最適輸送理論とリッチ曲率

～物を運ぶと曲率が分かる～

2015 年 2 月 20 日,21 日

目 次

1　モンジュの問題とモンジュ・カントロヴィッチ問題 (桑江)

本稿では最適輸送理論の歴史的経緯を交えながら基本となる考え方や方法を伝えたい. より詳しくは C. Villani 氏や Ambrosio-Gigli-Savaré 諸氏による教科書 [143, 144, 4] を, 日本語による解説では會田茂樹氏による講演スライド [147] や三上敏夫氏や太田慎一氏による論説 [153, 148, 149] を参照されたい.

1.1　モンジュの問題とモンジュ・カントロヴィッチ問題の歴史

次の問題を考える.

問題 1.1 ([103]) ある砂山をそれと同じ体積の穴に移したい. 砂粒一つ一つの移動には移動距離に依存したコストがかかるとき, 最適な移動のさせ方は何か？

この問題は仏の数学者・工学者である Gaspard Monge の 1781 年の論文 [103] の中で提唱され, モンジュ [1]の最適輸送問題, 略してモンジュの問題と呼ばれています. モンジュはフランス革命時に海軍大臣を務めたこともあり, モンジュの問題は物資の輸送におけるコストの節約の観点から現実的な要請が背景にあったと考えられます.

さて問題の文面のままでは何が問題なのか分りにくいですが, 実は砂山の砂粒はそれを移動させたら空中 (地中？) で静止してそこに留まっていることを暗に仮定しています. 解りやすい形に言い換えると次のようになります.

問題 1.2 n 個の工場と n 個の店舗があり, 各工場から各店舗にそれぞれ一個の製品のみを移動させ距離に応じてコストがかかるとき, 総コストを最小にする移動のさせ方は何か？

問題 1.2 [2]の解は工場と店の具体的な配置から決まります. しかも移動のさせ方は $n!$ 通りしかないので計算して具体的な解を (しらみつぶしかコンピュータで) 求めることが原理的に可能です. 最初の問題 1.1 は問題 1.2 において工場の個数が無限でしかも至る所密に詰っていて, さらに移動先の店舗も無限で密に詰っている状況の問題と解釈することができます. さらに一歩進めて「密に詰った無限の各工場から密に詰った無限の各店舗に総コストが最小になるような 1 対 1 上への写像を決める問題」と理解できます. 問題 1.1を数学的に述べると以下のようになります.

[1]モンジュはラプラス, フーリエと並んでナポレオンに仕えた数学者の一人である. 真面目で正義感の強い人物だったが, それが災いしてナポレオンを最後まで信奉したため悲惨な末路を迎えた逸話が残っている.

[2]この問題は, n 個の工場と m 個の店舗を考え, P_i が i 番目の工場の供給製品数, Q_j が j 番目の店舗の必要製品数を表し, X_{ij} を i 番目の工場の製品数 P_i のうち j 番目の店舗へ配送する供給数, その移動コストを C_{ij} として $\sum_{i=1}^{n} P_i = \sum_{j=1}^{m} Q_j = \sum_{i=1}^{n}\sum_{j=1}^{m} X_{ij}$ の条件下で総コスト $C(X) = \sum_{i=1}^{n}\sum_{j=1}^{m} X_{ij}C_{ij}$ を最小にする配送数行列 $X = (X_{ij})$ を求めよという問題に一般化できる. これをヒッチコック型輸送問題という. 問題 1.2は $n = m$ で C が単位行列の場合を考えている.

問題 1.3 D_1, D_2 は $\mathbb{R}^3$ の部分集合で同じ体積 1 をもつとする．位置 x の砂を位置 y に移動するのに要する価格は単位体積あたり $c(x,y)$ [3] だけかかるとする．D_1 から D_2 への写像 T で次の条件 (1), (2) を満たすものを考える:

(1) T は全単射である．すなわち $x \neq x'$ ならば $T(x) \neq T(x')$, かつ D_1 の T による像 $T(D_1) = \{T(x) \mid x \in D_1\}$ は D_2 である．

(2) D_1 の任意の部分集合 U に対してその像 $T(U)$ と U の体積は同じである．

これらの条件下で $C(T) := \int_{D_1} c(x, T(x)) \mathrm{d}x$ を最小にする写像 T を求めよ．

問題 1.3はたいへんな難問で解決されるまで長い年月がかかりました．Sudakov [135, 136] が 1979 年にモンジュの問題の解である最適輸送写像 T の存在証明を発表しましたが, それには一部誤りがありました．修復された証明は測度の台が互いに素でコンパクトという強い仮定の下で Evans-Gangbo [49] (1999), 密度をもつ仮定の下で Ambrosio [2] (2000), Trudinger-Wang [139] (2001), Caffarelli-Feldman-McCann [23] (2002) [4]等によりなされました．

1979 年の Sudakov [135, 136] 以前にこの問題に関して重要なアイディアを提出したのがロシアの数学者・経済学者 Leonid Kantorovich(レオニード・カントロヴィッチ)(1912-1986) です．カントロビッチはモンジュの問題を, 写像を決めるのでなく,「工場の散らばり (確率分布 μ) と店舗の散らばり (確率分布 ν) が与えられているときに x から y への移動コスト $c(x,y)$ の総コスト

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

が最小になるような工場と店舗の配置の結合分布 π(輸送計画) を決めよ」という問題に置き換えました．これをモンジュ・カントロビッチ問題 (以下 MK 問題) と呼びます．MK 問題はモンジュの問題よりは解決が数学的には易しく, モンジュの問題の解決に向けて大きな進展が得られました．カントロビッチは線形計画法 [5]において先駆的な業績を挙げており, 資源の適正配分に関する一連の研究により 1975 年にノーベル経済学賞を受賞しています．MK 問題の解から対応するモンジュの問題の解 T の満たすべき性質が予想できます．そのことに基づいてコスト関数 $c(x,y)$ が良い条件を満たすとき初めて厳密に解いたのが Yann Brenier です (1987 年).

定理 1.4 (Yann Brenier(1987), cf. [19, 20]) μ, ν は $\mathbb{R}^n$ 上の連続型確率分布で，2 次の積率をもつとし, μ は n-次元ルベーグ測度に関して絶対連続なものとする．コスト関数 $c(x,y) = \frac{1}{2}|x-y|^2$ に対する MK 問題の解 $\pi \in \Pi(\mu,\nu)$ に対して $\pi(A \times B) = \mu(A \cap T^{-1}(B))$,

[3] $c(x,y)$ はここでは必ずしも $\mathbb{R}^3$ のユークリッド距離 $|x-y|$ に依存しているとは限らないものとする．
[4] Brenier の定理の結果と異なり, 最適輸送写像の一意性は成立しない．
[5] ジョージ・ダンツィーグ (George Dantzig 1914–2005) が 1948 年に線形計画法を確立した．

$\forall A, B \in \mathcal{B}(\mathbb{R}^n)$ [6]を満たす写像 $T:\mathbb{R}^n \to \mathbb{R}^n$ が存在する．さらに凸関数 $\phi:\mathbb{R}^n \to \mathbb{R}$ で $T = \nabla\phi$ μ-a.s. をみたすものが存在する．ここで

$$\Pi(\mu,\nu) := \{\pi \in \mathcal{P}(\mathbb{R}^n \times \mathbb{R}^n) \mid \pi(A\times\mathbb{R}^n) = \mu(A), \pi(\mathbb{R}^n \times B) = \nu(B)\ \forall A, B \in \mathcal{B}(\mathbb{R}^n)\}$$

は μ, ν の結合分布 (カップリング) と呼ばれる $\mathbb{R}^n \times \mathbb{R}^n$ 確率測度の全体で, $\mathcal{P}(\mathbb{R}^n \times \mathbb{R}^n)$ は $\mathbb{R}^n \times \mathbb{R}^n = \mathbb{R}^{2n}$ 上のボレル確率測度の全体である.

証明　(定理 1.4の証明の概略) $\pi_0 \in \Pi(\mu,\nu)$ を

$$D^2[\pi] := \int_{\mathbb{R}^n\times\mathbb{R}^n} |x-y|^2 \pi(\mathrm{d}x\mathrm{d}y), \quad \pi \in \Pi(\mu,\nu)$$

を最小にする輸送計画とする (後述の定理 1.6からこの最小点 $\pi_0 \in \Pi(\mu,\nu)$ は存在する). すると任意の $x, y \in \mathbb{R}^n$ で $\varphi(x) + \psi(y) \le \frac{1}{2}|x-y|^2$ をみたす有界連続関数 $\varphi, \psi \in C_b(\mathbb{R}^n)$ [7]について

$$\begin{aligned}\frac{1}{2}D^2[\pi] = \frac{1}{2}\int_{\mathbb{R}^n\times\mathbb{R}^n} |x-y|^2\pi(\mathrm{d}x\mathrm{d}y) &\ge \int_{\mathbb{R}^n\times\mathbb{R}^n} (\varphi(x)+\psi(y))\pi(\mathrm{d}x\mathrm{d}y) \\ &= \int_{\mathbb{R}^n} \varphi(x)\mu(\mathrm{d}x) + \int_{\mathbb{R}^n} \psi(y)\nu(\mathrm{d}y)\end{aligned}$$

となって, これより

$$\begin{aligned}\tfrac{1}{2}\inf_{\pi\in\Pi(\mu,\nu)} D^2[\pi] \ge \sup\ &\left\{\int_{\mathbb{R}^n}\varphi(x)\mu(\mathrm{d}x) + \int_{\mathbb{R}^n}\psi(y)\nu(\mathrm{d}y)\ \middle|\right. \\ &\left.\quad \varphi,\psi\in C_b(\mathbb{R}^n),\ \varphi(x)+\psi(y) \le \tfrac{1}{2}|x-y|^2 \quad {}^{\forall}x, y \in \mathbb{R}^n\right\}\end{aligned} \tag{1.1}$$

を得る．右辺を最大化する関数の組 $(\varphi,\psi) \in L^1(\mu)\times L^1(\nu)$ [8]で, 任意の $x, y \in \mathbb{R}^n$ で $\varphi(x)+\psi(y) \le \frac{1}{2}|x-y|^2$ を満たし, かつ

$$\pi_0\text{-a.s. } (x,y) \in \mathbb{R}^n \times \mathbb{R}^n \text{ で } \varphi(x)+\psi(y) = \frac{1}{2}|x-y|^2 \tag{1.2}$$

をみたすものが存在することが (μ, ν の２次の積率有限性から後述の定理 1.24(2) が適用できて) わかり, (1.1) で等号が成立する (Kantorovich 双対性, 定理 1.24(3)). 特に (1.2) から μ-a.s. $x \in \mathbb{R}^n$ に対し, $\varphi(x)+\psi(y) = \frac{1}{2}|x-y|^2$ をみたす $y \in \mathbb{R}^n$ が存在する. 実際,

$$A := \left\{x \in \mathbb{R}^n \ \middle|\ \varphi(x)+\psi(y) = \frac{1}{2}|x-y|^2 \text{ を満たす } y \text{ が存在する}\right\}$$

[6] これは T のグラフ $G(T) := \{(x, T(x)) \in \mathbb{R}^{2n} \mid x \in \mathbb{R}^n\}$ に対して $\pi(G(T)) = 1$ と同値である.

[7] $C_b(\mathbb{R}^n)$ は $\mathbb{R}^n$ 上の有界連続関数の全体.

[8] この最大化する関数の組 (φ,ψ) の φ を **Kantorovich potential** という．また ψ はこの場合, $\psi(y) := \varphi^c(y) := \inf_{z\in\mathbb{R}^n}\left\{\frac{1}{2}|z-y|^2 - \varphi(z)\right\}$ と定め, その帰結として $\varphi(x)+\psi(y) \le \frac{1}{2}|x-y|^2$, $x, y \in \mathbb{R}^n$ となる.

とすると

$$
\begin{aligned}
\mu(A) &= \pi_0(A \times \mathbb{R}^n) \\
&= \pi_0\left(\left\{x \in \mathbb{R}^n \;\middle|\; \varphi(x)+\psi(y) = \frac{1}{2}|x-y|^2 \text{ を満たす } y \text{ が存在する}\right\} \times \mathbb{R}^n\right) \\
&\geq \pi_0\left(\left\{(x,y) \in \mathbb{R}^n \times \mathbb{R}^n \;\middle|\; \varphi(x)+\psi(y) = \frac{1}{2}|x-y|^2\right\}\right) = 1
\end{aligned}
$$

となる. $x \in A$ をとれば, $\varphi(x)+\psi(y) = \frac{1}{2}|x-y|^2$ をみたす y 毎に

$$
z \mapsto \frac{1}{2}|z-y|^2 - \varphi(z)
$$

は x で最小値をとることがわかる. 特に, $x \in A$ で

$$
\varphi(x) \leq \inf_{w \in \mathbb{R}^n}\left(\frac{1}{2}|x-w|^2 - \psi(w)\right) \leq \frac{1}{2}|x-y|^2 - \psi(y) = \varphi(x)
$$

から,

$$
\varphi(x) = \inf_{w \in \mathbb{R}^n}\left(\frac{1}{2}|x-w|^2 - \psi(w)\right) \tag{1.3}
$$

を得る.

$$
\varphi_k(z) := \inf_{|w|<k}\left\{\frac{1}{2}|z-w|^2 - \psi(w)\right\}
$$

とおくと, φ_k は局所 Lipschitz 連続であり, $x \in A$ として十分大きい $k \in \mathbb{N}$ に対して $|y| < k$ になることから $\varphi(x) = \varphi_k(x)$ を得る. これより μ-a.s. $x \in \mathbb{R}^n$ で $k \in \mathbb{N}$ がとれて $\varphi(x) = \varphi_k(x)$ となる. したがって Rademacher の定理 ([144, Theorem 10.8(ii)]) より μ-a.s. $x \in \mathbb{R}^n$ で φ は微分可能となる. そこで

$$
D(T) := A \cap \left\{x \in \mathbb{R}^n \mid \varphi \text{ は } x \text{ で微分可能}\right\}
$$

とする. 明らかに $\mu(D(T)) = 1$ である. $x \in D(T)$ において $\varphi(x)+\psi(y) = \frac{1}{2}|x-y|^2$ の両辺を微分して

$$
(x-y) - \nabla\varphi(x) = \nabla\left(\frac{1}{2}|\cdot - y|^2 - \varphi\right)(x) = 0,
$$

すなわち $y = x - \nabla\varphi(x)$ を得る. つまり $x \in D(T)$ 毎に $\varphi(x)+\psi(y) = \frac{1}{2}|x-y|^2$ をみたす y がこの表示から一意的に決まる. そこで写像 $T : D(T) \to \mathbb{R}^n$ を

$$
T(x) = x - \nabla\varphi(x) = \nabla\left(\frac{1}{2}|\cdot|^2 - \varphi\right)(x), \quad x \in D(T)
$$

で定めることにする. リーマン多様体での指数写像 $\exp_x$ をユークリッド空間の場合に適用すると

$$
T(x) = x - \nabla\varphi(x) = \exp_x(-\nabla\varphi(x))
$$

と表される. そこで

$$\phi(z) := \sup_{w\in\mathbb{R}^n}\left(\psi(w) - \frac{1}{2}|w|^2 + \langle z, w\rangle\right)$$

で関数 ϕ を定めると, ϕ は表示から明らかに $\mathbb{R}^n$ 上の凸関数であり [9], (1.3) から

$$\phi(x) = \frac{1}{2}|x|^2 - \varphi(x), \quad x \in A$$

となることから $T = \nabla\phi = \exp.(-\nabla\varphi)$ μ-a.s. を得る.

最後に $\pi_0 = (\mathrm{Id}, T)_\sharp\mu$ [10]を示そう. これから最適輸送計画 π_0 は一意に定まることがわかり, $\nu = T_\sharp\mu$ [11]も従う.[12] まず T のグラフを $G(T) := \{(x, T(x)) \mid x \in D(T)\}$ とする. $\Gamma_0 := \{(x,y) \in \mathbb{R}^n \times \mathbb{R}^n \mid \varphi(x) + \psi(y) = \frac{1}{2}|x-y|^2\}$ とすると写像 T の構成の手続きから $G(T) = \Gamma_0 \cap (D(T) \times \mathbb{R}^n)$ がわかるので $\pi_0(\Gamma_0) = 1$ から

$$\pi_0(G(T)) = \pi_0(\Gamma_0 \cap (D(T) \times \mathbb{R}^n)) = \pi_0(D(T) \times \mathbb{R}^n) = \mu(D(T)) = 1$$

となる. これより, 任意の $A, B \in \mathcal{B}(\mathbb{R}^n)$ に対して

$$\begin{aligned}
\pi_0(A \times B) &= \pi_0((A \times B) \cap G(T)) \\
&= \pi_0(\{(x, T(x)) \mid x \in D(T) \cap A, T(x) \in B\}) \\
&= \pi_0(\{(x, T(x)) \mid x \in D(T) \cap A \cap T^{-1}(B)\}) \\
&= \pi_0([(D(T) \cap A \cap T^{-1}(B)) \times \mathbb{R}^n] \cap G(T)) \\
&= \pi_0((D(T) \cap A \cap T^{-1}(B)) \times \mathbb{R}^n) \\
&= \mu(D(T) \cap A \cap T^{-1}(B)) = \mu(A \cap T^{-1}(B)) = (\mathrm{Id}, T)_\sharp(A \times B)
\end{aligned}$$

となり, $\pi_0 = (\mathrm{Id}, T)_\sharp\mu$ を得る. □

注意 1.5 (1) 定理 1.4の帰結として Brenier [19, 20] による極因子分解定理 **(Polar factorization Theorem)** がわかる: λ^n を n-次元ルベーグ測度とし, $\nu \in \mathcal{P}_2^{\mathrm{ac}}(\mathbb{R}^n)$ [13]を考える. 可測ベクトル場 $X : \mathbb{R}^n \to \mathbb{R}^n$ が

$$X_\sharp\lambda^n \ll \lambda^n$$

---

[9] $\phi((1-t)x + ty) = \sup_{w\in\mathbb{R}^n}\{\psi(w) - \frac{1}{2}|w|^2 + \langle(1-t)x + ty, w\rangle\} = \sup_{w\in\mathbb{R}^n}\{(1-t)(\psi(w) - \frac{1}{2}|w|^2 + \langle x, w\rangle) + t(\psi(w) - \frac{1}{2}|w|^2 + \langle y, w\rangle)\} \le (1-t)\phi(x) + t\phi(y)$.

[10] ここで $(\mathrm{Id}, T)_\sharp\mu$ は μ の直積写像 (Id, T) による押し出し測度 (確率論の教科書では像測度) と呼ばれるもので $(\mathrm{Id}, T)_\sharp\mu(C) := \mu((\mathrm{Id}, T)^{-1}(C)) = \mu(\{x \in \mathbb{R}^n \mid (x, T(x)) \in C\})$, $C \in \mathcal{B}(\mathbb{R}^n \times \mathbb{R}^n)$ で定義される.

[11] $T_\sharp\mu$ も写像 $T : \mathbb{R}^n \to \mathbb{R}^n$ による μ の押し出し測度で, $T_\sharp\mu(B) := \mu(T^{-1}(B)) = \mu(\{x \in \mathbb{R}^n \mid T(x) \in B\})$, $B \in \mathcal{B}(\mathbb{R}^n)$ で定義される. μ (resp. ν) が n-次元 Lebesgue 測度 λ^n に関して絶対連続で, その密度関数を $f(x)$ (resp. $g(x)$) とするとき, T が λ^n-a.e. で可微分ならば, $\nu = T_\sharp\mu$ は $|\det(\frac{\partial T_i}{\partial x_j}(x))|g(T(x)) = f(x)$ λ^n-a.e. $x \in \mathbb{R}^n$ と同等である. ここで $\det(\frac{\partial T_i}{\partial x_j}(x))$ は T のヤコビ行列式である. $T = \nabla\phi$ λ^n-a.e. と T が凸関数 ϕ の勾配で表示されていれば, T は λ^n-a.e. で微分可能で g が正なら, ϕ は Monge-Ampére 方程式 $\det\nabla^2\phi(x) = \frac{f(x)}{g(\nabla\phi(x))}$ の解となる.

[12] $\nu = T_\sharp\mu$ を直接的に示すこともできる. これは (φ, ψ) が (1.1) の右辺を最大化することから, $h \in C_0^1(\mathbb{R}^n)$ に対して $\int_{\mathbb{R}^n} h \circ T \mathrm{d}\mu - \int_{\mathbb{R}^n} h \mathrm{d}\nu = \lim_{\varepsilon\to 0}\frac{1}{\varepsilon}(J(\varphi + \varepsilon h \circ T, \psi - \varepsilon h) - J(\varphi, \psi)) = 0$ となることでわかる. ここで $J(f, g) := \int_{\mathbb{R}^n} f \mathrm{d}\mu + \int_{\mathbb{R}^n} g \mathrm{d}\nu$ である.

[13] $\mathcal{P}_2^{\mathrm{ac}}(\mathbb{R}^n) := \{\mu \in \mathcal{P}(\mathbb{R}^n) \mid \mu$ は Lebesgue 測度に関して絶対連続 $\}$.

を満たせば, ある凸関数 ψ と $\eta_\sharp\nu = \nu$ をみたす写像 $\eta : \mathbb{R}^n \to \mathbb{R}^n$ が存在して $X = \nabla\psi \circ \eta$ ν-a.s. と表せる.

証明　$\mu := X_\sharp\nu$ とすると仮定から $\mu \in \mathcal{P}_2^{\mathrm{ac}}(\mathbb{R}^n)$ であり, 定理 1.4からコスト関数 $c(x,y) = \frac{1}{2}|x-y|^2$ と μ, ν に関する最適輸送写像 T は下半連続な凸関数 ϕ を用いて $T = \nabla\phi$ μ-a.s. の表現を持ち

$$\nu = T_\sharp\mu = (\nabla\phi)_\sharp\mu = (\nabla\phi)_\sharp X_\sharp\nu = (\nabla\phi \circ X)_\sharp\nu$$

と表せる. $\eta := \nabla\phi \circ X$ とおくと上式は $\eta_\sharp\nu = \nu$ を表している. ϕ の Fenchel-Legendre 変換を $\phi^*(y) := \sup\{\langle x,y\rangle - \phi(x) \mid x \in \mathbb{R}^n\}$ とおくと, これは下半連続凸関数で $\nabla\phi^* \circ \nabla\phi = \mathrm{Id}$ μ-a.s. [14]となるので $X = \nabla\phi^* \circ \eta$ ν-a.s. を得る. □

(2) 定理 1.4は Mikami [99] によって確率論的な別証明が与えられている.

