Math-Net.Ru

Общероссийский математический портал

Д. А. Заев, О задаче Монжа–Канторовича с дополнительными линейными ограничениями, *Матем. заметки*, 2015, том 98, выпуск 5, 664–683

DOI: <https://doi.org/10.4213/mzm10896>

Использование Общероссийского математического портала Math-Net.Ru подразумевает, что вы прочитали и согласны с пользовательским соглашением <http://www.mathnet.ru/rus/agreement>

Параметры загрузки:

IP: 54.92.164.108

26 апреля 2023 г., 14:44:34

О задаче Монжа–Канторовича с дополнительными линейными ограничениями

Д. А. Заев

В работе рассматривается задача Монжа–Канторовича с дополнительным ограничением: допустимый транспортный план должен обращаться в нуль на некотором фиксированном подпространстве функций. Различный выбор подпространств порождает различные дополнительные условия на транспортные планы. Наши основные результаты сформулированы в общем виде и распространяются на ряд важных частных случаев. В том числе, они верны для задачи Монжа–Канторовича, решаемой в классе инвариантных или мартингалльных мер. Мы формулируем и доказываем критерий существования оптимального решения, утверждение о двойственности типа Канторовича и необходимое геометрическое условие на носитель оптимальной меры, которое аналогично стандартному условию c -монотонности.

Библиография: 15 названий.

DOI: 10.4213/mzm10896

Введение. Зафиксируем вероятностные меры μ_k на польских пространствах X_k и выберем подпространство W некоторого подходящего функционального пространства на $X_1 \times \cdots \times X_n$. Рассмотрим следующую задачу оптимизации, впервые предложенную Л. В. Канторовичем, но в некотором варианте восходящую к Г. Монжу (см. [1], [2]):

$$\inf \left\{ \int_{X_1 \times \cdots \times X_n} c \, d\pi : (\text{Pr}_k)_\# \pi = \mu_k, \int \omega \, d\pi = 0 \text{ для всех } \omega \in W \right\}$$

для некоторой функции стоимости $c: X_1 \times \cdots \times X_n \rightarrow \mathbb{R}$. Эту задачу можно назвать задачей Монжа–Канторовича (или же задачей Канторовича) с дополнительными линейными ограничениями. Согласно стандартной терминологии, принятой в теории Монжа–Канторовича, меры π на $X_1 \times \cdots \times X_n$ с заданными маргиналами $(\text{Pr}_k)_\# \pi = \mu_k$ называются *транспортными планами*. В нашем случае мы ограничиваем множество транспортных планов, добавляя следующее ограничение:

$$\int \omega \, d\pi = 0 \quad \text{для всех } \omega \in W.$$

Такие ограничения очевидно линейны.

В п. 1 мы приводим точную формулировку задачи и определяем специальный класс функций C_L , содержащий подпространство W , возникающее во всех наших примерах. Применяя стандартную технику теории меры, мы доказываем простой критерий существования решения. Более точно, доказывается, что при соответствующих предположениях о регулярности функции стоимости, оптимальный транспортный план существует тогда и только тогда, когда множество допустимых планов непусто.

Мотивацией для нашего исследования служат приложения в статистической физике и финансовой математике, которые приводят к формулировке задачи Монжа–Канторовича, где множество допустимых транспортных планов ограничено тем или иным образом. Примерами таких ограничений являются требование инвариантности по отношению к действию некоторой группы или требование быть мартингалом. Оба примера могут быть рассмотрены как частные случаи общей задачи, поставленной выше.

В работе разрабатывается общий подход к исследованию данной задачи и приводятся результаты, которые могут рассматриваться как обобщения соответствующих утверждений о классической задаче Монжа–Канторовича (современная теория классической задачи подробно описана, например, в [1], [2]). Одним из таких утверждений является следующий результат о двойственности:

$$\inf_{\pi \in \Pi_W} \int c \, d\pi = \sup_{\omega \in W} \left\{ \sum_{k=1}^n \int_{X_k} f_k(x_k) \, d\mu_k, \sum_{k=1}^n f_k + \omega \leq c \right\}.$$

Более точное утверждение и его доказательство представлены в п. 2.

Другой результат, касающийся рассматриваемой задачи, это геометрическое свойство носителя оптимального транспортного плана. Хорошо известно, что решение стандартной транспортной задачи должно лежать на c -монотонном множестве. Мы формулируем похожее свойство, определение которого будет зависеть не только от функции стоимости c , но и от пространства W , и называем его (c, W) -монотонностью. Необходимость подобного свойства для носителей оптимальных планов доказывается в п. 3 как следствие утверждения о двойственности. Независимо подобный результат был сформулирован и доказан в [3], но доказательство, приведенное там, отличается от нашего: оно независимо от утверждения двойственности и использует некоторые тонкие результаты абстрактной теории меры.

Один из наиболее интересных примеров линейного ограничения задается условием мартингалности, которое естественным образом возникает в финансовых приложениях. Любая вероятностная модель зависимости цены от времени для некоторого финансового инструмента может быть рассмотрена как транспортный план с несколькими фиксированными маргиналами (много-маргинальный план). Из теории финансовых рынков известно, что нужно ограничить множество допустимых распределений условием “честной игры”: в каждый конкретный момент времени ожидание цены в следующий момент равно ее текущему значению. Такое ограничение называется “условием мартингалности”, и оно дополнительно ограничивает множество допустимых транспортных планов, которые могут быть использованы для моделирования цены финансового инструмента. В качестве ссылки по этой тематике мы предлагаем работу [4], где описываются финансовые модели, основанные на оптимизации в множестве мартингалных транспортных планов. В [5], [6]

развита математическая теория для транспортной задачи с мартингалным ограничением, и получен ряд довольно сильных результатов. В п. 4 мы покажем, что условие мартингалности является линейным, и выведем ряд известных результатов о мартингалной задаче из общей теории.

Другой пример интересных линейных ограничений – условие инвариантности относительно непрерывного действия некоторой группы. Такие ограничения могут естественным образом возникать в эргодической теории (см., например, [7]) или в геометрии (см. [8]). Если функция стоимости инвариантна, то известно, что решения классической задачи Монжа–Канторовича также инвариантны [8]. В этом случае ограничение инвариантности никак не изменяет задачу. Однако в том случае, когда функция стоимости не инвариантна, такое условие существенно влияет на свойства задачи и ее решения. В п. 5 мы выводим из общей теории некоторые новые результаты, касающиеся этой задачи, и, в особенности, случая компактной группы симметрий.

1. Формулировка задачи. Пусть $X_1, \dots, X_n$ – польские пространства с борелевскими σ -алгебрами на них, $X := X_1 \times \dots \times X_n$, $\mu_1, \dots, \mu_n$ – фиксированные вероятностные меры на $X_1, \dots, X_n$ соответственно, $\mu := (\mu_1, \dots, \mu_n)$ – набор таких мер. Мы будем обозначать через $\mathcal{P}(X)$ множество борелевских вероятностных мер на X , а через $\Pi(\mu)$ множество мер на X с заданными маргиналами. Оба пространства снабдим топологией слабой сходимости.

Определим следующие функциональные пространства:

$$C_L(\mu_i) = \{f \in L^1(X_i, \mu_i) \cap C(X_i)\}$$

– пространство непрерывных интегрируемых функций с топологией, индуцированной из $L^1(X_i, \mu_i)$ нормой, а также

$$C_L(\mu) = \left\{ h \in C(X) : \text{найдется } f = \sum_{i=1}^n f_i(x_i) \in \bigoplus_{i=1}^n C_L(\mu_i) \text{ такие, что } |h| \leq f \right\},$$

снабженное следующей полунормой:

$$\|h\|_L := \sup_{\pi \in \Pi} \int |h| d\pi.$$

Нужно заметить, что очень похожее функциональное пространство определяется и рассматривается в работе Вершика, Петрова и Затитского [9]. Там вместо непрерывности требуется лишь измеримость функций, и такие функции получили название “почти непрерывных”. Многие результаты нашей работы могут быть сформулированы в терминах почти непрерывных функций, но мы будем придерживаться наших определений.

