Lecture 33: Optimal Transport. Duality

Lec 32 Optimal transport problems

- Transport plans
- Optimal transport problems

Lec 33

- One more concrete example. For distance squared cost.
- **Duality of optimal transport**

Example (distance squared cost example: two points to two points in $\mathbb{R}^1$.)

- $x_1 = 0, x_2 = 1, y_1 = 2, y_2 = 3 \in \mathbb{R}$.
- $\mu = (1/2, 1/2)$,
  mass $1/2$ at x_1 and mass $1/2$ at x_2.
- $\nu = (1/2, 1/2)$,
  mass $1/2$ at y_1 and mass $1/2$ at y_2.
- $c_{ij} = |x_i - y_j|^2$.

$$\min_{\pi \in \Pi(\mu,\nu)} \sum_{ij} c_{ij}\pi_{ij}.$$

What are the optimal solutions?

Distance squared cost example: two points to two points in $\mathbb{R}^1$

- $$[c_{ij}] = \begin{bmatrix} 2^2 & 3^2 \\ 1^2 & 2^2 \end{bmatrix}, \quad [\pi_{ij}] = \begin{bmatrix} \pi_{11} & \pi_{12} \\ \pi_{21} & \pi_{22} \end{bmatrix}$$
- $$\sum_{ij} c_{ij}\pi_{ij} = 2^2\pi_{11} + 3^2\pi_{12} + \pi_{21} + 2^2\pi_{22}$$
- Thus the LP problem is

$$\begin{array}{ll} \text{Minimize} & 2^2\pi_{11} + 3^2\pi_{12} + \pi_{21} + 2^2\pi_{22} \\ \text{subject to} & \pi_{11} + \pi_{12} = 1/2 \\ & \pi_{21} + \pi_{22} = 1/2 \\ & \pi_{11} + \pi_{21} = 1/2 \\ & \pi_{12} + \pi_{22} = 1/2 \\ & \pi_{11}, \pi_{12}, \pi_{21}, \pi_{22} \geq 0 \end{array}$$

Distace squared cost example: two points to two points in $\mathbb{R}^1$

$$
\begin{aligned}
\text{Minimize} \quad & 2^2\pi_{11} + 3^2\pi_{12} + \pi_{21} + 2^2\pi_{22} \\
\text{subject to} \quad & \pi_{11} + \pi_{12} = 1/2 \\
& \pi_{21} + \pi_{22} = 1/2 \\
& \pi_{11} + \pi_{21} = 1/2 \\
& \pi_{12} + \pi_{22} = 1/2 \\
& \pi_{11}, \pi_{12}, \pi_{21}, \pi_{22} \geq 0
\end{aligned}
$$

is equivalent to is equivalent to

$$
\begin{aligned}
\text{Minimize} \quad & 2^2\pi_{11} + 3^2(1/2 - \pi_{11}) + (1/2 - \pi_{11}) + 2^2\pi_{22} \\
\text{subject to} \quad & \pi_{11} \leq 1/2 \\
& 1/2 - \pi_{11} + \pi_{22} = 1/2 \\
& \pi_{11} \leq 1/2 \\
& 1/2 - \pi_{11} + \pi_{22} = 1/2 \\
& \pi_{11}, \pi_{22} \geq 0
\end{aligned}
$$

Distance squared cost: two points to two points in $\mathbb{R}^1$

$$
\begin{array}{ll}
\text{Minimize} & 2^2\pi_{11} + 3^2(1/2 - \pi_{11}) + (1/2 - \pi_{11}) + 2^2\pi_{22} \\
\text{subject to} & \pi_{11} \le 1/2 \\
& 1/2 - \pi_{11} + \pi_{22} = 1/2 \\
& \pi_{11} \le 1/2 \\
& 1/2 - \pi_{11} + \pi_{22} = 1/2 \\
& \pi_{11}, \pi_{22} \ge 0
\end{array}
$$

is equivalent to

$$
\begin{array}{ll}
\text{Minimize} & 2^2\pi_{11} + 3^2(1/2 - \pi_{11}) + (1/2 - \pi_{11}) + 2^2\pi_{11} \\
\text{subject to} & \pi_{11} \le 1/2 \\
& \pi_{11} \ge 0
\end{array}
$$

which is equivalent to

$$
\begin{array}{ll}
\text{Minimize} & 5 - 2\pi_{11} \\
\text{subject to} & 0 \le \pi_{11} \le 1/2
\end{array}
$$

Distance squared cost: two points to two points in $\mathbb{R}^1$

$$\begin{array}{ll} \text{Minimize} & 5 - 2\pi_{11} \\ \text{subject to} & 0 \leq \pi_{11} \leq 1/2 \end{array}$$

What does this mean?

- Answer: Optimal solution is $\pi_{11} = 1/2$.
  So, from the constraints

$$\pi_{11} + \pi_{12} = 1/2$$
$$\pi_{21} + \pi_{22} = 1/2$$
$$\pi_{11} + \pi_{21} = 1/2$$
$$\pi_{12} + \pi_{22} = 1/2$$

  we get
  $\pi_{12} = 0, \pi_{21} = 0, \pi_{22} = 1/2$.

