Multivariate Monge-Kantorovich Transportation Problem

Yinfang Shen (Corresponding author)

Institute of Mathematics and Statistics, Zhejiang University of Finance and Economics

PO box 310018, Hangzhou, China

Tel: 86-158-8889-6193 E-mail: fsilver@163.com

Received: January 6, 2011 Accepted: January 24, 2011 doi:10.5539/jmr.v3n2p66

The research is supported by Zhejiang Provincial Education Department research projects (Y201016421)

Abstract

Monge-Kantorovich transportation problem in a bounded region of Euclidean space is transferred into a partial differential equations group. And then the explicit formula of the optimal coupling is achieved. The proofs are based on variational method from probability point of view.

Keywords: Monge-Kantorovich problem, Multivariate, Optimal coupling, Partial differential equations group

1. Introduction

The optimal transport problem was first formulated by (Monge, 1781), and concerned finding the optimal way, in the sense of minimal transportation cost of moving a pile of soil from one site to another. This problem was given a modern formulation in the work of (Kantorovich, 1942) and so is now known as the Monge-Kontorovich problem.

This type of problem has appeared in economics, automatic control, transportation, fluid dynamics, statistical physics, shape optimization, expert system, meteorology and financial mathematics. So many mathematicians from different fields are interested in Monge-Kontorovich problem. This classical problem was revived in the mid eighties by the work of (Brenier, 1987, 1991), who characterized the optimal transfer plans in terms of gradients of convex functions, and then get the beautiful relationship with partial differential equations, fluid dynamics, geometry, probability theory and functional analysis.

Several (important) authors have been involved in literature of the optimal transportation theory, with remarkable applications, we mention here (Caffarelli, 1996, 2002), (Caffarelli, Feldman & McCann, 2002) and (Evans & Gangbo, 1999, 1996). We recommend Caffarelli's address to ICM2002 (Caffarelli, 2002), Ambrosio's address to ICM2002 (Ambrosio, 2002), Trudinger's invited lecture to ICM2006 (Trudinger, 2006), (Trudinger & Xujia, 2001) and (Evans & Gangbo, 1999) for major references from PDE point of view. We also recommend Rachev and Rüschendorf's book (Rachev & Rüschendorf, 1998), (Villani, 2002) for a major reference from probability point of view.

Suppose that we are given two probability distributions P and $\tilde{P}$ on R^2. A 4−dimensional random vector (X, Y) with P and $\tilde{P}$ as the marginal distributions is called a coupling of this pair $(P, \tilde{P})$. The minimum of the coupling distance $\|X - Y\|_{L_2}$ among all such possible couplings is called Kantorovich-Rubinstein-Wasserstein (abbr. KRW) L_2−distance between P and $\tilde{P}$, and the corresponding coupling is called the optimal coupling. It has important applications in both probability theory and mass transfer problems. In R^1, Kantorovich-Rubinstein-Wasserstein L_2−distance is just given by

$$\sqrt{\int_0^1 |F^{-1}(t) - \tilde{F}^{-1}(t)|^2 dt}, \tag{1}$$

where F and $\tilde{F}$ are distribution functions of P and $\tilde{P}$ respectively, $F^{-1}(t)$ and $\tilde{F}^{-1}(t)$, $(0 \le t \le 1)$ are their right inverses.

In R^2, see the refined result due to (Weian, 2000) and (Yinfang & Weian, 2010) in bounded real plane, and (Givens & Short, 1984) between normal random vectors. In our recent paper (Yinfang & Weian, 2010), we transformed the Monge-Kantorovich problem as $p = 2$ in a bounded region of Euclidean plane into a Dirichlet boundary problem. Denote by $\mathcal{H}$ the set of all density functions $p(x, y)$ on $[0, 1] \times [0, 1]$ such that $f_1(x) = \int_0^1 p(x, y)dy$ and $\tilde{f}_2(y) = \int_0^1 p(x, y)dx$. Let Z be the coordinate variable $Z(x, y) = (x, y)$ which has $p(x, y)$ as its density function. $\hat{X}(x, y) = (x, G(2, x, \int_0^y \frac{p(x,v)}{f_1(x)} dv))$ and $\hat{Y}(x, y) = (\tilde{G}(1, \int_0^x \frac{p(u,y)}{\tilde{f}_2(y)} du, y), y)$ have given density f and $\tilde{f}$, and $Z = (\hat{X}_1, \hat{Y}_2)$. Furthermore, if (X, Y) is the optimal coupling, then the above realization $(\hat{X}, \hat{Y})$ have the same optimal joint distribution. Thus we have

$$\begin{aligned} E|X - Y|^2 &= \int_0^1 \int_0^1 |t - G(2, x, \int_0^t \frac{p(x, v)}{f_1(x)} dv)|^2 p(x, t)dt\,dx \\ &+ \int_0^1 \int_0^1 |s - \tilde{G}(1, \int_0^s \frac{p(u, y)}{\tilde{f}_2(y)} du, y)|^2 p(s, y)ds\,dy. \end{aligned} \tag{2}$$

Lemma 1. *If $p(x, y)$ minimizes (2), then*

$$[G(2, x, \frac{1}{f_1(x)} H'_x(x, y))]'_x + [\tilde{G}(1, \frac{1}{\tilde{f}_2(y)} H'_y(x, y), y)]'_y = 0, \tag{3}$$

which is a quasi-linear elliptic equation with unknown $H(x, y)$, as $\tilde{G}_x(1, x, y) \geq 0$ and $G_y(2, x, y) \geq 0$. Moreover, the optimal mapping is given by $(\hat{Y}(\hat{X}_1, \hat{X}^{-1}), \hat{X}^{-1})$.

