Барицентры Фреше в факторизованном пространстве Монжа—Канторовича

Алексей Крошнин[1,2]

[1] Институт проблем передачи информации РАН
[2] Московский Физико-Технический Институт
kroshnin@phystech.edu

Аннотация. В статье рассматривается пространство $\mathcal{P}(X)$ вероятностных мер на произвольном польском пространстве X, снабженное транспортным функционалом $J(\cdot, \cdot)$. Исследуется факторпространство $\mathcal{P}(X)/G$, где G — некоторая группа, действующая на X, и транспортное расстояние в этом пространстве. Вводится понятие обобщенного среднего в факторпространстве $\mathcal{P}(X)/G$, называемое здесь *факторизованным барицентром Фреше*. Для барицентров доказываются достаточные условия сходимости, в частности, показывается их состоятельность.

Ключевые слова: пространство Монжа—Канторовича; транспортное расстояние; факторизация; барицентр Фреше; состоятельность.

1 Введение

Меры, в том числе вероятностные, являются естественными объектами во многих областях. Существуют различные формализации понятия близости распределений, например: дивергенция Кульбака—Лейблера, расстояние полной вариации или слабая сходимость. В этом плане оказывается весьма полезной задача Монжа—Канторовича и определяемое с ее помощью транспортное расстояние. Задача Монжа—Канторовича, или оптимального транспорта, состоит в нахождении отображения одной меры в другую с минимальной стоимостью, если задана ценовая функция — стоимость переноса единицы массы из одной точки в другую (см. [1], гл. 1). Соответственно, транспортным расстоянием между двумя мерами будем называть эту минимальную стоимость преобразования. Данное расстояние позволяет учитывать структуру пространства в терминах подходящим образом определенной ценовой функции, поэтому задача Монжа—Канторовича нашла широкое применение во многих областях, например, в анализе изображений и других сложных объектов. С другой стороны, при анализе изображений иногда нужно отождествлять те из них, которые получаются друг из друга с помощью сдвига или поворота. Это приводит к идее рассмотреть пространство Монжа—Канторовича, факторизованное под действием некоторой группы G, с индуцированным на нем транспортным расстоянием.

Возможность измерять расстояние между мерами позволяет также поставить другую важную задачу — о нахождении среднего в пространстве мер. В статье [2] вводится понятие *барицентра Вассерштейна* (для пространства 2-Вассерштейна) как среднего по Фреше, т. е. меры, которая минимизирует сумму стоимостей своего оптимального переноса в каждую из заданного набора мер. Определение барицентра очевидным образом можно обобщить на распределение на мерах, как в статьях [3,4], где также доказывается сходимость эмпирических барицентров к барицентру распределения (но, все так же, для пространств Вассерштейна).

В данной работе рассматривается общая постановка транспортной задачи: для мер над произвольным польским пространством X и широкого класса ценовых функций. Как было показано в [5], даже при достаточно слабых ограничениях $\mathcal{P}(X)$ (пространство вероятностных мер на X), снабженное транспортным расстоянием, обладает некоторыми полезными свойствами: сепарабельность, метризуемость и т. д. В частности, барицентр мер является состоятельным средним. Оказывается, факторпространство $\mathcal{P}(X)/G$ тоже обладает такими свойствами.

Работа построена следующим образом. В разделе 2 дается строгая формулировка транспортной задачи и приводятся некоторые свойства расстояния Монжа—Канторовича и топологии, порожденной транспортным расстоянием. В подразделе 2.3 рассматривается факторизованное пространство Монжа—Канторовича и свойства индуцированного транспортного расстояния. В разделе 3 вводится определение факторизованного барицентра Фреше, доказывается его существование и состоятельность (в том числе непрерывность барицентра от конечного набора мер и функциональный закон больших чисел для эмпирических барицентров).

2 Факторизованное пространство Монжа—Канторовича

2.1 Используемые обозначения

Пусть X — топологическое пространство, снабженное борелевской σ-алгеброй. Через $\mathcal{P}(X)$ будем обозначать пространство вероятностных мер на X.