ПРЕДЛОЖЕНИЕ 1.1. *Функция $\|h\|_L$ является полунормой.*

ДОКАЗАТЕЛЬСТВО. Эта функция очевидно принимает конечные значения на рассматриваемом пространстве, ибо

$$\sup_{\pi \in \Pi} \int |h| d\pi \leq \sum \int f_i d\mu_i < \infty,$$

неотрицательна и мультипликативна относительно неотрицательных скаляров. Необходимо проверить лишь субаддитивность: для любых $h, g \in C_L(\mu)$ имеем

$$\sup_{\pi \in \Pi} \int |h + g| d\pi \leq \sup_{\pi \in \Pi} \int |h| d\pi + \sup_{\pi \in \Pi} \int |g| d\pi = \|h\|_L + \|g\|_L,$$

что завершает доказательство.

ЗАМЕЧАНИЕ 1.2. Легко видеть, что любая функция f из $C_L(\mu)$ интегрируема относительно транспортного плана $\pi \in \Pi(\mu)$:

$$\int |f| d\pi \leq \sup_{\gamma \in \Pi} \int |f| d\gamma = \|f\| < \infty.$$

Введем обозначение $F := \bigoplus_{i=1}^n C_L(\mu_i) \subset C_L(\mu)$. Зафиксируем произвольное подпространство $W \subset C_L(\mu)$ и функцию $c \in C_L(\mu)$. Предметом нашего исследования является следующая модификация задачи Монжа–Канторовича:

$$\inf_{\pi \in \Pi_W} \left\{ \int_X c(x_1, \dots, x_n) d\pi \right\},$$

где инфимум вычисляется по множеству транспортных планов со свойством $\pi|_W = 0$:

$$\Pi_W = \left\{ \pi \in \mathcal{P}(X) : \text{для всех } w \in W \int w d\pi = 0, \Pr_{\#}(\pi) = \mu \right\},$$

$\Pr := (\Pr_1, \dots, \Pr_n)$ – набор естественных проекций из X на набор пространств $(X_1, \dots, X_n)$.

Предположим, что $W \ni w = f_i \circ \Pr_i$, где $f_i \in C_L(\mu_i)$. Тогда

$$0 = \int_X (f_i \circ \Pr_i) d\pi = \int_{X_i} f_i d\mu_i.$$

Следовательно, мы получаем необходимое условие на меры $\mu \in \mathcal{P}(X)$, чтобы существовал хотя бы один транспортный план в $\Pi_W(\mu)$:

$$\int_X f_i d\mu_i = 0, \quad \text{если } (f \circ \Pr_i) \in W.$$

Другими словами: $\mu_k|_{W \cap C_L(\mu_k)} = 0$ для каждого $k = 1, \dots, n$.

Теперь мы готовы дать формулировку центрального объекта нашего исследования.

ЗАДАЧА 1.3 (Задача Монжа–Канторовича с дополнительными линейными ограничениями). При некоторых заданных польских пространствах $X = X_1 \times \dots \times X_n$, мерах $\mu_i \in \mathcal{P}(X_i)$, функции стоимости $c \in C_L(\mu)$ и линейном подпространстве $W \subset C_L(\mu)$, где $F := \bigoplus_{i=1}^n C_L(\mu_i) \subset C_L(\mu)$, нужно найти

$$\inf_{\pi \in \Pi_W(\mu)} \left\{ \int_X c(x) d\pi \right\}.$$

Докажем важный факт о пространстве $C_L(\mu)$: естественное вложение $C_b(X)$ в $C_L(\mu)$ плотно.

ЛЕММА 1.4. Множество $C_b(X)$ плотно в $C_L(\mu)$.

ДОКАЗАТЕЛЬСТВО. Во-первых, можно показать, что $\|\cdot\|_{C_L} \leq \|\cdot\|_{C_b}$; следовательно, естественное вложение $C_b(X) \hookrightarrow C_L(\mu): f \mapsto f$ непрерывно. Для любых $h \in C_b(X)$

$$\|h\|_L = \sup_{\pi \in \Pi} \int |h| d\pi \leq \sup_{x \in X} |h(x)| \cdot \sup_{\pi \in \Pi} \int d\pi = \sup_{x \in X} |h(x)| = \|h\|_{C_b}.$$

Чтобы сделать следующий шаг в доказательстве, зафиксируем произвольную функцию $h \in C_L$. Пусть

$$|h| \leq f \in F = \bigoplus_{i=1}^n C_L(\mu_i), \quad k \in \mathbb{N},$$

$$h^k = \min\{k, h\}, \quad h_k^k = \max\{\min\{k, h\}, -k\} \in C_b(X).$$

Заметим, что $|h^k| \leq f$. Наша цель – показать, что $\|h - h_k^k\|_L \rightarrow 0$ при $k \rightarrow \infty$. Имеем

$$\|h - h_k^k\|_L \leq \|h - h^k\|_L + \|h^k - h_k^k\|_L.$$

Тот факт, что

$$\pm h - k \leq f - k = \sum_{i=1}^n \left(f_i - \frac{k}{n} \right),$$

вместе с положительностью оператора $(\cdot)_+ := \max\{\cdot, 0\}$ влечет, что

$$\|h - h^k\|_L = \sup_{\pi \in \Pi} \int ((h - k)_+) d\pi \leq \sum_{i=1}^n \int \left(f_i - \frac{k}{n} \right)_+ d\mu_i \rightarrow 0 \quad \text{при } k \rightarrow \infty,$$

$$\|h^k - h_k^k\|_L = \sup_{\pi \in \Pi} \int ((-h^k - k)_+) d\pi \leq \sum_{i=1}^n \int \left(f_i - \frac{k}{n} \right)_+ d\mu_i \rightarrow 0 \quad \text{при } k \rightarrow \infty.$$

Сходимость здесь обусловлена теоремой Лебега о мажорируемой сходимости: для любых $k \in \mathbb{N}$ $(f_i - k/n)_+ \leq |f_i| \in L_1(X_i, \mu_i)$ и $(f_i - k/n)_+ \rightarrow 0$ поточечно.

Известное следствие теоремы Прохорова состоит в том, что множество $\Pi(\mu)$ компактно в топологии слабой сходимости. Очевидно, что $\Pi_W = \{\pi : \pi|_W = 0\}$ замкнуто в такой топологии. Следовательно, множество $\Pi_W(\mu)$ компактно. Чтобы получить утверждение о существовании, нам нужно доказать непрерывность следующего функционала: $\pi \rightarrow \int h d\pi$. К счастью, этот факт напрямую следует из только что доказанной леммы.

СЛЕДСТВИЕ 1.5. Функционал $\pi \rightarrow \int h d\pi$, действующий из $\Pi(\mu)$ в $\mathbb{R}$, непрерывен для любой функции $h \in C_L(\mu)$.

ДОКАЗАТЕЛЬСТВО. Нам нужно проверить, что для всякой последовательности транспортных планов (π_k) , такой что $\lim_k \int \rho d\pi_k = \int \rho d\pi$ для любого $\rho \in C_b$, верно, что $\lim_k \int h d\pi_k = \int h d\pi$. Так как C_b плотно в C_L , существует такая последовательность, что $\rho_n \rightarrow h$ в $\|\cdot\|_L$ топологии и $\rho_n \in C_b$ для любого $n \in \mathbb{N}$. Заметим, что сходимость $\rho_n \rightarrow h$ в $\|\cdot\|_L$ означает, что последовательность сходится к пределу

равномерно в множестве всех транспортных планов. Этот факт вместе с существованием пределов $\lim_k \int \rho_N d\pi_k$ и $\lim_n \int \rho_n d\pi_K$ для любых достаточно больших N и K позволяет нам изменить порядок повторного перехода к пределу в следующем соотношении:

$$\lim_k \int h d\pi_k = \lim_k \lim_n \int \rho_n d\pi_k = \lim_n \lim_k \int \rho_n d\pi_k = \lim_n \int \rho_n d\pi = \int h d\pi.$$

Следствие доказано.

Следствием компактности и непрерывности является следующий критерий существования.

ПРЕДЛОЖЕНИЕ 1.6. *Задача Монжа–Канторовича с дополнительными линейными ограничениями имеет решение, если и только если множество $\Pi_W := \{\pi : \pi|_W = 0\}$ непусто.*

ЗАМЕЧАНИЕ 1.7. Ничего не изменится, если заменить ограничение $\pi|_W = 0$ на $\pi|_{\bar{W}} = 0$, где $\bar{W}$ – замыкание W в $\|\cdot\|_L$ топологии. Действительно, из определения такой топологии очевидно, что $\pi|_W = 0 \Leftrightarrow \pi|_{\bar{W}} = 0$, значит, замена ничего не меняет.

