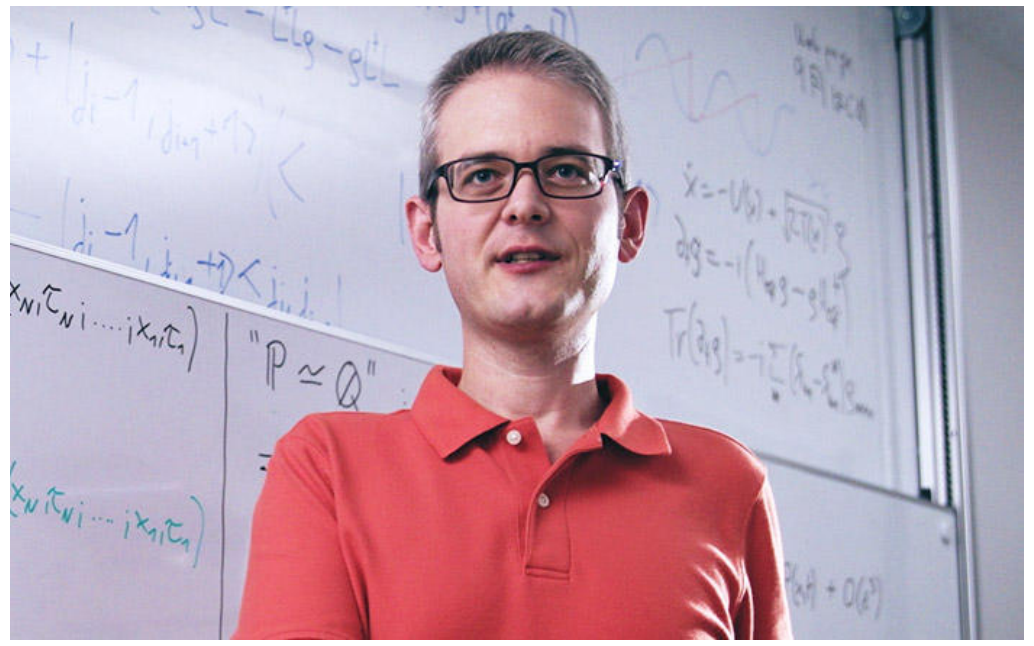
最適輸送と熱力学的最適化

Workshop OT 2023, 3/15, 2023

東京大学　理学系研究科附属生物普遍性研究機構　伊藤 創祐

共同研究者

- Andreas Dechant (京都大学), 佐々真一 (京都大学)




- 仲里夢叶 (旧ラボメンバー)、藤本悠雅 (旧ラボメンバー) 、
  吉村耕平（現ラボメンバー)、Artemy Kolchinsky (現ラボメンバー)

二日目の講演:
L^2-Wasserstein距離と離散ダイナミクス
(本講演の離散状態バージョンに相当)

Reference:

Nakazato, M., & Ito, S. (2021). *Physical Review Research*, *3*(4), 043093.

Fujimoto, Y., & Ito, S. (2021). *arXiv preprint arXiv:2112.14035*.

Dechant, A., Sasa, S. I., & Ito, S. (2022). *Physical Review Research*, *4*(1), L012034.

Dechant, A., Sasa, S. I., & Ito, S. (2022). *Physical Review E*, *106*(2), 024125.

Yoshimura, K., Kolchinsky, A., Dechant, A., & Ito, S. (2023). *Physical Review Research*, *5*(1), 013017.

Kolchinsky, A., Dechant, A., Yoshimura, K., & Ito, S. (2022). *arXiv preprint arXiv:2206.14599*.

Ito, S. (2023). *Information Geometry*, 1-42.

トーク内容

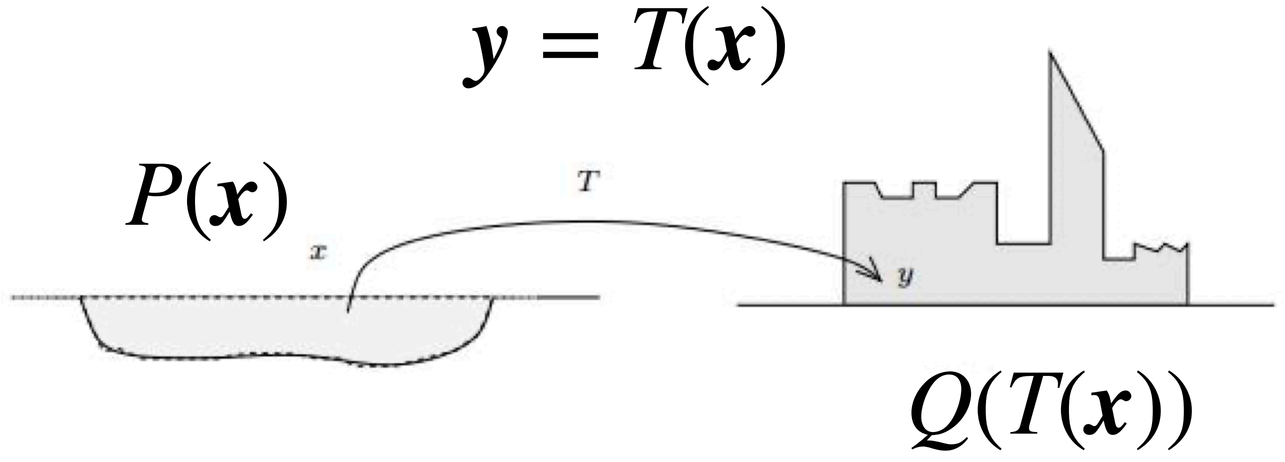
- 導入: 連続状態の確率密度分布の最適輸送
- ブラウン運動における熱力学的最適化と最適輸送
- 最適輸送の幾何と情報幾何

トーク内容

- 導入: 連続状態の確率密度分布の最適輸送
- ブラウン運動における熱力学的最適化と最適輸送
- 最適輸送の幾何と情報幾何

導入: Mongeの最適輸送問題

- Mongeの最適輸送問題 (1781)　$\boldsymbol{x} \in \mathbb{R}^d, \boldsymbol{y} \in \mathbb{R}^d, d \in \mathbb{N}$



確率分布　$P(\boldsymbol{x}), Q(\boldsymbol{y})$

$$P(\boldsymbol{x}) \geq 0, Q(\boldsymbol{y}) \geq 0 \qquad \int d\boldsymbol{x} P(\boldsymbol{x}) = 1, \int d\boldsymbol{y} Q(\boldsymbol{y}) = 1$$

