Vol.2021-AVM-115 No.13
2021/11/25

エントロピー正則付き半緩和最適輸送問題における Sinkhorn アルゴリズムに関する検討

福永 拓海[1,a)]　笠井 裕之[1,2,b)]

概要：確率分布間の距離を表現可能な最適輸送問題は機械学習の様々な分野で応用されている．しかし，大規模データに対して効率的に解くのが難しいことが知られている．当該問題の解決のために，制約条件を緩和することで高速に解くことを実現した半緩和最適輸送問題が提案されているが，収束は依然遅い．そこで，半緩和最適輸送問題に対してエントロピー正則化を施し，輸送行列の計算にスケーラビリティを利用した高速化アルゴリズムを提案する．具体的には，従来の最適輸送問題を効率的に解くために使用される Sinkhorn アルゴリズムと類似した交互最適化アルゴリズムを提案する．さらに，提案アルゴリズムが最適解に必ず最適解に収束することを示す．

A study on Sinkhorn algorithm about entropic semi-relaxed optimal transport

1. はじめに

最適輸送問題は始点から終点への最小な輸送コストを求める問題として定義される [1]．確率分布間の距離を表現可能な最適輸送問題は Wasserstein distance として知られており，グラフ最適輸送問題 [2], [3]，マルチビュー学習 [4]，クラスタリング [5]，ドメイン適応 [6] など様々な機械学習の分野で応用されている．最適輸送問題は凸線形計画問題として定式化されているため，ネットワーク・シンプレックス法 [7] などの数多くのソルバーが開発されてきた。しかし，線形計画法の計算量はデータの次元の三乗に比例して増加するため効率的に解くのが難しいことや厳密な質量保存を意味する最適輸送問題の制約条件が性能の低下を引き起こすことが知られている．

この問題の対処のために，最適輸送問題の制約条件を緩和した非平衡最適輸送問題や半緩和最適輸送問題が提案されており [8], [9], [10]，緩和罰則項として l_2 ノルムの二乗を採用した半緩和最適輸送問題に block-coordinate Frank-Wolfe を利用し，高速化を施したことも報告されている [11]．さらに，この緩和定式化は厳密な質量保存を表す最適輸送問題の制約条件を必ずしも満たす必要のない分野で大きな成功を収めている．最適輸送問題を用いた色相転写では，二つの画像の色の割合を保存したことにより生成した色転写画像に本来存在しない人工物が生成される問題が存在する．しかし，片側だけの制約条件を緩和した半緩和最適輸送問題により，色相転写における人工物の除去に成功している [12], [13]．さらに，グラフ問題を扱う際に使用される Gromov Wasserstein 距離の緩和した手法により，サブグラフ構造を捉えることを可能にしたグラフマッチングが実現されている [14]．PU 学習の分野 [15] では，最適輸送問題の制約条件を不等式制約にしたことにより，ラベル無しデータの予測精度が向上している．

緩和最適輸送問題は従来最適輸送問題よりも効率的に解くことを可能にし，多くの分野で有効性を示している．しかし，その収束は依然遅い．そこで，エントロピー正則化が従来の最適輸送問題で提案されている [16]．前述問題を解く Sinkhorn アルゴリズムは，並列計算可能であるため効率的である．さらに，付加したエントロピー項の強凸性と微分可能性により比較的容易な制約無し凸最適化問題を解くことを可能にする．このアルゴリズムによって生じた最適輸送問題もまた，機械学習の分野でその効果を発揮している．さらに，スケーラビリティを利用出来るため，従来最適輸送問題より，大規模データに対して効率的に解く

1　早稲田大学大学院基幹理工学研究科情報理工情報通信専攻
2　早稲田大学基幹理工学部情報通信学科
a)　f_takumi1997@suou.waseda.jp
b)　hiroyuki.kasai@waseda.jp

