Let $\mathbf{a}(t,x)$ and $\boldsymbol{\alpha}(t,x)$ be set-valued mappings from $[0,T] \times \mathbb{R}^n$ to $\mathbb{R}^n$ and to $\bar{S}_+(n)$, respectively. The system of the form

$$\begin{cases} D_*\xi(t) \in \mathbf{a}(t,\xi(t)), \\ D_2\xi(t) \in \boldsymbol{\alpha}(t,\xi(t)). \end{cases} \tag{3.2}$$

is called a first order differential inclusion with backward mean derivatives.

Definition 3.1. We say that (3.2) has a solution on $[0,T]$ with "inverse" Cauchy condition $\xi(T) = \xi_0$, if there exist a probability space $(\Omega, \mathcal{F}, \mathsf{P})$ and a process $\xi(t)$ given on $(\Omega, \mathcal{F}, \mathsf{P})$ and taking values in $\mathbb{R}^n$ such that $\xi(T) = \xi_0$ and P-a.s. and for almost all t inclusion (3.2) is satisfied.

Consider a solution $\eta(t)$, given on $t \in [0,T]$, with initial condition $\eta(0) = \xi_0$ of the following differential inclusion with forward mean derivatives

$$\begin{cases} D\eta(t) \in -\mathbf{a}(T-t,\eta(t)), \\ D_2\eta(t) \in \boldsymbol{\alpha}(T-t,\eta(t)). \end{cases} \tag{3.3}$$

Theorem 3.2. *The process $\xi(t) = \xi_0 - \eta(T) + \eta(T-t)$ is a solution of (3.2) with condition $\xi(T) = \xi_0$ where $\eta(t)$ is a solution of (3.3) with initial condition $\eta(0) = \xi_0$.*

Indeed, $D_*\xi(t) = -D\eta(T-t) \in \boldsymbol{a}(t,\eta(T-t)) = \boldsymbol{a}(t,\xi(t))$. For $D_2\xi(t)$ the arguments are analogous.

Theorem 3.3. *Specify an arbitrary final value* $\xi_0 \in \mathbb{R}^n$*. Let the set-valued mappings* $\mathbf{a}(t,x)$ *and* $\boldsymbol{\alpha}(t,x)$ *satisfy the conditions of Theorem 2.6. Then for every sequence* $\varepsilon_i \to 0$*,* $\varepsilon_i > 0$*, each pair of sequence* $a_i(t,x)$ *and* $\alpha_i(t,x)$ *of* ε_i*-approximations of* $\mathbf{a}(t,x)$ *and* $\boldsymbol{\alpha}(t,x)$*, respectively, generates a perfect solution of* (3.2) *with inverse initial condition* ξ_0*.*

Indeed, under the hypothesis of Theorem 3.3 inclusion (3.3) satisfies the condition of Theorem 2.6. Thus the assertion of Theorem 3.3 follows from Theorm 3.2.

Remark 3.4. Note that all sequences of ε-approximations for all sequences of $\varepsilon_i \to 0$, used in the proof of Theorem 2.6, satisfy (2.2) and (2.3) with the same K so that by corollary in Section III.2 [10] the set of measures $\{\mu_i\}$ (corresponding to all sequences and all i is weakly compact.

Let f be a continuous bounded real-valued function on $\mathbb{R} \times \mathbb{R}^n$. For solutions of (2.1) consider the cost criterion in the form

$$J(\xi(\cdot)) = E\int_0^T f(t,\xi(t))dt \tag{3.4}$$

We are looking for solutions, for which the value of the criterion is minimal.

Theorem 3.5. *Among the perfect solutions of* (3.2) *constructed in the proof of Theorem 3.3, there is a solution* $\xi(t)$ *on which the value of* J *is minimal.*

Proof. Since all the measures on $(\Omega,\mathcal{F})$, constructed in the proof of Theorem 3.3 for perfect solutions of (3.2), are probabilistic and the function f in (3.4) is bounded, the set of values of J on those solutions is bounded. If that set of values has a minimum, then the corresponding measure μ is the one we are looking for: the coordinate process on the space $(\Omega,\mathcal{F},\mu)$ is an optimal solution.

Suppose that the above-mentioned set of values has no minimum, but then it has a greatest lower bound $\aleph$ that is a limit point in that set. Let μ_i^* be a sequence of measures such that for the corresponding solutions $\xi_i^*(t)$ the values $J(\xi_i^*(t))$ converge to $\aleph$. Every μ_i^* is a weak limit of a sequence of measures μ_{ij} corresponding to some sequence of ε_j-approximations as $j \to \infty$. One can easily see that it is possible to select from the sequence a subsequence (for simplicity we denote it by the same symbol μ_{ij}) such that for the corresponding solutions $\xi_{ij}(t)$ and for all i we obtain the uniform convergence of $J(\xi_{ij}(\cdot))$ to $J(\xi_i^*(\cdot))$ as $j \to \infty$. Then $J(\xi_{ii}(\cdot)) \to \aleph$ as $i \to \infty$. Since the set of all measures corresponding to all approximations, is weakly compact (see above), we can select from μ_{ii} a subsequence (denote it by the same symbol μ_{ii}) that weakly converges to a certain measure μ^*. By the construction, for the coordinate process $\xi^*(t)$ on $(\Omega,\mathcal{F},\mu^*)$ we get $J(\xi^*(\cdot)) = \aleph$, i.e., the value is minimal. Since μ^* is a limit of μ_{ii}, $\xi^*(t)$ is a perfect solution of (2.1) that we are looking for. □

References

[1] Nelson, E.: Derivation of the Schrödinger equation from Newtonian mechanics, *Phys. Rev.*, **150**, no.4 (1966), 1079-1085.

[2] Nelson, E.: *Dynamical theory of Brownian motion*, Princeton Univ. Press, Princeton, NJ, 1967.

[3] Nelson, E.: *Quantum fluctuations*, Princeton Univ. Press, Princeton, NJ, 1985.
[4] Azarina, S.V, Gliklikh, Yu.E.: Differential inclusions with mean derivatives, *Dynam. Systems Appl.*, **16**, no. 1 (2007), 4971.
[5] Gliklikh, Yu.E.: *Global and stochastic analysis with applications to mathematical physics*, Springer-Verlag, London, 2011.
[6] Gliklikh, Yu.E., Zheltikova, O.O.: On existence of optimal solutions for stochastic differential inclusions with mean derivatives, *Applicable Analysis*, **93**, no. 1 (2014), 35 - 46.
[7] Obukhovskii, V., Gelman, D.: *Multivalued Maps and Differential Inclusions. Elements of Theory and Applications*, World Scientific Publishing Co., Hackensack, NJ, 2020)
[8] Parthasarathy, K.R.: *Introduction to Probability and Measure*, Springer-Verlag, New York, 1978.
[9] Elworthy, K.D.: *Stochastic differential equations on manifolds*, (Lect. Notes of London Math. Soc. **70**), Cambridge University Press, Cambridge, NJ, 1982.
[10] Gihman, I.I., Skorohod, A.V.: *Theory of Stochastic Processes. Vol.3*, Springer-Verlag, New York, 1979.

Yuri Gliklikh: Voronezh State University, Voronezh, Russia
E-mail address: yeg@math.vsu.ru

Alina Grankina: Voronezh State University, Voronezh, Russia
E-mail address: grankina@math.vsu.ru