

重なりを測るのか、隔たりを測るのか

— KL とワッサーズタインは競合しない：台が交わらないとき何が起きるか、そして収束率が $d = 3$ で性格を変えること — [論文139]

山岸 勇輝 (Yuuki YAMAGISHI)

Yamagishi Complexity Torus Lab, Japan | 山岸勇輝ラボ

2026年8月

— From thinking to syncing. —

要旨

論文137 は KL ダイバージェンスが距離でないことを書いた。本稿はその対になるワッサーズタイン距離を置き、二つが競合しないこと——別の問いに答えていること——を書く。新しい数学定理も新しい物理法則も主張しない。第一に、ワッサーズタインは距離である。標本三つ組を 3000 通り取り、非対称は 0 件、三角不等式の破れも 0 件であった。論文137 の KL は同じ試験で 25.4% 破れていた (第2節)。第二に、これが本稿の芯である。分ける実例は「台が交わらないとき」である。距離 d だけ離れた二つの点質量について、KL はどの d でも無限大であり、 W_2 はちょうど d である ($d = 0.5, 1, 5, 100$ で確認)。重なりがあるうちは両方が離れ具合を見る——等分散の正規分布なら KL は $d^2/2$ 、 W_2 は d ($d = 5$ で 12.500000 対 5、閉じた形と厳密一致)。「二つの分布はどれだけ違うか」は二つの問いである——重なっているか、そしてどれだけ離れているか (第3節)。第三に、ワッサーズタインが優れているのではない。前提が多いのである。ワッサーズタインは台の上の距離を要求する。順序のない三つのカテゴリに対して KL は 0.223805 と定まるが、ワッサーズタインは「赤と緑がどれだけ離れているか」を誰かが決めるまで形にならない (第4節)。第四に、熱方程式が二度書ける。点から出発した熱核について $W_2 = \sqrt{2t}$ である ($t = 0.25, 1, 4$ で確認)。 $[t] = L^2$ という論文119 の放物型の尺度がそのまま出る。そしてオットーの定理により、熱方程式はこの計量におけるエントロピーの勾配流である——同じ方程式が、二通りに書ける (第5節)。第五に、収束率が $d = 3$ で性格を変える。同一分布からの二標本間の W_1 (真値 0) を測ると、標本を 4 倍したときの改善は $d = 1$ で 2.18 倍、 $d = 3$ で 1.59 倍、 $d = 8$ で 1.22 倍である。率は $d \leq 2$ で $n^{-1/2}$ 、 $d \geq 3$ で $n^{-1/d}$ ——中心極限の率と被覆の率が $d = 3$ で入れ替わる (第6節)。

本稿の射程：新しい数学定理も新しい物理法則も主張しない。モンジュ、カントロヴィチ、オットーの勾配流、フルニエ＝ギヤンの収束率はいずれも標準的である。本稿が加えるのは、距離としての性質を KL と並べて数えたこと、台が変わらない場合を分離子として書いたこと、前提の非対称（地の距離が要るかどうか）を書いたこと、収束率の二つの領域を数値で確かめたことだけである。ワッサーズタインが KL より優れているとは言わない。数値について： $d=1$ の輸送費用は整列による厳密計算である。 $d \geq 2$ は対数領域シンクホーン法（エントロピー正則化）によった。実装は一次元の厳密解と突き合わせて検証した（相対誤差 **0.1~1.1%**）。スライス版は用いていない——それは一次元射影の平均であって収束率が次元に依らず、本稿が調べている効果を隠してしまうからである。

1. 記法と位置づけ

二つの確率測度 μ, ν について、 p -ワッサーズタイン距離を

$$W_p(\mu, \nu) = \left(\inf_{\gamma} \int \|x - y\|^p d\gamma(x, y) \right)^{1/p} \quad (1)$$

と書く（ γ は周辺分布が μ, ν である結合）。

位置づけ：既存の定理を並べ、その射程を書く。新しい定理は主張しない。論文137の続きであり、対である。

2. ワッサーズタインは距離である

検算 2.1：直線上の正規分布から取った標本三つ組を 3000 通り作り、 W_1 について対称性と三角不等式を調べた。

ワッサーズタイン（本稿）		KL（論文137）
非対称の件数	0 / 3000	常に非対称
三角不等式の破れ	0 / 3000	25.4%

（証明終）

観察 2.2

ワッサースタインは距離の公理を満たす。ゆえに「どちらがどれだけ遠いか」の一貫した順序を与える——論文137でKLが与えられなかったものである。

3. 分ける実例は、台が変わらないとき

検算 3.1（重なりがある場合）：等分散の正規分布を d だけずらすと、本稿の数値計算では

d	KL（数値）	$d^2/2$	W_2
0.5	0.125000	0.125000	0.5
1.0	0.500000	0.500000	1.0
5.0	12.500000	12.500000	5.0