ことが可能である．また，Sinkhorn アルゴリズムの計算量解析は盛んに行われており，l_1 ノルムにおける制約条件誤差に対し時間計算量が $\mathcal{O}(n^3/\epsilon^3)$ であると明らかにされている [17]．さらに，この計算量は $\mathcal{O}(n^2/\epsilon^2)$ に改善されている [18]．従来最適輸送問題だけでなく，緩和定式化である非平衡最適輸送問題に対するエントロピー正則化も提案されている [19]．当該問題に対しても従来 Sinkhorn アルゴリズムと類似したアルゴリズムを用いて解くことが可能であり，多くの分野でその有用性が報告されている．最近の研究では，非平衡最適輸送問題に対する Sinkhorn アルゴリズムの計算量解析も行われている．特に [20] では，最大値ノルムを利用してその収束性が保証し，時間計算量が $\mathcal{O}(n^2/\epsilon)$ であることを明らかにしている．

本稿では，半緩和最適輸送問題に対するエントロピー正則化を施し，その問題を効率的に解く Sinkhorn アルゴリズムを提案する．半緩和最適輸送問題は制約付き凸最適化問題であるため，基本的に解くのが難しい．エントロピー正則化を適用することで，解くのが容易である制約無し凸最適化問題に変換する．さらに提案アルゴリズムが必ず最適解に収束する結果を示す．

2. 関連研究と問題点

表記法について． $\mathbb{R}^n$ は n 次元ユークリッド空間を表す．$\mathbb{R}^{m\times n}$ は m 行 n 列の行列全体の集合を表す．ベクトルは $\boldsymbol{a}, \boldsymbol{b}, \ldots$ で表され，行列は $\mathbf{A}, \mathbf{B}, \ldots$ で表される．サイズが同じ行列 $\mathbf{A}, \mathbf{B}$ に対し，フロベニウスノルムは $\langle \mathbf{A}, \mathbf{C}\rangle = \mathrm{tr}(\mathbf{A}^T\mathbf{B})$ で示される．Kullback Leibler 情報量は $\mathrm{KL}(\boldsymbol{x}_i, \boldsymbol{y}_i) = \sum_i \boldsymbol{x}_i \log(\frac{\boldsymbol{x}_i}{\boldsymbol{y}_i}) - \boldsymbol{x}_i + \boldsymbol{y}_i$，エントロピーは $\mathrm{H}(\mathbf{T}) = -\sum_{i,j} \mathbf{T}_{i,j} \log \mathbf{T}_{i,j}$ で表される．同一空間内のベクトル $\boldsymbol{a}, \boldsymbol{b} \in \mathbb{R}^n$ に対し，その内積は $\boldsymbol{a}^T\boldsymbol{b}$ で表される．また，指数関数 $\exp(\cdot)$ や対数関数 $\log(\cdot)$ にベクトルが使用される場合，成分ごとに指数対数を取り計算を行う．さらに，ベクトル同士の乗算・除算も成分同士の乗算除算とみなす．

2.1 最適輸送問題

最適輸送問題は, 二つの経験的確率分布 $\boldsymbol{\nu} = \sum_{i=1}^m \boldsymbol{a}_i \boldsymbol{\delta}_{\boldsymbol{x}_i}$，$\boldsymbol{\mu} = \sum_{i=1}^n \boldsymbol{b}_i \boldsymbol{\delta}_{\boldsymbol{y}_i}$ 間の最適写像を求める Monge 問題として定義される．しかし，離散的な制約条件とその写像により，最適解を求めることは容易でない．これを対処するため，Kantorovich により連続的な定式化が提案された．その問題は次に定義される．

$$\min_{\mathbf{T} \in \mathcal{U}(\boldsymbol{a}, \boldsymbol{b})} \langle \mathbf{T}, \mathbf{C} \rangle \tag{1}$$

