確率測度の空間の幾何学

太田 慎一（京大・理）*

本講演では，距離空間上の確率測度のなす Wasserstein 空間の幾何学について概説する．まず Wasserstein 空間の測地線の性質とそれを用いたリーマン構造の導入法を解説し，応用としてリーマン幾何や偏微分方程式との関わりについて述べる．歴史的な背景や最適輸送理論全体の現状に関しては，Villani の２冊の本 [Vi1]，[Vi2] に詳しい．

1　最適輸送理論と Wasserstein 空間

最適輸送理論は，次のモンジュの問題にはじまるとされる．

問題 1.1 （モンジュの問題） 空間 X とその上のコスト関数 $c: X \times X \longrightarrow [0,\infty)$ が与えられたとする．このとき，X 上の２つの確率測度 μ,ν に対して，μ を ν に移す最もコストの少ない方法（最適輸送）を見つけよ．つまり，写像 $\Psi: X \longrightarrow X$ で $\Psi_\sharp\mu = \nu$ を満たし，かつ輸送コスト

$$\int_X c\bigl(x,\Psi(x)\bigr)\,d\mu(x)$$

をそのような写像の中で最小にするものを見つけよ．

ここで $\Psi_\sharp\mu$ は μ の Ψ による押し出しを表す．この問題は後にカントロビッチによって次のように拡張された．

問題 1.2 （モンジュ・カントロビッチの問題） 空間 X とコスト $c: X\times X \longrightarrow [0,\infty)$ が与えられたとする．このとき，X 上の２つの確率測度 μ,ν に対して，$X\times X$ 上の確率測度 π で各成分への射影が μ,ν であるもの（μ と ν の**カップリング**という）の中で，

$$\int_X c(x,y)\,d\pi(x,y)$$

を最小にするものを見つけよ．

*sohta@math.kyoto-u.ac.jp

モンジュの問題における写像 Ψ からは，$\pi = (\mathrm{Id}_X \times \Psi)_\sharp \mu$ として μ と ν のカップリングが得られる．よって，かかるコストの最小値はモンジュ・カントロビッチの問題の方が小さい．

コストとしては時間，費用，エネルギーなど様々なものが考えられるが，ここでは最も基本的な状況である距離空間 (X, d) と d から定まるコスト $c = d^2$（または $d^2/2$）のみを扱う．このとき，コストの最小値を μ と ν の間の距離と考えるのが，Wasserstein 空間である（Kantorovich-Rubinstein 空間ともいう）．距離空間 (X, d) 上のボレル確率測度全体を $\mathcal{P}(X)$ で表し，$\int_X d(x,y)^2\, d\mu(y) < \infty$ をある点 $x \in X$ で満たす $\mu \in \mathcal{P}(X)$ のなす部分集合を $\mathcal{P}_2(X)$ とおく．

定義 1.3 **（Wasserstein 空間）** $\mu, \nu \in \mathcal{P}_2(X)$ の間の $(L^2\text{-})$**Wasserstein 距離**を

$$d_2^W(\mu, \nu) := \inf_\pi \left(\int_{X \times X} d(x,y)^2\, d\pi(x,y) \right)^{1/2}$$

で定義する．ただし，下限は μ と ν のカップリング π 全体でとる．Wasserstein 距離を実現するカップリングを μ と ν の**最適カップリング**という．距離空間 $(\mathcal{P}_2(X), d_2^W)$ を X 上の $(L^2\text{-})$**Wasserstein 空間**という．

$(\mathcal{P}_2(X), d_2^W)$ の構造は (X, d) の構造と密接に関係する．例えば，$(\mathcal{P}_2(X), d_2^W)$ のコンパクト性は (X, d) のコンパクト性と同値である．

2　測地線とリーマン構造

この節では，簡単のため $X = \mathbb{R}^n$ の場合に限って，Wasserstein 空間の測地線の様子とそれを用いたリーマン構造の導入を述べる．

