

# Fast Unbalanced Optimal Transport on Tree

Thanh-Tung Nguyen, Thu-Le Tran\*, Tam Le, Clément Elvira

## Contents

1. Abstract .....	1
2. Preliminaries .....	1
2.1. Motivations .....	1
2.2. Balanced OT on tree .....	2
2.3. Unbalanced OT on tree .....	3
3. The proposed model for unbalanced OT based on Lasso .....	3
4. Algorithm for the proposed UOT .....	4
5. Future works .....	4
References .....	5

## 1. Abstract

In this paper we propose a Lasso-based model for unbalanced Optimal Transport (OT) on tree that can be efficiently computed in linear time. This improves the previous time complexity  $O(n \log^2 n)$  using Generalized Kantorovich Rubinstein (GKR) model [1].

## 2. Preliminaries

### 2.1. Motivations

OT is important in machine learning. OT problem aims to find a way to transport mass from supply locations to demand locations. In the standard model of OT, we assume the balance between supply and demand, this model is referred to as balanced OT.

However balanced condition does not hold in practice. To handle the unbalanced cases, there are multiple approaches:

1. **Relaxation Approaches:** The most well-known relaxation for UOT is Entropic OT [2].
2. **Regularization Approaches:** Observing that the optimal plan in Entropic OT is not sparse, other regularization methods has been introduced including L2 norm [3], k-norm [4].

3. **Tree-based Approaches:** Solving OT for a huge number of data points in  $\mathbb{R}^d$  is time consuming.

Recent approaches use line [5] and tree [6] to approximate the data points, then solve OT on line and tree. The main benefit of doing this is that balanced OT on tree (and thus, on line) has closed-form expression allowing the evaluation of balanced OT on tree in  $O(n)$ . For the unbalanced OT on tree, a infimum-convolution based approach has been proposed with time complexity  $O(n \log^2 n)$  using the Generalized Kantorovich Rubinstein (GKR) model [1].

**Contribution.** In this paper, we propose a Lasso-based model for unbalanced OT on tree allowing the computation in  $O(n)$  using Soft-thresholding operator.

## 2.2. Balanced OT on tree

Given a directed tree  $T$  with vertex set  $V = \{v_1, \dots, v_n\}$  and edge set  $E$  and root  $v_1$ . On  $T$ , there is a supply vector  $a \in \mathbb{R}_+^n$  and demand vector  $b \in \mathbb{R}_+^n$ . Each edge  $e \in E$  has a nonnegative length say  $\ell(e)$ . The edge lengths generate a distance between any two nodes of the tree.

Without loss of generality, we assume in this paper that<sup>1</sup>

*The intersection of support of  $a$  and support  $b$  is empty*

If the total supply is not less than total demand, i.e.

$$\sum_{i=1}^n a_i = \sum_{j=1}^n b_j$$

then we have *balanced OT* model

$$\min \sum_{i,j=1}^n D_{ij} T_{ij}$$

$$\text{s.t. } T \mathbf{1}_n = a$$

$$T^T \mathbf{1}_n = b$$

$$T \in \mathbb{R}_+^{n \times n}$$

where  $D_{ij}$  is the distance between  $v_i$  and  $v_j$  in  $V$  and  $T_{ij}$  is the mass transported from supply at  $v_i$  to demand at  $v_j$ .

We denote the minimum value by  $\text{OT}(a, b)$ . In [6], the authors showed that  $\text{OT}(a, b)$  has closed-form expression:

$$\text{OT}(a, b) = \sum_{i=2}^n \ell(e_i) |\hat{a}_i - \hat{b}_i|.$$

---

<sup>1</sup>Actually, we need to prove this claim.

where  $e_i$  is the edge connecting  $v_i$  and its parent,  $\hat{a}_i$  (and  $\hat{b}_i$ ) is the cumulative distribution, i.e.  $\hat{a}_i = \sum_{v_j \in T_i} a_j$  ( $\hat{b}_i = \sum_{v_j \in T_i} b_j$ ) with  $T_i$  the set of all descendants of  $v_i$  including itself. Notice that  $\hat{a}$  and  $\hat{b}$  can be computed recursively from the bottom-up approach, thus  $\text{OT}(a, b)$  can be computed in  $O(n)$ .

