

УДК 519.214

А. В. Крошнин^{1,2}, А. Н. Соболевский^{2,3}¹Московский физико-технический институт (государственный университет)²Институт проблем передачи информации им. А.А. Харкевича РАН³Национальный исследовательский университет «Высшая школа экономики»

Барицентры Фреше и закон больших чисел для мер на вещественной прямой

В статье рассматривается пространство $\mathcal{P}(\mathbb{R})$ вероятностных мер на вещественной прямой, снабженное транспортным функционалом $J(\mu, \nu)$, порожденным выпуклой ценовой функцией. Для распределения вероятности на $\mathcal{P}(\mathbb{R})$ вводится понятие среднего по отношению к транспортному расстоянию, называемое здесь *барицентром Фреше*, находится явный вид барицентра распределения и доказывается функциональный закон больших чисел для эмпирических барицентров.

Ключевые слова: задача Монжа–Канторовича, одномерный случай, барицентр, закон больших чисел.

A. V. Kroshnin^{1,2}, A. N. Sobolevski^{2,3}¹Moscow Institute of Physics and Technology (State University)²Institute for Information Transmission Problems of RAS (Kharkevich Institute)³National Research University «Higher School of Economics»

Fréchet Barycenters and the law of large numbers for measures on the real line

We endow the space $\mathcal{P}(\mathbb{R})$ of probability measures on \mathbb{R} with transportation cost $J(\mu, \nu)$ generated by a translation-invariant convex cost function. For the probability distribution on $\mathcal{P}(\mathbb{R})$ we introduce the notion of average with respect to this transportation cost called the *Fréchet barycenter* here, find an explicit formula for it, prove a version of the law of large numbers for Fréchet barycenters, and discuss the structure of $\mathcal{P}(\mathbb{R})$ with respect to the transportation cost J .

Key words: Monge–Kantorovich problem, one dimensional case, barycenter, law of large numbers.

1. Введение

В данной статье рассматривается транспортная задача Монжа–Канторовича и опирающаяся на нее задача нахождения среднего в пространстве мер. Транспортная задача состоит в нахождении отображения одной меры в другую с минимальной стоимостью, если задана ценовая функция $c(x, y)$ — стоимость переноса единицы массы из точки x в точку y . Про условия существования и единственности решения задачи Монжа–Канторовича известно достаточно много, но в большинстве случаев доказательства неконструктивные (см. [3]). Одномерный случай с выпуклой ценовой функцией примечателен тем, что позволяет найти решение задачи Монжа–Канторовича в явном виде, используя представление мер с помощью кумулятивных функций. Транспортное расстояние оказывается естественным во многих задачах, и возникает необходимость исследовать пространство мер, снабженное таким функционалом — в частности, задача об усреднении набора мер, учитывающем транспортную структуру.

Барицентр мер, как среднее в смысле Фреше (т.е. точка минимума взвешенной суммы транспортных расстояний до некоторого набора мер), рассматривается в [1, 2] для

пространства Вассерштейна W_2 , когда задан конкретный вид ценовой функции, — квадратичная. В одномерном случае удается найти барицентр как конечного набора, так и распределения мер в явном виде, при этом допустимый класс ценовых функций достаточно широк.

С точки зрения статистических приложений, естественно рассматривать случайные величины в пространстве мер. Например, если дана последовательность i.i.d. мер или мерозначный случайный процесс, то возникает вопрос о состоятельности эмпирических барицентров, т.е. барицентров первых n мер, в смысле транспортного расстояния. В данной статье доказывается результат, аналогичный теореме 6.1 из [2] (для пространства W_2) — сходимость барицентров по i.i.d. выборке мер к барицентру распределения, опять же, для достаточно широкого класса функций.

Отметим, что в вопросах, связанных с задачей Монжа–Канторовича, широко применяется теория двойственности. Тем не менее в данной статье мы не прибегаем к ней, т.к. в нашем случае все удобно описывать напрямую, используя кумулятивные функции.

Статья построена следующим образом. В разделе 2 дается строгая формулировка транспортной задачи и приводится ее решение в рассматриваемом нами случае. В секции 3 вводится определение барицентра конечного набора мер, доказывается его существование и приводится явная формула для его нахождения. Аналогичные результаты приводятся в разделе 4, но уже для барицентра распределения на пространстве мер. В разделе 5 доказывается вариант закона больших чисел, а именно: слабая сходимость эмпирических барицентров к барицентру распределения. Наконец, в секции 6 устанавливается сильная сходимость эмпирических барицентров, при некоторых дополнительных ограничениях на ценовую функцию. Также в этом разделе кратко рассматриваются свойства топологии, порожденной транспортным расстоянием.

1.1. Используемые обозначения

Введем некоторые обозначения, которые будут использоваться в статье.

Через $\mathcal{P}(X)$ будем обозначать пространство вероятностных мер над измеримым пространством X . Если не оговорено обратное, то подразумевается, что на X введена стандартная борелевская сигма-алгебра $\mathcal{B}(X)$.

Пусть даны измеримые пространства X, Y и измеримое отображение $T: X \rightarrow Y$. Пусть μ — мера над пространством X . Через $T_{\#}\mu$ будем обозначать образ меры (the image measure) при отображении T , т.е. меру над Y , определяемую следующим образом:

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

При этом $\int_Y f(y)d[T_{\#}\mu](y) = \int_X f(T(x))d\mu(x)$.

При интегрировании функций аргумент иногда будет опускаться, если это не может вызвать неоднозначности.

2. Задача Монжа–Канторовича в одномерном случае с выпуклой ценовой функцией

2.1. Формулировка задачи Монжа–Канторовича

Рассмотрим пространство вероятностных мер на числовой оси $\mathcal{P}(\mathbb{R})$. Введем множество транспортных планов (transport plan):

$$\Pi(\mu, \nu) := \left\{ \gamma \in \mathcal{P}(\mathbb{R} \times \mathbb{R}) \mid \pi_{\#}^x \gamma = \mu, \pi_{\#}^y \gamma = \nu \right\},$$

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

Пусть задана произвольная измеримая функция

$$c: \mathbb{R} \times \mathbb{R} \rightarrow \mathbb{R},$$

которую мы будем называть ценовой функцией (cost function).

Задача Монжа–Канторовича состоит в нахождении минимума транспортного функционала:

Задача 1.

$$\inf \left\{ K(\gamma) := \int_{\mathbb{R} \times \mathbb{R}} c(x, y) d\gamma \mid \gamma \in \Pi(\mu, \nu) \right\}.$$

Вообще говоря, задачу Монжа–Канторовича можно ставить для произвольных измеримых пространств, но мы будем рассматривать одномерный случай.

Определение 1. Оптимальным транспортным планом называется решение Задачи 1 $\gamma^* \in \Pi(\mu, \nu)$.

Определение 2. Пусть γ^* — оптимальный транспортный план. Если он имеет вид $\gamma^* = (id \times T)_{\#}\mu$, где $id(\cdot)$ — тождественное отображение, то $T: \mathbb{R} \rightarrow \mathbb{R}$ называется оптимальным транспортным отображением (transport map).

2.2. Случай выпуклой ценовой функции

Предположение 1. Ценовая функция имеет вид $c(x, y) = g(x - y)$, где $g(\cdot)$ — строго выпуклая неотрицательная функция, достигающая своего минимума.

