Optimal Couplings of Kantorovich-Rubinstein-Wasserstein L_p-distance

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: May 24, 2011 Accepted: June 9, 2011 Published: November 1, 2011
doi:10.5539/jmr.v3n4p3 URL: http://dx.doi.org/10.5539/jmr.v3n4p3

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

Abstract

We achieve that the optimal solutions according to Kantorovich-Rubinstein-Wasserstein L_p-distance ($p > 2$) (abbreviation: KRW L_p-distance) in a bounded region of Euclidean plane satisfy a partial differential equation. We can also obtain the similar results about Monge-Kantorovich problem with more general convex cost functions.

Keywords: Monge-Kantorovich Problem, KRW L_p-distance, Optimal coupling, Partial differential equation

1. Introduction

The classical mass transportation problem of Monge and its version of Kantorovich has found a lot of recent interest because of its applications in lots of fields. 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 $\|X - Y\|_{L_p}$ ($p \geq 1$) among all such possible couplings is called KRW L_p-distance between P and $\tilde{P}$, whose coupling is called optimal coupling. This problem has been only completely solved in one dimensional case. In R^1, KRW L_p-distance is just given by

$$\{\int_0^1 |F^{-1}(u) - \tilde{F}^{-1}(u)|^p du\}^{\frac{1}{p}}, \tag{1}$$

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

In our recent paper (Yinfang & Weian, 2010), we have transformed the Monge-Kantorovich problem as $p = 2$ into Dirichlet boundary problems, we have also obtained the corresponding partial differential equations group in (Yinfang, 2011), and we have achieved an explicit formula of Kantorovich-Rubinstein-Wasserstein L_p-distance ($p > 2$) in (Yinfang, 2011). Now we draw a conclusion that the optimal couplings according to KRW L_p-distance ($p > 2$) in a bounded region of Euclidean plane satisfies a partial differential equation. The proofs are based on variational method from probability point of view. We can also get the similar results about Monge-Kantorovich problem with more general convex cost functions.

2. Main results

Without losing generality, we may consider two probability measures P and $\tilde{P}$ on $[0, 1] \times [0, 1]$. Let X and Y be two random vectors defined on a same probability space with P and $\tilde{P}$ as their individual laws, and $p \geq 2$. Then

$$E|X - Y|^p = E(|X_1 - Y_1|^2 + |X_2 - Y_2|^2)^{\frac{p}{2}}. \tag{2}$$

We assume further their density functions $f(x, y)$ and $\tilde{f}(x, y)$ are smooth and strictly positive on their domains. Denote the marginal densities

$$f_1(x) = \int_0^1 f(x, y)dy, f_2(y) = \int_0^1 f(x, y)dx,$$

and

$$\tilde{f}_1(x) = \int_0^1 \tilde{f}(x, y)dy, \tilde{f}_2(y) = \int_0^1 \tilde{f}(x, y)dx.$$

Furthermore, denote the conditional distributions

$$F_{1|2}(x|y) = \frac{1}{f_2(y)}\int_0^x f(u,y)du, F_{2|1}(y|x) = \frac{1}{f_1(x)}\int_0^y f(x,u)du,$$

and

$$\tilde{F}_{1|2}(x|y) = \frac{1}{\tilde{f}_2(y)}\int_0^x \tilde{f}(u,y)du, \tilde{F}_{2|1}(y|x) = \frac{1}{\tilde{f}_1(x)}\int_0^y \tilde{f}(x,u)du,$$

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

Now Denote by $\mathcal{G}$ the set of all density functions $g(x,y)$ on $[0,1]\times[0,1]$ such that $f_1(x) = \int_0^1 g(x,y)dy$ and $\tilde{f}_1(y) = \int_0^1 g(x,y)dx$. Then we have

Lemma 1. *(Yinfang, 2011) Suppose that X, Y are the optimal coupling, hence*

$$\begin{aligned} E|X-Y|^p &= \int_0^1\int_0^1\int_0^1 [(x-y)^2 + (F_{2|1}^{-1}(\int_0^t \frac{g(x,u)}{f_1(x)}du|x) - \tilde{F}_{2|1}^{-1}(\int_0^t \frac{g(x,u)}{f_1(x)}du|y))^2]^{\frac{p}{2}} \\ &\quad g(x,y)g(x,t)\frac{1}{f_1(x)}dtdxdy. \end{aligned} \tag{3}$$

So we just need to look for a density function $g(x,y)\in\mathcal{G}$ minimizes (3). Actually, we have

