SPRINGER BRIEFS IN MATHEMATICS

Toshio Mikami

Stochastic Optimal Transportation

Stochastic Control with Fixed Marginals

Springer

SpringerBriefs in Mathematics

SpringerBriefs present concise summaries of cutting-edge research and practical applications across a wide spectrum of fields. Featuring compact volumes of 50 to 125 pages, the series covers a range of content from professional to academic. Briefs are characterized by fast, global electronic dissemination, standard publishing contracts, standardized manuscript preparation and formatting guidelines, and expedited production schedules.

Typical topics might include:

A timely report of state-of-the art techniques A bridge between new research results, as published in journal articles, and a contextual literature review A snapshot of a hot or emerging topic An in-depth case study A presentation of core concepts that students must understand in order to make independent contributions

SpringerBriefs in Mathematics showcases expositions in all areas of mathematics and applied mathematics. Manuscripts presenting new results or a single new result in a classical field, new field, or an emerging topic, applications, or bridges between new results and already published works, are encouraged. The series is intended for mathematicians and applied mathematicians. All works are peer-reviewed to meet the highest standards of scientific literature.

Titles from this series are indexed by Scopus, Web of Science, Mathematical Reviews, and zbMATH.

More information about this series at http://www.springer.com/series/10030

Toshio Mikami

Stochastic Optimal Transportation

Stochastic Control with Fixed Marginals

Springer

Toshio Mikami
Department of Mathematics
Tsuda University
Tokyo, Japan

ISSN 2191-8198 ISSN 2191-8201 (electronic)
SpringerBriefs in Mathematics
ISBN 978-981-16-1753-9 ISBN 978-981-16-1754-6 (eBook)
https://doi.org/10.1007/978-981-16-1754-6

This Springer imprint is published by the registered company Springer Nature Singapore Pte Ltd.
The registered company address is: 152 Beach Road, #21-01/04 Gateway East, Singapore 189721, Singapore

Preface

The construction of a stochastic process from given marginal distributions is an important part of the so-called marginal problem. A Markovian Bernstein process on [0, 1] constructed from two endpoint marginals at times $t = 0, 1$ and from the Bernstein transition density solves Schrödinger's functional equation. It is the so-called h-path process, provided the Markov process solves a stochastic differential equation. The Markov diffusion process and its construction from a solution of the Fokker–Planck equation are called the Nelson process and Nelson's problem, respectively. (In E. Nelson's original problem, he considered the Fokker–Planck equation which is satisfied by the square of the absolute value of a solution to Schrödinger's equation.) In the second part of our dissertation under the supervision of Professor Wendell H. Fleming, Brown University, we gave an approach to Nelson's problem via the continuum limit of a class of stochastic controls with given two endpoint marginals. It is closely related to the optimal transportation problem, in that we find optimal dynamics, in the minimization problem, among those who have the same partial information on marginal distributions. Nearly 15 years later, we gave a probabilistic proof to Monge's problem with a quadratic cost by the zero-noise limit of h-path processes, which encouraged us to consider optimal transportation for semimartingales which we call the **stochastic optimal transportation**. It can be considered a class of marginal problems.

In Chap. 1, we introduce Monge's problem and Schrödinger's to compare them so that one can see the relation between the nonstochastic and stochastic optimal transportations.

The Duality Theorems for stochastic optimal transportation problems are useful for considering marginal problems, including Nelson's. Indeed, we used them to construct a semimartingale from the Fokker–Planck equation under a general integrability condition. The construction of a semimartingale from the Fokker–Planck equation is also called the superposition principle and has been remarkably developed in the last several years.

In Chap. 2, we give two classes of stochastic optimal transportation problems. As an application, by the superposition principle, we give our recent progress on the Duality Theorems for stochastic optimal transportation problems with a convex cost function. We also give a sufficient condition for the finiteness of the minimum in stochastic optimal transportation problem and discuss the relation between the nonstochastic and stochastic optimal transportations by the zero-noise limit.

In Chap. 3, we consider the finiteness, the semiconcavity, and the continuity of the minimum in Schrödinger's problem. We also consider the regularity of the solution to Schrödinger's functional equation, marginal problems for stochastic processes, and stochastic optimal transportation with a nonconvex cost function in the one-dimensional case.

The generalizations of the results in Chap. 3 are our future projects.

Besides the page limit, since it is beyond our ability to introduce all topics in this rapidly developing field, we ask readers to find papers on missing topics such as stochastic mechanics, large deviations, entropic functional inequalities, martingale optimal transports, etc.

We would like to thank anonymous referees for constructive suggestions and for informing us about missing references, and thank Mr. Masayuki Nakamura, Springer for his support. We would also like to acknowledge the financial support by JSPS KAKENHI Grant Numbers JP16H03948 and 19K03548.