定理 2.1 **([Br])** $\mu, \nu \in \mathcal{P}_2(\mathbb{R}^n)$ で μ はルベーグ測度に絶対連続なものを考える．このとき，可測写像 $\Psi : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ で $\Psi_\sharp \mu = \nu$ かつ

$$\left(\int_{\mathbb{R}^n} d\big(x, \Psi(x)\big)^2\, d\mu(x) \right)^{1/2} = d_2^W(\mu, \nu)$$

なるものが（測度零の集合上の差を除いて）一意に存在する．さらに，Ψ はある凸関数 $f : \mathbb{R}^n \longrightarrow \mathbb{R}$ を用いて $\Psi(x) = \nabla f(x)$ と表せる．

つまり，Ψ（及び $\pi = (\mathrm{Id}_{\mathbb{R}^n} \times \Psi)_\sharp \mu$）はモンジュの問題とモンジュ・カントロビッチの問題双方の解を同時に与える．さらに，曲線 $\mu(t) = \{(1-t)\,\mathrm{Id}_X + t\nabla f\}_\sharp \mu$ $(t \in [0,1])$ は Wasserstein 空間での μ から ν への唯一つの最短測地線を与える．また逆に，任意の絶対連続な確率測度 μ と凸関数 $f : \mathbb{R}^n \longrightarrow \mathbb{R}$ に対し，$\Psi = \nabla f$ は μ から $\Psi_\sharp \mu$ への最適輸送である．

注意 2.2 定理 2.1 は McCann [Mc] によってリーマン多様体（上のコンパクトな台をもつ確率測度）に拡張された．McCann の手法は汎用性の高いものであり，同様の議論は曲率を下から押さえたアレクサンドロフ空間やフィンスラー多様体にも適用できる．

定理 2.1 では $\nabla f(x)$ を $\mathbb{R}^n$ の元と見なしているが，これは

$$\begin{aligned}\nabla f(x) = x + [\nabla f(x) - x] &= x + \nabla\left[f - \frac{|\cdot|^2}{2}\right](x) \\ &= \exp_x\left(\nabla\left[f - \frac{|\cdot|^2}{2}\right](x)\right)\end{aligned}$$

と書き直すことができる．すると，必要なのは $\varphi + |\cdot|^2/2$ が凸になるような関数 φ $(= f - |\cdot|^2/2)$ であり，特に任意の C^2 関数 φ に対し，$\varepsilon > 0$ を十分小さくとれば $\varepsilon\varphi$ はこの条件を満たす．これらを踏まえ，Otto [Ot] は $(\mathcal{P}_2(\mathbb{R}^n), d_2^W)$ の（形式的）リーマン構造を次のように導入した：$\mu \in \mathcal{P}_2(\mathbb{R}^n)$ での接空間を

$$T_\mu \mathcal{P}_2 := \overline{\{\nabla\varphi \,|\, \varphi \in C_c^\infty(\mathbb{R}^n)\}},$$

内積を

$$\langle \nabla\varphi, \nabla\psi \rangle_\mu := \int_{\mathbb{R}^n} \langle \nabla\varphi, \nabla\psi \rangle \, d\mu$$

で定義する．ただし，閉包は内積についてとっている．

このリーマン構造から定まる距離は Wasserstein 距離と一致する．また，この空間は平坦ではないが非負曲率を持つことが形式的な計算によって示される（[Lo] も参照）．この事実は，距離空間の幾何学の言葉を通して，次の命題によって説明される．距離空間の任意の 2 点が距離を実現する曲線で結べるとき，それを**測地空間**という．

命題 2.3 ([LV1], [St1]) コンパクト測地空間 (X, d) に対し，(X, d) が非負曲率アレクサンドロフ空間であることは，その上の Wasserstein 空間 $(\mathcal{P}_2(X), d_2^W)$ がそうであることと同値である．一方，(X, d) が非負曲率アレクサンドロフ空間でなければ，$(\mathcal{P}_2(X), d_2^W)$ の曲率は負定数でも下から押さえることはできない（つまり，$-\infty$ となる点がある）．

