Math 217: Topics in Applied Math
Optimal Transport
Spring 2022,
Instructor: Bo Li
Lectures: 1:00-1:50, MWF, AP&M 542

Lecture 1. Monday, 3/28/2022

- Brief description of the course.
- Discrete optimal transport (OT)
  Monge's formulation

About Optimal Transport: A subject of optimization, PDEs, the calculus of variations, probability, geometry, etc. with applications in economics, imaging science, molecular modeling, machine learning, etc.

About this course

- Applied math/comput. math aspects of optimal transport (OT)
- Introduction of the subject
- Highlight some research areas/projects
Topics to be covered

- Monge and Kantorovich formulations, equivalence, duality, existence, etc.
- Wasserstein metric
- Discrete OT, regularization, numerical methods, stability, and convergence
- Gradient flow with Wasserstein metric, Fokker-Planck equation and other evolution equations
- Applications in machine learning and molecular dynamics

References See the course web page.

Expect Participation in class discussions. Possibly read some papers/presentations.
Part I  Discrete Optimal Transport (OT)

Monge's formulation of discrete OT problem

\[ x_1 a_1 \rightarrow y_1 b_1 \]
\[ x_2 a_2 \rightarrow y_2 b_2 \]
\[ \vdots \]
\[ x_m a_m \rightarrow y_n b_n \]

Given/consider:
\[ X = \{x_1, \ldots, x_m\}, \]
\[ a_1, \ldots, a_m \in (0,1), \sum_i a_i = 1 \]
\[ Y = \{y_1, \ldots, y_n\}, \]
\[ b_1, \ldots, b_n \in (0,1), \sum_j b_j = 1. \]

Amount \(a_i\) of raw material at warehouse \(X_i\).
Amount \(b_j\) of raw material to be transported to factory \(Y_j\).

Possible/feasible transport map
\[ T: X \rightarrow Y \text{ such that} \]
\[ \begin{align*}
  b_j &= \sum_{i: T x_i = y_j} a_i \\
                  \quad (j=1,\ldots,n) \tag{*}
\end{align*} \]

Denote
\[ J = \{ \text{all } T: X \rightarrow Y \text{ satisfying (*)} \} \]
Cost function

\[ C : X \times Y \rightarrow [0, \infty) \]

\[ C(x_i, y_j) : \text{the cost to transport a unit product from } x_i \text{ to } y_j. \]

E.g., \( C(x_i, y_j) = \rho(x_i, y_j) \) if \( X, Y \)

are subsets of a metric space with metric \( \rho \).

The total cost for \( T \in T \) is

\[ E[T] = \sum_{i=1}^{m} a_i C(x_i, T(x_i)). \]

Monge's (discrete) OT problem

Find \( \hat{T} \in T \) s.t.

\[ E[\hat{T}] = \min_{T \in T} E[T]. \]

Call \( \hat{T} \) an optimal transport map.

Remarks The constraint (\( \rho \)) is crucial.

\( T \) soln \( \iff \) \( T \neq \emptyset \).

Soln may not be unique.

Exercise Example of nonuniqueness?

Since each \( b_j > 0 \), we have \( T \) is onto (i.e., surjective). Hence, \( m \geq n \).
Consider the case \( N = n \).

Each \( T \) is a bijection \((1-1, \text{onto})\).

If all \( a_i \) are distinct, then \( T \) is unique. \( T x_i = y_j \) with \( b_j = a_i \).

If some \( a_i \) are the same, then the problem can be decomposed into some subproblems.

\[
\begin{align*}
T_1 : & \{ a_1, x_1 \} \rightarrow \{ a_1, y_1 \} \\
T_2 : & \{ a_2, x_2 \} \rightarrow \{ a_2, y_2 \} \\
& \vdots \\
T_k : & \{ a_k, x_{k+1} \} \rightarrow \{ a_k, y_{k+1} \}
\end{align*}
\]

\[
\sum_{k=1}^{n} \text{Ne} = n. \quad \text{Find } T_j.
\]

Two steps:
1. Relabel \( x_i \) and \( y_j \).
2. For each group of \( x_i \) with same \( a \)-value, solve the OT problem.