2. Двойственность Канторовича. Следующая теорема обобщает известное утверждение двойственности Канторовича на случай дополнительных линейных ограничений.

ТЕОРЕМА 2.1. *Пусть $X_1, \dots, X_n, X = X_1 \times \dots \times X_n$ – польские пространства, $\mu = (\mu_k \in \mathcal{P}(X_k))$ для $k = 1, \dots, n$, W – подпространство $C_L(\mu)$, $c \in C_L(\mu)$. Тогда*

$$\inf_{\pi \in \Pi_W} \int c d\pi = \sup_{f + \omega \leq c} \sum_{k=1}^n \int_{X_k} f_k(x_k) d\mu_k.$$

где $f \in F = \bigoplus_{i=1}^n C_L(\mu_i)$, $f = \sum_{i=1}^n f_i(x_i)$, $\omega \in W$.

Чтобы доказать эту теорему, нам потребуются некоторые дополнительные результаты. Во-первых, нам нужна следующая версия теоремы двойственности Канторовича для задачи без дополнительных ограничений.

ТЕОРЕМА 2.2. *Пусть $X_1, \dots, X_n, X = X_1 \times \dots \times X_n$ – польские пространства, $\mu = (\mu_k \in \mathcal{P}(X_k))$ для $k = 1, \dots, n$, $c \in C_L(\mu)$. Тогда*

$$\inf_{\pi \in \Pi} \int c d\pi = \sup_{f \leq c} \sum_{k=1}^n \int_{X_k} f_k(x_k) d\mu_k,$$

где $f \in F = \bigoplus_{i=1}^n C_L(\mu_i)$.

Доказательство близкой версии теоремы может быть найдено в [10] или [11]. В Дополнении мы приводим полное доказательство данной формулировки теоремы.

Еще одно утверждение, которое мы собираемся использовать – обобщенная версия теоремы о минимаксе, сформулированная в [12; теорема 2.4.1]. Там же можно найти ее доказательство и все необходимые пояснения.

ТЕОРЕМА 2.3. Пусть K – компактное выпуклое подмножество хаусдорфова топологического векторного пространства, Y – выпуклое подмножество произвольного векторного пространства, а h – вещественнозначная функция ($\leq +\infty$) на $K \times Y$, которая полунепрерывна по x для каждого фиксированного y , выпукла на K и вогнута на Y . Тогда

$$\min_{x \in K} \sup_{y \in Y} h(x, y) = \sup_{y \in Y} \min_{x \in K} h(x, y).$$

Теперь мы готовы доказать анонсированный результат.

ДОКАЗАТЕЛЬСТВО ТЕОРЕМЫ ДВОЙСТВЕННОСТИ 2.1. Неравенство

$$\inf_{\pi \in \Pi_W} \int c \, d\pi \geq \sup_{f + \omega \leq c} \sum_{k=1}^n \int f_k \, d\mu_k$$

является почти очевидным:

$$\begin{aligned} \inf_{\pi \in \Pi_W} \int c \, d\pi &\geq \inf_{\pi \in \Pi_W} \sup_{f + \omega \leq c} \int (f + \omega) \, d\pi \\ &= \inf_{\pi \in \Pi_W} \sup_{f + \omega \leq c} \sum_{i=1}^n \int f_i \, d\mu_i = \sup_{f + \omega \leq c} \sum_{i=1}^n \int f_i \, d\mu_i. \end{aligned}$$

Докажем обратное неравенство:

$$\sup_{f + \omega \leq c} \sum_{k=1}^n \int f_k \, d\mu_k = \sup_{\omega \in W} \sup_{f \in F, f \leq (c - \omega)} \sum_{k=1}^n \int f_k \, d\mu_k = \sup_{\omega \in W} \inf_{\pi \in \Pi} \int (c - \omega) \, d\pi.$$

Здесь мы используем теорему (2.2) для функции стоимости $(c - \omega) \in C_L(\mu)$. Следующий шаг – использование теоремы о минимаксе 2.3, чтобы поменять местами инфимум и супремум. Используя обозначения этой теоремы, положим $K = \Pi(\mu)$, $Y = W$, $h(\pi, \omega) = \int (c - \omega) \, d\pi$. Заметим, что $h(\pi, \omega)$ линеен по обоим аргументам и непрерывен по π для каждого фиксированного ω (это уже было доказано нами, см. следствие 1.5). Следовательно, все предположения теоремы выполнены, и мы получаем, что

$$\sup_{\omega \in W} \inf_{\pi \in \Pi} \int (c - \omega) \, d\pi = \inf_{\pi \in \Pi} \sup_{\omega \in W} \int (c - \omega) \, d\pi.$$

Если $\pi \notin \Pi_W$, то существует $\omega_1 \in W$ такое, что $\int \omega_1 \, d\pi < 0$. Выбор $\omega = \alpha \omega_1$, $\alpha \rightarrow +\infty$ показывает, что супремум $\sup_{\omega \in W} \int (c - \omega) \, d\pi$ равен $+\infty$. Следовательно, можно заключить, что

$$\inf_{\pi \in \Pi} \sup_{\omega \in W} \int (c - \omega) \, d\pi = \inf_{\pi \in \Pi_W} \int c \, d\pi,$$

что и требовалось для завершения доказательства.

ЗАМЕЧАНИЕ 2.4. Заметим, что утверждение теоремы остается верным даже для случая пустого множества $\Pi_W(\mu)$, если мы положим $\inf(\emptyset) = +\infty$.

Используя в точности такое же рассуждение, можно доказать следующую версию теоремы двойственности, сформулированную в терминах пространства непрерывных ограниченных функций.

ТЕОРЕМА 2.5. Пусть $X_1, \dots, X_n$, $X = X_1 \times \dots \times X_n$ – польские пространства, $\mu = (\mu_k \in \mathcal{P}(X_k))$ для $k = 1, \dots, n$, W – подпространство $C_b(\mu)$, $c \in C_b(\mu)$. Тогда

$$\inf_{\pi \in \Pi_W} \int c d\pi = \sup_{f + \omega \leq c} \sum_{k=1}^n \int_{X_k} f_k(x_k) d\mu_k,$$

где $f \in F = \bigoplus_{i=1}^n C_b(X_i)$, $f = \sum_{i=1}^n f_i(x_i)$, $\omega \in W$.

Единственное, что нужно заменить в представленном выше доказательстве, состоит в использовании вместо 2.2 версии классической теоремы двойственности Канторовича для ограниченных непрерывных функций (например, теорему 5.10 из [2]).

3. Геометрия оптимальных транспортных планов. В этом разделе мы сформулируем аналог c -монотонности для нашей задачи и докажем его необходимость для носителей оптимальных транспортных планов. Ранее подобный результат был получен в работе [3]. Приведенное там доказательство отлично от нашего и основано на абстрактной теории меры в духе [11].

ОПРЕДЕЛЕНИЕ 3.1. Для двух мер α, β на $X = X_1 \times \dots \times X_n$ определим *отношение эквивалентности* $\sim_W$: $\alpha \sim_W \beta$, если и только если

- (1) $(\text{Pr}_k)_\#(\alpha) = (\text{Pr}_k)_\#(\beta)$ для всех $k = 1, \dots, n$,
- (2) $\int \omega d\alpha = \int \omega d\beta$ для всех $\omega \in W$.

Обозначим через $[\beta]_W$ класс эквивалентности β по отношению $\sim_W$. Пусть S_m – набор из m точек в пространстве X , β_s – мера с носителем S_m .

ОПРЕДЕЛЕНИЕ 3.2. Для измеримой по Борелю функции стоимости $c: X \rightarrow \mathbb{R}$ и линейного подмножества $W \subset C_L(\mu)$ множество $\Gamma \subset X$ называется (c, W) -*монотонным*, если и только если для любого $m \in \mathbb{N}$, любого $S_m \subset \Gamma$ и любой меры β_s такой, что $\text{supp}(\beta_s) = S_m$, а также любой меры $\alpha \sim_W \beta_s$ верно неравенство

$$\int c d\beta_s \leq \int c d\alpha.$$

ЗАМЕЧАНИЕ 3.3. В силу линейности интеграла ничего не изменится, если в определении мы будем рассматривать только вероятностные меры.

ПРЕДЛОЖЕНИЕ 3.4. Если $W = \{0\}$, то определение (c, W) -монотонности эквивалентно определению обычной c -монотонности.

