A modern application of Linear Programming: Optimal transport – Part I

The optimal transport problem as a Linear Program

Disclaimers

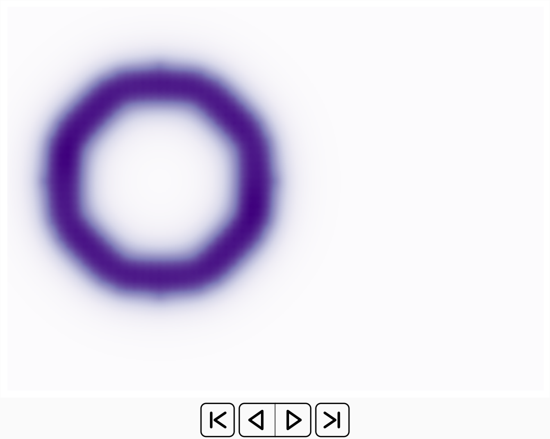
You will not be tested on that, and this will not help you to understand better Linear Programming.

Disclaimers

You will not be tested on that, and this will not help you to understand better Linear Programming.

The goal is to present you a modern direction of research where Linear Programming plays a role.

Teaser



The content of this lecture will explain how this movie was generated.

1. The optimal transport problem
2. A special case: Birkhoff Von Neumann theorem
3. Link with geometry: interpolating “shapes”

1. The optimal transport problem

An economic introduction 1

Coal mines extract coal, while factories need coal to operate. We want to find the most efficient *transportation* plan for the coal from mines to factories.

An economic introduction 1

Coal mines extract coal, while factories need coal to operate. We want to find the most efficient *transportation* plan for the coal from mines to factories.

Mine Number	Production	Factory number	Requirement
1	12	1	10
2	8	2	5
3	3	3	15
4	7		

As usual, values are made up.

An economic introduction 1

Coal mines extract coal, while factories need coal to operate. We want to find the most efficient *transportation* plan for the coal from mines to factories.

Mine Number	Production	Factory number	Requirement
1	12	1	10
2	8	2	5
3	3	3	15
4	7		

As usual, values are made up.

We make the important assumption that the total number of coal extracted (30) is equal to the total number of coal used by the factories (30), even though the number of mines and factories are different.

An economic introduction 2

We assume that the cost to transport one unit of coal from mine i to factory j is equal to C_{ij} with the following data:

Mine \ Factory	1	2	3
1	7	5	8
2	1	3	15
3	16	18	3
4	5	2	11

An economic introduction 2

We assume that the cost to transport one unit of coal from mine i to factory j is equal to C_{ij} with the following data:

Mine \ Factory	1	2	3
1	7	5	8
2	1	3	15
3	16	18	3
4	5	2	11

Importantly, we assume that the cost is *linear* in the amount of mass transported.

The Linear Program: Decision variables

Let's call P_{ij} the amount of coal transported from mine i to factory j. These are the decision variables and they are non negative. There are $4 \times 3 = 12$ of them.

The Linear Program: Decision variables

Let's call P_{ij} the amount of coal transported from mine i to factory j. These are the decision variables and they are non negative. There are $4 \times 3 = 12$ of them.

The total cost is

$$\sum_{i=1}^{4}\sum_{j=1}^{3} P_{ij}C_{ij} = 7P_{11} + 5P_{12} + 8P_{13} + P_{21} + 3P_{22} + 15P_{23} + 16P_{31} + 18P_{32} + 3P_{33} + 5P_{41} + 2P_{42} + 11P_{43}.$$

The Linear Program: Constraints

All the production of Mine 1 should be shipped:

$$\underbrace{12}_{\text{Coal produced by Mine 1}} = \underbrace{P_{11} + P_{12} + P_{13}}_{\text{Coal shipped from Mine 1}}$$

And 3 similar constraints for the 3 remaining mines.

The Linear Program: Constraints

All the production of Mine 1 should be shipped:

$$\underbrace{12}_{\text{Coal produced by Mine 1}} = \underbrace{P_{11} + P_{12} + P_{13}}_{\text{Coal shipped from Mine 1}}$$

And 3 similar constraints for the 3 remaining mines.

Factory 1 must receive its coal:

$$\underbrace{10}_{\text{Coal used by Factory 1}} = \underbrace{P_{11} + P_{21} + P_{31} + P_{41}}_{\text{Coal shipped to Factory 1}}$$

And 2 similar constraints for the two other factories.