Напомним, что слабая (narrow) сходимость в пространстве мер над топологическим пространством X определяется следующим образом: меры μ_n слабо сходятся к μ ($\mu_n \rightharpoonup \mu$), если для любой ограниченной непрерывной функции $f\colon X \to \mathbb{R}$ сходятся интегралы $\int_X f(x)\,d\mu_n(x) \to \int_X f(x)\,d\mu(x)$.

Пусть даны измеримые пространства X, Y и измеримое отображение $T\colon X \to Y$. Пусть μ — мера на пространстве X. Через $T_{\#}\mu$ будем обозначать образ меры μ при отображении T, т. е. меру на Y, определяемую следующим образом:

$$(T_{\#}\mu)(A) := \mu\big(T^{-1}(A)\big), \text{ где } Y \supset A \text{ — любое измеримое множество.}$$

При интегрировании обозначения аргумента функции и пространства иногда будут опускаться, если это не может вызвать неоднозначности.

2.2 Транспортное расстояние

Пусть X — польское пространство (т. е. полное метрическое и сепарабельное), и задана непрерывная *ценовая функция* $c\colon X \times X \to [0,\infty)$. Рассмотрим произвольные меры $\mu, \nu \in \mathcal{P}(X)$ и определим множество *транспортных планов*

$$\Pi(\mu,\nu) := \Big\{ \gamma \in \mathcal{P}(X \times X) : \pi^x_{\#}\gamma = \mu,\ \pi^y_{\#}\gamma = \nu) \Big\},$$

где π^x, π^y — операторы проекции на первую и вторую компоненту, соответственно.

Определение 1. Расстоянием Монжа—Канторовича*, или* транспортным функционалом*, будем называть*

$$J(\mu,\nu) := \inf_{\gamma \in \Pi(\mu,\nu)} \int c(x,y)\, d\gamma(x,y),\ \ \mu,\nu \in \mathcal{P}(X).$$

Наложим дополнительные условия на ценовую функцию: во-первых, потребуем *согласованности* ценовой функции с топологией: $x_n \to x \Leftrightarrow c(x,x_n) \to 0 \Leftrightarrow c(x_n,x) \to 0$ (в частности, $c(x,y) = 0 \Leftrightarrow x = y$). Также нам потребуется следующее предположение о ценовой функции.

Предположение 2 (ослабленное неравенство треугольника). Пусть для любого $\varepsilon > 0$ существует такая константа $C_\varepsilon > 0$, что для всех $x,y,z \in X$ выполняется

$$\begin{aligned} c(x,y) &\le \varepsilon + (1+\varepsilon)c(x,z) + C_\varepsilon c(y,z),\\ c(x,y) &\le \varepsilon + (1+\varepsilon)c(z,y) + C_\varepsilon c(z,x). \end{aligned}$$

Имеют место следующие свойства транспортного функционала (см. [5]):

1. Функционал $J(\cdot,\cdot)$ полунепрерывен снизу относительно слабой сходимости. Т.е. если $\mu_n \rightharpoonup \mu$ и $\nu_n \rightharpoonup \nu$, то $J(\mu,\nu) \le \liminf J(\mu_n,\nu_n)$.
2. Расстояние Монжа—Канторовича "наследует" слабое неравенство треугольника: для любых мер $\mu,\nu,\lambda \in \mathcal{P}(X)$ и $\varepsilon > 0$ выполняется

   $$\begin{aligned} J(\mu,\nu) &\le \varepsilon + (1+\varepsilon)J(\mu,\lambda) + C_\varepsilon J(\nu,\lambda),\\ J(\mu,\nu) &\le \varepsilon + (1+\varepsilon)J(\lambda,\nu) + C_\varepsilon J(\lambda,\mu). \end{aligned}$$

   В частности, $J(\mu_n,\nu_n) \to 0 \Leftrightarrow J(\nu_n,\mu_n) \to 0$, и из $J(\mu_n,\nu_n) \to 0$, $J(\nu_n,\lambda_n) \to 0$ следует $J(\mu_n,\lambda_n) \to 0$.
3. Семейство шаров $B^J_r(\mu) := \big\{\nu \in \mathcal{P}(X) : J(\mu,\nu) < r\big\}$, где $\mu \in \mathcal{P}(X)$, $r > 0$, образует базу *транспортной топологии* τ_J в $\mathcal{P}(X)$, и следующие условия эквивалентны:
   (a) $\nu_n \xrightarrow{J} \nu^*$, т. е. ν_n сходится к ν относительно τ_J;
   (b) $\nu_n \rightharpoonup \nu^*$ и $J(\mu,\nu_n) \to J(\mu,\nu^*)$ для любой меры $\mu \in \mathcal{P}(X)$;
   (c) $\nu_n \rightharpoonup \nu^*$ и $J(\mu,\nu_n) \to J(\mu,\nu^*) < \infty$ для какой-то меры $\mu \in C(\nu^*)$.

