Тема 1

Транспортные задачи. Случай конечных пространств.

Мы будем изучать задачи оптимальной транспортировки некоторым образом распределенной массы из заданного начального состояния в заданное конечное состояние. Имеется много различных постановок этой задачи. Мы приведем лишь две из них — задачу Монжа и задачу Канторовича, но более подробно будем разбирать лишь последнюю. Вот некоторая литература, посвященная этой проблеме: [8]–[12].

1.1 Постановки и формализации транспортных задач: случай конечных пространств

Одна из наглядных интерпретаций этих задач такова. Представим себе, что мы занимаемся перемещением добытых полезных ископаемых на склады некоторых предприятий. Пусть нам известно распределение этих ископаемых, размещение складов предприятий, а также стоимость доставки единицы добычи из одного пункта в другой. Требуется среди всех возможных перемещений выбрать наиболее дешевые.

Обозначим множество полезных ископаемых через X, а множество складов предприятий через Y. Мы разобьем всю область залегания полезных ископаемых на конечное число частей, множество которых обозначим через X. Кроме того, будем (вполне реалистично) предполагать, что множество складов Y также конечно.

Далее, для каждого $x \in X$ через $\lambda(x)$ обозначим отношение объема части x ко всему объему добытых ископаемых. В силу определения, числа $\lambda(x)$ неотрицательны и $\sum_{x\in X} \lambda(x) = 1$. Тем самым, λ можно рассматривать как вероятностную меру на X. Аналогично, для каждого $y \in Y$ через $\mu(y)$ обозначим отношение объема склада y к суммарному объему всех складов. Тогда μ также можно рассматривать как вероятностную меру на Y.

Пусть, кроме того, стоимость перемещения каждой единицы объема полезных ископаемых из части $x \in X$ на склад $y \in Y$ зависит лишь от x и y. Тогда эта стоимость задается некоторой неотрицательной функцией $c(x, y)$.

Наша задача — таким образом переместить добытые ископаемые на склады предприятий, чтобы суммарная стоимость перемещения была наименьшей возможной.

Приведем несколько вариантов постановки задачи оптимальной транспортировки.

1.1.1 Транспортная задача Монжа

Предположим, что каждая часть $x \in X$ добытых полезных ископаемых перемещается целиком на некоторый склад $y \in Y$, а также что каждый склад $y \in Y$ полностью заполняется перемещенными ископаемыми. В частности, суммарные объемы всех добытых ископаемых и всех имеющихся складов одинаковы. Тогда такое перемещение задается отображением $T\colon X \to Y$, для которого выполняется

$$\mu(y) = \sum_{x \in T^{-1}(y)} \lambda(x)$$

при всех $y \in Y$. Эти соотношения называются *граничными условиями*, а каждое такое отображение T — *транспортировкой типа* (λ, μ). Множество транспортировок типа (λ, μ) будем обозначать через $\mathcal{T}(\lambda, \mu)$.

Вычислим стоимость транспортировки $T \in \mathcal{T}(\lambda, \mu)$ используя функцию стоимости $c(x, y)$. Так как часть x полезных ископаемых перемещается на склад $T(x)$, стоимость такого перемещения равна $c\bigl(x, T(x)\bigr)\,\lambda(x)$ и, значит, полная стоимость равна

$$M(T) := \sum_{x \in X} c\bigl(x, T(x)\bigr)\,\lambda(x).$$

Определенное только что отображение $M\colon \mathcal{T}(\lambda, \mu) \to \mathbb{R}$ называется *функционалом Монжа*. Таким образом, *транспортная задача Монжа* состоит в минимизации функционала Монжа $M(T)$ по всем транспортировкам $T \in \mathcal{T}(\lambda, \mu)$ и поиске *оптимальной транспортировки* T, дающей этот минимум.

1.1.2 Транспортная задача Канторовича

Теперь мы не станем предполагать, что каждая часть $x \in X$ полезных ископаемых перемещается ровно на один склад, а разрешим распределять эту часть x по нескольким складам. Такое распределение будем задавать функцией $\pi\colon X \times Y \to \mathbb{R}$, причем величина $\pi(x, y)$ может быть интерпретирована как отношение объема ископаемых, перемещенных из части x на склад y, ко всему объему ископаемых. Будем предполагать, что отображение π удовлетворяет следующим *граничным условиям*:

(1) каждая часть x будет полностью перемещена:

$$\lambda(x) = \sum_{y \in Y} \pi(x, y) \quad \text{при всех } x \in X;$$

(2) каждый склад $y \in Y$ будет полностью заполнен:

$$\mu(y) = \sum_{x \in X} \pi(x, y) \quad \text{при всех } y \in Y.$$

Заметим, что если функция π неотрицательна (а это отвечает смыслу задачи), то π — вероятностная мера на $X \times Y$:

$$\sum_{x,y \in X} \pi(x,y) = \sum_{x \in X} \sum_{y \in Y} \pi(x,y) = \sum_{x \in X} \lambda(x) = 1.$$

Каждое неотрицательное π, удовлетворяющее граничным условиям, называется *транспортным планом типа* (λ, μ), а множество всех таких транспортных планов будем обозначать через $\Pi(\lambda, \mu)$.