I would like to thank Professors Wendell H. Fleming, Hitoshi Ishii, and Masayoshi Takeda for their academic influences and constant encouragement since I was a graduate student.

Lastly, I would like to thank my wife Mari for her unconditional support in my lifetime.

Tokyo, Japan
February 2021

Toshio Mikami

Contents

Notation

OT	Optimal transportation problem
SOT	Stochastic optimal transportation problem
FBSDE	Forward–backward stochastic differential equation
HJB	Hamilton–Jacobi–Bellman
$\mathbf{B}(S)$	Set of all Borel measurable subsets of a topological space S
$\mathcal{P}(S)$	Set of all Borel probability measures on a topological space S
$\mathcal{P}_{ac}(\mathbb{R}^d)$	$\{P \in \mathcal{P}(\mathbb{R}^d) \mid P(dx) \ll dx\}$
$\mathcal{P}_r(\mathbb{R}^d)$	$\{P \in \mathcal{P}(\mathbb{R}^d) \mid \int_{\mathbb{R}^d} \|x\|^r P(dx) < \infty\}$ for $r \geq 1$
$\mathcal{P}_{r,ac}(\mathbb{R}^d)$	$\mathcal{P}_{ac}(\mathbb{R}^d) \cap \mathcal{P}_r(\mathbb{R}^d)$ for $r \geq 1$
$\mathcal{A}(P_0, P_1)$	$\{\mu \in \mathcal{P}(\mathbb{R}^{2d}) \mid \mu(dx \times \mathbb{R}^d) = P_0(dx), \mu(\mathbb{R}^d \times dx) = P_1(dx)\}$
$AC(S)$	Set of all absolutely continuous functions on S
$USC(S)$	Set of all upper semicontinuous functions on S
$LSC(S)$	Set of all lower semicontinuous functions on S
$UC_b(S)$	Set of all uniformly continuous bounded functions on S
$C^{i,j}(S)$	Set of all functions on S, which are ith and jth continuously differentiable in the first and the second variables, respectively
$C_b^{i,j}(S)$	Set of all functions on S, which have bounded continuous derivatives up to the ith and jth orders in the first and the second variables, respectively
P^X	Probability distribution of a random variable X
$\dot{x}(t)$	$dx(t)/dt$
$\langle x, y \rangle$	$\sum_{i=1}^d x_i y_i$ for $x = (x_i)_{i=1}^d, y = (y_i)_{i=1}^d \in \mathbb{R}^d$
$\langle A, B \rangle$	$\sum_{i,j=1}^d a_{ij} b_{ij}$ for $A = (a_{ij})_{i,j=1}^d, B = (b_{ij})_{i,j=1}^d \in M(d, \mathbb{R})$
D_x	$(\partial/\partial x_i)_{i=1}^d$
D_x^2	$(\partial^2/\partial x_i \partial x_j)_{i,j=1}^d$
$\|\|f\|\|_\infty$	$\sup_{x \in S} \|f(x)\|, \quad f \in C(S)$
B_r	$\{x \in \mathbb{R}^d : \|x\| \leq r\}$