4. Если ценовая функция ограничена, то слабая сходимость эквивалентна транспортной.

Сделаем еще одно предположение — о локальной компактности пространства X.

Предположение 3 (локальная компактность). Пусть для некоторой (и, следовательно, для любой) точки $x_0 \in X$ любой замкнутый шар $\overline{B}_r^c(x_0) := \big\{y \in X : c(x_0, y) \le r\big\}$ является компактом.

Тогда верно следующее свойство.

5. Любой замкнутый шар $\overline{B}_r^J(\mu) := \big\{\nu \in \mathcal{P}(X) : J(\mu,\nu) \le r\big\}$ является слабым компактом в $\mathcal{P}(X)$.

2.3 Факторизация

Пусть задано измеримое действие некоторой группы G на X, т. е. для любого элемента $g \in G$ определена измеримая (борелевская) биекция $x \mapsto gx$. Действие на X, в свою очередь, порождает действие на $\mathcal{P}(X)$: $\mu \mapsto g\mu := (x \mapsto gx)_{\#}\mu$. Будем рассматривать такое действие, относительно которого ценовая функция инвариантна: для всех $x, y \in X$, $g \in G$ выполняется $c(gx, gy) = c(x, y)$. Отсюда немедленно следует, что расстояние Монжа—Канторовича $J(\cdot,\cdot)$ также является инвариантным относительно действия группы. На факторпространстве $\mathcal{P}(X)/G$ (пространстве орбит) можно определить индуцированное транспортное расстояние. Но нам удобнее работать с исходным пространством $\mathcal{P}(X)$, поэтому вместо классов эквивалентности определим новый функционал над мерами.

$$\tilde{J}(\mu,\nu) := J(G\mu, G\nu) := \inf_{g_1, g_2 \in G} J(g_1\mu, g_2\nu).$$

Из инвариантности следует, что

$$\tilde{J}(\mu,\nu) = \inf_{g \in G} J(g\mu,\nu) = \inf_{g \in G} J(\mu, g\nu).$$

Как несложно видеть, $\tilde{J}(\cdot,\cdot)$ тоже удовлетворяет ослабленному неравенству треугольника. Действительно, для любых мер $\mu, \nu, \lambda \in \mathcal{P}(X)$ и любого $\varepsilon > 0$ выполняется

$$\begin{aligned}
\tilde{J}(\mu,\nu) = \inf_{g \in G} J(g\mu,\nu) \le&\\
&\le \inf_{g, g_1 \in G}\big(\varepsilon + (1+\varepsilon)J(g\mu, g_1\lambda) + C_\varepsilon J(\nu, g_1\lambda)\big) =\\
&= \varepsilon + \inf_{g, g_1 \in G}\big((1+\varepsilon)J(\mu, g^{-1}g_1\lambda) + C_\varepsilon J(\nu, g_1\lambda)\big) =\\
&= \varepsilon + (1+\varepsilon)\inf_{g_2 \in G} J(\mu, g_2\lambda) + C_\varepsilon \inf_{g_1 \in G} J(\nu, g_1\lambda) =\\
&= \varepsilon + (1+\varepsilon)\tilde{J}(\mu,\lambda) + C_\varepsilon \tilde{J}(\nu,\lambda).
\end{aligned}$$

Аналогично,
$$\tilde{J}(\mu,\nu) \le \varepsilon + (1+\varepsilon)\tilde{J}(\lambda,\nu) + C_\varepsilon \tilde{J}(\lambda,\mu).$$

Предположение 4. Для любой меры μ ее орбита $G\mu := \{g\mu : g \in G\}$ является слабо замкнутой.

