最適輸送理論とその周辺

高津 飛鳥*

首都大学東京，2016 年 2 月

1　復習

講演では “物質をある場所から他の場所へ運ぶ費用を最小化” するための最適輸送理論とその周辺を紹介した．特に最適輸送理論に起源を持つ Wasserstein 幾何-確率測度のなす空間上の距離の幾何-を取り扱った．

定義 1.1. 完備可分距離空間 (X,d) に対し, $\mathcal{P}_2(X)$ を X 上の 2 次モーメントが有限なボレル確率測度のなす集合とする．すなわち

$$\mathcal{P}_2(X) := \left\{ \mu\text{: } X \text{ 上のボレル確率測度} \;\middle|\; \text{ある } x_0 \in X \text{ が存在し} \int_X d(x_0,x)^2 d\mu(x) < \infty \right\}$$

である．また X 上のボレル確率測度 μ,ν に対し, $X \times X$ 上のボレル確率測度 π が

$$\pi(B \times X) = \mu(B), \quad \pi(X \times B) = \nu(B), \quad \forall B \subset X\text{:ボレル集合}$$

を満たすとき π を μ,ν の**カップリング**と呼ぶ．さらに

$$W_2(\mu,\nu) := \inf_{\pi:\mu,\nu \text{ のカップリング}} \|d\|_{\mathrm{L}^2(\pi)} = \inf_{\pi:\mu,\nu \text{ のカップリング}} \left(\int_{X \times X} d(x,y)^2 d\pi(x,y) \right)^{\frac{1}{2}}$$

を μ,ν の *Wasserstein* 距離, 下限を達成するカップリングを**最適カップリング**と呼ぶ.

注記 1.2. 最適カップリングは一意とは限らないが常に存在する ([4, Theorem 4,1]). そして $(\mathcal{P}_2(X),W_2)$ もまた完備可分距離空間になり, (X,d) の最短線と $(\mathcal{P}_2(X),W_2)$ の最短線は密接に関係する ([4, Theorem 6.18]). 例えば W_2-最短測地線 $\{\mu_t\}_{t\in[0,1]}$ を考え, A_0,A_1 をそれぞれ μ_0,μ_1 の台とする．すると任意の $t \in [0,1]$ に対し

$$\mu_t(\{\gamma_t \mid \gamma : [0,1] \to X \text{ は最短測地線で } \gamma_0 \in A_0, \gamma_1 \in A_1 \text{ を満たす }\}) = 1$$

が成立つ ([4, Theorem 7.21]). これは標語的には “確率測度のなす空間内の測地線は測地線のなす空間上の確率測度” であることを意味する．

そしてこの距離構造を用いたリッチ曲率の下限・次元の上限の特徴付けが Lott–Villani[1] や Sturm[3] によって与えられた．この二条件は任意の最短線が分岐しない距離空間に対しては同値であり, ここでは Sturm による特徴付けを紹介する．

*asuka@tmu.ac.jp

そのために以下, 距離空間 (X, d) とその上のボレル測度 ω に対し

$$\mathcal{P}_2(X, \omega) := \{\mu \in \mathcal{P}_2(X) \mid \mu \text{ は } \omega \text{ に絶対連続}\}$$

と定義する. さらに $N \in (1, \infty)$ に対し, $\mathcal{P}_2(X, \omega)$ 上の汎関数 $S_{N,\omega}$ を

$$S_{N,\omega}(\mu) := -\int_X \rho(x)^{1-\frac{1}{N}} d\omega(x) = -\int_X \rho(x)^{-\frac{1}{N}} d\mu(x), \quad \forall \mu = \rho\omega \in \mathcal{P}_2(X, \omega)$$

と定める. (これは 2.2 節で定義される講演中に用いた汎関数 $\tilde{S}_{N,\omega}$ とは異なるが, 以下で考える汎関数の凸性の議論に関しては同値な対象となる.) さらに $K \in \mathbb{R}, N \in (1, \infty)$ と $t \in (0, 1)$ に対し

$$\beta^t_{K,N}(\theta) := \begin{cases} \infty & \text{if } K\theta^2 \geq (N-1)\pi^2, \\ \left(\dfrac{\sin(t\theta\sqrt{K/(N-1)})}{t\sin(\theta\sqrt{K/(N-1)})}\right)^{N-1} & \text{if } 0 < K\theta^2 < (N-1)\pi^2, \\ 1 & \text{if } K\theta^2 = 0, \\ \left(\dfrac{\sinh(t\theta\sqrt{-K/(N-1)})}{t\sinh(\theta\sqrt{-K/(N-1)})}\right)^{N-1} & \text{if } K\theta^2 < 0 \end{cases}$$