Here in this paper, we transfer Monge-Kantorovich problem in a bounded region of Euclidean space into a partial differential equations group. The proofs are based on variational method from probability point of view. Also we get the explicit formula of optimal coupling.

2. Main results

Without losing generality, we may consider two probability measures P and Q on $[0, 1]^n$, and n is even. Let X and Y be two random vectors defined on a same probability space with P and Q as their individual laws. Suppose that the coupling $X = (X_1, X_2, \ldots, X_n)$, $Y = (Y_1, Y_2, \ldots, Y_n)$ gives the desired Kantorovich-Wasserstein L_2-distance. If we denote

$Z^{(i)} = (Y_1, Y_2, \ldots, Y_i, X_{i+1}, \ldots, X_n) (1 \leq i \leq n-1)$, then

$$E|X - Y|^2 = E|Z^{(1)} - X|^2 + \sum_{i=1}^{n-2} E|Z^{(i+1)} - Z^{(i)}|^2 + E|Y - Z^{(n-1)}|^2. \tag{4}$$

So it is sufficient to find the distribution $H^{(i)} (1 \leq i \leq n-1)$ of $Z^{(i)}$ which are also supported in $[0, 1]^n$, where $h^{(i)}$ are their density functions, respectively. We assume further their density functions $f(x_1, x_2, \ldots, x_n)$ and $\tilde{f}(y_1, y_2, \ldots, y_n)$ are smooth and strictly positive on their domains. Denote the marginal densities as follows,

$$f_{2,3,\ldots,n}(x_2, x_3, \ldots, x_n) = \int_0^1 f(x_1, x_2, x_3, \ldots, x_n) dx_1,$$

$$h^{(1)}_{2,3,\ldots,n}(x_2, x_3, \ldots, x_n) = \int_0^1 h^{(1)}(y_1, x_2, x_3, \ldots, x_n) dy_1,$$

$$h^{(1)}_{1,3,\ldots,n}(y_1, x_3, \ldots, x_n) = \int_0^1 h^{(1)}(y_1, x_2, x_3, \ldots, x_n) dx_2,$$

$$h^{(n-1)}_{1,\ldots,n-2,n}(y_1, \ldots, y_{n-2}, x_n) = \int_0^1 h^{(n-1)}(y_1, \ldots, y_{n-2}, y_{n-1}, x_n) dy_{n-1},$$

$$h^{(n-1)}_{1,\ldots,n-1}(y_1, \ldots, y_{n-1}) = \int_0^1 h^{(n-1)}(y_1, \ldots, y_{n-1}, x_n) dx_n,$$

$$\tilde{f}_{1,\ldots,n-1}(y_1, \ldots, y_{n-1}) = \int_0^1 \tilde{f}(y_1, \ldots, y_{n-1}, y_n) dy_n,$$

when $2 \leq i \leq n-2$,

$$h^{(i)}_{1,\ldots,i-1,i+1,\ldots,n}(y_1, \ldots, y_{i-1}, x_{i+1}, \ldots, x_n) = \int_0^1 h^{(i)}(y_1, \ldots, y_{i-1}, y_i, x_{i+1}, \ldots, x_n) dy_i,$$

$$h^{(i)}_{1,\ldots,i,i+2,\ldots,n}(y_1, \ldots, y_i, x_{i+2}, \ldots, x_n) = \int_0^1 h^{(i)}(y_1, \ldots, y_i, x_{i+1}, x_{i+2}, \ldots, x_n) dx_{i+1}.$$

And denote the conditional distribution functions as follows,

$$F_{1|2,3,\ldots,n}(x_1|x_2, x_3, \ldots, x_n) = \frac{1}{f_{2,3,\ldots,n}(x_2, x_3, \ldots, x_n)} \int_0^{x_1} f(u_1, x_2, x_3, \ldots, x_n) du_1,$$

$$H^{(1)}_{1|2,3,\ldots,n}(y_1|x_2, x_3, \ldots, x_n) = \int_0^{y_1} \frac{h^{(1)}(u_1, x_2, x_3, \ldots, x_n)}{f_{2,\ldots,n}(x_2, x_3, \ldots, x_n)} du_1,$$

$$H^{(1)}_{2|1,3,\ldots,n}(x_2|y_1, x_3, \ldots, x_n) = \frac{1}{h^{(1)}_{1,3,\ldots,n}(y_1, x_3, \ldots, x_n)} \int_0^{x_2} h^{(1)}(y_1, u_2, x_3, \ldots, x_n) du_2,$$

$$H^{(n-1)}_{n-1|1,\ldots,n-2,n}(y_{n-1}|y_1, \ldots, y_{n-2}, x_n) = \frac{\int_0^{y_{n-1}} h^{(n-1)}(y_1, \ldots, y_{n-2}, u_{n-1}, x_n) du_{n-1}}{h^{(n-1)}_{1,\ldots,n-2,n}(y_1, \ldots, y_{n-2}, x_n)},$$