Вычислим теперь стоимость реализации транспортного плана. Из x в y перемещается $\pi(x,y)$ всего объема, поэтому стоимость этого перемещения равна $c(x,y)\pi(x,y)$, так что стоимость всего перемещения в соответствии с транспортным планом π имеет вид

$$K(\pi) := \sum_{(x,y) \in X \times Y} c(x,y)\pi(x,y).$$

Определенное только что отображение $K\colon \Pi(\lambda,\mu) \to \mathbb{R}$ называется *функционалом Канторовича*. *Транспортная задача Канторовича* состоит минимизации функционала Канторовича $K(\pi)$ по всем транспортным планам $\pi \in \Pi(\lambda, \mu)$ и поиске *оптимального транспортного плана* π, дающего этот минимум.

Замечание 1.1. Для каждой транспортировки $T\colon X \to Y$ типа (λ, μ) из задачи Монжа положим $\pi_T(x,y) = \lambda(x)$, если $x \in T^{-1}(y)$, и 0 в противном случае. Иными словами, $\pi_T\big(x, T(x)\big) = \lambda(x)$, а на остальных парах (x,y) функция π равна 0. Тогда

$$\mu(y) = \sum_{x \in T^{-1}(y)} \lambda(x) = \sum_{x \in X} \pi_T(x,y), \quad \sum_{y \in Y} \pi_T(x,y) = \pi_T\big(x, T(x)\big) = \lambda(x),$$

так что π_T удовлетворяет граничным условиям задачи Канторовича, т.е. $\pi_T \in \Pi(\lambda, \mu)$. При этом,

$$M(T) = \sum_{x \in X} c\big(x, T(x)\big)\lambda(x) = \sum_{x \in X} c\big(x, T(x)\big)\pi_T\big(x, T(x)\big) =$$
$$= \sum_{(x,y) \in X \times Y} c(x,y)\pi_T(x,y) = K(\pi).$$

Таким образом, задача Монжа может рассматриваться как задача Канторовича, в которой минимизация происходит ни по всему пространству $\Pi(\lambda,\mu)$, а по его части, составленной из транспортных планов вида π_T, $T \in \mathcal{T}(\lambda, \mu)$.

Сравнение задач Монжа и Канторовича

Задача Канторовича во многих случаях оказывается проще, чем задача Монжа.

- Для задачи Канторовича с граничными мерами μ и ν всегда существует хотя бы один транспортный план: в качестве него можно взять $\pi(x,y) = \lambda(x)\mu(x)$. Таким образом, $\Pi(\lambda, \mu) \neq \emptyset$ всегда.

- Для задачи Монжа может не существовать ни одной транспортировки.

  Пример 1.2. Обозначим через δ_a *вероятностную меру Дирака*, равную 1 на $a \in X$, и 0 на всех остальных точках из X, и возьмем ее в качестве λ. В качестве μ возьмем любую вероятностную меру на Y, отличную от меры Дирака. Это означает, что для каждого $y \in Y$ выполняется $\mu(y) < 1$ и, значит, существует $y_0 \in Y$ такой, что $0 < \mu(y_0) < 1$.

  Предположим, что существует некоторая транспортировка $T \in \mathcal{T}(\lambda, \mu)$. Тогда

  $$\mu(y_0) = \sum_{x \in T^{-1}(y_0)} \lambda(x) = \sum_{x \in T^{-1}(y_0)} \delta_a(x),$$

  но правая часть этого равенства равна или 0, или 1, что противоречит выбору y_0. Таким образом, $\mathcal{T}(\lambda, \mu) = \emptyset$.

- Задачи Монжа и Канторовича могут иметь разные решения при одних и тех же граничных условиях и функциях стоимости.