輸送コスト　$c(\boldsymbol{x}, \boldsymbol{y}) \geq 0$

砂山を瓦礫に輸送する総コスト
（コストの期待値)を,
輸送の仕方$T(x)$について最小化する

輸送写像　$\boldsymbol{y} = T(\boldsymbol{x}) \in \mathbb{R}^d$

ヤコビ方程式

$$Q(T(\boldsymbol{x})) \det |\nabla T(\boldsymbol{x})| = P(\boldsymbol{x})$$

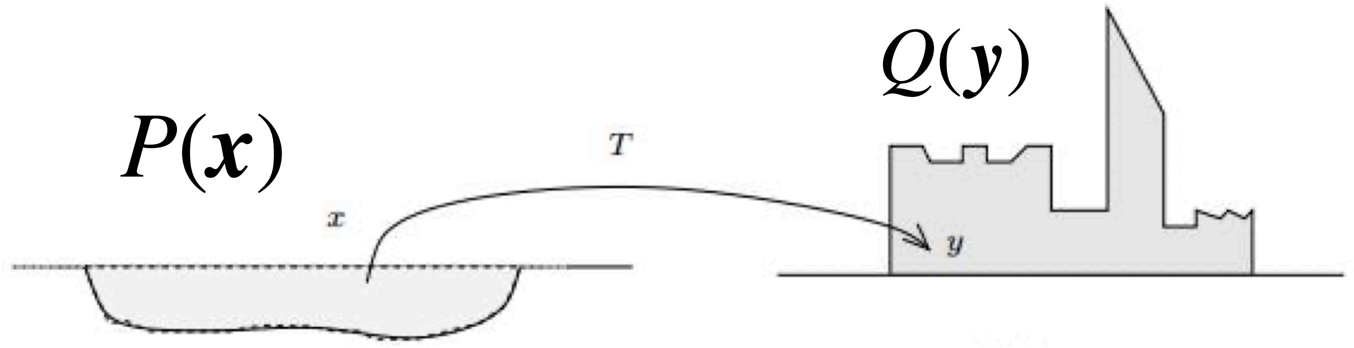
Mongeの問題

$$C_{\mathrm{M}}(P, Q) = \inf_T \int d\boldsymbol{x} c(\boldsymbol{x}, T(\boldsymbol{x})) P(\boldsymbol{x})$$

Villani, C. (2009). *Optimal transport: old and new* (Vol. 338, p. 23). Berlin: springer.

導入: Monge-Kantorovichの最適輸送問題

- Kantorovichの方法 (1939)



同時確率分布の集合

$$\Pi(P,Q) = \left\{ \pi(\boldsymbol{x},\boldsymbol{y}) \,\middle|\, \pi(\boldsymbol{x},\boldsymbol{y}) \geq 0, \int d\boldsymbol{y}\pi(\boldsymbol{x},\boldsymbol{y}) = P(\boldsymbol{x}), \int d\boldsymbol{y}\pi(\boldsymbol{x},\boldsymbol{y}) = Q(\boldsymbol{x}) \right\}$$

Monge-Kantorovichの問題

$$C(P,Q) = \inf_{\pi \in \Pi(P,Q)} \int d\boldsymbol{x} d\boldsymbol{y} c(\boldsymbol{x},\boldsymbol{y})\pi(\boldsymbol{x},\boldsymbol{y})$$

同時確率分布を導入することで,
総コスト(コストの期待値)
を最小化する.

Remark) Mongeの問題との関係

$$C_{\mathrm{M}}(P,Q) \geq C(P,Q) \qquad \because \qquad \pi_T \in \Pi(P,Q) \qquad \pi_T(\boldsymbol{x},\boldsymbol{y}) = \delta(\boldsymbol{y} - T(\boldsymbol{x}))P(\boldsymbol{x})$$

$$Q(T(\boldsymbol{x}))\det|\nabla T(\boldsymbol{x})| = P(\boldsymbol{x})$$

Villani, C. (2009). *Optimal transport: old and new* (Vol. 338, p. 23). Berlin: springer.

導入: Monge-Kantorovichの問題の具体例 - L2-Wasserstein距離

- L2-Wasserstein距離 (1969)

$$\sqrt{C(P,Q)} = \mathscr{W}_2(P,Q) \quad \text{s.t.} \quad c(\boldsymbol{x},\boldsymbol{y}) = ||\boldsymbol{x}-\boldsymbol{y}||^2$$

L2-Wasserstein距離

$$\mathscr{W}_2(P,Q) = \sqrt{\inf_{\pi\in\Pi(P,Q)} \int d\boldsymbol{x}d\boldsymbol{y}\,||\boldsymbol{x}-\boldsymbol{y}||^2\pi(\boldsymbol{x},\boldsymbol{y})}$$

二次モーメントが存在する($\int d\boldsymbol{x}\,||\boldsymbol{x}||^2 P(\boldsymbol{x}) < \infty$, $\int d\boldsymbol{y}\,||\boldsymbol{y}||^2 Q(\boldsymbol{y}) < \infty$)とき

最適解$\pi^*(\boldsymbol{x},\boldsymbol{y}) = \arg\min_{\pi\in\Pi(P,Q)} \int d\boldsymbol{x}d\boldsymbol{y}\,||\boldsymbol{x}-\boldsymbol{y}||^2\pi(\boldsymbol{x},\boldsymbol{y})$は存在し,

ヤコビ方程式を満たすある輸送写像$T^*(\boldsymbol{x})$によって, $\pi^*(\boldsymbol{x},\boldsymbol{y}) = \delta(\boldsymbol{y}-T^*(\boldsymbol{x}))P(\boldsymbol{x})$で表される. (Brenier, 1987)

$$\mathscr{W}_2(P,Q) = \sqrt{\int d\boldsymbol{x}\,||\boldsymbol{x}-T^*(\boldsymbol{x})||^2 p(\boldsymbol{x})} = \sqrt{C_M(P,Q)}$$

Monge-Kantorovichの問題がMongeの問題と一致する具体例になっている ($C(P,Q) = C_{\mathrm{M}}(P,Q)$)

Villani, C. (2009). *Optimal transport: old and new* (Vol. 338, p. 23). Berlin: springer.

導入: 最適輸送の流体ダイナミクス - Benamou-Brenier公式

- Benanou-Brenier公式 (2000)

連続の式

有限時間τ間で分布を輸送する流体ダイナミクスを考える

$$\partial_t P_t(\boldsymbol{x}) = -\nabla\cdot(\boldsymbol{\nu}_t(\boldsymbol{x})P_t(\boldsymbol{x})), \qquad P_0(\boldsymbol{x}) = P(\boldsymbol{x}), P_\tau(\boldsymbol{x}) = Q(\boldsymbol{x})$$

$$\mathscr{W}_2(P,Q) = \sqrt{\inf_{(\boldsymbol{\nu}_t(\boldsymbol{x}),P_t(\boldsymbol{x}))_{0\le t\le\tau}} \tau\int_0^\tau dt\int d\boldsymbol{x}\,||\boldsymbol{\nu}_t(\boldsymbol{x})||^2 P_t(\boldsymbol{x})} = \sqrt{\tau\int_0^\tau dt\int d\boldsymbol{x}\,||\boldsymbol{\nu}^*_t(\boldsymbol{x})||^2 P_t(\boldsymbol{x})}$$