$$H^{(n-1)}_{n|1,\dots,n-1}(y_n|y_1,\dots,y_{n-1}) = \frac{\int_0^{y_n} h^{(n-1)}(y_1,\dots,y_{n-1},u_n)du_n}{h^{(n-1)}_{1,\dots,n-1}(y_1,\dots,y_{n-1})},$$

$$\tilde{F}_{n|1,\dots,n-1}(y_n|y_1,\dots,y_{n-1}) = \frac{\int_0^{y_n} \tilde{f}(y_1,\dots,y_{n-1},u_n)du_n}{\tilde{f}_{1,\dots,n-1}(y_1,\dots,y_{n-1})},$$

when $2 \le i \le n-2$,

$$H^{(i)}_{i|1,\dots,i-1,i+1,\dots,n}(y_i|y_1,\dots,y_{i-1},x_{i+1},\dots,x_n) = \frac{\int_0^{y_i} h^{(i)}(y_1,\dots,y_{i-1},u_i,x_{i+1},\dots,x_n)du_i}{h^{(i)}_{1,\dots,i-1,i+1,\dots,n}(y_1,\dots,y_{i-1},x_{i+1},\dots,x_n)},$$

$$H^{(i)}_{i+1|1,\dots,i,i+2,\dots,n}(x_{i+1}|y_1,\dots,y_i,x_{i+2},\dots,x_n) = \frac{\int_0^{x_{i+1}} h^{(i)}(y_1,\dots,y_i,u_{i+1},x_{i+2},\dots,x_n)du_{i+1}}{h^{(i)}_{1,\dots,i,i+2,\dots,n}(y_1,\dots,y_i,x_{i+2},\dots,x_n)},$$

which are strictly increasing with respect to their first argument, so their inverse functions with respect to their first arguments exist.

Firstly, we have

Lemma 2.

$$\begin{aligned}&E|X-Z^{(1)}|^2\\ =\;& \int_0^1\cdots\int_0^1 |y_1 - F^{-1}_{1|2,\dots,n}(\int_0^{y_1}\frac{h^{(1)}(u_1,x_2,\dots,x_n)}{f_{2,\dots,n}(x_2,\dots,x_n)}du_1|x_2,\dots,x_n)|^2\\ &h^{(1)}(y_1,x_2,\dots,x_n)dy_1dx_2\dots dx_n,\end{aligned} \tag{5}$$

$$\begin{aligned}&E|Y-Z^{(n-1)}|^2\\ &=\int_0^1\int_0^1\cdots\int_0^1\int_0^1 |x_n-\tilde{F}^{-1}_{n|1,2,\dots,n-1}(\int_0^{x_n}\frac{h^{(n-1)}(y_1,y_2,\dots,y_{n-1},u_n)}{\tilde{f}_{1,2,\dots,n-1}(y_1,y_2,\dots,y_{n-1})}du_n|\\ &y_1,y_2,\dots,y_{n-1})|^2h^{(n-1)}(y_1,y_2,\dots,y_{n-1},x_n)dy_1dy_2\dots dy_{n-1}dx_n.\end{aligned} \tag{6}$$

Similarly, when $1 \le i \le n-2$,

$$\begin{aligned}&E|Z^{(i+1)}-Z^{(i)}|^2\\ &=\int_0^1\cdots\int_0^1\int_0^1\cdots\int_0^1 |y_{i+1}-(H^{(i)}_{i+1|1,\dots,i,i+2,\dots,n})^{-1}\\ &(\int_0^{y_{i+1}}\frac{h^{(i+1)}(y_1,\dots,y_i,u_{i+1},x_{i+2},\dots,x_n)}{h^{(i)}_{1,\dots,i,i+2,\dots,n}(y_1,\dots,y_i,x_{i+2},\dots,x_n)}du_{i+1}|y_1,\dots,y_i,x_{i+2},\dots,x_n)|^2\\ &h^{(i+1)}(y_1,\dots,y_{i+1},x_{i+2},\dots,x_n)dy_1\dots dy_{i+1}dx_{i+2}\dots dx_n,\end{aligned} \tag{7}$$

or

$$\begin{aligned}&E|Z^{(i+1)}-Z^{(i)}|^2\\ &=\int_0^1\cdots\int_0^1\int_0^1\cdots\int_0^1 |x_{i+1}-(H^{(i+1)}_{i+1|1,\dots,i,i+2,\dots,n})^{-1}\\ &(\int_0^{x_{i+1}}\frac{h^{(i)}(y_1,\dots,y_i,u_{i+1},x_{i+2},\dots,x_n)}{h^{(i+1)}_{1,\dots,i,i+2,\dots,n}(y_1,\dots,y_i,x_{i+2},\dots,x_n)}du_{i+1}|y_1,\dots,y_i,x_{i+2},\dots,x_n)|^2\\ &h^{(i)}(y_1,\dots,y_i,x_{i+1},\dots,x_n)dy_1\dots dy_idx_{i+1}\dots dx_n.\end{aligned} \tag{8}$$

Be connected with (4), then we get the exact formula of $E|X-Y|^2$.

So we just need to look for one pair of density functions $(h^{(1)}, h^{(2)}, \dots, h^{(n-1)})$ minimizes (4).

Further by some computations, we transfer it into a Dirichlet boundary problem associated to a partial differential equation system as follows.

