P. Del Moral, M. Doisy, Maslov idempotent probability calculus, I, *Teor. Veroyatnost. i Primenen.*, 1998, Volume 43, Issue 4, 735–751

DOI: https://doi.org/10.4213/tvp2029

ТЕОРИЯ ВЕРОЯТНОСТЕЙ И ЕЕ ПРИМЕНЕНИЯ
Том 43 Выпуск 4
1998

DEL MORAL P.[*], *DOISY M.*[**]

MASLOV IDEMPOTENT PROBABILITY CALCULUS, I[1)]

Идемпотентные меры Маслова позволяют вывести теорию оптимизации на такой же уровень общности, на котором находятся теория вероятностей и теория случайных процессов. Одна из целей настоящей работы — изложить основные концепции этой $(\max, +)$-версии теории вероятностей. Используя эту структуру, мы увидим, что принцип оптимальности Беллмана является идемпотентной версией классического принципа причинности Маркова. Во второй части исследования будут обсуждаться приложения к задачам оптимального управления, уравнениям Гамильтона–Якоби и математической морфологии.

Ключевые слова и фразы: теория оптимизации, идемпотентная мера Маслова, идемпотентное исчисление вероятностей, процессы Беллмана–Маркова.

Introduction

Maslov Idempotent Measure Theory has recently emerged as a new branch of mathematical analysis for studying deterministic control problems and first-order nonlinear partial differential equations such as Hamilton–Jacobi equations with discontinuous initial data and low-lying eigenfunctions of Schrödinger operator.

During the decade its application area has grown establishing unexpected connections with number of other fields. Litvinov and Maslov's presentation of the so-called Correspondence Principle for Idempotent Calculus [14] includes a very precise and useful summary on this subject. At the beginning of the 1990's several different approaches were suggested to construct an optimization theory at the same level of generality as probability and stochastic processes theory (see for instance [1], [3] or [6] and references therein).

*LSP-UMR C55830, CNRS, Bat. 1R1, Univ. Paul Sabatier, 31062 Toulouse, France; e-mail: delmoral@cict.fr

**ENSEEIHT, 2 Rue Camichel, 31071 Toulouse, France, and LSP-UMR C55830, CNRS; e-mail: doisy@len7.enseeiht.fr

1)This work was supported by the INTAS–RFBR grant 95-91.

The key idea consists in replacing in the structural axioms of Probability Theory the role of the classical semiring of positive real numbers $(\mathbf{R}_+, +, \cdot, 0, 1)$ by the idempotent semiring $(\mathbf{R}_- \cup \{-\infty\}, \oplus, \odot, \mathbb{0}, \mathbb{1}) \stackrel{\text{def}}{=} (\mathbf{R}_- \cup \{-\infty\}, \max, +, -\infty, 0)$, where $\mathbf{R}_-$ is the set of negative real numbers.

This new approach has the attribute of taking full advantage of the well-developed theory of Idempotent Analysis presented in [12], [13], [15]–[17]. The main purpose of this article is to discuss some topological aspects of this $(\max, +)$-version of Probability Theory and to present links between this theory and applications.

The paper is organized as follows. Before to present a formulation of Maslov optimization theory it is important to understand some of the basic principles of Idempotent Analysis. In the first section we follow fairly closely [12]. We introduce some notation, definitions and properties concerning idempotent measures on general Hausdorff topological spaces and integration theory on locally compact and separable metric spaces. In the second section we give a brief survey of the optimization theory developed in [6]. We also introduce some additional topological structure in the definition of an optimization or decision variable so that to present this material in an absolutely rigorous way. As we shall see this natural framework is dictated by the classical large deviation calculus as presented in [7] and [8]. The concepts of performance measure, optimization variables, independence and conditioning are the basic notions of our model. Using this framework we will see that the Bellman optimization principle is the $(\max, +)$-version of the classical Markov causality principle. The optimization processes satisfying this principle will be called Bellman–Maslov processes in reference to optimal control theory.

Some applications of Maslov optimization theory will be discussed in the second part of this study. We shall see that this new branch of functional analysis gives a natural probabilistic framework for formulating and studying Hamilton–Jacobi equations. In this connection we propose a law of large numbers and a central limit theorem which give some insight into the limit behavior of a class of Hamilton–Jacobi equations. Applications to mathematical morphology are also discussed.

1. Idempotent measures and integration

The elements of idempotent analysis (analysis of functions with values in a general idempotent semiring) are developed in [9], [11]–[17]. Let $\mathbf{R}_{\max} \stackrel{\text{def}}{=} [-\infty, +\infty[$ be the semiring of reals numbers endowed with the commutative semigroup laws $\oplus$, $\odot$, the neutrals elements $\mathbb{0}$, $\mathbb{1}$ and the exponential metric ρ such that

$$\mathbb{0} = -\infty, \quad a \oplus b = \sup(a, b), \quad \rho(a, b) = |e^a - e^b|,$$
$$\mathbb{1} = 0, \quad a \odot b = a + b, \quad \frac{a}{b}{}_{\circ} = a - b \quad (b \neq \mathbb{0}).$$

We also denote by $\rho(a,b) = \sup_{1\leqslant i\leqslant n} |e^{a_i} - e^{b_i}|$ the exponential metric on $\mathbf{R}^n_{\max}$.

Let Ω be a Hausdorff topological space equipped with a topology τ (i.e., $\varnothing, \Omega \in \tau$, any unions of sets of τ belongs to τ, and any finite intersections of elements of τ belongs to τ).

Let $(S, \cup, \cap)$ be a semiring of sets in Ω.

D e f i n i t i o n 1. A Maslov idempotent measure μ on $(S, \cup, \cap)$ is a mapping from S into $\mathbf{R}_{\max} \cup \{+\infty\}$ such that 1) $\mu(\varnothing) = \mathbb{0}$, 2) $\mu(A \cup B) = \mu(A) \oplus \mu(B)$ for all sets.

It is said to be bounded whenever $\mu(\Omega) < +\infty$.

As pointed out in [15] these measures fail to be continuous on the empty set and they have a nonunique extension to the σ-field spanned by S (i.e., the smallest σ-field that contains S) but there is a unique maximal extension to $\sigma(S)$. If S is taken as the set of all closed (or open) subsets of Ω the resulting σ-field is the Borel σ-field of Ω denoted by $\mathcal{B}(\Omega)$.

E x a m p l e 1. Assume that $f\colon \Omega \to \mathbf{R}_{\max}$ is an upper semicontinuous (u.s.c.) function with compact level sets, that is

$$\Omega(\alpha) = \{\omega \in \Omega\colon\ f(\omega) \geqslant \alpha\} \quad \text{are compact subsets of } \Omega.$$

Then the mapping

$$\mu\colon\ A \in \mathbf{B}(\Omega) \mapsto \mu(A) = \sup_{\omega\in A} f(\omega) \tag{1}$$

is a Maslov idempotent measure on $\mathbf{B}(\Omega)$. Let $(\Omega', \mathbf{B}(\Omega'))$ be an auxiliary Hausdorff topological space and, let $g\colon \Omega \to \Omega'$ be a continuous function. As we shall see in the forthcoming development the covariant image $m = \mu \circ g^{-1}$ of μ under the transformation g is the Maslov idempotent measure given by

$$m(A) = \mu\big(g^{-1}(A)\big) = \sup_{\omega'\in A} h(\omega') \quad \forall A \in \mathbf{B}(E),$$

where $h\colon \Omega' \longrightarrow \mathbf{R}_{\max}$ is the u.s.c. function with compact level sets given by $h(\omega') = \sup\{f(\omega) \mid \omega \in \Omega\colon\ g(\omega) = \omega'\}$.