ここでアレクサンドロフ空間とは，リーマン幾何における三角形の比較定理を用いる一般化された断面曲率の下限（や上限）の概念である．命題の後半の主張は，負に曲がった三角形がひとつでもあると，それをスケーリングして曲率のいくらでも小さい三角形が作れることによる．しかし，負定数で曲率を下から押さえたアレクサンドロフ空間であっても，スケーリングで不変な別の条件（バナッハ空間論における一様平滑性の一般化）を用いることにより，その上の Wasserstein

空間がリーマン構造をもつことが示される（[Oh1]）．その議論では，Otto の構成とは異なり，下のアレクサンドロフ空間の構造は使わずに Wasserstein 空間の三角形の性質のみを用いる．

3　曲率次元条件とリッチ曲率

(M,g) をコンパクト n 次元リーマン多様体，m をその上の体積要素とする．確率測度 $\mu \in \mathcal{P}_2(X)$ の**相対エントロピー**を

$$\mathrm{Ent}_m(\mu) := \begin{cases} \int_M \rho \log \rho \, dm & \text{if } \mu = \rho m, \\ \infty & \text{o.w.} \end{cases}$$

と定める．von Renesse と Sturm は，[CMS] や [OV] の結果をもとに，次を示した．

定理 3.1 **（[vRS]）** (M,g,m) をコンパクトリーマン多様体とその上の体積要素，K を実数とする．このとき，$(\mathcal{P}_2(X), d_2^W)$ 内の全ての測地線上で Ent_m の（弱い意味での）２階微分が K 以上であることは，M のリッチ曲率が K 以上であることと同値である．

より一般に，重み $V \in C^\infty(M)$ つきの場合に，

$$f(\mu) = \mathrm{Ent}_{e^{-V}m}(\mu) = \mathrm{Ent}_m(\mu) + \int_M V \, d\mu \tag{3.1}$$

の２階微分が K 以上であることは **Bakry-Émery テンソル** $\mathrm{Ric} + \mathrm{Hess}\, V$（[BE]）が K 以上であることと同値である．Bakry-Émery テンソルはある意味で無限次元のリッチ曲率と思えることが知られている（例えば，ユークリッド空間のガウス分布はこの意味で非負リッチ曲率をもつ）．Qian [Qi] によって導入された N $(\in [n,\infty))$ 次元のリッチ曲率についても，それが K 以上であることはある種のエントロピーの凸性と同値である（[St2]，[LV2] 参照）．

注意 3.2 これらの同値性は，フィンスラー多様体にも適切なリッチ曲率の概念を導入することで拡張される（[Oh2]）．

これらを踏まえて，Sturm [St1], [St2] と Lott と Villani [LV1], [LV2] は，一般の測度距離空間でのリッチ曲率の下限にあたる**曲率次元条件**を定義した．$N \in [0,\infty]$ と $K \in \mathbb{R}$ に対し，測度距離空間 (X,d,m) が $\mathsf{CD}(K,N)$ を満たすとは，上で述べた（リーマン多様体では N 次元リッチ曲率が K 以上であることと同値である）ある種のエントロピーの凸性で定義される．この条件は空間の測度つきグロモフ・ハウスドルフ収束で保たれる．また，曲率次元条件を満たす空間はリッチ曲率を下から押さえたリーマン多様体と同様の幾何学的な性質を持ち（ビショップ・グ

ロモフの体積比較定理，ボンネ・マイヤースの定理などが成立），種々の関数不等式（対数ソボレフ不等式，リヒネロビッツの不等式など）や測度の集中現象を満たす．さらに，一般化されたブルン・ミンコフスキの不等式などの，リーマン多様体でもこの手法を用いて初めて示された結果もある．