Theorem 1. *If* $(h^{(1)}, h^{(2)}, \ldots, h^{(n-1)})$ *minimize* $E|X-Y|^2$, *then* $\forall 0 < x_1, x_2, \ldots, x_n < 1$, *we have*

$$\frac{\partial^{n-1}}{\partial x_2 \ldots \partial x_n}[F^{-1}_{1|2,\ldots,n}(\int_0^{x_1} \frac{h^{(1)}(u_1, x_2, \ldots, x_n)}{f_{2,\ldots,n}(x_2, \ldots, x_n)} du_1 | x_2, \ldots, x_n)] + \frac{\partial^{n-1}}{\partial x_1 \partial x_3 \ldots \partial x_n}[(H^{(2)}_{2|1,3,\ldots,n})^{-1}(\int_0^{x_2} \frac{h^{(1)}(x_1, u_2, x_3, \ldots, x_n)}{h^{(2)}_{1,3,\ldots,n}(x_1, x_3, \ldots, x_n)} du_2 | x_1, x_3, \ldots, x_n)] = 0, \tag{9}$$

$$\frac{\partial^{n-1}}{\partial x_1 \ldots \partial x_{n-2} \partial x_n}[(H^{(n-2)}_{n-1|1,2,\ldots,n-2,n})^{-1}(\int_0^{x_{n-1}} \frac{h^{(n-1)}(x_1, x_2, \ldots, x_{n-2}, u_{n-1}, x_n)}{h^{(n-2)}_{1,2,\ldots,n-2,n}(x_1, x_2, \ldots, x_{n-2}, x_n)} du_{n-1} | x_1, \ldots, x_{n-2}, x_n)] + \frac{\partial^{n-1}}{\partial x_1 \ldots \partial x_{n-1}}[\tilde{F}^{-1}_{n|1,2,\ldots,n-1}(\int_0^{x_n} \frac{h^{(n-1)}(x_1, x_2, \ldots, x_{n-1}, u_n)}{\tilde{f}_{1,2,\ldots,n-1}(x_1, x_2, \ldots, x_{n-1})} du_n | x_1, x_2, \ldots, x_{n-1})] = 0. \tag{10}$$

As $2 \le i \le n-2$,

$$\frac{\partial^{n-1}}{\partial x_1 \ldots \partial x_{i-1} \partial x_{i+1} \ldots \partial x_n}[(H^{(i-1)}_{i|1,\ldots,i-1,i+1,\ldots,n})^{-1} (\int_0^{x_i} \frac{h^{(i)}(x_1, \ldots, x_{i-1}, u_i, x_{i+1}, \ldots, x_n)}{h^{(i-1)}_{1,\ldots,i-1,i+1,\ldots,n}(x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n)} du_i | x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n)] + \frac{\partial^{n-1}}{\partial x_1 \ldots \partial x_i \partial x_{i+2} \ldots \partial x_n}[(H^{(i+1)}_{i+1|1,\ldots,i,i+2,\ldots,n})^{-1} (\int_0^{x_{i+1}} \frac{h^{(i)}(x_1, \ldots, x_i, u_{i+1}, x_{i+2}, \ldots, x_n)}{h^{(i+1)}_{1,\ldots,i,i+2,\ldots,n}(x_1, \ldots, x_i, x_{i+2}, \ldots, x_n)} du_{i+1} | x_1, \ldots, x_i, x_{i+2}, \ldots, x_n)] = 0. \tag{11}$$

And we may easily obtain that the marginal distributions of $h^{(1)}, h^{(2)}, \ldots, h^{(n-1)}$ satisfy some boundary conditions in the region of $[0,1]^n$.

Also we may easily deduce that the explicit formula of the optimal coupling.

Theorem 2.

$$Y_1 = (H^{(1)}_{1|2,3,\ldots,n})^{-1}(\frac{1}{f_{2,3,\ldots,n}(X_2, X_3, \ldots, X_n)} \int_0^{X_1} f(u_1, X_2, X_3, \ldots, X_n) du_1),$$
$$Y_n = \tilde{F}^{-1}_{n|1,2,\ldots,n-1}(\int_0^{X_n} \frac{h^{(n-1)}(Y_1, Y_2, \ldots, Y_{n-1}, u_n)}{\tilde{f}_{1,2,\ldots,n-1}(Y_1, Y_2, \ldots, Y_{n-1})} du_n),$$

when $1 \le i \le n-2$,

$$Y_{i+1} = (H^{(i+1)}_{i+1|1,\ldots,i,i+2,\ldots,\ldots,n})^{-1}(\int_0^{X_{i+1}} \frac{h^{(i)}(Y_1, \ldots, Y_i, u_{i+1}, X_{i+2}, \ldots, X_n)}{h^{(i+1)}_{1,\ldots,i,i+2,\ldots,n}(Y_1, \ldots, Y_i, X_{i+2}, \ldots, X_n)} du_{i+1}).$$

3. Proof of Main Results

Proof of Lemma 2: Let's take (5) for example. In fact, by some properties of conditional expectation and the formula of Wasserstein distance in one dimension, see (Rachev & Rüschendorf, 1998, theorem 3.1.1, theorem 3.1.2, p. 107-109), (Villani, 2002, theorem 2.18, p. 74-78).

