Московский государственный университет им. М. В. Ломоносова
Механико-математический факультет
Кафедра Дифференциальной Геометрии и Приложений

Бардамов И. А.

Проблема Ферма–Штейнера в пространстве вероятностных мер на конечных 1-пространствах

Курсовая работа, 3 курс

Научный руководитель
профессор А. А. Тужилин

Москва, 2023

Пусть дано метрическое пространство, на котором заданы вероятностные меры $\mu_1, \mu_2, \dots, \mu_n$. Нужно найти такую вероятностную меру ν, что сумма расстояний в фиксированной метрике от неё до мер $\mu_1, \mu_2, \dots, \mu_n$ принимает наименьшее возможное значение.

В настоящей работе эта задача решается для k-точечного метрического пространства, в котором все расстояния равны 1, и мер на нём, расстояния между которыми задаются метриками Прохорова и Канторовича. В этих условиях получено явное описание всех возможных искомых мер ν.

1 Основные определения и предварительные результаты

Пусть X это k-точечное метрическое пространство. Расстояние между точками a и b будем обозначать через $|ab|$. Для любой точки $a \in X$ и числа $\varepsilon > 0$ через $U_\varepsilon(a)$ будем обозначать открытый шар с центром в точке a и радиусом ε; для каждого непустого $A \subset X$ и числа $\varepsilon > 0$ положим $U_\varepsilon(A) = \bigcup\limits_{\alpha \in A} U_\varepsilon(\alpha)$

Определение 1.1. [1] Пусть — $\mathcal{P}(X)$ множество вероятностных мер на X. Также пусть $\mathcal{B}_X$ борелевская σ-алгебра на X. Тогда можно ввести понятие *метрики Прохорова* на $\mathcal{P}(X)$. Положим для $\nu, \mu \in \mathcal{P}(X)$

$$d_P(\mu, \nu) = \inf\{\varepsilon > 0 : \mu(A) \leqslant \nu\big(U_\varepsilon(A)\big) + \varepsilon \ \& \ \nu(A) \leqslant \mu\big(U_\varepsilon(A)\big) + \varepsilon \ \ \forall A \in \mathcal{B}_X\}.$$

Теперь введём метрику Канторовича. Пусть ν и μ заданные вероятностные меры на X. Требуется найти такую неотрицательную функцию $\pi(a, b)$ на $X \times X$, чтобы выполнялись следующии условия:

1. $\sum\limits_{b \in X} \pi(a, b) = \nu(a)$ при всех $a \in X$;
2. $\sum\limits_{a \in X} \pi(a, b) = \mu(b)$ при всех $b \in X$;
3. $K(\pi) = \sum\limits_{a,b \in X} |ab|\pi(a, b) \to \min$.

Каждое π, удовлетворяющие условиям 1, 2 назовем транспортным планом типа (ν, μ), множество всех транспортных планов обозначим $\Pi(\nu, \mu)$.

Определение 1.2. [2] Пусть — $\mathcal{P}(X)$ множество вероятностных мер на X. Тогда можно ввести понятие *метрики Канторовича* на $\mathcal{P}(X)$. Положим для $\nu, \mu \in \mathcal{P}(X)$

$$k_\rho(\nu, \mu) = \inf\{K(\pi) : \pi \in \Pi(\nu, \mu)\}.$$

Также отметим, что метрику Канторовича в случае конечного метрического пространства можно наглядно описать следующим образом. Пусть в каждой точке a нашего пространства есть склад вместительности $\nu(a)$ и ресурсы в количестве $\mu(a)$. Чтобы перевезти груз массой x из точки a в точку b нужно "заплатить" $x|ab|$. Также отметим, что можно разделять ресурсы и перемещать на разные склады. Тогда расстояние между мерами ν и μ назовем минимальное количество денег, которое нужно затратить на перевозки, чтобы каждый склад был заполнен ровно полностью.

Определение 1.3. Пусть дано метрическое пространство, на котором заданы вероятностные меры $\mu_1, \mu_2, \dots, \mu_n$. Назовем меру ν *оптимальной*, если сумма расстояний в фиксированной метрике от неё до мер $\mu_1, \mu_2, \dots, \mu_n$ принимает наименьшее возможное значение.