ДОКАЗАТЕЛЬСТВО. Из $(c, \{0\})$ -монотонности очевидным образом следует c -монотонность. Обратное следует из известного факта (см. [2]): из c -монотонности $\text{supp}(\beta)$ следует, что β оптимальна в классе мер на X с теми же маргиналами, что и β . Действительно, если $W = \{0\}$, то этот класс совпадает с классом эквивалентности $[\beta]_W$. Получаем, что β оптимальна в $[\beta]_W$, где β – любая мера с носителем, состоящем из конечного числа точек и лежащем в c -монотонном множестве (здесь мы также используем известный факт, что подмножество c -монотонного множества также c -монотонно).

ОПРЕДЕЛЕНИЕ 3.5. Транспортный план $\pi \in \Pi_W(X)$ называется (c, W) -*монотонным*, если и только если существует (c, W) -монотонное множество Γ полной π -меры: $\pi(\Gamma) = 1$.

ТЕОРЕМА 3.6. Пусть $X_k, k = 1, \dots, n, X := X_1 \times \dots \times X_n$ – польские пространства, $\mu = (\mu_k \in \mathcal{P}(X_k))$, $c \in C_L(\mu)$ – функция стоимости, $W \subset C_L(\mu)$ – векторное подпространство, $\mu_k|_{W \cap C_L(\mu_k)} = 0$ и $\pi_* \in \Pi_W(\mu)$ – решение следующей задачи Монжа–Канторовича с дополнительными линейными ограничениями:

$$\inf_{\pi \in \Pi_W} \int_X c \, d\pi.$$

Тогда π_* – это (c, W) -монотонный транспортный план.

Представленное ниже доказательство основывается на теореме о двойственности, которая была доказана в предыдущем разделе.

ДОКАЗАТЕЛЬСТВО. Из теоремы о двойственности 2.1 имеем

$$\int_X c \, d\pi_* = \sup \left\{ \sum_{k=1}^n \int f_k \, d\mu_k \right\},$$

где супремум берется по множеству всех пар $(f, \omega) \in F \times W$ таких, что $f + \omega \leq c$. Напомним обозначения: $F = \bigoplus_{i=1}^n C_L(\mu_i)$, $f = \sum_{i=1}^n f_i(x_i)$. Пусть $(f^{(k)}, \omega_k)$ – максимизирующая последовательность для двойственной задачи и $c_k = c - f^{(k)} - \omega_k$. Так как

$$\int_X c_k \, d\pi_* = \int_X c \, d\pi_* - \sum_{i=1}^n \int_{X_i} f_i^{(k)} \, d\mu_i \rightarrow 0$$

и $c_k \geq 0$, мы можем найти такие подпоследовательность $c_{k(j)}$ и борелевское множество Γ , для которого $\pi_*(\Gamma) = 1$, что $c_{k(j)} \rightarrow 0$ на Γ . В дальнейшем условимся обозначать индексы $\{k(j)\}$ просто как $\{j\}$. Если $S = \{x_i\}_{i=1}^m \subset \Gamma$, β_s есть мера с носителем S и $\alpha \in [\beta_s]_W$, то мы получаем, что

$$\int c \, d\alpha \geq \sum_{i=1}^n \int f_i^{(k)} \, d\alpha + \int \omega_k \, d\alpha.$$

Так как

$$\begin{aligned} (\text{Pr}_i)_\#(\alpha) = (\text{Pr}_i)_\#(\beta) \text{ для каждого } i \in 1, \dots, m &\implies \int f_i^{(k)} \, d\alpha = \int f_i^{(k)} \, d\beta_s, \\ \int \omega \, d\alpha = \int \omega \, d\beta \text{ для каждого } \omega \in W &\implies \int \omega_k \, d\alpha = \int \omega_k \, d\beta_s, \end{aligned}$$

то можно получить, что

$$\int c \, d\alpha \geq \sum_{j=1}^n \int f_j^{(k)} \, d\beta_s + \int \omega_k \, d\beta_s = \int (c - c_k) \, d\beta_s$$

для любого k . Устремляя $k \rightarrow \infty$, мы получаем (c, W) -монотонность Γ .

Заметим, что в отличие от задачи без ограничений, условие (c, W) -монотонности не является достаточным для оптимальности в общем случае (см. пример 5.4). Однако в некоторых частных случаях достаточность все же имеет место быть, как, например, в случае мартингалных ограничений, которым посвящен следующий пункт.

4. Мартингальная задача Монжа–Канторовича. В этом разделе мы покажем, что так называемая мартингальная задача Монжа–Канторовича может быть рассмотрена как частный случай задачи Монжа–Канторовича с линейными ограничениями. Общая теория, развитая нами в предыдущих разделах, влечет утверждение двойственности и некоторые другие известные результаты о мартингальных оптимальных планах. Пусть $X_1 = X_2 = \dots = X_n = \mathbb{R}$, $X = \mathbb{R}^n$. Рассмотрим меры $\mu_k \in \mathcal{P}(\mathbb{R})$, $k \in \{1, \dots, n\}$, имеющие конечный первый момент. Положим

$$W = \left\{ \sum_{k=1}^{n-1} \rho(x_1, \dots, x_k)(x_{k+1} - x_k) : \rho(x_1, \dots, x_k) \in C_b(X) \right\}.$$

ПРЕДЛОЖЕНИЕ 4.1. Верно включение $W \subset C_L(\mu)$.

ДОКАЗАТЕЛЬСТВО. Мы имеем

$$\begin{aligned} \sum_{k=1}^{n-1} |\rho(x_1, \dots, x_k)(x_{k+1} - x_k)| &\leq \sum_{k=1}^{n-1} C_{k,\rho} |x_{k+1} - x_k| \leq \sum_{k=1}^{n-1} C_{k,\rho} (|x_{k+1}| + |x_k|), \\ \int C_{k,\rho} |x_k| d\mu_k &< \infty \quad \text{для каждого } k \in \{1, \dots, n\}. \end{aligned}$$

Последнее неравенство следует из того факта, что все μ_k имеют конечный первый момент.

Отметим следующий факт.

ПРЕДЛОЖЕНИЕ 4.2. Равенство

$$\Pi_W(\mu_1, \dots, \mu_n) = \left\{ \pi : \int_{\mathbb{R}} x_{k+1} d\pi_{x_1, \dots, x_k}(x_{k+1}) = x_k \right\}$$

верно для почти всех x_k относительно $(P_{1, \dots, k})\#(\pi)$, для любого $k \in \{1, \dots, n-1\}$, где $\pi_{x_1, \dots, x_k}$ – условная мера.

ДОКАЗАТЕЛЬСТВО. Действительно,

$$\begin{aligned} \int_{\mathbb{R}} x_{k+1} d\pi_{x_1, \dots, x_k}(x_{k+1}) &= x_k \quad \text{для почти всех } x_k \text{ относительно } (P_{1, \dots, k})\#(\pi) \\ \iff \int \rho(x_1, \dots, x_k) \int x_{k+1} d\pi_{x_1, \dots, x_k}(x_{k+1}) d\mu_{k+1} \\ &= \int x_k \rho(x_1, \dots, x_k) d\mu_{k+1} \quad \text{для любого } \rho \in C_b(X_1 \times \dots \times X_k) \\ \iff \int \rho(x_1, \dots, x_k)(x_{k+1} - x_k) d\pi &= 0 \quad \text{для любого } \rho \in C_b(X_1 \times \dots \times X_k). \end{aligned}$$

Так как непрерывность очевидна, доказательство окончено.

Такая задача имеет название *мартингальной задачи Монжа–Канторовича*. Заметим, что существование решения в этой задаче не гарантируется.

Подстановкой конкретного подпространства W в общее утверждение о двойственности 2.1 мы сразу получаем следующее утверждение.

ТЕОРЕМА 4.3. *В мартингальной задаче Монжа–Канторовича верно следующее утверждение о двойственности:*

$$\inf_{\pi \in \Pi_W} \int_{R^n} c(x_1, \dots, x_n) d\pi \\ = \sup \left\{ \sum_{k=1}^n \int_{\mathbb{R}} \phi(x_k) d\mu_k : \sum_{k=1}^n \phi_k(x_k) + \sum_{k=1}^{n-1} \rho(x_1, \dots, x_k)(x_{k+1} - x_k) \leq c \right\},$$

где $(\phi_k)_{k=1}^n \in (C_L(\mu_k))_{k=1}^n$, $\rho(x_1, \dots, x_k) \in C_b(\mathbb{R}^n)$.