$$
\begin{aligned}
E|X - Z^{(1)}|^2 &= E\{E[(X_1 - Y_1)^2|X_2, \ldots, X_n]\} \\
&= \int_0^1 \ldots \int_0^1 E[(X_1 - Y_1)^2|X_2 = x_2, \ldots, X_n = x_n] \\
&\quad h^{(1)}_{2,\ldots,n}(x_2, \ldots, x_n)dx_2 \ldots dx_n \\
&= \int_0^1 \int_0^1 \ldots \int_0^1 [F^{-1}_{1|2,\ldots,n}(u|x_2, \ldots, x_n) - (H^{(1)}_{1|2,\ldots,n})^{-1}(u|x_2, \ldots, x_n)]^2 \\
&\quad h^{(1)}_{2,\ldots,n}(x_2, \ldots, x_n)dudx_2 \ldots dx_n \\
&= \int_0^1 \int_0^1 \ldots \int_0^1 [F^{-1}_{1|2,\ldots,n}(H^{(1)}_{1|2,\ldots,n}(y_1|x_2, \ldots, x_n) - y_1]^2 (H^{(1)}_{1|2,\ldots,n}(y_1|x_2, \ldots, x_n))' \\
&\quad h^{(1)}_{2,\ldots,n}(x_2, \ldots, x_n)dy_1dx_2 \ldots dx_n \\
&= \int_0^1 \int_0^1 \ldots \int_0^1 [F^{-1}_{1|2,\ldots,n}(H^{(1)}_{1|2,\ldots,n}y_1|x_2, \ldots, x_n) - y_1]^2 \\
&\quad h^{(1)}(y_1, x_2, \ldots, x_n)dy_1dx_2 \ldots dx_n,
\end{aligned}
$$

thus we can get (5) at once.

Proof of Theorem 1: Take (9) as an example, for all $0 < a^{(i)}_1 < a^{(i)}_1 + \epsilon < a^{(i)}_2 < a^{(i)}_2 + \epsilon < 1, 1 \le i \le n, \epsilon, \delta(> 0)$ is small enough. Given $j_i = 1, 2, i = 1, 2, \ldots, n$ and denote as

$$
I_{i1} = I_{[}a^{(i)}_1, a^{(i)}_1 + \epsilon], \qquad I_{i2} = I_{[}a^{(i)}_2, a^{(i)}_2 + \epsilon],
$$

$$
\tilde{\xi}(u_1, u_2, \ldots, u_n) = \sum_{k=\sharp\{i;j_i=2,i=1,2,\ldots,n\}} (-1)^k I_{1j_1} \times I_{2j_2} \times \ldots \times I_{nj_n}. \tag{12}
$$

Thus if $(h^{(1)}, h^{(2)}, \ldots, h^{(n-1)})$ is the optimal solution, then

$$
\begin{aligned}
0 \le\ & \frac{1}{\epsilon^n}\{\int_0^1 \int_0^1 \ldots \int_0^1 |u_1 - F^{-1}_{1|2,\ldots,n}(\int_0^{u_1} \frac{h^{(1)}(s_1, u_2, \ldots, u_n) + \delta\tilde{\xi}(s_1, u_2, \ldots, u_n)}{f_{2,\ldots,n}(u_2, \ldots, u_n)} du_1 \\
& |u_2, \ldots, u_n)|^2 (h^{(1)}(u_1, u_2, \ldots, u_n) + \delta\tilde{\xi}(u_1, u_2, \ldots, u_n))du_1du_2 \ldots du_n \\
+\ & \int_0^1 \int_0^1 \int_0^1 \ldots \int_0^1 |u_2 - (H^{(2)}_{2|1,3,\ldots,n})^{-1}(\int_0^{u_2} \frac{h^{(1)}(u_1, s_2, u_3, \ldots, u_n) + \delta\tilde{\xi}(u_1, s_2, \ldots, u_n)}{h_{1,3,\ldots,n}(u_1, u_3, \ldots, u_n)} du_2 \\
& |u_1, u_3, \ldots, u_n)|^2 (h^{(1)}(u_1, u_2, u_3, \ldots, u_n) + \delta\tilde{\xi}(u_1, u_2, \ldots, u_n))du_1du_2 \ldots du_n \\
-\ & \int_0^1 \int_0^1 \ldots \int_0^1 |u_1 - F^{-1}_{1|2,\ldots,n}(\int_0^{u_1} \frac{h^{(1)}(s_1, u_2, \ldots, u_n)}{f_{2,\ldots,n}(u_2, \ldots, u_n)} du_1 \\
& |u_2, \ldots, u_n)|^2 h^{(1)}(u_1, u_2, \ldots, u_n)du_1du_2 \ldots du_n \\
-\ & \int_0^1 \int_0^1 \int_0^1 \ldots \int_0^1 |u_2 - (H^{(2)}_{2|1,3,\ldots,n})^{-1}(\int_0^{u_2} \frac{h^{(1)}(u_1, s_2, u_3, \ldots, u_n)}{h_{1,3,\ldots,n}(u_1, u_3, \ldots, u_n)} du_2 \\
& |u_1, u_3, \ldots, u_n)|^2 h^{(1)}(u_1, u_2, u_3, \ldots, u_n)du_1du_2 \ldots du_n\}.
\end{aligned}
$$