と定める. これは空間形のヤコビ方程式に関係する量である ([4, Definition 14.19] 参照). そして任意の $\mu_0 = \rho_0\omega, \mu_1 = \rho_1\omega \in \mathcal{P}_2(M, \omega)$ に対し, ある μ_0, μ_1 の最適カップリング π と μ_0, μ_1 を結ぶ $\mathcal{P}_2(X, \omega)$ に横たわる W_2-最短測地線 $\{\mu_t\}_{t\in[0,1]}$ が存在し

$$S_{N,\omega}(\mu_t) \leq -\int_{X\times X} \left[(1-t)\left(\frac{\rho_0(x)}{\beta^{1-t}_{K,N}(d(x,y))}\right)^{-\frac{1}{N}} + t\left(\frac{\rho_1(y)}{\beta^t_{K,N}(d(x,y))}\right)^{-\frac{1}{N}}\right] d\pi(x,y) \tag{1}$$

が任意の $t \in (0, 1)$ に対し成立つとき, (X, d, ω) は**条件** $\mathrm{CD}(K, N)$ **を満たす**と言う.

定理 1.3. ([3, Theorem 1.7]) 完備連結リーマン多様体 (M, g) と $K \in \mathbb{R}, N \in (1, \infty)$ を考える. このとき M のリッチ曲率が K 以上かつ次元が N 以下であることと, (M, d_g, vol_g) が条件 $\mathrm{CD}(K, N)$ を満たすことは同値である. ここで d_g と vol_g はそれぞれ (M, g) の距離関数と体積測度である.

注記 1.4. 条件 $\mathrm{CD}(K, N)$ ではある一つの最適カップリングと W_2-最短測地線について考察したが, 任意の $\mu_0, \mu_1 \in \mathcal{P}_2(M, \mathrm{vol}_g)$ に対しその最適カップリングとそれらを結ぶ W_2-最短測地線 $\{\mu_t\}_{t\in[0,1]}$ は一意的である ([4, Corollaries 9.3, 7.23]). さらにこの W_2-最短測地線は $\mathcal{P}_2(M, \mathrm{vol}_g)$ 内に横たわる ([4, Theorem 8.7]). よって $\mathcal{P}_2(M, \mathrm{vol}_g)$ は $(\mathcal{P}_2(M), W_2)$ の凸集合である.

このようにリーマン多様体のリッチ曲率の下限・次元の上限は条件 $\mathrm{CD}(K, N)$ で特徴付けられ, $\mathrm{CD}(K, N)$ は**曲率次元条件** (Curvature–Dimension condition) と呼ばれる. そして条件 $\mathrm{CD}(K, N)$ の定義には微分構造を必要としないので, 任意の最短線が分岐しない測度距離空間 (X, d, ω) に対しても論じることができる[1]. ここで**測度距離空間**とは任意の二点を最短線で結ぶことができる完備可分距離空間 (X, d) とその上の σ-有限なボレル測度 ω のなす三組 (X, d, ω) のことである. さらに $\mathrm{CD}(K, N)$ を満たす測度距離空間はリッチ曲率の下限が K・次元の上限が N であるリーマン多様体のように振舞うことが分かり, 測度距離空間上の幾何解析は著しく進展した.

また講演では $K = 0$ の場合のみを紹介した. このとき $\beta^t_{0,N} \equiv 1$ なので, (1) は

$$S_{N,\omega}(\mu_t) \leq -\int_{X\times X} \left[(1-t)\rho_0(x)^{-\frac{1}{N}} + t\rho_1(y)^{-\frac{1}{N}}\right] d\pi(x,y) = (1-t)S_{N,\omega}(\mu_0) + tS_{N,\omega}(\mu_1)$$

[1] 最短線が分岐する測度距離空間に対しては Lott–Villani [1] による条件を考えるが, 煩雑になるためここでは割愛する.

と $S_{N,\omega}$ の W_2-最短測地線に沿った凸性になる．ここで等式は π が μ_0, μ_1 のカップリングであることから従う．そして $\mathrm{CD}(0,N)$ を満たす測度距離空間がリッチ曲率の下限が 0・次元の上限が N であるリーマン多様体のように振舞う例として, 以下の *Brunn–Minkowski* **不等式**を紹介した.