$H(P\|Q)$	$= \int_S \{\log(P(dx)/Q(dx))\} P(dx)$ if $P \ll Q$; $= \infty$, otherwise.
$\mathscr{S}(P)$	$= \int_{\mathbb{R}^d} \{\log p(x)\} p(x)dx$ if $P(dx) = p(x)dx$; $= \infty$, otherwise.
$\|\|x\|\|_{L^2(P)}$	$\sqrt{\int_{\mathbb{R}^d} \|x\|^2 P(dx)}$, $P \in \mathscr{P}_2(\mathbb{R}^d)$
(A.0.0)	(i) $\sigma_{ij} \in C_b([0,1] \times \mathbb{R}^d)$, $i, j = 1, \cdots, d$. (ii) σ is nondegenerate.
(A.0)	$\sigma_{ij} \in C_b^1([0,1] \times \mathbb{R}^d)$, $i, j = 1, \cdots, d$.
(A.1)	(i) $L \in C([0,1] \times \mathbb{R}^d \times \mathbb{R}^d; [0,\infty))$. (ii) For $(t,x) \in [0,1] \times \mathbb{R}^d$, $u \mapsto L(t,x;u)$ is convex.
(A.2)	$\lim_{\|u\| \to \infty} \inf\{L(t,x;u) \| (t,x) \in [0,1] \times \mathbb{R}^d\}/\|u\| = \infty$.
(A.3)	(i) $\partial L(t,x;u)/\partial t$ and $D_x L(t,x;u)$ are bounded on $[0,1] \times \mathbb{R}^d \times B_R$ for all $R > 0$, where $B_R := \{x \in \mathbb{R}^d \| \|x\| \le R\}$. (ii) C_L is finite, where $C_L := \sup\left\{L(t,x;u)/(1 + L(t,y;u)) \| 0 \le t \le 1, x, y, u \in \mathbb{R}^d\right\}$.
(A.4)	(i) "σ is an identity" or "σ is uniformly nondegenerate, $\sigma_{ij} \in C_b^{1,2}([0,1] \times \mathbb{R}^d)$, $i, j = 1, \cdots, d$ and there exist functions L_1 and L_2 so that $L = L_1(t,x) + L_2(t,u)$". (ii) $L(t,x;u) \in C^1([0,1] \times \mathbb{R}^d \times \mathbb{R}^d; [0,\infty))$ and is strictly convex in u. (iii) $L \in C_b^{1,2,0}([0,1] \times \mathbb{R}^d \times B_R)$ for any $R > 0$.
(A.4)'	σ is uniformly nondegenerate, $\sigma_{ij} \in C_b^{1,2}([0,1] \times \mathbb{R}^d)$, $i, j = 1, \cdots, d$. $\xi \in C_b^{1,2}([0,1] \times \mathbb{R}^d; \mathbb{R}^d)$, $c \in C_b^{1,2}([0,1] \times \mathbb{R}^d)$, and for $(t,x,u) \in [0,1] \times \mathbb{R}^d \times \mathbb{R}^d$, $L(t,x;u) = \frac{1}{2}\langle a(t,x)^{-1}(u - \xi(t,x)), u - \xi(t,x)\rangle + c(t,x)$.
(A.5)	$L(t,x;\cdot) \in C^2(\mathbb{R}^d)$ for $(t,x) \in [0,1] \times \mathbb{R}^d$. $D_u^2 L(t,x;u)$ is bounded and uniformly nondegenerate on $[0,1] \times \mathbb{R}^d \times \mathbb{R}^d$.
(A.6)	$\sigma(t,x) = (\sigma^{ij}(t,x))_{i,j=1}^d$, $(t,x) \in [0,1] \times \mathbb{R}^d$, is a $d \times d$-matrix. $a := \sigma\sigma^t$ is uniformly positive definite, bounded, once continuously differentiable and uniformly Hölder continuous. $D_x a$ is bounded and the first derivatives of a are uniformly Hölder continuous in x uniformly in $t \in [0,1]$.
(A.7)	$\xi \in C_b([0,1] \times \mathbb{R}^d; \mathbb{R}^d)$ and is uniformly Hölder continuous in x uniformly in $t \in [0,1]$.
(A.8)	S is a compact metric space.
(A.8)'	S is a σ-compact metric space.
(A.8)"	S is a complete σ-compact metric space.
(A.9)	$q \in C(S \times S; (0,\infty))$.
(A.9.r)'	There exists $C_r > 0$ for which $x \mapsto C_r\|x\|^2 + \log q(x,y)$ and $y \mapsto C_r\|y\|^2 + \log q(x,y)$ are convex on B_r for any y and $x \in B_r$, respectively.
(A.10)	$L : \mathbb{R}^d \longrightarrow [0,\infty)$ is convex and $\liminf_{\|u\| \to \infty} L(u)/\|u\|^2 > 0$.
$V(P_0, P_1)$	$\inf\{E[\int_0^1 L(t, X(t); \beta_X(t,X))dt] \| X \in \mathscr{A}, P^{X(t)} = P_t, t = 0,1\}$
$\mathrm{v}(P_0, P_1)$	$\inf\{\int_{[0,1]\times\mathbb{R}} L(t,x;b(t,x))dt Q_t(dx) \| b \in \mathbf{A}(\{Q_t\}_{0 \le t \le 1}), Q_t = P_t, t = 0, 1\}$
$\tilde{\mathrm{v}}(P_0, P_1)$	$\inf\{\int_{[0,1]\times\mathbb{R}^d\times\mathbb{R}^d} L(t,x;u)\nu(dtdxdu) \| \nu \in \tilde{\mathscr{A}}, \nu_{1,t} = P_t, t = 0,1\}$

$\mathbf{V}(\{P_t\}_{0\le t\le 1})$	$\inf\{E[\int_0^1 L(t, X(t); \beta_X(t, X))dt] \mid X \in \mathscr{A}, P^{X(t)} = P_t, 0 \le t \le 1\}$
$\mathbf{v}(\{P_t\}_{0\le t\le 1})$	$\inf\{\int_{[0,1]\times\mathbb{R}} L(t, x; b(t, x))dt P_t(dx) \mid b \in \mathbf{A}(\{P_t\}_{0\le t\le 1})\}$
$\tilde{\mathbf{v}}(\{P_t\}_{0\le t\le 1})$	$\inf\{\int_{[0,1]\times\mathbb{R}^d\times\mathbb{R}^d} L(t, x; u)\nu(dtdxdu) \mid \nu \in \tilde{\mathscr{A}}, \nu_{1,t} = P_t, 0 \le t \le 1\}$