Пример 5 (пространства Вассерштейна над $\mathbb{R}^d$). Пусть $X = \mathbb{R}^d$, $c(x,y) := \|x-y\|_2^p$, т. е. мы имеем пространство Вассерштейна степени p. Ценовая функция и пространство удовлетворяют всем предположениям, которые были сделаны (см. например [5, Theorem 4.4]). В качестве группы можно взять группу поворотов и сдвигов на $\mathbb{R}^d$. Несложно убедиться, что орбита любой меры под действием этой группы является слабо замкнутой.

Лемма 6. *Для любых мер $\mu,\nu \in \mathcal{P}(X)$ существует такой элемент $g \in G$, что*
$$\tilde{J}(\mu,\nu) = J(\mu, g\nu).$$

Доказательство. Если $\tilde{J}(\mu,\nu) = \infty$, то утверждение, очевидно, верно. Пусть $\tilde{J}(\mu,\nu) < \infty$. Рассмотрим минимизирующую последовательность: $J(\mu, g_n\nu) \to \inf_{g\in G} J(\mu, g\nu) = \tilde{J}(\mu,\nu)$. По свойству 5 у нее найдется слабо сходящаяся подпоследовательность $g_{n_k}\nu \rightharpoonup \lambda$. В силу слабой замкнутости орбиты $\lambda = g^*\nu$ для некоторого $g^* \in G$. Из слабой полунепрерывности снизу транспортного расстояния получаем, что
$$J(\mu, g^*\nu) \le \liminf J(\mu, g_{n_k}\nu) = \tilde{J}(\mu,\nu),$$
т. е. $\tilde{J}(\mu,\nu) = J(\mu, g^*\nu)$.

Следствие 7. $\tilde{J}(\mu,\nu) = 0$ *тогда и только тогда, когда $G\mu = G\nu$. Таким образом, в факторпространстве транспортное расстояние тоже разделяет точки.*

Можно показать, что шары $\tilde{B}_r^J(\xi) := \left\{\eta \in \mathcal{P}(X)/G : J(\xi,\eta) < r\right\}$ образуют базу топологии на $\mathcal{P}(X)/G$, индуцированной τ_J. Через $\nu_n \xrightarrow{\tilde{J}} \nu$ будем обозначать сходимость $\tilde{J}(\nu_n,\nu) \to 0$.

Лемма 8. *Для любой меры $\mu \in \mathcal{P}(X)$ функционал $\tilde{J}(\mu,\cdot)$ является полунепрерывным снизу относительно слабой сходимости.*

Доказательство. Пусть $\nu_n \rightharpoonup \nu$. Достаточно рассмотреть случай, когда существует предел $\lim \tilde{J}(\mu,\nu_n) < \infty$. По Лемме 6 можно выбрать такие $g_n, g \in G$, что $\tilde{J}(\mu,\nu_n) = J(\mu, g_n\nu_n)$ и $\tilde{J}(\mu,\nu) = J(\mu, g\nu)$. По свойству 5 можно найти слабо сходящуюся подпоследовательность $g_{n_k}\nu_{n_k} \rightharpoonup \nu^*$, и по свойству 1 $J(\mu,\nu^*) \le \liminf J(\mu, g_{n_k}\nu_{n_k}) = \lim \tilde{J}(\mu,\nu_n)$.

Покажем, что $\nu^* \in G\nu$. Рассмотрим ценовую функцию $\bar{c}(x,y) := \min\{c(x,y), 1\}$. Несложно видеть, что она удовлетворяет Предположению 2 и остальным условиям на ценовую функцию, кроме, возможно, Предположения 3. Так как она ограничена, то сходимость относительно расстояния Монжа—Канторовича $\bar{J}(\cdot,\cdot)$ с этой ценовой функцией эквивалентна слабой сходимости по свойству 4, следовательно, $\bar{J}(g_n\nu, g_n\nu_n) = \bar{J}(\nu,\nu_n) \to 0$ и $\bar{J}(\nu^*, g_{n_k}\nu_{n_k}) \to 0$. Отсюда по свойству 2 получаем, что $\bar{J}(g_{n_k}\nu, \nu^*) \to 0$, т. е. $g_{n_k}\nu \rightharpoonup \nu^*$. Из слабой замкнутости орбиты следует, что $\nu^* \in G\nu$. Таким образом,