Этот результат впервые был получен в работе [5]. Также там показано (утверждение 4.1 в [5]), что экстремальное значение двойственной задачи в общем случае не достигается. Кроме того, в работе [6] доказано, что соответствующая формулировка (c, W) -монотонности для множества влечет оптимальность сосредоточенного на нем транспортного плана.

5. Инвариантная задача Монжа–Канторовича. Пусть G – некоторая группа, действующая непрерывно на $X_1, \dots, X_n$. Предположим также, что действие G определено на X и что оно диагонально:

$$g(x_1, \dots, x_n) = (g_1(x_1), \dots, g_n(x_n)).$$

Определим пространство W как подпространство C_b следующего вида:

$$W := \text{span}\{h \circ g - h : g \in G, h \in C_b(\mu)\},$$

где span означает пространство всевозможных конечных линейных комбинаций.

Можно доказать следующую характеристику мер, которые равны нулю на W .

ПРЕДЛОЖЕНИЕ 5.1. *Для каждого $\pi \in \mathcal{P}(X)$ $\pi|_W = 0$, если и только если π – инвариантная мера.*

ДОКАЗАТЕЛЬСТВО. Верны следующие импликации:

$$\begin{aligned} \pi \text{ инвариантна} &\iff \int h dg_{\#} \pi = \int h d\pi && \text{для всех } h \in C_L(\mu), \quad g \in G \\ &\iff \int h \circ g d\pi = \int h d\pi && \text{для всех } h \in C_L(\mu), \quad g \in G \\ &\iff \int h \circ g - h d\pi = 0 && \text{для всех } h \in C_L(\mu), \quad g \in G \\ &\iff \pi|_W = 0, \end{aligned}$$

что доказывает наше утверждение.

В случае, если функция W определена указанным выше способом, мы будем называть задачу

$$\inf_{\pi \in \Pi_W(\mu)} \left\{ \int_X c(x) d\pi \right\}$$

инвариантной задачей Монжа–Канторовича. В следствие диагональности действия G получаем, что $f \circ \text{Pr}_k \circ g = f \circ g_k \circ \text{Pr}_k$ и

$$F \cap W = \bigoplus_{k=1}^n \{f \in C_L(\mu_k) : f \circ g_k \circ P_k - f \circ P_k \in W\}.$$

Это означает, что необходимым условием для существования решения в инвариантной задаче является инвариантность заданных маргинальных мер μ_k относительно действия G на соответствующих X_k . Можно доказать, что эти условия также и достаточны.

ПРЕДЛОЖЕНИЕ 5.2. *Инвариантная задача Монжа–Канторовича с инвариантными маргиналами имеет решение.*

ДОКАЗАТЕЛЬСТВО. Достаточно показать, что $\mu \otimes \nu \in \Pi_W(\mu, \nu)$. Для любого $h \in W$ имеем

$$\begin{aligned} & \int (h \circ g - h) d\left(\bigotimes_{k=1}^n \mu_k\right) \\ &= \int_{X_2 \times \dots \times X_n} \left(\int_{X_1} h(g(x)) - h(x) d\mu_1(x_1) \right) d\left(\bigotimes_{k=2}^n \mu_k\right) \\ &= \int_{X_2 \times \dots \times X_n} \left(\int_{X_1} (h(g(x)) - h(g_1(x_1), x_2, \dots, x_n)) \right. \\ & \quad \left. + (h(g_1(x_1), x_2, \dots, x_n) - h(x)) d\mu_1(x_1) \right) d\left(\bigotimes_{k=2}^n \mu_k\right) \\ &= \int_{X_2 \times \dots \times X_n} \left(\int_{X_1} h(g(x)) - h(g_1(x_1), x_2, \dots, x_n) d\mu_1(x_1) \right) d\left(\bigotimes_{k=2}^n \mu_k\right) \\ &= \int_{X_1} \left(\int_{X_2 \times \dots \times X_n} h(g(x)) - h(g_1(x_1), x_2, \dots, x_n) d\bigotimes_{k=2}^n \mu_k \right) d\mu_1(x_1) = \dots \\ &= \int_{X_1} \int_{X_2} \dots \int_{X_n} (h(g(x)) - h(g_1(x_1), \dots, g_{n-1}(x_{n-1}), x_n)) d\mu_n(x_n) \dots d\mu_1(x_1) = 0, \end{aligned}$$

что завершает доказательство.

ПРИМЕР 5.3. Пусть $k = 2$, $X_1 = X_2 = \mathbb{R}^{\mathbb{N}}$ – прямое произведение счетного числа $\mathbb{R}^1$, $c(x, y) = |x_1 - y_1|^2$, $G = \mathfrak{S}^{\infty}$ – группа конечных перестановок, действующая перестановками координат. Меры, инвариантные относительно таких перестановок, называются *перестановочными* (exchangeable). Согласно общей теории оптимальный план в множестве таких инвариантных планов существует, если и только если маргинальные меры также инвариантны. Подробное описание и результаты исследования этой задачи могут быть найдены в [13]. Об альтернативном подходе к бесконечномерной задаче с использованием нормы Камерона–Мартина в качестве функции стоимости можно прочесть в [1], [14], где даны дополнительные ссылки.

ПРИМЕР 5.4. Пусть $k = 2$, $X = Y = \mathbb{S}^1$ – окружности единичного радиуса с нормированными мерами Лебега, $G = \mathbb{Z}$ действует поворотами (образующая группы

соответствует положительному повороту на один радиан), $c(x, y)$ – некоторая непрерывная ограниченная функция. Этот пример интересен тем, что диагональное множество $D := \{(x, x) : x \in X\} \subset X \times Y$ (как, впрочем, и любое другое непустое инвариантное множество) является (c, W) -монотонным, но, конечно, далеко не всегда обязано быть носителем оптимальной меры. Этот простой пример является контрпримером, показывающим, что (c, W) -монотонность множества в общем случае не влечет оптимальность сосредоточенной на нем меры.

То, что диагональ действительно (c, W) -монотонна, следует из того, что для любого $n \in \mathbb{N}$, любого набора n точек $S_n \subset D$, и любой вероятностной меры α с носителем S_n верно, что $[\alpha]_W = \alpha$.

Снабдим группу G такой топологией, что функционал $(f, x) : G \rightarrow \mathbb{R}, g \rightarrow (f \circ g)(x)$ будет измеримым относительно борелевской σ -алгебры на $(f, x) \in C_b(X) \times X$. Если топологическая группа G компактна, то можно определить отображение

$$\bar{f}(x) := \int_G (f \circ g)(x) d\chi(g),$$

где $f \in C_b(X)$ и χ – левоинвариантная вероятностная мера Хаара на группе G . Легко проверить, что $\bar{f} \in C_b(X)$: подынтегральное выражение непрерывно по переменной $x \in X$ и ограничено по $x \in X$ и $g \in G$, что влечет непрерывность $\bar{f}(x)$.

Определим еще одно подпространство в C_b :

$$W_1 = \text{span} \left\{ w \in W, f - \int_G f \circ g d\chi(g) \right\}.$$

Мы хотим доказать, что в нашей задаче нет никакой разницы между ограничениями, задаваемыми W и W_1 .

ПРЕДЛОЖЕНИЕ 5.5. *Верно равенство*

$$\pi|_{W_1} = 0 \iff \pi|_W = 0.$$

ДОКАЗАТЕЛЬСТВО. Следствие $\pi|_{W_1} = 0 \Rightarrow \pi|_W = 0$ очевидно. Докажем обратное. Из $\pi|_W = 0$ следует, что $\int_X h \circ g d\pi = \int_X h d\pi$ для любых $h \in C_b(X)$, $g \in G$. Тогда

$$\int_X \left(\int_G h \circ g d\chi(g) \right) d\pi = \int_G \left(\int_X h \circ g d\pi \right) d\chi(g) = \int_G \left(\int_X h d\pi \right) d\chi(g) = \int_X h d\pi.$$

Следовательно,

$$\int_X \left(h - \int_G h \circ g d\chi(g) \right) d\pi = 0$$

для каждого $h \in C_b(X)$.