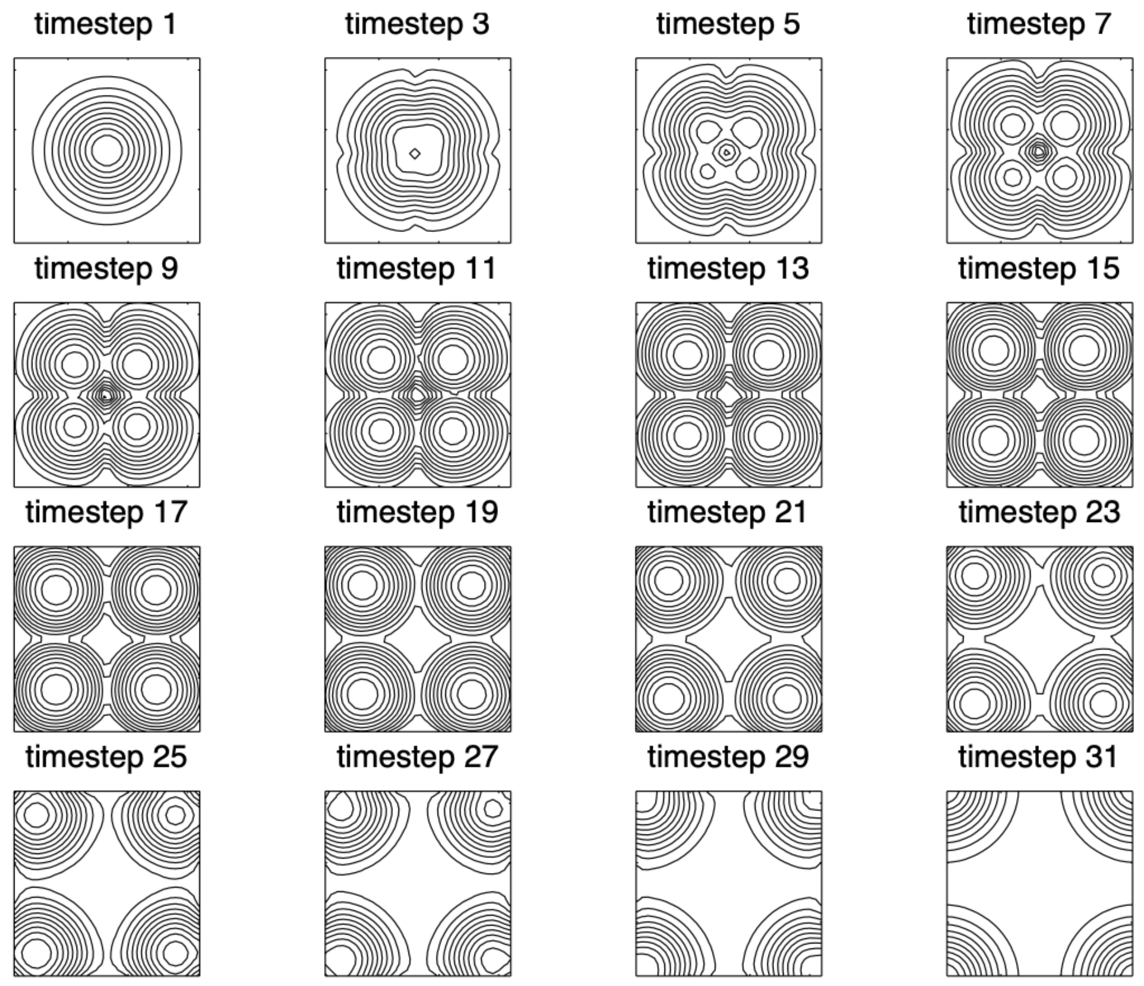
最適な速度場 $\boldsymbol{\nu}^*_t(\boldsymbol{x})$

$$\boldsymbol{\nu}^*_t(\boldsymbol{x}) = \nabla\phi_t(\boldsymbol{x})$$

$$\partial_t\phi_t(\boldsymbol{x}) + \frac{1}{2}||\nabla\phi_t(\boldsymbol{x})||^2 = 0$$

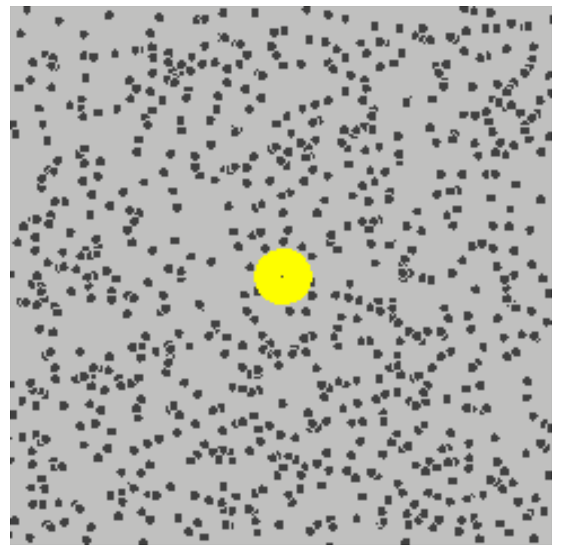
速度場が勾配で与えられ,
渦なしのEuler方程式で記述されるとき
最適輸送.

流体ダイナミクスによる最適輸送



Benamou, Jean-David, and Yann Brenier. (2000). *Numerische Mathematik* 84, 375-393.

導入: ブラウン運動の熱力学と最適輸送
- Jordan-Kinderlehrer-Otto scheme



- Jordan-Kinderlehrer-Otto scheme (1998)

Fokker-Planck方程式 = ポテンシャル$U(x)$で駆動されるブラウン運動の確率ダイナミクス

$$\partial_t P_t(\boldsymbol{x}) = -\nabla \cdot (\boldsymbol{\nu}_t(\boldsymbol{x}) P_t(\boldsymbol{x})), \quad \boldsymbol{\nu}_t(\boldsymbol{x}) = -\nabla U(\boldsymbol{x}) - \beta^{-1} \nabla \ln P_t(\boldsymbol{x})$$

β^{-1}:熱浴(溶媒)の温度

Jordan-Kinderlehrer-Otto scheme
(Fokker-Planck方程式の解$P_t(x)$は次の最小化問題から与えられる)

$$P_{t+\Delta t} = \arg\min_{Q} \left[\frac{1}{2} \mathcal{W}_2(P_t, Q)^2 + \Delta t F(Q) \right]$$

自由エネルギー: $F(Q) = \int d\boldsymbol{x} [U(\boldsymbol{x}) + \beta^{-1} \ln Q(\boldsymbol{x})] Q(\boldsymbol{x})$

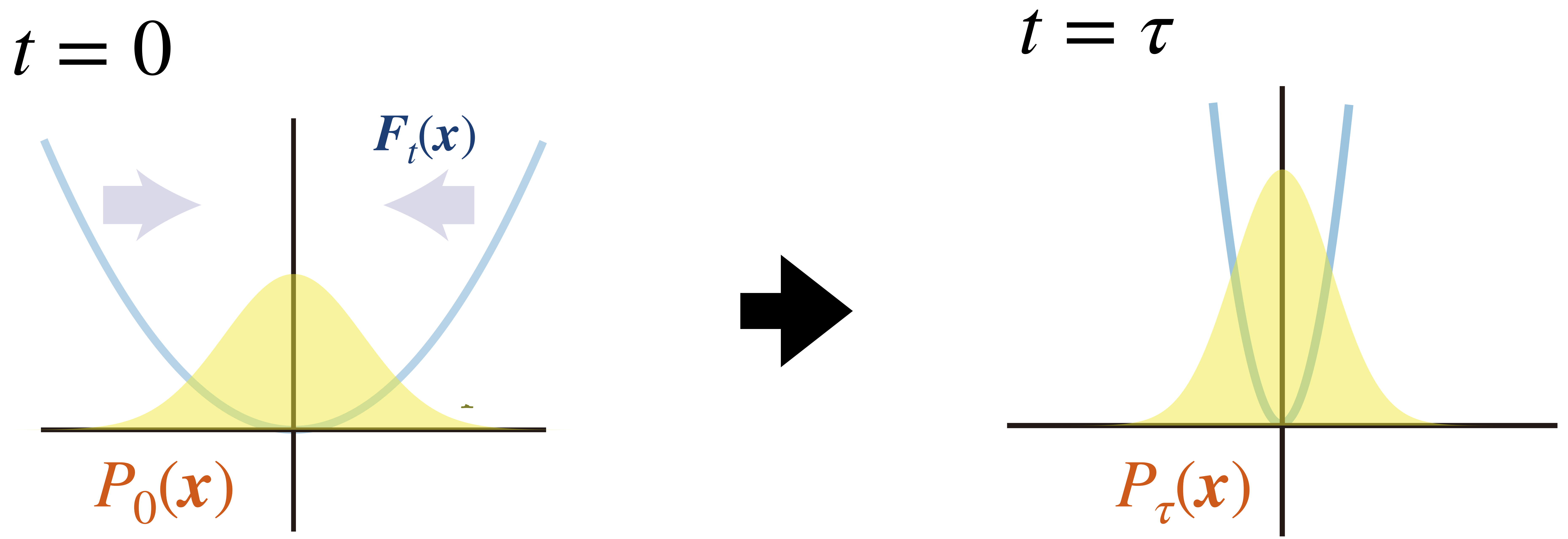
ブラウン運動のダイナミクスは, 熱力学的な散逸と最適な輸送コストを最小化するように時間発展する.