The optimal assignment problem (as a

Given: \( X = Y = \{ 1, 2, \ldots, n \} \).

\( a_i = b_i = \frac{1}{n} \) \((i = 1, \ldots, n)\). Marge's prob

of discret OT)
$$C = [C_{ij}] \subset \mathbb{R}^{m \times n}, \text{ all } C_{ij} \geq 0.$$  

Notation: $\mathbb{R}^{m \times n} = \{ \text{ all } m \times n \text{ real matrices} \}$

$S_n = \{ \text{ all permutations of } (1 \ldots n) \}$

E.g. $S_3 = \{ (123), (132), (213), (231), (312), (321) \}$

A permutation is a bijective function.

**The (optimal) assignment problem (in Monge form)**

Find $\hat{\sigma} \in S_n$ s.t.

$$\hat{\sigma} = \arg \min_{\sigma \in S_n} \frac{1}{n} \sum_{i=1}^{n} C_{i, \sigma(i)}$$

- Hence, the cost function is

  $$C(i, j) = C_{ij}, \quad i, j = 1, \ldots, n.$$  

- $c(x_i, T(x_i)) = c(x_i, y_{\sigma i}) = C_{i, \sigma i}.$

- So $\sigma$s exist but often non-unique. But $\text{card}(S_n) = n!$

  $$5! = 120, \quad 8! = 40,320, \quad 10! = 3,628,800$$  
  $$12! = 479,001,600, \quad 25! = 5,551,840, \quad 70! = 1,982,600,269,404,490,276,687,090,270,186,096,000$$

We will revisit this problem later.

The case $m > n$. More complicated.

But, interesting?

**Exercise** Solve this problem.
Kantorovich’s formulation of the 
(discrete) OT problem

- Allowing the split of $a_i$ into pieces
- Probabilistic approach

\[\begin{align*}
x_1 & \to b_1 \quad y_1 \\
x_2 & \to b_2 \quad y_2 \\
\vdots & \quad \vdots \\
x_m & \to b_n \quad y_n
\end{align*}\]

Given/consider:

- $X = \{x_1, \ldots, x_m\}$, $Y = \{y_1, \ldots, y_n\}$
  $a_i \geq 0$, $\sum a_i = 1$, $b_j \geq 0$, $\sum b_j = 1$
- $c_{ij} (\geq 0) =$ cost for transporting a unit product from $x_i$ to $y_j$
  $C = [C_{ij}] \in \mathbb{R}^{m \times n}$: cost matrix
- Feasible transport plans
  $P = [P_{ij}] \in \mathbb{R}^{m \times n}$, all $P_{ij} \geq 0$
  $P_{ij} =$ amount of product $i$ (i.e., part of $a_i$) at $x_i$ transported to $y_j$, becoming part of $b_j$. 
\[ \sum_{i=1}^{m} \sum_{j=1}^{n} p_{ij} = b_j \quad (j = 1, \ldots, n), \quad \text{col. sum of } P = b \]
\[ \sum_{j=1}^{n} p_{ij} = a_i \quad (i = 1, \ldots, m), \quad \text{row sum of } P = a \]

Note: \( \sum_{i,j} p_{ij} = 1 \).

The feasible set of transport plans:
\[ A(a,b) = \{ P = [p_{ij}] \in \mathbb{R}_{\geq 0}^{m \times n} : \text{ all } p_{ij} \geq 0, \quad \sum_i p_{ij} = b_j, \quad \forall j, \quad \sum_j p_{ij} = a_i, \quad \forall i \} \]

Call \( E(P) := \sum_{i,j} p_{ij} c_{ij} \) the total cost of the plan \( P \).

\underline{Kantorovich’s formulation}

Find \( \hat{P} \in A(a,b) \) s.t.
\[ \hat{P} = \arg \min_{P \in A(a,b)} \sum_{i,j} p_{ij} c_{ij} \]