



Title	A Simple Proof of Duality Theorem for Monge-Kantorovich Problem
Author(s)	Mikami, Toshio
Citation	Hokkaido University Preprint Series in Mathematics, 677, 1-6
Issue Date	2004-12-14
DOI	10.14943/83828
Doc URL	<a href="http://hdl.handle.net/2115/69482">http://hdl.handle.net/2115/69482</a>
Type	bulletin (article)
File Information	pre677.pdf



[Instructions for use](#)

# A Simple Proof of Duality Theorem for Monge-Kantorovich Problem

Toshio Mikami\*  
Hokkaido University

December 14, 2004

## Abstract

We give a simple proof of the duality theorem for the Monge-Kantorovich problem in the Euclidean setting. The selection lemma which is useful in the theory of stochastic optimal controls plays a crucial role.

## 1 Introduction.

Let  $P_0$  and  $P_1$  be Borel probability measures on  $\mathbf{R}^d$  and  $\mathcal{A}(P_0, P_1)$  denote the set of all  $\mu \in \mathcal{M}_1(\mathbf{R}^d \times \mathbf{R}^d)$  for which  $\mu(dx \times \mathbf{R}^d) = P_0(dx)$  and  $\mu(\mathbf{R}^d \times dx) = P_1(dx)$ , where  $\mathcal{M}_1(\mathbf{R}^d \times \mathbf{R}^d)$  denotes the complete separable metric space, with a weak topology, of Borel probability measures on  $\mathbf{R}^d \times \mathbf{R}^d$  (see e.g. [1]). Take also a Borel measurable  $c(\cdot, \cdot) : \mathbf{R}^d \times \mathbf{R}^d \mapsto [0, \infty)$ .

The study of a minimizer of the following  $\mathcal{T}(P_0, P_1)$  is called the Monge-Kantorovich problem which has been studied by many authors and which has been applied to many fields (see [2, 4, 6, 9, 10] and the references therein):

$$\mathcal{T}(P_0, P_1) := \inf \left\{ \int_{\mathbf{R}^d \times \mathbf{R}^d} c(x, y) \mu(dx dy) \mid \mu \in \mathcal{A}(P_0, P_1) \right\}. \quad (1.1)$$

---

\*Department of Mathematics, Hokkaido University, Sapporo 060-0810, Japan; mikami@math.sci.hokudai.ac.jp; Partially supported by the Grant-in-Aid for Scientific Research, No. 15340047, 15340051 and 16654031, JSPS.

The duality theorem for  $\mathcal{T}(P_0, P_1)$  plays a crucial role in the proof of the Monge-Kantorovich problem and has been proved for a wide class of functions  $c(\cdot, \cdot)$  (see [5, 8-10]).

They say that the duality theorem for  $\mathcal{T}(P_0, P_1)$  holds if

$$\mathcal{T}(P_0, P_1) = \sup \left\{ \int_{\mathbf{R}^d} \psi(1, y) P_1(dy) - \int_{\mathbf{R}^d} \psi(0, x) P_0(dx) \right\}, \quad (1.2)$$

where the supremum is taken over all  $\psi(t, \cdot) \in L^1(P_t)$  ( $t = 0, 1$ ) for which

$$\psi(1, y) - \psi(0, x) \leq c(x, y). \quad (1.3)$$

In [7] we obtained a stochastic control version of (1.2)-(1.3) and gave a new approach to  $h$ -path processes for diffusion processes as its application. We also showed that its zero noise limit yields (1.2)-(1.3) (see [8]).

In this paper we give a simple proof for (1.2)-(1.3) since known proofs for (1.2)-(1.3) are complicated. Indeed, Kellerer used a functional version of Choquet's capacitability theorem (see [5] and also [9]) and Villani did a minimax principle (see [10, sect. 1.1]).

Our proof relies on the Legendre duality of a lower semicontinuous convex function of Borel probability measures on  $\mathbf{R}^d$  and on the selection lemma which is useful in the theory of stochastic optimal controls (see the proof of Theorem 2.1).

## 2 A simple proof for Duality Theorem.

We first describe our assumption.

(A.1)  $c \in C(\mathbf{R}^d \times \mathbf{R}^d : [0, \infty))$  and  $c(x, y) \rightarrow \infty$  as  $|y - x| \rightarrow \infty$ , and  $\inf\{c(x, y) | y \in \mathbf{R}^d\}$  is bounded.

(A.2)  $\mathcal{T}(P_0, P_1)$  is finite.

(A.3)  $P_0$  is absolutely continuous with respect to the Lebesgue measure  $dx$ .

We give a simple proof to the following which can be obtained from the known result (see [5] and also [9, 10]).