定理 1.5. ([3, Proposition 2.1]) 測度距離空間 (X,d,ω) は $\mathrm{CD}(0,N)$ を満たすとする．このとき任意の正測度を持つコンパクト集合 $A_0, A_1 \subset X$ と $t \in (0,1)$ に対し,

$$A_t := \{\gamma_t \mid \gamma : [0,1] \to X \text{ は最短測地線で } \gamma_0 \in A_0, \gamma_1 \in A_1 \text{ を満たす} \}$$

と定めれば,

$$\omega(A_t)^{\frac{1}{N}} \geq (1-t)\omega(A_0)^{\frac{1}{N}} + t\omega(A_1)^{\frac{1}{N}}$$

が成立つ.

2 宿題の解答

宿題は「定理 1.5 を示せ」と「$N \to \infty$ のとき (定理 1.3,1.5 は) どうなるかを観察せよ」であった.

2.1 定理 1.5 の証明

ここでは [3, Proposition 2.1] の証明と同様に示す．$i = 0,1$ に対し A_i はコンパクトで ω は σ-有限なので $\omega(A_i) < \infty$ である．そこで

$$1_{A_i}(x) := \begin{cases} 1 & \text{if } x \in A_i, \\ 0 & \text{if } x \notin A_i, \end{cases} \qquad \mu_i := \frac{1_{A_i}}{\omega(A_i)}\omega$$

とおけば, A_i のコンパクト性より $\mu_i \in \mathcal{P}_2(X,\omega)$ である．またある μ_0, μ_1 の最適カップリング π と μ_0, μ_1 を結ぶ $\mathcal{P}_2(X,\omega)$ に横たわる W_2-最短測地線 $\{\mu_t\}_{t\in[0,1]}$ が存在し

$$S_{N,\omega}(\mu_t) \leq (1-t)S_{N,\omega}(\mu_0) + tS_{N,\omega}(\mu_1) \tag{2}$$

が成立つ．ここで

$$S_{N,\omega}(\mu_i) = -\int_X \rho_i(x)^{-\frac{1}{N}}\,d\mu(x) = -\omega(A_i)^{\frac{1}{N}}$$

であり, 注記 1.2 より $\mu_t(A_t) = 1$ なので

$$S_{N,\omega}(\mu_i) = -\int_{A_t} \rho_t(x)^{-\frac{1}{N}}\,d\mu_t(x) = -\int_{A_t} \rho_t(x)^{1-\frac{1}{N}}\,d\omega(x)$$

となる．そして (2) より $\omega(A_t) > 0$ が従い, $\omega(A_t) = \infty$ ならば主張は正しい．そこで $\omega(A_t) < \infty$ を仮定すると, Jensen の不等式 から

$$S_{N,\omega}(\mu_i) = -\omega(A_t)\int_{A_t} \rho_t(x)^{1-\frac{1}{N}}\frac{d\omega(x)}{\omega(A_t)} \geq -\omega(A_t)\left(\int_{A_t} \rho_t(x)\frac{d\omega(x)}{\omega(A_t)}\right)^{1-\frac{1}{N}} = -\omega(A_t)^{\frac{1}{N}}$$

が従う．これらの式を整理して $\omega(A_t)^{\frac{1}{N}} \geq (1-t)\omega(A_0)^{\frac{1}{N}} + t\omega(A_1)^{\frac{1}{N}}$ を得る.

注記 2.1. 一般の $\mathrm{CD}(K,N)$ を満たす測度距離空間に対する Brunn–Minkowski 不等式は例えば [3, Proposition 2.1][4, Theorem 30.7] を参照にして頂きたい．この曲がった空間に対する Brunn–Minkowski 不等式はリーマン多様体に対してさえも最適輸送を用いることで初めて示された．そしてこの Brunn–Minkowski 不等式からリッチ曲率の下限が K・元の上限が N であるリーマン多様体が満たす Bishop–Gromov 不等式と同様の不等式が導かれる ([3, Theorem 2.3][4, Theorem 30.11] 参照).

2.2 $N \to \infty$ でどうなるか?

定理 1.3 において $N \to \infty$ とすれば次元に関する制約がなくなり, 完備連結リーマン多様体に対しそのリッチ曲率が K 以上であることがある条件 $\mathrm{CD}(K,\infty)$ を満たすことに同値だと予測される. そして $\mathrm{CD}(K,N)$ は汎関数 $S_{N,\omega}$ の不等式 (1) で特徴付けられたので, $\mathrm{CD}(K,\infty)$ もある汎関数の不等式で特徴付けられると予想される. ここで $N \to \infty$ のとき $S_{N,\omega}$ は一般に意味のある量に収束しないが, $S_{N,\omega}$ に適切なスケールを施した汎関数 $\tilde{S}_{N,\omega}(\mu) := N(S_{N,\omega}(\mu)+1)$ は収束して

$$\lim_{N\to\infty} \tilde{S}_{N,\omega}(\mu) = S_\omega(\mu) := \int_X \rho(x)\log\rho(x)d\omega(x), \quad \forall\mu = \rho\omega \in \mathcal{P}_2(X,\omega)$$