В дальнейшем всегда полагаем, что Предположение 1 выполнено, если не оговорено обратное. Покажем, что при данном предположении решение одномерной Задачи 1 можно найти в явном виде.

Лемма 1. При Предположении 1 ценовая функция удовлетворяет неравенству Монжа, т.е. для любых $x_1 < x_2$, $y_1 < y_2$ выполняется

$$c(x_1, y_1) + c(x_2, y_2) < c(x_1, y_2) + c(x_2, y_1).$$

Доказательство. В силу строгой выпуклости $g(\cdot)$ для любого $\delta > 0$ функция

$$h_\delta(x) := g(x + \delta) - g(x), \quad x \in \mathbb{R}$$

строго возрастает. Выбрав $\delta = y_2 - y_1 > 0$ получаем, что

$$h_\delta(x_1 - y_2) = g(x_1 - y_1) - g(x_1 - y_2) < h_\delta(x_2 - y_2) = g(x_2 - y_1) - g(x_2 - y_2).$$

Т.е. $c(x_1, y_1) - c(x_1, y_2) < c(x_2, y_1) - c(x_2, y_2)$, откуда следует неравенство Монжа.

Теорема 1. Пусть инфимум транспортного функционала конечен. Тогда существует единственный оптимальный транспортный план $\gamma^* = (F_\mu^{-1} \times F_\nu^{-1})_{\#}\mathcal{L}$, где \mathcal{L} — стандартная мера Лебега на интервале $(0, 1)$, $F_\mu(x) := \mu((-\infty, x))$, $x \in \mathbb{R}$ — кумулятивная функция меры μ , а $F^{-1}(y) := \inf\{x \mid F(x) \geq y\}$, $y \in (0, 1)$.

Доказательство. Так как функция $c(\cdot, \cdot)$ непрерывна и ограничена снизу, то по Теореме 1.7 [3] существует оптимальный транспортный план γ^* . Допустим, существуют такие $x_1 < x_2$, $y_1 > y_2$, что (x_1, y_1) и (x_2, y_2) лежат в носителе γ^* . Из неравенства Монжа и непрерывности следует, что найдется $\varepsilon > 0$, для которого

$$\min_{U_1} c(x, y) + \min_{U_2} c(x, y) > \max_{\tilde{U}_1} c(x, y) + \max_{\tilde{U}_2} c(x, y),$$

где $U_1 := [x_1 - \varepsilon, x_1 + \varepsilon] \times [y_1 - \varepsilon, y_1 + \varepsilon]$, $U_2 := [x_2 - \varepsilon, x_2 + \varepsilon] \times [y_2 - \varepsilon, y_2 + \varepsilon]$, $\tilde{U}_1 := [x_1 - \varepsilon, x_1 + \varepsilon] \times [y_2 - \varepsilon, y_2 + \varepsilon]$, $\tilde{U}_2 := [x_2 - \varepsilon, x_2 + \varepsilon] \times [y_1 - \varepsilon, y_1 + \varepsilon]$.

Рассмотрим γ_1 — сужение меры γ^* на U_1 , $\mu_1 := \pi_{\#}^x \gamma_1$ и $\nu_1 := \pi_{\#}^y \gamma_1$. Аналогично определим γ_2 , μ_2 , ν_2 . Будем считать, что $\gamma_1(U_1) = \gamma_2(U_2) = p > 0$. Тогда мера $\tilde{\gamma} = \gamma^* - \gamma_1 - \gamma_2 + \mu_1 \otimes \nu_2 + \mu_2 \otimes \nu_1 \in \Pi(\mu, \nu)$. При этом

$$K(\gamma^*) - K(\tilde{\gamma}) = \int_{U_1} c(x, y) d\gamma_1 + \int_{U_2} c(x, y) d\gamma_2 - \int_{\tilde{U}_1} c(x, y) d(\mu_1 \otimes \nu_2) - \\ - \int_{\tilde{U}_2} c(x, y) d(\mu_2 \otimes \nu_1) \geq (\min_{U_1} c(x, y) + \min_{U_2} c(x, y) - \max_{\tilde{U}_1} c(x, y) - \max_{\tilde{U}_2} c(x, y))p > 0,$$

что противоречит оптимальности γ^* .

Таким образом, носитель γ^* имеет вид $\text{supp } \gamma^* = \{(x(t), y(t)) \mid t \in [a, b]\}$, где функции $x(t)$, $y(t)$ неубывающие. Так как $\gamma^* \in \Pi(\mu, \nu)$, то существует единственное решение:

$$\gamma^* = (F_{\mu}^{-1} \times F_{\nu}^{-1})_{\#} \mathcal{L}.$$

Замечание 1. Если существует оптимальное отображение $T(\cdot)$ (в частности, если мера μ безатомная), то $T(x) = F_{\nu}^{-1}(F_{\mu}(x))$ μ -п.в.

Замечание 2. В случае, когда функция $g(\cdot)$ выпуклая, но не строго, транспортный план из теоремы 1 (и отображение, если оно существует) будет оптимальным, но, возможно, не единственным.

Определим «расстояние» между мерами как минимум транспортного функционала

$$J(\mu, \nu) := \min \left\{ K(\gamma) \mid \gamma \in \Pi(\mu, \nu) \right\} = \int_0^1 g(F_{\mu}^{-1} - F_{\nu}^{-1}) dx.$$

($J(\mu, \nu) = \int_{\mathbb{R}} g(x - T(x)) d\mu(x)$, если существует оптимальное отображение.)

Отметим, что без дополнительных ограничений на ценовую функцию и множество мер функция $J(\cdot, \cdot)$ может принимать любые значения, в том числе $+\infty$, и не удовлетворять неравенству треугольника.

3. Барицентр конечного набора мер

В этом разделе мы хотим определить среднее (в каком-то смысле) для конечного набора мер. Для этого мы поставим задачу минимизации некоторого функционала. Например, для квадратичной функции это задача минимизации момента инерции, т.е. нахождение барицентра.

Пусть заданы меры μ_1, \dots, μ_n из $\mathcal{P}(\mathbb{R})$ и веса $\lambda_1, \dots, \lambda_n$: $\lambda_i > 0$, $i = \overline{1, n}$, $\sum_{i=1}^n \lambda_i = 1$. Рассмотрим следующую задачу:

Задача 2.

$$\min \left\{ \sum_{i=1}^n \lambda_i J(\mu_i, \nu) \mid \nu \in \mathcal{P}(\mathbb{R}) \right\}.$$

Определение 3. Барицентром конечного набора мер $(\mu_i, \lambda_i)_{i=\overline{1, n}}$ будем называть решение Задачи 2 $\nu^* \in \mathcal{P}(\mathbb{R})$, обозначая его $\nu^* = \text{bar}(\mu_i, \lambda_i)_{i=\overline{1, n}}$.

Теорема 2. Пусть существует мера $\nu \in \mathcal{P}(\mathbb{R})$, такая, что $J(\mu_i, \nu) < +\infty \quad \forall i = \overline{1, n}$. Тогда Задача 2 имеет единственное решение $\nu^* = \psi_{\#} \mathcal{L}$, где

$$\psi(x) := \operatorname{argmin}_{y \in \mathbb{R}} \sum_{i=1}^n \lambda_i g(F_{\mu_i}^{-1}(x) - y), \quad x \in (0, 1).$$

Доказательство. Заметим, что $J(\cdot, \cdot)$ ограничена снизу, т.к. $g(\cdot)$ имеет минимум.