Let us specialize to the situation described in [12], [13] and [15], [16]. Namely, a more restrictive setting is to assume that Ω is a locally compact and separable space. In this situation Maslov has presented in [15] a new idempotent measure theory. Here some basic and well-known results are briefly presented. Their proof can be found for instance in [12], [13], [15], [16].

Let us introduce some additional notation. We denote by $\mathscr{C}_{\mathbb{0}}(\Omega)$ the semimodule of continuous $\mathbf{R}_{\max}$-valued functions on Ω tending to $\mathbb{0}$ at infinity, and $\mathscr{C}_{\mathrm{K}}(\Omega)$ is the sub-semimodule of $\mathscr{C}_{\mathbb{0}}(\Omega)$ such that the support of any $f \in \mathscr{C}_{\mathrm{K}}(\Omega)$ is compact (i.e., $\operatorname{Supp} f \stackrel{\text{def}}{=} \overline{\{\omega \in \Omega\colon\ f(\omega) > \mathbb{0}\}}$ is compact).

We denote by $\mathscr{C}'_{\mathbb{0}}(\Omega)$ ($\mathscr{C}'_{\mathrm{K}}(\Omega)$) the semimodule of continuous and $\mathbf{R}_{\max}$-valued $(\oplus, \odot)$-linear forms on $\mathscr{C}_{\mathbb{0}}(\Omega)$ (respectively on $\mathscr{C}_{\mathrm{K}}(\Omega)$).

By $\mathrm{IM}(\Omega)$ we denote the semimodule of upper semicontinuous functions on Ω. If (Ω, d_Ω) is a metric space, we denote by $\mathrm{Lip}(\Omega)$ the semimodule of upper bounded and Lipschitz continuous functions $f\colon \Omega \to \mathbf{R}_{\max}$, i.e., upper bounded functions f such that

$$\rho\big(f(x), f(y)\big) \leqslant C\, d_\Omega(x, y) \quad \forall\, x, y \in \Omega.$$

The more general fact about the semimodules $\mathscr{C}'_{\mathrm{K}}(\Omega)$ and $\mathscr{C}'_{\mathbb{0}}(\Omega)$ which will be helpful to have at our disposal is the following

Theorem 1 [Kolokoltsov–Maslov [12]. *The duality relation*

$$\mathscr{C}_{\mathrm{K}}(\Omega) \times \mathrm{IM}(\Omega) \longrightarrow \mathbf{R}_{\max}$$

$$(\varphi, f) \longmapsto \int_\Omega^\oplus \varphi(\omega) \odot f(\omega) \otimes d\omega \stackrel{\text{def}}{=} \sup_{\omega \in \Omega} \varphi(\omega) + f(\omega)$$

determines a representation of $\mathscr{C}'_{\mathrm{K}}(\Omega)$ *as* $\mathrm{IM}(\Omega)$.

To functionals $m \in \mathscr{C}'_{\mathbb{0}}(\Omega)$ *there correspond bounded functions* f *such that*

$$m(\varphi) = \int_\Omega^\oplus \varphi(\omega) \odot f(\omega) \otimes d\omega \stackrel{\text{def}}{=} m_f(\varphi).$$

Moreover for any $m_{f_1}, m_{f_2} \in \mathscr{C}'_{\mathbb{0}}(\Omega)$ *we have* $m_{f_1} = m_{f_2} \Longleftrightarrow f_1^\star = f_2^\star$ *where* $f^\star$ *is the upper semicontinuous envelope of* f, *i.e.*,

$$f^\star(x) \stackrel{\text{def}}{=} \sup\big\{\varphi(x) \mid \varphi \in \mathscr{C}_{\mathbb{0}}(\Omega)\colon f \leqslant \varphi\big\}.$$

For any $m_1, m_2 \in \mathscr{C}'_{\mathbb{0}}(\Omega)$ *we have*

$$\forall \varphi \in \mathrm{Lip}\,(\Omega) \quad m_1(\varphi) = m_2(\varphi) \Longrightarrow m_1 = m_2.$$

Many constructions and results can be borrowed to idempotent analysis from usual analysis and functional analysis, the interested reader is recommended to consult [14] and references therein.

2. Idempotent probability calculus

A first study of the parallelism between Probability and Stochastic Processes on the one hand and Performance Measures and Optimization Processes on the other hand can be founded in [1] and [6]. But to present this material in an absolutely rigorous way some additional topological structure has to be added. These notions are strongly related to Large Deviations Calculus.

As far as applications are concerned, this branch of idempotent analysis has been motivated by control theoretic applications and nonlinear estimation problems.

The aim of this section is to present a brief survey of this duality and to introduce the basic concepts and the mathematical tools upon which the theory dwells.

2.1. Performance spaces and optimization variables. As in probability theory, the basic notion is the (max, +)-version of the definition of a probability space.

D e f i n i t i o n 2. A performance space or an optimization basis is a triplet $(\Omega, \mathscr{B}(\Omega), \mathbf{P})$ where

1) $(\Omega, \mathscr{B}(\Omega))$ is a Hausdorff topological space equipped with the Borel σ-field $\mathscr{B}(\Omega)$;

2) $\mathbf{P}$ is a Maslov idempotent measure on $(\Omega, \mathscr{B}(\Omega))$ such that $\mathbf{P}(\Omega) = 1$ and

$$\forall A \in \mathscr{B}(\Omega) \quad \mathbf{P}(A) = \sup_{\omega \in A} \mathbf{p}(\omega),$$

where $\mathbf{p}\colon \Omega \to \mathbf{R}_{\max}$ is an upper semicontinuous function such that for any $\alpha \in \mathbf{R}$ the level sets $\Omega(\alpha, \mathbf{P}) = \{\omega \in \Omega\colon \mathbf{p}(\omega) \geqslant \alpha\}$ are compact subsets of Ω.

The measure $\mathbf{P}$ is called a performance or cost measure on $(\Omega, \mathscr{B}(\Omega))$ and the upper semicontinuous function $\mathbf{p}$ is called the performance density of the measure $\mathbf{P}$.

The density $\mathbf{p}$ of a performance measure $\mathbf{P}$ on $(\Omega, \mathscr{B}(\Omega))$ is similar to the definition of a good rate function in classical large deviation calculus. Let us recall some elementary properties of such functions.

Lemma 1. *Let $(\Omega, \mathscr{B}(E), \mathbf{P})$ be a performance space.*

1) *For any closed subset $A \subset \Omega$ there exists some $\omega^\star \in A$ such that $\mathbf{p}(\omega^\star) = \sup_{\omega \in A} \mathbf{p}(\omega)$.*

2) *If $\{F_\varepsilon\colon \varepsilon > 0\}$ is a nested family of closed sets (i.e., $F_{\varepsilon_1} \subset F_{\varepsilon_2}$ for any $\varepsilon_1 < \varepsilon_2$) then $\mathbf{P}(\bigcap_{\varepsilon > 0} F_\varepsilon) = \lim_{\varepsilon \to 0} \mathbf{P}(F_\varepsilon)$.*

The proof of this lemma is standard, the interested reader is referred to Lemma 2.1.2 in [8] or Lemma 4.1.6 in [7].

The second basic notion is the (max, +)-version of a random variable.

D e f i n i t i o n 3. Let $(\Omega, \mathscr{B}(\Omega), \mathbf{P})$ be a performance space and let $(E, \mathscr{B}(E))$ be a Hausdorff topological space. A measurable mapping $X\colon \Omega \to E$ is an E-valued optimization variable provided its restrictions to the compact level sets $\Omega(\alpha, \mathbf{P})$, $\alpha \in \mathbf{R}_-$, are continuous.

Let us give some comments on the above notions.