**Theorem 2.1 (Duality Theorem)** *Suppose that (A.1)-(A.3) hold. Then (1.2)-(1.3) holds.*

(Proof) We prove (1.2), where the supremum is taken over all  $\psi(t, \cdot) \in C_b(\mathbf{R}^d)$  ( $t = 0, 1$ ) for which (1.3) holds. This implies the duality theorem for  $\mathcal{T}(P_0, P_1)$  since (1.2)-(1.3) with “=” replaced by “ $\geq$ ” holds and since  $C_b(\mathbf{R}^d) \subset L^1(P_t)$  ( $t = 0, 1$ ).

The proof is divided into the following (2.1)-(2.3):

$$\mathcal{T}(P_0, P_1) = \sup_{f \in C_b(\mathbf{R}^d)} \left\{ \int_{\mathbf{R}^d} f(y) P_1(dy) - \mathcal{T}_{P_0}^*(f) \right\}, \quad (2.1)$$

where for  $f \in C_b(\mathbf{R}^d)$ ,

$$\mathcal{T}_{P_0}^*(f) := \sup_{P \in \mathcal{M}_1(\mathbf{R}^d)} \left\{ \int_{\mathbf{R}^d} f(y) P(dy) - \mathcal{T}(P_0, P) \right\}.$$

For  $f \in C_b(\mathbf{R}^d)$ ,

$$\varphi(x; f) := \sup_{y \in \mathbf{R}^d} \{f(y) - c(x, y)\} \in C_b(\mathbf{R}^d), \quad (2.2)$$

$$\mathcal{T}_{P_0}^*(f) = \int_{\mathbf{R}^d} \varphi(x; f) P_0(dx). \quad (2.3)$$

We first prove (2.1). We only have to prove  $\mathcal{T}(P_0, \cdot) : \mathcal{M}_1(\mathbf{R}^d) \mapsto [0, \infty)$  is lower semicontinuous and convex. Indeed, this and (A.2) implies (2.1) from [1, Theorem 2.2.15 and Lemma 3.2.3], by putting  $\mathcal{T}(P_0, P) = \infty$  for  $P \notin \mathcal{M}_1(\mathbf{R}^d)$ .

Suppose that  $Q_n \rightarrow Q$  weakly as  $n \rightarrow \infty$ . Then it is easy to see that  $\cup_{n \geq 1} \mathcal{A}(P_0, Q_n)$  is tight in  $\mathcal{M}_1(\mathbf{R}^d)$ . Take  $\mu_n \in \mathcal{A}(P_0, Q_n)$  ( $n \geq 1$ ) for which

$$\mathcal{T}(P_0, Q_n) + \frac{1}{n} > \int_{\mathbf{R}^d \times \mathbf{R}^d} c(x, y) \mu_n(dx dy) \geq \mathcal{T}(P_0, Q_n). \quad (2.4)$$

For any convergent subsequence  $\{\mu_{k(n)}\}_{n \geq 1}$  of  $\{\mu_n\}_{n \geq 1}$  and its weak limit  $\mu_0$ ,  $\mu_0 \in \mathcal{A}(P_0, Q)$  and

$$\liminf_{n \rightarrow \infty} \int_{\mathbf{R}^d \times \mathbf{R}^d} c(x, y) \mu_{k(n)}(dx dy) \geq \int_{\mathbf{R}^d \times \mathbf{R}^d} c(x, y) \mu_0(dx dy) \quad (2.5)$$

since  $c \geq 0$  from (A.1). Hence  $\mathcal{T}(P_0, \cdot) : \mathcal{M}_1(\mathbf{R}^d) \mapsto [0, \infty)$  is lower semicontinuous.  $\mathcal{T}(P_0, \cdot) : \mathcal{M}_1(\mathbf{R}^d) \mapsto [0, \infty)$  is also convex since for any  $P, Q \in \mathcal{M}_1(\mathbf{R}^d)$  and  $\lambda \in (0, 1)$ ,

$$\{\lambda\mu + (1 - \lambda)\nu \mid \mu \in \mathcal{A}(P_0, P), \nu \in \mathcal{A}(P_0, Q)\} \subset \mathcal{A}(P_0, \lambda P + (1 - \lambda)Q).$$

We next prove (2.2). From (A.1), for  $f \in C_b(\mathbf{R}^d)$  and  $x \in \mathbf{R}^d$ ,

$$-\infty < \inf_{y \in \mathbf{R}^d} f(y) - \sup_{y \in \mathbf{R}^d} \{ \inf_{y \in \mathbf{R}^d} c(x, y) \} \leq \varphi(x; f) \leq \sup_{y \in \mathbf{R}^d} f(y) < \infty, \quad (2.6)$$

which implies that  $\varphi(\cdot; f)$  is bounded.