Зафиксируем произвольные числа a_1, \dots, a_n и рассмотрим функцию

$$f(x) := \sum_{i=1}^n \lambda_i g(a_i - x), \quad x \in \mathbb{R}.$$

Очевидно, что эта функция также строго выпуклая и имеет минимум. Введем функцию

$$\psi(x) := \operatorname{argmin}_{y \in \mathbb{R}} \sum_{i=1}^n \lambda_i g(F_{\mu_i}^{-1}(x) - y), \quad x \in (0, 1)$$

Покажем, что эта функция неубывающая. Действительно, допустим, $\exists 0 < x_1 < x_2 < 1$: $\psi(x_1) > \psi(x_2)$. Тогда, в силу неравенства Монжа для $g(\cdot)$ и монотонности $F_{\mu_i}^{-1}(\cdot)$,

$$\begin{aligned} \sum_{i=1}^n \lambda_i [g(F_{\mu_i}^{-1}(x_2) - \psi(x_2)) - g(F_{\mu_i}^{-1}(x_2) - \psi(x_1))] &\geq \\ &\geq \sum_{i=1}^n \lambda_i [g(F_{\mu_i}^{-1}(x_1) - \psi(x_2)) - g(F_{\mu_i}^{-1}(x_1) - \psi(x_1))] > 0, \end{aligned}$$

что противоречит определению функции $\psi(\cdot)$.

Тогда $F(\cdot) := \psi^{-1}(\cdot)$ — неубывающая, непрерывная слева (следует из определения), $F(-\infty) = 0$, $F(+\infty) = 1$, следовательно, это кумулятивная функция меры $\nu^* = \psi_{\#} \mathcal{L} \in \mathcal{P}(\mathbb{R})$. При этом $\psi(\cdot)$ непрерывна слева, поэтому $F_{\nu^*}^{-1}(x) = \psi(x) \quad \forall x \in (0, 1)$.

$$\begin{aligned} \forall \nu \in \mathcal{P}(\mathbb{R}) \quad \sum_{i=1}^n \lambda_i J(\mu_i, \nu) &= \sum_{i=1}^n \lambda_i \int_0^1 g(F_{\mu_i}^{-1} - F_{\nu}^{-1}) dx = \\ &= \int_0^1 \sum_{i=1}^n \lambda_i g(F_{\mu_i}^{-1} - F_{\nu}^{-1}) dx \geq \int_0^1 \sum_{i=1}^n \lambda_i g(F_{\mu_i}^{-1} - \psi) dx = \\ &= \int_0^1 \sum_{i=1}^n \lambda_i g(F_{\mu_i}^{-1} - F_{\nu^*}^{-1}) dx = \sum_{i=1}^n \lambda_i J(\mu_i, \nu^*) < +\infty \end{aligned}$$

Следовательно, $\nu^* = \operatorname{bar}(\mu_i, \lambda_i)_{i=1, \dots, n}$, и, в силу единственности минимума строго выпуклой функции, барицентр единственен.

Замечание 3. Если функция $g(\cdot)$ дифференцируема, то, очевидно, $\psi(\cdot)$ можно определять из условия $\sum_{i=1}^n \lambda_i g'(F_{\mu_i}^{-1}(x) - \psi(x)) = 0$. В частности, для случая квадратичной функции $g(x) = x^2$ выполняется $\psi(x) = \sum_{i=1}^n \lambda_i F_{\mu_i}^{-1}(x)$.

Замечание 4. Также можно показать, что если ценовая функция нестрого выпуклая, и $g(x) \rightarrow +\infty$ при $x \rightarrow \infty$, то, переопределив функцию $\psi(\cdot)$, например

$$\psi(x) := \inf_{y \in \mathbb{R}} \operatorname{Argmin}_{i=1}^n \lambda_i g(F_{\mu_i}^{-1}(x) - y), \quad x \in (0, 1)$$

мы аналогично получим меру $\tilde{\nu}$, которая будет являться барицентром, но, возможно, не единственным.

4. Барицентр распределения мер

В пространстве $\mathcal{P}(\mathbb{R})$ введем борелевскую сигма-алгебру $\mathcal{B}(\mathcal{P}(\mathbb{R}))$, порожденную топологией слабой сходимости.

Утверждение 1. *Функция*

$$K(\mu, x) := F_\mu^{-1}(x), \quad \mu \in \mathcal{P}(\mathbb{R}), \quad x \in (0, 1)$$

измерима относительно произведения сигма-алгебр $\mathcal{B}(\mathcal{P}(\mathbb{R})) \otimes \mathcal{B}((0, 1))$.

Доказательство. Зафиксируем произвольный луч $(a, +\infty)$. Надо показать, что $U_a := K^{-1}((a, +\infty))$ — измеримое множество.

Зафиксируем $x \in (0, 1)$ и рассмотрим функцию

$$f_x(\mu) := K(\mu, x) = F_\mu^{-1}(x), \quad \mu \in \mathcal{P}(\mathbb{R}).$$

Для любого $y > a$ рассмотрим произвольную меру $\mu_y: F_{\mu_y}^{-1}(x) = y$. В силу непрерывности слева $\exists \delta > 0: f_{x-\delta}(\mu_y) = F_{\mu_y}^{-1}(x - \delta) > (a + y)/2$. Рассмотрим функцию $v \in BC(\mathbb{R})$, такую, что

$$v(\mathbb{R}) = [0, 1], \quad v(t) = 0, \quad t \leq a, \quad v(t) = 1, \quad t \geq \frac{(a + y)}{2}.$$

Функция $v(\cdot)$ порождает окрестность из слабой топологии

$$V(\mu_y) = \left\{ \nu \in \mathcal{P}(\mathbb{R}) \mid \left| \int_{\mathbb{R}} v d\nu - \int_{\mathbb{R}} v d\mu_y \right| < \delta \right\},$$

для которой $\forall \nu \in V(\mu_y) \quad f_x(\nu) > a$. Следовательно, $f_x^{-1}((a, +\infty))$ — открытое множество.

Из монотонности обратной кумулятивной функции следует, что если $A \times \{x\} \subset U_a$, то и $A \times [x, 1) \subset U_a$. Из непрерывности слева получаем, что $\forall (\mu, x) \in U_a \exists s \in (0, x] \cap \mathbb{Q}: F_\mu^{-1}(s) > a$, значит, $(\mu, s) \in U_a$. Следовательно,

$$U_a = \bigcup_{x \in (0, 1)} f_x^{-1}((a, +\infty)) \times \{x\} = \bigcup_{s \in (0, 1) \cap \mathbb{Q}} f_s^{-1}((a, +\infty)) \times [s, 1).$$

Отсюда U_a — измеримое множество, как счетное объединение измеримых. Таким образом, $K(\cdot, \cdot)$ — измеримая функция.

Пусть μ — случайная величина в $\mathcal{P}(\mathbb{R})$ с распределением P_μ . Тогда, в частности, $\mathbb{E} J(\mu, \nu) = \int_{\mathcal{P}(\mathbb{R})} J(\mu, \nu) dP_\mu \in [0, +\infty]$. Рассмотрим следующую задачу.