An optimization problem is described by a Hausdorff topological space $(\Omega, \mathscr{B}(\Omega))$ where Ω is the set of all possible control actions or policies and $\mathscr{B}(\Omega)$ is its natural Borel σ-field. $X(\omega) \in E$ is the output of a given controlled system $X\colon \Omega \to E$ under the control $\omega \in \Omega$.

The performance measure $\mathbf{P}$ given on $(\Omega, \mathscr{B}(\Omega))$ describes the quality of each control action.

As in Probability Theory the minimal assumption on X is to be measurable. In this situation we can define the covariant image $\mathbf{P}^X = \mathbf{P} \circ X^{-1}$ of $\mathbf{P}$ under the transformation X. Namely, $\mathbf{P}^X(A) = \mathbf{P}(X^{-1}(A))\ \forall A \in \mathscr{B}(E)$. Note that

$$\mathbf{P}^X(A) = \sup_{x \in A} \mathbf{p}^X(x) \quad \text{with } \mathbf{p}^X(x) \stackrel{\text{def}}{=} \sup\{\mathbf{p}(\omega) \mid \omega\colon X(\omega) = x\}.$$

Unfortunately the function $\mathbf{p}^X\colon E \to \mathbf{R}_{\max}$ is not necessarily u.s.c. and the compact level set condition is not always true. The first claim follows from the fact that a pointwise supremum of an arbitrary collection of u.s.c. functions is not necessarily u.s.c.

The continuity assumption in the optimization variable definition is a natural condition on X which ensures that $\mathbf{p}^X$ is a performance.

Our claim that this framework is a natural framework for formulating optimization problems will be amply justified by the results which will follow.

Proposition 1. *Let $(\Omega, \mathcal{B}(\Omega), \mathbf{P})$ be a performance space and let $(E, \mathcal{B}(E))$ be a Hausdorff topological space. Let $X\colon \Omega \to E$ be an optimization variable. Then*

1) $(E, \mathcal{B}(E), \mathbf{P}^X)$ is a performance space and the density of $\mathbf{P}^X$ is the upper semicontinuous function $\mathbf{p}^X\colon E \to \mathbf{R}_{\max}$ given by

$$\mathbf{p}^X(x) = \sup\{\mathbf{p}(\omega) \mid \omega \in \Omega\colon X(\omega) = x\} \quad \forall x \in E; \tag{2}$$

2) If E is a metric space, X is a continuous optimization variable and the function $\mathbf{p}^X$ is bounded, then there exists a nonincreasing sequence of Lipschitz continuous functions $\mathbf{p}_n^X\colon E \to \mathbf{R}_{\max}$ such that

$$\mathbf{p}^X(x) = \inf_{n \geqslant 1} \mathbf{p}_n^X(x).$$

P r o o f. 1) When the optimization variable is continuous this result is nothing else than the classical contraction principle of Large Deviation Theory. The extension to general optimization variables taking values in a metric space can be found in the classical texts [7, Theorem 4.2.23, p. 117] or [8, Lemma 2.1.4, p. 37]. The general result is contained in [2].

2) We note that if $\mathbf{p}^X$ is bounded we can define $a = \inf_E \mathbf{p}^X \in \mathbf{R}_-$. Then for all $x \in E$

$$\mathbf{p}^X(x) = \sup\{\mathbf{p}(\omega) \mid \omega\colon X(\omega) = x \quad \text{and} \quad \mathbf{p}(\omega) \geqslant a\}. \tag{3}$$

Now, let $\mathbf{p}_n^X$, $n \geqslant 1$, be the sequence of functions defined by $\mathbf{p}_n^X(x) = \sup\{\mathbf{p}(\omega) - n d_E(x, X(\omega)) \mid \omega\colon \mathbf{p}(\omega) \geqslant a\}$. Using the standard inequality $\sup\{(1)\} - \sup\{(2)\} \leqslant \sup\{(1) - (2)\}$ one gets easily

$$\mathbf{p}_n^X(x) - \mathbf{p}_n^X(y) \leqslant n \sup\Big\{d_E\big(y, X(\omega)\big) - d_E\big(y, X(\omega)\big) \mid \omega\colon \mathbf{p}(\omega) \geqslant a\Big\}$$

for any $x, y \in E$. Using the triangle inequality one can check that $|\mathbf{p}_n^X(x) - \mathbf{p}_n^X(y)| \leqslant n\, d_E(x, y)$. To complete the proof of 2) it suffices to show that

$$\mathbf{p}^X(x) = \inf_{n \geqslant 1} \mathbf{p}_n^X(x). \tag{4}$$

First for each $x \in E$ we note that $\inf_n\{-n d_E(x, X(\omega))\} = \delta_x \circ X(\omega)$, where $\delta_\alpha(\mathbf{B}) = 0$ if $\alpha = \mathbf{B}$ and $-\infty$ otherwise. Therefore

$$\mathbf{p}^X(x) = \sup_{\omega \in \Omega(a, \mathbf{P})} \Big(\mathbf{p}(\omega) + \delta_x\big(X(\omega)\big)\Big).$$

This yields the inequality

$$\mathbf{p}^X \leqslant \inf_n \mathbf{p}_n^X. \tag{5}$$

Finally, since the subsets $X^{-1}\{x\} \cap \Omega(a, \mathbf{P})$, $x \in E$, are compact there exists a sequence $(\omega_n)_{n\geqslant 1}$, $\omega_n \in \Omega(a, \mathbf{P})$, such that

$$\mathbf{p}_n^X(x) = \mathbf{p}(\omega_n) - n\, d_E(x, X(\omega_n)). \tag{6}$$

As usually there exists a subsequence $(\omega_{n_k})_{k\geqslant 1}$, $\omega_{n_k} \in \Omega(a, \mathbf{P})$, such that $\lim_{k\to\infty} \omega_{n_k} \stackrel{\text{def}}{=} \omega \in \Omega(a, \mathbf{P})$. Assume that $X(\omega) \neq x$. In this situation (6) implies that for any $b \in \mathbf{R}_-$ there exists $k \geqslant 1$ such that $\mathbf{p}_{n_k}^X(x) \leqslant b$. A contradiction with $\mathbf{p}_{n_k}^X(x) \geqslant a$ for all $k \geqslant 1$ and $x \in E$. This implies that $X(\omega) = x$ and in view of (6) it follows that

$$\inf_n \mathbf{p}_n^X(x) \leqslant \inf_k \mathbf{p}_{n_k}^X(x) \leqslant \lim_{k\to\infty} \mathbf{p}(\omega_{n_k}) \leqslant \mathbf{p}(\omega). \tag{7}$$

Combining (5) and (7) one gets the desired equality (4). Proposition 1 is proved.

It is important to note that the form of the density (2) is really justified only if the mapping satisfies the nearly continuity assumption of definition 3. In this situation the performance space $(E, \mathscr{B}(E), \mathbf{P}^X)$ is constructed in a completely rigorous way. By analogy with the mean-value operator in Probability Theory we define the $(\max, +)$-expectation of the optimization variable.

D e f i n i t i o n 4. Let X be an $\mathbf{R}_{\max}$-valued optimization variable defined on an optimization basis $(\Omega, \mathscr{B}(\Omega), \mathbf{P})$. We denote by $\mathbf{E}(X)$ the $(\max, +)$-expectation or mean value of the optimization variable X:

$$\mathbf{E}(X) = \int_\Omega^\oplus X(\omega) \odot \mathbf{p}(\omega) \odot d\omega = \int_{\mathbf{R}}^\oplus x \odot \mathbf{p}^X(x) \odot dx.$$

We list some basic properties of the $(\max, +)$-expectation (in what follows (X, Y) is a pair optimization variable taking values in $\mathbf{R}^2$ and $a, b \in \mathbf{R}_{\max}$):

$$\mathbf{E}(a \odot X \oplus b \oplus Y) = a \odot \mathbf{E}(X) \oplus b \oplus \mathbf{E}(Y),$$
$$X \geqslant Y \Rightarrow \mathbf{E}(X) \geqslant \mathbf{E}(Y),$$
$$Y = \mathbb{1}_A(X) \Rightarrow \mathbf{E}(Y) = \mathbf{P}^X(A) = \int_A^\oplus \mathbf{p}^X(x) \odot dx,$$

where $\mathbb{1}_A(x) = \mathbb{1}(= 0)$ if $x \in A$ and $\mathbb{0}(= -\infty)$ otherwise.