As $\epsilon \to 0$,

$$
\begin{aligned}
0 \leq \quad & 2\delta \sum_{k=\sharp\{i;j_i=2,i=2,\dots,n\}} (-1)^k \int_{a_1^{(1)}}^{a_2^{(1)}} (u_1 - F^{-1}_{1|2,\dots,n}(\int_0^{u_1} \frac{h^{(1)}(s_1, a^{(2)}_{j_2}, \dots, a^{(n)}_{j_n})}{f_{2,\dots,n}(a^{(2)}_{j_2}, \dots, a^{(n)}_{j_n})} ds_1 \\
& |a^{(2)}_{j2}, \dots, a^{(n)}_{j_n}))(F^{-1}_{1|2,\dots,n})'_1(\int_0^{u_1} \frac{h^{(1)}(s_1, a^{(2)}_{j_2}, \dots, a^{(n)}_{j_n})}{f_{2,\dots,n}(a^{(2)}_{j_2}, \dots, a^{(n)}_{j_n})} ds_1 | a^{(2)}_{j_2}, \dots, a^{(n)}_{j_n}) \\
& \frac{h^{(1)}(u_1, a^{(2)}_{j_2}, \dots, a^{(n)}_{j_n})}{f_{2,\dots,n}(a^{(2)}_{j_2}, \dots, a^{(n)}_{j_n})} du_1 \\
& +2\delta \sum_{k=\sharp\{i;j_i=2,i=1,3,\dots,n\}} (-1)^k \int_{a_1^{(2)}}^{a_2^{(2)}} (u_2 - (H^{(2)}_{2|1,3,\dots,n})^{-1}(\int_0^{u_2} \frac{h^{(1)}(a^{(1)}_{j_1}, s_2, a^{(3)}_{j_3}, \dots, a^{(n)}_{j_n})}{h^{(2)}_{1,3,\dots,n}(a^{(1)}_{j_1}, a^{(3)}_{j_3}, \dots, a^{(n)}_{j_n})} ds_2 \\
& |a^{(1)}_{j_1}, a^{(3)}_{j_3}, \dots, a^{(n)}_{j_n}))((H^{(2)}_{2|1,3,\dots,n})^{-1})'_2(\int_0^{u_2} \frac{h^{(1)}(a^{(1)}_{j_1}, s_2, a^{(3)}_{j_3}, \dots, a^{(n)}_{j_n})}{h^{(2)}_{1,3,\dots,n}(a^{(1)}_{j_1}, a^{(3)}_{j_3}, \dots, a^{(n)}_{j_n})} ds_2 | a^{(1)}_{j_1}, a^{(3)}_{j_3}, \dots, a^{(n)}_{j_n}) \\
& \frac{h^{(1)}(a^{(1)}_{j_1}, u_2, a^{(3)}_{j_3}, \dots, a^{(n)}_{j_n})}{h^{(1)}_{1,3,\dots,n}(a^{(1)}_{j_1}, a^{(3)}_{j_3}, \dots, a^{(n)}_{j_n})} du_2 \\
& +\delta \sum_{k=\sharp\{i;j_i=2,i=2,\dots,n\}} (-1)^k |a_1^{(1)} - F^{-1}_{1|2,\dots,n}(\int_0^{a_1^{(1)}} \frac{h^{(1)}(u_1, a^{(2)}_{j_2}, \dots, a^{(n)}_{j_n})}{f_{2,\dots,n}(a^{(2)}_{j_2}, \dots, a^{(n)}_{j_n})} du_1 | a^{(2)}_{j_2}, \dots, a^{(n)}_{j_n})|^2 \\
& -\delta \sum_{k=\sharp\{i;j_i=2,i=2,\dots,n\}} (-1)^k |a_2^{(1)} - F^{-1}_{1|2,\dots,n}(\int_0^{a_2^{(1)}} \frac{h^{(1)}(u_1, a^{(2)}_{j_2}, \dots, a^{(n)}_{j_n})}{f_{2,\dots,n}(a^{(2)}_{j_2}, \dots, a^{(n)}_{j_n})} du_1 | a^{(2)}_{j_2}, \dots, a^{(n)}_{j_n})|^2 \\
& +\delta \sum_{k=\sharp\{i;j_i=2,i=1,3,\dots,n\}} (-1)^k |a_1^{(2)} - (H^{(2)}_{2|1,3,\dots,n})^{-1}(\int_0^{a_1^{(2)}} \frac{h^{(1)}(a^{(1)}_{j_1}, u_2, a^{(3)}_{j_3}, \dots, a^{(n)}_{j_n})}{h^{(2)}_{1,3,\dots,n}(a^{(1)}_{j_1}, a^{(3)}_{j_3}, \dots, a^{(n)}_{j_n})} du_2 | \\
& a^{(1)}_{j_1}, a^{(3)}_{j_3}, \dots, a^{(n)}_{j_n})|^2 \\
& -\delta \sum_{k=\sharp\{i;j_i=2,i=1,3,\dots,n\}} (-1)^k |a_2^{(2)} - (H^{(2)}_{2|1,3,\dots,n})^{-1}(\int_0^{a_2^{(2)}} \frac{h^{(1)}(a^{(1)}_{j_1}, u_2, a^{(3)}_{j_3}, \dots, a^{(n)}_{j_n})}{h^{(2)}_{1,3,\dots,n}(a^{(1)}_{j_1}, a^{(3)}_{j_3}, \dots, a^{(n)}_{j_n})} du_2 | \\
& a^{(1)}_{j_1}, a^{(3)}_{j_3}, \dots, a^{(n)}_{j_n})|^2. \qquad (13)
\end{aligned}
$$