Задача 3.

$$\min \left\{ \mathbb{E} J(\mu, \nu) \mid \nu \in \mathcal{P}(\mathbb{R}) \right\}.$$

Определение 4. Барицентром распределения P_μ будем называть решение Задачи 3 $\nu^* \in \mathcal{P}(\mathbb{R})$, обозначая его $\nu^* = \text{bar}(P_\mu)$.

Лемма 2. Пусть $X \in \mathbb{R}$ — случайная величина с распределением P_X . Рассмотрим функцию

$$\xi(x) := \mathbb{E} g(X - x) \in [0, +\infty], \quad x \in \mathbb{R}.$$

Тогда если функция $\xi(\cdot) \not\equiv +\infty$, то она имеет единственный минимум.

Доказательство. Рассмотрим множество

$$I := \left\{ x \in \mathbb{R} \mid \xi(x) < +\infty \right\} \neq \emptyset.$$

Очевидно, функция $\xi(\cdot)$ строго выпуклая, т.к. $g(\cdot)$ строго выпуклая. Следовательно, I представляет собой некоторый промежуток $I = \langle a, b \rangle$.

Покажем, что $\xi(\cdot)$ полунепрерывна снизу. Рассмотрим произвольную точку x_0 .

Случай 1. $\xi(x_0) = +\infty$. Зафиксируем произвольное число C . Тогда существует отрезок $[-A, A]$, для которого $\int_{-A}^A g(t - x_0) dP_X(t) > C$. Функция $g(\cdot)$ непрерывна, следовательно, равномерно непрерывна на $[-A, A]$, поэтому $\exists \delta > 0$, такая, что

$$|x - x_0| < \delta \Rightarrow \xi(x) = \mathbb{E} g(\mathbf{X} - x) \geq \int_{-A}^A g(t - x) dP_X(t) > C,$$

т.е. $\lim_{x \rightarrow x_0} \xi(x) = +\infty = \xi(x_0)$.

Случай 2. $\xi(x_0) < +\infty$. Зафиксируем $\varepsilon > 0$. Тогда $\exists A: \int_{-A}^A g(t - x_0) dP_X(t) > \xi(x_0) - \varepsilon$. Из непрерывности $g(\cdot)$, $\exists \delta > 0$, такая, что

$$|x - x_0| < \delta \Rightarrow \xi(x) = \mathbb{E} g(\mathbf{X} - x) \geq \int_{-A}^A g(t - x) dP_X(t) > \xi(x_0) - \varepsilon.$$

Следовательно, $\lim_{x \rightarrow x_0} \xi(x) \geq \xi(x_0)$.

Отметим, что т.к. $\lim_{x \rightarrow \infty} g(x) = +\infty$, то и $\lim_{x \rightarrow \infty} \xi(x) = +\infty$.

Из выпуклости и полунепрерывности снизу следует, что функция $\xi(\cdot)$ непрерывна на I . Также если промежуток I открыт с какого-то края, то функция $\xi(\cdot)$ там неограниченно растет. Следовательно, $\xi(\cdot)$ достигает на I минимума, причем он единственен в силу строгой выпуклости.

Теорема 3. Пусть $\mu \in \mathcal{P}(\mathbb{R})$ — случайная величина с распределением P_μ , и существует мера $\nu \in \mathcal{P}(\mathbb{R})$, для которой $\mathbb{E} J(\mu, \nu) < +\infty$. Тогда Задача 3 имеет единственное решение $\nu^* = \psi_{\#} \mathcal{L}$, где

$$\psi(x) := \operatorname{argmin}_{y \in \mathbb{R}} \mathbb{E} g(F_\mu^{-1}(x) - y), \text{ для п.в. } x \in (0, 1).$$

Доказательство. Из условий теоремы следует существование множества $M \subset (0, 1)$, такого, что $\mathcal{L}(M) = 1$ и

$$\forall x \in M \exists y_x: \mathbb{E} g(F_\mu^{-1}(x) - y_x) < +\infty.$$

Тогда по Лемме 2 можно определить функцию

$$\psi(x) := \operatorname{argmin}_{y \in \mathbb{R}} \mathbb{E} g(F_\mu^{-1}(x) - y), \quad x \in M,$$

которая, аналогично случаю конечного набора мер, является неубывающей. В остальных точках интервала $(0, 1)$ функцию можно доопределить, например, по непрерывности слева. При этом $\psi(\cdot)$ порождает меру $\nu^* = \psi_{\#} \mathcal{L} \in \mathcal{P}(\mathbb{R})$, и $F_{\nu^*}^{-1}(\cdot) = \psi(\cdot)$ п.в.

$$\begin{aligned} \forall \nu \in \mathcal{P}(\mathbb{R}) \quad \int_{\mathcal{P}(\mathbb{R})} J(\mu, \nu) dP_\mu &= \int_{\mathcal{P}(\mathbb{R})} \int_0^1 g(F_\mu^{-1}(x) - F_\nu^{-1}(x)) dx dP_\mu = \\ &= \int_0^1 \int_{\mathcal{P}(\mathbb{R})} g(F_\mu^{-1}(x) - F_\nu^{-1}(x)) dP_\mu dx \geq \int_0^1 \int_{\mathcal{P}(\mathbb{R})} g(F_\mu^{-1}(x) - \psi(x)) dP_\mu dx = \\ &= \int_{\mathcal{P}(\mathbb{R})} \int_0^1 g(F_\mu^{-1}(x) - F_{\nu^*}^{-1}(x)) dx dP_\mu = \int_{\mathcal{P}(\mathbb{R})} J(\mu, \nu^*) dP_\mu < +\infty. \end{aligned}$$

Таким образом, $\nu^* = \operatorname{var}(P_\mu)$, и барицентр единственен в силу единственности (почти всюду) $\psi(\cdot)$.

5. Слабая сходимость барицентров

Лемма 3. Пусть $\{\mathbf{X}_n\}_{n \in \mathbb{N}} \subset \mathbb{R}$ — последовательность *i.i.d.* случайных величин с распределением P_X , таким, что $\xi(\cdot) \not\equiv +\infty$, где

$$\xi(x) := \mathbb{E} g(\mathbf{X} - x), \quad x \in \mathbb{R}.$$

Определим точку

$$x^* := \operatorname{argmin}_{x \in \mathbb{R}} \xi(x)$$

и случайные величины

$$\bar{\mathbf{X}}_n := \operatorname{argmin}_{x \in \mathbb{R}} \sum_{i=1}^n g(\mathbf{X}_i - x), \quad n \in \mathbb{N}.$$

Тогда $\bar{\mathbf{X}}_n \rightarrow x^*$ с вероятностью 1.