$$\tilde{J}(\mu,\nu) \le J(\mu,\nu^*) \le \liminf \tilde{J}(\mu,\nu_n).$$

3 Барицентры Фреше в факторпространстве

Транспортное расстояние позволяет ввести усреднение в пространстве мер с помощью конструкции среднего по Фреше. Такое среднее, называемое в данном случае барицентром Фреше (или Вассерштейна, для пространств Вассерштейна), определяется как точка, которая минимизирует среднее отклонение $\mathbb{E}\, J(\boldsymbol{\mu},\nu)$, где $\boldsymbol{\mu}$ — случайная мера [2,6,3]. Регуляризованные барицентры Вассерштейна рассматриваются также в [7]. По аналогии с барицентрами Фреше и Вассерштейна определим факторизованный барицентр Фреше.

Определение 9. *Пусть дано борелевское распределение P на $\mathcal{P}(X)$ и функционал $H\colon \mathcal{P}(X) \to [0,+\infty)$ (регуляризатор), инвариантный относительно действия G.* H-регуляризованным факторизованным барицентром Фреше распределения P *будем называть решение задачи*

$$\int \tilde{J}(\mu,\nu)\, dP(\mu) + H(\nu) \to \inf_{\nu\in\mathcal{P}(X)} \tag{1}$$

обозначая его $\mathrm{bar}_H(P)$.

Заметим, что можно рассматривать распределения и искать барицентр сразу в пространстве $\mathcal{P}(X)/G$, но задача, поставленная выше, кажется более естественной, так как скорее мы наблюдаем исходные меры, а не орбиты. Кроме того, любая мера $P \in \mathcal{P}\big(\mathcal{P}(X)\big)$ естественным образом индуцирует распределение $(\mu \mapsto G\mu)_\# P$ на $\mathcal{P}(X)/G$.

Теорема 10. *Пусть дано распределение P на $\mathcal{P}(X)$ и регуляризатор $H(\cdot)$, полунепрерывный снизу относительно слабой сходимости. Если*

$$\inf_{\nu\in\mathcal{P}(X)} \Big(\int \tilde{J}(\mu,\nu)\, dP(\mu) + H(\nu)\Big) < +\infty,$$

то существует барицентр распределения P. Также, если $\{\nu_n\}_{n\in\mathbb{N}}$ — минимизирующая последовательность для задачи (1)*, то у нее найдется такая подпоследовательность $\{\nu_{n_k}\}_{k\in\mathbb{N}}$ и мера $\nu^* = \mathrm{bar}_H(P)$, что $\nu_{n_k} \xrightarrow{\tilde{J}} \nu^*$.*

Доказательство. Зафиксируем произвольную меру ν_0 такую, что

$$\int \tilde{J}(\mu,\nu_0)\, dP(\mu) + H(\nu_0) < +\infty.$$

По Лемме 6 можно выбрать такие $g_n \in G$, что $\tilde{J}(\nu_0, \nu_n) = J(\nu_0, g_n\nu_n)$. Из ослабленного неравенства треугольника следует, что

$$\tilde{J}(\nu_0, \nu_n) \le 1 + 2\int \tilde{J}(\mu, \nu_n)\, dP(\mu) + C_1 \int \tilde{J}(\mu, \nu_0)\, dP(\mu),$$

поэтому $\limsup \tilde{J}(\nu_0, \nu_n) < \infty$. По свойству 5 найдется слабо сходящаяся подпоследовательность (без переобозначения) $g_n\nu_n \rightharpoonup \nu^*$. Из слабой полунепрерывности снизу и леммы Фату получаем

$$\int \tilde{J}(\mu, \nu^*)\, dP(\mu) + H(\nu^*) \le \int \liminf \tilde{J}(\mu, g_n\nu_n)\, dP(\mu) + \liminf H(g_n\nu_n) \le$$

$$\le \liminf\Big(\int \tilde{J}(\mu, \nu_n)\, dP(\mu) + H(\nu_n)\Big) = \inf_{\nu \in \mathcal{P}(X)} \Big(\int \tilde{J}(\mu, \nu)\, dP(\mu) + H(\nu)\Big).$$