When X is an optimization variable taking values in a locally compact and separable metric space E and for any test function $\varphi \in \mathscr{C}_{\mathrm{K}}(E)$ the expectation of $\varphi(X)$ can be rewritten using the duality $\mathscr{C}'_{\mathrm{K}}(E) = \mathbb{M}(E)$

$$\mathbf{E}(\varphi(X)) = \int_E^\oplus \varphi(x) \odot \mathbf{p}^X(x) \odot dx = \langle \varphi, \mathbf{P}^X \rangle_{\mathscr{C}_{\mathrm{K}}(E), \mathbb{M}(E)}.$$

At this point it is already useful to give some examples of optimization variables and performance spaces which often appear in practical situations.

Example 2. $\Omega = \mathbf{R}^{d_1+d_2}$, d_1, d_2 are integers,

$$\mathbf{p}(\omega) = -\frac{1}{p_1} \left\|\pi_1(\omega)\right\|_{p_1}^{p_1} - \frac{1}{p_2} \left\|\pi_2(\omega) - h\big(\pi_1(\omega)\big)\right\|_{p_2}^{p_2}$$

where p_1, p_2 are real numbers $\geqslant 1$ and h: $\mathbf{R}^{d_1} \to \mathbf{R}^{d_2}$ is a continuous function, $\pi_1(\omega) \stackrel{\text{def}}{=} (\omega_1, \ldots, \omega_{p_1})$ and $\pi_2(\omega) \stackrel{\text{def}}{=} (\omega_{p_1+1}, \ldots, \omega_{p_1+p_2})$. In this situation the pair variable (X, Y) given by

$$\begin{aligned}(X, Y)\colon \Omega &\longrightarrow \mathbf{R}^{d_1} \times \mathbf{R}^{d_2}\\ \omega &\mapsto \big(X(\omega), Y(\omega)\big) = \big(\pi_1(\omega), \pi_2(\omega)\big)\end{aligned}$$

is an optimization variable with density

$$\mathbf{p}^{(X,Y)}(x, y) = -\frac{1}{p_1} \|x\|_{p_1}^{p_1} - \frac{1}{p_2} \|y - h(x)\|_{p_2}^{p_2}.$$

In addition one can easily check that the single mapping X: $\Omega \to \mathbf{R}^{d_1}$ is an optimization variable with density

$$\mathbf{p}^X(x) = -\frac{1}{p_1} \|x\|_{p_1}^{p_1}.$$

Another important example of optimization variable is the $(\max, +)$-version of the classical Gaussian random variables in Probability Theory. Since we will now be dealing with optimization variables defined on Banach space let us recall some basic properties. Let E be a separable Banach space. The Borel σ-field $\mathscr{B}(E)$ coincides with the smallest σ-field of subsets of E containing all subsets of the form $\{s \in E\colon f(x) \leqslant \alpha\}$, $\alpha \in \mathbf{R}$, $f \in E^\star$. Let E be a separable Hilbert space with inner product $\langle \cdot\,, \cdot \rangle$, norm $\|\bullet\|$ and equipped with the weak topology. In this situation the function

$$x \in E \longrightarrow \|x\| = \sup_{h \in E} \langle h, x \rangle \in \mathbf{R}_+$$

is lower semicontinuous as a limit of the continuous functions $x \in E \to \langle h, x\rangle \in \mathbf{R}$ and all subsets $\{x \in E\colon \|x\| \leqslant \alpha\}$, $\alpha \in \mathbf{R}$, are compact for the weak topology.

Definition 5. Let us assume that H is a separable Hilbert space with inner product $\langle \cdot\,, \cdot \rangle$, norm $\|\bullet\|$ and equipped with the weak topology. Let X be an optimization variable on a performance space $(\Omega, \mathscr{B}(\Omega), \mathbf{P})$ and taking values in H. It is called a quadratic optimization variable when the induced performance density $\mathbf{p}^X$ is given by $\mathbf{p}^X(x) = -\frac{1}{2}\|x\|^2$.

It is noteworthy that a linear transformation of a quadratic optimization variable is still quadratic.

Proposition 2. *Let X be a quadratic optimization variable taking values in a separable Hilbert space H with norm $\|\bullet\|$ and equipped with the weak topology. Let E be an auxiliary separable Hilbert space and let T: $H \to E$ be a bounded linear mapping such that $T(H) = E$.*

1) *The mapping* $Y = TX$ *is a quadratic optimization variable and*

$$\mathbf{p}^Y(y) = -\frac{1}{2}\,\|T^{-1}y\|^2$$

where T^{-1} *is the pseudoinverse of the transformation* T.

2) *Let* $T^\star\colon E \to H$ *be the adjoint of* T. *If* $TT^\star$ *is invertible then*

$$\mathbf{p}^Y(y) = -\frac{1}{2}\,\|T^\star(TT^\star)^{-1}y\|^2.$$

3) *Assume that* $E = \mathbf{R}^d$ *with* $d \geqslant 1$ *and* $T\colon H \to \mathbf{R}^d$ *is given by* $Tx = (T_1x, \dots, T_dx)$ *for any* $x \in H$ *where* $T_n\colon H \to \mathbf{R}$ *are bounded linear transformations satisfying*

$$T_m T_n^\star = \delta_m(n)\,\mathrm{Id} \quad \forall\, m, n \in \{1, \dots, d\}. \tag{8}$$

Then we have the formulae

$$\mathbf{p}^Y(y) = -\frac{1}{2}\sum_{m=1}^{d} \|T^\star y_m\|^2 \quad \forall\, y = (y_1, \dots, y_d) \in \mathbf{R}^d.$$

P r o o f. By the definition of the pseudoinverse T^{-1}, if $y \in \operatorname{Im} T$ then $T^{-1}y$ is the element of $T^{-1}\{y\}$ of the minimal norm. This implies the first assertion.

Let us prove 2). Using 1) the second assertion is clear.

To prove the last part of the proposition it suffices to note that (8) implies that $TT^\star = \mathrm{Id}$ and for any $y \in \operatorname{Im} T$

$$\|T^\star y\|^2 = \langle T^\star y, T^\star y\rangle = \left\langle \sum_{m=1}^{d} T_m^\star y_m, \sum_{m=1}^{d} T_m^\star y_m \right\rangle$$
$$= \sum_{m=1}^{d} \langle T_m^\star y_m, T_m^\star y_m\rangle = \sum_{m=1}^{d} \|T_m^\star y_m\|^2.$$

Proposition 2 is proved.

Next example initiate the study of optimization processes. Although these objects will be introduced in the end of this section this example shows that the above proposition gives a natural and simple way to calculate the densities of integral functions of a quadratic optimization variables.