(3) 定理 1.4はユークリッド空間の設定で Gangbo-McCann [56] においていくつかの条件をみたす狭義凸関数 $h : \mathbb{R}^n \to [0,+\infty[$ [15]を用いたコスト関数 $c(x,y) = h(x-y)$, $x, y \in \mathbb{R}^n$ の場合に拡張されている. 詳細は [153] の解説を参照されたい. モンジュの問題を解決した Trudinger-Wang [139], Caffarelli-Feldman-McCann [23] の証明は Gangbo-McCann [56] の結果に立脚している.[16]

(4) 定理 1.4は McCann [97] によってコンパクトリーマン多様体の場合に, その後 Fathi-Figalli [52], Figalli-Gigli [53] によって非コンパクトリーマン多様体 (M,g) の場合に拡張された. その場合もリーマン距離 $d_g(x,y)$ を用いた $c(x,y) = \frac{1}{2}d_g^2(x,y)$ をコスト関数とする $\mu_0 \in \mathcal{P}_2^{\mathrm{ac}}(M)$ と $\mu_1 \in \mathcal{P}_2(M)$ に関する最適輸送写像 T は局所弱凸関数 φ が存在して

$$T = \exp_\cdot(\nabla\varphi) \quad \mu\text{-a.s.}$$

と表現される. また勾配ベクトル場に沿う μ_0 の押し出し測度

$$\mu_t := (T_t)_\sharp\mu_0, \quad T_t(x) := \exp_x(t\nabla\varphi(x)), \quad t \in [0,1]$$

が後述の L^2-Wasserstein 空間 $(\mathcal{P}_2(M), W_2)$ [17]における μ_0 から μ_1 への最短測地線を与える. すなわち

$$W_2(\mu_0,\mu_t) = tW_2(\mu_0,\mu_1), \quad W_2(\mu_t,\mu_1) = (1-t)W_2(\mu_0,\mu_1), \quad t \in [0,1].$$

---

[14] 凸関数の組 (ϕ,ϕ^*) は $\mu := X_\sharp\nu$ と ν に対してコスト関数 $c(x,y) = -\langle x,y\rangle$ に後述の Kantorovich 双対性 (定理 1.24) を適用して, $(-\phi,-\phi^*)$ が $\sup_{\pi\in\Pi(\mu,\nu)}\int_{\mathbb{R}^n\times\mathbb{R}^n}\langle x,y\rangle\pi(\mathrm{d}x\mathrm{d}y) = \inf_{(\varphi,\psi)\in\Phi_c\cap C_b} J(-\varphi,-\psi)$ の右辺を最小化するものであり, 左辺の最大点 $\pi_0 \in \Pi(\mu,\nu)$ について等式 $\langle x,y\rangle = \phi(x) + \phi^*(y)$ π_0-a.s. $(x,y) \in \mathbb{R}^n\times\mathbb{R}^n$ をみたすものになっている. $\phi(x)$ (resp. $\phi^*(y)$) は Lebesgue 測度に関して a.e. x (resp. a.e. y) で微分可能であるので π_0-a.s. (x,y) で $\phi(x)$ と $\phi^*(y)$ は微分可能となり, 等式 $\langle x,y\rangle = \phi(x)+\phi^*(y)$ の両辺を y について微分することで $x = \nabla\phi^*(y)$, x について微分することで $y = \nabla\phi(x)$ を得て, $x = \nabla\phi^*\circ\nabla\phi(x)$ π_0-a.s. (x,y) となる, したがって $x = \nabla\phi^*\circ\nabla\phi(x)$ μ-a.s. $x \in \mathbb{R}^n$ となる.

[15] $h(x) = |x|^p$, $x \in \mathbb{R}^n$, $p \in]1,+\infty[$ などが一例.

[16] コスト関数 $c_p(x,y) = |x-y|^p$ に対するモンジュ・カントロヴィッチ問題の一意的な解において極限 $p \to 1$ をとることで $c(x,y) = |x-y|$ に対するモンジュ・カントロヴィッチ問題の解を構成している.

[17] $\mathcal{P}_2(M)$ はリーマン多様体 (M,g) における２次積率有限なボレル確率測度の全体である. すなわち $\mathcal{P}_2(M) := \{\mu \in \mathcal{P}(M) \mid \exists x_0 \in M \text{ s.t. } \int_M d_g^2(x,x_0)\mu(\mathrm{d}x) < \infty\}$ (d_g は (M,g) 上のリーマン距離).

1.2 モンジュ・カントロヴィッチ問題の解の存在

定理 1.6 (最小点の存在) X, Y をポーランド空間[18]とし, $\mu \in \mathcal{P}(X)$, $\nu \in \mathcal{P}(Y)$ をそれぞれの空間上の確率測度とする．コスト関数 $c : X \times Y \to [0,\infty]$ [19] は下半連続とする．$\int_{X\times Y} c(x,y)\pi(\mathrm{d}x\mathrm{d}y) < \infty$ をみたす $\pi \in \Pi(\mu,\nu)$ が存在すれば, モンジュ・カントロヴィッチ問題

$$\inf_{\pi\in\Pi(\mu,\nu)} \int_{X\times Y} c(x,y)\pi(\mathrm{d}x\mathrm{d}y) < \infty \tag{1.4}$$

の最小点 $\pi_0 \in \Pi(\mu,\nu)$ が存在する．この π_0 を μ から ν への**最適輸送計画**, あるいは μ と ν の**最適カップリング**と呼ぶ.

注意 1.7 (1) (1.4) の 最小点である最適輸送計画 π_0 は一般に一意ではない.

(2) MK 問題は確率論的には

$$\inf \{\mathbf{E}[c(U,V)] \mid U_\sharp \mathbf{P} = \mu, V_\sharp \mathbf{P} = \nu\} \tag{1.5}$$

を最小化する $X\times Y$-値確率変数 (U,V) を求めることでもある．(U,V) の分布 $(U,V)_\sharp\mathbf{P}$ は U の分布 $U_\sharp\mathbf{P} = \mu$, V の分布 $V_\sharp\mathbf{P} = \nu$ のカップリングに他ならない.

定義 1.8 (半連続性) $(E,\mathcal{O}(E))$ を位相空間とする．関数 $f : E \to \overline{\mathbb{R}}$ の $x \in E$ における下極限 (lower limit) を

$$\varliminf_{y\to x} f(y) := \sup_{U\in\mathcal{U}(x)} \inf_{y\in U} f(y) \tag{1.6}$$

で定める．同様に上極限 (upper limit) を

$$\varlimsup_{y\to x} f(y) := \inf_{U\in\mathcal{U}(x)} \sup_{y\in U} f(y) \tag{1.7}$$

で定める．ここで $\mathcal{U}(x)$ は $x \in E$ での位相の近傍系である．$(E,\mathcal{O}(E))$ が第一可算で, 点列 $\{x_n\}$ が x に収束しているときは

$$\varliminf_{y\to x} f(y) = \varliminf_{n\to\infty} f(x_n) := \sup_{n\in\mathbb{N}} \inf_{k\geq n} f(x_k), \tag{1.8}$$

$$\varlimsup_{y\to x} f(y) = \varlimsup_{n\to\infty} f(x_n) := \inf_{n\in\mathbb{N}} \sup_{k\geq n} f(x_k) \tag{1.9}$$

となる．$(E,\mathcal{O}(E))$ が距離 d から定まる位相空間のときは

$$\varliminf_{y\to x} f(y) = \lim_{\varepsilon\to 0} \inf_{d(x,y)<\varepsilon} f(y), \tag{1.10}$$

$$\varlimsup_{y\to x} f(y) = \lim_{\varepsilon\to 0} \sup_{d(x,y)<\varepsilon} f(y) \tag{1.11}$$

---

[18]ポーランド空間とは完備可分な距離空間と位相同型な位相空間である．この講演で扱う空間は全てポーランド空間である.

[19]ここでは簡単のためコスト c は非負としているが, 非負の条件を弱めて上半連続な $a \in L^1(X;\mu)$, $b \in L^1(Y;\nu)$ がとれて $c(x,y) \geq a(x) + b(y)$ ${}^\forall(x,y) \in X \times Y$ が成立するとしても同じ結論を得る.

となる． 関数 $f : E \to \overline{\mathbb{R}}$ が $x \in E$ で**下半連続** (lower semi continuous at $x \in E$) (resp. $x \in E$ で**上半連続** (upper semi continuous at $x \in E$)) であるとは $f(x) \le \underline{\lim}_{y\to x} f(y)$ (resp. $f(x) \ge \overline{\lim}_{y\to x} f(y)$) のこととする．また $f : E \to \overline{\mathbb{R}}$ が E **上で下半連続** (lower semi continuous on E) (resp. E **上で上半連続** (upper semi continuous on E) とは f が全ての点 $x \in E$ で下半連続 (resp. 全ての点 $x \in E$ で下半連続) のこととする．

補題 1.9 (E, d) を距離空間とする．下に有界な下半連続関数 $f : E \to \overline{\mathbb{R}}$ に対して Lipschitz 連続な有界関数からなる単調非減少列 $\{f_n\}$ で $f(x) = \lim_{n\to\infty} f_n(x)$ となるものが存在する．

証明 f_n, g_n を

$$g_n(x) := \inf_{y\in E} \{nd(x, y) + f(y)\}, \quad f_n(x) := g_n(x) \wedge n$$

とおくと，明らかに $g_n(x) \le g_{n+1}(x) \le f(x)$ で，g_n は n-Lipschitz であるので f_n も同じ性質をもつ．f が下に有界であることから g_n, f_n も下に一様に有界で同じ下界をもつ．$g_n(x) \to f(x)$ as $n \to \infty$ のみ示せばよい．$f(x) - \sup_{n\in\mathbb{N}} g_n(x) > 0$ として $0 < \varepsilon < f(x) - \sup_{n\in\mathbb{N}} g_n(x)$ をとる．これから点列 $\{y_n\}$ で $nd(x, y_n) + f(y_n) < f(x) - \varepsilon$ をみたすものがとれる．したがって

$$d(x, y_n) < \frac{f(x) - f(y_n) - \varepsilon}{n} \le \frac{f(x) - \inf_{y\in E} f(y) - \varepsilon}{n} \to 0 \quad n \to \infty$$

となるので $\lim_{n\to\infty} y_n = x$ となる．これと f の下半連続性から $f(x) \le \underline{\lim}_{n\to\infty} f(y_n) \le f(x) - \varepsilon$ を得て矛盾するので，$f(x) = \sup_{n\in\mathbb{N}} g_n(x)$ となる． □

定義 1.10 (弱収束) (E, d) を距離空間とする． $\{\mu_k\} \subset \mathcal{P}(E)$ が $\mu \in \mathcal{P}(E)$ に**弱収束** (weakly convergent) するとは，任意の $\varphi \in C_b(E)$ [20]に対して

$$\lim_{k\to\infty} \int_E \varphi(x)\mu_k(\mathrm{d}x) = \int_E \varphi(x)\mu(\mathrm{d}x) \tag{1.12}$$

が成立することとする．このとき $\mu_k \xrightarrow{w} \mu$ と記す．これらは次の条件と同値である．

(1) 任意の有界な下半連続関数 φ に対し，$\displaystyle\int_E \varphi(x)\mu(\mathrm{d}x) \le \underline{\lim}_{k\to\infty} \int_E \varphi(x)\mu_k(\mathrm{d}x)$.

(2) 任意の下に有界な下半連続関数 φ に対し，$\displaystyle\int_E \varphi(x)\mu(\mathrm{d}x) \le \underline{\lim}_{k\to\infty} \int_E \varphi(x)\mu_k(\mathrm{d}x)$.

(3) 任意の有界な上半連続関数 φ に対し，$\displaystyle\int_E \varphi(x)\mu(\mathrm{d}x) \ge \overline{\lim}_{k\to\infty} \int_E \varphi(x)\mu_k(\mathrm{d}x)$.

(4) 任意の上に有界な上半連続関数 φ に対し，$\displaystyle\int_E \varphi(x)\mu(\mathrm{d}x) \ge \overline{\lim}_{k\to\infty} \int_E \varphi(x)\mu_k(\mathrm{d}x)$.

[20] $C_b(E)$ は E 上の有界連続関数の全体.

(5) 任意の開集合 G に対し, $\mu(G) \leq \varliminf_{k\to\infty} \mu_k(G)$.

(6) 任意の閉集合 F に対し, $\mu(F) \geq \varlimsup_{k\to\infty} \mu_k(F)$.

(7) 任意のボレル集合 A で $\mu(\partial A) = 0$ をみたすものに対し, $\mu(A) = \lim_{k\to\infty} \mu_k(A)$.

証明 (定義 1.10での主張の同値性の証明) (5), (6), (7) と (1.12) との同値性は標準的なテキスト (例えば, 小谷 [151, 命題 9.2]) に記述があるので略する. (1) と (3) との同値性は φ を $-\varphi$ にすることで自明. したがって (1) を仮定すると (3) とあわせて (1.12) が得られる. 逆に (1.12) を仮定すると, 補題 1.9から任意の有界な下半連続関数 φ が有界な連続関数列 $\{\varphi_n\}$ の単調増大極限であることから

$$\int_E \varphi \mathrm{d}\mu =\uparrow \lim_{n\to\infty} \int_E \varphi_n \mathrm{d}\mu \leq \lim_{n\to\infty} \lim_{k\to\infty} \int_E \varphi_n \mathrm{d}\mu_k \leq \varliminf_{k\to\infty} \int_E \varphi \mathrm{d}\mu_k$$

となり (3) を得る. (1)$\Longrightarrow$(2) と (3)$\Longrightarrow$(4) は自明, (2)$\Longrightarrow$(1) は $\varphi_n := \varphi \wedge n$ として

$$\int_E \varphi_n \mathrm{d}\mu \leq \varliminf_{k\to\infty} \int_E \varphi_n \mathrm{d}\mu_k \leq \varliminf_{k\to\infty} \int_E \varphi \mathrm{d}\mu_k$$

から従う. (4)$\Longrightarrow$(3) も同様である. □

注意 1.11 (1) 距離空間 (E, d) において $\mathcal{P}(E)$ 上の弱収束は距離化可能な位相から決まる. この位相を弱位相 **(weak topology)** という ([4, Remark 5.1.1]).

(2) 距離空間 (E, d) において確率測度の全体 $\mathcal{P}(E)$ は有界連続関数全体 $C_b(E)$ の位相的双対空間 $(C_b(E))^*$ の単位球とみなせることから $\mathcal{P}(E)$ の弱位相は $(C_b(E))^*$ の weak* 位相とみなせる.

定義 1.12 (緊密性 (tightness)) (E, d) を距離空間, $\mathcal{P}(E)$ を E 上のボレル確率測度の全体とする. E 上の確率測度の族 $\Pi(\subset \mathcal{P}(E))$ が緊密 **(tight)** であるとは $\forall\varepsilon > 0$ に対し, コンパクト集合 $\exists K_\varepsilon$ で $\sup_{\mu\in\Pi} \mu(K_\varepsilon^c) < \varepsilon$ を満たすものがとれることとする.

定理 1.13 (Prokhorov's Theorem)
(E, d) を可分距離空間とする. E 上の確率測度の族 $\Pi(\subset \mathcal{P}(E))$ に対し,

(1) Π は緊密 (tight) である.

(2) Π は $\mathcal{P}(E)$ の弱収束で相対点列コンパクトである.

としたとき, (1)$\Rightarrow$(2) が成立する. さらに (E, d) が完備なら (2)$\Rightarrow$(1) が成立する.

注意 1.14 (1) 定理 1.13の証明については Billingsley [16, Theorems 5.1 and 5.2], Ikeda-Watanabe [73, Theorem 2.6], Dellacherie-Meyer [41, III-59] や小谷 [151, 定理 9.4] 等の確率論の教科書を参照されたい.

(2) 定理 1.13から E がポーランド空間なら Π の緊密性と Π の弱位相での相対コンパクト性が同値になる．このような性質をもつ位相空間をラドン空間 **(Radon space)** という ([4, Definition 5.1.4])．定理 1.13はポーランド空間ならラドン空間であることを主張している．

証明　(定理 1.6の証明) $\{\pi_n\} \subset \Pi(\mu,\nu)$ を (1.4) の 最小化列とすると, $\{\pi_n\}$ は $\mathcal{P}(X\times Y)$ において緊密 (tight) な族になる．実際, X, Y がポーランド空間であることから Prokhorov の定理(定理 1.13) より $\forall\varepsilon>0$ に対してコンパクト集合 $K_1\subset X, K_2\subset Y$ で $\mu(K_1^c), \nu(K_2^c) < \varepsilon/2$ となるものがとれるので $\pi_n((K_1\times K_2)^c) \le \pi_n(K_1^c\times Y)+\pi_n(X\times K_2^c) = \mu(K_1^c)+\nu(K_2^c) < \varepsilon$ から $\sup_{n\in\mathbb{N}}\pi_n((K_1\times K_2)^c) < \varepsilon$ となるので $\{\pi_n\}$ は緊密な族である．これと Prokhorov の定理 (定理 1.13) から $\{\pi_n\}$ から部分列 $\{\pi_{n_k}\}$ がとれて, ある $\pi_0\in\mathcal{P}(X\times Y)$ に弱収束する．コスト関数の下半連続性から

$$\begin{aligned}\int_{X\times Y} c(x,y)\pi_0(\mathrm{d}x\mathrm{d}y) &\le \lim_{n\to\infty}\int_{X\times Y}\{c(x,y)\wedge n\}\pi_0(\mathrm{d}x\mathrm{d}y)\\ &= \lim_{n\to\infty}\varliminf_{k\to\infty}\int_{X\times Y}\{c(x,y)\wedge n\}\pi_{n_k}(\mathrm{d}x\mathrm{d}y)\\ &= \varliminf_{k\to\infty}\int_{X\times Y} c(x,y)\pi_{n_k}(\mathrm{d}x\mathrm{d}y)\\ &= \inf_{\pi\in\Pi(\mu,\nu)}\int_{X\times Y} c(x,y)\pi(\mathrm{d}x\mathrm{d}y)\end{aligned}$$

から π_0 が最小点であることがわかる．次に $\pi_0\in\Pi(\mu,\nu)$ になることを示す．$A\in\mathcal{B}(X)$ に対して $\pi_0(A\times Y)=\mu(A)$ のみ示せば十分である．$f\in C_b(X)$ に対して $f\in C_b(X\times Y)$ でもあるので $\pi_{n_k}\in\Pi(\mu,\nu)$ とルベーグ積分の変数変換公式から

$$\int_X f(x)\pi_{n_k}(\mathrm{d}x\mathrm{d}y) = \int_X f(x)\mu(\mathrm{d}x)$$

を得る．これより

$$\int_X f(x)\pi_0(\mathrm{d}x\mathrm{d}y) = \lim_{k\to\infty}\int_X f(x)\pi_{n_k}(\mathrm{d}x\mathrm{d}y) = \int_X f(x)\mu(\mathrm{d}x)$$

である．これから開集合 $G\in\mathcal{O}(X)$ に対して $\mathbf{1}_G(x)=\lim_{n\to\infty} nd(x,G^c)\wedge 1$ から

$$\begin{aligned}\pi_0(G\times Y) &= \int_X \mathbf{1}_G(x)\pi_0(\mathrm{d}x\mathrm{d}y)\\ &= \lim_{n\to\infty}\int_X (nd(x,G^c)\wedge 1)\pi_0(\mathrm{d}x\mathrm{d}y)\\ &= \lim_{n\to\infty}\int_X (nd(x,G^c)\wedge 1)\mu(\mathrm{d}x)\\ &= \int_X \mathbf{1}_G(x)\mu(\mathrm{d}x) = \mu(G)\end{aligned}$$

を得る. 補集合をとることで X の閉集合 F に対しても $\pi_0(F\times Y)=\mu(F)$ を得る.

$$\mathcal{A}=\{A\in\mathcal{B}(X)\mid \pi_0(A\times Y)=\mu(A)\}$$

と置くと上記のことから $\mathcal{O}(X)\subset\mathcal{A}\subset\mathcal{B}(X)$ であり, $\mathcal{A}$ が X 上の σ-加法族になることから $\mathcal{A}=\mathcal{B}(X)$, すなわち 任意の $A\in\mathcal{B}(X)$ に対して $\pi_0(A\times Y)=\mu(A)$ を得る. $\square$

ボレル可測写像 $T:X\to Y$ で $T_\sharp\mu=\nu$ をみたすものに対して $\pi_T:=(\mathrm{Id},T)_\sharp\mu$ とおくと π_T は μ と ν のカップリングになる. 実際, $A\in\mathcal{B}(X)$, $B\in\mathcal{B}(Y)$ に対し $\pi_T(A\times Y)=\mu((\mathrm{Id},T)^{-1}(A\times Y))=\mu(A\times T^{-1}(Y))=\mu(A)$, $\pi_T(X\times B)=\mu((\mathrm{Id},T)^{-1}(X\times B))=\mu(X\times T^{-1}(B))=\mu(T^{-1}(B))=\nu(B)$ である. このような型の輸送計画において最小化を考えるのが (一般化された) モンジュの問題である：

定義 1.15 ((一般化された) モンジュの問題)

$$\begin{aligned}&\inf\left\{\int_{X\times Y}c(x,y)\pi_T(\mathrm{d}x\mathrm{d}y)\,\middle|\, T:X\to Y \text{ ボレル可測},\ T_\sharp\mu=\nu\right\}\\&=\inf\left\{\int_X c(x,T(x))\mu(\mathrm{d}x)\,\middle|\, T:X\to Y \text{ ボレル可測},\ T_\sharp\mu=\nu\right\}\end{aligned}$$

を最小にする $T_\sharp\mu=\nu$ をみたすボレル可測写像 $T:X\to Y$ を求めることを (一般化された) モンジュの問題という. “一般化” とは必ずしも T に全単射性を要請しないことを意味する. 以後, 一般化された意味でのモンジュの問題を単にモンジュの問題と呼ぶことにする.

例 1.16 (ディラック測度のケース I) X,Y を位相空間とする. $\mu\in\mathcal{P}(X)$, $\nu\in\mathcal{P}(Y)$ において $\nu=\delta_b$, $b\in Y$ と仮定する. このとき μ と ν のカップリングは直積測度 $\mu\times\delta_b$ のみである. 実際, $b\notin B$ なら $\pi(A\times B)\le\pi(X\times B)=\delta_b(B)=0$ なので $b\in B$ なら $\pi(A\times B)=\pi(A\times Y)=\mu(A)$ となることから, $\pi(A\times B)=\mu(A)\delta_b(B)$ がわかる. また $T_\sharp\mu=\delta_b$ をみたす輸送写像 T は $T(x):=b$, $x\in X$ となる定値写像なので $\pi_T=\mu\times\delta_b$ がわかり

$$\int_{X\times Y}c(x,y)\pi_T(\mathrm{d}x\mathrm{d}y)=\int_X c(x,b)\mu(\mathrm{d}x)$$

が MK 問題かつモンジュの問題の最小値である. T は定値写像なので全射でも単射でもないことに注意されたい. 特に $\mu=\delta_a$, $\nu=\delta_b$ $(a\in X, b\in Y)$ のとき, μ と ν のカップリングは直積測度 $\delta_a\times\delta_b$ のみで $c(a,b)$ が MK 問題かつモンジュの問題の最小値である.