Jordan, R., Kinderlehrer, D., & Otto, F. (1998). *SIAM journal on mathematical analysis*, *29*(1), 1-17.

トーク内容

- 導入: 連続状態の確率密度分布の最適輸送
- ブラウン運動における熱力学的最適化と最適輸送
- 最適輸送の幾何と情報幾何

アイディア: ブラウン運動における最適輸送



Fokker-Planck方程式

$$\partial_t P_t(\boldsymbol{x}) = -\nabla \cdot (\boldsymbol{\nu}_t(\boldsymbol{x}) P_t(\boldsymbol{x}))$$

$$\boldsymbol{\nu}_t(\boldsymbol{x}) = \mu(\boldsymbol{F}_t(\boldsymbol{x}) - \beta^{-1} \nabla \ln P_t(\boldsymbol{x}))$$

μ: 易動度 (粘性の逆数)

$\boldsymbol{F}_t(\boldsymbol{x})$: 時間依存する外力 (非ポテンシャル力を含む)

外力$\boldsymbol{F}_t(\boldsymbol{x})$をコントロールすることで, 有限時間$\tau$で

最も熱力学的な散逸が少ない(最小エントロピー生成の)最適な輸送を考える.

エントロピー生成と最小エントロピー生成

エントロピー生成

$$\Sigma_\tau = \int_0^\tau dt \sigma_t$$

熱力学的散逸の指標

(非負性$\Sigma_\tau \geq 0$は熱力学第二法則に相当.)

エントロピー生成率

$$\sigma_t = \frac{\beta}{\mu} \int d\boldsymbol{x} \|\boldsymbol{\nu}_t(\boldsymbol{x})\|^2 P_t(\boldsymbol{x})$$

Fokker-Planck方程式

$$\partial_t P_t(\boldsymbol{x}) = -\nabla \cdot (\boldsymbol{\nu}_t(\boldsymbol{x}) P_t(\boldsymbol{x}))$$

$P_0(x)$と$P_\tau(x)$を固定した元での最小エントロピー生成 (熱力学的速度限界)

$$\Sigma_\tau \geq \frac{\beta(\mathscr{W}_2(P_0, P_\tau))^2}{\mu\tau}$$

($\because$ Benamou-Brenierの公式)

(早く状態変化させようとするほど, 散逸が必要になる)

最小エントロピー生成を達成するプロトコル

$$\boldsymbol{F}_t(\boldsymbol{x}) = -\nabla U_t(\boldsymbol{x})$$

$$\boldsymbol{\nu}_t(\boldsymbol{x}) = \nabla \phi_t(\boldsymbol{x})$$

$$\phi_t(\boldsymbol{x}) = \mu(-U_t(\boldsymbol{x}) - \beta^{-1} \ln P_t(\boldsymbol{x}))$$

$$\partial_t \phi_t(\boldsymbol{x}) + \frac{1}{2} || \nabla \phi_t(\boldsymbol{x}) ||^2 = 0$$

Aurell, E., Mejía-Monasterio, C., & Muratore-Ginanneschi, P. (2011). *Physical review letters*, *106*(25), 250601.
Aurell, E., Gawędzki, K., Mejía-Monasterio, C., Mohayaee, R., & Muratore-Ginanneschi, P. (2012). *Journal of statistical physics*, *147*, 487-505.

エントロピー生成率の下限と過剰エントロピー生成率

無限小時間のBenamou-Brenierの公式

$$\sigma_t \geq \sigma_t^{\mathrm{ex}}$$

過剰エントロピー生成

$$\Sigma_\tau^{\mathrm{ex}} = \int_0^\tau dt \sigma_t^{\mathrm{ex}} (\leq \Sigma_\tau)$$

過剰エントロピー生成率

$$\sigma_t^{\mathrm{ex}} = \frac{\beta}{\mu} \lim_{\Delta t \to 0} \frac{\mathscr{W}_2(P_t, P_{t+\Delta t})^2}{\Delta t^2} := \frac{\beta}{\mu} \left(\frac{d\mathscr{L}_t}{dt} \right)^2$$

$\mathscr{L}_t$: 時刻0からのL2-Wasserstein距離による経路長

$$\mathscr{L}_t = \lim_{\Delta t \to 0} \sum_{n=1}^{\mathrm{floor}(t/\Delta t)} \mathscr{W}_2(P_{n\Delta t}, P_{(n-1)\Delta t})$$

時間発展$(P_t(\boldsymbol{x}))_{0 \leq t \leq \tau}$を固定した下での最小過剰エントロピー生成 (熱力学的速度限界)

$$\Sigma^{\mathrm{ex}} \geq \frac{\beta \mathscr{L}_\tau^2}{\mu\tau} \left(\geq \frac{\beta(\mathscr{W}_2(P_0, P_\tau))^2}{\mu\tau} \right)$$

$$\frac{d\mathscr{L}_t}{dt} = \frac{\mathscr{W}_2(P_0, P_\tau)}{\tau} = \mathrm{const.}$$

で時間発展したとき, 全ての等号が成立.

(測地線に沿って分布を時間発展させたとき, 過剰な熱力学的散逸が最小)

Nakazato, M., & Ito, S. (2021). *Physical Review Research*, *3*(4), 043093.

Dechant, A., Sasa, S. I., & Ito, S. (2022). *Physical Review Research*, *4*(1), L012034.

Ito, S. (2023). *Information Geometry*, 1-42.

最適輸送理論に基づいた定常状態熱力学: エントロピー生成率の分解

２つの非負な成分への分解

(定常状態熱力学)

$$\sigma_t = \sigma_t^{\mathrm{ex}} + \sigma_t^{\mathrm{hk}}$$

過剰エントロピー生成率の定義

$$\sigma_t^{\mathrm{ex}} = \frac{\beta}{\mu} \lim_{\Delta t \to 0} \frac{\mathscr{W}_2(P_t, P_{t+\Delta t})^2}{\Delta t^2} (\geq 0)$$

維持エントロピー生成率の定義

$$\sigma_t^{\mathrm{hk}} = \sigma_t - \sigma_t^{\mathrm{ex}} (\geq 0)$$

Fokker-Planck方程式と同じ時間発展を与える速度場 $\nabla \phi_t(\boldsymbol{x})$

$$\partial_t P_t(\boldsymbol{x}) = -\nabla \cdot (\boldsymbol{\nu}_t(\boldsymbol{x}) P_t(\boldsymbol{x})) = -\nabla \cdot (\nabla \phi_t(\boldsymbol{x}) P_t(\boldsymbol{x}))$$