Пусть заданы вероятностные меры $\mu_1, \mu_2, \dots, \mu_n$ и ν на конечном метрическом пространстве. Обозначим упорядоченный набор значений мер $\mu_1, \mu_2, \dots, \mu_n$ в точке a следующим образом:

$$\mu(a)^{(1)} \leqslant \mu(a)^{(2)} \leqslant \dots \leqslant \mu(a)^{(n)}.$$

Дополнительно положим $\mu(a)^{(0)} = 0$ и $\mu(a)^{(n+1)} = 1$ для всех точек a.

Определение 1.4. Пусть заданы вероятностные меры $\mu_1, \mu_2, \dots, \mu_n$ и ν на конечном метрическом пространстве. Тогда точке a сопоставим тройку $(l_a(\nu), m_a(\nu), r_a(\nu))$ по следующему принципу:

$$\mu(a)^{(1)} \leqslant \dots \leqslant \mu(a)^{(l_a)} < \nu(a) = \mu(a)^{(l_a+1)} = \dots = \mu(a)^{(l_a+m_a)} < \\ < \mu(a)^{(l_a+m_a+1)} \leqslant \dots \leqslant \mu(a)^{(l_a+m_a+r_a)}.$$

Неформально говоря, $l_a(\nu)$ — это количество мер, которые в точке a меньше ν, $m_a(\nu)$ — количество равных ν в точке a, и $r_a(\nu)$ — количество мер, больших ν в точке a. Имеем $l_a(\nu) + m_a(\nu) + r_a(\nu) = n$ для любого a.

Определение 1.5. Пусть заданы вероятностные меры $\mu_1, \mu_2, \dots, \mu_n$ на конечном метрическом пространстве. Положим $S^j = \sum_{a=1}^{k} \mu(a)^{(j)}$.

2 Основные результаты

Далее полагаем, что X это k-точечное метрическое пространство, в котором все расстояния равны 1. Нам понадобится следующая

Лемма 2.1. *Пусть μ и ν вероятностные меры на X. Тогда если $\mu(a)$ и $\nu(a)$ — значения соответствующих вероятностных мер в точке a, то*

$$d_P(\mu, \nu) = k_\rho(\mu, \nu) = \frac{1}{2} \sum_{a=1}^{k} |\mu(a) - \nu(a)|.$$

Доказательство. 1. Сначала докажем, что формула верна для метрики Прохорова. Для начала заметим, что если X — конечное метрическое пространство в котором все расстояния равны 1, то $\mathcal{B}_X = 2^X$. Также для любых таких X верно $U_\varepsilon(A) = A$ при $\varepsilon \leqslant 1$, и $U_\varepsilon(A) = X$ при $\varepsilon > 1$. Отсюда для мер получаем равенства $\nu\big(U_\varepsilon(A)\big) = \nu(A)$ при $\varepsilon \leqslant 1$, и $\nu\big(U_\varepsilon(A)\big) = 1$ при $\varepsilon > 1$, и поэтому

$$d_P(\mu, \nu) = \min\{\varepsilon > 0 : \mu(A) \leqslant \nu(A) + \varepsilon \ \& \ \nu(A) \leqslant \mu(A) + \varepsilon \ \ \forall A \subset X\} = \\ = \max\{|\mu(A) - \nu(A)| : A \subset X\}.$$

Максимум достигается в случае, когда A содержит все точки, для которых $\mu(a) \geqslant \nu(a)$, либо когда A содержит все точки в которых $\nu(a) \geqslant \mu(a)$. Заметим, что

$$\sum_{\mu(a) \geqslant \nu(a)} \big(\mu(a) - \nu(a)\big) = \sum_{\mu(a) \leqslant \nu(a)} \big(\nu(a) - \mu(a)\big).$$

Таким образом,

$$d_P(\mu,\nu) = \sum_{\mu(a)\geqslant\nu(a)} \big(\mu(a)-\nu(a)\big) = \frac{1}{2}\sum_{a=1}^{k} |\mu(a)-\nu(a)|.$$