例 1.17 (ディラック測度のケース II) X,Y を位相空間とする. $\mu\in\mathcal{P}(X)$, $\nu\in\mathcal{P}(Y)$ において $\mu=\delta_a$, $a\in X$ のみを仮定する. このとき例 1.16 と同様に μ と ν のカップリングは直積測度 $\delta_a\times\nu$ のみである. したがって

$$\int_X c(a,y)\nu(\mathrm{d}x)$$

が MK 問題の最小値である． しかしながら $T_\sharp\delta_a = \delta_{T(a)}$ なので $\nu = \delta_{T(a)}$ の場合を除いて $T_\sharp\delta_a = \nu$ とはできない．したがって $\nu = \delta_{T(a)}$ の場合を除いてモンジュの問題を定式化することはできない．

例 1.18 (有限集合のケース) X, Y を位相空間とする．$\mu \in \mathcal{P}(X)$ $\nu \in \mathcal{P}(Y)$ が同じ個数の有限集合をそれぞれ台とする一様離散確率測度とする．すなわち

$$\mu = \frac{1}{n}\sum_{i=1}^{n}\delta_{x_i}, \quad \nu = \frac{1}{n}\sum_{i=1}^{n}\delta_{y_i}$$

とする． この場合, μ と ν のカップリング π は双確率行列 $\Pi = (\pi_{ij})$ を用いて $\pi = \frac{1}{n}\sum_{i,j=1}^{n}\pi_{ij}\delta_{(x_i,y_j)}$ と表示される．ここで双確率行列とは

$$\forall j, \quad \sum_{i=1}^{n}\pi_{ij} = 1, \quad \forall i, \quad \sum_{j=1}^{n}\pi_{ij} = 1$$

をみたす行列 $\Pi = (\pi_{ij})$ のことである．したがって MK 問題は

$$\inf\left\{\frac{1}{n}\sum_{i,j=1}^{n}\pi_{ij}c(x_i, y_j) \;\middle|\; \Pi \in \mathcal{B}_n\right\}$$

を最小化する問題となる．ここで $\mathcal{B}_n$ は $n \times n$ の双確率行列の全体である．この問題は $n \times n$ 行列全体の有界凸部分集合 $\mathcal{B}_n$ 上の線形最小化問題そのものである．Choquet の定理 (定理 1.48) により，この MK 問題の最小点は $\mathcal{B}_n$ の端点集合の全体である． ここで $\mathcal{B}_n$ の端点とは $\mathcal{B}_n$ 内の任意の異なる 2 点を結ぶ線分の内点でない点を指す．Birkhoff の定理 (定理 1.49) により, $\mathcal{B}_n$ の端点 $\Pi = (\pi_{ij})$ は置換 $\sigma \in S_n$ を用いて $\pi_{ij} = \delta_{j\sigma(i)}$ と表現される．写像 T を $T(x_i) := y_{\sigma(i)}$, $i \in \{1, 2, \cdots, n\}$ をみたすようにとると $T_\sharp\mu = \nu$ が確認できる．したがって MK 問題は離散モンジュ問題

$$\inf\left\{\frac{1}{n}\sum_{i=1}^{n}c(x_i, y_{\sigma(i)}) \;\middle|\; \sigma \in S_n\right\}$$

を最小化する問題となる．

1.3 Kantorovich 双対性

Kantorovich 双対性 (定理 1.24) と Kantorovich-Rubinstein Theorem (定理 1.27) の説明と証明のため以下の諸概念を準備する．

定義 1.19 (c-変換, c-凹性, c-劣微分) X, Y を空でない集合とする．$c : X \times Y \to \overline{\mathbb{R}}$ を任意の関数とする．

(1) 関数 $u : X \to \overline{\mathbb{R}}$ に対して, その c-変換 (c-transform) $u^c : Y \to \overline{\mathbb{R}}$ を

$$u^c(y) := \inf_{x \in X}(c(x,y) - u(x))$$

で定める．ここで $c(x,y) = u(x) = +\infty$ のときは $c(x,y) - u(x) := +\infty$ と規約する．同様に, 関数 $v : Y \to \overline{\mathbb{R}}$ に対して, その c-変換 (c-transform) $v^c : X \to \overline{\mathbb{R}}$ を

$$v^c(x) := \inf_{y \in X}(c(x,y) - v(y))$$

で定める．$c \equiv +\infty$ のときは上記の規約性から $u^c \equiv +\infty$ となり, $u^{cc} \equiv +\infty$ となる．

(2) 関数 $u : X \to \overline{\mathbb{R}}$ が c-凹 (c-concave) とは, ある関数 $v : Y \to \overline{\mathbb{R}}$ を用いて $u = v^c$ と表されることとする．同値な言い換えとして u が c-凹であるとは, ある $\{(y_i, t_i)\}_{i \in I} \subset Y \times \overline{\mathbb{R}}$ で

$$u(x) = \inf_{i \in I}(c(x, y_i) + t_i) \quad \forall x \in X \tag{1.13}$$

が成立することである．同様に, 関数 $v : Y \to \overline{\mathbb{R}}$ が c-凹 (c-concave) とはある関数 $u : X \to \overline{\mathbb{R}}$ を用いて $v = u^c$ と表されることとする．

(3) $\varphi : X \to \overline{\mathbb{R}}$ が c-凸関数 (c-convex) [21]であるとは $-\varphi$ が c-凹関数であることとする．

(4) c-凸関数 φ の $x \in X$ での c-劣微分 $\partial^c\varphi(x)$ を

$$\partial^c\varphi(x) := \{y \in Y \mid \varphi(z) + c(z,y) \geq \varphi(x) + c(x,y) \text{ for any } z \in X\}$$

で定める．また c-凸関数 φ の c-劣微分 $\partial^c\varphi$ を

$$\partial^c\varphi := \{(x,y) \in X \times Y \mid y \in \partial^c\varphi(x)\}$$

で定める．

注意 1.20 $X = Y = \mathbb{R}^n$ で $c(x,y) = -\langle x, y\rangle = -\sum_{i=1}^n x_i y_i$, $x, y \in \mathbb{R}^n$ のとき, φ の c-凸性と通常の凸性は一致し, $\partial^c\varphi(x) = \partial\varphi(x)$, $\partial^c\varphi = \partial\varphi$ となる．このとき, $u : \mathbb{R}^n \to \overline{\mathbb{R}}$ の c-変換は Fenchel-Legendre 変換

$$f^*(y) := \sup\{\langle x, y\rangle - f(x) \mid x \in \mathbb{R}^n\}$$

を用いると $u^c(y) = -(-u)^*(y)$ と表される．ここで

$$\partial\varphi(x) := \{y \in \mathbb{R}^n \mid \varphi(z) - \varphi(x) \geq \langle y, z - x\rangle \text{ for any } z \in \mathbb{R}^n\}$$

は凸関数 φ の x における劣微分 **(subdifferential at** $x \in \mathbb{R}^n$**)** で, $\partial\varphi := \{(x,y) \mid y \in \partial\varphi(x)\}$ は φ の劣微分 **(subdifferential)** である．φ が x で微分可能なときは $\partial\varphi(x) = \{\nabla\varphi(x)\}$ となるので φ が微分可能関数のときは $\partial\varphi = \{\nabla\varphi(x) \mid x \in \mathbb{R}^n\}$ となる．

[21]後述の K-凸性の概念とは意味が異なるので注意を要する．

補題 1.21 X, Y を空でない集合とする．$c: X \times Y \to [0, +\infty]$ を任意の関数とする．次が成立する．

(1) 関数 $u: X \to \overline{\mathbb{R}}$ に対し, $u^{cc} \geq u$ が成立する．等号が成立するのは u が c-凹のときである．同様に, 関数 $v: Y \to \overline{\mathbb{R}}$ に対し, $v^{cc} \geq v$ が成立する．等号が成立するのは v が c-凹のときである．

(2) c が連続なら関数 $u: X \to \overline{\mathbb{R}}$ に対し, u^c は上半連続である．関数 $v: Y \to \overline{\mathbb{R}}$ に対しても同様の主張が成立する．

証明 (1): $u^{cc}(x) = -\infty$ のときは, ある $\{y_i\} \subset Y$ がとれて $c(x, y_i) - u^c(y_i) \overset{i\to\infty}{\to} -\infty$ となる．部分列をとることで全ての i で $c(x, y_i) < \infty$ としてよい．実際, 有限個の i を除いて $c(x, y_i) = +\infty$ とすると $c(x, y_i) - u^c(y_i) = +\infty$ が $u^c(y_i)$ の値に依らず成立するので矛盾である．このとき, $u(x) = c(x, y_i) - (c(x, y_i) - u(x)) \leq c(x, y_i) - u^c(y_i) \to -\infty$ から $u(x) = -\infty$ となり主張が成立する．$u^{cc}(x) = +\infty$ のときは自明な主張である．したがって $u^{cc}(x) \in \mathbb{R}$ と仮定してよい．任意の $\varepsilon > 0$ に対して $\exists y \in Y$ s.t. $u^{cc}(x) \leq c(x, y) - u^c(y) < u^{cc}(x) + \varepsilon$ となる．$u^c(y) = +\infty$ とすると $c(x, y) < \infty$ なら $u^{cc}(x) = -\infty$ となり矛盾, $c(x, y) = +\infty$ なら $c(x, y) - u^c(y) = +\infty < u^{cc}(x) + \varepsilon$ となるので矛盾．したがって, $u^c(y) < +\infty$ となり, $c(x, y) < +\infty$ を得る．ゆえに

$$\begin{aligned} u^{cc}(x) + \varepsilon &\geq c(x, y) - u^c(y) \\ &= c(x, y) - \inf_{z \in X}(c(z, y) - u(z)) \\ &\geq (c(x, y) - c(x, y) + u(x)) = u(x) \end{aligned}$$

において $\varepsilon \to 0$ として $u^{cc}(x) \geq u(x)$ を得る．等号成立ならあきらかに u は c-凹である: $u = u^{cc} \iff u = (u^c)^c$．逆に u が c-凹とする．すなわち $u = v^c$ となる関数 $v: Y \to \overline{\mathbb{R}}$ があるとする．$v^{ccc} \leq v^c$ を示せばよい．これは

$$v^{ccc}(x) \leq c(x, y) - v^{cc}(y) \leq c(x, y) - v(y)$$

から自明である．
(2) は自明な主張である． □

定義 1.22 (c-巡回的単調性) X, Y をポーランド空間とし, $c: X \times Y \to [0, \infty]$ を任意の関数とする．$\Gamma \subset X \times Y$ が c-巡回的単調 (c-cyclically monotone) であるとは任意の $n \in \mathbb{N}$, $\{(x_i, y_i)\}_{i=1}^n \subset \Gamma$ と置換 $\sigma \in S_n$ に対し,

$$\sum_{i=1}^n c(x_i, y_{\sigma(i)}) = \sum_{i=1}^n c(x_{\sigma(i)}, y_i) \geq \sum_{i=1}^n c(x_i, y_i)$$

が成立することとする．

補題 1.23 c-凸関数 $\varphi: X \to \mathbb{R}$ に対し, $\partial^c \varphi$ は c-巡回的単調である．

証明 $n \in \mathbb{N}, \{(x_i, y_i)\}_{i=1}^n \subset \partial^c \varphi, \sigma \in S_n$ に対し, $y_i \in \partial^c \varphi(x_i)$ なので $\varphi(x_{\sigma(i)}) + c(x_{\sigma(i)}, y_i) \geq \varphi(x_i) + c(x_i, y_i)$ が成立する. 両辺を $i = 1, 2, \cdots, n$ について和をとれば主張を得る. □

定理 1.24 (Kantorovich 双対性) X, Y をポーランド空間とし, $c : X \times Y \to [0, \infty]$ は proper で下半連続とする. $\pi \in \mathcal{P}(X \times Y)$, $(\varphi, \psi) \in L^1(X; \mu) \times L^1(Y; \nu)$ に対して

$$C[\pi] := \int_{X \times Y} c(x, y) \mathrm{d}\pi(x, y), \quad J(\varphi, \psi) := \int_X \varphi \mathrm{d}\mu + \int_Y \psi \mathrm{d}\nu$$

とし,

$$\begin{aligned} \Phi_c \cap L^1 := \Big\{ (\varphi, \psi) \in L^1(X; \mu) \times L^1(Y; \nu) \ \Big| \ & \varphi(x) + \psi(y) \leq c(x, y) \\ & \text{for } \mu\text{-a.s. } x \in X \quad \nu\text{-a.s. } y \in Y \Big\}, \end{aligned} \tag{1.14}$$

$$\begin{aligned} \Phi_c \cap C_b := \Big\{ (\varphi, \psi) \in C_b(X) \times C_b(Y) \ \Big| \ & \varphi(x) + \psi(y) \leq c(x, y) \\ & \text{for } (x, y) \in X \times Y \Big\} \end{aligned} \tag{1.15}$$

とする. このとき次が成立する.

(1) π_0 が最適輸送計画で $\int_{X \times Y} c \, \mathrm{d}\pi_0 < \infty$ をみたすとする. このとき, ある c-巡回的単調なボレル集合 $\Gamma \subset X \times Y$ で $\pi_0(\Gamma) = 1$ をみたすものがとれる. さらに c が連続なら $\Gamma \supset \mathrm{supp}[\pi_0]$ となり $\mathrm{supp}[\pi_0]$ は c-巡回的単調になる.

(2) $c < \infty$ とする. ある c-巡回的単調なボレル集合 $\Gamma \subset X \times Y$ で $\pi_0(\Gamma) = 1$ をみたすものがとれるとする. さらに

$$\mu\left(\left\{ x \in X \ \middle| \ \int_Y c(x, y) \nu(\mathrm{d}y) < \infty \right\}\right) > 0, \tag{1.16}$$

$$\nu\left(\left\{ y \in Y \ \middle| \ \int_X c(x, y) \mu(\mathrm{d}x) < \infty \right\}\right) > 0 \tag{1.17}$$

を仮定する. このとき, π_0 は最適輸送計画となり, $\int_{X \times Y} c \, \mathrm{d}\pi_0 < \infty$ となる. さらに $\sup_{(\varphi, \psi) \in \Phi_c \cap C_b} J(\varphi, \psi)$ を最大化する関数の組 $(\varphi, \psi) \in L^1(X; \mu) \times L^1(Y; \nu)$ [22]で φ が μ-a.s. で c-凹かつ $\psi = \varphi^c$ で $\varphi(x) + \psi(y) = c(x, y)$ π_0-a.s. $(x, y) \in X \times Y$ となるものがとれる. さらに次の等式が成立する.

$$\inf_{\pi \in \Pi(\mu, \nu)} C[\pi] = \sup_{(\varphi, \psi) \in \Phi_c \cap L^1} J(\varphi, \psi) = \sup_{(\varphi, \psi) \in \Phi_c \cap C_b} J(\varphi, \psi). \tag{1.18}$$

(3) (1.18) が常に成立する.

[22] 組 (φ, φ^c) は (1.18) の中辺の最大点ではあるが, 有界連続になるとは限らないので (φ, φ^c) は (1.18) の右辺の最大点とは限らない.

注意 1.25 Caffarelli による Kantorovich 双対性の解釈：φ の替わりに $-\varphi$ を考えると, (1.18) の等式

$$\inf_{\pi\in\Pi(\mu,\nu)}\int_{X\times Y} c\,\mathrm{d}\pi = \sup\left\{\int_Y \psi\,\mathrm{d}\nu - \int_X \varphi\,\mathrm{d}\mu \;\middle|\; (-\varphi,\psi)\in\Phi_c\cap C_b\right\} \tag{1.19}$$

は以下のように解釈できる:

- $\varphi(x)$: x における単位当たりの供給商品の価格.
- $\psi(y)$: y における単位当たりの需要商品の価格.
- 輸送業者は x から y に商品を輸送するに当たり, 価格 $\varphi(x)$ で供給源から購入して, 価格 $\psi(y)$ で需要先に売却する.
- $\psi(y)-\varphi(x)$ は x から y への商品の輸送に関する輸送業者の収入を表し, $J(-\varphi,\psi) = \int_Y \psi\,\mathrm{d}\nu - \int_X \varphi\,\mathrm{d}\mu$ は輸送業者の総収入を表す.
- 条件 $(-\varphi,\psi)\in\Phi_c\cap C_b$ は, 商品を輸送するにあたり全ての x, y に対し輸送業者の収入 $\psi(y)-\varphi(x)$ が x から y への輸送コスト $c(x,y)$ 以下に制限されることを意味する.
- 商品の供給と需要を統括する会社は輸送会社に最小の総コストで支払うことを希望し, 輸送業者の方は輸送コストを越えない制限下で最大収入を得ることを希望する. それらが一致することを表現するのが双対性の式 (1.19) である.

系 1.26 (有界な距離 d での Kantorovich-Rubinstein Theorem)
(E,d) をポーランド空間とし, d を有界な距離とする. $\mu,\nu\in\mathcal{P}(E)$ と $\pi\in\Pi(\mu,\nu)$ に対し, $D[\pi] := \int_{X\times X} d(x,y)\pi(\mathrm{d}x\mathrm{d}y)$ として

$$\begin{aligned}\inf_{\pi\in\Pi(\mu,\nu)} D[\pi] &= \sup_{\varphi\in L^1(\mu)} J(\varphi^{dd},\varphi^d)\\ &= \sup_{\varphi\in L^1(\mu)} J(-\varphi^d,\varphi^d) = \sup_{f\in 1\text{-Lip}(E)} J(f,-f).\end{aligned} \tag{1.20}$$

証明　d-凹な φ に対しては $\varphi=\varphi^{dd}$ なので定理 1.24 から, あきらかに $\inf_{\pi\in\Pi(\mu,\nu)} D[\pi] \le \sup_{\varphi\in L^1(\mu)} J(\varphi^{dd},\varphi^d)$. $\varphi\in L^1(E;\mu)$ に対し, $\varphi^d(y) := \inf_{x\in E}(d(x,y)-\varphi(x))$ は 1-Lipschitz 連続で, d の有界性から $\varphi^d\in L^1(E;\nu)$ となる. また, φ^d の 1-Lipschitz 連続性から

$$-\varphi^d(x) \le \inf_{y\in E}[d(x,y)-\varphi^d(y)] \le -\varphi^d(x)$$

となって, $\varphi^{dd} = -\varphi^d$ を得る. したがって

$$\begin{aligned}\sup_{\varphi\in L^1(\mu)} J(\varphi^{dd},\varphi^d) &= \sup_{\varphi\in L^1(\mu)} J(-\varphi^d,\varphi^d) \le \sup_{f\in 1\text{-Lip}(E)} J(f,-f)\\ &\le \sup_{(\varphi,\psi)\in\Phi_d} J(\varphi,\psi) = \inf_{\pi\in\Pi(\mu,\nu)} D[\pi].\end{aligned}$$

□

証明　(定理 1.24の証明) (1): $\pi_0 \in \Pi(\mu,\nu)$ を最適輸送計画とする．$\{(\varphi_n,\psi_n)\} \subset C_b(X)\times C_b(Y)$, $\varphi_n(x)+\psi_n(y) \le c(x,y)$ $(n\in\mathbb{N})$ を (1.18) の右辺を最大化する関数列とする．φ_n, ψ_n ともに各点で n に関して単調増加していると仮定していても一般性を失わない．$c_n := c-\varphi_n-\psi_n \ge 0$ on $X\times Y$ から $\int_{X\times Y} c_n\,\mathrm{d}\pi_0 = \int_{X\times Y} c\,\mathrm{d}\pi_0 - \int_X \varphi_n\,\mathrm{d}\mu - \int_Y \psi_n\,\mathrm{d}\nu \to 0$ as $n\to\infty$ なので部分列 $\{n_k\}$ で $\{c_{n_k}\}$ は 0 に π_0-a.s. で収束する．そこで

$$\Gamma := \{(x,y)\in X\times Y \mid \lim_{k\to\infty} c_{n_k}(x,y)=0, \quad c(x,y)<\infty\}$$

とおくと明らかに $\pi_0(\Gamma)=1$ となる．c が連続の場合, Γ 上で $\{c_n\}$ が c に各点において単調減少で収束することから, Dini の定理を適用することで Γ が閉集合になり, $\Gamma \supset \mathrm{supp}[\pi_0]$ を得る．

$\{(x_i,y_i)\}_{1\le i\le n}\subset\Gamma$ に対し，

$$\begin{aligned}\sum_{i=1}^n c(x_i,y_{\sigma(i)}) &\ge \sum_{i=1}^n (\varphi_{n_k}(x_i)+\psi_{n_k}(y_{\sigma(i)}))\\ &= \sum_{i=1}^n (\varphi_{n_k}(x_i)+\psi_{n_k}(y_i))\\ &= \sum_{i=1}^n (c(x_i,y_i)-c_{n_k}(x_i,y_i))\end{aligned}$$

から $k\to\infty$ として $\sum_{i=1}^n c(x_i,y_{\sigma(i)}) \ge \sum_{i=1}^n c(x_i,y_i)$, すなわち Γ の c-巡回的単調性を得る．c が連続の場合での $\mathrm{supp}[\pi_0]$ の c-巡回的単調性は $\Gamma\supset\mathrm{supp}[\pi_0]$ から明らかである．

(2): Γ を c-巡回的単調なボレル集合で $\pi_0(\Gamma)=1$ を仮定する．一般性を失うことなく, $\Gamma=\bigcup_{k=1}^\infty \Gamma_k$, 各 Γ_k はコンパクトで $c|_{\Gamma_k}$ は連続としてよい．また c_ℓ を補題 1.9で構成された c に単調増大で収束する有界連続な非負関数列とする．まず固定された $(x_0,y_0)\in\Gamma_1$ に対し, φ を

$$\varphi(x) := \inf\left\{ \sum_{i=0}^n (c(x_{i+1},y_i)-c(x_i,y_i)) \,\middle|\, n\in\mathbb{N}, x=x_{n+1}, \{(x_i,y_i)\}_{i=1}^n \subset \Gamma \right\}$$

とおく．

$$\varphi_{n,m,\ell}(x) := \inf\left\{ \sum_{i=0}^n (c_\ell(x_{i+1},y_i)-c(x_i,y_i)) \,\middle|\, n\in\mathbb{N}, x=x_{n+1}, \{(x_i,y_i)\}_{i=1}^n \subset \Gamma_m \right\}$$

とすると

$$\varphi(x) = \lim_{n\to\infty}\lim_{m\to\infty}\lim_{\ell\to\infty}\varphi_{n,m,\ell}(x)$$

