Gauss 測度族の Wasserstein 幾何

高津飛鳥（東北大理）
記：野田知宣（大歯大）

1. 主結果

d 次対称行列全体を $\mathrm{Sym}(d,\mathbb{R})$ で表し、特に正定値のものを $\mathrm{Sym}^+(d,\mathbb{R})$ で表す。$m \in \mathbb{R}^d$, $V \in \mathrm{Sym}^+(d,\mathbb{R})$ に対し密度が

$$\left(\frac{1}{2\pi}\right)^{d/2} \frac{1}{\sqrt{\det V}} \exp\left(-\frac{1}{2}\langle x-m, V^{-1}(x-m)\rangle\right)$$

で与えられる $\mathbb{R}^d$ 上の測度を **Gauss 測度**と呼び $N(m,V)$ で表す。（m は平均、V は共分散行列。）

Gauss 測度の族を $\mathcal{N}^d$ とする。このとき

$$\begin{aligned}\mathcal{N}^d &= \{N(m,V)\ ;\ m \in \mathbb{R}^d,\ V \in \mathrm{Sym}^+(d,\mathbb{R})\} \\ &\cong \mathbb{R}^d \times \mathrm{Sym}^+(d,\mathbb{R})\end{aligned}$$

であり、これは $d+\frac{1}{2}d(d+1)$ 次元多様体である。$\mathcal{N}^d$ 上の Riemann 計量としては後述の Wasserstein 計量を用いる。

次に $\mathcal{N}^d$ に対する指数写像を定義する。$\rho := N(m,V) \in \mathcal{N}^d \cong \mathbb{R}^d \times \mathrm{Sym}^+(d,\mathbb{R})$ における $\mathcal{N}^d$ の接空間と $\mathbb{R}^d \times \mathrm{Sym}(d,\mathbb{R})$ とを同一視する：$T_\rho\mathcal{N}^d \cong \mathbb{R}^d \times \mathrm{Sym}(d,\mathbb{R})$. いま $\xi \in \mathbb{R}^d$, $X \in \mathrm{Sym}(d,\mathbb{R})$ とする。このとき指数写像 $T_\rho\mathcal{N}^d \to \mathcal{N}^d$ を

$$\exp_\rho t(\xi,X) = N(l(t),W(t)), \qquad \begin{cases} l(t) = m + t\xi, \\ W(t) = (E+tX)V(E+tX) \end{cases}$$

で定める。ここで E は単位行列。（一般に、接空間とその上の指数写像の定義される length space を Riemannian length space と呼ぶ。詳しくは [2] 参照。）ここで定義した指数写像の像は Wasserstein 計量に関し測地線となる事に注意する。

本講演の主結果は次である。

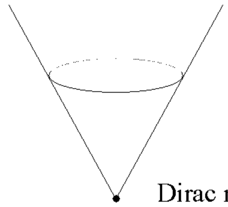
主定理. $\mathcal{N}^d$ 上の Wasserstein 計量の断面曲率を K とする。$\rho \in \mathcal{N}^d$, $\xi,\eta \in \mathbb{R}^d$, $X,Y \in \mathrm{Sym}(d,\mathbb{R})$ に対し

・$K_\rho(\xi,\eta) = K_\rho(\xi,X) = 0$;

・$K_\rho(X,Y)q_\rho(X,Y) = \frac{3}{4}\mathrm{tr}([Y,X]-Z)V^t([Y,X]-Z)$.

ここで $q_\rho(X,Y) := \mathrm{tr}(XVY)\mathrm{tr}(YVY) - (\mathrm{tr}(XVY))^2$ であり、$Z \in \mathrm{Sym}(d,\mathbb{R})$ は任意の $x \in \mathbb{R}^d$ に対し $\langle x, ([Y,X]-Z)Vx\rangle = 0$ を満たすもの。

注意. i) $K_\rho \geq 0$ であり、$\mathcal{N}^d$ は cone の構造を持つ。しかし完備でない。



ii) Otto [6] において $\mathrm{Diff}(\mathbb{R}^d)$ から $\mathcal{P}_2^{\mathrm{ac}}(\mathbb{R}^d)$ への isometric submersion と $\mathrm{Diff}(\mathbb{R}^d)$ 上の平坦計量を用い、$\mathcal{P}_2^{\mathrm{ac}}(\mathbb{R}^d)$ の断面曲率の explicit expression が与えられた。更に O'Neill の公式 [5] を用い $\mathcal{P}_2^{\mathrm{ac}}(\mathbb{R}^d)$ は nonnegative sectional curvature である事が示された。上の主定理はこの Otto による結果の厳密化である。

2.