The Linear Program: final result

Minimizing:

$$7P_{11} + 5P_{12} + 8P_{13} + P_{21} + 3P_{22} + 15P_{23} \\ + 16P_{31} + 18P_{32} + 3P_{33} + 5P_{41} + 2P_{42} + 11P_{43}$$

Under the constraints

Mines

$$\begin{array}{cccccc} P_{11} & +P_{12} & +P_{13} & = & 12 \\ P_{21} & +P_{22} & +P_{23} & = & 8 \\ P_{31} & +P_{32} & +P_{33} & = & 3 \\ P_{41} & +P_{42} & +P_{43} & = & 7 \end{array}$$

Factories

$$\begin{array}{cccccc} P_{11} & +P_{21} & +P_{31} & +P_{41} & = & 10 \\ P_{12} & +P_{22} & +P_{32} & +P_{42} & = & 5 \\ P_{13} & +P_{23} & +P_{33} & +P_{43} & = & 15 \end{array}$$

And the non negativity constraints $P_{ij} \geqslant 0$ for $i \in \{1,2,3,4\}$ and $j \in \{1,2,3\}$.

The optimal transport: general case 1

You have n initial locations indexed by i and m final locations indexed by j. You have a number a_i produced at location i, and a number b_j consumed at location j. The cost of transporting one unit of mass from i to j is C_{ij}.

The optimal transport: general case 1

You have n initial locations indexed by i and m final locations indexed by j. You have a number a_i produced at location i, and a number b_j consumed at location j. The cost of transporting one unit of mass from i to j is C_{ij}.

We assume balance of mass condition: $\sum_{i=1}^{n} a_i = \sum_{j=1}^{m} b_j$.

The optimal transport: general case 1

You have n initial locations indexed by i and m final locations indexed by j. You have a number a_i produced at location i, and a number b_j consumed at location j. The cost of transporting one unit of mass from i to j is C_{ij}.

We assume balance of mass condition: $\sum_{i=1}^{n} a_i = \sum_{j=1}^{m} b_j$.

The decision variables are P_{ij} the amount of mass transported from i to j.

The optimal transport problem: general case 2

Optimal transport problem

We want to find the cheapest way to transport, that is we want to minimize

$$\sum_{i=1}^{n}\sum_{j=1}^{m} P_{ij}C_{ij}$$

subject to

$$\begin{array}{rcll} \sum_{j=1}^{m} P_{ij} & = & a_i & \forall i \in \{1, 2, \ldots, n\} \\ \sum_{i=1}^{n} P_{ij} & = & b_j & \forall j \in \{1, 2, \ldots, m\} \end{array}$$

and the non negativity constraints $P_{ij} \geqslant 0$ for all i, j.

The optimal transport problem: general case 2

Optimal transport problem

We want to find the cheapest way to transport, that is we want to minimize

$$\sum_{i=1}^{n} \sum_{j=1}^{m} P_{ij} C_{ij}$$

subject to

$$\begin{array}{rcll} \sum_{j=1}^{m} P_{ij} & = & a_i & \forall i \in \{1, 2, \ldots, n\} \\ \sum_{i=1}^{n} P_{ij} & = & b_j & \forall j \in \{1, 2, \ldots, m\} \end{array}$$

and the non negativity constraints $P_{ij} \geqslant 0$ for all i, j.

There are $n \times m$ decision variables and $n + m$ equality constraints. This is a problem in standard equality form.

Note that the constraints are about the sums along rows and columns of the matrix P.

The dual problem

Let u_i be the dual variable associated to the constraint that Mine i ships all of its coal. Let v_j be the dual variable associated to the constraint that Factory j gets all of its coal.

The dual problem

Let u_i be the dual variable associated to the constraint that Mine i ships all of its coal. Let v_j be the dual variable associated to the constraint that Factory j gets all of its coal.

The dual is the maximization of

$$\sum_{i=1}^{n} a_i u_i + \sum_{j=1}^{m} b_j v_j$$

subject to

$$\underbrace{u_i + v_j \leqslant C_{ij}}_{\text{The company is cheaper than C}} \qquad \forall (i,j) \in \{1,2,\ldots,n\} \times \{1,2,\ldots,m\}$$

The dual problem

Let u_i be the dual variable associated to the constraint that Mine i ships all of its coal. Let v_j be the dual variable associated to the constraint that Factory j gets all of its coal.

The dual is the maximization of

$$\sum_{i=1}^{n} a_i u_i + \sum_{j=1}^{m} b_j v_j = \text{[Profit of the company]}$$

subject to

$$\underbrace{u_i + v_j \leqslant C_{ij}}_{\text{The company is cheaper than } C} \qquad \forall (i,j) \in \{1,2,\ldots,n\} \times \{1,2,\ldots,m\}.$$

This is the LP solved by a transport company which charges u_i to load one unit of mass at location i and charges v_j to unload one unit of mass at location j.

Historical time frame

The optimal transport LP was proposed by Kantorovich in 1942 with an infinite number of decision variables (sums are replaced by integrals). He also introduced the dual problem. It was not known only a decade later in the West.