E x a m p l e 3. Let U be a quadratic optimization variable with values in $H = L^2([0,1], \mathbf{R})$ (equipped with the weak topology) and with performance density $\mathbf{p}^U(u) = -\frac{1}{2}\int_0^1 u_t^2\,dt$. Denote by $X_t\colon H \to \mathbf{R}$, $t \in (0,1)$, the linear operator given by $X_t(u) = \int_0^t u_s\,ds$. It is easy to see that $\operatorname{Im} X_t = \mathbf{R}$ so that the adjoint operator $X_t^\star\colon \mathbf{R} \to H$ is defined by $X_t^\star(x)(s) = x\,1_{[0,t]}(s) \quad \forall\, s \in [0,1]$. It is clear that $X_tX_t^\star = t\,\mathrm{Id}$ is invertible for any $t \in (0,1)$ and therefore, by proposition 2, the mapping X_t is an quadratic optimization variable and its performance is given by

$$\mathbf{p}^X(x) = -\frac{1}{2}\int_0^t \left(\frac{x}{t}\right)^2 dt = -\frac{x^2}{2t} \quad \forall\, t \in (0,1).$$

Let $t_1 < \cdots < t_n$, $n \geqslant 1$, be a collection of times in $(0,1)$ and let $\sigma = (t_1,\ldots,t_n)$ be the corresponding n-uple of times. We note $X_\sigma\colon H \to \mathbf{R}^n$ the linear operator given by $X_\sigma(u) = (X_{t_1}(u),\ldots,X_{t_n}(u))\ \ \forall\, u \in H$. In view of proposition 2, X_σ is a quadratic optimization variable taking values in $\mathbf{R}^n$ and for any $x = (x_1,\ldots,x_n) \in \operatorname{Im} X_\sigma$

$$\mathbf{p}^{X_\sigma}(x) = \sup\left\{\mathbf{p}^U(u) \mid u \in H\colon X_\sigma(u) = x\right\}.$$

Note that $\mathbf{p}^{X_\sigma}$ may also be written $\mathbf{p}^{X_\sigma}(x) = \sup\{\mathbf{p}^U(u) \mid u \in H\colon R_\sigma(u) = \Delta_\sigma(x)\}$ with $R_\sigma = (R_1,\ldots,R_n)$, $R_i = (X_i - X_{i-1})/(t_i - t_{i-1})\ \forall\, 1 \leqslant i \leqslant n$ with the convention $X_0 = t_0 = 0$ and

$$\Delta_\sigma(x) = \left(\frac{x_1}{t_1}, \frac{x_2 - x_1}{t_2 - t_1}, \ldots, \frac{x_n - x_{n-1}}{t_n - t_{n-1}}\right).$$

Note that the adjoint operators $R_i^\star$, $1 \leqslant i \leqslant n$, are given by

$$R_i^\star(z)(s) = \frac{z}{t_i - t_{i-1}}\, 1_{[t_{i-1},t_i]}(s) \quad \forall\, 1 \leqslant i \leqslant n.$$

Clearly condition (8) holds and proposition 2 implies that for any $\sigma = (t_0,\ldots,t_n)$ the optimization variable X_σ is quadratic and its performance is given by

$$\mathbf{p}^{X_\sigma}(x) = -\frac{1}{2}\sum_{i=1}^n \frac{(x_i - x_{i-1})^2}{t_i - t_{i-1}}$$

with the convention $x_0 = t_0 = 0$.

2.2. Independence and conditioning. Our next aim is to transfer probabilistic axioms to optimization theory. The independence concept in such a framework is given by

Definition 6. Let $(\Omega, \mathscr{B}(\Omega), \mathbf{P})$ be a performance space. Two subsets $A, B \subset \mathscr{B}(\Omega)$ are said to be independent if

$$\mathbf{P}(A \cap B) = \mathbf{P}(A) \odot \mathbf{P}(B).$$

Let $(E, \mathscr{B}(E))$ and $(F, \mathscr{B}(F))$ be two Hausdorff topological spaces and let (X,Y) be a pair optimization variable taking values in $E \times F$. We assume that (X,Y) is defined on an optimization basis $(\Omega, \mathscr{B}(\Omega), \mathbf{P})$ and we denote by $\mathbf{p}^{X,Y}$ its performance density. We say that they are $\mathbf{P}$-independent if $\mathbf{p}^{X,Y}(x,y) = \mathbf{p}^X(x) \odot \mathbf{p}^Y(y)$ for any $x, y \in E$.

At the opposite of the latter concept of independence, we introduce the extension of Bayes's formula to performance measure. The conditional performance of event A assuming an event B (denoted by $\mathbf{P}(A \mid B)$) is by definition the ratio

$$\mathbf{P}(A \mid B) = \frac{\mathbf{P}(A \cap B)}{\mathbf{P}(B)}_{\circ} = \mathbf{P}(A \cap B) - \mathbf{P}(B)$$

with the convention $\mathbf{P}(A \mid B) = \mathbb{0}$ when $\mathbf{P}(B) = \mathbb{0}$. Continuing in the same vein, let $(E, \mathscr{B}(E))$ be a Hausdorff topological space and let (X, Y) be a pair of E-valued optimization variables with density $\mathbf{p}^{(X,Y)}$. The function $\mathbf{p}^{X|Y}$ defined by

$$\mathbf{p}^{X|Y}(x \mid y) = \frac{\mathbf{p}^{(X,Y)}(x, y)}{\mathbf{p}^{Y}(y)}_{\circ} = \mathbf{p}^{X,Y}(x, y) - \mathbf{p}^{Y}(y)$$

with the convention $\mathbf{p}^{X|Y}(x \mid y) = \mathbb{0}$ if $\mathbf{p}^{Y}(y) = \mathbb{0}$ is called the conditional performance density of X with respect to Y.

One can easily check that for any $y \in E$ such that $\mathbf{p}^{Y}(y) \neq \mathbb{0}$, the function $x \mapsto \mathbf{p}^{X|Y}(x \mid y)$ is a well-defined performance density.

Let us give some comments on this definition. Using the framework of the previous development we discuss now the meaning of the conditional density $\mathbf{p}^{X|Y}$. Recall that the densities $\mathbf{p}^{(X,Y)}$ and $\mathbf{p}^{Y}$ may be written

$$\mathbf{p}^{(X,Y)}(x, y) = \sup \{\omega \in \Omega \mid \omega\colon X(\omega) = x \text{ and } Y(\omega) = y\},$$
$$\mathbf{p}^{Y}(y) = \sup \{\omega \in \Omega \mid \omega\colon Y(\omega) = y\}.$$

On the other hand

$$\mathbf{p}^{Y}(y) = \sup_{x \in E} \mathbf{p}^{(X,Y)}(x, y). \tag{9}$$

To see this claim we note that

$$\begin{aligned} \sup_{x \in E} \mathbf{p}^{(X,Y)}(x, y) &= \sup_{x \in E} \sup_{\omega \in Y^{-1}(y) \cap X^{-1}(x)} \mathbf{p}(\omega) \\ &= \sup_{\omega \in Y^{-1}(y)} \mathbf{p}(\omega) = \mathbf{p}^{Y}(y). \end{aligned} \tag{10}$$

This indicates that the conditional performance $\mathbf{p}^{X|Y}(\bullet \mid y)$ is indeed a normalized idempotent measure on the subsets $A \cap Y^{-1}(y)$, $A \in \mathscr{B}(\Omega)$.

Clearly we have also the following properties.

Proposition 3. *Let (X, Y, Z) be a triplet of optimization variable with performance density $\mathbf{p}^{(X,Y,Z)}$. The following properties hold.*

For any z with $\mathbf{p}^{Z}(z) \neq \mathbb{0}$ we have

$$\mathbf{p}^{X|Z}(x \mid z) = \int_{E}^{\oplus} \mathbf{p}^{X|Y,Z}(x \mid y, z) \odot \mathbf{p}^{Y|Z}(y \mid z) \odot dz$$

and therefore

$$\mathbf{p}^{X}(x) = \int_{E}^{\oplus} \mathbf{p}^{X|Y}(x \mid y) \odot \mathbf{p}^{Y}(y) \odot dy.$$

In addition, for any (y, z) such that $\mathbf{p}^{Y,X}(y,z) \neq \mathbb{0}$

$$\mathbf{p}^{X|Y,X}(x \mid y,\, z) = \delta_z(x) \odot \mathbf{p}^{X|Y}(x \mid y),$$

where $\delta_z(x) = 0$ if $z = x$ and $-\infty$ otherwise.