Theorem 1. *When $p > 2$, $g\in\mathcal{G}$ minimize (3), then*

$$\begin{aligned} &\frac{\partial^2}{\partial x\partial y}\{\int_0^1 [(x-y)^2 + (F_{2|1}^{-1}(\int_0^t \frac{g(x,u)}{f_1(x)}du|x) - \tilde{F}_{2|1}^{-1}(\int_0^t \frac{g(x,u)}{f_1(x)}du|y))^2]^{\frac{p}{2}}\frac{g(x,t)}{f_1(x)}dt \\ &+\int_0^1 (x-t)^p\frac{g(x,y)}{f_1(x)}dt\} = 0. \end{aligned} \tag{4}$$

Proof: For $0 < a_1 < a_2 < 1$ and $0 < b_1 < b_2 < 1$ when ϵ is small enough, s.t. $a_1+\epsilon < a_2 < a_2+\epsilon < 1, b_1+\epsilon < b_2 < b_2+\epsilon < 1$. Define

$$\xi(s,t) = I_{([a_1,a_1+\epsilon]\times[b_1,b_1+\epsilon])\cup([a_2,a_2+\epsilon]\times[b_2,b_2+\epsilon])}(s,t) - I_{([a_1,a_1+\epsilon]\times[b_2,b_2+\epsilon])\cup([a_2,a_2+\epsilon]\times[b_1,b_1+\epsilon])}(s,t). \tag{5}$$

and then $g(s,t)+\delta\xi(s,t)\in\mathcal{G}$ when both ϵ,δ are small. Since g is the minimum,

$$\begin{aligned} 0 \le\ & \frac{1}{\epsilon^2}\{\int_0^1\int_0^1\int_0^1 [(x-y)^2 + (F_{2|1}^{-1}(\int_0^t \frac{g(x,u)+\delta\xi(x,u)}{f_1(x)}du|x) - \tilde{F}_{2|1}^{-1}(\int_0^t \frac{g(x,u)+\delta\xi(x,u)}{f_1(x)}du|y))^2]^{\frac{p}{2}} \\ & (g(x,y)+\delta\xi(x,y))(g(x,t)+\delta\xi(x,t))\frac{1}{f_1(x)}dtdxdy \\ & -\int_0^1\int_0^1\int_0^1 [(x-y)^2 + (F_{2|1}^{-1}(\int_0^t \frac{g(x,u)}{f_1(x)}du|x) - \tilde{F}_{2|1}^{-1}(\int_0^t \frac{g(x,u)}{f_1(x)}du|y))^2]^{\frac{p}{2}} g(x,y)g(x,t)\frac{1}{f_1(x)}dtdxdy\} \end{aligned}$$

Denote $\int_0^t \frac{g(x,v)}{f_1(x)}dv = \phi(x,t)$, letting $\epsilon\to 0$, we get

$$\begin{aligned} 0 \le\ & \int_0^1\int_{b_1}^{b_2}((a_1-y)^2 + (F_{2|1}^{-1}(\int_0^t \frac{g(a_1,u)}{f_1(a_1)}du|x) - \tilde{F}_{2|1}^{-1}(\int_0^t \frac{g(a_1,u)}{f_1(a_1)}du|y))^2)^{\frac{p}{2}-1} \\ & (F_{2|1}^{-1}(\int_0^t \frac{g(a_1,u)}{f_1(a_1)}du|x) - \tilde{F}_{2|1}^{-1}(\int_0^t \frac{g(a_1,u)}{f_1(a_1)}du|y)) \\ & (\frac{\partial}{\partial\phi(a_1,t)}F_{2|1}^{-1}(\int_0^t \frac{g(a_1,u)}{f_1(a_1)}du|x) - \frac{\partial}{\partial\phi(a_1,t)}\tilde{F}_{2|1}^{-1}(\int_0^t \frac{g(a_1,u)}{f_1(a_1)}du|y))p\delta g(a_1,y)g(a_1,t)\frac{1}{f_1^2(a_1)}dtdy \end{aligned}$$