4　勾配流と熱方程式

Jordan, Kinderlehrer と Otto [JKO] は，離散近似を用いて，ユークリッド空間上の Wasserstein 空間内の相対エントロピーの勾配流（の密度関数）は熱流と一致することを示した．より一般に $V \in C_c^\infty(\mathbb{R}^n)$ に対し，(3.1) で定義した f の勾配流（の密度関数）は方程式

$$\frac{\partial \rho}{\partial t} = \Delta \rho + \mathrm{div}(\rho \nabla V)$$

の解と一致する．これらの結果は，前々節で述べた Wasserstein 空間のリーマン構造を用いたより厳密な手法によって，コンパクトリーマン多様体に拡張される（[Oh1]）．また，フィンスラー多様体にも，[AGS] や [Vi2] のような下の空間の構造を使う方法で拡張できる（[OS]）．

これを前節の結果と合わせると，リッチ曲率が K 以上のコンパクトリーマン多様体の熱核の収縮性

$$d_2^W\big(p(t,x,\cdot)m, p(t,y,\cdot)m\big) \le e^{-Kt} d(x,y)$$

の別証明が得られる（より詳細な結果は，[vRS] など参照）．ここでは，Wasserstein 空間がリーマン構造をもっていることが本質的である．なぜなら，収縮性の証明には（勾配流を生成する関数及び距離関数の）第一変分公式を用いるため,「角度」を本質的に使う必要がある．実際，バナッハ空間では凸関数の勾配流の収縮性すら知られていない（[AGS] 参照）．

参考文献

[AGS] L. Ambrosio, N. Gigli and G. Savaré, Gradient flows in metric spaces and in the space of probability measures, Birkhäuser Verlag, Basel, 2005.

[BE] D. Bakry and M. Émery, *Diffusions hypercontractives* (French), Séminaire de probabilités, XIX, 1983/84, 177–206, Lecture Notes in Math., 1123, Springer, Berlin, 1985.

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

[CMS] D. Cordero-Erausquin, R. J. McCann and M. Schmuckenschläger, *A Riemannian interpolation inequality á la Borell, Brascamp and Lieb*, Invent. Math. **146** (2001), 219–257.

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

[Lo] J. Lott, *Some geometric calculations on Wasserstein space*, Comm. Math. Phys. **277** (2008), 423–437.

[LV1] J. Lott and C. Villani, *Ricci curvature for metric-measure spaces via optimal transport*, to appear in Ann. of Math.

[LV2] J. Lott and C. Villani, *Weak curvature conditions and functional inequalities*, J. Funct. Anal. **245** (2007), 311–333.

[Mc] R. J. McCann, *Polar factorization of maps on Riemannian manifolds*, Geom. Funct. Anal. **11** (2001), 589–608.

[Oh1] S. Ohta, *Gradient flows on Wasserstein spaces over compact Alexandrov spaces*, to appear in Amer. J. Math.

[Oh2] S. Ohta, *Finsler interpolation inequalities*, preprint (2008).

[OS] S. Ohta and K.-T. Sturm, *Heat flow on Finsler manifolds*, in preparation.

[Ot] F. Otto, *The geometry of dissipative evolution equation*: *the porous medium equation*, Comm. Partial Differential Equations **26** (2001), 101–174.

[OV] F. Otto and C. Villani, *Generalization of an inequality by Talagrand and links with the logarithmic Sobolev inequality*, J. Funct. Anal. **173** (2000), 361–400.

[Qi] Z. Qian, *Estimates for weighted volumes and applications*, Quart. J. Math. Oxford Ser. (2) **48** (1997), 235–242.

[vRS] M.-K. von Renesse and K.-T. Sturm, *Transport inequalities, gradient estimates, entropy and Ricci curvature*, Comm. Pure Appl. Math. **58** (2005), 1–18.

[St1] K.-T. Sturm, *On the geometry of metric measure spaces*, Acta Math. **196** (2006), 65–131.

[St2] K.-T. Sturm, *On the geometry of metric measure spaces.* II, Acta Math. **196** (2006), 133–177.

[Vi1] C. Villani, Topics in optimal transportation, American Mathematical Society, Providence, RI, 2003.

[Vi2] C. Villani, Optimal transport, old and new, to appear.