Доказательство. Из Леммы 2 следует, что x^* определен корректно. Допустим, что $\bar{\mathbf{X}}_n \not\rightarrow x^*$ с ненулевой вероятностью. Без ограничения общности будем считать, что существует подпоследовательность $\bar{\mathbf{X}}_{n_k} > x^* + 2\delta$ для некоторого $\delta > 0$. Определим функцию

$$\varepsilon(x) := g(x - x^*) - 2g(x - x^* - \delta) + g(x - x^* - 2\delta) > 0, \quad x \in \mathbb{R}.$$

Заметим, что точки $\mathbf{X}_i - x^* - 2\delta$, $\mathbf{X}_i - \bar{\mathbf{X}}_{n_k} + \delta$ лежат между $\mathbf{X}_i - \bar{\mathbf{X}}_{n_k}$ и $\mathbf{X}_i - x^* - \delta$. Тогда по Лемме 1

$$g(\mathbf{X}_i - x^* - \delta) - g(\mathbf{X}_i - x^* - 2\delta) > g(\mathbf{X}_i - \bar{\mathbf{X}}_{n_k} + \delta) - g(\mathbf{X}_i - \bar{\mathbf{X}}_{n_k}).$$

Следовательно,

$$\begin{aligned} g(\mathbf{X}_i - x^*) - g(\mathbf{X}_i - x^* - \delta) &= g(\mathbf{X}_i - x^* - \delta) - g(\mathbf{X}_i - x^* - 2\delta) + \varepsilon(\mathbf{X}_i) > \\ &> g(\mathbf{X}_i - \bar{\mathbf{X}}_{n_k} + \delta) - g(\mathbf{X}_i - \bar{\mathbf{X}}_{n_k}) + \varepsilon(\mathbf{X}_i). \end{aligned}$$

Суммируя неравенства по $1 \leq i \leq n_k$, получаем

$$\begin{aligned} \frac{1}{n_k} \sum_{i=1}^{n_k} g(\mathbf{X}_i - x^*) - \frac{1}{n_k} \sum_{i=1}^{n_k} g(\mathbf{X}_i - x^* - \delta) &> \frac{1}{n_k} \sum_{i=1}^{n_k} g(\mathbf{X}_i - \bar{\mathbf{X}}_{n_k} + \delta) - \\ &- \frac{1}{n_k} \sum_{i=1}^{n_k} g(\mathbf{X}_i - \bar{\mathbf{X}}_{n_k}) + \frac{1}{n_k} \sum_{i=1}^{n_k} \varepsilon(\mathbf{X}_i) \geq \frac{1}{n_k} \sum_{i=1}^{n_k} \varepsilon(\mathbf{X}_i). \end{aligned}$$

Согласно усиленному ЗБЧ,

$$\frac{1}{n_k} \sum_{i=1}^{n_k} \varepsilon(\mathbf{X}_i) \rightarrow \mathbb{E} \varepsilon(\mathbf{X}) = \varepsilon > 0 \quad \text{п.н.}$$

Следовательно, для достаточно больших k выполняется

$$\frac{1}{n_k} \sum_{i=1}^{n_k} \varepsilon(\mathbf{X}_i) \geq \varepsilon/2.$$

Отсюда,

$$\frac{1}{n_k} \sum_{i=1}^{n_k} g(\mathbf{X}_i - x^* - \delta) \leq \frac{1}{n_k} \sum_{i=1}^{n_k} g(\mathbf{X}_i - x^*) - \frac{\varepsilon}{2}.$$

Но т.к. для любого x

$$\frac{1}{n_k} \sum_{i=1}^{n_k} g(\mathbf{X}_i - x) \rightarrow \mathbb{E} g(\mathbf{X} - x) = \xi(x) \in [0, +\infty] \quad \text{п.н.,}$$

то $\xi(x^* + \delta) \leq \xi(x^*) - \varepsilon/2$, что противоречит определению x^* .

Таким образом, $\bar{\mathbf{X}}_n \rightarrow x^*$ почти наверное.

Теорема 4. Пусть $\{\mu_n\}_{n \in \mathbb{N}} \subset \mathcal{P}(\mathbb{R})$ — последовательность *i.i.d.* случайных величин с распределением P_μ , для которых выполнены условия Теоремы 3. Тогда для эмпирических барицентров $\nu_n := \text{bar}(\mu_i, 1/n)_{i=1, \dots, n}$ имеет место слабая сходимость

$$\nu_n \rightharpoonup \nu^* \quad \text{п.н.},$$

где $\nu^* := \text{bar}(P_\mu)$.

Доказательство. Покажем, что эмпирические барицентры существуют с вероятностью 1. По условию Теоремы 3, существует такая мера $\nu \in \mathcal{P}(\mathbb{R})$, что $\mathbb{E} J(\mu, \nu) < +\infty$. Допустим, для какого-то $n \in \mathbb{N}$ выполняется

$$\mathbb{P}\left\{\sum_{i=1}^n J(\mu_i, \nu) = +\infty\right\} > 0.$$

Тогда, очевидно, $\mathbb{P}\{J(\mu, \nu) = +\infty\} > 0$. Следовательно, $\mathbb{E} J(\mu, \nu) = +\infty$ — противоречие. Таким образом, для любого n имеем

$$\mathbb{P}\left\{\sum_{i=1}^n J(\mu_i, \nu) = +\infty\right\} = 0,$$

откуда по Теореме 2 следует, что барицентр $\nu_n := \text{bar}(\mu_i, 1/n)_{i=1, \dots, n}$ существует почти наверное.

Определим функцию $\psi(\cdot)$ как в Теореме 3:

$$\psi(x) := \underset{y \in \mathbb{R}}{\text{argmin}} \mathbb{E} g(F_\mu^{-1}(x) - y), \quad \text{для п.в. } x \in (0, 1).$$

Из Леммы 3 следует, что для почти всех $x \in (0, 1)$ выполняется $F_{\nu_n}^{-1}(x) \xrightarrow{\text{п.н.}} \psi(x)$. Так как $\psi(\cdot) = F_{\nu^*}^{-1}(\cdot)$ п.в., можно выбрать $M \subset (0, 1)$ — всюду плотное счетное подмножество, такое, что

$$F_{\nu_n}^{-1}(x) \xrightarrow{\text{п.н.}} F_{\nu^*}^{-1}(x) \quad \forall x \in M.$$

В силу счетной аддитивности вероятностной меры, отсюда следует, что почти наверное

$$F_{\nu_n}^{-1}(x) \rightarrow F_{\nu^*}^{-1}(x) \quad \forall x \in M. \quad (1)$$

Будем рассматривать те исходы, когда утверждение (1) выполнено. Покажем, что тогда $F_{\nu_n}(x) \rightarrow F_{\nu^*}(x)$, если $F_{\nu^*}(\cdot)$ непрерывна в точке x . Допустим, это не так, т.е. существует точка x_0 , в которой функция $F_{\nu^*}(\cdot)$ непрерывна и $F_{\nu_n}(x_0) \not\rightarrow F_{\nu^*}(x_0)$. Тогда возможны два случая.

Случай 1. Существует подпоследовательность $\{\nu_{n_k}\}_{k \in \mathbb{N}}$, такая, что $F_{\nu_{n_k}}(x_0) > F_{\nu^*}(x_0) + \varepsilon$ для некоторого $\varepsilon > 0$. Выберем $y \in M$: $F_{\nu^*}(x_0) < y < F_{\nu^*}(x_0) + \varepsilon$. В силу непрерывности $F_{\nu^*}(\cdot)$ существует $\gamma > 0$: $x < x_0 + \gamma \Rightarrow F_{\nu^*}(x) < y$. Тогда $F_{\nu_{n_k}}^{-1}(y) \geq x_0 + \gamma$, а $F_{\nu_{n_k}}^{-1}(y) \leq x_0$ — противоречие с (1).

Случай 2. Существует подпоследовательность ν_{n_k} , такая, что $F_{\nu_{n_k}}^{-1}(x_0) < F_{\nu^*}^{-1}(x_0) - \varepsilon$ для некоторого $\varepsilon > 0$. Аналогично первому случаю, это противоречит (1).