となる. なぜなら正数上の関数

$$s_N(r) := -N(r^{-\frac{1}{N}} - 1)$$

を用いれば

$$\tilde{S}_{N,\omega}(\mu) = \int_X \rho(x)s_N(\rho(x))d\omega(x) = \int_X s_N(\rho(x))d\mu(x), \quad \forall\mu = \rho\omega \in \mathcal{P}_2(X,\omega)$$

と表現され, そして $\tau < N$ に対し定義される s_N の逆関数 $\sigma_N(\tau)$ は

$$\sigma_N(\tau) = \left(1 - \frac{\tau}{N}\right)^{-N} \xrightarrow{N\to\infty} e^\tau$$

という性質を持ち, $\lim_{N\to\infty} s_N(t) = \log t$ となるからである. また (1) は

$$\begin{aligned}
\tilde{S}_{N,\omega}(\mu_t) &\le -N\int_{X\times X}\left[(1-t)\left(\frac{\rho_0(x)}{\beta^{1-t}_{K,N}(d(x,y))}\right)^{-\frac{1}{N}} + t\left(\frac{\rho_1(y)}{\beta^{t}_{K,N}(d(x,y))}\right)^{-\frac{1}{N}}\right]d\pi(x,y) + N \\
&= \int_{X\times X}\left[(1-t)s_N\left(\frac{\rho_0(x)}{\beta^{1-t}_{K,N}(d(x,y))}\right) + ts_N\left(\frac{\rho_1(y)}{\beta^{t}_{K,N}(d(x,y))}\right)\right]d\pi(x,y)
\end{aligned}$$

と同値である. さらに $0 < K\theta^2 < (N-1)\pi^2$ ならば sin のテイラー展開を用いることで

$$\beta^t_{K,N}(\theta) = \left(\frac{\sin(t\theta\sqrt{K/(N-1)})}{t\sin(\theta\sqrt{K/(N-1)})}\right)^{N-1} = \left(1 + \frac{K(1-t^2)\theta^2}{3!(N-1)} + o\left(\frac{1}{N-1}\right)\right)^{N-1} \quad \text{as } N\to\infty$$

を得るので, $\lim_{N\to\infty}\beta^t_{K,N}(\theta) = \exp\left(K(1-t^2)\theta^2/6\right)$ となる. 同様に一般の $K\in\mathbb{R}, N\in(1,\infty)$ と $t\in(0,1]$ そして $\theta\in\mathbb{R}$ に対しても

$$\beta^t_{K,\infty}(\theta) := \lim_{N\to\infty}\beta^t_{K,N}(\theta) = \exp\left(\frac{K}{6}(1-t^2)\theta^2\right)$$

を得る. ここで $\pi(\Omega) = 1$ となる適切な集合 $\Omega\subset X\times X$ 上で $N\to\infty$ としたときに $\beta^t_{K,N}(d(x,y))$ の収束が s_N の収束に較べて一様に速ければ,

$$\begin{aligned}
&\lim_{N\to\infty}\int_{X\times X}\left[(1-t)s_N\left(\frac{\rho_0(x)}{\beta^{1-t}_{K,N}(d(x,y))}\right) + ts_N\left(\frac{\rho_1(y)}{\beta^{t}_{K,N}(d(x,y))}\right)\right]d\pi(x,y) \\
=&\int_{X\times X}\left[(1-t)\log\left(\frac{\rho_0(x)}{\beta^{1-t}_{K,\infty}(d(x,y))}\right) + t\log\left(\frac{\rho_1(y)}{\beta^{t}_{K,\infty}(d(x,y))}\right)\right]d\pi(x,y) \\
=&(1-t)\int_{X\times X}\log\left(\rho_0(x)\right)d\pi(x,y) + t\int_{X\times X}\log\left(\rho_1(y)\right)d\pi(x,y) \\
&\quad - \int_{X\times X}\left[(1-t)\log\left(\beta^{1-t}_{K,\infty}(d(x,y))\right) + t\log\left(\beta^{t}_{K,\infty}(d(x,y))\right)\right]d\pi(x,y) \\
=&(1-t)S_\omega(\mu_0) + tS_\omega(\mu_1) - \frac{K}{2}t(1-t)\int_{X\times X}d(x,y)^2d\pi(x,y)
\end{aligned}$$