ここで $\mathbf{C}$ は輸送コストを表し，制約条件 $\mathcal{U}(\boldsymbol{a}, \boldsymbol{b})$ は

$$\mathcal{U}(\boldsymbol{a}, \boldsymbol{b}) = \{\mathbf{T} \in \mathbb{R}_+^{m\times n} : \mathbf{T}\mathbf{1}_n = \boldsymbol{a}, \mathbf{T}^T\mathbf{1}_m = \boldsymbol{b}\} \tag{2}$$

で定義される．実行領域 $\mathcal{U}(\boldsymbol{a}, \boldsymbol{b})$ は二つの確率 $\boldsymbol{a}, \boldsymbol{b}$ の厳密な質量保存を表す．この最適解 $\mathbf{T}^*$ を用いた確率分布間の距離は次に定義される．

$$\mathcal{W}(\boldsymbol{\mu}, \boldsymbol{\nu}) = (\langle \mathbf{T}^*, \mathbf{C} \rangle)^{\frac{1}{p}}. \tag{3}$$

式 (3) は p 次 Wasserstein 距離と呼ばれる．最適輸送問題は線形計画問題であるため，基本的に解くことが難しい．しかし，与えられた二つの分布が正規分布であるとき，最適輸送問題は閉形式となり，その解析解が存在するため，行列演算だけで最適解を計算することが可能である [21]．最適輸送問題の更なる詳細について [22] を参照されたい．

2.2 非平衡最適輸送問題と半緩和最適輸送問題

前章で示したように，最適輸送問題は凸線形計画法で定められるため，大規模データに対して効率的に解くことが困難である．さらに，特定の分野では，最適輸送問題における厳密な質量保存制約を表す制約条件がうまく機能しない．そこで，実行領域 (2) の制約条件を緩和する手法が提案されている．具体的には，次の非平衡最適輸送問題が定義される [10]．

$$\min_{\mathbf{T} \geq \mathbf{0}} \langle \mathbf{T}, \mathbf{C} \rangle + \tau \Phi(\mathbf{T}\mathbf{1}_n, \boldsymbol{a}) + \tau \Phi(\mathbf{T}^T\mathbf{1}_m, \boldsymbol{b}). \tag{4}$$

ここで $\Phi(\boldsymbol{x}, \boldsymbol{y})$ は滑らかな情報量である．さらに最適輸送問題の制約条件のうち片方のみを緩和する手法も存在する．この半緩和最適輸送問題は次に定義される [10]．

$$\min_{\mathbf{T} \geq \mathbf{0}, \mathbf{T}^T\mathbf{1}_m = \boldsymbol{b}} \langle \mathbf{T}, \mathbf{C} \rangle + \tau \Phi(\mathbf{T}\mathbf{1}_n, \boldsymbol{a}). \tag{5}$$

非平衡最適輸送問題は様々な機械学習の分野で利用されており，ドメイン適応 [23], [24]，生成モデル [25] やシーケンスマッチング [26] などの分野で成功を収めている．半緩和最適輸送問題はサブグラフ構造を意識したグラフマッチング [14] や人工物を除去した色相転写 [13] を可能にしている．

2.3 エントロピー正則化と Sinkhorn アルゴリズム

大規模データに対する最適輸送問題を効率的に解くために，エントロピー正則化が提案されている [16]．その問題は次に定義される．

$$\min_{\mathbf{T} \in \mathcal{U}(\boldsymbol{a}, \boldsymbol{b})} \sum_{i=1}^m \sum_{j=1}^n \mathbf{T}_{i,j}\mathbf{C}_{i,j} - \epsilon \mathrm{H}(\mathbf{T}). \tag{6}$$

エントロピー項の強凸性により，問題 (6) には大域的最適解が存在する．この最適解を求める Sinkhorn アルゴリズムはスケーラビリティを有しているため，大規模データに対する線形計画問題より効率的に解くことができる．また，その微分可能性により制約無し凸最適化問題を解くことを可能にしている．さらに，並列実行可能であるため実用的な実行も効果的である．最適輸送問題の近似解を得る

ために，多くの機械学習の分野でこのアルゴリズムが使用されている [27].