Отсюда следует слабая сходимость $\nu_n \rightarrow \nu^*$ п.н., т.е. с вероятностью 1 выполняется

$$\int_{\mathbb{R}} f d\nu_n \rightarrow \int_{\mathbb{R}} f d\nu^* \quad \forall f \in BC(\mathbb{R}).$$

6. Сильная сходимость барицентров

Потребуем, чтобы функция $g(\cdot)$ достигала минимума в нуле: $g(0) = 0$. Пусть также выполнено следующее предположение.

Предположение 2. Существуют константы $A, B > 0$, такие, что

$$g(x - y) \leq A + B(g(x) + g(y)) \quad \forall x, y, \quad (2)$$

Утверждение 2. Пусть функция $g(\cdot) = \Theta(|x|^p)$ для некоторого $p \geq 1$, т.е. существуют такие $0 < C_1 < C_2$, $x_0 > 0$, что

$$C_1|x|^p < g(x) < C_2|x|^p, \quad |x| > x_0.$$

Тогда выполнено Предположение 2.

Доказательство. Так как $g(\cdot)$ непрерывна, существует $C_0 > 0$, для которой $g(x) < C_0$ при $|x| \leq x_0$. Для любых x, y имеем:

$$\begin{aligned} g(x - y) &< \max\{C_0, C_2|x - y|^p\} < C_0 + C_2|x - y|^p, \\ |x - y|^p &\leq |2x|^p + |2y|^p = 2^p(|x|^p + |y|^p), \\ |x|^p &\leq \max\{x_0^p, \frac{g(x)}{C_1}\} \leq x_0^p + \frac{g(x)}{C_1}. \end{aligned}$$

Таким образом,

$$g(x - y) \leq (C_0 + 2^{p+1}C_2x_0^p) + 2^p \frac{C_2}{C_1} (g(x) + g(y)).$$

Обозначив $A = C_0 + 2^{p+1}C_2x_0^p$ и $B = 2^p C_2/C_1$, получаем условие 2.

Введем следующее отношение между мерами: для любых μ, ν из $\mathcal{P}(\mathbb{R})$ $\mu \sim \nu$, если и только если $J(\mu, \nu) < +\infty$.

Лемма 4. Отношение \sim является отношением эквивалентности на $\mathcal{P}(\mathbb{R})$.

Доказательство

1) Для любого $\mu \in \mathcal{P}(\mathbb{R})$

$$J(\mu, \mu) = \int_0^1 g(F_\mu^{-1} - F_\mu^{-1}) dx = 0 < +\infty.$$

2) Пусть $J(\mu, \nu) < +\infty$, т.к. $g(-x) \leq A + Bg(x)$, то

$$J(\nu, \mu) = \int_0^1 g(F_\nu^{-1} - F_\mu^{-1}) dx \leq A + BJ(\mu, \nu) < +\infty.$$

3) Пусть $J(\mu, \nu) < +\infty$, $J(\nu, \lambda) < +\infty$. Тогда

$$\begin{aligned} J(\mu, \lambda) &= \int_0^1 g(F_\mu^{-1} - F_\lambda^{-1}) dx \leq \int_0^1 [A + Bg(F_\mu^{-1} - F_\nu^{-1}) + \\ &\quad + Bg(F_\lambda^{-1} - F_\nu^{-1})] dx = A + BJ(\mu, \nu) + BJ(\lambda, \nu) < +\infty. \end{aligned}$$

Следовательно, отношение \sim рефлексивно, симметрично и транзитивно, т.е. это отношение эквивалентности.

Таким образом, пространство $\mathcal{P}(\mathbb{R})$ разбивается на классы эквивалентности

$$C(\nu) := \left\{ \mu \in \mathcal{P}(\mathbb{R}) \mid \mu \sim \nu \right\}.$$

При этом, т.к. для любого $\nu \in \mathcal{P}(\mathbb{R})$ функция $f(\mu) := J(\nu, \mu)$ измерима относительно борелевской сигма-алгебры, порожденной топологией слабой сходимости в $\mathcal{P}(\mathbb{R})$, то и любой класс $C(\cdot)$ — измеримое множество. Отметим, что для ценовой функции $g(x) := |x|^p$, $p \geq 1$ соответствующее пространство Вассерштейна $W_p(\mathbb{R}) = C(\delta(0))$, т.е. это меры, эквивалентные дельта-мере.

Утверждение 3. *Классов эквивалентности несчетное число.*

Доказательство. Рассмотрим меру ν , для которой $\int_0^1 |F_\nu^{-1}| dx = +\infty$. Определим меры ν_α для $\alpha \geq 0$, такие, что $F_{\nu_\alpha}^{-1}(\cdot) := \alpha F_\nu^{-1}(\cdot)$. Пусть $\alpha \neq \beta$, тогда

$$\int_0^1 |F_{\nu_\alpha}^{-1} - F_{\nu_\beta}^{-1}| dx = |\alpha - \beta| \int_0^1 |F_\nu^{-1}| dx = +\infty.$$

Так как функция $g(\cdot)$ строго выпукла, то она растет не медленнее линейной. Значит, существуют такие константы $C_1 > 0$, C_2 , что $g(x) \geq C_1|x| + C_2$ для всех $x \in \mathbb{R}$. Следовательно, $J(\nu_\alpha, \nu_\beta) = +\infty$, т.е. $C(\nu_\alpha) \neq C(\nu_\beta)$.

Лемма 5. 1) Пусть заданы меры μ_1, \dots, μ_n из $\mathcal{P}(\mathbb{R})$ и положительные веса $\lambda_1, \dots, \lambda_n$. Баричесентр данного набора $\text{bar}(\mu_i, \lambda_i)_{1 \leq i \leq n}$ существует тогда и только тогда, когда меры лежат в одном классе эквивалентности, т.е. $\mu_1 \sim \mu_2 \sim \dots \sim \mu_n$.

2) Пусть $\mu \in \mathcal{P}(\mathbb{R})$ — случайная величина с распределением P_μ . Тогда для существования баричесентра $\text{bar}(P_\mu)$ необходимо, чтобы нашелся класс $C(\nu)$, такой, что $P_\mu(C(\nu)) = 1$, т.е. вся вероятностная мера сосредоточена в одном классе.

Доказательство

1) Пусть меры μ_1, \dots, μ_n лежат в разных классах эквивалентности. Тогда, очевидно,

$$\sum_{i=1}^n \lambda_i J(\mu_i, \nu) \equiv +\infty, \quad \nu \in \mathcal{P}(\mathbb{R}),$$

и баричесентра не существует.

Если же $\mu_1 \sim \mu_2 \sim \dots \sim \mu_n$, то

$$\sum_{i=1}^n \lambda_i J(\mu_i, \mu_1) < +\infty,$$

и по Теореме 2 существует единственный баричесентр $\text{bar}(\mu_i, \lambda_i)_{1 \leq i \leq n}$.

2) Пусть для любой меры ν выполняется $P_\mu(C(\nu)) < 1$. Тогда

$$P_\mu(\mathcal{P}(\mathbb{R}) \setminus C(\nu)) = P\{\mu \notin C(\nu)\} = P\{J(\mu, \nu) + \infty\} > 0.$$

Следовательно,

$$\mathbb{E} J(\mu, \nu) \equiv +\infty, \quad \nu \in \mathcal{P}(\mathbb{R}),$$

поэтому не существует баричесентра распределения $\text{bar}(P_\mu)$.

