Тема 9

Двойственность по Канторовичу и теорема существования

Двойственность по Канторовичу — это далеко идущее обобщение теорем о двойственности из линейного программирования.

9.1 Формулировка теоремы и схема доказательства

Напомним, что *польским* называется полное сепарабельное метрическое пространство. Через $\mathcal{P}(X)$ обозначим множество борелевских вероятностных мер на топологическом пространстве X. Через $L^1(\mu)$ обозначим пространство μ-измеримых функций с интегрируемым модулем.

Напомним, что функция $f\colon \mathcal{X} \to \mathbb{R}$ называется *полунепрерывной снизу в точке* x, если $f(x) \leq \liminf_{y\to x} f(y)$. Это равносильно следующему условию: для каждого $\varepsilon > 0$ найдется такая окрестность $U(x)$ точки x, что для любого $y \in U(x)$ выполнено неравенство $f(y) \geq f(x) - \varepsilon$, т.е. образ окрестности U содержится в луче $(x - \varepsilon, +\infty)$.

Следующий результат содержит в себе как принцип двойственности, так и теорему существования.

Теорема 9.1. *Пусть X и Y — польские пространства, $\mu \in \mathcal{P}(X)$, $\nu \in \mathcal{P}(Y)$ — борелевские вероятностные меры на них, и $c\colon X \times Y \to \mathbb{R} \cup \{+\infty\}$ — полунепрерывная снизу функция стоимости. Для любой $\pi \in \mathcal{P}(X \times Y)$ и любых $(\varphi, \psi) \in L^1(\mu) \times L^1(\nu)$ положим*

$$K(\pi) = \int_{X\times Y} c(x,y)\, d\pi(x,y), \qquad J(\varphi,\psi) = \int_X \varphi(x)\, d\mu + \int_Y \varphi(y)\, d\nu.$$

Обозначим через $\Pi(\mu,\nu)$ множество борелевских вероятностных мер на $X \times Y$, удовлетворяющих граничным условиям

$$\pi(A \times Y) = \mu(A), \quad \pi(X \times B) = \nu(D), \qquad \textit{для всех измеримых } A \subset X,\ B \subset Y$$

и пусть Φ_c — множество всех пар измеримых функций (φ, ψ), $\varphi \in L^1(\mu)$, $\psi \in L^1(\nu)$ таких, что

$$\varphi(x) + \psi(y) \le c(x, y) \qquad \text{для } \mu\text{-почти всех } x \in X \text{ и } \nu\text{-почти всех } y \in Y.$$

Тогда

$$\inf_{\Pi(\mu,\nu)} K(\pi) = \sup_{\Phi_c} J(\varphi, \psi).$$

Более того, инфимум в левой части равенства достигается. Кроме того, супремум в правой части равенства не изменится, если брать его только по непрерывным ограниченным функциям (φ, ψ).

Замечание 9.2. Можно представить себе такую модель. Как уже отмечалось, функция стоимости $c(x, y)$ может интерпретироваться как стоимость перевозки единицы продукта из $x \in X$ в $y \in Y$, а сама задача Канторовича — задача минимизации стоимости перераспределения этого продукта из имеющегося распределения μ в требуемое распределение ν. Можно подойти к этой задаче по-другому: вместо того, чтобы перевозить самому, Вы нанимаете перевозчика, которому будете платить за погрузку единицы продукта в точке $x \in X$ величину $\varphi(x)$, а за разгрузку продукта в точке $y \in Y$ — величину $\psi(y)$, причем предполагается, что $\varphi(x) + \psi(y) \le c(x, y)$. Так вот, теорема утверждает, что при этих условиях перевозчик не сможет взять с Вас больше, чем Вы сами потратили бы при «умном» планировании перевозок.

Мы разобьем доказательство теоремы 9.1 на несколько этапов. Начнем с проверки следующего неравенства.

Лемма 9.3. *В сделанных обозначениях,*

$$\sup_{\Phi_c} J(\varphi, \psi) \le \inf_{\Pi(\mu,\nu)} K(\pi).$$

При этом супремум в левой части неравенства можно брать только по непрерывным ограниченным функциям из Φ_c.

Доказательство. По определению функционалов I, J и семейства мер $\Pi(\mu, \nu)$ для любой пары (φ, ψ) и любой меры $\pi \in \Pi(\mu, \nu)$ имеем:

$$J(\varphi, \psi) = \int_X \varphi \, d\mu + \int_Y \psi \, d\nu = \int_{X\times Y} \varphi(x) \, d\pi(x, y) + \int_{X\times Y} \psi(y) \, d\pi(x, y),$$

так как для любого измеримого $A \subset X$ и любого измеримого $B \subset Y$ выполнены равенства $\pi(A \times Y) = \mu(A)$ и $\pi(X \times B) = \nu(B)$.