がわかるので, $\varphi_{n,m,\ell}$ の上半連続性から φ のボレル可測性がわかる．

一方で $n=1$ として

$$\varphi(x) \le c(x,y_1)-c(x_1,y_1)+c(x_1,y_0)-c(x_0,y_0)$$

となるので $(x_1, y_1) = (x_0, y_0)$ として $\varphi(x) \leq c(x, y_0) - c(x_0, y_0)$ を得る．Γ が c-巡回的単調なので $\varphi(x) \geq 0$ である．また $x' \in X$ に対し，$x'_{n+1} := x'$, $x_{n+1} := x$, $x'_i = x_i$, $(x_i, y_i) \in \Gamma$, $(1 \leq i \leq n)$ とすると，

$$\sum_{i=0}^{n}(c(x'_{i+1}, y_i) - c(x'_i, y_i)) - \sum_{i=0}^{n}(c(x_{i+1}, y_i) - c(x_i, y_i)) \leq c(x', y_n) - c(x, y_n)$$

から，$x_{n+1} = x_n = x$, $y_n = y$ とおくと

$$\begin{aligned}\varphi(x') &\leq \sum_{i=0}^{n}(c(x_{i+1}, y_i) - c(x_i, y_i)) + c(x', y_n) - c(x, y_n) \\ &= \sum_{i=0}^{n-1}(c(x_{i+1}, y_i) - c(x_i, y_i)) + c(x', y) - c(x, y) + c(x, y) - c(x, y) \\ &= \sum_{i=0}^{n-1}(c(x_{i+1}, y_i) - c(x_i, y_i)) + c(x', y) - c(x, y)\end{aligned}$$

が任意の $n \in \mathbb{N}$, $(x_i, y_i) \in \Gamma$, $(1 \leq i \leq n-1)$ で成立し，

$$\varphi(x') \leq \varphi(x) + c(x', y) - c(x, y) < \infty, \quad (x, y) \in \Gamma \tag{1.21}$$

を得る．(1.21) に $x' = x_0$ を代入することで $0 \leq \varphi(x_0) \leq c(x_0, y_1) - c(x_1, y_1) + c(x_1, y_0) - c(x_0, y_0) = 0$ for $(x_1, y_1) = (x_0, y_0)$ から $\varphi(x_0) = 0$ を得る．これを再び (1.21) に代入して $\varphi > -\infty$ on $p_1(\Gamma)$ を得て $\varphi(x') < \infty$ と合わせて $\varphi \in \mathbb{R}$ μ-a.s. on X が成立する．

以下 $\psi := \varphi^c$ [23] とする．$\varphi(x') > -\infty$ なら $c(x', y) - \varphi(x') \geq c(x, y) - \varphi(x)$ から $y \in p_2(\Gamma)$ なら

$$\begin{aligned}c(x, y) - \varphi(x) &\leq \inf_{x' \in X, \varphi(x') > -\infty}(c(x', y) - \varphi(x')) \\ &\leq \inf_{x' \in X}(c(x', y) - \varphi(x')) =: \psi(y) \\ &\leq c(x, y) - \varphi(x)\end{aligned}$$

すなわち $\varphi(x) + \varphi^c(y) = c(x, y)$ が $(x, y) \in \Gamma$ で成立する．特に $\varphi + \varphi^c = c$ π_0-a.s. on $X \times Y$ となる．これより，μ-a.s. $x \in X$ に対し，$\varphi(x) + \psi(y) = c(x, y)$ をみたす y が存在する．そこで

$$A := \{x \in X \mid \varphi(x) + \psi(y) = c(x, y) \text{ をみたす } y \in Y \text{ が存在する }\}$$

とすると $\mu(A) = 1$ であり，$x \in A$ をとれば，$\varphi(x) + \psi(y) = c(x, y)$ をみたす y 毎に

$$z \mapsto c(z, y) - \varphi(z)$$

[23] $\psi := \varphi^c$ とすることで，$\varphi(x) + \psi(y) \leq c(x, y)$ が任意の $(x, y) \in X \times Y$ で成立する．

は x で最小値をとることがわかる．特に，

$$\varphi(x) \leq \inf_{z \in Y} (c(x,z) - \psi(z)) \leq c(x,y) - \psi(y) = \varphi(x)$$

から，

$$\varphi(x) = \inf_{z \in Y} (c(x,z) - \psi(z)),$$

すなわち $\varphi(x) = \psi^c(x) = \varphi^{cc}(x)$, $x \in A$ となり, φ の μ-a.s. な c-凹性を得る．

次に $\pi_0 = \int_Y \pi_y \nu(\mathrm{d}y)$ を測度 π_0 の測度分解 (後述の定理 1.32) とすると $\pi_y(\Gamma_y) = 1$ ν-a.s. から

$$\psi(y) = \int_X (c(x,y) - \varphi(x)) \pi_y(\mathrm{d}x) \quad \nu\text{-a.s. } y \in Y.$$

$y \mapsto \pi_y$ がボレル写像をなす (定理 1.32の直前を参照) ことから $\psi = \varphi^c$ は ν-可測である．

(1.16) と $\varphi \in \mathbb{R}$ μ-a.s. から $\exists x \in X$ s.t. $\int_Y c(x,y)\nu(\mathrm{d}y) < \infty$ and $\varphi(x) \in \mathbb{R}$. また (1.17) と $\psi \in \mathbb{R}$ ν-a.s. から $\exists y \in Y$ s.t. $\int_X c(x,y)\mu(\mathrm{d}y) < \infty$ and $\psi(x) \in \mathbb{R}$ なので

$$\begin{aligned}
\psi^+ &\leq c(x,\cdot) + \varphi^-(x) \in L^1(Y;\nu) \quad \text{for some } x \in X, \\
\varphi^+ &\leq c(\cdot,y) + \psi^-(y) \in L^1(X;\mu) \quad \text{for some } y \in Y
\end{aligned}$$

を得る．したがって

$$\begin{aligned}
\int_{X \times Y} c \, \mathrm{d}\pi_0 &= \int_\Gamma c \, \mathrm{d}\pi_0 = \int_\Gamma (\varphi + \psi) \mathrm{d}\pi_0 = \int_{X \times Y} (\varphi + \psi) \mathrm{d}\pi_0 \\
&= \int_X \varphi \, \mathrm{d}\mu + \int_Y \psi \, \mathrm{d}\nu < \infty
\end{aligned}$$

から

$$\begin{aligned}
\varphi^- &= (\psi - c)^+ \leq (\psi^+ - c)_+ \leq \psi^+ \in L^1(X;\mu), \\
\psi^- &= (\varphi - c)^+ \leq (\varphi^+ - c)_+ \leq \varphi^+ \in L^1(Y;\nu)
\end{aligned}$$

なので $\varphi \in L^1(X;\mu)$, $\psi = \varphi^c \in L^1(Y;\nu)$ を得る．したがってこの $\varphi \in L^1(X;\mu)$ と任意の $(\tilde{\varphi}, \tilde{\psi}) \in \Phi_c \cap L^1$ に対して

$$\begin{aligned}
J(\varphi, \varphi^c) &= \int_X \varphi \, \mathrm{d}\mu + \int_Y \varphi^c \, \mathrm{d}\nu = \int_{X \times Y} (\varphi + \varphi^c) \mathrm{d}\pi_0 \\
&= \int_{X \times Y} c(x,y) \pi_0(\mathrm{d}x\mathrm{d}y) = C[\pi_0] \quad (\because \varphi + \varphi^c = c \quad \pi_0\text{-a.s.}) \\
&\geq \int_{X \times Y} (\tilde{\varphi} + \tilde{\psi}) \mathrm{d}\pi_0 = J(\tilde{\varphi}, \tilde{\psi}) \quad (\because \tilde{\varphi} + \tilde{\psi} \leq c \quad \pi_0\text{-a.s.})
\end{aligned}$$

から

$$J(\varphi, \varphi^c) \geq \inf_{\pi \in \Pi(\mu,\nu)} C[\pi] \geq \sup_{(\tilde{\varphi}, \tilde{\psi}) \in \Phi_c \cap L^1} J(\tilde{\varphi}, \tilde{\psi}) \geq J(\varphi, \varphi^c)$$

となり, $\inf_{\pi\in\Pi(\mu,\nu)} C[\pi] = \sup_{(\tilde{\varphi},\tilde{\psi})\in\Phi_c\cap L^1} J(\tilde{\varphi},\tilde{\psi})$ を得る. 同様にして $\inf_{\pi\in\Pi(\mu,\nu)} C[\pi] = \sup_{(\tilde{\varphi},\tilde{\psi})\in\Phi_c\cap C_b} J(\tilde{\varphi},\tilde{\psi})$ も示せる. π_0 の最適性も, 任意の $\pi\in\Pi(\mu,\nu)$ に対し

$$\begin{aligned}\int_{X\times Y} c\,\mathrm{d}\pi_0 &= \int_{X\times Y} (\varphi+\varphi^c)\,\mathrm{d}\pi_0 = \int_X \varphi\,\mathrm{d}\mu + \int_Y \varphi^c\,\mathrm{d}\nu\\ &= \int_{X\times Y} (\varphi+\varphi^c)\,\mathrm{d}\pi \le \int_{X\times Y} c\,\mathrm{d}\pi\end{aligned}$$

から自明である.
(3): c が有界連続のときは (1), (2) から主張は成立する. 一般の proper 下半連続な $c: X\times Y\to[0,+\infty]$ のときは補題 1.9から c を有界連続な $c_n: X\times Y\to[0,+\infty[$ の単調増加極限で表現できる. 前半の議論から c_n に対して等式

$$\inf_{\pi\in\Pi(\mu,\nu)} C_n[\pi] = \sup_{(\varphi,\psi)\in\Phi_{c_n}\cap L^1} J(\varphi,\psi) = \sup_{(\varphi,\psi)\in\Phi_{c_n}\cap C_b} J(\varphi,\psi) \tag{1.22}$$

が成立する.

$$\inf_{\pi\in\Pi(\mu,\nu)} C[\pi] \le \lim_{n\to\infty} \inf_{\pi\in\Pi(\mu,\nu)} C_n[\pi] \tag{1.23}$$

が示されれば, (1.22) から

$$\begin{aligned}\inf_{\pi\in\Pi(\mu,\nu)} C_n[\pi] &= \sup_{(\varphi,\psi)\in\Phi_{c_n}\cap C_b} J(\varphi,\psi)\\ &= \sup\left\{\int_E \varphi(x)\mu(\mathrm{d}x) + \int_E \psi(y)\nu(\mathrm{d}y) \,\middle|\, (\varphi,\psi)\in\Phi_{c_n}\cap C_b\right\}\\ &\le \sup\left\{\int_E \varphi(x)\mu(\mathrm{d}x) + \int_E \psi(y)\nu(\mathrm{d}y) \,\middle|\, (\varphi,\psi)\in\Phi_{c}\cap C_b\right\}\end{aligned}$$

なので

$$\inf_{\pi\in\Pi(\mu,\nu)} C[\pi] \le \sup_{(\varphi,\psi)\in\Phi_c\cap C_b} J(\varphi,\psi) \le \sup_{(\varphi,\psi)\in\Phi_c\cap L^1} J(\varphi,\psi) \tag{1.24}$$

となる. 逆向きの不等式は自明なので主張を得る. (1.23) を示そう. 定理 1.6 の証明で見たように $\Pi(\mu,\nu)$ も緊密なので Prokhorov の定理 (定理 1.13) から確率測度の弱収束の位相で相対コンパクトである. $\{\pi_n^k\}$ を $\inf_\pi C_n[\pi]$ の最小化列とすると, 部分列をとることである $\pi_n\in\mathcal{P}(E\times E)$ に弱収束する. 特にこれから $f,g\in C_b(E)$ に対し $\int_E (f(x)+g(y))\pi_n(\mathrm{d}x\mathrm{d}y) = \lim_{k\to\infty}\int_E (f(x)+g(y))\pi_n^k(\mathrm{d}x\mathrm{d}y) = \int_E f(x)\mu(\mathrm{d}x) + \int_E g(y)\nu(\mathrm{d}y)$ となって $\pi_n\in\Pi(\mu,\nu)$ がわかり (この議論から $\Pi(\mu,\nu)$ はコンパクトである)

$$\inf_{\pi\in\Pi(\mu,\nu)} C_n[\pi] = \lim_{k\to\infty} C_n[\pi_n^k] = C_n[\pi_n]$$

を得る．$\Pi(\mu,\nu)$ のコンパクト性から部分列をとることで $\{\pi_n\}$ がある $\pi_* \in \Pi(\mu,\nu)$ に弱収束するとしてよい．このとき，

$$C_m[\pi_*] = \lim_{n\to\infty} C_m[\pi_n] \le \varliminf_{n\to\infty} C_n[\pi_n]$$

から

$$\inf_{\pi\in\Pi(\mu,\nu)} C[\pi] \le C[\pi_*] = \lim_{m\to\infty} C_m[\pi_*] \le \varliminf_{n\to\infty} C_n[\pi_n] = \lim_{n\to\infty} \inf_{\pi\in\Pi(\mu,\nu)} C_n[\pi]$$

となり，(1.23) を得る． □

$X = Y = E$ でコスト関数が E 上の距離関数 d のときは次のことが成立する．

定理 1.27 (Kantorovich-Rubinstein Theorem) (E,d) をポーランド空間とし，$\mu,\nu \in \mathcal{P}(E)$ と $\pi \in \Pi(\mu,\nu)$ に対し，$D[\pi] := \int_{X\times X} d(x,y)\pi(\mathrm{d}x\mathrm{d}y)$ とする．このとき，

$$\inf_{\pi\in\Pi(\mu,\nu)} D[\pi] = \sup_{f\in 1\text{-Lip}(E)} \left\{ \int_E f(x)\mu(\mathrm{d}x) - \int_E f(y)\nu(\mathrm{d}y) \right\} \tag{1.25}$$

が成立する．ここで 1-Lip(E) は距離 d に関する 1-Lipschitz 関数の全体である．

証明 d が有界な距離の場合は系 1.26で示した．実際 $d_n := d/(1+n^{-1}d)$ は $d_n \le d$ であり，各点で単調増大で下の距離に収束する．特に d_n に関する 1-Lipschitz 関数は d 関する 1-Lipschitz 関数になる．d_n で主張が成立したとする．$D_n[\pi] := \int_{E\times E} d_n(x,y)\,\pi(\mathrm{d}x\mathrm{d}y)$ として

$$\inf_{\pi\in\Pi(\mu,\nu)} D[\pi] \le \lim_{n\to\infty} \inf_{\pi\in\Pi(\mu,\nu)} D_n[\pi] \tag{1.26}$$

が示されれば，

$$\begin{aligned}\inf_{\pi\in\Pi(\mu,\nu)} D_n[\pi] &= \sup\left\{ \int_E f(x)\mu(\mathrm{d}x) - \int_E f(y)\nu(\mathrm{d}y) \,\middle|\, |f(x)-f(y)| \le d_n(x,y) \right\} \\ &\le \sup\left\{ \int_E f(x)\mu(\mathrm{d}x) - \int_E f(y)\nu(\mathrm{d}y) \,\middle|\, |f(x)-f(y)| \le d(x,y) \right\}\end{aligned}$$

から

$$\inf_{\pi\in\Pi(\mu,\nu)} D[\pi] \le \sup_{f\in 1\text{-Lip}(E)} \left\{ \int_E f(x)\mu(\mathrm{d}x) - \int_E f(y)\nu(\mathrm{d}y) \right\} \tag{1.27}$$

となる．逆向きの不等式は自明なので (1.25) を得る．(1.26) の証明は (1.23) の証明と同様なので省略する． □

1.4 Wasserstein 距離

定義 1.28 (Wasserstein 距離) (E,d) をポーランド空間, $p \in [1,\infty[$ とする. E 上の確率測度 μ,ν に対し, その p-次の Wasserstein 距離 W_p を

$$W_p(\mu,\nu) := \inf_{\pi\in\Pi(\mu,\nu)} \left(\int_{E\times E} d^p(x,y)\pi(\mathrm{d}x\mathrm{d}y)\right)^{1/p} (\leq \infty), \tag{1.28}$$

$$= \inf\left\{\mathbf{E}[d^p(X_1,X_2)] \mid (X_1)_\sharp\mathbf{P} = \mu,\ (X_2)_\sharp\mathbf{P} = \nu\right\} (\leq \infty) \tag{1.29}$$

で定める. ここで $\Pi(\mu,\nu)$ は $\mu,\nu \in \mathcal{P}(E)$ のカップリングの全体であり, X_1, X_2 は確率空間 $(\Omega,\mathcal{F},\mathbf{P})$ 上の E-値確率変数で $(X_1)_\sharp\mathbf{P} = \mu$ は X_1 の分布が μ であること [24] を表す.

注意 1.29 (1) 後の講演では我々は主に $p = 2$ の場合を扱い, 単に Wasserstein といえば 2-次の Wasserstein 空間 $(\mathcal{P}_2(E), W_2))$ 等のことを指すことにする.

(2) W_1 を Kantorovich-Rubinstein 距離とも呼ぶ. Wasserstein 距離 W_p への本質的な寄与は Kantorovich(-Rubinstein) に負う ([141, 142]). 後年, この距離は様々な研究者によって独立に発見もしくは再発見された. 確率解析における田中の公式で有名な慶応大学名誉教授・田中洋先生 [137] もそのうちの一人である. 年代順に列挙すると, Gini [64, (1914)], Salvemini [123, (1943)], Kantorovich [77, (1942)], Kantorovich–Rubinshtein [78, (1958)], Dall'Aglio [39, (1956)], Fréchet [55, (1957)], Vasershtein [140, (1969)], Mallows [96, (1972)], Tanaka [137, (1973)] となる. Gini [64, (1914)] の研究対象は $\mathbb{R}$ 上の離散分布のカップリングで W_1, W_2 を扱い, Salvemini [123, (1943)] も離散の設定で扱った. Dall'Aglio [39, (1956)] は離散分布とは限らない一次元確率分布 μ,ν の場合を扱い,

$$W_p(\mu,\nu) = \left(\int_0^1 |F_\mu^{-1}(t) - F_\nu^{-1}(t)|^p \mathrm{d}t\right)^{\frac{1}{p}} \tag{1.30}$$

を最初に示した. ここで F_μ, F_ν はそれぞれ μ,ν の分布関数である.[25] Gini [64, (1914)] は (1.30) を μ,ν が経験分布で $p = 1,2$ の場合で示している. Fréchet [55, (1957)] は W_p の距離としての性質を扱った. Mallows [96, (1972)] も統計学の範疇で W_2 のみを扱った. 一方で Tanaka [137, (1973)] は Boltzmann 方程式での平衡状態への収束の記述のため W_2 を扱った.

(3) Wasserstein 距離の名称は Vasershtein [140] に起源をもつ. この名称を最初に名付けたのは Dobrushin [43, (1970)] である. Vasershtein [140] 自身は (Kantorovich の研究を知らずに) Kantorovich-Rubinstein 距離 W_1 が研究対象であった. Vasershtein [26] は

---

[24] $\mathbf{P}(X_1 \in C) = \mu(C), \forall C \in \mathcal{B}(E)$

[25] $F_\mu(t) := \mu(]-\infty,t]), F_\nu(t) := \nu(]-\infty,t])$

[26] Leonid Vasershtein, Pennsylvania State University 数学教室教授, 専門は代数学と力学系, Moscow State University で lya I. Piatetski-Shapiro の下で 1969 年に学位を取得した. 1978 年以後に欧州・合衆国に移籍した.

英語の文献では Wasserstein とするのが通例である. これは Vaserstein がドイツ語圏を起源とする名称であることに由来する[27](Rubinstein の名称も昔は Rubinshtein であった).

(3) 確率測度間の距離として W_p 以外に以下の距離が知られている:

 (i) Lévy–Prokhorov 距離 (あるいは単に Prokhorov 距離):

$$d_P(\mu,\nu) := \inf\left\{\varepsilon>0 \;\middle|\; \exists\pi\in\Pi(\mu,\nu)\text{ s.t. }\int_{E\times E}\mathbf{1}_{\{d(x,y)>\varepsilon\}}\pi(\mathrm{d}x\mathrm{d}y)\le\varepsilon\right\}.$$

 (ii) 有界 Lipschitz 距離:

$$d_{bP}(\mu,\nu) := \sup\left\{\int_E\varphi\,\mathrm{d}\mu-\int_E\varphi\,\mathrm{d}\nu \;\middle|\; \|\varphi\|_\infty+\|\varphi\|_{\mathrm{Lip}}\le 1\right\}.$$

 (iii) 局所コンパクト距離空間 E 上の確率測度空間での weak*距離: $\{\varphi_k\}_{k=1}^\infty$ を $(C_0(E),\|\cdot\|_\infty)$ 内の可算稠密集合として

$$d_{w*}(\mu,\nu) := \sum_{k=1}^\infty 2^{-k}\left|\int_E\varphi_k\,\mathrm{d}\mu-\int_E\varphi_k\,\mathrm{d}\nu\right|.$$

 (iv) Boltzmann 方程式に有用な距離として Toscani によって導入された $\mu,\nu\in\mathcal{P}_2(\mathbb{R}^n)$ 間の Toscani 距離:

$$d_T(\mu,\nu) := \sup_{\xi\in\mathbb{R}^n\setminus\{0\}}\frac{1}{|\xi|^2}\left|\int_{\mathbb{R}^n}e^{-i\langle x,\xi\rangle}\mu(\mathrm{d}x)-\int_{\mathbb{R}^n}e^{-i\langle x,\xi\rangle}\nu(\mathrm{d}x)\right|.$$

 ただし $\int_{\mathbb{R}^n}x\mu(\mathrm{d}x)=\int_{\mathbb{R}^n}x\nu(\mathrm{d}x)$ とする.

(4) $\mu,\nu\in\mathcal{P}_p(E)$ ならば $W_p(\mu,\nu)<\infty$ である. 実際, $\pi\in\Pi(\mu,\nu)$ に対し

$$\begin{aligned}
&\left(\int_{E\times E}d^p(x,y)\pi(\mathrm{d}x\mathrm{d}y)\right)^{1/p}\\
&\le\left(\int_{E\times E}d^p(x,x_0)\pi(\mathrm{d}x\mathrm{d}y)\right)^{1/p}+\left(\int_{E\times E}d^p(x_0,y)\pi(\mathrm{d}x\mathrm{d}y)\right)^{1/p}\\
&=\left(\int_E d^p(x,x_0)\mu(\mathrm{d}x)\right)^{1/p}+\left(\int_E d^p(x_0,y)\nu(\mathrm{d}y)\right)^{1/p}<\infty.
\end{aligned}$$

(5) (1.28) の最適輸送計画を π_0 と置く. π_0 を $E\times E$ 上の分布にもつ $E\times E$-値確率変数 (X_1,X_2) を (1.29) に対する最適輸送計画と呼ぶ. 一般に, 可分距離空間 (E,d) 上の確率測度を分布にもつ E-値確率変数 X が適当な確率空間 $(\Omega,\mathcal{F},\mathbf{P})$ を構成することで常に存在する (Skorokhod の定理 [73, Theorem 2.7]).

[27] Vaserstein 自身が自分のホームページ `http://www.math.psu.edu/vstein/` で Wasserstein 距離の呼称を使用している.