定義 2.1. $\mathcal{P}_2$ で $\mathbb{R}^d$ 上の確率測度で 2 次モーメントの存在するもの全体を表す。即ち

$$\mathcal{P}_2 := \{\mu : \mathbb{R}^d \text{ の確率密度 };\ \int_{\mathbb{R}^d} |x|^2 d\mu(x) < +\infty\}.$$

任意の $\mu, \nu \in \mathcal{P}_2$ に対し

$$W_2(\mu,\nu)^2 := \inf_{\pi \in \Pi} \int_{\mathbb{R}^d \times \mathbb{R}^d} d(x,y)^2\ d\pi(x,y)\ \ < +\infty$$

を **Wasserstein 距離**と呼ぶ。ここで Π は周辺分布が μ, ν であるような $\mathbb{R}^d \times \mathbb{R}^d$ 上の確率測度の空間を表す。即ち π は任意の $A \subset \mathbb{R}^d$ に対し

$$\pi(A \times \mathbb{R}^d) = \mu(A),\ \pi(\mathbb{R}^d \times A) = \nu(A)$$

を満たす $\mathbb{R}^d \times \mathbb{R}^d$ 上の確率測度。また $d(x,y)$ は $\mathbb{R}^d$ の Euclid 距離。

このとき $(\mathcal{P}_2, W_2)$ は距離空間となり、これを $(L^2\text{-})$Wasserstein 空間と呼ぶ。

注意 2.2. (i) 背景は Monge-Kantorovich 問題。

(ii) これらは完備可分距離空間上で定義される。

(iii) $(\mathcal{P}_2, W_2)$ の代わりに $(\mathcal{P}_r, W_r)$, $r \in [1, +\infty)$, でも同じ事が云える。

(iv) inf は常に存在する ([10]) : $\mu \in \mathcal{P}_2^{\mathrm{ac}} := \{\mu \in \mathcal{P}_2\ ;\ \mu \ll dx\}$ のとき minimizer は push-forward measure で書かれる ([1],[4])。但し $\mu \ll dx$ は μ が Lubesgue 測度 dx に関し絶対連続（absolutely continuous）を表す。

これらについては 4 節 参照。

Monge-Kantorovich 問題に対し次が成立する。

定理 2.3 (Brenier[1], McCann[4])**.** $\mathbb{R}^d$（resp. (M,g) : 連結 Riemann 多様体, $\partial M = \emptyset$), $\mu, \nu \in \mathcal{P}_2$, $\mu \ll dx$（resp. $\mu \ll d\mathrm{vol}_g$）とする。このとき $\psi : \mathbb{R}^d \to \mathbb{R}$ は次を満たす :

(i) $\frac{1}{2}|\cdot|^2 + \psi$ は convex（resp. $-\psi$ は $\frac{d^2}{2}$-concave）;

(ii) $^{\exists!}\psi$ up to additive constant;

(iii) $T_t(x) = x + t\nabla\psi(x)$ (resp. $T_t(x) = \exp_x t\nabla\psi(x)$), $(\mathrm{id}\times T)_\sharp\mu$ は minimizer.
特に $T_\sharp\mu = \nu$. 更に $\{T_{t\sharp}\mu\}_{t\in[0,1]}$ は μ から ν への測地線。