2. Для начала заметим, что в силу того, что все расстояния равны 1 и нужно переместить ресурсы из тех точек, где $\mu(a) > \nu(a)$ верно

$$k_\rho(\mu,\nu) \geqslant \sum_{\mu(a)>\nu(a)} \big(\mu(a)-\nu(a)\big).$$

Построим план, на котором будет достигаться показанная выше оценка. Пусть для любой вершины c выполнено $\pi(c,c) = \min\big(\nu(c),\mu(c)\big)$. Далее рассмотрим вершину b, такую что $\mu(b) > \nu(b)$. Так как $\mu(b)-\nu(b) \leqslant \sum\limits_{\mu(a)>\nu(a)} \big(\mu(a)-\nu(a)\big) = \sum\limits_{\mu(a)<\nu(a)} \big(\nu(a)-\mu(a)\big)$ можно перевезти $\mu(b)-\nu(b)$ ресурсов на склады так, чтобы они не "переполнились". Аналогичные рассуждения будут верны для всех следующих вершин c, таких что $\mu(c) > \nu(c)$. Таким образом через конечное число шагов мы перевезём все ресурсы не "переполнив" склады. Окончательно, в силу того, что $\sum\limits_{\mu(a)>\nu(a)} \big(\mu(a)-\nu(a)\big) = \sum\limits_{\mu(a)<\nu(a)} \big(\nu(a)-\mu(a)\big)$, все склады будут заполнены ровно полностью. □

Может показаться, что искомую меру можно построить следующим образом:

1. Для каждого $a \in X$ и найдем число x_a, наименее удаленное от чисел $\mu_1,\mu_2,\ldots,\mu_n$;
2. В качестве оптимальной меры возьмем $\nu(a) = x_a$.

Однако ν не обязательно является мерой, что показывает следующий

Пример 2.2.

$$\mu_1(1) = 0.1,\ \mu_1(2) = 0.3,\ \mu_1(3) = 0.6$$

$$\mu_2(1) = 0.2,\ \mu_2(2) = 0,\ \mu_2(3) = 0.8$$

$$\mu_3(1) = 0.3,\ \mu_3(2) = 0.4,\ \mu_3(3) = 0.3$$

Заметим, что следуя выше указанному алгоритму, мы получим числа 0.2, 0.3 и 0.6. Но их сумма не равна 1.

Теперь докажем две основные леммы.

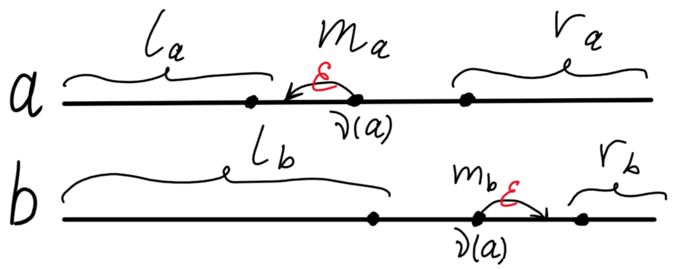
Лемма 2.3. *Пусть $\mu_1,\mu_2,\ldots,\mu_n$ — вероятностные меры на X, и мера ν является для них оптимальной. Тогда для любых вершин a и b выполнено неравенство $l_b(\nu)+m_b(\nu) \geqslant l_a(\nu)$.*

Доказательство. Заметим, что если $\nu(a) = 0$, то $l_a(\nu) = 0$, а значит требуемое равенство выполнено. Также при $\nu(b) = 1$ выполнено $l_b(\nu)+m_b(\nu) = n$, откуда следует, что равенство верно. Теперь рассмотрим оставшиеся случаи.

Построим новую меру $\overline{\nu}$ следующим образом:

$$\overline{\nu}(c) = \begin{cases} \nu(c), & c \neq a,b, \\ \nu(a)-\varepsilon, & c = a, \\ \nu(b)+\varepsilon, & c = b, \end{cases}$$

где $\varepsilon > 0$ выбирается достаточно малым, чтобы были верны равенства $l_a(\nu) = l_a(\overline{\nu})$ и $r_b(\nu) = r_b(\overline{\nu})$ (такой ε всегда существует, например $\varepsilon = \frac{1}{2}\min\big(\nu(a)-\mu(a)^{(l_a(\nu))}, \mu(b)^{(l_b(\nu)+m_b(\nu)+1)}-\nu(b)\big)$.