Пусть $\bar{W}_1 \in C_b(X)$ – замыкание W_1 в равномерной топологии. Так как эта топология сильнее, чем $L_1(\pi)$ -топология для любого $\pi \in \Pi$, верно, что ограничения, задаваемые W_1 и $\bar{W}_1$, эквивалентны: $\pi|_{\bar{W}_1} = 0 \Leftrightarrow \pi|_{W_1} = 0$ и $\Pi_{W_1} = \Pi_{\bar{W}_1}$.

В том случае, если топологическая группа G компактна, возможно определить линейный оператор $\text{Pr}_{\bar{W}_1} : C_b(X) \rightarrow \bar{W}_1$ как

$$\text{Pr}_{\bar{W}_1}(f) := f - \int_G (f \circ g) d\chi(g),$$

где χ – это левоинвариантная вероятностная мера Хаара на группе G .

ПРЕДЛОЖЕНИЕ 5.6. Если G – компактная группа, то линейный оператор $\text{Pr}_{\overline{W}_1}$ – непрерывная проекция на подпространство $\overline{W}_1$.

ДОКАЗАТЕЛЬСТВО. Сначала докажем непрерывность оператора относительно равномерной топологии на $C_b(X)$. Пусть $\sup_{x \in X} |f(x) - h(x)| < \varepsilon/2$ для некоторого фиксированного $\varepsilon > 0$. Тогда

$$\begin{aligned} \sup_{x \in X} |\text{Pr}_{\overline{W}_1}(f - h)(x)| &= \sup_{x \in X} |f(x) - h(x) + \int_G ((f - h) \circ g)(x) d\chi(g)| \\ &\leq \sup_{x \in X} |f(x) - h(x)| + \sup_{x \in X} \left| \int_G (f - h)(g(x)) d\chi(g) \right| \\ &\leq \frac{\varepsilon}{2} + \sup_{x \in X} \int_G |(f - h)(g(x))| d\chi(g) \leq \frac{\varepsilon}{2} + \sup_{x \in X} \int_G \frac{\varepsilon}{2} d\chi(g) = \varepsilon. \end{aligned}$$

Таким образом, непрерывность доказана.

Очевидно, что для любого $h \in C_b(X)$ $\text{Pr}_{\overline{W}_1}(h) = h - \int_G (f \circ g) d\chi(g) \in \overline{W}_1$. Чтобы доказать, что $\text{Pr}_{\overline{W}_1}$ – проекция, нам осталось показать следующее: $\text{Pr}_{\overline{W}_1}(h) = h$ для каждого $h \in \overline{W}_1$. Если $h = f \circ g_0 - f$, то

$$\text{Pr}_{\overline{W}_1}(f \circ g_0 - f) = f \circ g_0 - f - \int_G (f \circ g_0 \circ g) d\chi(g) + \int_G (f \circ g) d\chi(g) = f \circ g_0 - f.$$

Если же $h = f - \int_G (f \circ g) d\chi(g)$, то

$$\begin{aligned} \text{Pr}_{\overline{W}_1}(f - \int_G (f \circ g) d\chi(g)) &= f - \int_G (f \circ g) d\chi(g) - \int_G (f \circ g) d\chi(g) + \int_G \left(\int_G (f \circ g) d\chi(g) \right) \circ g d\chi(g) \\ &= f \circ g - \int_G (f \circ g) d\chi(g). \end{aligned}$$

Используя линейность, получаем, что $\text{Pr}_{\overline{W}_1} = \text{Id}$ на W_1 . Из непрерывности следует, что оператор единственным образом продолжается с подпространства на его замыкание, так что $\text{Pr}_{\overline{W}_1} = \text{Id}$ на $\overline{W}_1$, что завершает доказательство.

Соотношение двойственности для инвариантной задачи может быть получено подстановкой соответствующего подпространства W в равенство из теоремы 2.5. Однако в случае компактной группы G двойственность может быть сформулирована в более удобной форме.

Обозначим через $\hat{h} := \text{Pr}_{\overline{W}_1}(h)$ непрерывную проекцию $h \in C_b(X)$ на $\overline{W}_1$,

$$\bar{h} := (\text{Id} - \text{Pr}_{\overline{W}_1})(h) = h - \hat{h} = \int_G (h \circ g) d\chi(g).$$

Пусть V – образ $C_b(X)$ относительно этой проекции. Ясно, что $C_b(X) = \overline{W}_1 \oplus V$ в смысле прямой суммы банаховых пространств. Подобные объекты могут быть определены и для функций $f_k \in C_b(X_k)$ или $f \in F$. Роль подпространства $\overline{W}_1$ в этих случаях будет играть равномерное замыкание подпространств $W_1 \cap C_b(X_k)$ и $W_1 \cap F$ соответственно.

ТЕОРЕМА 5.7. Для инвариантной задачи с компактной группой симметрий G , функцией стоимости $c \in C_b(X)$ и инвариантными маргинальными мерами μ_k верна следующая формулировка утверждения о двойственности:

$$\inf_{\pi \in \Pi_W} \int_X c(x) d\pi = \sup_{f \leq \bar{c}} \left(\sum_{k=1}^n \int f_k(x) d\mu_k \right) = \sup_{\bar{f} \leq \bar{c}} \left(\sum_{k=1}^n \int f_k(x) d\mu_k \right),$$

где $f \in F = \bigoplus_{k=1}^n C_b(X_k)$.

ДОКАЗАТЕЛЬСТВО. Из теоремы 2.5 получаем, что

$$\inf_{\pi \in \Pi_W} \int c d\pi = \sup_{f + \omega \leq c} \sum_{k=1}^n \int_{X_k} f_k(x_k) d\mu_k,$$

где $f \in F$, $\omega \in W$. Кроме того,

$$\begin{aligned} \sup_{f + \omega \leq c} \sum_{k=1}^n \int_{X_k} f_k(x_k) d\mu_k &= \sup_{\bar{f}, \omega} \left(\sum_{k=1}^n \int_{X_k} f_k(x_k) d\mu_k : \bar{f} + \hat{f} + \omega \leq \bar{c} + \hat{c} \right) \\ &= \sup_{\bar{f}, \hat{f}, \omega} \left(\sum_{k=1}^n \int_{X_k} f_k(x_k) d\mu_k : \bar{f} \leq \bar{c} + (\hat{c} - \omega - \hat{f}) \right). \end{aligned}$$

Заметим, что максимизируемый функционал не зависит от $\bar{W}_1$ -составляющей f , так что мы можем выбирать $\hat{f}$ произвольным образом. Следовательно, $\tilde{c} := \hat{c} - \hat{\omega} - \hat{f}$ — произвольная функция из $\bar{W}_1$. Неравенство $\bar{f}(x) \leq \bar{c}(x) + \tilde{c}(x)$ выполняется поточечно, поэтому, действуя на него произвольным элементом $g \in G$, получаем, что

$$(\bar{f} \circ g)(x) \leq (\bar{c} \circ g + \tilde{c} \circ g)(x) \iff \bar{f}(x) \leq (\bar{c} + \tilde{c} \circ g)(x)$$

для каждого фиксированного $x \in X$. Таким образом,

$$\bar{f}(x) \leq (\bar{c} + \tilde{c})(x) \iff \bar{f}(x) \leq \left(\bar{c} + \inf_{g \in G} (\tilde{c} \circ g) \right)(x).$$

Из определения $\bar{W}_1$ следует, что $\inf_{g \in G} (\tilde{c} \circ g)(x) \leq 0$ для каждой фиксированной точки x . Для элементов W_1 это очевидно, а так как равномерная сходимость влечет поточечную, это верно также для элементов замыкания. Следовательно, супремум достигается при $\tilde{c} \equiv 0$. Наконец, мы получаем

$$\begin{aligned} &\sup_{\bar{f}, \hat{f}, \omega} \left(\sum_{k=1}^n \int_{X_k} f_k(x_k) d\mu_k : \bar{f} \leq \bar{c} + (\hat{c} - \omega - \hat{f}) \right) \\ &= \sup_{\bar{f}, \hat{f}, \omega} \left(\sum_{k=1}^n \int_{X_k} f_k(x_k) d\mu_k : \bar{f} \leq \bar{c} + \inf_{g \in G} ((\hat{c} - \omega - \hat{f}) \circ g) \right) \\ &= \sup_{\bar{f}, \hat{f}} \left(\sum_{k=1}^n \int_{X_k} f_k(x_k) d\mu_k : \bar{f} \leq \bar{c} + \inf_{g \in G} (\hat{f} \circ g) \right) \\ &= \sup_{\bar{f}, \hat{f}} \left(\sum_{k=1}^n \int_{X_k} f_k(x_k) d\mu_k : \bar{f} \leq \bar{c} + \hat{f} \right) = \sup_{\bar{f}} \left(\sum_{k=1}^n \int_{X_k} f_k(x_k) d\mu_k : \bar{f} \leq \bar{c} \right). \end{aligned}$$

что завершает доказательство.