内積表現:

$$\langle \boldsymbol{a}, \boldsymbol{b} \rangle_t := \frac{\beta}{\mu} \int d\boldsymbol{x} \boldsymbol{a}(\boldsymbol{x}) \cdot \boldsymbol{b}(\boldsymbol{x}) P_t(\boldsymbol{x})$$

$$\sigma_t^{\mathrm{ex}} = \langle \nabla \phi_t, \nabla \phi_t \rangle_t$$

時間発展に起因する熱力学的散逸

$$\sigma_t^{\mathrm{hk}} = \langle \boldsymbol{\nu}_t - \nabla \phi_t, \boldsymbol{\nu}_t - \nabla \phi_t \rangle_t$$

時間発展に影響を与えない”渦”の自由度($-\nabla \cdot ([\boldsymbol{\nu}_t(\boldsymbol{x}) - \nabla \phi_t(\boldsymbol{x})] P_t(\boldsymbol{x})) = 0$**)**
からくる熱力学的散逸

直交性 $\langle \nabla \phi_t, \boldsymbol{\nu}_t - \nabla \phi_t \rangle_t = 0$ **を用いた, ピタゴラスの定理による分解に相当**

$$\sigma_t = \langle \boldsymbol{\nu}_t, \boldsymbol{\nu}_t \rangle_t = \langle \nabla \phi_t, \nabla \phi_t \rangle_t + \langle \boldsymbol{\nu}_t - \nabla \phi_t, \boldsymbol{\nu}_t - \nabla \phi_t \rangle_t = \sigma_t^{\mathrm{ex}} + \sigma_t^{\mathrm{hk}}$$

Nakazato, M., & Ito, S. (2021). *Physical Review Research*, *3*(4), 043093.

Dechant, A., Sasa, S. I., & Ito, S. (2022). *Physical Review Research*, *4*(1), L012034.

cf.) Maes, C., & Netočný, K. (2014). *Journal of Statistical Physics*, *154*, 188-203.

熱力学的不確定性関係

$R(\boldsymbol{x}) \in \mathbb{R}$:任意の観測量

観測量の期待値

$$\mathbb{E}[R]_t = \int d\boldsymbol{x} R(\boldsymbol{x}) P_t(\boldsymbol{x})$$

Cauchy-Schwartzの不等式

$$(\langle \nabla \phi_t, \nabla R \rangle_t)^2 \leq \langle \nabla \phi_t, \nabla \phi_t \rangle_t \langle \nabla R, \nabla R \rangle_t$$

$$\partial_t \mathbb{E}[R]_t = \frac{\mu}{\beta} \langle \nabla \phi_t, \nabla R \rangle_t$$

熱力学的不確定性関係 (cf. Cramér–Raoの不等式)

$$\sigma_t^{\mathrm{ex}} \geq \frac{\beta |\partial_t \mathbb{E}[R]_t|^2}{\mu \int d\boldsymbol{x} \|\nabla R(\boldsymbol{x})\|^2 P_t(\boldsymbol{x})}$$

時間発展に起因する熱力学的散逸σ_t^{ex}と,

任意の観測量期待値の変化の速さ$|\partial_t \mathbb{E}[R]_t|$と

観測量の空間ゆらぎ$\int d\boldsymbol{x} \|\nabla R(\boldsymbol{x})\|^2 P_t(\boldsymbol{x})$の間のトレードオフ関係

(早く期待値を変化させようとするほど散逸が必要になる)

Dechant, A., & Sasa, S. I. (2018). *Journal of Statistical Mechanics: Theory and Experiment*, *2018*(6), 063209.

Ito, S., & Dechant, A. (2020). *Physical Review X*, *10*(2), 021056.

Dechant, A., Sasa, S. I., & Ito, S. (2022). *Physical Review Research*, *4*(1), L012034.

Dechant, A., Sasa, S. I., & Ito, S. (2022). *Physical Review E*, *106*(2), 024125.

具体例: 1次元ブラウン粒子系の最適な有限時間熱機関

1次元Fokker-Planck方程式

$$\partial_t P_t(x) = -\partial_x(\nu_t(x)P_t(x))$$

$$\nu_t(x) = \mu(-\nabla U_t(x) - \beta_t^{-1}\nabla \ln P_t(x))$$

$$U_t(x) = \frac{1}{2}k_t(x-a_t)^2$$

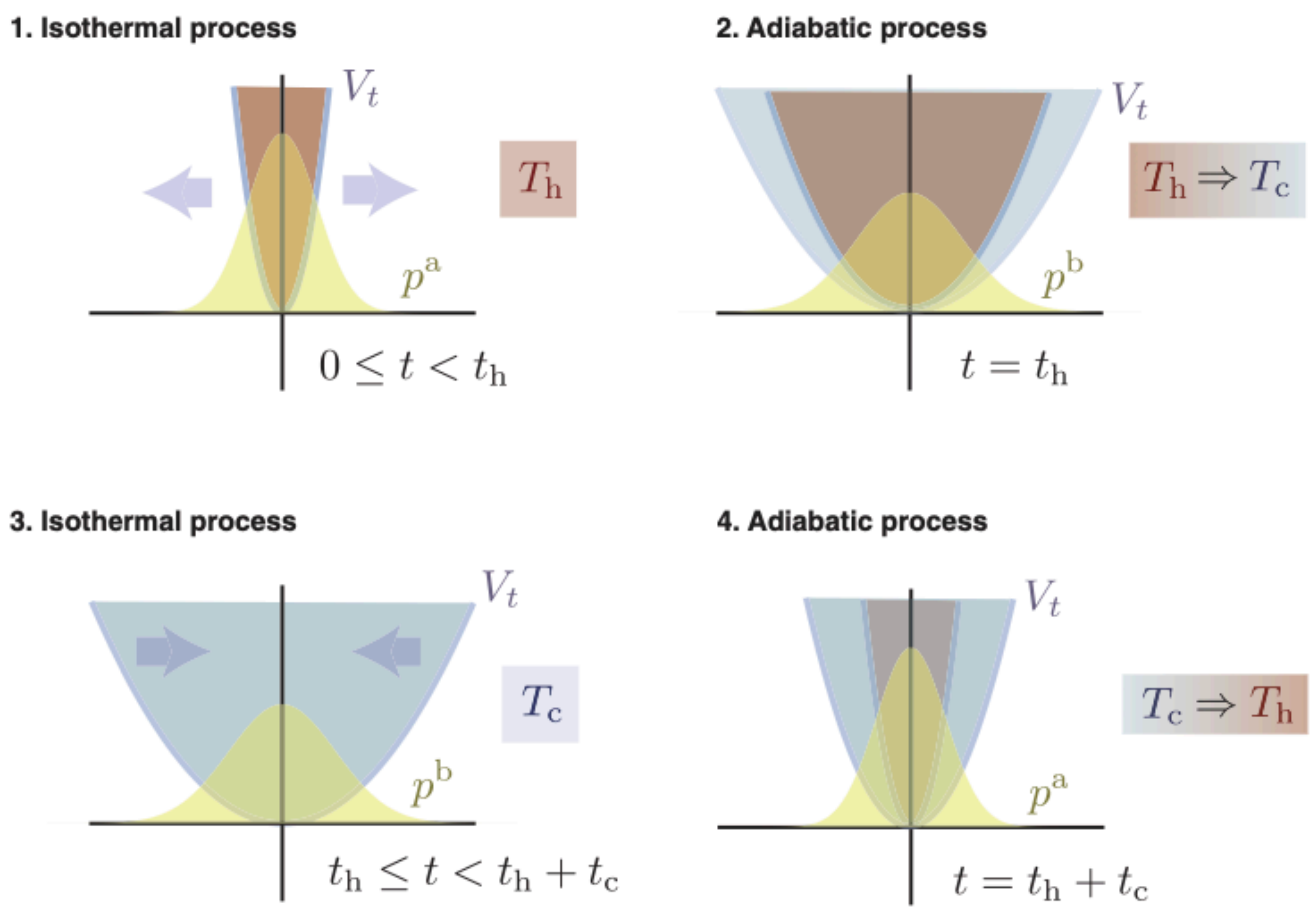
$$T_{\mathrm{h}} > T_{\mathrm{c}}$$

$$\beta_t^{-1} = T_{\mathrm{h}}\ (0 \le t \le t_{\mathrm{h}})$$

$$\beta_t^{-1} = T_{\mathrm{c}}\ (t_{\mathrm{h}} \le t < t_{\mathrm{h}} + t_{\mathrm{c}})$$