2.3. Optimization processes. After the introduction of the basic axioms of performance measure analysis the next step consists in introducing optimization processes. As it will be obvious in the forthcoming development the objects on which the optimization processes are sought may vary considerably. It is therefore necessary to undertake the study in an abstract setting.

To achieve this program we need to introduce some additional notation. The basic definition of an optimization process with time space $\mathcal{T}$ ($\mathcal{T} = (s,t) \subset \mathbf{R}$ or $\mathbf{Z}$ with $s \leqslant t$) will consist of a sequence of optimization variables $\{X_t; t \in \mathcal{T}\}$ taking values in a Polish space $(S, \mathcal{B}(S))$.

Definition 7. An S-valued optimization (or decision) process with time space $\mathcal{T}$ is a system $(\Omega,\, \mathcal{B}(\Omega),\, \mathbf{P},\, X = (X_t)_{t \in \mathcal{T}})$ defined by: 1) a performance space $(\Omega, \mathcal{B}(\Omega), \mathbf{P})$ and 2) an optimization variable $X\colon \Omega \to \mathcal{E}(\mathcal{T}, S)$, where $\mathcal{E}(\mathcal{T}, S)$ is a metric space of mappings $x\colon\ \mathcal{T} \to S$ furnished with a topology such that the natural restriction mappings $\pi_{[s,t]}\colon\ \mathcal{E}(\mathcal{T}, S) \to \mathcal{E}([s,t], S)$, $s \leqslant t$, and $\pi_{[s,t],[s',t']}\colon\ \mathcal{E}([s,t],S) \longrightarrow \mathcal{E}([s',t'],S)$, $s \leqslant s' \leqslant t' \leqslant t$, are continuous.

The optimization variable X_t is called the state of X at the time $t \in \mathcal{T}$ and the curve $t \in \mathcal{T} \mapsto X_t(\omega) \in S$ is called the trajectory of the path $\omega \in \Omega$.

Before presenting some examples of optimization processes we remark that the above definition may cover continuous or discontinuous optimization processes. More precisely, let $\mathcal{T} = \mathbf{R}$ and let $\mathcal{E}(\mathcal{T}, S) = \mathcal{D}(\mathcal{T}, S)$ be the space of right continuous paths which have a left limit. There exists a metric d on this space such that the natural restriction mappings are continuous. Moreover the relative topology which $\mathcal{D}(\mathcal{T}, S)$ inherits as a subset of the latter coincide with the uniform convergence on compact sets (see for instance Lemma 5.4.29 in [8]).

When X is an optimization process on $\mathcal{T}$ the function $X_t\colon\ \Omega \to S$ is given by $X_t = \pi_t \circ X$, where π_t is the projection on the time $t \in \mathcal{T}$. According to the definition of an optimization variable and π_t being continuous the mapping X_t satisfies the continuity assumption of Definition 3. This implies that X_t is an optimization variable with performance density

$$\mathbf{p}_t^X(x) = \sup \{\mathbf{p}(\omega) \mid \omega\colon\ X_t(\omega) = x\}.$$

More generally the composite mappings $X_{[s,t]} = \pi_{[s,t]} \circ X\colon\ \Omega \longrightarrow \mathcal{E}([s,t], S)$ can be regarded as optimization variables with performance density

$$\mathbf{p}^X_{[s,t]}(x) = \sup \{\mathbf{p}^X(z) \mid z \in \mathcal{E}(\mathcal{T},\, S)\colon\ \pi_{[s,t]}(z) = x\}.$$

Summarizing we have the following proposition.

Proposition 4. *Let $(\Omega, \mathscr{B}(\Omega), \mathbf{P}, X = (X_t)_{t \in \mathscr{T}})$ be an optimization process with time space T and taking values in a metric space (S, d_S). The following properties are satisfied.*

1) *For any $s \leqslant t$, $s, t \in \mathscr{T}$, the composite mapping $X_{[s,t]} = \pi_{[s,t]} \circ X \colon \Omega \longrightarrow \mathscr{E}([s,t], S)$ is an optimization variable and its performance density is the upper continuous function $\mathbf{p}^X_{[s,t]} \colon \mathscr{E}([s,t], S) \longrightarrow \mathbf{R}_{\max}$ (with compact level sets) given by*

$$\mathbf{p}^X_{[s,t]}(x) = \sup\left\{\mathbf{p}(\omega) \mid \omega \in \Omega \colon X(\omega) = x\right\} \quad \forall x \in \mathscr{E}([s,\, t],\, S).$$

2) *For any $s \leqslant s' \leqslant t' \leqslant t$ the performance densities $\mathbf{p}^X_{[s,t]}$ and $\mathbf{p}^X_{[s',t']}$ satisfy*

$$\mathbf{p}^X_{[s',t']}(x_2) = \sup\left\{\mathbf{p}^X_{[s,t]}(x_1) \mid x_1 \in \mathscr{E}([s,\, t],\, S) \colon \pi_{[s,t],[s',t']}(x_1) = x_2\right\}$$

for any $x_2 \in \mathscr{E}([s', t'], S)$.

We now present a number of examples which can be handled in our framework.

E x a m p l e 4. Let us assume that $\mathscr{T} = [0,1]$, $S = \mathbf{R}^d$, $d \geqslant 1$, and

$$\Omega = \mathscr{C}_0([0,\, 1],\, \mathbf{R}^d) = \left\{\omega \in \mathscr{C}([0,\, 1],\, \mathbf{R}^d) \colon \omega(0) = 0\right\},$$

$$\mathscr{E}(\mathscr{T},\, S) = \mathscr{C}([0,\, 1],\, \mathbf{R}^d)$$

equipped with the uniform norm. For any $\omega \in \Omega$ we define $X \colon \Omega \longrightarrow \mathscr{C}([0,1], \mathbf{R}^d)$ by $X_t(\omega) = \omega_t + \int_0^t f(X_s(\omega))\, ds$ $\forall t \in T$, where $f \colon \mathbf{R}^d \to \mathbf{R}^d$ satisfies the Lipschitz condition $\|f(x) - f(y)\| \leqslant C\, \|x - y\|$, $C < \infty$. One can check that X is a continuous mapping. Now, let $\mathbf{p} \colon \Omega \to \mathbf{R}_{\max}$ be defined by

$$\mathbf{p}(\omega) = \begin{cases} -\dfrac{1}{2} \displaystyle\int_0^1 \|\dot{\omega}_t\|^2\, dt & \text{if } \omega \in H^1([0,\, 1],\, \mathbf{R}^d), \\ -\infty & \text{otherwise.} \end{cases}$$

The function $\mathbf{p}$ is u.s.c. on $\mathscr{C}_0([0,1], \mathbf{R}^d)$. On the other hand using the fact that bounded subsets of $H^1([0,1], \mathbf{R}^d)$ are relatively compact in Ω, the subsets

$$\left\{\omega \in \Omega \colon \mathbf{p}(\omega) \geqslant \alpha\right\} = \left\{\omega \in H^1([0,\, 1],\, \mathbf{R}^d) \colon \|\omega\|_{H^1} \leqslant \sqrt{2|\alpha|}\right\}, \quad \alpha \in \mathbf{R}_-,$$

are compacts. Hence, $\mathbf{p}$ is a well-defined performance density on $(\Omega, \mathscr{B}(\Omega))$ (cf. p. 12, 13 in [8]).