閉じた形と厳密に一致する。重なりがあるうちは、どちらも離れ具合を見ている（KLは d^2 で、 W_2 は d で）。（証明終）

検算 3.2（重なりが無い場合）：距離 d だけ離れた二つの点質量について

d	KL	W_2
0.5	無限大	0.5
1.0	無限大	1.0
5.0	無限大	5.0
100.0	無限大	100.0

KLはどの d でも同じ値（無限大）を返し、 d を区別しない。 W_2 は d をそのまま返す。（証明終）

観察 3.3（これが芯である）

「二つの分布はどれだけ違うか」という問いは、実は二つの問いである。

問い	答える道具	台が変わらないとき
重なっているか	KL (論文137)	無限大 (それ以上言えない)
どれだけ離れているか	ワッサースタイン	距離をそのまま返す

一語「隔たり」が、この二つを一緒にしている (論文17)。

4. ただし、前提が違う

観察 4.1

(1) には $\|x - y\|$ が現れる。ワッサースタインは台の上の距離を要求する。KL の定義 $\sum p \log(p/q)$ にはそのようなものが現れない。

検算 4.2：順序のない三つのカテゴリ上の分布 $p = (0.5, 0.3, 0.2)$ 、 $q = (0.2, 0.5, 0.3)$ について、本稿では $D(p||q) = 0.223805$ を得た。一方ワッサースタインは、赤と緑がどれだけ離れているかを誰かが決めるまで形にならない。(証明終)

観察 4.3 (優劣ではない)

ワッサースタインは単純に「より良い」対象ではない。要求する入力が多く、その入力が存在しない場では作れない。二つの道具、二つの前提である (論文95)。

5. 熱方程式は、二度書ける

検算 5.1：点から出発した熱核は $\rho(\cdot, t) = N(0, 2t)$ だから $W_2(\delta_0, \rho_t) = \sqrt{2t}$ である。本稿では $t = 0.25, 1, 4$ について 0.708606、1.414860、2.829112 を得た (理論値 0.707107、1.414214、2.828427)。(証明終)

観察 5.2 (尺度)

$W_2 \propto \sqrt{t}$ は $[t] = L^2$ である——論文119 が第四の時計について書いた放物型の尺度がそのまま出る。

観察 5.3 （オットーの定理）

熱方程式は、ワッサーズタイン計量におけるエントロピーの勾配流である（オットー 2001、ジョーダン＝キンダーレーラー＝オットー 1998）。同じ方程式が二通りに書ける——偏微分方程式として、そして最急降下として。後者の読みは、この計量を用意して初めて意味を持つ。

そして体系の中に同じ形がある——論文110 はリッチ・フローが繰り込み群の流れであることを書いた。「同じ方程式の二つの記述」は、そこでも起きていた。ただし住所は別である（あちらは弦の世界面、こちらは確率測度の空間）。

6. 収束率は、 $d = 3$ で性格を変える

検算 6.1：同一の標準正規分布から取った二つの標本のあいだの W_1 ——真の値は 0 だから、測っているのは推定誤差だけである。

d	$n = 64$	$n = 256$	比	予測
1	0.22411	0.10263	2.18	2.00
2	0.55029	0.27474	2.00	2.00
3	0.82414	0.51940	1.59	1.59
5	1.48053	1.06582	1.39	1.32
8	2.26564	1.86390	1.22	1.19

予測は $4^{\min(1/2, 1/d)}$ である。（証明終）

観察 6.2 （二つの領域）

率は $d \leq 2$ で $n^{-1/2}$ 、 $d \geq 3$ で $n^{-1/d}$ である。すなわち「次元の呪い」は d^{-1} が $1/2$ を下回る $d = 3$ から始まる——中心極限の率と被覆の率が、そこで入れ替わる。

観察 6.3（これは次元が選ぶ側である）

論文114 は次元を選ぶ事実と選ばない事実を分けた。ここでの $d = 3$ は、選ぶ側の閾値である——ただし選んでいるのは幾何ではなく、二つの率のどちらが勝つかである。論文133 のジェームズ＝スタインの $d \geq 3$ とは、閾値が同じ数でも根が違う（あちらは許容性、こちらは被覆数）。