例 1.30 $W_p(\delta_x,\delta_y)=d(x,y)$ が p に依らず任意の $x,y\in E$ で常に成立する.

証明　$\pi\in\Pi(\delta_x,\delta_y)$ は常に $\pi=\delta_{(x,y)}$ になる. 実際, $\pi(A\times E)=\delta_x(A)$ と $\pi(E\times B)=\delta_y(B)$ から $\pi(\{(x,y)\}^c)=\pi((\{x\}^c\times E)\cup(E\times\{y\}^c))\le\delta_x(\{x\}^c)+\delta_y(\{y\}^c)=0$. □

定義 1.31 (ボレル写像) 一般に可分距離空間 E_1,E_2 が与えられているとき, $E_1\ni x\mapsto\mu_x\in\mathcal{P}(E_2)$ を $\mathcal{P}(E_2)$-値写像とする. μ_x がボレル写像 (Borel map) とは $B\in\mathcal{B}(E_2)$ 毎に $x\mapsto\mu_x(B)$ がボレル可測関数のこととする. 測度論の単調族定理から

$$E_1\ni x\mapsto\int_{E_2}f(x,y)\mu_x(\mathrm{d}y)\tag{1.31}$$

は任意の有界 (もしくは非負) ボレル関数 $f:E_1\times E_2\to\mathbb{R}$ に対してボレル可測である. (1.31) から $\nu\in\mathcal{P}(E_1)$ に対し,

$$\langle\boldsymbol{\mu},f\rangle:=\int_{E_1}\left(\int_{E_2}f(x,y)\mu_x(\mathrm{d}y)\right)\nu(\mathrm{d}x)\tag{1.32}$$

は確率測度 $\boldsymbol{\mu}\in\mathcal{P}(E_1\times E_2)$ を定める. この $\boldsymbol{\mu}$ を $\int_{E_1}\mu_x\nu(\mathrm{d}x)$ と表す. 逆に任意の確率測度 $\boldsymbol{\mu}\in\mathcal{P}(E_1\times E_2)$ は E_1 への周辺分布 $\nu:=(p_1)_\sharp\boldsymbol{\mu}$ を用いて書けるのが次の測度分解定理である.

補題 1.32 (測度分解定理 (Disintegration Theorem)) $\boldsymbol{E}$, E をポーランド空間で $\boldsymbol{\mu}\in\mathcal{P}(\boldsymbol{E})$ とし, $\Pi:\boldsymbol{E}\to E$ をボレル写像で $\nu:=\Pi_\sharp\boldsymbol{\mu}$ とおく. このとき ν-a.s. に一意的で, かつ全ての $x\in E$ で定義されるボレル写像をなす確率測度の族 $\{\mu_x\}_{x\in E}\subset\mathcal{P}(\boldsymbol{E})$ が存在して

$$\mu_x(\boldsymbol{E}\setminus\Pi^{-1}(\{x\}))=0\quad\text{for }\nu\text{-a.s. }x\in E\tag{1.33}$$

かつ

$$\int_{\boldsymbol{E}}f(\boldsymbol{x})\boldsymbol{\mu}(\mathrm{d}\boldsymbol{x})=\int_E\left(\int_{\Pi^{-1}(\{x\})}f(\boldsymbol{x})\mu_x(d\boldsymbol{x})\right)\nu(\mathrm{d}x)\tag{1.34}$$

が任意のボレル関数 $f:\boldsymbol{E}\to[0,\infty]$ で成立する. 特に $\boldsymbol{E}:=E_1\times E_2$, $E:=E_1$, $\boldsymbol{\mu}\in\mathcal{P}(E_1\times E_2)$, $p_1(x_1,x_2):=x_1$, $\nu:=\mu_1:=(p_1)_\sharp\boldsymbol{\mu}$ なら, それぞれのファイバー $(p_1)^{-1}(\{x_1\})$ と E_2 を同一視し, μ_1-a.s. $x_1\in E_1$ に一意的に定義されるボレル写像をなす確率測度の族 $\{\mu_{x_1}\}_{x_1\in E_1}\subset\mathcal{P}(E_2)$ で $\boldsymbol{\mu}=\int_{E_1}\mu_{x_1}\mu_1(\mathrm{d}x_1)$ となるものが存在する.

証明　証明は Dellacherie-Meyer [41] の III-70 に準ずる.
$\boldsymbol{E}$ がコンパクト距離空間のとき: $f\in C(\boldsymbol{E})$ 毎に E 上の符号値測度 $\Pi_\sharp(f\boldsymbol{\mu})$ を考える. この測度は $\nu:=\Pi_\sharp(\boldsymbol{\mu})$ について絶対連続なので Radon-Nikodym の密度関数 d_f が ν-a.s. に定義される. $\mathcal{H}(\subset C(\boldsymbol{E}))$ を有理数体 $\mathbb{Q}$ を係数体とする可算線形空間で $\min,\max$ の演算で閉じていて 1 を含み $C(\boldsymbol{E})$ 内で稠密となるものとする. このような $\mathcal{H}$ は, $\{\boldsymbol{x}_i\}$ を $\boldsymbol{E}$ 内の可算稠密集合として, 1 と全ての $f_{ij}(\boldsymbol{x}):=(d(\boldsymbol{x}_i,\boldsymbol{x}_j)/2-d(\boldsymbol{x},\boldsymbol{x}_i))_+$ を含む最小の $\mathbb{Q}$-ベクトル束として構成される. $\{\boldsymbol{x}_i\}$ が稠密であるので $f_{ij}(\boldsymbol{x})$ から任意の 2 点を分離する関

数がとれる．そこでStone-Weierstrass Theorem のベクトル束版から $\mathcal{H}$ は $C(\boldsymbol{E})$ において一様位相で稠密となる．E の部分集合 A を

$$A := \{x \in E \mid f \mapsto d_f(x) \text{ が } \mathcal{H} \text{ 上の非減少 } \mathbb{Q}\text{-線形汎関数かつ } d_1(x) = 1\}.$$

と定める．$A \in \mathcal{B}(E)$ で $\nu(A) = 1$ が簡単にわかる．$x \in A$ ならば $f \mapsto d_f(x)$ はノルムが1の $C(\boldsymbol{E})$ 上の非減少 $\mathbb{R}$-線形汎関数に拡張される, すなわち $\boldsymbol{E}$ 上の確率測度となる．これを μ_x と記す．$x \notin A$ のときは μ_x は E 上の適当な確率測度で定める．
関数 $E \ni x \mapsto \langle \mu_x, f \rangle$ は $f \in \mathcal{H}$ なら $\mathcal{B}(E)$-可測であり, $f \in C(\boldsymbol{E})$ でも同じことが一様近似からわかる．したがって f が単に有界 $\mathcal{B}(\boldsymbol{E})$-可測でも $\mathcal{B}(E)$-可測となることが単調族定理から得られる．

まず, $f \in C(\boldsymbol{E})$ に対して

$$\int_{\boldsymbol{E}} f(\boldsymbol{x})\boldsymbol{\mu}(\mathrm{d}\boldsymbol{x}) = \int_E \left(\int_{\boldsymbol{E}} f(\boldsymbol{x})\mu_x(\mathrm{d}\boldsymbol{x}) \right) \nu(\mathrm{d}x) \tag{1.35}$$

を示す．これと (1.33) が示されれば, (1.34) がわかる．$\mathcal{H}$ が $C(\boldsymbol{E})$ で稠密であることから $f \in \mathcal{H}$ で (1.35) を示せばよい．このとき

$$\begin{aligned}
\int_{\boldsymbol{E}} f(\boldsymbol{x})\boldsymbol{\mu}(\mathrm{d}\boldsymbol{x}) &= \int_E \Pi_\sharp(f\boldsymbol{\mu})(\mathrm{d}x) = \int_E \frac{\mathrm{d}\Pi_\sharp(f\boldsymbol{\mu})}{\mathrm{d}\Pi_\sharp(\boldsymbol{\mu})}(x)\Pi_\sharp(\boldsymbol{\mu})(\mathrm{d}x) \\
&= \int_E \frac{\mathrm{d}\Pi_\sharp(f\boldsymbol{\mu})}{\mathrm{d}\Pi_\sharp(\boldsymbol{\mu})}(x)\nu(\mathrm{d}x) \\
&= \int_A d_f(x)\nu(\mathrm{d}x) = \int_E \langle \mu_x, f \rangle \nu(\mathrm{d}x) \\
&= \int_E \left(\int_{\boldsymbol{E}} f(\boldsymbol{x})\mu_x(\mathrm{d}\boldsymbol{x}) \right) \nu(\mathrm{d}x).
\end{aligned}$$

つぎに $\boldsymbol{E}$ が一般のポーランド空間のときに (1.35) を示す．$\boldsymbol{\mu}$ の緊密性から, あるコンパクト集合の増大列 $\{\boldsymbol{K}_n\}$ がとれて $\boldsymbol{J} := \bigcup_{n=1}^{\infty} \boldsymbol{K}_n$ は $\boldsymbol{\mu}(\boldsymbol{J}) = 1$ をみたす．$\boldsymbol{E}$ はコンパクト距離空間 $\boldsymbol{K}$ に埋め込めるので, 測度 $\boldsymbol{\mu}$ をコンパクト距離空間 $\boldsymbol{K}$ 上の測度で $\boldsymbol{\mu}(\boldsymbol{J}^c) = 0$ となるものと考えてよい．写像 $\Pi : \boldsymbol{K} \to E$ は $\boldsymbol{K} \setminus \boldsymbol{E}$ 上では適当な定値で定めておく．μ_x を $\boldsymbol{K}$ 上の測度として構成して,

$$\int_{\boldsymbol{K}} f(\boldsymbol{x})\boldsymbol{\mu}(\mathrm{d}\boldsymbol{x}) = \int_E \left(\int_{\boldsymbol{K}} f(\boldsymbol{x})\mu_x(\mathrm{d}\boldsymbol{x}) \right) \nu(\mathrm{d}x) \tag{1.36}$$

が既に得られている．$\boldsymbol{\mu}(\boldsymbol{K} \setminus \boldsymbol{J}) = 0$ と (1.36) から ν-a.s. $x \in E$ で $\mu_x(\boldsymbol{K} \setminus \boldsymbol{J}) = 0$ となるので (1.35) を得る．

最後に (1.33) を示す．$(\mathrm{Id}, \Pi)(\boldsymbol{J})$ の $x \in E$ での切り口 $(\mathrm{Id}, \Pi)(\boldsymbol{J})_x := \{\boldsymbol{y} \in \boldsymbol{E} \mid (\boldsymbol{y}, x) \in (\mathrm{Id}, \Pi)(\boldsymbol{J})\}$ は $(\mathrm{Id}, \Pi)(\boldsymbol{J})_x \subset \Pi^{-1}(\{x\})$ を満たすことと, Fubini の定理から $(\mu_x \times \delta_x)(A) =$

$\mu_x(A_x)$, $A \in \mathcal{B}(\boldsymbol{E} \times E)$ なので

$$\begin{aligned}
\int_E \mu_x(\boldsymbol{E} \setminus \Pi^{-1}(\{x\}))\nu(\mathrm{d}x) &\le \int_E \mu_x((\mathrm{Id}, \Pi)(\boldsymbol{J}^c)_x)\nu(\mathrm{d}x) \\
&= \int_E (\mu_x \times \delta_x)(\mathrm{Id}, \Pi)(\boldsymbol{J}^c)\nu(\mathrm{d}x) \\
&= \int_E (\mu_x \times \delta_x)((\boldsymbol{J}^c \times \Pi(\boldsymbol{J}^c)))\nu(\mathrm{d}x) \\
&= \int_E \mu_x(\boldsymbol{J}^c)\delta_x(\Pi(\boldsymbol{J}^c))\nu(\mathrm{d}x) \\
&= 0
\end{aligned}$$

となる．ここで ν-a.s. $x \in E$ で $\mu_x(\boldsymbol{J}^c) = 0$ を用いた． □

補題 1.33 (接着補題 (Gluing Lemma)) (E_i, μ_i) を確率測度 μ_i をともなったポーランド空間とする $(i = 1, 2, 3)$. (X_1, X_2) が (μ_1, μ_2) のカップリング, (Y_2, Y_3) が (μ_2, μ_3) のカップリングならば 3 変数確率変数 (Z_1, Z_2, Z_3) で, (Z_1, Z_2) と (X_1, X_2) が同分布[28], (Z_2, Z_3) が (Y_2, Y_3) と同分布[29] であるものが存在する．

証明　π_{12} を (X_1, X_2) の分布[30], π_{23} を (X_2, X_3) の分布[31] とする．すなわち $\pi_{12} \in \Pi(\mu_1, \mu_2)$, $\pi_{23} \in \Pi(\mu_2, \mu_3)$ とする．測度分解定理 (定理 1.32) より, μ_2-a.s. $x_2 \in E_2$ に一意的に定義されるボレル写像をなす確率測度の族 $\{(\pi_{12})_{x_2}\}_{x_2 \in E_2} \subset \mathcal{P}(E_1)$ (resp. $\{(\pi_{23})_{x_2}\}_{x_2 \in E_2} \subset \mathcal{P}(E_3)$) が存在して

$$\pi_{12} = \int_{E_2} (\pi_{12})_{x_2} \mu_2(\mathrm{d}x_2) \quad (\text{resp. } \pi_{23} = \int_{E_2} (\pi_{23})_{x_2} \mu_2(\mathrm{d}x_2))$$

をみたす．そこで

$$\pi_{123}(\mathrm{d}x_1 \mathrm{d}x_2 \mathrm{d}x_3) = (\pi_{12})_{x_2}(\mathrm{d}x_1)(\pi_{23})_{x_2}(\mathrm{d}x_3)\mu_2(\mathrm{d}x_2).$$

とおいて π_{123} を分布とする確率変数 (Z_1, Z_2, Z_3) をとればよい． □

命題 1.34 W_p は擬距離[32]である．特に $(\mathcal{P}_p(E), W_p)$ は距離空間になる．

証明　(1) $W_p(\mu, \nu) = W_p(\nu, \mu)$ の証明: $\pi \in \Pi(\mu, \nu)$ に対して

$$\overline{\pi}(A) := \pi(\{(x, y) \mid (y, x) \in A\})$$

で $\overline{\pi}$ を定めると $\overline{\pi} \in \Pi(\nu, \mu)$ となり

$$\int_{E \times E} d^p(x, y)\overline{\pi}(\mathrm{d}x\mathrm{d}y) = \int_{E \times E} d^p(y, x)\pi(\mathrm{d}x\mathrm{d}y) = \int_{E \times E} d^p(x, y)\pi(\mathrm{d}x\mathrm{d}y)$$

から $W_p(\mu, \nu) = W_p(\nu, \mu)$ が従う．

[28] $\mathbf{P}((Z_1, Z_2) \in A \times B) = \mathbf{P}((X_1, X_2) \in A \times B)$, $A \in \mathcal{B}(E_1)$, $B \in \mathcal{B}(E_2)$.
[29] $\mathbf{P}((Z_2, Z_3) \in B \times C) = \mathbf{P}((Y_2, Y_3) \in B \times C)$, $B \in \mathcal{B}(E_2)$, $C \in \mathcal{B}(E_3)$.
[30] $\pi_{12}(A \times B) = \mathbf{P}(X_1 \in A, X_2 \in B)$, $A \in \mathcal{B}(E_1)$, $B \in \mathcal{B}(E_2)$,
[31] $\pi_{23}(B \times C) = \mathbf{P}(X_2 \in B, X_3 \in C)$, $B \in \mathcal{B}(E_2)$, $C \in \mathcal{B}(E_3)$.
[32] ここでいう擬距離とは $W_p(\mu, \nu) = +\infty$ を許すが, それ以外は距離の公理系をみたすものとして扱う．

(2) $W_p(\mu,\nu) \geq 0$, $W_p(\mu,\nu) = 0 \iff \mu = \nu$ の証明: $W_p(\mu,\nu) \geq 0$ は自明. 写像 $p := (\mathrm{Id},\mathrm{Id}) : E \to \mathsf{diag}(\subset E \times E))$ を $p(x) := (x,x)$ で定めて $\pi_0 := p_\sharp\mu$ とおくと $\pi_0 \in \Pi(\mu,\mu)$ である. 実際, $\pi_0(A \times E) = \mu(p^{-1}(A \times E)) = \mu(A) = \mu(p^{-1}(E \times A)) = \pi_0(E \times A)$ である. $\pi_0(E \times E \setminus \mathsf{diag}) = \mu(p^{-1}(E \times E \setminus \mathsf{diag})) = \mu(\emptyset) = 0$ から

$$\int_{E\times E} d^p(x,y)\pi_0(\mathrm{d}x\mathrm{d}y) = \int_{E\times E\setminus\mathsf{diag}} d^p(x,y)\pi_0(\mathrm{d}x\mathrm{d}y) = 0$$

である. したがって, ${W_p}^p(\mu,\mu) \leq \int_{E\times E} d^p(x,y)\pi_0(\mathrm{d}x\mathrm{d}y) = 0$.
次に, $W_p(\mu,\nu) = 0$ とする. E がポーランド空間であることから最適輸送計画 $\pi_0 \in \Pi(\mu,\nu)$ が存在して

$$\int_{E\times E} d^p(x,y)\pi_0(\mathrm{d}x\mathrm{d}y) = {W_p}^p(\mu,\nu) = 0$$

なので, $d(x,y) = 0$ π_0-a.s. $(x,y) \in E \times E$, すなわち $\pi_0(E \times E \setminus \mathsf{diag}) = \pi_0(\{(x,y) \mid d(x,y) > 0\}) = 0$ (i.e. $\pi_0(\mathsf{diag}) = 1$) である. これより,

$$\begin{aligned}\mu(A) &= \pi_0(A \times E) = \pi_0((A \times E) \cap \mathsf{diag})) = \pi_0(\{(x,x) \mid x \in A\}) \\ &= \pi_0((E \times A) \cap \mathsf{diag})) = \pi_0(E \times A) = \nu(A).\end{aligned}$$

(3) 三角不等式の証明: $\mu_1,\mu_2,\mu_3 \in \mathcal{P}(E)$ をとる. (X_1,X_2) を (μ_1,μ_2) に対する最適輸送計画, (Y_2,Y_3) を (μ_2,μ_3) に対する最適輸送計画とする. 補題 1.33 により, 確率変数 (Z_1,Z_2,Z_3) で (Z_1,Z_2) と (X_1,X_2) が同分布[33], (Z_2,Z_3) と (Y_2,Y_3) が同分布[34] となるものが存在する. 特に (Z_1,Z_3) は (μ_1,μ_3) の輸送計画 [35] になる. その結果,

$$\begin{aligned}W_p(\mu_1,\mu_3) &\leq \mathbf{E}\left[d^p(Z_1,Z_3)\right]^{1/p} \\ &\leq \mathbf{E}\left[d^p(Z_1,Z_2)\right]^{1/p} + \mathbf{E}\left[d^p(Z_2,Z_3)\right]^{1/p} \\ &= \mathbf{E}\left[d^p(X_1,X_2)\right]^{1/p} + \mathbf{E}\left[d^p(Y_2,Y_3)\right]^{1/p} \\ &= W_p(\mu_1,\mu_2) + W_p(\mu_2,\mu_3).\end{aligned}$$

□

定義 1.35 ($\mathcal{P}_p(E)$ での収束) (E,d) を距離空間とする, $p \in [1,\infty[$ とする. $\{\mu_k\} \subset \mathcal{P}_p(E)$ が $\mu \in \mathcal{P}_p(E)$ に $\mathcal{P}_p(E)$ において弱収束することを, ある $x_0 \in E$ (全ての $x_0 \in E$) に対して成立する次の 4 つの同値な条件で定める.

(1) $\mu_k \xrightarrow{w} \mu$ かつ $\displaystyle\lim_{k\to\infty} \int_E d^p(x,x_0)\mu_k(\mathrm{d}x) = \int_E d^p(x,x_0)\mu(\mathrm{d}x)$.

(2) $\mu_k \xrightarrow{w} \mu$ かつ $\displaystyle\varlimsup_{k\to\infty} \int_E d^p(x,x_0)\mu_k(\mathrm{d}x) \leq \int_E d^p(x,x_0)\mu(\mathrm{d}x)$.

[33] $\mathbf{P}((Z_1,Z_2) \in A \times B) = \mathbf{P}((X_1,X_2) \in A \times B)\ \forall A,B \in \mathcal{B}(E)$
[34] $\mathbf{P}((Z_2,Z_3) \in A \times B) = \mathbf{P}((Y_2,Y_3) \in A \times B)\ \forall A,B \in \mathcal{B}(E)$
[35] $(Z_1)_\sharp\mathbf{P} = \mu_1$, $(Z_3)_\sharp\mathbf{P} = \mu_3$

(3) $\mu_k \xrightarrow{w} \mu$ かつ $\displaystyle\lim_{R\to\infty}\varlimsup_{k\to\infty}\int_{d(x,x_0)\geq R} d^p(x,x_0)\mu_k(\mathrm{d}x) = 0.$

(4) $|\varphi(x)| \leq C(1+d^p(x,x_0))$, $C>0$ をみたす連続関数 φ に対して (1.12) が成立する.

証明 **((1),(2),(3),(4) の同値性)** (4)$\Longrightarrow$(1)$\Longrightarrow$(2) は自明. (2)$\Longrightarrow$(3) は

$$\begin{aligned}\varlimsup_{k\to\infty}\int_{d(x,x_0)\geq R} d^p(x,x_0)\mu_k(\mathrm{d}x) &= \varlimsup_{k\to\infty}\int_E d^p(x,x_0)\mu_k(\mathrm{d}x) - \varliminf_{k\to\infty}\int_{d(x,x_0)<R} d^p(x,x_0)\mu_k(\mathrm{d}x)\\ &\leq \int_E d^p(x,x_0)\mu(\mathrm{d}x) - \int_{d(x,x_0)>R} d^p(x,x_0)\mu(\mathrm{d}x)\\ &= \int_{d(x,x_0)\geq R} d^p(x,x_0)\mu(\mathrm{d}x)\end{aligned}$$

から従う. (3)$\Longrightarrow$(4) は

$$\begin{aligned}\varlimsup_{k\to\infty}\left(\int_E \varphi\,\mathrm{d}\mu_k - \int_E\varphi\,\mathrm{d}\mu\right) &\leq \varlimsup_{k\to\infty}\left|\int_{d>R}\varphi\,\mathrm{d}\mu_k - \int_{d>R}\varphi\,\mathrm{d}\mu\right| + \varlimsup_{k\to\infty}\int_{d\leq R}\varphi\,\mathrm{d}\mu_k - \int_{d\leq R}\varphi\,\mathrm{d}\mu\\ &\leq \varlimsup_{k\to\infty}\int_{d>R}|\varphi|\,\mathrm{d}\mu_k + \int_{d>R}|\varphi|\,\mathrm{d}\mu \to 0 \text{ as } R\to\infty\end{aligned}$$

と

$$\begin{aligned}\varlimsup_{k\to\infty}\left(\int_E \varphi\,\mathrm{d}\mu - \int_E\varphi\,\mathrm{d}\mu_k\right) &\leq \varlimsup_{k\to\infty}\left|\int_{d\geq R}\varphi\,\mathrm{d}\mu - \int_{d\geq R}\varphi\,\mathrm{d}\mu_k\right| + \int_{d<R}\varphi\,\mathrm{d}\mu - \varliminf_{k\to\infty}\int_{d<R}\varphi\,\mathrm{d}\mu_k\\ &\leq \varlimsup_{k\to\infty}\int_{d\geq R}|\varphi|\,\mathrm{d}\mu_k + \int_{d\geq R}|\varphi|\,\mathrm{d}\mu \to 0 \text{ as } R\to\infty\end{aligned}$$

から従う. □

補題 1.36 (E,d) をポーランド空間, $p\in[1,\infty[$ とする. $\{\mu_k\}\subset\mathcal{P}_p(E)$ が W_p-コーシー列なら, それは緊密である. 特に Prokhorov の定理 (定理 1.13) から弱位相で相対コンパクトである.