It is important to notice that OT formula can be rewritten in vector form

$$\text{OT}(a, b) = \|A(a - b)\|_1$$

where  $A = \Lambda H \in \mathbb{R}^{n \times n}$  where  $\Lambda = \text{diag}(0, \ell(e_2), \dots, \ell(e_n))$  and  $H$  is the matrix mapping  $a$  to  $\hat{a}$ , i.e.  $Ha = \hat{a}$ .

### 2.3. Unbalanced OT on tree

One way to model the unbalanced OT is to use the Generalized Kantorovich Rubinstein (GKR) distance defined as follows

$$\text{OT}(a, b) = \min_{T \in U} \sum_{i,j=1}^n D_{ij} T_{ij} + \langle \lambda_d, a - T \mathbf{1}_n \rangle + \langle \lambda_c, b - T^T \mathbf{1}_n \rangle$$

where

$$U = \{T \in \mathbb{R}_+^{n \times n} : T \mathbf{1}_n \leq a, T^T \mathbf{1}_n \leq b\}.$$

Loosely speaking, GRK model finds a plan  $T \in U$  such and then penalize when there is any destruction of supply or any create of demand.

### 3. The proposed model for unbalanced OT based on Lasso

Now, assuming that there is no balance and without loss of generality, we assume that the total supply is not less than total demand

$$\sum_{i=1}^n a_i \geq \sum_{j=1}^n b_j$$

We assume that

- There are destructions of supply denoted by  $x \in \mathbb{R}_+^n$ , so the new supply is  $a - x$
- There are creations of demand denoted by  $y \in \mathbb{R}_+^n$ , so the new demand is  $b + y$
- There is a balance between new supply and new demand, i.e.

$$\sum_{i=1}^n a_i - x_i = \sum_{j=1}^n b_j + y_j$$

Which can be rewritten as

$$\sum_{i=1}^n x_i + y_i = \Delta := \sum_{i=1}^n a_i - b_i$$

We propose to consider the following problem

$$\min_{x, y \in \mathbb{R}_+^n} \text{OT}(a - x, b + y) + \langle \lambda_d, x \rangle + \langle \lambda_c, y \rangle$$

$$\text{s.t.} \quad \sum_{i=1}^n x_i + y_i = \Delta.$$

Since  $a - x$  and  $b - y$  satisfy the balanced condition on tree, thus  $\text{OT}(a - x, b + y)$  can be computed in closed-form

$$\text{OT}(a - x, b + y) = \|A((a - x) - (b + y))\|_1 = \|c - Az\|_1.$$

where  $c = A(a - b)$  and  $z = x + y$ . We have the problem

- If  $\lambda = \lambda_d = \lambda_c$ , we have L1 norm regression.

$$\min \|c - Az\|_1.$$

#### 4. Algorithm for the proposed UOT

In the current version of the paper, we assume that  $\lambda_d = \lambda_c \in \mathbb{R}$ , then we have the problem

$$\min_{z \in \mathbb{R}_+^n} \sum_{i=2}^n \ell(e_i) |\hat{a}_i - \hat{b}_i - \hat{z}_i|.$$

#### 5. Future works

- When  $\lambda_d \neq \lambda_c$ , how to solve the problem?

## References

- [1] R. Sato, M. Yamada, and H. Kashima, “Fast unbalanced optimal transport on a tree,” *Advances in neural information processing systems*, vol. 33, pp. 19039–19051, 2020.
- [2] M. Cuturi, “Sinkhorn distances: Lightspeed computation of optimal transport,” *Advances in neural information processing systems*, vol. 26, 2013.
- [3] M. Blondel, V. Seguy, and A. Rolet, “Smooth and sparse optimal transport,” in *International conference on artificial intelligence and statistics*, 2018, pp. 880–889.
- [4] T. Liu, J. Puigcerver, and M. Blondel, “Sparsity-constrained optimal transport,” *arXiv preprint arXiv:2209.15466*, 2022.
- [5] N. Bonneel and D. Coeurjolly, “Spot: sliced partial optimal transport,” *ACM Transactions on Graphics (TOG)*, vol. 38, no. 4, pp. 1–13, 2019.
- [6] T. Le, M. Yamada, K. Fukumizu, and M. Cuturi, “Tree-sliced variants of Wasserstein distances,” *Advances in neural information processing systems*, vol. 32, 2019.