アルゴリズムの計算量解析も研究されており，その時間計算量は $\mathcal{O}(n^2/\epsilon^2)$ として知られている [18]．正則化の微小な値に対する不安定性や非ロバスト性が問題であったが，これらを改善したアルゴリズムも提案されている [19]．最近の研究では，非平衡最適輸送問題に対してもエントロピー正則化項を付加し，類似した Sinkhorn アルゴリズムも提案され，その計算量解析も行われている [19], [20]．理論解析において，入力する確率分布が正規分布であるとき，エントロピー正則化を施した最適輸送問題と非平衡最適輸送問題が閉形式を有することが示されている [28].

3. 提案手法

本章では，エントロピー正則化付き半緩和最適輸送問題を定義し，それを解く Sinkhorn アルゴリズムを与える．具体的には，エントロピー正則化付き半緩和最適輸送問題を定式化を行い，その問題に対するフェンシェル双対問題を導出する．どう出した双対問題を基にアルゴリズムを提案する．最後に，提案したアルゴリズムの性質を示す.

3.1 エントロピー正則化付き半緩和最適輸送問題

半緩和最適輸送問題 (5) にエントロピー正則化を加えた問題は次に定義される.

$$\min_{\mathbf{T} \geq \mathbf{0}, \mathbf{T}^T \mathbf{1}_n = \boldsymbol{b}} \langle \mathbf{T}, \mathbf{C} \rangle + \tau \mathrm{KL}(\mathbf{T}\mathbf{1}_n, \boldsymbol{a}) - \eta \mathrm{H}(\mathbf{T}). \quad (7)$$

主問題 (7) のフェンシェル双対問題を考える．双対問題の目的関数は次で与えられる．次の式変形で得られる.

$$\min_{\substack{\mathbf{T}^T \mathbf{1}_n = \boldsymbol{b} \\ \boldsymbol{y} \in \mathbb{R}_+^n}} \langle \mathbf{T}, \mathbf{C} \rangle + \tau \mathrm{KL}(\boldsymbol{y}, \boldsymbol{a}) - \eta \mathrm{H}(\mathbf{T}) + \boldsymbol{u}^T (\boldsymbol{y} - \mathbf{T}\mathbf{1}_n).$$

これは $\mathbf{T}, \boldsymbol{y}$ のそれぞれに関する最小問題として定められる.

$$\min_{\mathbf{T}^T \mathbf{1}_n = \boldsymbol{b}} \langle \mathbf{T}, \mathbf{C} \rangle - \boldsymbol{u}^T \mathbf{T} \mathbf{1}_n - \eta \mathrm{H}(\mathbf{T}) \quad (8)$$

$$\min_{\boldsymbol{y} \in \mathbb{R}_+^n} \boldsymbol{u}^T \boldsymbol{y} + \tau \mathrm{KL}(\boldsymbol{y}, \boldsymbol{a}). \quad (9)$$

第一項 (8) の最小化問題は KKT 条件を考えることで 求められる．その関数は次で目的関数が定義される.

$$-\eta \sum_{i,j} \exp\left(\frac{\boldsymbol{u}_i + \boldsymbol{\nu}_j - \mathbf{C}_{i,j}}{\eta}\right) + \boldsymbol{v}^T \boldsymbol{b}.$$

ここで，$\boldsymbol{v} \in \mathbb{R}^n$ である．第二項 (9) は凸関数であるため，勾配が $\mathbf{0}$ であるベクトルを求めれば，その最小値を求められる．この計算を施すと第二項の関数は次で与えられる.

$$-\tau \left(\boldsymbol{a}^T \exp\left(-\frac{\boldsymbol{u}}{\eta}\right) - \boldsymbol{a}^T \mathbf{1}_n\right).$$

以上を統合すると，双対問題は $\boldsymbol{u}, \boldsymbol{v}$ の最大化問題として定式化される．本稿では，この最大化問題の符号を反転し，最小化問題を解く．この問題は次に定義する.