熱方程式は通常 L^2 上のエネルギーの勾配流と考えられていた。Otto は $(\mathcal{P}_2^{\mathrm{ac}}, W_2)$ を Riemann 多様体とみなし、熱方程式をエントロピーの勾配流とみなした：$\mathcal{P}_2^{\mathrm{ac}}$ 上で指数写像を

$$\exp_\rho t\nabla\psi(x) := [\mathrm{id} + t\nabla\psi]_\sharp\rho$$

とするとこれは測地線であり、

$$g_\rho(t\nabla\psi, t\nabla\psi) := W_2(\rho, \exp_\rho t\nabla\psi(x))^2 = \int_{\mathbb{R}^d} t^2|\nabla\psi(x)|^2\rho(x)dx,$$

$$(\text{すなわち、 } g_\rho(\nabla\psi, \nabla\psi) = \int_{\mathbb{R}^d} |\nabla\psi^2|^2\rho(x)dx\ )$$

$$T_\rho\mathcal{P}_2^{\mathrm{ac}} := \{\psi\ ;\ \int_{\mathbb{R}^d} |\nabla\psi|^2\rho(x)dx < +\infty\}$$

とすると、接空間と Riemann 計量が与えられた（特にこの計量は W_2 を誘導する）。エントロピー

$$E(\rho) = \int_{\mathbb{R}^d} \rho\log\rho dx$$

に対し $E(\rho)$ の勾配流は $\frac{\partial\rho}{\partial t} = \Delta\rho$ を満たす。

$p_t(x,\cdot)$ を熱核とする。このとき次が成立する。

定理 2.4 (Sturm and von Renesse [9])**.** (M, g) を滑らかな連結 Riemann 多様体、$K \in \mathbb{R}$ とする。このとき次は同値：
(i) $\mathrm{Ric}(M) \geq K$;
(ii) $W_2(p_t(x,\cdot), p_t(y,\cdot)) \leq e^{-K}W_2(p_0(x,\cdot), p_0(y,\cdot)) = e^{-K}d(x,y)$.
($x \in M$ に対し $\delta_x \in \mathcal{P}_2$ で定まる写像 $M \to \mathcal{P}_2$ は等長的：$W_2(\delta_x, \delta_y) = d(x,y)$.)

Sturm and von Renesse [9] では Riemann 多様体上の Ricci 曲率の lower bound（Ric$\geq K$）とエントロピーの凸性（$\mathcal{P}_2$ 上のエントロピーは K-convex）の同値性を示している。熱核を $(\mathcal{P}_2, W_2)$ 上のエントロピーの勾配流と解釈する事で上の同値性を得る。またこの定理で $M = \mathbb{R}^d$ なら等号が成立する事を注意しておく。

$\mathbb{R}^d$ 上の Gauss 測度族 $\mathcal{N}^d$ に話を戻す。次が成立する。

定理 2.5 (McCann[3])**.** $\mathcal{N}^d \subset \mathcal{P}_2^{\mathrm{ac}}$ は全測地的。

これより $\rho := N(m, V) \in \mathcal{N}^d$ から出る測地線も Gauss 測度になる事が判る。いま

$$\exp_\rho(\nabla\psi(x)) = N(n, V)$$

と表し、

$$W = U^{1/2}(U^{1/2}VU^{1/2})^{-1/2}U^{1/2},$$

$$\psi(x) = \frac{1}{2}\langle x - m, (W - E)(x - m)\rangle + \langle x, n\rangle$$

と置くと

$$\begin{aligned} W_2(N(m,V),N(n,U)) &= |m-n|^2 + \mathrm{tr}((E-W)V(E-W)) \\ &= |m-n|^2 + \mathrm{tr}V + \mathrm{tr}U - 2\mathrm{tr}(U^{1/2}VU^{1/2})^{1/2} \end{aligned}$$

が成立する。

3. 主定理の概証

Otto の結果から次の図式を得る：

$$\begin{array}{ccccc}
\mathrm{Diff}(\mathbb{R}^d) \ni \Phi & & \xrightarrow{\quad\text{submersion}\quad} & & \Phi_\sharp \rho \in \mathcal{P}_2^{\mathrm{ac}} \\
\cup & & & & \cup \\
\mathrm{Gl}(d,\mathbb{R}) & \xrightarrow{\text{submersion}} & \mathrm{Sym}^+(d,\mathbb{R}) & \xrightarrow{\text{isometric}} & \mathcal{N}_0^d = \{N(0,V) \in \mathcal{N}^d\} \\
\Cup & & \Cup & & \Cup \\
A & \longrightarrow & V := A^tA & \longrightarrow & V_\sharp^{1/2} N(0,E) = N(0,V)
\end{array}$$