Chapter 1
Introduction

Abstract Starting from Monge's problem in 1781, the theory of optimal mass transportation (OT for short) has been studied by many authors in many fields of research. Partly as a stochastic analog of the OT, we have been studying the so-called stochastic optimal transportation problem (SOT for short). It is a stochastic optimal control problem with fixed marginal distributions. One of the important purposes is to study the OT in the framework of the SOT. It is also a generalization of Schrödinger's problem and is related to Nelson's stochastic mechanics. We briefly describe the OT and Schrödinger's problems in such a way that one can compare the similarities between them.

1.1 Background

We are interested in the mathematical analysis of phenomena. First, we would like to find a variational problem in which a minimizer describes a phenomenon and then study the phenomenon itself via the variational problem from a mathematical point of view.

To consider a variational problem, we first construct a function S and a set A over which we minimize S. Then we find a condition under which A is not empty. Indeed, the sets under consideration in this monograph can be empty. Our problem can be described as in the following manner:

$$\inf\{S(x)|x \in A\}. \tag{1.1}$$

As a typical example, we consider the distance between two points $x_0, x_1 \in \mathbb{R}^d$ via a variational problem. Let $C([0, 1]; \mathbb{R}^d)$ and $AC([0, 1]; \mathbb{R}^d)$ denote the set of all continuous and absolutely continuous functions from $[0, 1]$ to $\mathbb{R}^d$, respectively.

$$S_{0,1}(x) := \begin{cases} \int_0^1 |\dot{x}(t)|^2 dt, & x \in AC([0, 1]; \mathbb{R}^d), \\ \infty, & x \in C([0, 1]; \mathbb{R}^d) \cap AC([0, 1]; \mathbb{R}^d)^c, \end{cases} \tag{1.2}$$

T. Mikami, *Stochastic Optimal Transportation*, SpringerBriefs in Mathematics,
https://doi.org/10.1007/978-981-16-1754-6_1

where $\dot{x}(t) := dx(t)/dt$. $S_{0,1}(x)$ can be considered a cost to move from $x(0)$ to $x(1)$ along $x = (x(t))_{0\le t\le 1}$. (1.2) means that we do not move along $x \notin AC([0,1];\mathbb{R}^d)$ which is too zigzag!

It seems that we move from one point to another along the line segment which connects the two points since we know that the following holds:

$$|x_1 - x_0|^2 = \inf\left\{S_{0,1}(x) | x(0) = x_0, x(1) = x_1, x \in C([0,1];\mathbb{R}^d)\right\}, \tag{1.3}$$

where the minimizer is

$$x_0 + t(x_1 - x_0), \quad 0 \le t \le 1. \tag{1.4}$$

$S_{0,1}(x)$ and $\{x \in C([0,1];\mathbb{R}^d) | x(0) = x_0, x(1) = x_1\}$ play roles of $S(x)$ and $A(\neq \emptyset)$ in (1.1), respectively.

We show that (1.3)–(1.4) hold. By Schwartz's inequality,

$$|x(1) - x(0)|^2 = \left|\int_0^1 \dot{x}(t)dt\right|^2 \le \int_0^1 |\dot{x}(t)|^2 dt, \quad x \in AC([0,1];\mathbb{R}^d),$$

where the equality holds if and only if $\dot{x}(t)$ is a constant, which implies (1.3)–(1.4).

We consider Euler's equation for the right-hand side of (1.3), that is, the following:

$$\text{the first variation of the right-hand side of (1.3)} = 0.$$

Let $x \in AC([0,1];\mathbb{R}^d)$ be a minimizer of (1.3). Then for any $y \in AC([0,1];\mathbb{R}^d)$ for which $y(0) = y(1) = 0$ and $\delta \in \mathbb{R}$,

$$x(0) + \delta y(0) = x_0, \quad x(1) + \delta y(1) = x_1, \quad S_{0,1}(x + \delta y) \ge S_{0,1}(x),$$

which implies that the Gâteaux derivative vanishes:

$$\lim_{\delta \to 0} \frac{S_{0,1}(x + \delta y) - S_{0,1}(x)}{\delta} = 0, \tag{1.5}$$

provided it exists. That is,

$$\frac{d}{d\delta} \int_0^1 |\dot{x}(t) + \delta \dot{y}(t)|^2 \, dt \Big|_{\delta=0} = 2\int_0^1 \langle \ddot{x}(t), y(t)\rangle dt = 0, \tag{1.6}$$

since $y(0) = y(1) = 0$. Equation (1.6) implies that the following holds:

$$\ddot{x}(t) = 0, \tag{1.7}$$

which implies (1.3)–(1.4).