証明

$$\int_E d^p(x,x_0)\mu_k(\mathrm{d}x) = {W_p}^p(\mu_k,\delta_{x_0})$$

は $k\in\mathbb{N}$ について有界である. $W_1\leq W_p$ より $\{\mu_k\}$ は W_1-コーシー列なので, 任意の $\varepsilon>0$ に対し, $n_0\in\mathbb{N}$ がとれ $k\geq n_0$ なら $W_1(\mu_k,\mu_{n_0})<\varepsilon^2$ とできる. 特に $j\in\{1,2,\cdots,n_0\}$ がとれ, $W_1(\mu_k,\mu_j)<\varepsilon^2$ とできる. $\{\mu_1,\mu_2,\cdots,\mu_{n_0}\}$ の緊密性からコンパクト集合 K で $\sup_{1\leq j\leq n_0}\mu_j(K^c)<\varepsilon$ をみたすものがとれる. K を有限個の開球で覆い: $K\subset\bigcup_{i=1}^m B_\varepsilon(x_i)=:U$.

$$U_\varepsilon := B_\varepsilon(U)\subset\bigcup_{i=1}^m B_{2\varepsilon}(x_i)$$

$$\phi(x) := \left(1-\frac{d(x,U)}{\varepsilon}\right)_+$$

とおくと $\mathbf{1}_U \le \phi \le \mathbf{1}_{U_\varepsilon}$ と ϕ の $1/\varepsilon$-Lipschitz 連続性がわかる．これと Kantorovich-Rubinstein Theorem (定理 1.27) から $j \le n_0$ と任意の $k \in \mathbb{N}$ に対し,

$$\begin{aligned}\mu_k(U_\varepsilon) &\ge \int_E \phi \,\mathrm{d}\mu_k = \int_E \phi \,\mathrm{d}\mu_j + \left(\int_E \phi \,\mathrm{d}\mu_k - \int_E \phi \,\mathrm{d}\mu_j\right) \\ &\ge \int_E \phi \,\mathrm{d}\mu_j - \frac{W_1(\mu_k, \mu_j)}{\varepsilon} \\ &\ge \mu_j(U) - \frac{W_1(\mu_k, \mu_j)}{\varepsilon}.\end{aligned}$$

一方で $j \le n_0$ なら $\mu_j(U) \ge \mu_j(K) \ge 1-\varepsilon$ なので $\mu_k(U_\varepsilon) \ge 1 - \varepsilon - \frac{W_1(\mu_k,\mu_j)}{\varepsilon}$ となる．$k \in \mathbb{N}$ に対し, $j = j(k)$ がとれて $W_1(\mu_k, \mu_j) < \varepsilon^2$ だったことから,

$$\mu_k(U_\varepsilon) \ge 1 - \varepsilon - \frac{\varepsilon^2}{\varepsilon} = 1 - 2\varepsilon$$

となる．以上をまとめると $\varepsilon > 0$ に対し, 有限個の点 $\{x_i\}_{1 \le i \le m}$ がとれて $\mu_k\left(\bigcup_{i=1}^m B_{2\varepsilon}(x_i)\right) \ge 1-2\varepsilon$ が任意の $k \in \mathbb{N}$ で成立する．ε を $\varepsilon/2^{\ell+1}$, $\ell \in \mathbb{N}$ に置き換えて条件をみたす $\{x_i\}_{1 \le i \le m(\ell)}$ を構成する:

$$\mu_k\left(E \setminus \bigcup_{i=1}^{m(\ell)} B_{\varepsilon/2^\ell}(x_i)\right) < 2^{-\ell}\varepsilon.$$

これより $\mu_k(E \setminus S) < \varepsilon$ とできる．ここで

$$S := \bigcap_{\ell=1}^{\infty} \bigcup_{i=1}^{m(\ell)} \overline{B_{\varepsilon/2^\ell}(x_i)}.$$

S の構成方法から S は任意の $\delta > 0$ に対して半径 δ の有限個の開球で被覆される．すなわち S は全有界な閉集合となりコンパクトである．したがって $\{\mu_k\}$ は緊密な族である． □

定理 1.37 (W_p による位相との一致) (E, d) をポーランド空間, $p \in [1, \infty[$ とする．$\{\mu_k\} \subset \mathcal{P}_p(E)$ が $\mu \in \mathcal{P}_p(E)$ に $\mathcal{P}_p(E)$ において弱収束することと $\lim_{k\to\infty} W_p(\mu_k, \mu) = 0$ は同値である．

証明 $W_p(\mu_k, \mu) \to 0$ as $k \to \infty$ とする．$\{\mu_k\}$ の緊密性から $\{\pi_k\} (\subset \Pi(\mu_k, \mu))$ も緊密となり, Prokhorov の定理 (定理 1.13) より相対コンパクトなので共通の部分列 $\{\mu_{k_i}\}$, $\{\pi_{k_i}\}$, がとれ, $\nu \in \mathcal{P}(E)$, $\pi \in \mathcal{P}(E \times E)$ にぞれぞれ弱収束する．特に $\pi \in \Pi(\nu, \mu)$ を得る．このとき輸送計画の弱収束に関する総コストの下半連続性から

$$W_p(\mu, \nu) \le \varliminf_{i\to\infty} W_p(\mu, \mu_{k_i}) = 0$$

となって $\mu = \nu$ を得て $\{\mu_k\}$ が μ に弱収束することもわかる．これだけでは, まだ $\mathcal{P}_p(E)$ における弱収束を示したことにはならない．これを示そう．

$p>1$ のときは初等的な不等式: $\forall\varepsilon>0$, $\exists C_\varepsilon>0$ [36] s.t.

$$(a+b)^p \le (1+\epsilon)a^p + C_\varepsilon b^p \quad a,b\ge 0$$

から

$$d^p(x,x_0) \le (1+\epsilon)d^p(x_0,y) + C_\varepsilon d^p(x,y)$$

を得る. $p=1$ のときは $\varepsilon=0$, $C_\varepsilon=1$ として適用する. $\pi_k\in\Pi(\mu_k,\mu)$ を最適輸送計画とすると,

$$\begin{aligned}\int_E d^p(x,x_0)\mu_k(\mathrm{d}x) &\le (1+\epsilon)\int_E d^p(x_0,y)\mu(\mathrm{d}y) + C_\varepsilon\int_{E\times E} d^p(x,y)\pi_k(\mathrm{d}x\mathrm{d}y)\\ &= (1+\epsilon)\int_E d^p(x_0,y)\mu(\mathrm{d}y) + C_\varepsilon {W_p}^p(\mu_k,\mu)\end{aligned}$$

となる. したがって

$$\varlimsup_{k\to\infty}\int_E d^p(x,x_0)\mu_k(\mathrm{d}x) \le (1+\epsilon)d^p(x_0,y)\mu(\mathrm{d}y)$$

となり, $\varepsilon\to 0$ とすることで 定義 1.35(2) を得る. 逆に $\{\mu_k\}\subset\mathcal{P}_p(E)$ が $\mu\in\mathcal{P}_p(E)$ に $\mathcal{P}_p(E)$ で弱収束したとする. $\pi_k\in\Pi(\mu_k,\mu)$ を最適輸送計画とする:

$$\int_{E\times E} d^p(x,y)\pi_k(\mathrm{d}x\mathrm{d}y) = {W_p}^p(\mu_k,\mu).$$

$\{\mu_k\}$ が μ に弱収束するので, これは緊密であり, 一点集合 $\{\mu\}$ も緊密なので $\{\pi_k\}$ も緊密になる. したがって Prokhorov の定理 (定理 1.13) から部分列をとることで π_k はある π に弱収束する. したがって $\pi\in\Pi(\mu,\mu)$ である. 特に $\int_{E\times E} d^p(x,y)\pi(\mathrm{d}x\mathrm{d}y)=0$ を得て, $\pi=(\mathrm{Id},\mathrm{Id})_\sharp(\mu)$ となり, π の最適性と同時に, 部分列に依らずに $\pi=(\mathrm{Id},\mathrm{Id})_\sharp\mu$ に $\{\pi_k\}$ が弱収束することがわかる.

　$x_0\in E$ と $R>0$ を固定する. $d(x,y)\ge R$ とすると, $d(x,x_0)$, $d(y,x_0)$ のいずれかは $R/2$ 以上で, ともに $d(x,y)/2$ 未満とはならないので

$$\begin{aligned}\mathbf{1}_{\{d(x,y)\ge R\}} &\le \mathbf{1}_{\{d(x,y)\ge R\ \&\ d(x,x_0)\ge d(x,y)/2\}} + \mathbf{1}_{\{d(x,y)\ge R\ \&\ d(y,x_0)\ge d(x,y)/2\}}\\ &\le \mathbf{1}_{\{d(x,x_0)\ge R/2\ \&\ d(x,x_0)\ge d(x,y)/2\}} + \mathbf{1}_{\{d(y,x_0)\ge R/2\ \&\ d(y,x_0)\ge d(x,y)/2\}}\end{aligned}$$

となる. これより,

$$\begin{aligned}(d^p(x,y)-R^p)_+ &\le d^p(x,y)\mathbf{1}_{\{d(x,x_0)\ge R/2\ \&\ d(x,x_0)\ge d(x,y)/2\}}\\ &\qquad + d^p(x,y)\mathbf{1}_{\{d(y,x_0)\ge R/2\ \&\ d(y,x_0)\ge d(x,y)/2\}}\\ &\le 2^p d^p(x,x_0)\mathbf{1}_{\{d(x,x_0)\ge R/2\}} + 2^p d^p(y,x_0)\mathbf{1}_{\{d(y,x_0)\ge R/2\}}\end{aligned}$$

[36] $C_\varepsilon = \frac{1+\varepsilon}{\left((1+\varepsilon)^{1/(p-1)}-1\right)^{p-1}}$.

となる．したがって

$$
\begin{aligned}
{W_p}^p(\mu_k,\mu) &= \int_{E\times E} d^p(x,y)\pi_k(\mathrm{d}x\mathrm{d}y) \\
&= \int_{E\times E}[d(x,y)\wedge R]^p\pi_k(\mathrm{d}x\mathrm{d}y) + \int_{E\times E}[d^p(x,y)-R^p]_+\pi_k(\mathrm{d}x\mathrm{d}y) \\
&\le \int_{E\times E}[d(x,y)\wedge R]^p\pi_k(\mathrm{d}x\mathrm{d}y) + 2^p\int_{d(x,x_0)\ge R/2} d^p(x,x_0)\pi_k(\mathrm{d}x\mathrm{d}y) \\
&\qquad + 2^p\int_{d(y,x_0)\ge R/2} d^p(y,x_0)\pi_k(\mathrm{d}x\mathrm{d}y) \\
&\le \int_{E\times E}[d(x,y)\wedge R]^p\pi_k(\mathrm{d}x\mathrm{d}y) + 2^p\int_{d(x,x_0)\ge R/2} d^p(x,x_0)\mu_k(\mathrm{d}x) \\
&\qquad + 2^p\int_{d(y,x_0)\ge R/2} d^p(y,x_0)\mu(\mathrm{d}y)
\end{aligned}
$$

これと定義 1.35(3) から

$$
\begin{aligned}
\varlimsup_{k\to\infty}{W_p}^p(\mu_k,\mu) &\le \lim_{R\to\infty}2^p\varlimsup_{k\to\infty}\int_{d(x,x_0)\ge R/2} d^p(x,x_0)\mu_k(\mathrm{d}x) \\
&\qquad + \lim_{R\to\infty}2^p\varlimsup_{k\to\infty}\int_{d(y,x_0)\ge R/2} d^p(y,x_0)\mu(\mathrm{d}y) \\
&= 0.
\end{aligned}
$$

□

系 1.38 (W_p の連続性) (E,d) をポーランド空間, $p\in[1,\infty[$ とする．このとき W_p は $\mathcal{P}_p(E)$ 上で連続である．すなわち, $\{\mu_k\}$ (resp. $\{\nu_k\}$) が μ (resp. ν) に $\mathcal{P}_p(E)$ において $k\to\infty$ で弱収束すれば

$$\lim_{k\to\infty}W_p(\mu_k,\nu_k) = W_p(\mu,\nu)$$

となる．

系 1.39 ($(\mathcal{P}_p(E),W_p)$ の完備性) $p\in[1,\infty[$ とする．(E,d) をポーランド空間で d を完備な距離とする．このとき $(\mathcal{P}_p(E),W_p)$ は完備である．

証明　$\{\mu_k\}$ を W_p-コーシー列とする．補題 1.36 から部分列 $\{\mu_{k_i}\}$ と $\mu\in\mathcal{P}(E)$ がとれ $\{\mu_{k_i}\}$ は μ に弱収束する．このとき

$$\int_E d^p(x,x_0)\mu(\mathrm{d}x) \le \varliminf_{i\to\infty}\int_E d^p(x,x_0)\mu_{k_i}(\mathrm{d}x) < \infty$$

から $\mu\in\mathcal{P}_p(E)$ となる．さらに W_p の下半連続性から

$$W_p(\mu,\mu_{k_j}) \le \varliminf_{i\to\infty}W_p(\mu_{k_i},\mu_{k_j})$$

より

$$\overline{\lim_{j\to\infty}} W_p(\mu, \mu_{k_j}) \le \overline{\lim_{i,j\to\infty}} W_p(\mu_{k_i}, \mu_{k_j}) = 0$$

を得る．これは部分列 $\{\mu_{k_i}\}$ が W_p で μ に収束することを示している．つまり，コーシー列が収束する部分列を含んでいることを示した．したがって収束列である． □

補題 1.40 (E,d) をポーランド空間，$p \in [1, +\infty[$ とする．$x_0 \in E$ 毎に

$$W_p(\mu,\nu) \le 2^{\frac{1}{q}} \left(\int_E d^p(x_0, x)|\mu-\nu|(\mathrm{d}x) \right)^{\frac{1}{p}}, \quad \frac{1}{p} + \frac{1}{q} = 1$$

が成立する．

証明　π を

$$\begin{aligned}
&\pi := (\mathrm{Id}, \mathrm{Id})_\sharp(\mu\wedge\nu) + \frac{1}{a}(\mu-\nu)_+ \times (\mu-\nu)_-, \\
&\mu\wedge\nu := \mu - (\mu-\nu)_+, \quad a := (\mu-\nu)_+(E) = (\mu-\nu)_-(E)
\end{aligned}$$

で定めると [37] $\pi \in \Pi(\mu,\nu)$ となる．

実際，$A, B \in \mathcal{B}(E)$ に対し

$$\begin{aligned}
\pi(A\times E) &= (\mu\wedge\nu)(A) + \frac{1}{a}(\mu-\nu)_+(A)(\mu-\nu)_-(E) \\
&= (\mu-(\mu-\nu)_+)(A) + (\mu-\nu)_+(A) = \mu(A)
\end{aligned}$$

となる．$\pi(E\times B) = \nu(B)$ も同様に確認できる．これから

$$\begin{aligned}
W_p^p(\mu,\nu) &\le \int_E d^p(x,y)\pi(\mathrm{d}x\mathrm{d}y) \\
&= \frac{1}{a}\int_E d^p(x,y)(\mu-\nu)_+(\mathrm{d}x)(\mu-\nu)_-(\mathrm{d}y) \\
&\le \frac{2^{p-1}}{a}\int_E (d^p(x,x_0) + d^p(x_0,y))(\mu-\nu)_+(\mathrm{d}x)(\mu-\nu)_-(\mathrm{d}y) \\
&= 2^{p-1}\left[\int_E d^p(x,x_0)(\mu-\nu)_+(\mathrm{d}x) + \int_E d^p(y,x_0)(\mu-\nu)_-(\mathrm{d}y)\right] \\
&= 2^{p-1}\left[\int_E d^p(x,x_0)\{(\mu-\nu)_+ + (\mu-\nu)_-\}(\mathrm{d}x)\right] \\
&= 2^{p-1}\left[\int_E d^p(x,x_0)|\mu-\nu|(\mathrm{d}x)\right].
\end{aligned}$$

□

[37] $|\mu-\nu| := (\mu-\nu)_+ + (\mu-\nu)_-$, $(\mu-\nu)_- := (\nu-\mu)_+$, $(\mu-\nu)_+(A) := \sup\{(\mu-\nu)(B) \mid B\in\mathcal{B}(E), B\subset A\}$ for $A\in\mathcal{B}(E)$.

定理 1.41 $p \in [1,+\infty[$ とする．(E,d) をポーランド空間とすると $(\mathcal{P}_p(E), W_p)$ は可分である．

証明　D を E の可算稠密集合とする．$\mathcal{P}$ を

$$\mathcal{P} := \left\{ \sum_{j=1}^{N} a_j \delta_{x_j} \;\middle|\; a_j \in \mathbb{Q}, a_j \in D, 1 \leq j \leq N, N \in \mathbb{N} \right\}$$

で定め，$\mathcal{P}$ が $(\mathcal{P}_p(E), W_p)$ で稠密であることを示そう．$\mu \in \mathcal{P}_p(E)$ をとる．任意の $\varepsilon > 0$ に対して E のコンパクト部分集合 K で

$$\int_{K^c} d^p(x_0, x)\mu(\mathrm{d}x) < \varepsilon$$

となるものをとる．K を有限個の開球 $\{B_{\varepsilon/2}(x_k)\}_{k=1}^N$ で被覆する．ここで $x_k \in D$, $1 \leq k \leq N$ である．集合 B'_k を

$$B'_k := B_\varepsilon(x_k) \setminus \bigcup_{j=1}^{k-1} B_\varepsilon(x_j), \quad 1 \leq k \leq N$$

で定めると，$\{B'_k\}_{k=1}^N$ は互いに素な K の開被覆となる．関数 $f : E \to \mathbb{R}$ を

$$f(x) := \begin{cases} x_k, & x \in B'_k \cap K, \\ x_0, & x \in K^c \end{cases}$$

と定めると，$x \in K$ で $d(x, f(x)) \leq \varepsilon$ となる．そのことから

$$\int_E d^p(x, f(x))\mu(\mathrm{d}x) \leq \varepsilon^p \mu(K) + \int_{K^c} d^p(x, f(x))\mu(\mathrm{d}x) < 2\varepsilon^p$$

となる．$(\mathrm{Id}, f)_\sharp\mu$ は μ と $f_\sharp\mu$ のカップリングなので，$W_p^p(\mu, f_\sharp\mu) < 2\varepsilon^p$ を得る．測度 $f_\sharp\mu$ は f の定め方から $\sum_{k=0}^N a_k \delta_{x_k}$ $(a_k := \mu(B'_k \cap K),\ 1 \leq k \leq N,\ a_0 := \mu(K^c))$ の形をしている．あとは a_k を有理数にしたもので μ が $\sum_{k=0}^N a_k \delta_{x_k}$ で近似できればよい．それは補題 1.40を用いて

$$\begin{aligned} W_p\left(\sum_{j=0}^N a_j \delta_{x_j}, \sum_{j=0}^N b_j \delta_{x_j}\right) &\leq 2^{\frac{1}{q}} \left(\int_E d^p(x_k, x)\mathrm{d}\left| \sum_{j=0}^N a_j \delta_{x_j} - \sum_{j=0}^N b_j \delta_{x_j} \right| \right)^{\frac{1}{p}} \\ &\leq 2^{\frac{1}{q}} \left(\int_E d^p(x_k, x)\mathrm{d} \sum_{j=0}^N |a_j - b_j| \delta_{x_j} \right)^{\frac{1}{p}} \\ &\leq 2^{\frac{1}{q}} \max_{0 \leq k, \ell \leq N} d(x_k, x_\ell) \sum_{j=0}^N |a_j - b_j|^{\frac{1}{p}} \end{aligned}$$

から示される．□

定義 1.42 (曲線, 曲線の長さ, 測地線) 距離空間 (E,d) において曲線とは連続写像 $\gamma: I \to E$ のこととする．ここで $I=[a,b]\subset\mathbb{R}$ は閉区間である．曲線 $\gamma: I\to E$ の長さ $L(\gamma)$ とは

$$L(\gamma) := \sup\left\{ \sum_{i=1}^{n} d(\gamma(t_{i-1}),\gamma(t_i)) \,\middle|\, a=t_0<t_1<\cdots<t_{n-1}<t_n=b \right\}$$

のこととする．曲線 $\gamma: I\to E$ が最短測地線とは $L(\gamma|_{[s,t]})=d(\gamma_s,\gamma_t)$ が任意の $s,t\in I, s<t$ で成立することとする．これは任意の $r<s<t$ に対して $d(\gamma_r,\gamma_t)=d(\gamma_r,\gamma_s)+d(\gamma_s,\gamma_t)$ が成立することと同値である．最短測地線は一般に一意ではないことに注意しよう．曲線 $\gamma: I\to E$ が測地線とは $|t-s|$ が十分小さい任意の $s,t\in I$, $s<t$ に対して $L(\gamma|_{[s,t]})=d(\gamma_s,\gamma_t)$ が成立することとする．

定義 1.43 (測地空間) 距離空間 (E,d) が測地空間であるとは E の任意の 2 点 $x,y\in E$ が最短測地線 $\gamma:[0,1]\to E$, $\gamma_0=x,\gamma_1=y$ で結ばれることとする．

定理 1.44 ($(\mathcal{P}_p(E),W_p)$ の測地空間の性質) (E,d) を完備可分な測地空間とする．このとき $(\mathcal{P}_p(E),W_p)$ も完備可分な測地空間になる．