観察 6.4（厳密性と可測性）

距離の公理を満たすことは、推定しやすいことを意味しない。論文136 で見た形がここでも現れる——正しいことと、測れることは別である。

7. 結論

1. ワッサースタインは距離である（検算2.1）。3000 組で非対称 0 件、三角不等式の破れ 0 件。論文137 の KL は 25.4% 破れていた。
2. 分ける実例は台が交わらないとき（検算3.2）。KL は $d = 0.5$ でも 100 でも無限大、 W_2 は d をそのまま返す。
3. 「どれだけ違うか」は二つの問いである（観察3.3）。重なっているかとどれだけ離れているか。
4. ただし前提が違う（検算4.2）。ワッサースタインは台の上の距離を要求し、KL は要求しない。優劣ではなく、前提の違いである。
5. 熱方程式は二度書ける（検算5.1・観察5.3）。 $W_2 = \sqrt{2t}$ すなわち $[t] = L^2$ （論文119）、そしてオットーの定理により勾配流。
6. 収束率は $d = 3$ で性格を変える（検算6.1）。 $d \leq 2$ で $n^{-1/2}$ 、 $d \geq 3$ で $n^{-1/d}$ 。距離であることは、測りやすさを意味しない（論文136）。

結び：KL とワッサースタインは競合していない。片方は重なりを、もう片方は隔たりを測っている。台が交わらない場所で、その差がいちばんはっきり出る——片方は「無限に違う」としか言えず、もう片方は距離を答える。どちらを使うかは、こちらがどちらの問いを持っているかで決まる。

参考文献

1. Kantorovich, L. V. (1942). *On the translocation of masses*. Doklady Akademii Nauk SSSR, 37.
2. Jordan, R., Kinderlehrer, D. & Otto, F. (1998). *The variational formulation of the Fokker-Planck equation*. SIAM Journal on Mathematical Analysis, 29.

3. Otto, F. (2001). *The geometry of dissipative evolution equations: the porous medium equation*. Communications in Partial Differential Equations, 26.
4. Villani, C. (2009). *Optimal Transport: Old and New*. Springer.
5. Fournier, N. & Guillin, A. (2015). *On the rate of convergence in Wasserstein distance of the empirical measure*. Probability Theory and Related Fields, 162.
6. Cuturi, M. (2013). *Sinkhorn distances: lightspeed computation of optimal transport*. Advances in Neural Information Processing Systems, 26.
7. 山岸勇輝 (2026). 等号は六つの別根を持つ [論文16] 第3版. Zenodo. 10.5281/zenodo.21600471
8. 山岸勇輝 (2026). 別根と言うのは半分である——住所を書け [論文17] 第3版. Zenodo. 10.5281/zenodo.21600813
9. 山岸勇輝 (2026). 二つのエントロピー [論文88]. Zenodo. 10.5281/zenodo.21747809
10. 山岸勇輝 (2026). フラクタル次元 [論文94]. Zenodo. 10.5281/zenodo.21758581
11. 山岸勇輝 (2026). 規約と事実 [論文95]. Zenodo. 10.5281/zenodo.21758621
12. 山岸勇輝 (2026). カルデロンの逆問題 [論文99]. Zenodo. 10.5281/zenodo.21758790
13. 山岸勇輝 (2026). 王道の点呼 [論文100]. Zenodo. 10.5281/zenodo.21758864
14. 山岸勇輝 (2026). ウィーナー・フィルタ [論文106]. Zenodo. 10.5281/zenodo.21836984
15. 山岸勇輝 (2026). リッチ・フローは繰り込み群の流れである [論文110]. Zenodo. 10.5281/zenodo.21837476
16. 山岸勇輝 (2026). 極限としてのモジュール [論文113]. Zenodo. 10.5281/zenodo.21962996
17. 山岸勇輝 (2026). 尺度不変性は四つの根を持つ [論文117]. Zenodo. 10.5281/zenodo.22005515
18. 山岸勇輝 (2026). 第四の時計 [論文119]. Zenodo. 10.5281/zenodo.22036353
19. 山岸勇輝 (2026). フィッシャー計量が一意なのは、古典の側だけである [論文133]. Zenodo. 10.5281/zenodo.22080550
20. 山岸勇輝 (2026). 等式は厳密でも、平均は裾が握っている [論文136]. Zenodo. 10.5281/zenodo.22081230
21. 山岸勇輝 (2026). KL は距離ではない。だが、その二階微分は計量である [論文137]. Zenodo. 10.5281/zenodo.22081369
22. 山岸勇輝 (2026). どこまでがハミルトニアンかは、規約である [論文138]. Zenodo. 10.5281/zenodo.22081512

— *From thinking to syncing.* —

山岸 勇輝 (Yuuki YAMAGISHI)

© 2026 Yamagishi Complexity Torus Lab, Japan · Licensed under Creative Commons Attribution 4.0 International (CC BY 4.0)