Заметим, что в случае инвариантной функции стоимости ($c(x) = c(g(x))$) для любого $g \in G$), инвариантная двойственная задача совпадает с двойственной задачей для случая без дополнительных ограничений. Действительно, если c инвариантна, то $\bar{c} = c$ и

$$\inf_{\pi \in \Pi_W(\mu)} \left\{ \int c \, d\pi \right\} = \sup_{f \leq c} \left\{ \sum_{k=1}^n \int f_k \, d\mu_k \right\}.$$

Известно (см., например, теорему 2.1.1 в [10]), что максимум в двойственной задаче для случая без ограничений достигается на некотором наборе функций. Эти функции инвариантны и являются максимизирующими для двойственной задачи с ограничением инвариантности.

СЛЕДСТВИЕ 5.8. *Если функция стоимости c и все маргинальные меры μ_k инвариантны относительно действия группы G , то*

$$\inf_{\pi \in \Pi_W(\mu)} \left\{ \int c \, d\pi \right\} = \sup_{f \leq c} \left(\sum_{k=1}^n \int f_k \, d\mu_k \right) = \inf_{\pi \in \Pi(\mu)} \left\{ \int c \, d\pi \right\},$$

причем решение инвариантной двойственной задачи совпадает с решением классической двойственной задачи Канторовича.

Подобный результат ранее был получен в работе [8] с использованием других соображений.

Дополнение. Доказательство теоремы о двойственности Канторовича для случая без дополнительных ограничений. Наша цель – доказать следующее утверждение.

ТЕОРЕМА. *Пусть $X_1, \dots, X_n$, $X = X_1 \times \dots \times X_n$ – польские пространства, $\mu = (\mu_k \in \mathcal{P}(X_k))$ для $k = 1, \dots, n$, $c \in C_L(\mu)$. Тогда*

$$\inf_{\pi \in \Pi} \int c \, d\pi = \sup_{f \leq c} \sum_{k=1}^n \int_{X_k} f_k(x_k) \, d\mu_k,$$

где $f \in F = \bigoplus_{i=1}^n C_L(\mu_i)$.

ДОКАЗАТЕЛЬСТВО. Зададим линейный функционал $T: F \rightarrow \mathbb{R}$ формулой

$$T(f) = \sum_{i=1}^n \int f_i \, d\mu_i.$$

Он положителен и непрерывен относительно полунормы $\|\cdot\|_L$. На $C_L(\mu)$ зададим функционал

$$U(h) = \inf\{T(f) : f \geq h\}.$$

Можно доказать, что U субаддитивен, т.е.

$$\begin{aligned} U(h+g) &= \inf_{f \in F} \{T(f) : f \geq (h+g)\} \\ &\leq \inf\{T(f) : f \geq h\} + \inf\{T(f) : f \geq g\} = U(h) + U(g), \end{aligned}$$

так как для любого $f_1 > h$, $f_2 > g$ верно, что $f_1 + f_2 > h + g$. Также U положительно однороден: для любого $\alpha \in \mathbb{R}^+$ имеем

$$U(\alpha h) = \inf_{f \in F} \{T(f) : f \geq (\alpha h)\} = \inf_{f \in F} \{T(\alpha f) : f \geq h\} = \alpha \inf_{f \in F} \{T(f) : f \geq h\} = \alpha U(h).$$

Таким образом, U — сублинейный функционал, поэтому U может быть использован в теореме Хана–Банаха. Также нам пригодится следующее свойство U : для всякого $t \in \mathbb{R}$ имеем

$$U(t \cdot h) \geq t \cdot U(h).$$

Действительно,

$$\begin{aligned} U(-h) &= \inf_{f \in F} \{T(f) : f \geq -h\} = \inf_{f \in F} \{T(f) : -f \leq h\} = \inf_{f \in F} \{T(-f) : f \leq h\} \\ &= -\sup_{f \in F} \{T(f) : f \leq h\} \geq -\inf_{f \in F} \{T(f) : f \geq h\} = -U(h). \end{aligned}$$

Комбинируя этот результат с положительной однородностью, мы получаем желаемое неравенство. Соотношение

$$-\sup_{f \in F} \{T(f) : f \leq h\} \geq -\inf_{f \in F} \{T(f) : f \geq h\}$$

следует из положительности функционала T . Так как он положителен, он сохраняет порядок; следовательно, любой элемент образа $\{f : f \leq h\}$ при отображении T не превосходит любой из элементов образа множества $\{f : f \geq h\}$ при том же отображении. Получаем, что

$$\sup_{f \in F} \{T(f) : f \leq h\} \leq \inf_{f \in F} \{T(f) : f \geq h\},$$

что нам и нужно с точностью до умножения на -1 .

Используя тот факт, что $T \leq U$ на F , мы можем применить теорему Хана–Банаха о продолжении, чтобы продолжить функционал T с подпространства F на все пространство $C_L(\mu)$. Обозначим такое продолжение через P и докажем, что свойство $P \leq U$ влечет положительность P .

Предположим, что P — не положительный функционал. Следовательно, существует функция $h \in C_L$ такая, что $h \geq 0$ и $P(h) < 0$. Это приводит к противоречию:

$$0 < P(-h) \leq U(-h) = \inf \{T(f) : f \geq -h\} \leq 0.$$

Определим еще один линейный оператор $T_c : \{f + tc : t \in \mathbb{R}, f \in F\} \rightarrow \mathbb{R}$ так, чтобы он совпадал с T на F : $T_c|_F = T$ и совпадал с U в точке $-c$: $T_c(-c) = U(-c)$. Из линейности T_c и свойств U следует, что $T_c(t \cdot c) = t \cdot U(c) \leq U(t \cdot c)$. Получаем, что $T_c \leq U$ всюду на области определения T_c . По теореме Хана–Банаха можно продолжить T_c до линейного функционала $P_c : C_L(\mu) \rightarrow \mathbb{R}$, так что $P_c|_F = T_c$, $P_c(-c) = U(-c)$, $P_c \leq U$.

Из построения линейных продолжений следует, что

$$\sup_P P(-c) \leq \inf_{f \in F} \{T(f) : f \geq -c\},$$

где супремум (и инфимум) берутся по множеству всех возможных линейных продолжений, удовлетворяющих сформулированным выше условиям (а именно продолжающих T и мажорируемых U). Умножая неравенство на -1 и используя линейность T и P , получаем

$$\inf_P P(c) \geq \sup_{f \in F} \{T(f) : f \leq c\}.$$

Аналогично, из равенства $P_c(-c) = U(-c)$, применяя умножение на -1 и линейность T и P , получаем

$$P_c(c) = \sup_{f \in F} \{T(f) : f \leq c\}.$$

Так как P_c продолжает T и мажорируется U , то

$$\inf_P P(c) = \sup_{f \in F} \{T(f) : f \leq c\}.$$

Это равенство отличается от доказываемого утверждения о двойственности тем, что инфимум здесь берется по множеству линейных операторов, которые априори не являются мерами. Поэтому в оставшейся части доказательства мы покажем, что на самом деле эти функционалы – транспортные планы.

Определим для каждого P его ограничение $l := P|_{C_b} \in (C_b(X))^*$ на двойственное пространство к пространству ограниченных непрерывных функций на X . Согласно соответствующей версии теоремы Рисса о представлении

$$(C_b(X))^* \simeq (C(\beta X))^* \simeq \mathcal{M}(\beta X),$$

где βX – компактификация Стоуна – Чеха пространства X , а $\mathcal{M}(\beta X)$ – множество знакопеременных мер на нем. Оба изоморфизма сохраняют соответствующие нормы и конусы положительных элементов (см. [15; теоремы 7.10.4 и 7.10.5]). Так как l положителен и $\langle l, 1 \rangle = 1$, то π – вероятностная мера.

Также известно, что для любого $f_i \in C_L(\mu_i)$ имеем $\int f_i d\pi = \langle l, f_i \rangle = \int f_i d\mu_i$. Так как $C_L(\mu_i)$ плотно в $L_1(X_i, \mu_i)$, то это равенство верно для всех интегрируемых (относительно μ_i) функций, в частности, для индикаторных функций измеримых множеств на X_i .