  Пример 1.3. Пусть $X = \{x_1, x_2\}$, $Y = \{y_1, y_2\}$. Положим $\lambda = \frac{1}{3}\delta_{x_1} + \frac{2}{3}\delta_{x_2}$, $\mu = \frac{1}{3}\delta_{y_1} + \frac{2}{3}\delta_{y_2}$, $c_{ij} = c(x_i, y_j)$. Тогда единственной транспортировкой в задаче Монжа может быть отображение $T\colon x_i \mapsto y_i$, $i = 1, 2$, и ее стоимость равна $c_{11}/3 + 2c_{22}/3$.

  Рассмотрим теперь задачу Канторовича. Пусть π — транспортный план, в соответствии с которым из x_i в y_j перемещается π_{ij} массы. Тогда стоимость такой транспортировки равна $\sum_{i,j} c_{ij}\pi_{ij}$. Запишем теперь граничные условия. Имеем $\pi_{11}+\pi_{12} = \pi_{11}+\pi_{21} = 1/3$, $\pi_{12}+\pi_{22} = \pi_{21}+\pi_{22} = 2/3$, откуда $\pi_{12} = \pi_{21} = 1/3-\pi_{11}$, $\pi_{22} = 1/3 + \pi_{11}$, $0 \leq \pi_{11} \leq 1/3$. Таким образом, стоимость такого транспортного плана равна

  $$K(\pi) = K(\pi_{11}) = \frac{c_{12} + c_{21} + c_{22}}{3} + \pi_{11}(c_{11} + c_{22} - c_{12} - c_{21}).$$

  Пусть функция стоимости выбрана так, что $c_{11} + c_{22} - c_{12} - c_{21} > 0$, тогда оптимальным транспортным планом будет тот, в котором $\pi_{11} = 0$, но тогда стоимость такого плана равна $(c_{12} + c_{21} + c_{22})/3$. Заметим, что, в силу сделанных предположений, $c_{11}/3 + 2c_{22}/3 \neq (c_{12} + c_{21} + c_{22})/3$, поэтому задачи Монжа и Канторовича в данном случае имеют разные ответы.

В дальнейшем мы будем заниматься важным случаем задачи Канторовича, когда $X = Y$ — метрическое пространство, и в качестве функции стоимости $c(x, y)$ выбрана заданная на X функция расстояния. Начнем со случая конечного X.

1.2 Транспортная задача Канторовича на метрическом пространстве

Пусть $X = \{e_1, \ldots, e_n\}$ — конечное метрическое пространство с метрикой ρ. Обозначим через $V = V(X)$ векторное пространство всевозможных вещественных функций на X. Каждая функция $v\colon X \to \mathbb{R}$ однозначно определяется своими значениями

$v_i = v(e_i)$ на всех точках $e_i \in X$, и ее удобно записывать в виде $\sum_i v_i e_i$. Таком образом, точки $e_i \in X$ можно рассматривать как базисные векторы пространства V, функции из V — как линейные комбинации этих базисных векторов, а числа v_i — как декартовы координаты в V.

Пусть $\Delta = \Delta(X)$ — стандартный $(n-1)$-мерный симплекс с вершинами — базисными векторами e_i. Тогда точки из Δ можно рассматривать как вероятностные меры, определенные на множестве X.

Рассмотрим транспортную задачу Канторовича на пространствах $X = Y$, выбрав в качестве функции стоимости метрику ρ. Итак, для заданных двух вероятностных мер $\lambda, \mu \in \Delta$ требуется построить такую неотрицательную функцию π на $X \times X$, что выполняются следующие условия:

(1) $\sum_{y\in X} \pi(x, y) = \lambda(x)$ при всех $x \in X$;

(2) $\sum_{x\in X} \pi(x, y) = \mu(x)$ при всех $y \in X$;

(3) $K_\rho(\pi) = \sum_{x,y\in X} \rho(x, y)\pi(x, y) \to \min$.

Условия (1) и (2) мы назвали *граничными*, каждое π, удовлетворяющее граничным условиям, — *транспортным планом типа* (λ, μ), а множество всех таких транспортных планов мы обозначили через $\Pi(\lambda, \mu)$. Отметим, что, как уже отмечалось выше, функция π на самом деле является вероятностной мерой на $X \times X$ (это мгновенно вытекает из граничных условий).

Напомним также, что отображение $K_\rho \colon \Pi(\lambda, \mu) \to \mathbb{R}$ мы назвали *функционалом Канторовича*, а транспортный план, дающий минимальное значение функционала K_ρ — *оптимальным транспортным планом.*

Замечание 1.4. Имеется много классических способов решения этой задачи, самым известным является симплекс-метод. Также имеется много различных обобщений этой задачи, часть из которым мы рассмотрим в дальнейших лекциях.

