# On the Computation of Kantorovich-Wasserstein Distances between 2D-Histograms by Uncapacitated Minimum Cost Flows

Maria Despoina Siampou (sdi1600151), George Michas (sdi1400109)

HELLENIC REPUBLIC
**National and Kapodistrian University of Athens**

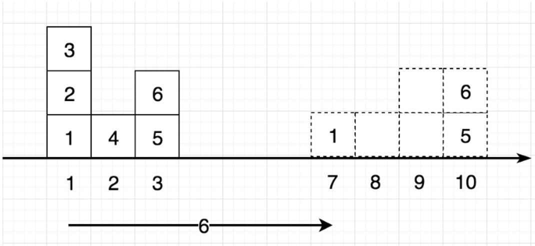
May 26, 2019

# Contents

# Contents

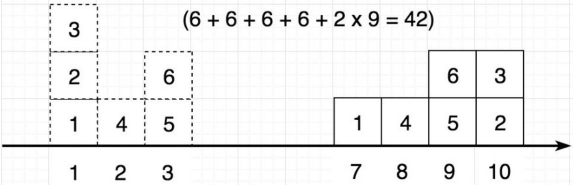
# Example

Let's try to move box 1, from location 1 to location 7. The moving cost equals to its weight times the distance. For simplicity, we will set the weight to be 1. Therefore the cost to move box 1 equals to 6 (7–1).



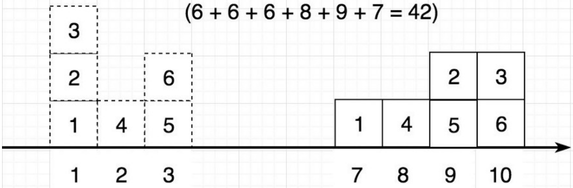
# Example

The tables in the right illustrates how boxes are moved.




$\gamma_1$

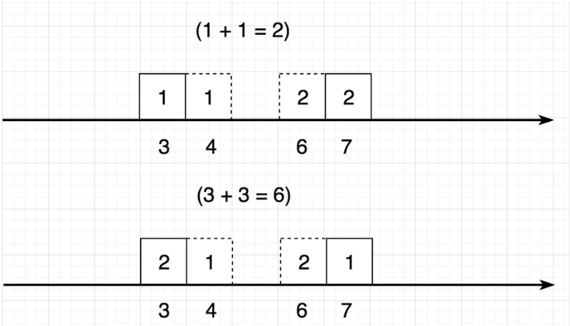
| | 7 | 8 | 9 | 10 |
|---|---|---|---|---|
| 1 | 1 | 0 | 0 | 2 |
| 2 | 0 | 1 | 0 | 0 |
| 3 | 0 | 0 | 2 | 0 |




$\gamma_2$

| | 7 | 8 | 9 | 10 |
|---|---|---|---|---|
| 1 | 1 | 0 | 1 | 1 |
| 2 | 0 | 1 | 0 | 0 |
| 3 | 0 | 0 | 1 | 1 |

However, not all transport plans bear the same cost. The Wasserstein distance (or the EM distance) is the cost of the cheapest transport plan. In the example below, both plans have different cost and the Wasserstein distance (minimum cost) is two.



# Introduction

In mathematics, the Wasserstein or Kantorovich-Rubinstein metric or distance is a distance function defined between probability distributions on a given metric space M.

# Intuition and connection to optimal transport

One way to understand the motivation of the above definition is to consider the optimal transport problem. That is, for a distribution of mass $\mu(x)$ on a space X, we wish to transport the mass in such a way that it is transformed into the distribution $\nu(x)$ on the same space, transforming the 'pile of earth' $\mu(x)$ to the pile $\nu(x)$. This problem only makes sense if the pile to be created has the same mass as the pile to be moved; therefore without loss of generality assume that $\mu(x)$ and $\nu(x)$ are probability distributions containing a total mass of 1.