Using the above observations one can check that X is a well-defined optimization variable on $(\Omega, \mathscr{B}(\Omega))$ with performance density

$$\mathbf{p}^X(x) = -\frac{1}{2} \int_0^1 \|\dot{x}_t - f(x_t)\|^2\, dt$$

if the integral exists and $\mathbb{0} = -\infty$ otherwise.

Example 5. Let us assume that $\mathcal{T} = [0,1]$, $S = \mathbf{R}^d$, $d \geqslant 1$, $p \geqslant 1$ and $\Omega = L^p([0, 1], \mathbf{R}^d)$, $\mathcal{E}([0, 1], \mathbf{R}^d) = \mathcal{C}([0, 1], \mathbf{R}^d)$ equipped respectively with the weak convergence topology and the uniform norm. For any $\omega \in \Omega$ define $X\colon \Omega \longrightarrow \mathcal{C}([0,1],\mathbf{R}^d)$ by setting

$$\begin{cases} X^{\bullet}(\omega)_t = A(X(\omega)_t) + B\,\omega_t \\ X(\omega)_0 = x_0 \in \mathbf{R}^d \end{cases}$$

where $B \in \mathbf{R}^{d\times d}$ and $A\colon \mathbf{R}^d \to \mathbf{R}^d$ is a bounded and uniformly Lipschitz mapping. Under our assumptions one can check that X is continuous.

Now, let $\mathbf{p}\colon \Omega \to \mathbf{R}_{\max}$ be defined as $\mathbf{p}(\omega) = -p^{-1}\|\omega\|_p^p$. It is clear that $(\Omega, \mathcal{B}(\Omega), \mathbf{P})$ is an optimization basis. Using proposition 1 we conclude that X is an optimization process with performance density

$$\mathbf{p}^X(x) = \sup\left\{ -p^{-1}\|\omega\|_p^p;\ \omega \in L^p\big([0, 1], \mathbf{R}^d\big)\colon X(\omega) = x\right\}$$

$\forall x \in \mathcal{C}([0, 1], \mathbf{R}^d)$. Hence

$$\mathbf{p}^X(x) = -\frac{1}{p}\int_0^1 \big|\overset{\bullet}{x}_t - A(x_t)\big|^p\, dt$$

when the integral is defined and $\mathbb{0} = -\infty$ otherwise.

Example 6. As an example of discrete time optimization process let us consider the case $\mathcal{T} = \{0,\dots,T\}$, $T \geqslant 1$, and $\Omega = U^T$ where (U, d_U) is a metric space. For each $1 \leqslant t \leqslant T$ we denote by $\mathbf{p}_t\colon U \to \mathbf{R}_{\max}$ a performance measure on $(U, \mathcal{B}(U))$. Assume that U^T is equipped with the metric $d(\omega,\omega') = \sup_{1\leqslant t\leqslant T} d_U(\omega_t,\omega'_t)$ and the performance densities $\{\mathbf{p}_t; 0 \leqslant t \leqslant T\}$ are choosen so that the mapping $\mathbf{p}(\omega) = \sum_{t=1}^T \mathbf{p}_t(\omega_t)$ is a well-defined performance density on $(\Omega, \mathbf{B}(\Omega))$. Let (S, d_S) be another metric space. For any $\omega \in \Omega$ we define $X(\omega) \in S^{T+1}$ by setting

$$\begin{cases} X(\omega)_t = F\big(X(\omega)_{t-1}, \omega_t\big), & 1 \leqslant t \leqslant T, \\ X(\omega)_0 = x_0 \in S, \end{cases}$$

where $F\colon S\times U \to S$ is a continuous function such that $d_S(F(x,u), F(y,v)) \leqslant d_S(x,z) + d_U(u,v)$. Under our assumptions the mapping $X\colon U^T \to S^{T+1}$ can be regarded as an optimization variable with performance density

$$\mathbf{p}^X(z) = \sup\left\{ \sum_{t=1}^T \mathbf{p}_t(\omega_t);\ \omega \in U^T\colon X(\omega) = z\right\}.$$

Using Lemma 1 for any $x_1, \dots, x_T \in S$ we can find some $\omega(x) \in U^T$ such that $\mathbf{p}^X(x_0, x_1, \dots, x_T) = \sum_{t=1}^T \mathbf{p}_t(\omega_t(x))$. Finally, if we denote by $\mathbf{p}^X_{t|r,\dots,s}$, $r \leqslant s \leqslant t$, the conditional performance of X_t with respect to $(X_r, \dots, X_s)$ one can check that $\mathbf{p}^X_{t|0,\dots,t-1}(x_t \mid x_0, \dots, x_{t-1}) = \sup\{\mathbf{p}_t(u);\ u \in U\colon F(x_{t-1}, u) =$

$x_t\} = \mathbf{p}^X_{t|t-1}(x_t \mid x_{t-1})$ and therefore the performance density of X can be rewritten in the following form:

$$\mathbf{p}^X(x_0, x_1, \ldots, x_T) = \sum_{t=1}^{T} \mathbf{p}^X_{t|t-1}(x_t \mid x_{t-1}).$$

2.4. Bellman–Maslov processes. The essence of the Bellman–Hamilton–Jacobi Theory can be introduced in forward time with initial penalty. This shows the central role played by the concatenation semigroup of optimization transition performances in Bellman's optimality principle as probability transitions in Markov systems.

We state that this principle may be viewed as a basic definition of optimization processes like Markov's property rather than a conclusion. In other words, a Maslov process is an optimization process that satisfies the (max, +)-version of Markov's causality principle. The reversal of time yields optimal control processes of regulation type. The groundwork for the theory of Markov stochastic processes was laid in 1906 by A. A. Markov. In his investigation of connected experiments, he formulated the principle that the «future» is independent of the «past» when the «present» is known. This principle is the causality principle of classical physics state carried over to stochastic dynamical systems. It specifies that the knowledge of the state of a system at a given time is sufficient to determine its state at any future time.

The following concept is the extension of the Markov causality principle in the Maslov optimization framework.

Definition 8. Let $(\Omega, \mathscr{B}(\Omega), \mathbf{P},\ X = \{X_t\}_{t\in\mathscr{T}})$ be an S-valued optimization process with time space $\mathscr{T}$. It is called a *Bellman–Maslov process* whenever its future and its past are independent when its present is known. In other words, X is a Bellman–Maslov process if for every $r \leqslant s \leqslant t$ we have the relation

$$\mathbf{p}^X_{[r,s],[s,t]|s}(x_1, x_2 \mid x) = \mathbf{p}^X_{[r,s]|s}(x_1 \mid x) \odot \mathbf{p}^X_{[s,t]|s}(x_2 \mid x) \tag{11}$$

for any $x_1 \in \mathscr{E}([r,s], S)$, $x_2 \in \mathscr{E}([s,t], S)$ and $x \in S$; where $\mathbf{p}^X_{[r,s],[s,t]|s}$ denotes the conditional performance density of the pair $X_{[r,s],[s,t]} = (X_{[r,s]}, X_{[s,t]})$ with respect to X_s.

One consequence of (11) is that for any $s \leqslant t$ the marginal performance density of X at time t can be computed from the marginal of X at time s: $\mathbf{p}^X_t(x) = \int^{\oplus}_S \mathbf{p}^X_{t|s}(x \mid z) \odot \mathbf{p}^X_s(z) \odot dz$. In other words

$$\mathbf{p}^X_t(x) = \sup_{z\in S} \mathbf{p}^X_{t|s}(x \mid z) + \mathbf{p}^X_s(z) \quad \text{(Bellman's Principle)}.$$

This means that Bellman's causality principle is the (max, +)-version of the classical Chapman–Kolmogorov's equation for Markov stochastic processes.