$$\min_{\boldsymbol{u}, \boldsymbol{v}} \eta \sum_{i,j} \exp\left(\frac{\boldsymbol{u}_i + \boldsymbol{\nu}_j - \mathbf{C}_{i,j}}{\eta}\right) - \boldsymbol{v}^T \boldsymbol{b} + \tau \boldsymbol{a}^T \exp\left(-\frac{\boldsymbol{u}}{\eta}\right). \quad (10)$$

以下双対問題の目的関数を $h(\boldsymbol{u}, \boldsymbol{v})$ と記す.

3.2 提案アルゴリズム

双対問題 (10) を解くための交互最適化アルゴリズムを提案する．双対問題の目的関数 h は $\boldsymbol{u}, \boldsymbol{v}$ に関して両凸関数であるので，交互最適化を施しても大域的最適解に収束する．反復数 k の時に提案アルゴリズムで得られる解 $\boldsymbol{u}^{(k)}, \boldsymbol{v}^{(k)}$ を固定して，$k+1$ 回目の解を求める．まず $\boldsymbol{u}^{(k+1)}$ を $\boldsymbol{u}$ に関する h の勾配を用いて求める．その勾配が零ベクトルとなる時，次の等式が満たされる.

$$\boldsymbol{a}_i \exp\left(-\frac{\boldsymbol{u}_i^{(k+1)}}{\tau}\right) = \eta \sum_j \exp\left(\frac{\boldsymbol{u}_i^{(k+1)} + \boldsymbol{v}_j^{(k)} - \mathbf{C}_{ij}}{\eta}\right).$$

ここで k 回目の輸送行列 $\mathbf{T}_{i,j}^{(k)} = \exp\left(\frac{\boldsymbol{u}_i^{(k)} + \boldsymbol{v}_j^{(k)} - \mathbf{C}_{ij}}{\eta}\right)$, $\boldsymbol{a}^{(k)} = \mathbf{T}^{(k)} \mathbf{1}_n$ を定める．上記等式に対して $\frac{\boldsymbol{u}_i^{(k)}}{\eta}$ を両辺に掛けると，右辺の総和を $\boldsymbol{a}_i^{(k)}$ に置換できる．その後，両辺の対数を取ることにより $\boldsymbol{u}^{(k+1)}$ に対する更新式が得られる．その更新式は次に定められる.

$$\boldsymbol{u}^{(k+1)} = \frac{\tau\eta}{\tau+\eta}\left(\frac{\boldsymbol{u}^{(k)}}{\eta} + \log(\boldsymbol{a}) - \log(\boldsymbol{a}^{(k)})\right) \quad (11)$$

同様に $\boldsymbol{v}$ に対する勾配を利用して $\boldsymbol{v}^{(k+1)}$ の更新式も求められる．$\boldsymbol{b}^{(k)} = (\mathbf{T}^{(k)})^T \mathbf{1}_n$ を定めると $\boldsymbol{v}^{(k+1)}$ に関する更新式は次に定義される.

$$\boldsymbol{v}^{(k+1)} = \boldsymbol{v}^{(k)} + \eta(\log(\boldsymbol{b}) - \log(\boldsymbol{b}^{(k)})) \quad (12)$$

更新式 (11)，(12) を統合したものを，アルゴリズム 1 に示す.

3.3 偶数時における更新式の性質

提案アルゴリズムにおける偶数時の更新式は主問題の制約条件の射影である.

Theorem 3.1. *アルゴリズム 1* の偶数時の更新後により得られる輸送行列 $\mathbf{T}^{(k+1)}$ は主問題の制約条件 $\mathbf{T}^T \mathbf{1}_n = \boldsymbol{b}$ を満たす.