Следовательно, $\nu^* = \mathrm{bar}_H(P)$. Кроме того, для P-почти всех μ выполняется $\tilde{J}(\mu, \nu^*) = \liminf \tilde{J}(\mu, \nu_n) < \infty$. Зафиксируем произвольную такую меру μ_0. Без ограничения общности можно считать, что существует предел $\lim \tilde{J}(\mu_0, \nu_n) = \tilde{J}(\mu_0, \nu^*)$. Опять же, выберем такие элементы $\bar{g}_n \in G$, что $\tilde{J}(\mu_0, \nu_n) = J(\mu_0, \bar{g}_n\nu_n)$. Снова можно выбрать сходящуюся подпоследовательность такую, что (без переобозначения) $\bar{g}_n\nu_n \rightharpoonup \lambda$. Как было показано при доказательстве Леммы 8, $\lambda = g^*\nu^*$ для какого-то $g^* \in G$. Отсюда получаем, что

$$J(\mu_0, g^*\nu^*) \ge \tilde{J}(\mu_0, \nu^*) = \lim \tilde{J}(\mu_0, \nu_n) = \lim J(\mu_0, \bar{g}_n\nu_n).$$

Тогда по свойству 3 $\bar{g}_n\nu_n \xrightarrow{J} g^*\nu^*$. Таким образом, $\tilde{J}(\nu^*, \nu_n) \le J(g^*\nu^*, \bar{g}_n\nu_n) \to 0$, следовательно, $\nu_n \xrightarrow{\tilde{J}} \nu^* = \mathrm{bar}_H(P)$.

Теперь мы можем показать, что факторизованные барицентры состоятельны в том же смысле, что и обычные барицентры Фреше или Вассерштейна (см. [5, Theorem 5.5], [3, Theorem 2], [4, Theorem 6.1]). Для этого введем расстояние Монжа—Канторовича над мерами на $\mathcal{P}(X)$ с ценовой функцией $\tilde{J}(\cdot, \cdot)$:

$$\mathcal{J}(P, P') := \inf_{F \in \Pi(P, P')} \int \tilde{J}(\mu, \nu)\, dF(\mu, \nu),\ P, P' \in \mathcal{P}\big(\mathcal{P}(X)\big).$$

Теорема 11 (состоятельность барицентров). *Пусть выполнены условия Теоремы 10 и дана последовательность распределений* $\{P_n\}_{n\in\mathbb{N}} \subset \mathcal{P}\big(\mathcal{P}(X)\big)$, $\mathcal{J}(P_n, P) \to 0$ *для некоторого распределения* P*, у которого существует* H*-регуляризованный факторизованный барицентр Фреше.*

Тогда, начиная с некоторого n*, существуют барицентры распределений* P_n*, и у любой последовательности барицентров* $\nu_n = \mathrm{bar}_H(P_n)$ *существует такая подпоследовательность* $\{\nu_{n_k}\}_{k\in\mathbb{N}}$ *и мера* $\nu^* = \mathrm{bar}_H(P)$*, что* $\nu_n \xrightarrow{\tilde{J}} \nu^*$*. В частности, если существует единственный (с точностью до действия группы* G*) барицентр* $\nu^* = \mathrm{bar}_H(P)$*, то* $\nu_n \xrightarrow{\tilde{J}} \nu^*$*.*

Доказательство. Достаточно показать, что

$$\int \tilde{J}(\mu,\nu_n)\,dP + H(\nu_n) \to \inf_{\nu\in\mathcal{P}(X)} \Big(\int \tilde{J}(\mu,\nu)\,dP + H(\nu) \Big),$$

и применить Теорему 10.

Во-первых, рассмотрим произвольные распределения P, P' и меру ν. Для любого транспортного плана $F \in \Pi(P',P)$ и $\varepsilon > 0$ выполняется

$$\int \tilde{J}(\mu,\nu)\,dP(\mu) \le \int \big(\varepsilon + (1+\varepsilon)\tilde{J}(\mu',\nu) + C_\varepsilon \tilde{J}(\mu',\mu)\big)\,dF(\mu',\mu) =$$
$$= \varepsilon + (1+\varepsilon)\int \tilde{J}(\mu,\nu)\,dP'(\mu) + C_\varepsilon \int \tilde{J}(\mu',\mu)\,dF(\mu',\mu).$$