If the time space $\mathcal{T} = \mathrm{N}$, then an optimization process satisfying (11) will be called a Bellman–Maslov chain in reference to Markov chain theory. If the time the transitions performance density $\mathbf{p}^X_{t|s}$ only depends on $|t-s|$, then the optimization process is said to be time homogeneous.

Example 7. Let X be a discrete time optimization process taking values in a metric space S. Assume that its performance density $\mathbf{p}^X\colon \mathscr{E}(\mathrm{N}, S) \to \mathbf{R}_{\max}$ is given by $\mathbf{p}^X(x) = c_0(x_0) + \sum_{t\geqslant 1} c_t(x_{t-1}, x_t)$, where $c_0\colon S \to \mathbf{R}_{\max}$ and $c_t(x, \bullet)\colon S \to \mathbf{R}_{\max}$ are performance densities on S. By proposition 1 for any $t \geqslant 0$ the sequence $(X_0, \dots, X_t)$ can be regarded as an S^{t+1}-valued optimization variable with performance density $\mathbf{p}^X_{0,\dots,t}(x_0, \dots, x_t) = c_0(x_0) + \sum_{s=1}^t c_s(x_{s-1}, x_s)$ and therefore the conditional density of X_t with respect to $(X_0, \dots, X_{t-1})$ is given by

$$\mathbf{p}^X_{t|0,\dots,t-1}(x_t \mid x_0, \dots, x_{t-1}) = c_t(x_{t-1}, x_t) = \mathbf{p}^X_{t|t-1}(x_t \mid x_{t-1}).$$

This yields

$$\mathbf{p}^X_t(x) = \sup_{z\in S} \left(\mathbf{p}^X_{t|t-1}(x \mid z) + \mathbf{p}^X_{t-1}(z)\right) = \int_S^{\oplus} \mathbf{p}^X_{t|t-1}(x \mid z) + \mathbf{p}^X_{t-1}(z) \odot dz.$$

Example 8. Assume that Ω is a Polish space of controls on $(0,1)$ and taking values in a metric space U. For each $\omega \in \Omega$ we define $X\colon \omega \in \Omega \longrightarrow X(\omega) \in \mathscr{C}((0,1), \mathbf{R}^d)$ by setting

$$\begin{cases} \overset{\bullet}{X}(\omega)_t = f(X_t(\omega), \omega_t), \\ X(\omega)_0 = x_0 \in \mathbf{R}^d, \end{cases} \tag{12}$$

where $f\colon \mathbf{R}^d \times U \longrightarrow \mathbf{R}^d$ is a sufficiently regular mapping such that the mapping $X\colon \omega \in \Omega \longrightarrow \mathscr{C}((0,1), S)$ is continuous (as usually the space $\mathscr{C}((0,1), \mathbf{R}^d)$ is furnished with the uniform topology). Finally suppose we are given a performance measure $\mathbf{P}$ on $\mathscr{B}(E)$ with a density of the form $\mathbf{p}(\omega) = \int_0^1 L_t(X_t(\omega), \omega_t)\, dt$ for some $L_t\colon \mathbf{R}^d \times U \longrightarrow \mathbf{R}_{\max}$. If the following condition is met

(H) $\forall x \in \mathbf{R}^d\ \forall t \in (0,1) \exists \omega \in \Omega$ such that $\int_t^1 L_t(X^x_s(\omega), \omega_s)\, ds = 0$ where $\{X^x_s(\omega)\colon t \leqslant s \leqslant 1\}$ is the solution of (12) such that $X_t(\omega) = x$,

then $X_{[0,t]} = \pi_{[0,t]} \circ X$ is an optimization variable with performance density

$$\mathbf{p}^X_{[0,t]}(x) = \sup\left\{ \int_0^t L_s(X_s(\omega), \omega_s)\, ds;\ \omega \in \Omega\colon X_{[0,t]}(\omega) = x \right\}.$$

Moreover if we denote by $\mathbf{p}^X_t$ the performance density of X_t we have for any $s \leqslant t$

$$\mathbf{p}^X_t(x) = \sup_{z\in \mathbf{R}^d} \left(\mathbf{p}^X_{t|s}(x \mid z) + \mathbf{p}^X_s(z)\right) = \int_{\mathbf{R}^d}^{\oplus} \left(\mathbf{p}^X_{t|s}(x \mid z) + \mathbf{p}^X_s(z)\right) \odot dz$$

where $\mathbf{p}^X_{t|s}(x \mid z) = \sup\{\int_s^t L_r(X_r(\omega), \omega_r)\, dr;\ \omega \in \Omega\colon\ X_s = z$ and $X_t(\omega) = x\}$.

REFERENCES

1. *Akian M., Quadrat J. P., Viot M.* Bellman processes. — Lecture Notes in Control and Inform. Sci., 1994, v. 199, p. 302–311.
2. *Azencott R.* Grandes déviations et applications. — Lecture Notes in Math., 1980, v. 774, p. 1–176.
3. *Baccelli F. L., Cohen G., Olsder G. J., Quadrat J. P.* Synchronization and Linearity. An Algebra for Discrete Event Systems. Chichester: Wiley, 1992. 489 p.
4. *Cox H.* Estimation of state variables via dynamic programming. — Proceedings of the 1964 Joint Automatic Control Conference (Stanford, California), p. 376–381.
5. *Cuninghame-Green R.* Minimax algebra. — Lecture Notes in Econ. and Math. Systems, 1979, v. 166, p. 1–258.
6. *Del Moral P.* Maslov optimization theory: optimality versus randomness. — Appendix in *Kolokoltsov V. N., Maslov V. P.* Idempotent Analysis and Applications. Dordrecht: Kluwer, 1997.
7. *Dembo A., Zeitouni O.* Large Deviations Techniques and Applications. Boston: Jones and Bartlett, 1993. 346 p.
8. *Deuschel J.-D., Stroock D. W.* Large Deviations. Boston: Academic Press, 1989. 307 p.
9. *Дудников П. И., Самборский С. Н.* Эндорморфизмы полумодулей над полукольцами с идемпотентной операцией. Препринт № 87-48. Киев: Ин-т матем. АН Украины, 1987.
10. *Freidlin M.I., Wentzell A. D.* Random Perturbations of Dynamical Systems. Berlin etc.: Springer-Verlag, 1984.
11. *Колокольцов В. Н.* Полукольцевые аналоги линейных эквивалентных пространств. — Тезисы Всесоюзной школы «Оптимальное управление, геометрия и анализ», Кемерово, 1988, с. 30.
12. *Колокольцов В. Н., Маслов В. П.* Идемпотентный анализ как аппарат теории управления и оптимального синтеза. — Функц. анализ и прилож., 1989, т. 23, № 1, с. 1–14; № 4, с. 53–62.
13. *Колокольцов В. Н., Маслов В. П.* Общий вид эндоморфизмов в пространстве непрерывных функций со значениями в числовом коммутативном полукольце (с операцией $\oplus = \max$). — Докл. АН СССР, 1987, т. 295, № 2, с. 283–287.
14. *Litvinov G. L., Maslov V. P.* Correspondence Principle for Idempotent Calculus and Some Computer Applications. Preprint IHES-M-95-33. Institut des Hautes Etudes Scientifiques: Bures-sur-Yrette, 1995.
15. *Маслов В.* Операционные методы. М.: Мир, 1987.
16. *Maslov V.* Quasilinear systems which are linear in some semimoduli. — Congrès international sur les problèmes hyperboliques, 13–17 Janvier 1986, Saint Etienne, France.
17. *Maslov V. P., Samborskii S. N.* (eds). Idempotent Analysis. Providence, RI: Amer. Math. Soc., 1992. 210 p.

Поступила в редакцию
10.VI.1998