いま $G(X,Y) := \mathrm{tr}(X^tY)$ とすると $(\mathrm{Gl}(d,\mathbb{R}),G)$ は平坦。

ここで次を用いる：

定理 3.1 (O'Neill の公式 [5])**.** $\pi : (M,G) \to B$ を Riemannian 沈込とする。このとき、M 上の水平なベクトル場 X,Y (すなわち $X,Y \in \mathrm{Ker}^\perp d\pi$) に対し、$M$ の X,Y が張る接平面の断面曲率 K_M と B の $d\pi(X), d\pi(Y)$ が張る接平面の断面曲率 K_B には以下の関係がある:

$$K_B = K_M + \frac{3}{4} G([X,Y]_v,[X,Y]_v)/Q(X,Y).$$

ここで、$[X,Y]_v$ は B の接空間に対する $[X,Y]$ の垂直成分を表す。また、$Q(X,Y) = G(X,X)G(Y,Y) - G(X,Y)^2$ である。

これより求める結果を得る。 □

Dirac 測度 δ における tangent cone $\overline{\mathcal{N}^d}$ は曲率 ≥ 0 の Aleksandrov 空間。底空間と $\mathcal{P}_2$ の幾何構造に関し次が成立する。詳しくは [8] を見よ。

定理 3.2. X を cone の構造を持つ完備可分空間とすると、$\mathcal{P}_2(X)$ は cone の構造を持つ。

4. おまけ

ここでは 2 節の注意 2.2 に関し、若干の補足を記しておく。

(i) Monge-Kantorovich 問題は Monge による最適輸送問題（1781 年）の一般化である。面積の等しい（形は一般に異なる） 2 つの土地の一方の芝生を他方に移したい。芝生の移動には距離に依存するコストが掛かる。このときどのように芝生を移すのが最良かを問うのがオリジナルの Monge の問題である。現在では以下のように定式化したものを **Monge の問題**と呼ぶ。$D_1, D_2 \subset \mathbb{R}^d$ を $|D_1| = |D_2|$ である

（有界）領域、$c(x,y)$ で $x\in\mathbb{R}^d$ を $y\in\mathbb{R}^d$ に移動する際の単位（体積当りの）コストを表す。このとき $T(D_1)=D_2$ と任意の $U\subset D_1$ に対し $|T(U)|=|U|$ を満たす単射 $T:\mathbb{R}^d\to\mathbb{R}^d$ に対し

$$C(T)=\int_{D_1}c(x,T(x))dx$$

を最小にする T を求めよ。

$\mathcal{T}=\{T:\mathbb{R}^d\to\mathbb{R}^d\,;\,T$ は測度を保つ単射$\}$ の中でコスト $C(T)$ を最小にする T を探すのであるが、$\mathcal{T}$ はコンパクトでなく C は連続でない（Weierstrass の定理が使えない)。そこで次のような一般化を考える。

$T:\mathbb{R}^d\to\mathbb{R}^d$ のグラフを $G(T)$, μ を $\mathbb{R}^d$ 上の確率測度とする。$\mathrm{id}\times T: x\mapsto (x,T(x))\in G(T)$ による μ の像測度 $(\mathrm{id}\times T)_\sharp\mu$ は $G(T)$ の上の確率測度を定める（一般に $T_\sharp\mu(A):=\mu(T^{-1}(A))$ で定まる測度 $T_\sharp\mu$ を像測度と呼ぶ)。

$\mathbb{R}^d$ 上の確率測度 μ,ν に対し

$$\begin{aligned}\Pi(\mu,\nu):=\{\pi:\mathbb{R}^d\times\mathbb{R}^d\ \text{上の確率測度}\,;\\ {}^\forall A\subset\mathbb{R}^d\ \text{に対し}\ \pi(A\times\mathbb{R}^d)=\mu(A),\ \pi(\mathbb{R}^d\times A)=\nu(A)\}\end{aligned}$$