Зафиксируем меру $\nu_0 \in \mathcal{P}(\mathbb{R})$ и в дальнейшем будем рассматривать меры из класса эквивалентности $C(\nu_0)$.

Лемма 6. Пусть дана мера $\mu^* \in C(\nu_0)$ и последовательность $\{\mu_n\}_{n \in \mathbb{N}} \subset C(\nu_0)$, такая, что $J(\mu_n, \mu^*) \rightarrow 0$. Тогда μ_n слабо сходятся к μ^* .

Доказательство. Допустим, μ_n не сходятся слабо к μ^* . Тогда существует $x_0 \in (0, 1)$ — точка непрерывности $F_{\mu^*}^{-1}(\cdot)$, в которой $F_{\mu_n}^{-1}(x_0) \not\rightarrow F_{\mu^*}^{-1}(x_0)$. Без ограничения общности будем считать, что существует подпоследовательность $\{\mu_{n_k}\}$ и число $\delta > 0$, такие, что $F_{\mu_{n_k}}^{-1}(x_0) \geq F_{\mu^*}^{-1}(x_0) + 2\delta$. Из монотонности $F_{\mu_{n_k}}^{-1}(\cdot)$ и непрерывности $F_{\mu^*}^{-1}(\cdot)$ в x_0 следует, что в некоторой окрестности $[x_0, x_0 + \gamma]$ выполняется $F_{\mu_{n_k}}^{-1}(x) \geq F_{\mu^*}^{-1}(x) + \delta$. Отсюда получаем, что

$$J(\mu_{n_k}, \mu^*) \geq \int_{x_0}^{x_0+\gamma} g(F_{\mu_{n_k}}^{-1} - F_{\mu^*}^{-1}) dx \geq \int_{x_0}^{x_0+\gamma} g(\delta) dx > 0,$$

что противоречит условию $J(\mu_n, \mu^*) \rightarrow 0$.

Теорема 5. Пусть дана последовательность $\{\mu_n\}_{n \in \mathbb{N}} \subset C(\nu_0)$ и мера $\mu^* \in C(\nu_0)$. Тогда следующие условия эквивалентны:

- 1) $\mu_n \rightharpoonup \mu^*$ и для любой меры $\nu \in C(\nu_0)$ выполняется $J(\mu_n, \nu) \rightarrow J(\mu^*, \nu)$.
- 2) $\mu_n \rightharpoonup \mu^*$ и существует мера $\nu \in C(\nu_0)$, такая, что $J(\mu_n, \nu) \rightarrow J(\mu^*, \nu)$.
- 3) $J(\mu_n, \mu^*) \rightarrow 0$.

Доказательство

- 1) Покажем, что 2) \Rightarrow 3).

В силу слабой сходимости для произвольного $\varepsilon > 0$

$$\int_{\varepsilon}^{1-\varepsilon} g(F_{\mu_n}^{-1} - F_{\nu}^{-1}) dx \rightarrow \int_{\varepsilon}^{1-\varepsilon} g(F_{\mu^*}^{-1} - F_{\nu}^{-1}) dx.$$

Так как

$$\begin{aligned} J(\mu_n, \nu) &= \int_0^{\varepsilon} g(F_{\mu_n}^{-1} - F_{\nu}^{-1}) dx + \int_{\varepsilon}^{1-\varepsilon} g(F_{\mu_n}^{-1} - F_{\nu}^{-1}) dx + \int_{1-\varepsilon}^1 g(F_{\mu_n}^{-1} - F_{\nu}^{-1}) dx \rightarrow \\ &\rightarrow J(\mu^*, \nu) = \int_0^{\varepsilon} g(F_{\mu^*}^{-1} - F_{\nu}^{-1}) dx + \int_{\varepsilon}^{1-\varepsilon} g(F_{\mu^*}^{-1} - F_{\nu}^{-1}) dx + \int_{1-\varepsilon}^1 g(F_{\mu^*}^{-1} - F_{\nu}^{-1}) dx, \end{aligned}$$

то

$$\begin{aligned} \int_0^{\varepsilon} g(F_{\mu_n}^{-1} - F_{\nu}^{-1}) dx + \int_{1-\varepsilon}^1 g(F_{\mu_n}^{-1} - F_{\nu}^{-1}) dx \rightarrow \\ \rightarrow \int_0^{\varepsilon} g(F_{\mu^*}^{-1} - F_{\nu}^{-1}) dx + \int_{1-\varepsilon}^1 g(F_{\mu^*}^{-1} - F_{\nu}^{-1}) dx \rightarrow 0 \text{ при } \varepsilon \rightarrow 0. \end{aligned}$$

Откуда получаем, что

$$\begin{aligned} \int_0^{\varepsilon} g(F_{\mu_n}^{-1} - F_{\mu^*}^{-1}) dx + \int_{1-\varepsilon}^1 g(F_{\mu_n}^{-1} - F_{\mu^*}^{-1}) dx \leq \\ \leq \int_0^{\varepsilon} [A + Bg(F_{\mu_n}^{-1} - F_{\nu}^{-1}) + Bg(F_{\mu^*}^{-1} - F_{\nu}^{-1})] dx + \\ + \int_{1-\varepsilon}^1 [A + Bg(F_{\mu_n}^{-1} - F_{\nu}^{-1}) + Bg(F_{\mu^*}^{-1} - F_{\nu}^{-1})] dx \xrightarrow{n \rightarrow \infty} \\ \xrightarrow{n \rightarrow \infty} 2A\varepsilon + 2B \left[\int_0^{\varepsilon} g(F_{\mu^*}^{-1} - F_{\nu}^{-1}) dx + \int_{1-\varepsilon}^1 g(F_{\mu^*}^{-1} - F_{\nu}^{-1}) dx \right] \rightarrow 0 \text{ при } \varepsilon \rightarrow 0. \end{aligned}$$

Опять же, благодаря слабой сходимости

$$\int_{\varepsilon}^{1-\varepsilon} g(F_{\mu_n}^{-1} - F_{\mu^*}^{-1}) dx \rightarrow 0$$

для любого $\varepsilon > 0$.

Таким образом, получаем

$$J(\mu_n, \mu^*) = \int_0^1 g(F_{\mu_n}^{-1} - F_{\mu^*}^{-1}) dx \rightarrow 0.$$

2) Покажем, что 3) \Rightarrow 1). Зафиксируем произвольную меру $\nu \in C(\nu_0)$.