$$P_0 = P_{t_{\mathrm{h}}+t_{\mathrm{c}}} = p^{\mathrm{a}}$$

$$P_{t_{\mathrm{h}}} = p^{\mathrm{b}}$$



$$P_t(x) = \frac{1}{\sqrt{2\pi \mathrm{Var}_t}} \exp\left[-\frac{(x-\mathrm{E}_t)^2}{2\mathrm{Var}_t}\right]$$:ガウス分布を仮定

ガウスの場合: $$\left(\frac{d\mathscr{L}_t}{dt}\right)^2 = \left(\frac{d\mathrm{E}_t}{dt}\right)^2 + \left(\frac{d\sqrt{\mathrm{Var}_t}}{dt}\right)^2$$

最適なプロトコル(測地線)： $\dfrac{d\mathrm{E}_t}{dt} = \mathrm{const.},\ \dfrac{d\sqrt{\mathrm{Var}_t}}{dt} = \mathrm{const.}$

有限時間での最大熱効率

$$\eta \le \frac{T_{\mathrm{h}} - T_{\mathrm{c}} - \dfrac{\mathscr{W}_2(p^{\mathrm{a}}, p^{\mathrm{b}})^2}{\mu\Delta S(t_{\mathrm{h}}^{-1} + t_{\mathrm{c}}^{-1})^{-1}}}{T_{\mathrm{h}} - \dfrac{\mathscr{W}_2(p^{\mathrm{a}}, p^{\mathrm{b}})^2}{\mu\Delta S t_{\mathrm{h}}}} \left(\le \frac{T_{\mathrm{h}} - T_{\mathrm{c}}}{T_{\mathrm{h}}}\right)$$

$$(t_{\mathrm{c}} \to \infty, t_{\mathrm{h}} \to \infty)$$

$$\Delta S = \int dx p^{\mathrm{a}}(x) \ln p^{\mathrm{a}}(x) - \int dx p^{\mathrm{b}}(x) \ln p^{\mathrm{b}}(x)$$

Nakazato, M., & Ito, S. (2021). *Physical Review Research*, *3*(4), 043093.

最適輸送理論からの熱力学的最適化への直感的な教え

<u>熱力学的な散逸を減らし, 熱力学的な最適化を達成するには</u>

① 状態を大きく変化させない ($\mathscr{W}_2(P_0, P_\tau)$, $\mathscr{L}_\tau$を小さくする)

② ゆっくり動かす (τを大きくする, $|\partial_t \mathbb{E}[R]_t|$を小さくする)

③ ポテンシャル力で動かし, 非ポテンシャル力で発生する
渦のモードによる散逸をなくす
($\boldsymbol{F}_t(\boldsymbol{x}) = -\nabla U_t(\boldsymbol{x})$で動かす. $\sigma_t^{\mathrm{hk}} = 0$とする)

④ L2-Wasserstein距離の空間で, 測地線に沿って確率分布を動かす
($\frac{d\mathscr{L}_t}{dt} = \frac{\mathscr{W}_2(P_0, P_\tau)}{\tau} = \mathrm{const.}$ で確率分布を時間発展させる)

エントロピー生成率の分解

$$\sigma_t = \sigma_t^{\mathrm{ex}} + \sigma_t^{\mathrm{hk}} \quad \Sigma_\tau^{\mathrm{ex}} = \int_0^\tau dt \sigma_t^{\mathrm{ex}} (\leq \Sigma_\tau)$$

速度限界と等号達成条件

$$\Sigma^{\mathrm{ex}} \geq \frac{\beta \mathscr{L}_\tau^2}{\mu\tau} \left(\geq \frac{\beta(\mathscr{W}_2(P_0, P_\tau))^2}{\mu\tau} \right)$$

($\frac{d\mathscr{L}_t}{dt} = \frac{\mathscr{W}_2(P_0, P_\tau)}{\tau} = \mathrm{const.}$
のとき等号が成立.)

熱力学的不確定性関係

$$\sigma_t^{\mathrm{ex}} \geq \frac{\beta |\partial_t \mathbb{E}[R]_t|^2}{\mu \int d\boldsymbol{x} \|\nabla R(\boldsymbol{x})\|^2 P_t(\boldsymbol{x})}$$

トーク内容

- 導入: 連続状態の確率密度分布の最適輸送
- ブラウン運動における熱力学的最適化と最適輸送
- 最適輸送の幾何と情報幾何

二つの情報幾何と最適輸送(L2-Wasserstein距離)の幾何

確率分布P_tの情報幾何

時間に関するFisher情報量

$$\frac{ds^2}{dt^2} = \int d\boldsymbol{x} P_t(\boldsymbol{x})(\partial_t \ln P_t(\boldsymbol{x}))^2$$

P_t, P_{t+dt}に関する同時確率分布$\mathbb{P}$の情報幾何

Fisher計量

$$g_{\theta\theta}(\mathbb{P}) = \int d\boldsymbol{x}_t d\boldsymbol{x}_{t+dt} \mathbb{P}(\boldsymbol{x}_t, \boldsymbol{x}_{t+dt})(\partial_\theta \ln \mathbb{P}(\boldsymbol{x}_t, \boldsymbol{x}_{t+dt}))^2$$

最適輸送の幾何 (L2-Wasserstein距離の幾何)

$$\sigma_t^{\mathrm{ex}} = \frac{\beta}{\mu} \lim_{\Delta t \to 0} \frac{\mathscr{W}_2(P_t, P_{t+\Delta t})^2}{\Delta t^2} (\geq 0)$$

Ito, S. (2023). *Information Geometry*, 1-42.

二つの情報幾何と最適輸送(L2-Wasserstein距離)の幾何

確率分布P_tの情報幾何

時間に関するFisher情報量

$$\frac{ds^2}{dt^2} = \int d\boldsymbol{x} P_t(\boldsymbol{x})(\partial_t \ln P_t(\boldsymbol{x}))^2$$

P_t, P_{t+dt}に関する同時確率分布$\mathbb{P}$の情報幾何

Fisher計量

$$g_{\theta\theta}(\mathbb{P}) = \int d\boldsymbol{x}_t d\boldsymbol{x}_{t+dt} \mathbb{P}(\boldsymbol{x}_t, \boldsymbol{x}_{t+dt})(\partial_\theta \ln \mathbb{P}(\boldsymbol{x}_t, \boldsymbol{x}_{t+dt}))^2$$

最適輸送の幾何 (L2-Wasserstein距離の幾何)

$$\sigma_t^{\mathrm{ex}} = \frac{\beta}{\mu} \lim_{\Delta t \to 0} \frac{\mathscr{W}_2(P_t, P_{t+\Delta t})^2}{\Delta t^2} (\geq 0)$$

Ito, S. (2023). *Information Geometry*, 1-42.