$$
\begin{aligned}
&-\int_0^1\int_{b_1}^{b_2}((a_2-y)^2+(F_{2|1}^{-1}(\int_0^t\frac{g(a_2,u)}{f_1(a_2)}du|x)-\tilde{F}_{2|1}^{-1}(\int_0^t\frac{g(a_2,u)}{f_1(a_2)}du|y))^2)^{\frac{p}{2}-1}\\
&(F_{2|1}^{-1}(\int_0^t\frac{g(a_2,u)}{f_1(a_2)}du|x)-\tilde{F}_{2|1}^{-1}(\int_0^t\frac{g(a_2,u)}{f_1(a_2)}du|y))\\
&(\frac{\partial}{\partial\phi(a_2,t)}F_{2|1}^{-1}(\int_0^t\frac{g(a_2,u)}{f_1(a_2)}du|x)-\frac{\partial}{\partial\phi(a_2,t)}\tilde{F}_{2|1}^{-1}(\int_0^t\frac{g(a_2,u)}{f_1(a_2)}du|y))p\delta g(a_2,y)g(a_2,t)\frac{1}{f_1^2(a_2)}dtdy\\
&+\int_0^1((a_1-b_1)^2+(F_{2|1}^{-1}(\int_0^t\frac{g(a_1,u)}{f_1(a_1)}du|x)-\tilde{F}_{2|1}^{-1}(\int_0^t\frac{g(a_1,u)}{f_1(a_1)}du|y))^2)^{\frac{p}{2}}\delta g(a_1,t)\frac{1}{f_1(a_1)}dt\\
&-\int_0^1((a_2-b_1)^2+(F_{2|1}^{-1}(\int_0^t\frac{g(a_2,u)}{f_1(a_2)}du|x)-\tilde{F}_{2|1}^{-1}(\int_0^t\frac{g(a_2,u)}{f_1(a_2)}du|y))^2)^{\frac{p}{2}}\delta g(a_2,t)\frac{1}{f_1(a_2)}dt\\
&-\int_0^1((a_1-b_2)^2+(F_{2|1}^{-1}(\int_0^t\frac{g(a_1,u)}{f_1(a_1)}du|x)-\tilde{F}_{2|1}^{-1}(\int_0^t\frac{g(a_1,u)}{f_1(a_1)}du|y))^2)^{\frac{p}{2}}\delta g(a_1,t)\frac{1}{f_1(a_1)}dt\\
&+\int_0^1((a_2-b_2)^2+(F_{2|1}^{-1}(\int_0^t\frac{g(a_2,u)}{f_1(a_2)}du|x)-\tilde{F}_{2|1}^{-1}(\int_0^t\frac{g(a_2,u)}{f_1(a_2)}du|y))^2)^{\frac{p}{2}}\delta g(a_2,t)\frac{1}{f_1(a_2)}dt\\
&+\int_0^1((a_1-y)^2+(F_{2|1}^{-1}(\int_0^{b_1}\frac{g(a_1,u)}{f_1(a_1)}du|x)-\tilde{F}_{2|1}^{-1}(\int_0^{b_1}\frac{g(a_1,u)}{f_1(a_1)}du|y))^2)^{\frac{p}{2}}\delta g(a_1,y)\frac{1}{f_1(a_1)}dt\\
&-\int_0^1((a_1-y)^2+(F_{2|1}^{-1}(\int_0^{b_2}\frac{g(a_1,u)}{f_1(a_1)}du|x)-\tilde{F}_{2|1}^{-1}(\int_0^{b_2}\frac{g(a_1,u)}{f_1(a_1)}du|y))^2)^{\frac{p}{2}}\delta g(a_1,y)\frac{1}{f_1(a_1)}dt\\
&-\int_0^1((a_2-y)^2+(F_{2|1}^{-1}(\int_0^{b_1}\frac{g(a_2,u)}{f_1(a_2)}du|x)-\tilde{F}_{2|1}^{-1}(\int_0^{b_1}\frac{g(a_2,u)}{f_1(a_2)}du|y))^2)^{\frac{p}{2}}\delta g(a_2,y)\frac{1}{f_1(a_2)}dt\\
&+\int_0^1((a_2-y)^2+(F_{2|1}^{-1}(\int_0^{b_2}\frac{g(a_2,u)}{f_1(a_2)}du|x)-\tilde{F}_{2|1}^{-1}(\int_0^{b_2}\frac{g(a_2,u)}{f_1(a_2)}du|y))^2)^{\frac{p}{2}}\delta g(a_2,y)\frac{1}{f_1(a_2)}dt
\end{aligned}
$$