Multiply by $\frac{1}{(a_2^{(1)}-a_1^{(1)})(a_2^{(2)}-a_1^{(2)})\dots(a_2^{(n)}-a_1^{(n)})}$, let $(a_2^{(1)}-a_1^{(1)})(a_2^{(2)}-a_1^{(2)})\dots(a_2^{(n)}-a_1^{(n)}) \to 0$, then

$$\frac{\partial^n}{\partial x_1 \partial x_2 \dots \partial x_n} \tilde{M}(x_1, x_2, \dots, x_n) \geq 0, \qquad (14)$$

where

$$
\begin{aligned}
& \tilde{M}(x_1, x_2, \ldots, x_n) \\
= \; & -2\int_0^{x_1}(u_1 - F^{-1}_{1|2,\ldots,n}(\int_0^{u_1}\frac{h^{(1)}(s_1, x_2, \ldots, x_n)}{f_{2,\ldots,n}(x_2, \ldots, x_n)}ds_1|x_2, \ldots, x_n)) \\
& (F^{-1}_{1|2,\ldots,n})'_1(\int_0^{u_1}\frac{h^{(1)}(s_1, x_2, \ldots, x_n)}{f_{2,\ldots,n}(x_2, \ldots, x_n)}ds_1|x_2, \ldots, x_n)\frac{h^{(1)}(u_1, x_2, \ldots, x_n)}{f_{2,\ldots,n}(x_2, \ldots, x_n)}du_1 \\
& -2\int_0^{x_2}(u_2 - (H^{(2)}_{2|1,3,\ldots,n})^{-1}(\int_0^{u_2}\frac{h^{(1)}(x_1, s_2, x_3, \ldots, x_n)}{f_{1,3,\ldots,n}(x_1, x_3, \ldots, x_n)}ds_2|x_1, x_3, \ldots, x_n)) \\
& ((H^{(2)}_{2|1,3,\ldots,n})^{-1})'_2(\int_0^{u_2}\frac{h^{(1)}(x_1, s_2, x_3, \ldots, x_n)}{h^{(2)}_{1,3,\ldots,n}(x_1, x_3, \ldots, x_n)}ds_2|x_1, x_3, \ldots, x_n))\frac{h^{(1)}(x_1, u_2, x_3, \ldots, x_n)}{h^{(2)}_{1,3,\ldots,n}(x_1, x_3, \ldots, x_n)}du_2 \\
& +|x_1 - (F_{1|2,\ldots,n})^{-1}(\int_0^{x_1}\frac{h^{(1)}(u_1, x_2, \ldots, x_n)}{f_{2,\ldots,n}(x_2, \ldots, x_n)}du_1|x_2, \ldots, x_n)|^2 \\
& +|x_2 - (H^{(2)}_{2|1,3,\ldots,n})^{-1}(\int_0^{x_2}\frac{h^{(1)}(x_1, u_2, x_3, \ldots, x_n)}{h^{(2)}_{1,3,\ldots,n}(x_1, x_3, \ldots, x_n)}du_2|x_1, x_3, \ldots, x_n)|^2 \\
= \; & 2\int_0^{x_1}(u_1 - F^{-1}_{1|2,\ldots,n}(\int_0^{u_1}\frac{h^{(1)}(s_1, x_2, \ldots, x_n)}{f_{2,\ldots,n}(x_2, \ldots, x_n)}ds_1|x_2, \ldots, x_n))du_1 \\
& +\delta\sum_{k=\sharp\{i;j_i=2,i=1,3,\ldots,n\}}(-1)^k|a^{(2)}_1 - (H^{(2)}_{2|1,3,\ldots,n})^{-1}(\int_0^{a^{(2)}_1}\frac{h^{(1)}(a^{(1)}_{j_1}, u_2, a^{(3)}_{j_3}, \ldots, a^{(n)}_{j_n})}{h^{(2)}_{1,3,\ldots,n}(a^{(1)}_{j_1}, a^{(3)}_{j_3}, \ldots, a^{(n)}_{j_n})}du_2| \\
& a^{(1)}_{j_1}, a^{(3)}_{j_3}, \ldots, a^{(n)}_{j_n})|^2 \\
& +2\sum_{k=\sharp\{i;y_i=0,i=1,3,\ldots,n\}}(-1)^k\int_0^{x_2}(H^{(2)}_{2|1,3,\ldots,n})^{-1}(\int_0^{u_2}\frac{h^{(1)}(y_1, s_2, y_3, \ldots, y_n)}{h^{(2)}_{1,3,\ldots,n}(y_1, y_3, \ldots, y_n)}ds_2 \\
& |y_1, y_3, \ldots, y_n)du_2 \\
& \quad +(-1)^{n-1}\int_0^{x_n}\cdots\int_0^{x_2}\int_0^{x_1}\{\frac{\partial^{n-1}}{\partial t_2\ldots\partial t_n}[F^{-1}_{1|2,\ldots,n}(\int_0^{t_1}\frac{h^{(1)}(s_1, t_2, \ldots, t_n)}{f_{2,\ldots,n}(t_2, \ldots, n)}ds_1|t_2\ldots, t_n)] \\
+ \; & \frac{\partial^{n-1}}{\partial t_1\partial t_3\ldots\partial t_n}[(H^{(2)}_{2|1,3,\ldots,n})^{-1}(\int_0^{t_2}\frac{h^{(1)}(t_1, s_2, t_3, \ldots, t_n)}{h^{(2)}_{1,3,\ldots,n}(t_1, t_3, \ldots, t_n)}ds_2|t_1, t_3, \ldots, t_n)]\}dt_1dt_2\ldots dt_n,
\end{aligned}
\tag{15}
$$