# Intuition and connection to optimal transport

Assume also that there is given some cost function: $c(x,y) \mapsto [0,\infty)$ That gives the cost of transporting a unit mass from the point x to the point y. A transport plan to move $\mu$ into $\nu$ can be described by a function $\gamma(x,y)$ which gives the amount of mass to move from x to y. In order for this plan to be meaningful, it must satisfy the following properties:

$$\int \gamma(x',x)\mathrm{d}x' = \nu(x)$$

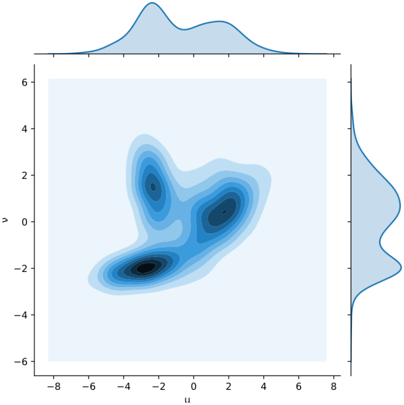
$$\int \gamma(x,x')\mathrm{d}x' = \mu(x)$$

# Intuition and connection to optimal transport

That is, that the total mass moved out of an infinitesimal region around x must be equal to $\mu(x)\mathrm{d}x$ and the total mass moved into a region around x must be $\nu(x)\mathrm{d}x$. This is equivalent to the requirement that $\gamma$ be a joint probability distribution with marginals $\mu$ and $\nu$ . Thus, the infinitesimal mass transported from x to y is $\gamma(x,y)\mathrm{d}x\mathrm{d}y$, and the cost of moving is $c(x,y)\gamma(x,y)\mathrm{d}x\mathrm{d}y$, following the definition of the cost function. Therefore, the total cost of a transport plan $\gamma$ is :

$$\iint c(x,y)\gamma(x,y)\mathrm{d}x\mathrm{d}y = \int c(x,y)\mathrm{d}\gamma(x,y)$$

The plan $\gamma$ is not unique; the optimal transport plan is the plan with the minimal cost out of all possible transport plans. If the cost of a move is simply the distance between the two points, then the optimal cost is identical to the definition of the $W_1$ distance.



# Contents

# Introduction

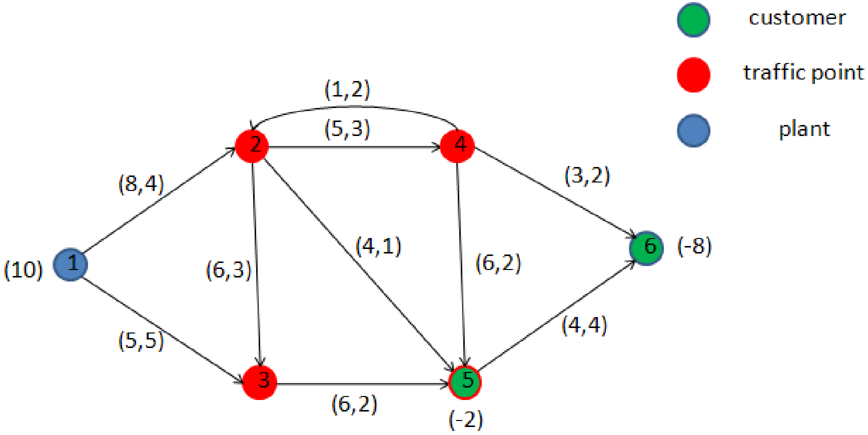
The minimum-cost flow problem (MCFP) is an optimization and decision problem to find the cheapest possible way of sending a certain amount of flow through a flow network.

A typical application of this problem involves finding the best delivery route from a factory to a warehouse where the road network has some capacity and cost associated.
The minimum cost flow problem is one of the most fundamental among all flow and circulation problems because most other such problems can be cast as a minimum cost flow problem and also that it can be solved efficiently using the network simplex algorithm.

# Definition

- D = (V, A): directed graph
- $\omega_{uv}$ unit flow cost through arc (u, v) $\in$ A.
- $c_{uv}$ : capacity of arc (u, v) $\in$ A.
- $b_u$: Value associated with each node u $\in$ V, called balance of u.
  - if $b_u > 0$, node u is called a source and $b_u$ its production or supply
  - if $b_u < 0$, node u is called a sink and $b_u$ its demand
  - if $b_u = 0$, node u is called an intermediate node.

# Example


customer
traffic point
plant
(1,2)
(5,3)
(8,4)
(3,2)
(10)
1
2
4
(6,3)
(4,1)
(6,2)
6
(-8)
(5,5)
(4,4)
3
5
(6,2)
(-2)