Consequently we can say

$$\frac{\partial^2}{\partial x\partial y}N(x,y)\geq 0, \tag{6}$$

where

$$
\begin{aligned}
N(x,y) &= -\int_0^1\int_0^y[(x-t)^2+(F_{2|1}^{-1}(\int_0^w\frac{g(x,u)}{f_1(x)}du|x)-\tilde{F}_{2|1}^{-1}(\int_0^w\frac{g(x,u)}{f_1(x)}du|t))^2]^{\frac{p}{2}-1}\\
&\quad(F_{2|1}^{-1}(\int_0^w\frac{g(x,u)}{f_1(x)}du|x)-\tilde{F}_{2|1}^{-1}(\int_0^w\frac{g(x,u)}{f_1(x)}du|t))\frac{g(x,t)g(x,w)}{f_1^2(x)}\\
&\quad(\frac{\partial}{\partial\phi(x,w)}F_{2|1}^{-1}(\int_0^w\frac{g(x,u)}{f_1(x)}du|x)-\frac{\partial}{\partial\phi(x,w)}\tilde{F}_{2|1}^{-1}(\int_0^w\frac{g(x,u)}{f_1(x)}du|t))dwdt\\
&\quad+\int_0^1[(x-y)^2+(F_{2|1}^{-1}(\int_0^t\frac{g(x,u)}{f_1(x)}du|x)-\tilde{F}_{2|1}^{-1}(\int_0^t\frac{g(x,u)}{f_1(x)}du|y))^2]^{\frac{p}{2}}\frac{g(x,t)}{f_1(x)}dt\\
&\quad+\int_0^1[(x-t)^2+(F_{2|1}^{-1}(\int_0^y\frac{g(x,u)}{f_1(x)}du|x)-\tilde{F}_{2|1}^{-1}(\int_0^y\frac{g(x,u)}{f_1(x)}du|t))^2]^{\frac{p}{2}}\frac{g(x,y)}{f_1(x)}dt\\
&= -\int_0^1[(x-t)^2+(F_{2|1}^{-1}(\int_0^w\frac{g(x,u)}{f_1(x)}du|x)-\tilde{F}_{2|1}^{-1}(\int_0^w\frac{g(x,u)}{f_1(x)}du|t))^2]^{\frac{p}{2}}|_{w=0}^{w=y}\frac{g(x,y)}{f_1(x)}dt\\
&\quad+\int_0^1[(x-y)^2+(F_{2|1}^{-1}(\int_0^t\frac{g(x,u)}{f_1(x)}du|x)-\tilde{F}_{2|1}^{-1}(\int_0^t\frac{g(x,u)}{f_1(x)}du|y))^2]^{\frac{p}{2}}\frac{g(x,t)}{f_1(x)}dt\\
&\quad+\int_0^1[(x-t)^2+(F_{2|1}^{-1}(\int_0^y\frac{g(x,u)}{f_1(x)}du|x)-\tilde{F}_{2|1}^{-1}(\int_0^y\frac{g(x,u)}{f_1(x)}du|t))^2]^{\frac{p}{2}}\frac{g(x,y)}{f_1(x)}dt\\
&= \int_0^1[(x-y)^2+(F_{2|1}^{-1}(\int_0^t\frac{g(x,u)}{f_1(x)}du|x)-\tilde{F}_{2|1}^{-1}(\int_0^t\frac{g(x,u)}{f_1(x)}du|y))^2]^{\frac{p}{2}}\frac{g(x,t)}{f_1(x)}dt\\
&\quad+\int_0^1(x-t)^p\frac{g(x,y)}{f_1(x)}dt.
\end{aligned}
$$

On the other hand, if one replace $g + \delta\xi$ by $g - \delta\xi$, the same computation leads

$$\frac{\partial^2}{\partial x \partial y} N(x, y) \leq 0. \tag{7}$$

Thus we deduce that

$$\frac{\partial^2}{\partial x \partial y} N(x, y) = 0, \qquad \forall\, 0 < x, y < 1. \tag{8}$$

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. (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.

Givens, C. R. & Short, R. M. (1984). A Class of Wasserstein Metrics for Probability Distributions, *J Michigan Math.*, 31, 231-140. http://dx.doi.org/10.1307/mmj/1029003026

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. http://dx.doi.org/10.1007/BF00941290

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. (1985). The MongeCKantorovich mass transference problem and its stochastic applications, *Theory of Probability and Its Applications*, 29, 647C676.

Rachev, S.T. & Rüschendorf, L. (1990). A characterization of random variables with minimum L_2 distance, *J. Multivariate Anal.*, 32, 48-54. http://dx.doi.org/10.1016/0047-259X(90)90070-X

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

Rüschendorf, L. (1985). The Wasserstein distance and approximation theorems, *Z. Wahrs. verw. Ge-biete*, 70, 117-129.

Rüschendorf, L. (1995). Optimal solutions of multivariate coupling problems, *Appl. Mathematics*, 23, 325C338.

Rüschendorf, L. (1996). On c-optimal random variables, *Statistics Prob. Letters*, 27, 267C 270.

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. http://dx.doi.org/10.1007/PL00009922

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

Vallender, S. (1973). Calculation of the Wasserstein distance between probability distributions on the line, *Theory prob.Appl.*, 18, 785-786.

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.

Yinfang, S. (2011). Multivariate Monge-Kantorovich Transportation Problem, *Journal of Mathematics Research*, 3, No. 2, 66-73.

Yinfang, S. (2011). Kantorovich-Rubinstein-Wasserstein L_p-distance ($p > 2$), *Science* & *Technology Information*, No. 15, 22.