と置く。$T_\sharp\mu=\nu$ とすると $(\mathrm{id}\times T)_\sharp\mu\in\Pi(\mu,\nu)$ であり、これで定まる写像 $\mathcal{T}\to\Pi(\mu,\nu)$ は単射。更に

$$C(T)=\int_{\mathbb{R}^d\times\mathbb{R}^d}c(x,y)(\mathrm{id}\times T)_\sharp\mu$$

が成立する。よって $\Pi(\mu,\nu)$ 上の函数 I を

$$I(\pi):=\int_{\mathbb{R}^d\times\mathbb{R}^d}c(x,y)d\pi(x,y)$$

で定めると

$$\mathcal{T}\subset\Pi,\ I:\Pi(\mu,\nu)\to\mathbb{R},\ I|_{\mathcal{T}}=C$$

が成立する。よって Monge の問題を解くには $\Pi(\mu,\nu)$ で I を最小にする T を求め、その T が $\mathcal{T}$ に属する事を示せば良い。I を最小にする $\pi\in\Pi$ を探す問題を **Monge-Kantorovich 問題**と呼ぶ。Kantorovich は資源の最適配分に関する理論の研究で 1975 年にノーベル経済学賞を受賞している。

いま直積 $\mu\times\nu\in\Pi(\mu,\nu)$ が成立し、特に $\Pi(\mu,\nu)\neq\emptyset$. また weak convergence の位相で $\Pi(\mu,\nu)$ はコンパクトで $I:\Pi(\mu,\nu)\to\mathbb{R}$ は連続。よって I を最小にする $\pi_0\in\Pi$ が存在する。

特に $c(x,y)=d(x,y)^2$（Euclid 距離の自乗）で μ ,ν の 2 次モーメントが存在すれば $\pi_0=(\mathrm{id}\times T_0)_\sharp\mu$ と表せ $T_0\in\mathcal{T}$ は Monge の問題の解を与える（Brenier)。

(ii) (M,g) を連結 Riemann 多様体、$d(x,y)$ を Riemann 距離函数とする（一般に完備可分距離空間で良い)。M 上の測度 μ, ν に対し (i) と同様に $\Pi(\mu,\nu)$ を考える。

このとき $r \in [1,\infty)$ に対し

$$W_r(\mu, r)^r := \int_{\pi\in\Pi} \left(\int_{M\times M} d(x,y)^r \ d\pi(x,y) \right),$$

$$\mathcal{P}_r(M) := \{\mu : \text{prob. measure} \ ; \ \int_M d(x,y)^r d\mu(y) < +\infty, \ ^\forall x \in M\}$$

とすると $(\mathcal{P}_r, W_r)$ は geodesic space.

References

[1] Y. Brenier, *Polar factorization and monotone rearrangement of vector-valued functions*, Comm. Pure Appl. Math. **44** (1991), no. 4, 375–417.

[2] J.A. Carrillo, R.J. McCann, and C. Villani, *Contractions in the 2-Wasserstein length space and thermalization of granular media*, Arch. Ration. Mech. Anal. **179** (2006), no. 2, 217–263.

[3] R.J. McCann, A convexity principle for interacting gases, Adv. Math. 128 (1997), no. 1, 153.179.

[4] R.J. McCann, *Polar factorization of maps on Riemannian manifolds*, Geom. Funct. Anal. **11** (2001), no. 3, 589–608.

[5] B. O'Neill, The fundamental equations of a submersion, Michigan Math. J. 13 (1966), 459.469.

[6] F. Otto, The geometry of dissipative evolution equations: the porous medium equation, Comm. Partial Differential Equations 26 (2001), no. 1-2, 101.174.

[7] Asuka Takatsu, On Wasserstein geometry of the space of Gaussian measures, arXiv:0801.2250 [math.DG].

[8] Asuka Takatsu, Cone structure of L^2-Wasserstein spaces, arXiv:0812.2752 [math.MG].

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

[10] C. Villani, *Optimal transport, new and old*, Grundlehren der mathematischen Wissenschaften, vol. 338, Springer, Berlin, 2008.