Duality

Primal:

$$\begin{array}{ll} \text{Minimize} & \sum_{i=1}^{m} \sum_{j=1}^{n} c_{ij}\pi_{ij} \\ \text{subject to} & \sum_{j=1}^{n} \pi_{ij} = \mu_i \\ & \sum_{i=1}^{m} \pi_{ij} = \nu_j \\ & \pi_{ij} \geq 0 \\ & \forall i = 1, ..., m, \forall j = 1, ..., n \end{array}$$

Dual:

$$\begin{array}{ll} \text{Maximize} & \sum_{i=1}^{m} \phi_i \mu_i + \sum_{j=1}^{n} \psi_j \nu_j \\ \text{subject to} & \phi_i + \psi_j \leq c_{ij} \\ & \forall i = 1, ..., m, \forall j = 1, ..., n \end{array}$$

We will explain:

Primal:

$$\begin{array}{rl} \text{Minmize} & \vec{c}^T\vec{x} \\ \text{subject to} & A\vec{x} = \vec{b} \\ & \vec{x} \geq \vec{0} \end{array}$$

Dual:

$$\begin{array}{rl} \text{Maximize} & \vec{b}^T\vec{y} \\ \text{subject to} & A^T\vec{y} \leq \vec{c} \end{array}$$

Explanation of duality

Primal:

$$\begin{aligned} \text{Minmize} \quad & \vec{c}^T \vec{x} \\ \text{subject to} \quad & A\vec{x} = \vec{b} \\ & \vec{x} \geq \vec{0} \end{aligned}$$

In standard (minimization) form

$$\begin{aligned} \text{Minmize} \quad & \vec{c}^T \vec{x} \\ \text{subject to} \quad & A\vec{x} \geq \vec{b} \\ & -A\vec{x} \geq -\vec{b} \\ & \vec{x} \geq \vec{0} \end{aligned}$$

Explanation of duality

$$\begin{array}{ll} \text{Minmize} & \vec{c}^T\vec{x} \\ \text{subject to} & A\vec{x} \geq \vec{b} \quad \text{(corresponds to } \vec{y}^+\text{)} \\ & -A\vec{x} \geq -\vec{b} \quad \text{(corresponds to } \vec{y}^-\text{)} \\ & \vec{x} \geq \vec{0} \end{array}$$

Dual variables are $\vec{y}^+, \vec{y}^- \geq \vec{0}$.

Primal

$$\begin{array}{ll} \text{Min.} & \vec{c}^T\vec{x} \\ \text{subj.} & A\vec{x} \geq \vec{b} \\ & -A\vec{x} \geq -\vec{b} \\ & \vec{x} \geq \vec{0} \end{array}$$

Dual

$$\begin{array}{ll} \text{Max.} & \vec{b}^T\vec{y}^+ - \vec{b}^T\vec{y}^- \\ \text{subj.} & A^T\vec{y}^+ - A^T\vec{y}^- \leq \vec{c} \\ & \vec{y}^+, \vec{y}^- \geq \vec{0} \end{array}$$

Explanation of duality

Primal:

$$\begin{array}{ll} \text{Minmize} & \vec{c}^T\vec{x} \\ \text{subject to} & A\vec{x} = \vec{b} \\ & \vec{x} \geq \vec{0} \end{array}$$

In standard (minimization) form:

$$\begin{array}{ll} \text{Min.} & \vec{c}^T\vec{x} \\ \text{subj.} & A\vec{x} \geq \vec{b} \\ & -A\vec{x} \geq -\vec{b} \\ & \vec{x} \geq \vec{0} \end{array}$$

Dual

$$\begin{array}{ll} \text{Max.} & \vec{b}^T\vec{y}^+ - \vec{b}^T\vec{y}^- \\ \text{subj.} & A^T\vec{y}^+ - A^T\vec{y}^- \leq \vec{c} \\ & \vec{y}^+, \vec{y}^- \geq \vec{0} \end{array}$$

By letting $\vec{y} = \vec{y}^+ - \vec{y}^-$, we have

Dual :
$$\begin{array}{ll} \text{Max.} & \vec{b}^T\vec{y} \\ \text{subj.} & A^T\vec{y} \leq \vec{c} \end{array}$$

Explanation of duality

So, we can (roughly) see

Primal:

$$\begin{array}{ll} \text{Min.} & \vec{c}^T\vec{x} \\ \text{subj.} & A\vec{x} = \vec{b} \\ & \vec{x} \geq \vec{0} \end{array}$$

Dual:

$$\begin{array}{ll} \text{Max.} & \vec{b}^T\vec{y} \\ \text{subj.} & A^T\vec{y} \leq \vec{c} \end{array}$$

Primal:

$$\begin{array}{ll} \text{Min.} & \sum_{i=1}^m \sum_{j=1}^n c_{ij}\pi_{ij} \\ \text{subj.} & \sum_{j=1}^n \pi_{ij} = \mu_i \\ & \sum_{i=1}^m \pi_{ij} = \nu_j \\ & \pi_{ij} \geq 0 \\ & \forall i = 1, ..., m, \\ & \forall j = 1, ..., n \end{array}$$

Dual:

$$\begin{array}{ll} \text{Max} & \sum_{i=1}^m \phi_i\mu_i + \sum_{j=1}^n \psi_j\nu_j \\ \text{subj.} & \phi_i + \psi_j \leq c_{ij} \\ & \forall i = 1, ..., m, \\ & \forall j = 1, ..., n \end{array}$$

More precise explanation in this direction is possible. Just a bit tedious and complicated.