From (A.1), the following set is not empty for any  $x \in \mathbf{R}^d$  and is bounded on every bounded subset of  $\mathbf{R}^d$ :

$$D_x := \{y \in \mathbf{R}^d \mid \varphi(x; f) = f(y) - c(x, y)\}. \quad (2.7)$$

Suppose that  $x_n \rightarrow x$  as  $n \rightarrow \infty$ . Take  $y_n \in D_{x_n}$  and  $y \in D_x$ . Then there exist a convergent subsequence  $\{y_{k(n)}\}_{n \geq 1}$  and  $\tilde{y}$  such that  $y_{k(n)} \rightarrow \tilde{y}$  as  $n \rightarrow \infty$  and such that

$$\begin{aligned} \limsup_{n \rightarrow \infty} \varphi(x_n; f) &= \lim_{n \rightarrow \infty} \{f(y_{k(n)}) - c(x_{k(n)}, y_{k(n)})\} \\ &= f(\tilde{y}) - c(x, \tilde{y}) \leq \varphi(x; f). \end{aligned} \quad (2.8)$$

The following together with (2.8) implies that  $\varphi(\cdot; f) \in C(\mathbf{R}^d)$ :

$$\liminf_{n \rightarrow \infty} \varphi(x_n; f) \geq \lim_{n \rightarrow \infty} \{f(y) - c(x_n, y)\} = \varphi(x; f). \quad (2.9)$$

We prove (2.3) to complete the proof. For  $f \in C_b(\mathbf{R}^d)$ ,

$$\begin{aligned} &\mathcal{T}_{P_0}^*(f) \\ &= \sup_{P \in \mathcal{M}_1(\mathbf{R}^d)} \left\{ \sup \left\{ \int_{\mathbf{R}^d \times \mathbf{R}^d} (f(y) - c(x, y)) \mu(dx dy) \mid \mu \in \mathcal{A}(P_0, P) \right\} \right\} \\ &\leq \int_{\mathbf{R}^d} \varphi(x; f) P_0(dx). \end{aligned} \quad (2.10)$$

(A.1) implies that the set  $\cup_{|x| \leq r} \{x\} \times D_x$  is compact for any  $r > 0$ . Indeed, this set is closed from (2.2) and is bounded as we mentioned in (2.7). Hence

there exists a measurable function  $u : \mathbf{R}^d \mapsto \mathbf{R}^d$  such that  $u(x) \in D_x$ ,  $dx$ -a.e. by the selection lemma (see [3, p. 199]). In particular, from (A.3) and (2.10),

$$\mathcal{T}_{P_0}^*(f) \leq \int_{\mathbf{R}^d \times \mathbf{R}^d} \{f(y) - c(x, y)\} P_0(dx) \delta_{u(x)}(dy) \leq \mathcal{T}_{P_0}^*(f). \quad (2.11)$$

Q.E.D.

## References

- [1] Deuschel JD and Stroock DW, *Large deviations*. Pure and Applied Mathematics Vol. 137 (Academic Press, Inc., Boston, MA, 1989).
- [2] Evans LC, Partial differential equations and Monge-Kantorovich mass transfer, *Current developments in mathematics, 1997 (Cambridge, MA)* (Int. Press, Boston, MA, 1999) 65–126.
- [3] Fleming WH and Rishel RW, *Deterministic and stochastic optimal control* (Springer-Verlag, Berlin, Heidelberg, New York, Tokyo, 1975).
- [4] Gangbo W and McCann RJ, The geometry of optimal transportation, *Acta Math.* **177** (1996) 113-161.
- [5] Kellerer HG, Duality theorems for marginal problems. *Z. Wahrsch. Verw. Gebiete* **67** (1984) 399–432.
- [6] Mikami T, Monge’s problem with a quadratic cost by the zero-noise limit of  $h$ -path processes. *Probab. Theory Related Fields* **129** (2004) 245–260.
- [7] Mikami T and Thieullen M, Duality theorem for stochastic optimal control problem, Hokkaido University preprint series, #652 (2004).
- [8] Mikami T and Thieullen M, Optimal Transportation Problem by Stochastic Optimal Control, in preparation.
- [9] Rachev ST and Rüschendorf L, *Mass transportation problems, Vol. I: Theory, Vol. II: Application* (Springer-Verlag, Berlin, Heidelberg, New York, Tokyo, 1998).

- [10] Villani C, *Topics in Optimal Transportation*. Graduate Studies in Mathematics Vol. 58 (Amer. Math. Soc., Providence, RI, 2003).