OPTIMAL TRANSPORT PROBLEMS

JIAFENG CHEN AND FRANCISCO RIVERA

1. Theoretical foundations

Let $\mathbf{a} \in \mathbb{R}^m_{\geq 0}$ and $\mathbf{b} \in \mathbb{R}^n_{\geq 0}$ be two vectors of nonnegative real numbers such that $\mathbf{a}^T \mathbb{1} = \mathbf{b}^T \mathbb{1} = 1$, where a_i and b_j represents the amount of "mass" at different positions. We are interested in a matrix $\mathbf{P} \in \mathbb{R}^{m\times n}_{\geq 0}$ where P_{ij} prescribes the amount of mass flowing from a_i to b_j. Naturally, we must have $\sum_j P_{ij} = a_i$ and $\sum_i P_{ij} = b_j$, so that mass is preserved. Let $\mathbf{C} \in \mathbb{R}^{m\times n}$ be a *cost matrix*, where C_{ij} represents the cost of transporting one unit of a_i to b_j. The *Kantorovich optimal transport problem* is the following linear program:

$$\begin{aligned} \min_{\mathbf{P}\in\mathbb{R}^{m\times n}_{\geq 0}} & \sum_{i,j} P_{ij} C_{ij} \\ \text{subjected to } & \mathbf{P}\mathbb{1} = \mathbf{a} \\ & \mathbf{P}^T \mathbb{1} = \mathbf{b} \end{aligned}$$

Remark 1 (A probabilistic view). Let X, Y be discrete random variables with support $\mathbf{x} \in \mathcal{X}^m$ and $\mathbf{y} \in \mathcal{Y}^n$, such that their marginal distributions are prescribed by $P(X = x_i) = a_i$ and $P(Y = y_j) = b_j$. Let $P_{ij} = P(X = x_i, Y = y_j)$, making $\mathbf{P} = (P_{ij})_{ij}$ a joint distribution. Let $C : \mathcal{X} \times \mathcal{Y} \to \mathbb{R}$ be a cost function. Then the Kantorovich problem is equivalent to the following:

$$\min_{\mathbf{P}} \mathbb{E}_{(X,Y)\sim\mathbf{P}} \left[C(X, Y)\right],$$

where the expectation is taken over valid joint distributions of (X, Y) respecting the marginal distributions. Note that when $\mathcal{X} = \mathcal{Y} = \mathbb{R}^d$ and $C(x, y) = \|x - y\|^p$, then the optimal value of the above problem for X, Y is the p-Wasserstein distance between the marginal distributions μ_X, μ_Y (up to a power of $1/p$), for discrete, d-dimensional distributions with finite support.

Date: March 24, 2018.

It can be shown that the dual problem of the above LP takes the following form:

$$\max_{(\mathbf{f},\mathbf{g})\in\mathbb{R}^m\times\mathbb{R}^n} \sum_i a_i f_i + \sum_j b_j g_j$$
$$\text{subjected to } f_i + g_j \leq C_{ij} \text{ for all } i, j.$$

The dual problem has a neat economic intuition. Suppose Kevin needs to transport $\mathbf{a}$ to $\mathbf{b}$ but does not understand how. Francisco, his profit-maximizing colleague, offers him a deal, where Kevin will pay f_i for Francisco to pick up a unit of mass at a_i and pay g_j for Francisco to dropoff a unit of mass at b_j. At minimum, Kevin knows that if any $f_i + g_j > C_{ij}$, then Francisco is ripping him off, as transporting one unit from a_i to b_j costs exactly C_{ij}. If, on the other hand, $f_i + g_j \leq C_{ij}$ is satisfied, then given any $\mathbf{P}$, we have

$$\text{(Weak duality)} \quad \sum_{i,j} C_{ij}P_{ij} \geq \sum_{i,j} f_i P_{ij} + \sum_{i,j} g_j P_{ij} = \sum_i f_i a_i + \sum_j g_j b_j,$$

which means that Kevin cannot lose by taking Francisco's deal. Strong duality, which indeed holds, here imply that Francisco's optimal profit is zero, and the trade is fair.

Given optimal solutions to the primal and dual, $\mathbf{P}$, $(\mathbf{f}, \mathbf{g})$, respectively, *complementary slackness* implies that

$$P_{ij}(C_{ij} - f_i - g_j) = 0$$

holds, which means that $\mathbf{P}_{ij} \neq 0 \implies C_{ij} = f_i - g_j$ and vice versa.