Historical time frame

The optimal transport LP was proposed by Kantorovich in 1942 with an infinite number of decision variables (sums are replaced by integrals). He also introduced the dual problem. It was not known only a decade later in the West.

If you ask the question: “when was duality theory invented in Linear Programming?”, you have different answers.

Historical time frame

The optimal transport LP was proposed by Kantorovich in 1942 with an infinite number of decision variables (sums are replaced by integrals). He also introduced the dual problem. It was not known only a decade later in the West.

If you ask the question: "when was duality theory invented in Linear Programming?", you have different answers.

- 1928. MinMax theorem by Von Neumann.
- 1942. Duality in optimal transport by Kantorovich. Statement of a LP with an infinite number of variables.
- 1947. Simplex algorithm, together with duality theory, by Dantzig.

2. A special case: Birkhoff Von Neumann theorem

Assignment problem

We consider an *assignment problem*, that is an optimal transport problem where:

- There is the same number of initial and final locations, $n = m$.

Assignment problem

We consider an *assignment problem*, that is an optimal transport problem where:

- There is the same number of initial and final locations, $n = m$.
- The mass to be shipped from or to a location does not depend on the location, that is $a_i = b_j = 1$ for all i, j. By convention we set $a_i = b_j = 1$.

Assignment problem

We consider an *assignment problem*, that is an optimal transport problem where:

- There is the same number of initial and final locations, $n = m$.
- The mass to be shipped from or to a location does not depend on the location, that is $a_i = b_j = 1$ for all i, j. By convention we set $a_i = b_j = 1$.

In short, we take n an integer, $C \in \mathbb{R}^{n \times n}$ is the matrix with the transportation cost between locations and we look at the minimization of

$$\sum_{i=1}^{n} \sum_{j=1}^{n} P_{ij} C_{ij}$$

subject to

$$\begin{array}{rcll} \sum_{j=1}^{n} P_{ij} & = & 1 & \forall i \in \{1, 2, \dots, n\} \\ \sum_{i=1}^{n} P_{ij} & = & 1 & \forall j \in \{1, 2, \dots, n\} \end{array}$$

and the non negativity constraints $P_{ij} \geqslant 0$ for all i, j.

Permutations

It makes sense to assign each initial location to a single final location.

Permutations

It makes sense to assign each initial location to a single final location.

Permutation

A permutation of $\{1, 2, \ldots, n\}$ is a one to one function of $\{1, 2, \ldots, n\}$ into $\{1, 2, \ldots, n\}$.

The set of permutation of $\{1, 2, \ldots, n\}$ is denoted by $\mathcal{S}_n$, there are $n! = n \times (n-1) \times \ldots \times 2 \times 1$ of them.

Permutations

It makes sense to assign each initial location to a single final location.

Permutation

A permutation of $\{1, 2, \ldots, n\}$ is a one to one function of $\{1, 2, \ldots, n\}$ into $\{1, 2, \ldots, n\}$.

The set of permutation of $\{1, 2, \ldots, n\}$ is denoted by $\mathcal{S}_n$, there are $n! = n \times (n-1) \times \ldots \times 2 \times 1$ of them.

If you have a permutation $\sigma \in \mathcal{S}_n$, you can say that the mass at location i is transported only to location $\sigma(i)$. This corresponds to the following feasible P:

$$P_{ij} = \begin{cases} 1 & \text{if } j = \sigma(i) \\ 0 & \text{otherwise.} \end{cases}$$

Main result

Birkhoff Von Neumann theorem

Take an assignment problem. Then a feasible solution P is a basic feasible solution (it can obtained by setting non basic variables to 0 in a feasible dictionary) if and only if there exists a permutation $\sigma \in \mathcal{S}_n$ such that

$$P_{ij} = \begin{cases} 1 & \text{if } j = \sigma(i) \\ 0 & \text{otherwise.} \end{cases} \qquad \forall (i,j) \in \{1, 2, \ldots, n\}^2$$

Main result

Birkhoff Von Neumann theorem

Take an assignment problem. Then a feasible solution P is a basic feasible solution (it can obtained by setting non basic variables to 0 in a feasible dictionary) if and only if there exists a permutation $\sigma \in \mathcal{S}_n$ such that

$$P_{ij} = \begin{cases} 1 & \text{if } j = \sigma(i) \\ 0 & \text{otherwise.} \end{cases} \qquad \forall (i,j) \in \{1, 2, \ldots, n\}^2$$

This result makes the link between combinatorics (permutations) and linear programming. Indeed, the combinatorial problem

$$\min_{\sigma \in \mathcal{S}_n} \sum_{i=1}^{n} C_{i,\sigma(i)}$$

can be solved by finding the associated assignment problem which is a linear program.

Number of dictionaries

In particular, the number of dictionaries of the assignment problem is $n!$.