証明　$(\mathcal{P}_p(E),W_p)$ の完備性は系 1.39で, 可分性は定理 1.41で示したので任意の $\mu_0,\mu_1\in\mathcal{P}_p(E)$ を結ぶ $(\mathcal{P}_p(E),W_p)$ での最短測地線 $(\mu_t)_{0\le t\le 1}$ の存在のみを示せばよい．$\mu_0,\mu_1\in\mathcal{P}_p(E)$ をとり, $\pi_0\in\Pi(\mu_0,\mu_1)$ をコスト関数 $c(x,y)=d^p(x,y)$ に関する最適輸送計画とする．写像 $\Phi_t: E\times E\to E$ を $\Phi_t(x,y):=\gamma_t^{x,y}$ で定める.[38] ここで $\gamma^{x,y}:[0,1]\to E$ は x から y への最短測地線とする．そこで $\mu_t:=(\Phi_t)_\sharp\pi_0$ とすると $(\mu_t)_{0\le t\le 1}$ が μ_0 から μ_1 への $(\mathcal{P}_p(E),W_p)$ での最短測地線を与える．実際, Kantorovich 双対性 (定理 1.24(3)) から, $0\le s<t\le 1$ に対し

$$\begin{aligned}
W_p^p(\mu_s,\mu_t) &= \sup\left\{ \int_E \varphi\,\mathrm{d}\mu_s + \int_E \psi\,\mathrm{d}\mu_t \,\middle|\, \varphi,\psi\in C_b(E),\ \varphi(x)+\psi(y)\le d^p(x,y)\ {}^\forall x,y\in E \right\} \\
&= \sup\Bigg\{ \int_{E\times E} (\varphi(\Phi_s(x,y))+\psi(\Phi_t(x,y))\,\mathrm{d}\pi_0 \,\Bigg| \\
&\qquad\qquad \varphi,\psi\in C_b(E),\ \varphi(x)+\psi(y)\le d^p(x,y)\ {}^\forall x,y\in E \Bigg\} \\
&\le \int_{E\times E} d^p(\Phi_s(x,y),\Phi_t(x,y))\pi_0(\mathrm{d}x\mathrm{d}y) = |t-s|^p \int_{X\times X} d^p(x,y)\pi_0(\mathrm{d}x\mathrm{d}y) \\
&= |t-s|^p W_p^p(\mu_0,\mu_1)
\end{aligned}$$

[38] E 上の最短測地線 $\gamma:[0,1]\to E$ の全体に一様距離 $d_\Gamma(\gamma,\eta):=\max_{t\in[0,1]} d(\gamma_t,\eta_t)$ をいれた空間を $\Gamma(E)$ とし, 端点 $x,y\in E$ にそれを結ぶ最短測地線を対応させる写像を $\Phi: E\times E\to\Gamma(E)$ とするとこれは Aumann の可測選択定理 ([88, Proposition 1] や [26, Theorem III.22] を参照) から可測であり γ に γ_t を対応させる写像 $e_t:\Gamma(E)\to E$, $e_t(\gamma):=\gamma_t$ が連続であることから $\Phi_t=\Phi\circ e_t$ は可測になる．

なので $W_p(\mu_s,\mu_t) \le (t-s)W_p(\mu_0,\mu_1)$ を得る．またこれと三角不等式から

$$\begin{aligned}W_p(\mu_0,\mu_1) &\le W_p(\mu_0,\mu_s)+W_p(\mu_s,\mu_t)+W_p(\mu_t,\mu_1)\\ &\le sW_p(\mu_0,\mu_1)+W_p(\mu_s,\mu_t)+(1-t)W_p(\mu_0,\mu_1)\end{aligned}$$

となるので $(t-s)W_p(\mu_0,\mu_1) \le W_p(\mu_s,\mu_t)$ を得て, $W_p(\mu_s,\mu_t) = (t-s)W_p(\mu_0,\mu_1)$ を得る．
□

命題 1.45 $p \in [1,+\infty[$ とする．(E,d) がコンパクト距離空間なら, $(\mathcal{P}_p(E), W_p)$ はコンパクト距離空間になる．

証明　E のコンパクト性から $\mathcal{P}_p(E) = \mathcal{P}(E)$ であり, これは常に緊密になるので Prokhorov の定理 (定理 1.13) から $\mathcal{P}_p(E) = \mathcal{P}(E)$ から弱収束する部分列, したがって W_p で収束する部分列を取り出せる．したがって, $(\mathcal{P}_p(E), W_p)$ は点列コンパクトになり $(\mathcal{P}_p(E), W_p)$ が距離空間であることからこれはコンパクトである． □

1.5 Appendix

定義 1.46 (端集合, 端点) E を線形空間, $A \subset K \subset E$ とする．A が K の端集合 (extremal set) であるとは $x, y \in K$, $t \in]0,1[$ について $(1-t)x + ty \in A$ ならば $x, y \in A$ となることとする．K の端集合 A が一点集合 $\{z\}$ のとき, z を端点 (extremal point) という．換言すると z が K の端点 (extremal point) とは $z \in K$ が K の任意の異なる 2 点を結ぶ線分の内点でないことである．

定理 1.47 (Krein-Milman's Theorem, [154, 定理 6.3, 6.4], [44, pp. 439–440]) E を Banach 空間, K を空でない E のコンパクト部分集合とする．このとき次が成立する．

(1) K には端点が存在する．

(2) K の端点の全体を $\mathcal{E}(K)$ とすると, K は $\mathcal{E}(K)$ の閉凸包 (closed convex hull) $\overline{\mathrm{co}}\,\mathcal{E}(K)$ に含まれる．特に K の任意の点は $\mathcal{E}(K)$ の元の有限凸結合で近似される．

定理 1.48 (Choquet's Theorem) E を Banach 空間, K を E のコンパクト凸集合とする．$\mathcal{E}(K)$ で K の端点の全体とする．$\ell : K \to \mathbb{R}$ を E 上の連続線形汎関数の K への制限としたとき ℓ の K 上の最小点は $\mathcal{E}(K)$ に属する．

証明　$\mathcal{E}(K)$ は閉集合なのでコンパクトである．実際, $\{z_n\} \subset \mathcal{E}(K)$ が $z \in K$ に収束したとして $x, y \in K$, $t \in]0,1[$ が $(1-t)x + ty = z$ をみたすとする．このとき, $y_n := y$, $x_n := y + \frac{1}{1-t}(z_n - y)$ とおくと $(1-t)x_n + ty_n = z_n$ となるので $x_n = y_n = z_n$ である．一方で $\{x_n\}$ は x に収束するので, これより $x = y = z$．したがって $y \in \mathcal{E}(K)$ である．Krein-Milman の定理から $x \in K$ は $\mathcal{E}(K)$ の元の有限凸結合で表現される列 $\{x_n\}$ の極限

である. すなわち x は $x_n = \sum_{k=1}^{N_n} \lambda_k^n x_k^n$, $(x_k^n \in \mathcal{E}(K), \sum_{k=1}^{N_n} \lambda_k^n = 1)$ で定まる列 $\{x_n\}$ の極限である. したがって

$$\begin{aligned} \ell(x_n) &= \ell\left(\sum_{k=1}^{N_n} \lambda_k^n x_k^n\right) = \sum_{k=1}^{N_n} \lambda_k^n \ell(x_k^n) \\ &\geq \inf_{y \in \mathcal{E}(K)} \ell(y) \sum_{k=1}^{N_n} \lambda_k^n = \inf_{y \in \mathcal{E}(K)} \ell(y) \end{aligned}$$

となり, $n \to \infty$ として $\inf_{x \in K} \ell(x) = \inf_{x \in \mathcal{E}(K)} \ell(x)$ となり, $\mathcal{E}(K)$ のコンパクト性から $\ell|_K$ の最小点は $\mathcal{E}(K)$ の点である. □

定理 1.49 (Birkhoff's Theorem) (n, n)-双確率行列の全体 $\mathcal{B}_n$ の端点は置換行列である. すなわち $\mathcal{B}_n$ の端点 $\Pi = (\pi_{ij})$ はある置換 $\sigma \in S_n$ を用いて $\pi_{ij} = \delta_{j\sigma(i)}$ と表される.

証明 Horn-Johnson [72] の Theorem 8.7.1 を参照せよ. □

2　リーマン幾何（塩谷）

最適輸送理論を用いて測度距離空間にリッチ曲率の下限条件が定義されるが，このように曲率の概念を一般化することの有用性は，リーマン多様体の収束・崩壊理論に起因する．ここでは，曲率概念の説明から初めて，リーマン幾何の研究の歴史を概観しつつ，リーマン多様体の収束・崩壊理論を解説する．これにより，曲率概念を一般の空間へ拡張することの意義を伝えたい．

2.1　曲率の大小を測る

$\mathbb{R}^3$ の曲面またはより一般にリーマン多様体 M を考える．曲率の大小をどのように測ったらよいだろうか？ 一つの方法として，球体

$$B_r(p) := \{\, q \in M \mid d(p,q) \le r \,\} \qquad (p \in M,\ r > 0)$$

の体積 $\operatorname{vol} B_r(p)$ を見てみよう．ここで，距離 d は測地的距離（リーマン距離）とする．M の曲率が大きいほど，この体積は小さくなる．また，M の測地三角形（３辺が最短測地線から成る三角形）を見る方法もある．曲率が大きいほど，測地三角形は太くなる．このような考察から，以下のことが読み取れる．

$$\text{曲率が大（正）} \Longrightarrow \text{空間が小さい}$$
$$\text{曲率が小（負）} \Longrightarrow \text{空間が大きくなりうる}$$

ここで，曲率が小さいからといって，空間が必ずしも大きいとは限らない．例えば，２次元の平坦トーラスに漏斗を接着した２次元リーマン多様体を考えると，これは大域的には円錐と同じ形をしていて，空間の広がり具合は円錐と同じと言えるが，曲率が非正である．つまり，曲率 $\ge \kappa$（κ：定数）は条件として意味があるが，曲率 $\le \kappa$ は空間の大域的な形にあまり制約を与えないのである．

話が前後するが，リーマン多様体 M の曲率について，より詳しく説明しよう．代表的な曲率として，断面曲率とリッチ曲率がある．断面曲率（sec で表す）は，M の一点 p における接空間の２次元線形部分空間 σ に対して決まる実数である．p から出て σ に接する測地線全体からなる束を考えると，これは p の近くで２次元の曲面を成すが，この曲面の p におけるガウス曲率が断面曲率に一致する．リッチ曲率（Ric で表す）は，M の一点 p における単位接ベクトル u に対して決まる．u に接するような接空間の２次線形部分空間を動かして断面曲率の平均をとった値の $n-1$ 倍がリッチ曲率である．ここで，n は M の次元とする．（正確には，リッチ曲率 Ric は $(2,0)$ 型テンソル場として定義され，ここに述べた値は $\operatorname{Ric}(u,u)$ と一致する．）

$M^n(\kappa)$ を定曲率 κ の n 次元完備単連結空間形とする．つまり，$\kappa > 0$ のとき，$M^n(\kappa)$ は半径 $1/\sqrt{\kappa}$ の球面，$\kappa = 0$ のときはユークリッド空間 $\mathbb{R}^n$，$\kappa < 0$ のときは双曲空間である．

三角形の比較として次の定理が知られている．

定理 2.1 (Toponogov の比較定理) 完備リーマン多様体 M と実定数 κ に対して，次の (1),(2) は同値である．

(1) $\sec \geq \kappa$

(2) M の任意の測地三角形 $\triangle pqr$ と，辺 qr 上の任意の点 s に対して，$M^2(\kappa)$ の測地三角形 $\triangle\tilde{p}\tilde{q}\tilde{r}$ が存在して，$d(p,q) = d(\tilde{p},\tilde{q})$, $d(q,r) = d(\tilde{q},\tilde{r})$, $d(r,p) = d(\tilde{r},\tilde{p})$ が成り立ち，$\tilde{s}$ を辺 $\tilde{q}\tilde{r}$ 上の点で $d(q,s) = d(\tilde{q},\tilde{s})$ をみたすものとすると，

$$d(p,s) \geq d(\tilde{p},\tilde{s})$$

が成り立つ．

任意の 2 点を結ぶ長さ最短の曲線が存在するような距離空間を測地空間と呼ぶが，定理 2.1(2) の条件は測地空間でも意味をもつ仮定である．

定義 2.2 (Alexandrov 空間) 定理 2.1(2) をみたす測地空間を曲率 $\geq \kappa$ の *Alexandrov* 空間という．

$M^n(\kappa)$ の半径 r の閉球体を $B^n_r(\kappa)$ とおく．その体積 $\mathrm{vol}\, B^n_r(\kappa)$ は中心の点によらない．体積の比較として以下の 2 つの不等式がある．

定理 2.3 (Bishop と Bishop-Gromov の不等式) $n \geq 2$ を整数，κ を実数とする．M が n 次元完備リーマン多様体で, $\mathrm{Ric} \geq \kappa$ をみたすならば, 任意の点 $p \in M$ と任意の $R \geq r > 0$ に対して

(1) $$\mathrm{vol}\, B_r(p) \leq \mathrm{vol}\, B^n_r(\kappa) \qquad \text{(Bishop の不等式)}$$

(2) $$\frac{\mathrm{vol}\, B_r(p)}{\mathrm{vol}\, B_R(p)} \geq \frac{\mathrm{vol}\, B^n_r(\kappa)}{\mathrm{vol}\, B^n_R(\kappa)} \qquad \text{(Bishop-Gromov の不等式)}$$

が成り立つ．

それでは，この体積の比較定理の 2 つの不等式でリッチ曲率を定義すればよいのでは？という発想が思い浮かぶが，これではあまり面白い帰結が得られない．条件としては弱すぎるのである．

2.2 リーマン幾何の歴史

リーマン多様体の研究は，局所理論が整備された後，大域的な性質にその研究の中心が移っていった．代表的な研究として球面定理がある．これは種々の幾何学的条件のもとリーマン多様体が球面に等長同型，微分同相，または同相になることを結論する形の定理である．以下に代表的な球面定理を幾つか列挙しよう．

ユークリッド空間 $\mathbb{R}^{n+1}$ 内の n 次元単位球面を $S^n(1)$ で表すことにする．

定理 2.4 (最大直径定理；Myers-Cheng [35]) $n \geq 2$ とする．n 次元閉リーマン多様体 M のリッチ曲率が $\mathrm{Ric} \geq n-1$ をみたすならば，その直径は

$$\operatorname{diam} M \leq \pi$$

をみたし，等号成立は M が n 次元単位球面 $S^n(1)$ に等長同型であるときに限る．

この定理の系として，リッチ曲率の下限が正の完備リーマン多様体はコンパクトであり，その基本群が有限であることが従う．

定理 2.5 (1/4 ピンチング定理；Berger-Klingenberg-Brendle-Schoen [18]) 単連結な閉リーマン多様体 M の断面曲率が $1/4 < \sec \leq 1$ をみたすとき，M は球面 $S^n(1)$ に微分同相となる．

これは長い歴史をもつ定理である．最初に Hopf がこの定理を予想し，Rauch が $3/4 \leq \sec \leq 1$ のときに同相版を証明し，その後 Berger と Klingenberg が同相版を証明した．微分同相版について，Im Hof-Ruh, 杉本（後藤）・塩濱，陶山らによる様々な研究を経て，最終的にリッチ・フローを用いて Brendle-Schoen が解決した．

定理 2.6 (Lichnerowicz-小畠の定理 [90, 104]) $n \geq 2$ とし，M を n 次元閉リーマン多様体で $\mathrm{Ric} \geq n-1$ と仮定する．M のラプラシアンの非ゼロ第 1 固有値 $\lambda_1(M)$ について，

$$\lambda_1(M) \geq n$$

が成り立ち，等号成立は M が n 次元単位球面 $S^n(1)$ に等長同型であるときに限る．

定理 2.7 (Grove-塩濱 [69]) n 次元閉リーマン多様体が $\sec \geq 1$ かつ $\operatorname{diam} M > \pi/2$ をみたすとき，M は n 次元球面に同相である．

このように様々な興味深い定理が生み出されていく中で，条件を少し摂動してみようというのは自然な発想である．そのような摂動についての初期の試みとして，例えば以下がある．

定理 2.8 (勝田 [79]) 任意の整数 $n \geq 2$ と任意の実数 $\Delta > 0$ に対して，ある実数 $\varepsilon(n, \Delta) > 0$ が存在して，n 次元閉リーマン多様体 M が $\mathrm{Ric} \geq n-1$，$|\sec| \leq \Delta$，かつ $\operatorname{vol} M \geq \operatorname{vol} S^n(1) - \varepsilon(n, \Delta)$ をみたすとき，M は n 次元球面 $S^n(1)$ に微分同相である．

この定理以前に関連する幾つかの結果（山口，糸川，塩濱らによる）があったが，それについては省略する．この定理は Gromov によるリーマン多様体の収束理論によって得られた．以下にこれを説明する．

定義 2.9 (Hausdorff 距離) X を距離空間とする．２つのコンパクト部分集合 $A, B \subset X$ の間の *Hausdorff* 距離 $d_{\mathrm{H}}(A, B)$ を

$$d_{\mathrm{H}}(A, B) := \inf\{\, \varepsilon > 0 \mid A \subset B_\varepsilon(B),\ B \subset B_\varepsilon(A) \,\}$$

で定義する．

定義 2.10 (Gromov-Hausdorff 距離 [67]) ２つのコンパクト距離空間 X, Y の間の *Gromov-Hausdorff* 距離 $d_{\mathrm{GH}}(X,Y)$ を

$$d_{\mathrm{GH}}(X,Y) := \inf\{\, d_{\mathrm{H}}(\varphi(X), \psi(Y)) \mid \varphi : X \to Z,\ \psi : Y \to Z \text{ は} \\ \text{ある距離空間 } Z \text{ への等長埋め込み} \,\}$$

で定義する．コンパクト距離空間の列が d_{GH} について収束するとき，*Gromov-Hausdorff* 収束するという．

定理 2.11 (Gromov の収束定理 [68]) $n \geq 1$ を整数, $\Delta, v > 0$ を実数とする．$i = 1, 2, \ldots$ に対して，M_i を n 次元閉リーマン多様体で $|\sec| \leq \Delta$, $\mathrm{vol}\, M_i \geq v$ とする．もし M_i が距離空間 X へ Gromov-Hausdorff 収束するならば，以下の (1),(2) が成り立つ．

(1) X は C^∞ 級微分可能多様体で $C^{1,\alpha}$ 級リーマン計量（$0 < \alpha < 1$ は任意）をもつ．

(2) ある番号 i_0 が存在して，$i \geq i_0$ ならば M_i は X と微分同相である．

定理 2.8の証明の概略　もし定理 2.8をみたす $\varepsilon(n, \Delta)$ が存在しなかったと仮定すると，ある閉リーマン多様体の列 M_i で $\mathrm{Ric} \geq n-1$, $|\sec| \leq \Delta$, かつ $\mathrm{vol}\, M_i \to \mathrm{vol}\, S^n(1)$ $(i \to \infty)$ をみたすものが存在する．このとき，M_i は $S^n(1)$ へ Gromov-Hausdorff 収束することが分かる．そこで，Gromov の収束定理を用いて矛盾を得る．□

さらに以下が知られている．

定理 2.12 (Cheeger-Colding [31]) 任意の整数 $n \geq 2$ に対して，ある実数 $\varepsilon(n) > 0$ が存在して，n 次元閉リーマン多様体 M が $\mathrm{Ric} \geq n-1$ かつ $\mathrm{vol}\, M \geq \mathrm{vol}\, S^n(1) - \varepsilon(n)$ をみたすとき，M は n 次元球面 $S^n(1)$ に微分同相である．

この定理の証明においても定理 2.8のように背理法から多様体の列を得るが，これが球面に Gromov-Hausdorff 収束することを証明することは遥かに難しい．

リーマン多様体の収束・崩壊を考える上で重要となるのは，何時与えられたリーマン多様体の列が収束部分列をもつか？ である．これについて次の章で詳しく説明する．

2.3 Gromov のプレコンパクト性定理

距離空間 X の離散部分集合 $\mathcal{N} \subset X$ が ε-ネットであるとは,

$$X = \bigcup_{x \in \mathcal{N}} B_\varepsilon(x)$$

が成り立つときをいう．距離空間 X がプレコンパクトまたは全有界であるとは，任意の $\varepsilon > 0$ に対して X が有限な ε-ネットをもつときをいう．言い換えると，ある有限個の点 $x_1, x_2, \ldots, x_N \in X$ が存在して

$$X = \bigcup_{i=1}^{N} B_\varepsilon(x_i)$$

をみたすことである．距離空間がコンパクトであることと，プレコンパクトかつ完備であることは必要十分であることに注意しておく．

$n \geq 2$, $\kappa \in \mathbb{R}$, $D > 0$ とする．$\mathcal{M}_{\mathrm{Ric}}(n, \kappa, D)$ を n 次元閉リーマン多様体で $\mathrm{Ric} \geq (n-1)\kappa$ かつ $\operatorname{diam} M \leq D$ をみたすものの等長同型類全体の集合とする．

$\mathcal{H}$ をコンパクト距離空間の等長同型類全体の集合とすると，これは Gromov-Hausdorff 距離に関して完備であることが知られている．

定理 2.13 (Gromov のプレコンパクト性定理 [67]) 任意の $n \geq 2$, $\kappa \in \mathbb{R}$, $D > 0$ に対して，$\mathcal{M}_{\mathrm{Ric}}(n, \kappa, D)$ は Gromov-Hausdorff 距離に関してプレコンパクトである．即ち $\mathcal{M}_{\mathrm{Ric}}(n, \kappa, D)$ の $\mathcal{H}$ における閉包はコンパクトである．特に，$\mathcal{M}_{\mathrm{Ric}}(n, \kappa, D)$ の中の任意のリーマン多様体の列は，あるコンパクト距離空間へ Gromov-Hausdorff 収束するような部分列をもつ．

この章ではこの定理の証明を簡単に紹介する．証明のキーポイントは以下の補題である．

補題 2.14 部分集合 $\mathcal{C} \subset \mathcal{H}$ について以下の (1),(2) を仮定する．

(1) ある実数 $D > 0$ が存在して，任意の $X \in \mathcal{C}$ に対して $\operatorname{diam} X \leq D$ が成り立つ．

(2) 任意の $\varepsilon > 0$ に対してある自然数 $N(\varepsilon)$ が存在して，任意の $X \in \mathcal{C}$ に対して X のある ε-ネット $\mathcal{N}$ が存在して $\mathcal{N}$ の元の個数が $\#\mathcal{N} \leq N(\varepsilon)$ をみたす．

このとき，$\mathcal{C}$ は Gromov-Hausdorff 距離に関してプレコンパクトである．

証明 証明のアイディアは，アスコリ・アルツェラの定理の証明と同じように有限次元近似をすることである．

自然数 N と実数 $D > 0$ に対して，

$$\begin{aligned}\mathcal{F}(N, D) := \{\, X \mid\ & X \text{ は距離空間の等長同型類で} \\ & \#X \leq N,\ \operatorname{diam} X \leq D \text{ をみたす} \,\}\end{aligned}$$

とおくとき，$\mathcal{F}(N, D)$ は Gromov-Hausdorff 距離に関してコンパクトである．$\mathcal{F}(N, D)$ は有限次元であり，そのコンパクト性は直感的には明らかなので，証明は省略する．