В силу оптимальности меры ν имеем

$$\frac{1}{2}\sum_{a=1}^{k}\sum_{i=1}^{n}|\overline{\nu}(a)-\mu_i(a)| \geqslant \frac{1}{2}\sum_{a=1}^{k}\sum_{i=1}^{n}|\nu(a)-\mu_i(a)|,$$

поэтому

$$\sum_{a=1}^{k}\sum_{i=1}^{n}|\overline{\nu}(a)-\mu_i(a)|-\sum_{a=1}^{k}\sum_{i=1}^{n}|\nu(a)-\mu_i(a)| = \big(r_a(\nu)+m_a(\nu)\big)\varepsilon-l_a(\nu)\varepsilon+\big(m_b(\nu)+l_b(\nu)\big)\varepsilon-r_b(\nu)\varepsilon =$$
$$= \big(n-l_a(\nu)-m_a(\nu)+m_a(\nu)\big)\varepsilon-l_a(\nu)\varepsilon+\big(m_b(\nu)+l_b(\nu)\big)\varepsilon-\big(n-l_b(\nu)-m_b(\nu)\big)\varepsilon =$$
$$= 2\varepsilon\big(l_b(\nu)+m_b(\nu)-l_a(\nu)\big) \geqslant 0.$$

Отсюда $l_b(\nu)+m_b(\nu) \geqslant l_a(\nu)$. □

Лемма 2.4. *Пусть $\mu_1, \mu_2, ..., \mu_n$ — вероятностные меры на X, и мера ν является для них оптимальной. Тогда существует такое целое $q \in [0, n]$, что для любого с выполнено $\nu(c) \in [\mu(c)^{(q)}, \mu(c)^{(q+1)}]$.*

Доказательство. Выберем точку a так, чтобы значение $l_a(\nu)+m_a(\nu)$ было минимальным. Тогда $\mu(b)^{(l_a(\nu)+m_a(\nu))} \leqslant \mu(b)^{(l_b(\nu)+m_b(\nu))} = \nu(b) \leqslant \mu(b)^{(l_b(\nu)+1)} \leqslant \mu(b)^{(l_a(\nu)+m_a(\nu)+1)}$, где последнее неравенство выполнено в силу леммы 2.2. Таким образом подходящим q будет $l_a(\nu)+m_a(\nu)$. □

Теорема 2.5. *Пусть $\mu_1, \mu_2, \ldots, \mu_n$ — вероятностые меры на X, и q выбрано так, что $S^q < 1$ и $S^{q+1} \geqslant 1$. Тогда следующие 2 утверждения эквивалентны:*

1. *ν — оптимальная мера.*
2. *$\nu(a) \in [\mu(a)^{(q)}, \mu(a)^{(q+1)}]$ и $\sum\limits_{a=1}^{k}\nu(a) = 1$.*

Доказательство. Так как пространство вероятностных мер компактно и функция расстояний до фиксированных мер непрерывна, минимум этой функции достигается хотя бы в одной точке, являющейся оптимальной мерой. Обозначим её ν. Тогда в силу леммы 2.3 существует такое p, что $\nu(a) \in [\mu(a)^{(p)}, \mu(a)^{(p+1)}]$. Так как $\sum\limits_{a=1}^{k}\nu(a) = 1$, то $\nu(a)$ обязана принадлежать всем отрезкам $[\mu(a)^{(q)}, \mu(a)^{(q+1)}]$, таким образом, импликация $1 \Rightarrow 2$ уже доказана. Теперь покажем что любая мера $\overline{\nu}$ удовлетворяющая условиям 2 является оптимальной. Положим $h(a) = \nu(a) - \overline{\nu}(a)$. Тогда

$$\sum_{a=1}^{k}\sum_{i=1}^{n}|\nu(a)-\mu_i(a)|-\sum_{a=1}^{k}\sum_{i=1}^{n}|\overline{\nu}(a)-\mu_i(a)| = \sum_{a=1}^{k}\big(qh(a)-(n-q)h(a)\big) = (2q-n)\sum_{a=1}^{k}h(a) = 0.$$

Значит, мера $\overline{\nu}$ тоже является оптимальной. □

Дополнительно отметим, что если для оптимальной меры ν известно, что $\nu(a) \in [\mu(a)^{(q)}, \mu(a)^{(q+1)}]$, то q может быть определено неоднозначно. Например когда все меры совпадают.

Список литературы

[1] А.О.Иванов, А.А.Тужилин: Материалы курса 2018-2019 года "Транспортная задача Канторовича и геометрия пространств вероятностных мер."

[2] А.О.Иванов, А.А.Тужилин: Материалы курса 2019-2020 года (весенний семестр) "Геометрия пространств компактов с метриками Хаусдорфа и Громова–Хаусдорфа. Метрические тройки Громова."