В дальнейшем нам понадобится ряд обозначений, которые мы сейчас введем:

$$\rho_{ij} = \rho(e_i, e_j),\ \pi_{ij} = \pi(e_i, e_j),\ \lambda_i = \lambda(e_i),\ \mu_i = \mu(e_i).$$

Также для каждой матрицы A через A^t будем обозначать транспонированную матрицу A.

Следующая теорема утверждает, что оптимальный транспортный план существует всегда.

Теорема 1.5 (Канторович [8])**.** *Определим на $\Delta \times \Delta$ функцию k_ρ, положив*

$$k_\rho(\lambda, \mu) = \inf\bigl\{K_\rho(\pi) : \pi \in \Pi(\lambda, \mu)\bigr\}.$$

Тогда

(1) *для каждых $\lambda, \mu \in \Delta$ существует оптимальный план $\pi \in \Pi(\lambda, \mu)$, т.е. план, для которого $k_\rho(\lambda, \mu) = K_\rho(\pi)$;*

(2) *функция k_ρ — метрика на Δ, совпадающая с ρ на вершинах симплекса Δ и линейно меняющаяся на ребрах этого симплекса.*

Доказательство. (1) Отметим, что множество $\Pi(\lambda,\mu)$ представляет собой пересечение симплекса $\Delta(X\times X)$ и аффинного подпространства в $V(X\times X)$, заданного граничными условиями (этих условий имеется $2n$ штук). Таким образом, $\Pi(\lambda,\mu)$ — компакт в $V(X\times X)=\mathbb{R}^{n^2}$. Функция K_ρ является ограничением на $\Delta(X\times X)$ соответствующего линейного функционала и, потому, непрерывна, так что она ограничена и принимает свое наименьшее значение.

(2) Покажем, что для каждого λ выполняется $k_\rho(\lambda,\lambda)=0$. Действительно, пусть $\pi_{ij}=\lambda_i\delta_{ij}$, тогда $\pi\in\Pi(\lambda,\lambda)$ и $K_\rho(\pi)=\sum_{i,j}\rho_{ij}\lambda_i\delta_{ij}=0$, откуда $k_\rho(\lambda,\lambda)=0$.

Докажем невырожденность функции k_ρ. Для этого достаточно проверить, что при $\lambda\neq\mu$ выполняется $k_\rho(\lambda,\mu)>0$. Предположим противное, т.е. существует последовательность $\pi^k\in\Pi(\lambda,\mu)$ такая, что $K_\rho(\pi^k)\to0$ при $k\to\infty$. Так как $\Pi(\lambda,\mu)$ компактно, из этой последовательности можно выбрать сходящуюся подпоследовательность, поэтому, без ограничения общности, сразу будем считать, что последовательность π^k сходится. Положим $\pi^k_{ij}=\pi^k(e_i,e_j)$, тогда $\sum_{ij}\rho_{ij}\pi^k_{ij}\to0$. Так как все ρ_{ij} при $i\neq j$ положительны, а все π^k_{pq} неотрицательны, то при всех $i\neq j$ выполняется $\pi^k_{ij}\to0$. С другой стороны, при всех k имеем $\lambda_i=\sum_j\pi^k_{ij}$ и $\mu_i=\sum_j\pi^k_{ji}$, откуда $\lambda_i=\lim_{k\to\infty}\sum_j\pi^k_{ij}=\lim_{k\to\infty}\pi^k_{ii}=\mu_i$ при всех i, что противоречит условию $\lambda\neq\mu$.

Докажем симметричность функции k_ρ. Для каждого $\pi\in\Pi(\lambda,\mu)$ имеем $\pi^t\in\Pi(\mu,\lambda)$, и для каждого $\pi\in\Pi(\mu,\lambda)$ выполняется $\pi^t\in\Pi(\lambda,\mu)$, поэтому если π "пробегает" все $\Pi(\lambda,\mu)$, то π^t "пробегает" все $\Pi(\mu,\lambda)$. С другой стороны, в силу симметричности метрики ρ, имеем

$$K_\rho(\pi)=\sum_{i,j}\rho_{ij}\pi_{ij}=\sum_{i,j}\rho_{ji}\pi^t_{ji}=K_\rho(\pi^t),$$

откуда, с учетом сказанного выше, получаем $k_\rho(\lambda,\mu)=k_\rho(\mu,\lambda)$.

Докажем теперь неравенство треугольника.