По Лемме 6 из сходимости $J(\mu_n, \mu^*) \rightarrow 0$ всегда следует слабая сходимость. Рассмотрим

$$\begin{aligned} \int_0^\varepsilon g(F_{\mu_n}^{-1} - F_\nu^{-1}) dx &\leq \int_0^\varepsilon [A + Bg(F_{\mu_n}^{-1} - F_{\mu^*}^{-1}) + Bg(F_\nu^{-1} - F_{\mu^*}^{-1})] dx \leq \\ &\leq \int_0^\varepsilon [A + Bg(F_{\mu_n}^{-1} - F_{\mu^*}^{-1}) + AB + B^2g(F_{\mu^*}^{-1} - F_\nu^{-1})] dx \xrightarrow{n \rightarrow \infty} \\ &\xrightarrow{n \rightarrow \infty} A(1+B)\varepsilon + B^2 \int_0^\varepsilon g(F_{\mu^*}^{-1} - F_\nu^{-1}) dx \rightarrow 0 \text{ при } \varepsilon \rightarrow 0. \end{aligned}$$

Аналогичный результат имеем для $\int_{1-\varepsilon}^1 g(F_{\mu_n}^{-1} - F_\nu^{-1}) dx$. Используя слабую сходимость, получаем

$$\int_\varepsilon^{1-\varepsilon} g(F_{\mu_n}^{-1} - F_\nu^{-1}) dx \xrightarrow{n \rightarrow \infty} \int_\varepsilon^{1-\varepsilon} g(F_{\mu^*}^{-1} - F_\nu^{-1}) dx \rightarrow J(\mu^*, \nu) \text{ при } \varepsilon \rightarrow 0.$$

Суммируя интегралы по отрезкам $[0, \varepsilon]$, $[\varepsilon, 1 - \varepsilon]$ и $[1 - \varepsilon, 1]$, имеем

$$J(\mu_n, \nu) \rightarrow J(\mu^*, \nu).$$

3) Очевидно, 1) \Rightarrow 2). Следовательно, все три условия эквивалентны.

Замечание 5. В любом из условий 1) – 3) аргументы $J(\cdot, \cdot)$ можно поменять местами, не уменьшая общности. В частности, $J(\mu_n, \mu^*) \rightarrow 0$ эквивалентно $J(\mu^*, \mu_n) \rightarrow 0$.

Замечание 6. В условиях 1) и 2) достаточно $\overline{\lim}_{n \rightarrow \infty} J(\mu_n, \nu) \leq J(\mu^*, \nu)$ (за счет слабой сходимости это эквивалентно $J(\mu_n, \nu) \rightarrow J(\mu^*, \nu)$).

Замечание 7. Так как $J(\mu_n, \mu^*) \rightarrow 0 \Rightarrow J(\mu_n, \nu) \rightarrow J(\mu^*, \nu)$ для любого $\nu \in C(\nu_0)$, то шары

$$B_r(\nu) := \left\{ \mu \in C(\nu_0) \mid J(\mu, \nu) < r \right\}$$

образуют базу некоторой топологии τ_g в $C(\nu_0)$. Очевидно, эта топология не слабее, чем топология слабой сходимости.

Теорема 6. Пусть $\{\mu_n\}_{n \in \mathbb{N}} \subset C(\nu_0)$ – последовательность *i.i.d.* случайных величин с распределением P_μ , для которых выполнены условия Теоремы 3. Тогда для эмпирических барцентров $\nu_n := \text{bar}(\mu_i, 1/n)_{i=1, \overline{n}}$ имеет место сильная сходимость

$$\lim_{n \rightarrow \infty} J(\nu_n, \nu^*) = \lim_{n \rightarrow \infty} J(\nu^*, \nu_n) = 0 \quad \text{п.н.},$$

где $\nu^* := \text{bar}(P_\mu)$.

Доказательство. Будем рассматривать функцию $\psi(\cdot)$ из Теоремы 4. Тогда

$$\mathbb{E} J(\mu, \nu^*) = \int_0^1 \mathbb{E} g(F_\mu^{-1} - \psi) dx < +\infty.$$

Следовательно,

$$\lim_{k \rightarrow \infty} \int_0^{\frac{1}{k}} \mathbb{E} g(F_\mu^{-1} - \psi) dx = \lim_{k \rightarrow \infty} \int_{1-\frac{1}{k}}^1 \mathbb{E} g(F_\mu^{-1} - \psi) dx = 0.$$

По УЗБЧ, почти наверное для всех $k \in \mathbb{N}$

$$\frac{1}{n} \sum_{i=1}^n \int_0^{\frac{1}{k}} g(F_{\mu_i}^{-1} - \psi) dx \xrightarrow{n \rightarrow \infty} \mathbb{E} \int_0^{\frac{1}{k}} g(F_\mu^{-1} - \psi) dx < +\infty.$$

Из Предположения 2 и вида эмпирических барицентров (Теорема 2) получаем, что п.н.

$$\begin{aligned} \int_0^{\frac{1}{k}} g(F_{\nu_n}^{-1} - \psi) dx &\leq \frac{A}{k} + B \int_0^{\frac{1}{k}} \frac{1}{n} \sum_{i=1}^n (g(F_{\mu_i}^{-1} - F_{\nu_n}^{-1}) + g(F_{\mu_i}^{-1} - \psi)) dx \leq \\ &\leq \frac{A}{k} + \frac{2B}{n} \sum_{i=1}^n \int_0^{\frac{1}{k}} g(F_{\mu_i}^{-1} - \psi) dx \xrightarrow{n \rightarrow \infty} \frac{A}{k} + 2B \mathbb{E} \int_0^{\frac{1}{k}} g(F_{\mu}^{-1} - \psi) dx \xrightarrow{k \rightarrow \infty} 0. \end{aligned}$$

Аналогично,

$$\begin{aligned} \int_{1-\frac{1}{k}}^1 g(F_{\nu_n}^{-1} - \psi) dx &\leq \frac{A}{k} + \frac{2B}{n} \sum_{i=1}^n \int_{1-\frac{1}{k}}^1 g(F_{\mu_i}^{-1} - \psi) dx \xrightarrow{n \rightarrow \infty} \\ &\xrightarrow{n \rightarrow \infty} \frac{A}{k} + 2B \mathbb{E} \int_{1-\frac{1}{k}}^1 g(F_{\mu}^{-1} - \psi) dx \xrightarrow{k \rightarrow \infty} 0. \end{aligned}$$

Так как по Теореме 4 $\nu_n \rightarrow \nu^*$ п.н., то с вероятностью 1

$$\int_{\frac{1}{k}}^{1-\frac{1}{k}} g(F_{\nu_n}^{-1} - \psi) dx \xrightarrow{n \rightarrow \infty} \int_{\frac{1}{k}}^{1-\frac{1}{k}} g(F_{\nu^*}^{-1} - \psi) dx = 0 \quad \forall k \in \mathbb{N}.$$

Таким образом, получаем, что почти наверное

$$J(\nu_n, \nu^*) = \int_0^1 g(F_{\nu_n}^{-1} - \psi) dx \rightarrow 0,$$

и аналогично,

$$J(\nu^*, \nu_n) \rightarrow 0.$$

Литература

1. *Agueh M., Carlier G.* Barycenters in the Wasserstein space // SIAM Journal on Mathematical Analysis. 2011. V. 43(2). P. 904–924.
2. *Bigot J., Klein T.* Characterization of barycenters in the Wasserstein space by averaging optimal transport maps. arXiv:1212.2562v5. 2015.
3. *Santambrogio F.* Optimal Transport for Applied Mathematicians. Basel: Birkhauser, 2015.

References

1. *Agueh M., Carlier G.* Barycenters in the Wasserstein space. SIAM Journal on Mathematical Analysis. 2011. V. 43(2). P. 904–924.
2. *Bigot J., Klein T.* Characterization of barycenters in the Wasserstein space by averaging optimal transport maps. arXiv:1212.2562v5. 2015.
3. *Santambrogio F.* Optimal Transport for Applied Mathematicians. Basel: Birkhauser, 2015.

Поступила в редакцию 27.05.2016