となる. さらに π は最適なので

$$\int_{X\times X} d(x,y)^2 d\pi(x,y) = W_2(\mu_0,\mu_1)^2$$

である. 以上を鑑みて (任意の最短線が分岐しない) 測度距離空間 (X,d,ω) が**条件** $\mathrm{CD}(K,\infty)$ **を満たす**ことを, 任意の $\mu_0=\rho_0\omega, \mu_1=\rho_1\omega \in \mathcal{P}_2(M,\omega)$ に対し, ある μ_0,μ_1 の最適カップリング π と μ_0,μ_1 を結ぶ $\mathcal{P}_2(X,\omega)$ に横たわる W_2-最短測地線 $\{\mu_t\}_{t\in[0,1]}$ が存在し

$$S_\omega(\mu_t) \le (1-t)S_{N,\omega}(\mu_0) + tS_{N,\omega}(\mu_1) - \frac{K}{2}t(1-t)W_2(\mu_0,\mu_1)^2$$

が任意の $t\in(0,1)$ に対し成立つこと, と定義することは妥当である. そして実際に次が成立つ.

定理 2.2. ([2, Theorem 1.1]) 完備連結リーマン多様体 (M,g) と $K\in\mathbb{R}$ を考える. このとき M のリッチ曲率が K 以上であることと, (M,d_g,vol_g) が条件 $\mathrm{CD}(K,\infty)$ を満たすことは同値である.

注記 2.3. (1) $\mathrm{CD}(1,\infty)$ を満たす測度距離空間の最もたる例は, ユークリッド空間とその上の標準ガウス測度がなす測度距離空間である. ここでユークリッド空間の次元は任意である.
(2) $\mathbb{R}^n$ 上の滑らかな関数 f と $K\in\mathbb{R}$ に対し, f のヘッセ行列の固有値が下から一様に K で抑えられること (このことを $\mathrm{Hess} f\ge K$ と書く) と, 任意の $x_0,x_1\in\mathbb{R}^n$ と $t\in(0,1)$ に対し

$$f((1-t)x_0+tx_1) \le (1-t)f(x_0)+tf(x_1) - \frac{K}{2}t(1-t)|x_0-x_1|^2$$

が成立つことは同値である ([4, Proposition 16.2] 参照). よって測度距離空間 (X,d,ω) に対し, もし $(\mathcal{P}_2(X,\omega),W_2)$ が可微分構造を許容し汎関数 S_ω が滑らかならば, (X,d,ω) が条件 $\mathrm{CD}(K,\infty)$ を満たすことは $\mathrm{Hess}S_\omega\ge K$ と同値である. よって $K=0$ のとき, 上述の議論は

(X,d,ω) が条件 $\mathrm{CD}(0,N)$ を満たす	$\Longleftrightarrow$	$\mathrm{Hess}\tilde{S}_{N,\omega}\ge 0$	
$N\to\infty\downarrow$		$N\to\infty\downarrow$	というようにみなせる.
(X,d,ω) が条件 $\mathrm{CD}(0,\infty)$ を満たす	$\Longleftrightarrow$	$\mathrm{Hess}S_\omega\ge 0$	

また 2.1 節にある定理 1.5 の証明において $S_{N,\omega}$ を S_ω に置き換え同様の議論を行う, あるいは

$$\begin{aligned} &\omega(A_t)^{\frac{1}{N}} \ge (1-t)\omega(A_0)^{\frac{1}{N}} + t\omega(A_1)^{\frac{1}{N}} \\ \Longleftrightarrow &N(\omega(A_t)^{\frac{1}{N}}-1) \ge (1-t)N(\omega(A_0)^{\frac{1}{N}}-1) + tN(\omega(A_1)^{\frac{1}{N}}-1) \end{aligned}$$

と変形して $N\to\infty$ とすることで, $\mathrm{CD}(0,\infty)$ を満たす測度距離空間 (X,d,ω) 上の Brunn–Minkowski 不等式は

$$\log\omega(A_t) \ge (1-t)\log\omega(A_0) + t\log\omega(A_1)$$

と定式化されることも分かる.

参考文献

[1] J. Lott and C. Villani, Ricci curvature for metric-measure spaces via optimal transport, Ann. of Math. **169** (2009), 903–991.

[2] M.-K. von Renesse and K.-T. Sturm, Transport inequalities, gradient estimates, entropy and Ricci curvature, Comm. Pure Appl. Math. **58** (2005), 923–940.

[3] K.-T. Sturm, On the geometry of metric measure spaces. II, Acta Math. **196** (2006), 133–177.

[4] C. Villani, Optimal transport, old and new, Springer-Verlag, Berlin, 2009.