任意に $\varepsilon > 0$ をとる．条件 (2) のような $N(\varepsilon)$ が存在する．任意の $X \in \mathcal{C}$ に対して，ある ε-ネット $\mathcal{N} \subset X$ が存在して，$\#\mathcal{N} \leq N(\varepsilon)$ をみたす．$d_{\mathrm{GH}}(X, \mathcal{N}) \leq \varepsilon$ かつ $\mathcal{N} \in \mathcal{F}(N(\varepsilon), D)$ なので，

$$\mathcal{C} \subset B_{\varepsilon}(\mathcal{F}(N(\varepsilon), D))$$

が成り立つ．$\mathcal{F}(N(\varepsilon), D)$ のプレコンパクト性から，ある有限個の距離空間 $X_1, \ldots, X_n \in \mathcal{F}(N(\varepsilon), D)$ が存在して

$$\mathcal{F}(N(\varepsilon), D) \subset \bigcup_{i=1}^{n} B_{\varepsilon}(X_i)$$

をみたす．従って，三角不等式から

$$\mathcal{C} \subset \bigcup_{i=1}^{n} B_{2\varepsilon}(X_i)$$

を得る．補題が証明された．□

Gromov のプレコンパクト性定理 2.13の証明　$\mathcal{M}(n,\kappa,D)$ が補題 2.14の条件 (1),(2) をみたすことをチェックすればよい．(1) は明らかなので，(2) を確かめればよい．

任意の $\varepsilon > 0$ と $M \in \mathcal{M}(n,\kappa,D)$ をとる．$\mathcal{N} \subset M$ を ε-離散ネットとする．つまり，任意の異なる 2 点 $p, q \in \mathcal{N}$ に対して $d(p,q) > \varepsilon$ をみたすような集合とする．このとき，$B_{\varepsilon/2}(p)$ と $B_{\varepsilon/2}(q)$ は交わらないことに注意する．Bishop-Gromov の不等式から

$$\operatorname{vol} B_{\varepsilon/2}(p) \geq \frac{\operatorname{vol} B^n_{\varepsilon/2}(\kappa)}{\operatorname{vol} B^n_D(\kappa)} \operatorname{vol} M$$

が得られる．故に

$$\operatorname{vol} M \geq \sum_{p\in\mathcal{N}} \operatorname{vol} B_{\varepsilon/2}(x) \geq \frac{\operatorname{vol} B^n_{\varepsilon/2}(\kappa)}{\operatorname{vol} B^n_D(\kappa)} \operatorname{vol} M \cdot \#\mathcal{N}$$

よって

$$\#\mathcal{N} \leq \frac{\operatorname{vol} B^n_D(\kappa)}{\operatorname{vol} B^n_{\varepsilon/2}(\kappa)}$$

となる．$\mathcal{N}$ を M の極大な ε-離散ネットとおくと，これは ε-ネットとなる．実際，もしそうでないなら，ある点 $p \in M$ が存在して $d(p,\mathcal{N}) > \varepsilon$ をみたす．$\mathcal{N} \cup \{p\}$ も ε-離散ネットとなるので，$\mathcal{N}$ の極大性に反する．

$\mathcal{N}$ は M の ε-ネットで，その元の個数が M によらない数で評価されたので，補題 2.14 の条件 (2) が得られた．従って Gromov のプレコンパクト性定理が証明された．□

定理 2.13の証明と全く同じ方法で，以下が得られる．$\mathcal{A}(n,\kappa,D)$ を n 次元コンパクト Alexandrov 空間 X で曲率 $\geq \kappa$，$\operatorname{diam} X \leq D$ をみたすものの等長同型類全体の集合とする．

定理 2.15 (Gromov のプレコンパクト性定理（Alexandrov 空間版）) 任意の $n \geq 2$, $\kappa \in \mathbb{R}$, $D > 0$ に対して，$\mathcal{A}(n,\kappa,D)$ は Gromov-Hausdorff 距離に関してプレコンパクトである．

2.4　リーマン多様体の収束・崩壊とその極限

リーマン多様体の収束理論の応用例として以下を挙げる．$\mathcal{M}_{\mathrm{sec}}(n,\kappa,D,v)$ を n 次元リーマン多様体 M で $\sec \geq \kappa$, $\operatorname{diam} M \leq D$, $\operatorname{vol} M \geq v$ をみたすものの等長同型類全体の集合とする．

定理 2.16 (有限性定理；**Cheeger-Grove-Petersen-Wu-Perelman [120]**) 任意の $n \geq 1$，$\kappa \in \mathbb{R}$，$D, v > 0$ に対して，$\mathcal{M}_{\mathrm{sec}}(n, \kappa, D, v)$ は高々有限個の位相同型類しか含まない．

これは最初，ホモトピー版が証明された．最終的には，Perelman が Gromov-Hausdorff 距離と Alexandrov 空間の理論を用いて証明した．以下の定理が主要な役割を果たす．

定理 2.17 (**Perelman** の安定性定理 [**120**]) 任意の Alexandrov 空間 $X \in \mathcal{A}(n, \kappa, D)$ に対して，ある $\varepsilon(X) > 0$ が存在して $Y \in \mathcal{A}(n, \kappa, D)$ が $d_{\mathrm{GH}}(X, Y) < \varepsilon(X)$ をみたすならば，X と Y は互いに同相である．

有限性定理の証明の概略　$\mathcal{A}(n, \kappa, D, v)$ を $X \in \mathcal{A}(n, \kappa, D)$ で n 次元 Hausdorff 測度が $\mathcal{H}^n(X) \geq v$ なるもの全体の集合とする．$\mathcal{M}_{\mathrm{sec}}(n, \kappa, D, v)$ は $\mathcal{A}(n, \kappa, D, v)$ に含まれるので，$\mathcal{A}(n, \kappa, D, v)$ に含まれる位相同型類が高々有限であることを示せば十分である．Gromov のプレコンパクト性定理と少しの議論により，$\mathcal{A}(n, \kappa, D, v)$ は Gromov-Hausdorff 距離に関してコンパクトであることが分かる．よって，ある有限個の Alexandrov 空間 $X_1, \ldots, X_N \in \mathcal{A}(n, \kappa, D, v)$ が存在して，

$$\mathcal{A}(n, \kappa, D, v) \subset \bigcup_{i=1}^{N} B_{\varepsilon(X_i)}(X_i)$$

が成り立つ．ここで，$\varepsilon(X)$ は Perelman の安定性定理のものである．安定性定理により，$B_{\varepsilon(X_i)}(X_i) \cap \mathcal{A}(n, \kappa, D, v)$ の任意の元は X_i に同相なので，$\mathcal{A}(n, \kappa, D, v)$ は高々 N 個の位相同型類しか含まない．□

ここに紹介した定理たちはほんの一部であり，他にも多くの重要な結果がある．ここでは，Gromov-Hausdorff 収束における極限の次元が，列の多様体の次元と同じだが，極限の次元が列の多様体の次元より真に小さくなるときは，問題がより難しくなり，そのような収束を崩壊とよぶ．崩壊理論の研究は Cheeger, Fukaya, Gromov, Perelman，山口らによって整備された．断面曲率が下に有界な完備リーマン多様体の列の Gromov-Hausdorff 極限は Alexandrov 空間となるが，そのような崩壊現象の解明は，Alexandrov 空間の理解が本質的であった．その中で一つの代表的な定理が以下である．

定理 2.18 (体積崩壊定理；塩谷・山口 [**129, 130**]) ある定数 $\varepsilon_0 > 0$ が存在して，3 次元閉リーマン多様体 M で $\sec \geq -1$ かつ $\mathrm{vol}\, M \leq \varepsilon_0$ をみたすものは，グラフ多様体に限る．

この定理は最初 Perelman [121] により証明なしで主張され，塩谷・山口が証明した．これは Perelman によるポアンカレ予想および幾何化予想の解決における一つの重要なステップであった．

このように断面曲率の下限条件において，Alexandrov 空間を解析することで，収束・崩壊現象を解明する手段はかなり有効な研究手段であった．同じことがリッチ曲率に対して期待できるが，リッチ曲率の下限条件を備えた距離空間は最適輸送理論を用いて定義される．

また他方で，Cheeger-Colding [31, 32, 33] はリッチ曲率が下に有界かつ次元が上に有界なリーマン多様体の列の極限空間の構造を詳しく研究した．現在，最適輸送の意味でリッチ曲率が下に有界な距離空間について，Cheeger-Colding の結果を一般化するという結果も得られつつある．

3　曲率次元条件（太田）

この講演の内容については，[148] に（情報は若干古いが）概説がある．興味を持たれた方はそちらも参照されたい．

3.1　リッチ曲率の下限の特徴づけ

まずリーマン多様体において，リッチ曲率がある定数以上であるという性質が，Wasserstein 空間上のある汎関数（エントロピー）の凸性で特徴づけられることを述べる．このエントロピーの凸性を曲率次元条件と呼ぶ．

(M,g) を n 次元完備連結リーマン多様体（$n \geq 2$，$\partial M = \emptyset$），vol_g をその体積測度とし，リーマン距離関数を単に d で表す．簡単のため，M は常にコンパクトであるとする．

3.1.1　重みつきリッチ曲率

曲率次元条件における「次元」の役割を理解するためには，vol_g に M 上の正値関数による重みをつけた測度

$$\mathfrak{m} = \mathrm{e}^{-V}\,\mathrm{vol}_g, \qquad V \in C^\infty(M)$$

を考えた方が良い．すると，リッチ曲率は測度の振る舞いを制御するものであるので，V に応じて変形する必要がある．

定義 3.1 (重みつきリッチ曲率)　$N \in (n,\infty)$ と $v \in T_xM$ に対し，

$$\mathrm{Ric}_N(v) := \mathrm{Ric}_g(v,v) + \mathrm{Hess}\,V(v,v) - \frac{\langle \nabla V(x), v\rangle^2}{N-n}$$

（$\mathrm{Ric}_g : TM \times TM \longrightarrow \mathbb{R}$ は g についてのリッチテンソル）．極限として，

$$\mathrm{Ric}_\infty(v) := \mathrm{Ric}_g(v,v) + \mathrm{Hess}\,V(v,v),$$

$$\mathrm{Ric}_n(v) := \begin{cases} \mathrm{Ric}_g(v,v) + \mathrm{Hess}\,V(v,v) & \text{if } \langle \nabla V(x), v\rangle = 0, \\ -\infty & \text{if } \langle \nabla V(x), v\rangle \neq 0. \end{cases}$$

定義から直ちに次の性質がわかる．

補題 3.2　(i) $v \in TM$ と $c \in \mathbb{R}$ に対し，$\mathrm{Ric}_N(cv) = c^2\,\mathrm{Ric}_N(v)$.

(ii) 重みがない場合（$\mathfrak{m} = \mathrm{vol}_g$，つまり $V \equiv 0$），全ての $N \in [n,\infty]$ に対して $\mathrm{Ric}_N(v) = \mathrm{Ric}_g(v,v)$.

(iii) Ric_N は N について単調非減少である．つまり，$n \leq N < N' \leq \infty$ ならば，$\mathrm{Ric}_N(v) \leq \mathrm{Ric}_{N'}(v)$.

注意 3.3 (a) Ric_∞ を下から押さえた状況での幾何・解析は，Lichnerowicz [89] による先駆的な仕事の後，Bakry–Émery [11] らによって線形半群や関数不等式の枠組で組織的に研究された．Ric_∞ は **Bakry–Émery** テンソル，または **Bakry–Émery–Ricci** テンソルとも呼ばれる．その後，Ric_N が Bakry [10] や Qian [122] により導入され，解析的・幾何的な興味から研究が進んだ（[91] など参照）．Bakry らの理論では，ラプラシアンを一般化したような線形作用素について，$\mathrm{Ric}_N \ge K$ での Bochner 不等式に対応する不等式を曲率次元条件と呼び，その下での解析を展開した（**Γ-calculus** と呼ばれる，桒田氏の講演参照）．この講演で扱う最適輸送理論に基づく曲率次元条件は彼らの理論からその名前を流用しており，区別するために Bakry らのものを「解析的曲率次元条件」，最適輸送理論に基づくものを「幾何的曲率次元条件」と呼ぶことがある．

(b) 以前は $N < n$ の場合は考えない，若しくは形式的に $\mathrm{Ric}_N(v) = -\infty$ とされていたが，最近 $N < 0$（または更に $N < 1$）の場合にも（同じ定義による）Ric_N を考えることに意味があることがわかってきている．例えば，Bochner 不等式が $N < 0$ でも成り立つ（[83, 108]）．

以降，実数 $K \in \mathbb{R}$ に対し $\mathrm{Ric}_N(v) \ge K|v|^2$ が全ての $v \in TM$ で成り立つことを，「$\mathrm{Ric}_N \ge K$」で表す．$\mathrm{Ric}_N \ge K$ を満たす $(M, g, \mathfrak{m})$ は，「リッチ曲率が K 以上，かつ次元が N 以下」であるように振る舞うことが知られている．例えば，$N \in [n, \infty)$ に対し $\mathrm{Ric}_N \ge K$ であるとき，Bishop–Gromov 体積比較定理の拡張：

$$\frac{\mathfrak{m}(B_R(x))}{\mathfrak{m}(B_r(x))} \le \frac{\int_0^R \mathbf{s}_{K/(N-1)}(t)^{N-1}\,dt}{\int_0^r \mathbf{s}_{K/(N-1)}(t)^{N-1}\,dt} \qquad \forall x \in M,\ 0 < r < R$$

が成り立つ．ここで，

$$\mathbf{s}_\kappa(r) := \begin{cases} \sqrt{1/\kappa}\sin(r\sqrt{\kappa}) & (\kappa > 0), \\ r & (\kappa = 0), \\ \sqrt{-1/\kappa}\sinh(r\sqrt{-\kappa}) & (\kappa < 0) \end{cases} \tag{3.1}$$

は方程式

$$f'' + \kappa f = 0, \quad f(0) = 0, \quad f'(0) = 1$$

の解で，比較定理で頻繁に現れる．

3.1.2 エントロピー

次に，曲率次元条件で主役を務める Wasserstein 空間上の汎関数（エントロピー）を定義する．

定義 3.4 (エントロピー) (1) $\mu \in \mathcal{P}(M)$ の $\mathfrak{m}$ についての相対エントロピーを，μ が $\mathfrak{m}$ と絶対連続である場合（$\mu = \rho\mathfrak{m}$ と表せる）には

$$\mathrm{Ent}_{\mathfrak{m}}(\mu) := \int_M \rho \log \rho \, d\mathfrak{m},$$

それ以外のときは $\mathrm{Ent}_{\mathfrak{m}}(\mu) := \infty$ と定義する．

(2) $N \in (1,\infty)$ に対し，$\mu \in \mathcal{P}(M)$ の $\mathfrak{m}$ についての **Rényi（–Tsallis）**エントロピーを，μ を $\mu = \rho\mathfrak{m} + \mu_s$ と $\mathfrak{m}$ と絶対連続な部分 $\rho\mathfrak{m}$ と特異な部分 μ_s にルベーグ分解して，

$$S_N(\mu) := -\int_M \rho^{(N-1)/N}\,d\mathfrak{m}$$

と定義する．

上の (1) において，Jensen の不等式より

$$\mathrm{Ent}_{\mathfrak{m}}(\mu) = \mathfrak{m}(M)\int_M \rho\log\rho\,\frac{d\mathfrak{m}}{\mathfrak{m}(M)} \geq \mathfrak{m}(M)\frac{1}{\mathfrak{m}(M)}\log\frac{1}{\mathfrak{m}(M)} = -\log\mathfrak{m}(M).$$

従って $\mathrm{Ent}_{\mathfrak{m}}(\mu)$ は well-defined である（$\mathfrak{m}(M)=\infty$ の場合には，$\int_{\{\rho>1\}}\rho\log\rho\,d\mathfrak{m}=\infty$ なら $\mathrm{Ent}_{\mathfrak{m}}(\mu) := \infty$ とする）．

相対エントロピーは，熱力学や情報理論で基本的な概念であるボルツマン・シャノンエントロピーの符号を逆にしたものである（従って，自然な拡散で相対エントロピーは減少する）．実際，集合 $A \subset M$ 上の一様分布 $\mu_A = \mathfrak{m}(A)^{-1}\cdot\mathfrak{m}|_A$ については

$$\mathrm{Ent}_{\mathfrak{m}}(\mu_A) = \int_A \frac{1}{\mathfrak{m}(A)}\log\frac{1}{\mathfrak{m}(A)}\,d\mathfrak{m} = -\log\mathfrak{m}(A) \tag{3.2}$$

となり，$\mathfrak{m}(A)$ が大きいほど（広範囲に一様に分布しているほど）$\mathrm{Ent}_{\mathfrak{m}}(\mu_A)$ は小さくなる．S_N は相互作用のあるより複雑な系についての統計力学などに関連するものである．

3.1.3　リッチ曲率の下限の特徴づけ

定義3.4のエントロピーを用いると，重みつきリッチ曲率がある定数以上であるという性質が次のように特徴づけられる．次の定理の (i) で $\mathfrak{m} = \mathrm{vol}_g$ の場合は Renesse–Sturm [126]，それ以外は Sturm [131, 132, 133] と Lott–Villani [92, 93] による（$\mathrm{Ric}_\infty \geq K$ が $\mathrm{Ent}_{\mathfrak{m}}$ の K 凸性を導くことについては，Cordero-Erausquin et al [37, 38] による先行研究がある）．

定理 3.5 (リッチ曲率の下限の特徴づけ) $K \in \mathbb{R}$ とする．

(i) $(M,g,\mathfrak{m})$ が $\mathrm{Ric}_\infty \geq K$ を満たすことは，$\mathrm{Ent}_{\mathfrak{m}}$ が K 凸であること，つまり任意の $\mu_0,\mu_1 \in \mathcal{P}(M)$ とその間の W_2 についての最短測地線 $(\mu_t)_{t\in[0,1]}$ に対し

$$\mathrm{Ent}_{\mathfrak{m}}(\mu_t) \leq (1-t)\,\mathrm{Ent}_{\mathfrak{m}}(\mu_0) + t\,\mathrm{Ent}_{\mathfrak{m}}(\mu_1) - \frac{K}{2}(1-t)tW_2^2(\mu_0,\mu_1) \tag{3.3}$$

が全ての $t \in (0,1)$ で成り立つこと，と同値である．

(ii) $N \in [n,\infty)$ に対し，$(M,g,\mathfrak{m})$ が $\mathrm{Ric}_N \geq K$ を満たすことは，任意の $\mathfrak{m}$ に絶対連続な $\mu_0,\mu_1 \in \mathcal{P}(M)$ とその間の W_2 についての最短測地線 $(\mu_t)_{t\in[0,1]}$ に対し，

$$\begin{aligned} S_N(\mu_t) \leq -\int_{M\times M} \Big\{ &(1-t)\beta_{K,N}^{1-t}\big(d(x,y)\big)^{1/N}\rho_0(x)^{-1/N} \\ &+ t\beta_{K,N}^{t}\big(d(x,y)\big)^{1/N}\rho_1(x)^{-1/N}\Big\}\,\pi(dxdy) \end{aligned} \tag{3.4}$$

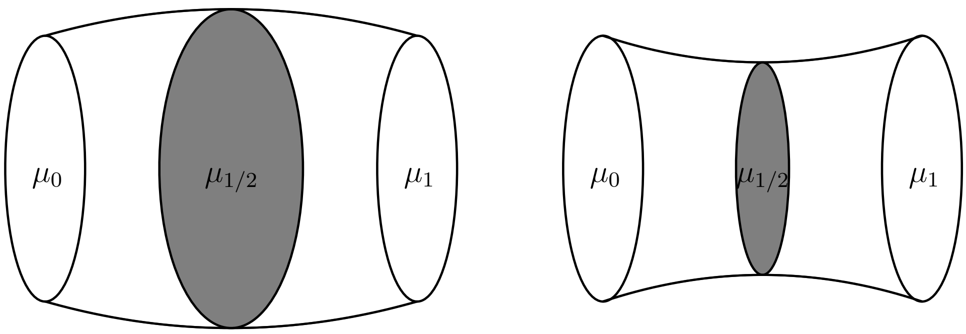
が全ての $t \in (0,1)$ で成り立つことと同値である．ここで，$i=0,1$ に対し $\mu_i = \rho_i \mathfrak{m}$ とし，π を μ_0 と μ_1 の最適カップリングとした．

不等式 (3.4) 中の $\beta^t_{K,N}$ は，(3.1) の $\mathbf{s}_\kappa$ を用いて次のように定義される：$t \in (0,1)$ に対し，

$$\beta^t_{K,N}(r) := \left(\frac{\mathbf{s}_{K/(N-1)}(tr)}{t\mathbf{s}_{K/(N-1)}(r)}\right)^{N-1}, \qquad \beta^t_{K,\infty}(r) := \mathrm{e}^{K(1-t^2)r^2/6}. \tag{3.5}$$

(3.3) は $d^2[\mathrm{Ent}_{\mathfrak{m}}(\mu_t)]/dt^2 \geq KW_2^2(\mu_0,\mu_1)$ を弱い意味で（積分した形で）表したものであり，$\mathcal{P}(M)$ をリーマン多様体と見なした場合には $\mathrm{Hess}\,\mathrm{Ent}_{\mathfrak{m}} \geq K$ に当たる．(3.4) は絶対連続でない μ_0, μ_1 に対しても書くことはできるが，煩雑になるので割愛する．

幾何的なイメージでは，曲率が大きいほど μ_t は μ_0, μ_1 に比べて「広がる」ため，エントロピーが小さくなり，よってより凸になる（下図）．



$\mathrm{Ric} > 0$ $\qquad\qquad$ $\mathrm{Ric} < 0$

エントロピーの凸性のイメージ

注意 3.6 $K=0$ のときは $\beta_{0,N} \equiv 1$ となり，よって (3.4) は S_N の凸性（0 凸性）：

$$S_N(\mu_t) \leq (1-t)S_N(\mu_0) + tS_N(\mu_1)$$

に他ならない．しかし，$K \neq 0$ については S_N の K 凸性は良い条件ではない．具体的には，$K>0$ に対し S_N が K 凸になることはなく，$K<0$ の場合に S_N が K 凸であることは $\mathrm{Ric}_N \geq 0$ であることと同値である（[131, 113]）．

$\mathrm{Ric}_N \geq K$ からエントロピーの凸性 (3.3), (3.4) を得る方法の概略を述べる．$N \in [n,\infty)$ とする（$N=\infty$ の場合も同様）．簡単のため，μ_0, μ_1 が共に $\mathfrak{m}$ に絶対連続なときのみ考える．ユークリッド空間の場合と同様に，μ_0 から μ_1 への最短測地線（最適輸送）は一意であり，M の測地線に沿った押し出しとなる．具体的には，ある種の凸性を満たす関数 $\varphi : M \longrightarrow \mathbb{R}$ が存在し，

$$T_t(x) := \exp_x\big(t\nabla\varphi(x)\big)$$