Another explanation of duality of OT

(Similar to Lecture 12 when we used the concept of Lagrange multiplier.)

- $$\begin{array}{ll} \text{Minimize} & \sum_{i=1}^{m} \sum_{j=1}^{n} c_{ij}\pi_{ij} \\ \text{subject to} & \sum_{j=1}^{n} \pi_{ij} = \mu_i \\ & \sum_{i=1}^{m} \pi_{ij} = \nu_j \\ & \pi_{ij} \geq 0 \end{array}$$

  is equivalent to

- $$\min_{\pi \geq \mathbf{0}} \max_{\phi,\psi} \quad \sum_{i=1}^{m} \sum_{j=1}^{n} c_{ij}\pi_{ij} + \sum_{i} \phi_i \left[\mu_i - \sum_{j=1}^{n} \pi_{ij} \right] + \sum_{j} \psi_i \left[\nu_j - \sum_{i=1}^{m} \pi_{ij} \right]$$

Another explanation of duality: continued

- $$\begin{aligned}\min_{\pi\geq\mathbf{0}}\max_{\phi,\psi} \quad & \sum_{i=1}^{m}\sum_{j=1}^{n} c_{ij}\pi_{ij} + \sum_{i}\phi_i\left[\mu_i - \sum_{j=1}^{n}\pi_{ij}\right] \\ & + \sum_{j}\psi_i\left[\nu_j - \sum_{i=1}^{m}\pi_{ij}\right]\end{aligned}$$

- is equivalent to

$$\min_{\pi\geq\mathbf{0}}\max_{\phi,\psi}\left[\sum_{i}\phi_i\mu_i + \sum_{j}\psi_i\nu_j + \sum_{i=1}^{m}\sum_{j=1}^{n}\left[c_{ij} - \phi_i - \psi_j\right]\pi_{ij}\right]$$

- In fact, in this case one can reorder $\min\max$ to $\max\min$. (This is a bit of cheating for we know we can do this becasue the duality holds.)

Another explanation of duality: continued

- $$\max_{\phi,\psi} \min_{\pi \geq \mathbf{0}} \left[\sum_i \phi_i \mu_i + \sum_j \psi_j \nu_j + \sum_{i=1}^{m} \sum_{j=1}^{n} \left[c_{ij} - \phi_i - \psi_j \right] \pi_{ij} \right]$$
- If $c_{i'j'} - \phi_{i'} - \psi_{j'} < 0$ for some i', j', one can choose
  $$\pi_{ij} = \begin{cases} \lambda & \text{for } i = i', j = j' , \\ 0 & \text{otherwise.} \end{cases}$$
  and such that the **minimum** $\to -\infty$ as $\lambda \to \infty$.
- For those ϕ, ψ with $c_{ij} - \phi_i - \psi_j \geq 0$, the minimum occurs when $\pi = \mathbf{0}$.
- Thus, the problem can be reduced to
  $$\max_{\phi_i + \psi_j \leq c_{ij}} \left[\sum_i \phi_i \mu_i + \sum_j \psi_j \nu_j \right] .$$

c-transforms

c-transform.

$$\phi^c(y) = \min_x [c(x, y) - \phi(x)]$$
$$\psi^c(x) = \min_y [c(x, y) - \psi(y)]$$

Or the discrete version:

$$\phi_j^c = \min_i [c_{ij} - \phi_i] \qquad \& \qquad \psi_i^c = \min_j [c_{ij} - \psi_j]$$

Note

- $\phi(x) + \phi^c(y) \leq c(x, y) \quad \forall x, y$
- $\psi^c(x) + \psi(y) \leq c(x, y) \quad \forall x, y$

Special case: $c(\vec{x}, \vec{y}) = \vec{x}^T \vec{y}$ for $\vec{x}, \vec{y} \in \mathbb{R}^d$

$$\phi^*(\vec{y}) = \min_{\vec{x} \in \mathbb{R}^d} \left[\vec{x}^T \vec{y} - \phi(\vec{x}) \right]$$

[This is a version of the so-called Legendre transform in convex analysis.]
The function ϕ^* is a concave function.
See board for a picture.

Example

[Your exercises] For $x, y \in \mathbb{R}$.

- If $\phi(x) = 0$, then $\phi^*(y) = -\infty$ if $y \neq 0$, and $\phi^*(0) = 0$.
- If $\phi(x) = x$, then $\phi^*(y) = -\infty$ if $y \neq 1$, and $\phi^*(1) = 0$.
- If $\phi(x) = -\frac{1}{2}x^2$, then $\phi^*(y) = -\frac{1}{2}y^2$.