確率分布P_tの情報幾何と最適輸送理論: 勾配流によるFokker-Planck方程式の表現

Fokker-Planck方程式

$$\partial_t P_t(\boldsymbol{x}) = -\nabla\cdot(\boldsymbol{\nu}_t(\boldsymbol{x})P_t(\boldsymbol{x})) = -\nabla\cdot(\nabla\phi_t(\boldsymbol{x})P_t(\boldsymbol{x}))$$

$$\boldsymbol{\nu}_t(\boldsymbol{x}) = \mu(\boldsymbol{F}_t(\boldsymbol{x}) - \beta^{-1}\nabla\ln P_t(\boldsymbol{x})) \qquad \nabla\phi_t(\boldsymbol{x}) = \mu(-\nabla U_t^*(\boldsymbol{x}) - \beta^{-1}\nabla\ln P_t(\boldsymbol{x}))$$

$U_t^*(\boldsymbol{x})$: 最適輸送を与える最適なポテンシャル (pseudo potential)

$P_t^{\mathrm{pcan}}(\boldsymbol{x}) = \dfrac{\exp[-\beta U_t^*(\boldsymbol{x})]}{\int d\boldsymbol{x}\exp[-\beta U_t^*(\boldsymbol{x})]}$: $U_t^*(\boldsymbol{x})$で与えられる仮想的なカノニカル分布 (pseudo canonical dist.)

Fokker-Planck方程式 (勾配流による表現)：

$$\partial_t P_t(\boldsymbol{x}) = -\nabla\cdot\left(\mu\beta^{-1}P_t(\boldsymbol{x})\nabla\ln\frac{P_t(\boldsymbol{x})}{P_t^{\mathrm{pcan}}(\boldsymbol{x})}\right)$$

$\ln\dfrac{P_t(\boldsymbol{x})}{P_t^{\mathrm{pcan}}(\boldsymbol{x})}$の勾配に沿って, 緩和していくダイナミクス

Yoshimura, K., Kolchinsky, A., Dechant, A., & Ito, S. (2023). *Physical Review Research*, *5*(1), 013017.

Ito, S. (2023). *Information Geometry*, 1-42.

確率分布P_tの情報幾何と最適輸送理論: 過剰エントロピー生成率とKullback-Leiblerダイバージェンス

$$\sigma_t^{\rm ex} = \langle \nabla \phi_t, \nabla \phi_t \rangle_t = \langle \nabla \phi_t, \mu\beta^{-1} \nabla \ln \frac{P_t}{P_t^{\rm pcan}} \rangle_t = \int d\boldsymbol{x} \nabla \cdot (\nabla \phi_t P_t(\boldsymbol{x})) \ln \frac{P_t}{P_t^{\rm pcan}}$$

$$\sigma_t^{\rm ex} = -\int d\boldsymbol{x} [\partial_t P_t(\boldsymbol{x})] \ln \frac{P_t(\boldsymbol{x})}{P_t^{\rm can}(\boldsymbol{x})} = -\partial_t D_{\rm KL}(P_t || P_s^{\rm pcan}) \Big|_{s=t}$$

KLダイバージェンス: $$D_{\rm KL}(P_t || P_s^{\rm pcan}) = \int d\boldsymbol{x} P_t(\boldsymbol{x}) \ln \frac{P_t(\boldsymbol{x})}{P_s^{\rm pcan}(\boldsymbol{x})}$$

勾配流に沿ったKLダイバージェンスの減少速度が,

L2-Wasserstein距離の空間での速度の二乗$\sigma_t^{\rm ex} = \frac{\beta}{\mu} \lim_{\Delta t \to 0} \frac{\mathscr{W}_2(P_t, P_{t+\Delta t})^2}{\Delta t^2}$を与える.

Dechant, A., Sasa, S. I., & Ito, S. (2022). *Physical Review E*, *106*(2), 024125.

Ito, S. (2023). *Information Geometry*, 1-42.

確率分布P_tの情報幾何と最適輸送理論: 過剰エントロピー生成率と時間に関するFisher情報量

確率の対数での仮想的なカノニカル分布からの離れ具合 $\theta_t(\boldsymbol{x}) = \ln P_t(\boldsymbol{x}) - \ln P_t^{\mathrm{pcan}}(\boldsymbol{x})$

θ_tの分散 $\mathrm{Var}_{P_t}[\theta_t] = \int d\boldsymbol{x} P_t(\boldsymbol{x})(\theta_t(\boldsymbol{x}) - \mathbb{E}_{P_t}[\theta_t])^2$ $\quad \mathbb{E}_{P_t}[\theta_t] = \int d\boldsymbol{x} P_t(\boldsymbol{x})\theta_t(\boldsymbol{x})$

時間に関するFisher情報量の平方根
(確率単体上の情報幾何学的な速度)

$$\frac{ds}{dt} = \sqrt{\int d\boldsymbol{x} P_t(\boldsymbol{x})(\partial_t \ln P_t(\boldsymbol{x}))^2}$$

情報幾何による制限
(cf. Cramér–Raoの不等式)

$$\sigma_t^{\mathrm{ex}} \leq \sqrt{\mathrm{Var}_{P_t}[\theta_t]}\frac{ds}{dt}$$

L2-Wasserstein距離の空間での速度の二乗σ_t^{ex}は
情報幾何学的な速度と, 確率の対数の
離れ具合の分散を用いて上から抑えられる.

Ito, S. (2023). *Information Geometry*, 1-42.

(Remark) 論文のノーテーション$\eta_t(x)$は適切でないです. $\theta_t(x)$は情報幾何的にθ座標相当.

二つの情報幾何と最適輸送(L_2-Wasserstein距離)の幾何

確率分布P_tの情報幾何

時間に関するFisher情報量

$$\frac{ds^2}{dt^2} = \int d\boldsymbol{x} P_t(\boldsymbol{x})(\partial_t \ln P_t(\boldsymbol{x}))^2$$

P_t, P_{t+dt}に関する同時確率分布$\mathbb{P}$の情報幾何

Fisher計量

$$g_{\theta\theta}(\mathbb{P}) = \int d\boldsymbol{x}_t d\boldsymbol{x}_{t+dt} \mathbb{P}(\boldsymbol{x}_t, \boldsymbol{x}_{t+dt})(\partial_\theta \ln \mathbb{P}(\boldsymbol{x}_t, \boldsymbol{x}_{t+dt}))^2$$

最適輸送の幾何 (L_2-Wasserstein距離の幾何)

$$\sigma_t^{\mathrm{ex}} = \frac{\beta}{\mu} \lim_{\Delta t \to 0} \frac{\mathscr{W}_2(P_t, P_{t+\Delta t})^2}{\Delta t^2} (\geq 0)$$

Ito, S. (2023). *Information Geometry*, 1-42.