Выберем произвольные $\lambda,\mu,\nu\in\Delta$. В силу пункта (1), существуют оптимальные планы $\pi^1\in\Pi(\lambda,\mu)$ и $\pi^2\in\Pi(\mu,\nu)$. Положим $\pi^p_{ij}=\pi^p(e_i,e_j)$ и определим числа η_{ijk} следующим образом:

$$\eta_{ijk}=\begin{cases}\frac{\pi^1_{ij}\pi^2_{jk}}{\mu_j}, \text{ если } \mu_j\neq0,\\ 0 \text{ в противном случае.}\end{cases}$$

Далее, положим $\pi_{ik}=\sum_j\eta_{ijk}$. Так как $\mu_j=\sum_i\pi^1_{ij}=\sum_k\pi^2_{jk}$, и все компоненты матриц π^p неотрицательные, то $\mu_j=0$, если и только если $\pi^1_{ij}=0$ при всех i, а также $\pi^2_{jk}=0$ при всех k. Отсюда вытекает, что $\sum_i\eta_{ijk}=\pi^2_{jk}$ и $\sum_k\eta_{ijk}=\pi^1_{ij}$, откуда $\sum_i\pi_{ik}=\sum_j\pi^2_{jk}=\nu_k$ и $\sum_k\pi_{ik}=\sum_j\pi^1_{ij}=\lambda_i$, поэтому $\pi\in\Pi(\lambda,\nu)$.

Следовательно,

$$k_\rho(\lambda,\nu)\le K_\rho(\pi)=\sum_{i,k}\rho_{ik}\pi_{ik}=\sum_{i,j,k}\rho_{ik}\eta_{ijk}\le\sum_{i,j,k}(\rho_{ij}+\rho_{jk})\eta_{ijk}=$$
$$=\sum_{i,j,k}\rho_{ij}\eta_{ijk}+\sum_{i,j,k}\rho_{jk}\eta_{ijk}=\sum_{i,j}\rho_{ij}\pi^1_{ij}+\sum_{j,k}\rho_{jk}\pi^2_{jk}=k_\rho(\lambda,\mu)+k_\rho(\mu,\nu).$$

Проверим, что $k_\rho(e_i,e_j)=\rho_{ij}$. Случай $i=j$ разобран выше, поэтому сразу предположим, что $i\neq j$. Заметим, что если $\pi\in\Pi(e_i,e_j)$, то при каждом $p\notin\{i,j\}$ имеем

$0 = \sum_k \pi_{pk} = \sum_k \pi_{kp}$, откуда ненулевыми могут быть лишь π_{ii}, π_{ij}, π_{ji}, π_{jj}. Далее, имеем $\pi_{ii} + \pi_{ij} = 1$, $\pi_{ji} + \pi_{jj} = 0$, $\pi_{ii} + \pi_{ji} = 0$ и $\pi_{ij} + \pi_{jj} = 1$, откуда мгновенно получаем, что $\pi_{ij} = 1$, а все остальные π_{pq} равны 0. Итак, $\Pi(e_i, e_j)$ состоит из одного элемента, который, тем самым, является оптимальным планом, поэтому $k_\rho(e_i, e_j) = K_\rho(\pi) = \rho_{ij}\pi_{ij} = \rho_{ij}$.

Наконец, для $0 \le t \le 1$ и $i \ne j$ положим $\lambda_t = (1-t)e_i + t\,e_j$, и для $0 \le t \le s \le 1$ вычислим $k_\rho(\lambda_t, \lambda_s)$. Пусть $\pi \in \Pi(\lambda_t, \lambda_s)$. Тогда граничные условия на π снова приводят к выводу, что при каждом $p \notin \{i, j\}$ имеем $\pi_{pk} = \pi_{kp} = 0$ при всех k, так что ненулевыми могут быть, как и раньше, только π_{uv} для $u, v \in \{i, j\}$. Напишем оставшиеся краевые условия. Имеем

$$\pi_{ii} + \pi_{ij} = 1 - t, \quad \pi_{ji} + \pi_{jj} = t, \quad \pi_{ii} + \pi_{ji} = 1 - s, \quad \pi_{ij} + \pi_{jj} = s.$$

Решая эту систему, получаем однопараметрическое семейство $\pi_{ji} = a$, $\pi_{ij} = s - t + a$, $\pi_{jj} = t - a$, $\pi_{ii} = 1 - s - a$. Таким образом, $K_\rho(\pi) = \rho_{ij}(s - t + 2a)$, причем $0 \le a \le \min(t, 1-s)$. При $t \le s$ наименьшее значение функции $K_\rho(\pi)$ приходится на $a = 0$, откуда $k_\rho(\lambda_t, \lambda_s) = \rho_{ij}(s - t)$, что и требовалось. □

Определение 1.6. Метрику k_ρ на симплексе Δ вероятностных мер, определенную в теореме 1.5, будем называть *транспортным расстоянием* или *метрикой Канторовича.*