4. 収束解析

本章では，提案アルゴリズムの収束解析の結果を示す．本収束解析では，従来研究である非平衡最適輸送問題に対する Sinkhorn アルゴリズムの収束解析 [20] を参考にし，

Algorithm 1 Entropy Semi-Relaxed Sinkhorn Algorithm

Require: $\boldsymbol{a}, \boldsymbol{b}, \eta, \tau, \boldsymbol{u}^{(0)} = \boldsymbol{0}, \boldsymbol{v}^{(0)} = \boldsymbol{0}$
Ensure: $\boldsymbol{u}^{(k)}, \boldsymbol{v}^{(k)}$

for $k = 0 \ldots K$ **do**
  $\mathbf{T}^{(k)} = \mathrm{diag}(\exp(\frac{\boldsymbol{u}^{(k)}}{\eta}))\exp(-\frac{\mathbf{C}}{\eta})\mathrm{diag}(\exp(\frac{\boldsymbol{v}^{(k)}}{\eta}))$
  $\boldsymbol{a}^{(k)} = \mathbf{T}^{(k)}\mathbf{1}_n$
  if k is even **then**
    $\boldsymbol{u}^{(k+1)} = \frac{\tau\eta}{\tau+\eta}(\frac{\boldsymbol{u}^{(k)}}{\eta} + \log(\boldsymbol{a}) - \log(\boldsymbol{a}^{(k)}))$
    $\boldsymbol{v}^{(k+1)} = \boldsymbol{v}^{(k)}$
  else
    $\boldsymbol{u}^{(k+1)} = \boldsymbol{u}^{(k)}$
    $\boldsymbol{v}^{(k+1)} = \boldsymbol{v}^{(k)} + \eta(\log(\boldsymbol{b}) - \log(\boldsymbol{b}^{(k)}))$
  end if
end for

提案アルゴリズムの収束性を保証している．その内容は次に示される．

Theorem 4.1. アルゴリズム *1* で得られる反復数 k に対する解 $\boldsymbol{u}^{(k+1)}, \boldsymbol{v}^{(k+1)}$ は次の不等式を満たす．

$$\max\left\{\| \boldsymbol{u}^{(k+1)} - \boldsymbol{u}^* \|_\infty, \| \boldsymbol{v}^{(k+1)} - \boldsymbol{v}^* \|_\infty\right\} \leq (2\tau+\eta)\left(\frac{\tau}{\tau+\eta}\right)^{\frac{k+2}{2}} R. \tag{13}$$

定理 4.1 に示される不等式の右辺は反復するに連れて 0 に収束していく．さらに，不等式の左辺も 0 以上である．そのため，$\boldsymbol{u}^{(k+1)}, \boldsymbol{v}^{(k+1)}$ は最適解に収束する．

5. まとめ

本稿では，エントロピー正則化付き半緩和最適輸送問題の定式化を行い，その問題の最適解を求める交互最適化アルゴリズムを提案した．さらに提案アルゴリズムが必ず最適化に収束することを示した．

今後の課題は二つ存在する．一つ目は提案アルゴリズムの収束割合を明らかにすることである．提案アルゴリズムが少なくとも収束することは示されたが，具体的な停止条件に対して収束するのに必要な具体的な回数は未だ明らかになっていない．二つ目は，提案アルゴリズムによって得られる近似解と半緩和最適輸送問題の最適解を任意精度で近似できるかどうかである．二つの課題を解決することで，任意精度で近似するのに必要な時間計算量を担保できる．