確率分布P_tの情報幾何と最適輸送理論: Fokker-Planckダイナミクスのinterpolationと経路の確率

Fokker-Planck方程式

$$\partial_t P_t(\boldsymbol{x}) = -\nabla \cdot (\boldsymbol{\nu}_t(\boldsymbol{x}) P_t(\boldsymbol{x})) = -\nabla \cdot (\nabla \phi_t(\boldsymbol{x}) P_t(\boldsymbol{x}))$$

$$\boldsymbol{\nu}_t(\boldsymbol{x}) = \mu(\boldsymbol{F}_t(\boldsymbol{x}) - \beta^{-1} \nabla \ln P_t(\boldsymbol{x}))$$

Interpolated dynamics

($\theta = 0$で普通のFokker-Planckダイナミクス, $\theta = 1$で速度場$\nu_t'(x)$のダイナミクス)

$$\partial_t P_t(\boldsymbol{x}) = -\nabla \cdot (\boldsymbol{\nu}_t^\theta(\boldsymbol{x}) P_t(\boldsymbol{x})) \qquad \boldsymbol{\nu}_t^\theta(\boldsymbol{x}) = (1-\theta)\boldsymbol{\nu}_t(\boldsymbol{x}) + \theta \boldsymbol{\nu}_t'(\boldsymbol{x})$$

Interpolated dynamicsのP_t, P_{t+dt}に関する同時確率分布 (Onsager-Machlup関数による表現)

$$\mathbb{P}_{\boldsymbol{\nu}'}^\theta(\boldsymbol{x}_t, \boldsymbol{x}_{t+dt}) = \frac{1}{(4\pi\mu T dt)^{\frac{d}{2}}} \exp\left[-\frac{\|\boldsymbol{x}_{t+dt} - \boldsymbol{x}_t - \mu \boldsymbol{F}_t(\boldsymbol{x}_t)dt - \theta(\boldsymbol{\nu}_t'(\boldsymbol{x}_t) - \boldsymbol{\nu}_t(\boldsymbol{x}_t))dt\|^2}{4\mu T dt}\right] P_t(\boldsymbol{x}_t)$$

θは指数型分布族のパラメータであり,
Interpolated dynamicsはe-測地線を与える.

$$\ln \mathbb{P}_{\boldsymbol{\nu}'}^\theta = (1-\theta)\ln \mathbb{P}_{\boldsymbol{\nu}'}^0 + \theta \ln \mathbb{P}_{\boldsymbol{\nu}'}^1 + O(dt)$$

(ただし$\mathbb{P}_{\boldsymbol{\nu}'}^0$は普通のFokker-Planck ダイナミクスの同時確率分布)

Dechant, A., Sasa, S. I., & Ito, S. (2022). *Physical Review E*, *106*(2), 024125.
Ito, S. (2023). *Information Geometry*, 1-42.

確率分布P_tの情報幾何と最適輸送理論:
同時確率分布のKullback-Leiblerダイバージェンスと最適輸送

Kullback-Leiblerダイバージェンスによる表記

$$\sigma_t^{\mathrm{ex}} = \lim_{dt\to 0} \frac{4D_{\mathrm{KL}}(\mathbb{P}^0_{\nu_t-\nabla\phi_t} || \mathbb{P}^\theta_{\nu_t-\nabla\phi_t})}{\theta^2 dt}$$

L2-Wasserstein距離の空間での速度の

二乗$\sigma_t^{\mathrm{ex}} = \frac{\beta}{\mu} \lim_{\Delta t\to 0} \frac{\mathcal{W}_2(P_t, P_{t+\Delta t})^2}{\Delta t^2}$

は最適輸送方向$\nabla\phi_t(x)$への変化させた時の

KLダイバージェンスで与えられる.

($\mathbb{P}^\theta_{\nu_t-\nabla\phi_t}$に対する速度場: $\nu_t^\theta(x) = \nu_t(x) + \theta\nabla\phi_t(x)$)

Fisher計量による表記

$$\sigma_t^{\mathrm{ex}} = \frac{2}{dt} \int d\boldsymbol{x}_t d\boldsymbol{x}_{t+dt} \mathbb{P}^\theta_{\boldsymbol{\nu}_t-\nabla\phi_t}(\boldsymbol{x}_t, \boldsymbol{x}_{t+dt}) [\partial_\theta \ln \mathbb{P}^\theta_{\boldsymbol{\nu}_t-\nabla\phi_t}(\boldsymbol{x}_t, \boldsymbol{x}_{t+dt})]^2 \bigg|_{\theta=0} + O(dt)$$

すなわち, L2-Wasserstein距離の空間での速度の二乗σ_t^{ex}は同時確率分布の空間における情報幾何の計量である

Fisher計量を与える. (L2-Wasserstein距離の空間を, 情報幾何で取り扱える!)

Ito, S. (2023). *Information Geometry*, 1-42.

(熱力学的不確定性関係はこの世界でのCramér–Raoの不等式に相当)

確率分布P_tの情報幾何と最適輸送理論:
同時確率分布における射影と最適輸送

情報幾何による射影

$$\sigma_t^{\mathrm{ex}} = \lim_{dt\to 0} \inf_{\mathbb{Q}\in\mathscr{M}_{\mathrm{ZD}}(\mathbb{P})} \frac{4D_{\mathrm{KL}}(\mathbb{P}||\mathbb{Q})}{dt}$$

$$\sigma_t^{\mathrm{hk}} = \lim_{dt\to 0} \inf_{\mathbb{Q}\in\mathscr{M}_{\mathrm{G}}(\mathbb{P})} \frac{4D_{\mathrm{KL}}(\mathbb{P}||\mathbb{Q})}{dt}$$

連続の式の時間発展が0となる速度場による dynamicsの同時確率分布の集合

$$\mathscr{M}_{\mathrm{ZD}}(\mathbb{P}) = \{\mathbb{P}^1_{\boldsymbol{\nu}'} \,|\, \nabla\cdot(\boldsymbol{\nu}'(\boldsymbol{x})P_t(\boldsymbol{x})) = 0\}$$

グラディエントでかける速度場による dynamicsの同時確率分布の集合

$$\mathscr{M}_{\mathrm{G}}(\mathbb{P}) = \{\mathbb{P}^1_{\nabla r} \,|\, r(\boldsymbol{x}) \in \mathbb{R}\}$$

$\mathbb{P} = \mathbb{P}^0_{\boldsymbol{\nu}'}$: 普通のFokker-Planckダイナミクスの経路の確率

L_2-Wasserstein距離の空間での速度の二乗σ_t^{ex}は

KLダイバージェンスの最小化問題として, 情報幾何学的に求められる

Kolchinsky, A., Dechant, A., Yoshimura, K., & Ito, S. (2022). *arXiv preprint arXiv:2206.14599*.

Ito, S. (2023). *Information Geometry*, 1-42.

本トークのまとめ

・連続状態における最適輸送の導入(L2-Wasserstein距離)と物理との繋がり(流体力学、熱力学)についてのレビューを行った.

・特にブラウン運動を記述するFokker-Planck方程式と最適輸送理論が熱力学でつながっていることに基づき, 熱力学的な散逸(最小エントロピー生成)を達成する最適輸送を考えた.

・L2-Wasserstein距離の空間での速度の二乗に比例する散逸(過剰エントロピー生成率)が, さまざまな熱力学的な最適化に関係する不等式(熱力学的速度限界, 熱力学的不確定性関係)を与えることを示した.

・最適輸送の幾何と情報幾何が結びつくことを, 勾配流による表現による確率分布の情報幾何と, interpolated dynamicsと呼ぶ速度場を変更したダイナミクスの経路の同時確率分布の情報幾何の二つの視点から導出した.

詳しい内容はIto, S. (2023). *Information Geometry*, 1-42.をご覧ください