which $y_i = 0, x_i, 1 \le i \le n$. If we replace $(h^{(1)} + \delta\tilde{\xi}, h^{(2)}, \ldots, h^{(n)})$ by $(h^{(1)} - \delta\tilde{\xi}, h^{(2)}, \ldots, h^{(n)})$, we may get

$$\frac{\partial^n}{\partial x_1\partial x_2\ldots\partial x_n}\tilde{M}(x_1, x_2, \ldots, x_n) \le 0, \tag{16}$$

therefore $0 < x_1, x_2, \ldots, x_n < 1$,

$$\frac{\partial^n}{\partial x_1\partial x_2\ldots\partial x_n}\tilde{M}(x_1, x_2, \ldots, x_n) = 0, \tag{17}$$

or

$$
\begin{aligned}
& \frac{\partial^{n-1}}{\partial x_2\ldots\partial x_n}F^{-1}_{1|2,\ldots,n}(\int_0^{x_1}\frac{h^{(1)}(u_1, x_2, \ldots, x_n)}{f_{2,\ldots,n}(x_2, \ldots, x_n)}du_1|x_2, \ldots, x_n) + \\
& \frac{\partial^{n-1}}{\partial x_1\partial x_3\ldots\partial x_n}(H^{(2)}_{2|1,3,\ldots,n})^{-1}(\int_0^{x_2}\frac{h^{(1)}(x_1, u_2, x_3, \ldots, x_n)}{h^{(2)}_{1,3,\ldots,n}(x_1, x_3, \ldots, x_n)}du_2|x_1, x_3, \ldots, x_n) = 0.
\end{aligned}
\tag{18}
$$

Then theorem 1 follows.

References

Ambrosio, L. (2002). Optimal transport maps in Monge-Kantorovich problem. *Proceedings of ICM Beijing*, vol. iii, 131-140.

Brenier, Y. (1987). Décomposition polaire et réarrangement monotone des champs de vecteurs, *C.R. Acad. Sci. Paris, Sér I Math*, 305, 805-808.

Brenier, Y. (1991). Polar factorization and monotone rearrangement of vector-valued functions, *Comm. Pure Appl. Math.* 44, 375-417.

Caffarelli, L. (1996). Allocation maps with general cost functions, *Lecture Notes in Pure and Appl. Math.* 177, 29-35.

Caffarelli, L., Feldman, M. & McCann R. J. (2002). Constructing optimal maps for Monge's transport problem as a limit of strictly convex costs, *J. Amer. Math. Soc.* 15, 1-26.

Caffarelli, L. (2002). Non linear elliptic theory and the Monge-Ampere equation, *Proceedings of the ICM Beijing*, vol. i, 179-187.

Durrett, R. (2005). *Probability, theory and examples*, Thomson, Brooks/Cole.

Evans, L. C. & Gangbo, W. (1999). Differential equations methods for the Monge-Kantorovich mass transfer problem, *Mem. Amer. Math. Soc.* 137, no. 653.

Gangbo, W. & McCann, R. J. (1996). The geometry of optimal transportation. *Acta Math.*, 177, 113-161.

Givens, C. R. & Short, R. M. (1984). A Class of Wasserstein Metrics for Probability Distributions, *J Michigan Math.*, 31, 231-140.

Gilbarg, D. & Trudinger, N. S. (1983). *Elliptic Partial Differential Equations of Second Order*, Springer-Verlag.

Jordan, R., Kinderlehrer, D. & Otto, F. (1998). The variational formulation of the Fokker-Planck equation, *SIAM J Math. Anal.*, 29, 1-17.

Kantorovich, L. V. (1942). On the transfer of masses, *Dokl. Akad. Nauk. SSSR*, 37, 227-229.

Knott, M. & Smith, C. S. (1987). Note on the optimal transportation of distributions, *J Optim. Theory Appl.*, 52, 323-329.

Monge, G. (1781). Memoire sur la Theorie des Dblais et des Remblais, *Histoire de l'Acad. des Sciences de Paris*.

Mufa, C. (1992). *From Markov Chains to Non-Equilibrium Particle Systems*, World Scientific.

Rachev, S. T. & Rüschendorf, L. (1998). *Mass transportation problems*, Vol I:Theory, Vol. II: Applications, Springer.

Trudinger, N. S. (2006). Recent developments in elliptic partial differential equations of Monge-Ampère type, *ICM*, invited lecture.

Trudinger, N. S. & Xujia, W. (2001). On the Monge mass transfer problem, *Calc. Var. PDE*, 13, 19-31.

Villani, C. (2003). *Topics in optimal transportation*, AMS.

Weian, Z. (2000). Rate Estimation for Weak Convergence of Diffusion Processes, *Skorohod's Ideas in Probability Theory*, National Academy of Science of Ukraine, Institute of Mathematics, Kyiv, 318-332.

Yinfang, S. & Weian, Z. (2010). On Monge-Kantorovich Problem in the Plane, *C. R. Acad. Sci. Paris*, Ser. I, 348, 267-271.