Пусть $\pi|_X$ – мера на X , задаваемая формулой $\pi|_X(A) = \pi(A \cap X)$ для любого $A \in \beta X$, измеримого относительно π . Мы хотим доказать, что $\pi|_X$ – вероятностная мера, так как из этого сразу следует, что X – множество полной π -меры. Очевидно, что полная вариация $\pi|_X$ не больше единицы, значит, нам остается доказать, что она не меньше единицы. Вспомним, что X имеет структуру топологического прямого произведения: $X = X_1 \times \cdots \times X_n$, следовательно, существует корректно определенная проекция $\text{Pr}_i : X \rightarrow X_i$, которая переводит меру $\pi|_X$ в некоторую меру на X_i . На самом деле, мера-образ при таком отображении есть в точности μ_i : для любого A_i , измеримого относительно μ_i имеем

$$((\text{Pr}_i)_\# \pi|_X)(A_i) = \pi|_X(\text{Pr}_i^{-1}(A_i)) = \pi(\text{Pr}_i^{-1}(A_i)) = \int \text{Ind}(A_i) d\mu_i = \mu_i(A_i).$$

В частности, $\pi|_X(X) = \pi|_X(\text{Pr}_i^{-1}(X_i)) = \mu(X_i) = 1$.

Таким образом, мы получили, что $l = P|_{C_b} \simeq \pi$ – вероятностная мера на X с маргиналами μ_i . Наша следующая цель: показать, что функционал P сам является мерой.

Мы будем использовать следующую полунорму на пространстве $C_L(\mu)$:

$$\|h\|_D = \inf_{f \in F} \{T(f) : f \geq |h|\}.$$

Можно проверить, что эта полунорма корректно определена (см. детали в [10]). Порождаемая этой нормой топология не слабее, чем топология, порождаемая $\|\cdot\|_L$:

$$\inf_{f \in F} \{T(f) : f \geq |h|\} \geq \sup_{\pi \in \Pi} \int |h| d\pi = \|h\|_L,$$

но на подпространстве $C_b \subset C_L$ она слабее, чем топология равномерной сходимости:

$$\inf_{f \in F} \{T(f) : f \geq |h|\} \geq \inf_{f \in F} \left\{ T(f) : f \geq \sup_{x \in X} |h(x)| \right\} = T\left(\sup_{x \in X} |h(x)|\right) = \sup_{x \in X} |h(x)|.$$

Заметим, что функционал P непрерывен относительно $\|h\|_D$:

$$P(|h|) \leq U(|h|) = \|h\|_D,$$

причем ограничение P на $C_b(X)$: $P|_{C_b} = l$ также непрерывно.

Важно, что C_b плотно в C_L относительно полунормы $\|\cdot\|_D$. Выберем $g \in C_L$. Пусть

$$|g| \leq f \in F = \bigoplus_{i=1}^n C_L(\mu_i), \quad k \in \mathbb{N},$$

$$g^k = \min\{k, g\}, \quad g_k^k = \max\{\min\{k, g\}, -k\} \in C_b(X).$$

Заметим, что $|g^k| \leq f$. Покажем, что $\|g - g_k^k\|_D \rightarrow 0$ при $k \rightarrow \infty$. Действительно,

$$\|g - g_k^k\|_D \leq \|g - g^k\|_D + \|g^k - g_k^k\|_D.$$

Тот факт, что $\pm g - k \leq f - k = \sum_{i=1}^n (f_i - k/n)$, вместе с положительностью отображений $(\cdot)_+ := \max\{\cdot, 0\}$ и U , а также теоремой Лебега о мажорируемой сходимости влечет соотношения

$$\|g - g^k\|_D = U((g - k)_+) \leq U\left(\sum_{i=1}^n \left(f_i - \frac{k}{n}\right)_+\right) \rightarrow 0 \quad \text{при } k \rightarrow \infty,$$

$$\|g^k - g_k^k\|_D = U((-g^k - k)_+) \leq U\left(\sum_{i=1}^n \left(f_i - \frac{k}{n}\right)_+\right) \rightarrow 0 \quad \text{при } k \rightarrow \infty.$$

Заметим, что $l := P|_{C_b}$ действует на каждой функции из $C_L(\mu)$ интегрированием, значит, может рассматриваться как линейный оператор на C_L . Так как интегрирование непрерывно в $\|\cdot\|_D$ -топологии и $\|\cdot\|_D$ не слабее, чем $\|\cdot\|_L$, мы заключаем, что оба функционала – продолженный l и P – являются непрерывными линейными функционалами на $(C_L(\mu), \|\cdot\|_D)$ и совпадают на C_b . Используя тот факт, что C_b плотно в $(C_L, \|\cdot\|_D)$, мы получаем, что продолженный l и P совпадают на всем пространстве C_L . Итак, $P \simeq \pi \in \Pi(\mu)$. Заметим, что на самом деле P – не просто мера, но также и транспортный план с маргиналами μ_i . Таким образом, теорема доказана.

Благодарю своего научного руководителя А.В. Колесникова за ценные идеи, которыми он делился со мной, и за долгие обсуждения, которые помогли мне в написании этой работы.

СПИСОК ЦИТИРОВАННОЙ ЛИТЕРАТУРЫ

- [1] В. И. Богачев, А. В. Колесников, “Задача Монжа–Канторовича: достижения, связи и перспективы”, *УМН*, **67**:5 (2012), 3–110.
- [2] C. Villani, *Optimal Transport. Old and New*, Grundlehren Math. Wiss., **338**, Springer-Verlag, Berlin, 2009.
- [3] M. Beiglböck, C. Griessler, *An Optimality Principle with Applications in Optimal Transport*, 2014, arXiv: 1404.7054.
- [4] D. Hobson, “The Skorokhod embedding problem and model-independent bounds for option prices”, *Paris–Princeton Lectures on Mathematical Finance 2010*, Lecture Notes in Math., **2003**, Springer-Verlag, Berlin, 2011, 267–318.
- [5] M. Beiglböck, P. Henry-Labordère, F. Penkner, “Model-independent bounds for option prices – a mass transport approach”, *Finance Stoch.*, **17**:3 (2013), 477–501.
- [6] M. Beiglböck, N. Juillet, *On a Problem of Optimal Transport Under Marginal Martingale Constraints*, 2012, arXiv: 1208.1509.
- [7] A. O. Lopes, J. K. Mengue, “Duality theorems in ergodic transport”, *J. Stat. Physics*, **149**:5 (2012), 921–942.
- [8] A. Moameni, *Invariance Properties of the Monge–Kantorovich Mass Transport Problem*, 2013, arXiv: 1311.7051.
- [9] A. M. Vershik, P. B. Zatitskiy, F. V. Petrov, “Virtual continuity of measurable functions of several variables and embedding theorems”, *Funct. Anal. Appl.*, **47**:3 (2013), 165–173.
- [10] S. T. Rachev, L. Ruschendorf, *Mass Transportation Problems. V. I. Theory*, Probab. Appl. (N. Y.), Springer-Verlag, New York, 1998; S. T. Rachev, L. Ruschendorf, *Mass Transportation Problems. V. II. Applications*, Probab. Appl. (N. Y.), Springer-Verlag, New York, 1998.
- [11] H. G. Kellerer, “Duality theorems for marginal problems”, *Z. Wahrsch. Verw. Gebiete*, **67**:4 (1984), 399–432.
- [12] D. R. Adams, L. I. Hedberg, *Function Spaces and Potential Theory*, Grundlehren Math. Wiss., **314**, Springer-Verlag, Berlin, 1996.
- [13] A. V. Kolesnikov, D. A. Zaev, *Optimal Transportation of Processes with Infinite Kantorovich Distance. Independence and Symmetry*, 2013, arXiv: 1303.7255.
- [14] D. B. Bukin, “On the Monge and Kantorovich problems for distributions of diffusion processes”, *Math. Notes*, **96**:5 (2014), 864–870.
- [15] V. I. Bogachev, *Measure Theory*, Vol. I, II, Springer-Verlag, Berlin, 2007.

Д. А. Заев

Национальный исследовательский университет
 “Высшая школа экономики”, г. Москва
 E-mail: zaev.da@gmail.com

Поступило

17.06.2015