Number of dictionaries

In particular, the number of dictionaries of the assignment problem is $n!$.

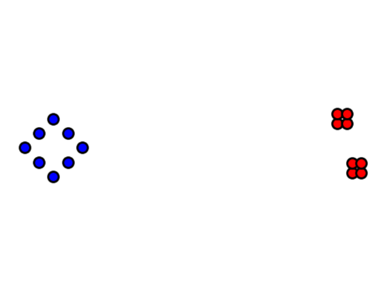
If for instance $n = 20$, so 400 decision variables and 40 constraints, the number of dictionaries is

$$20! \simeq 2.43 \times 10^{18}.$$

Hopefully the simplex algorithm (in this case the “network simplex” which exploits the structure of the problem) does not visit all dictionaries.

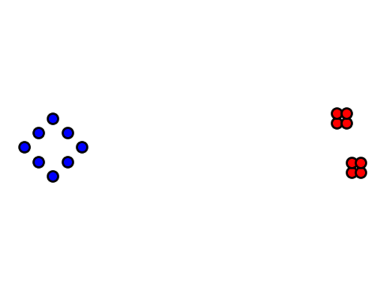
3. Link with geometry: interpolating "shapes"

Geometric embedding



Initial locations and final locations, here $n = m = 8$: this is an assignment problem.

Geometric embedding



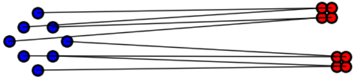
Initial locations and final locations, here $n = m = 8$: this is an assignment problem.

Let x_i be the location of the i-th point and y_j being the one of the j-th point. The transportation cost between i and j is

$$C_{ij} = \|x_i - y_j\|^2.$$

Other functions of the distance work but the quadratic case gives nice movies.

Geometric embedding



Initial locations and final locations, here $n = m = 8$: this is an assignment problem.

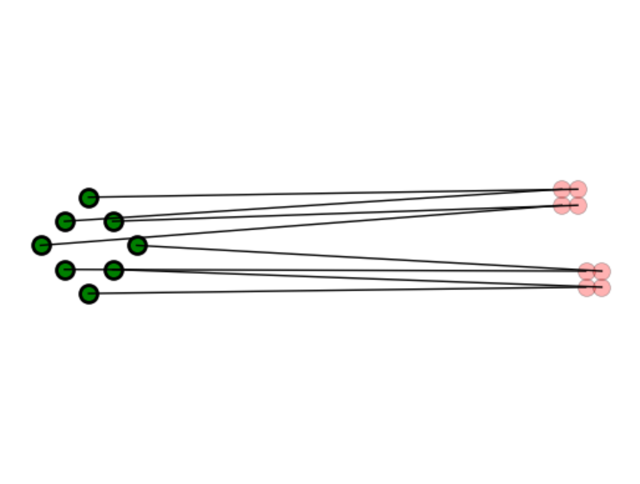
Let x_i be the location of the i-th point and y_j being the one of the j-th point. The transportation cost between i and j is

$$C_{ij} = \|x_i - y_j\|^2.$$

Other functions of the distance work but the quadratic case gives nice movies.

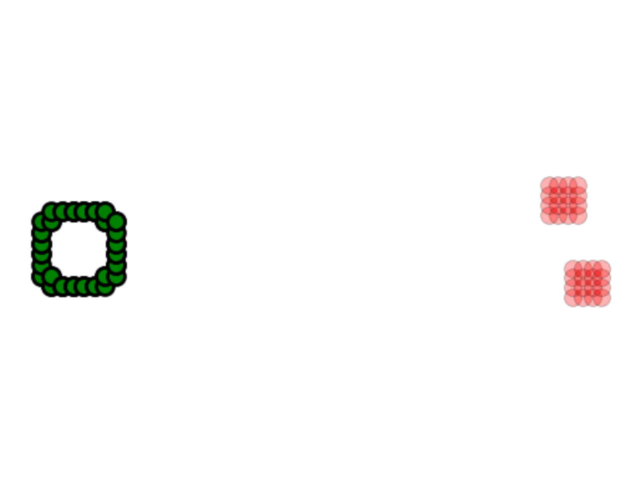
Optimal P: by Birkhoff Von Neumann theorem each point is transported onto a single one: a black line means that two points are connected.

Interpolation



If $\color{blue}{x_i}$ is transported onto $\color{red}{y_j}$, we have a *point* which travels at a constant speed.

More points



$n = 28, m = 32$ and I have removed the black lines.

Even more points ($n = 312$, $m = 394$), shapes are closer, and rendering.

Next time

Numerical questions: how to solve (approximately) large scales LP coming from the optimal transport problem?

Next time

Numerical questions: how to solve (approximately) large scales LP coming from the optimal transport problem?

Applications:

- More on shape interpolation,
- color transfer,
- fluid mechanics,
- etc.