参考文献

[1] Villani, C.: *Optimal transport: Old and new*, Springer (2008).
[2] Huang, J. and Kasai, H.: Graph embedding using multi-layer adjacent point merging model, *ICASSP* (2021).
[3] Huang, J. and et al., Z. F.: LCS graph kernel based on Wasserstein distance in longest common subsequence metric space, *Signal Processing*, Vol. 189, p. 108281 (2021).
[4] Kasai, H.: Multi-View Wasserstein Discriminant Analysis with Entropic Regularized Wasserstein Distance, *ICASSP* (2020).
[5] Fukunaga, T. and H., K.: Wasserstein k-means with sparse simplex projection, *ICPR* (2020).
[6] Kerdoncuff, T. and Emonet, R. e. a.: Metric Learning in Optimal Transport for Domain Adaptation, *IJCAI* (2020).
[7] Bonneel, N. and van de Panne, M. e. a.: Displacement Interpolation Using Lagrangian Mass Transport, *ACM Trans. Graph.*, Vol. 30, pp. 1–12 (2011).
[8] Benamou, Jean-David: Numerical resolution of an "unbalanced" mass transport problem, *ESAIM: M2AN*, Vol. 37, No. 5, pp. 851–868 (2003).
[9] Frogner, C. and et al., C. Z.: Learning with a Wasserstein Loss, *arXiv preprint arXiv:1506.05439* (2015).
[10] Blondel, M. and et al., V. S.: Smooth and Sparse Optimal Transport, *AISTATS* (2018).
[11] Fukunaga, T. and Kasai, H.: Fast block-coordinate Frank-Wolfe algorithm for semi-relaxed optimal transport, *arXiv preprint: arXiv:2103.05857* (2021).
[12] Ferradans, S. and Papadakis, N. e. a.: Regularized Discrete Optimal Transport, *SIAM Journal on Imaging Sciences*, Vol. 7, No. 3, pp. 1853–1882 (2013).
[13] Julien, R. and et al., F. S.: Adaptive Color Transfer With Relaxed Optimal Transport, *ICIP* (2014).
[14] Vincent-Cuaz, C. and et al., R. F.: Semi-relaxed Gromov Wasserstein divergence with applications on graphs, *arXiv preprint: arXiv:2110.02753* (2021).
[15] Chapel, L. and Alaya, M. Z. e. a.: Partial Optimal Tranport with applications on Positive-Unlabeled Learning, *NIPS* (2020).
[16] Cuturi, M.: Sinkhorn Distances: Lightspeed Computation of Optimal Transport, *NIPS* (2013).
[17] Altschuler, J. and Niles-Weed, J. e. a.: Near-linear time approximation algorithms for optimal transport via Sinkhorn iteration, *NIPS* (2017).
[18] Dvurechensky, P. E. and et al., A. V. G.: Computational Optimal Transport: Complexity by Accelerated Gradient Descent Is Better Than by Sinkhorn's Algorithm, *ICML* (2018).
[19] Chizat, L. and et al., G. P.: Scaling Algorithms for Unbalanced Transport Problems, *arXiv preprint: arXiv:1607.05816* (2017).
[20] Pham, K. and Le, K. e. a.: On Unbalanced Optimal Transport: An Analysis of Sinkhorn Algorithm, *ICML* (2020).
[21] Takatsu, A.: Wasserstein geometry of Gaussian measures, *Osaka J.Math*, Vol. 48, No. 4, pp. 1005–1026 (2011).
[22] Peyre, G. and Cuturi, M.: Computational Optimal Transport, *arXiv preprint: arXiv:1803.00567* (2020).
[23] Fatras, K. and et al., Y. Z.: Minibatch optimal transport distances; analysis and applications, *ICML* (2021).
[24] Courty, N. and Flamary, R. e. a.: Optimal Transport for Domain Adaptation, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 39, No. 9, pp. 1853–1865 (2017).
[25] Yang, K. D. and Uhler, C.: Scalable Unbalanced Optimal Transport using Generative Adversarial Networks, *arXiv preprint: arXiv:1810.11447* (2019).
[26] Janati, H. and Cuturi, M. e. a.: Spatio-temporal alignments: Optimal transport through space and time, *AISTATS* (2020).
[27] Cuturi, M. and Doucet, A.: Fast Computation of Wasserstein Barycenters, *ICML* (2014).
[28] Janati, H. and Muzellec, B. e. a.: Entropic Optimal Transport between (Unbalanced) Gaussian Measures has a Closed Form, *NIPS* (2020).