Минимизируя по транспортным планам, получаем

$$\int \tilde{J}(\mu,\nu)\,dP(\mu) \le \varepsilon + (1+\varepsilon)\int \tilde{J}(\mu,\nu)\,dP'(\mu) + C_\varepsilon \mathcal{J}(P',P).$$

Аналогично, $\mathcal{J}(P',P) \le \varepsilon + C_\varepsilon \mathcal{J}(P,P')$. В частности, отсюда следует, что $\mathcal{J}(P,P_n) \to 0$ и

$$\int \tilde{J}(\mu,\nu)\,dP_n(\mu) \to \int \tilde{J}(\mu,\nu)\,dP(\mu).$$

Следовательно, для любой меры $\nu \in \mathcal{P}(X)$ верно

$$\int \tilde{J}(\mu,\nu_n)\,dP(\mu) + H(\nu_n) \le$$
$$\le \varepsilon + (1+\varepsilon)\Big(\int \tilde{J}(\mu,\nu_n)\,dP_n(\mu) + H(\nu_n)\Big) + C_\varepsilon \mathcal{J}(P_n,P) \le$$
$$\le \varepsilon + (1+\varepsilon)\Big(\int \tilde{J}(\mu,\nu)\,dP_n(\mu) + H(\nu)\Big) + C_\varepsilon \mathcal{J}(P_n,P) \to$$
$$\to \varepsilon + (1+\varepsilon)\Big(\int \tilde{J}(\mu,\nu)\,dP(\mu) + H(\nu)\Big) \to$$
$$\to \int \tilde{J}(\mu,\nu)\,dP(\mu) + H(\nu) \text{ при } \varepsilon \to 0.$$

Таким образом, $\{\nu_n\}_{n\in\mathbb{N}}$ — минимизирующая последовательность, откуда следует утверждение теоремы.

Пример 12. Пусть даны последовательности мер $\mu_n^i \xrightarrow[n\to\infty]{\tilde{J}} \mu^i$ и весов $\lambda_n^i \xrightarrow[n\to\infty]{} \lambda^i$, $i = 1,\ldots,m$, $\lambda_n^i \ge 0$, $\sum_{i=1}^m \lambda_n^i = 1$. Пусть также $\tilde{J}(\nu^i,\nu^j) < \infty$ для любых i, j. Рассмотрим меры $P_n := \sum_{i=1}^m \lambda_n^i \delta_{\mu_n^i}$ и $P := \sum_{i=1}^m \lambda^i \delta_{\mu^i}$. Тогда $\mathcal{J}(P_n,P) \to 0$, следовательно, по Теореме 11 их барицентры сходятся к барицентру P.

Пример 13. Рассмотрим i. i. d. случайные меры $\boldsymbol{\mu}_n \sim P$. Пусть $\boldsymbol{P}_n := \sum_{i=1}^n \boldsymbol{\mu}_i$ — эмпирические меры. Тогда почти наверное $\mathcal{J}(P_n,P) \to 0$, и эмпирические барицентры $\boldsymbol{\nu}_n := \mathrm{bar}_H(\boldsymbol{P}_n)$ сходятся к барицентру распределения P.

Список литературы

1. F. Santambrogio. *Optimal Transport for Applied Mathematicians.* Birkhäuser, Basel, 2015.
2. M. Agueh and G. Carlier. Barycenters in the Wasserstein Space. *SIAM Journal on Mathematical Analysis*, 43(2):904–924, 2011.
3. T. Le Gouic and J.-M. Loubes. Existence and Consistency of Wasserstein Barycenters. *ArXiv e-prints*, 2015.
4. J. Bigot and T. Klein. Consistent Estimation of a Population Barycenter in the Wasserstein Space. *ArXiv e-prints*, 2015.
5. A. Kroshnin. Fréchet Barycenters in the Monge–Kantorovich spaces. *ArXiv e-prints*, 2016.
6. Y.-H. Kim and B. Pass. Wasserstein Barycenters over Riemannian manifolds. *ArXiv e-prints*, 2014.
7. J. Bigot, E. Cazelles, and N. Papadakis. Regularization of barycenters in the Wasserstein space. *ArXiv e-prints*, 2016.