Заметим теперь, что так как неравенство $\varphi(x) + \psi(y) \le c(x, y)$ выполнено для μ-почти всех $x \in X$ и ν-почти всех $y \in Y$, то оно выполнено и для π-почти всех $(x, y) \in X \times Y$. Действительно, пусть оно не выполняется при $x \in N_x \subset X$, $\mu(N_x) = 0$ и $y \in N_y \subset Y$, $\nu(N_y) = 0$. Тогда оно выполнено для всех $(x, y) \in (X \setminus N_x) \times (Y \setminus N_y)$. Принимая во внимание, что

$$(X \setminus N_x) \times (Y \setminus N_y) = (X \times Y) \setminus \big((X \times N_y) \cup (N_x \times Y)\big),$$

а также, что $\pi(X \times N_y) = \nu(N_y) = 0$ и $\pi(N_x \times Y) = \mu(N_x) = 0$, заключаем, что $(X \setminus N_x) \times (Y \setminus N_y)$ — множество полой меры.

Таким образом,

$$J(\varphi, \psi) = \int_{X\times Y} \big(\varphi(x) + \psi(y)\big)\, d\pi(x,y) \le \int_{X\times Y} c(x,y)\, d\pi(x,y) = K(\pi).$$

Теперь требуемое неравенство получается переходом к супремуму слева и инфимуму справа. Замечание про непрерывные ограниченные функции очевидно, так супремум, взятый по меньшему множеству, может быть только меньше. Лемма доказана. $\square$

Замечание 9.4. Обратное неравенство достаточно доказать для супремума, взятого по непрерывным ограниченным функциям из Φ_c.

Пример 9.5. Продемонстрируем, как работает минимакс-принцип на примере классической задачи линейного программирования. Напомним, что в линейном программировании двойственность устроена следующим образом:

$$\inf_{\substack{y\ge 0\\ A^T y = c}} \langle b, y\rangle = \sup_{Ax \le b} \langle x, c\rangle.$$

Это равенство можно вывести с помощью минимакс-принципа так:

$$\inf_{\substack{y\ge 0\\ A^T y = c}} \langle b, y\rangle = \inf_{y\ge 0} \sup_{x} \big[\langle b, y\rangle + \langle x, c - A^T y\rangle\big] = \sup_{x} \inf_{y\ge 0} \big[\langle x, c\rangle + \langle b, y\rangle - \langle x, A^T y\rangle\big] =$$

$$= \sup_{x} \big[\langle x, c\rangle + \inf_{y\ge 0} \langle b - Ax, y\rangle\big] = \sup_{b \ge Ax} \langle x, c\rangle.$$

Чтобы превратить эти рассуждения в аккуратное доказательство, нам понадобится разить некоторую технику.

9.2 Преобразование Лежандра–Юнга–Фенхеля

Нам понадобятся некоторые свойства выпуклых множеств и функций.

9.3 Выпуклость

Пусть E — линейное топологическое векторное пространство, и $f\colon E \to \mathbb{R} \cup +\infty$ — выпуклая (вниз) функция на E. Последнее, напомним, означает, что для любых $z_1, z_2 \in E$ и любого $\lambda \in [0,1]$ имеет место неравенство

$$f\big(\lambda z_1 + (1-\lambda) z_2\big) \le \lambda f(z_1) + (1-\lambda) f(z_2).$$

Множество $\operatorname{dom} f = \big\{x \in E : f(x) < +\infty\big\}$ называют *эффективным множеством* функции f, а множество $\operatorname{epi} f = \big\{(x,\alpha) \in E \times \mathbb{R} : x \in \operatorname{dom} f, f(x) \le \alpha\big\}$ — ее *надграфиком*.

Упражнение 9.6. Проверьте, что функция выпукла тогда и только тогда, когда ее надграфик — выпуклое подмножество.

Пример 9.7. Если $A \subset E$ произвольное не пустое подмножество, то его *индикаторной функцией* в выпуклом анализе принято называть функцию, определенную так:

$$I_A(x) = \begin{cases} 0, & x \in A, \\ +\infty, & x \notin A. \end{cases}$$

Проверьте, что индикаторная функция выпуклого множества выпукла.

Пример 9.8. Еще один пример выпуклой функции — это квазиполунорма. Она обладает дополнительным свойством — положительной полуоднородностью. В частности, функционал Минковского определенный выше — выпуклая функция.

Следующее упражнение позволяет конструировать новые примеры выпуклых функций из уже имеющихся.

Упражнение 9.9. Проверьте, что

- сумма конечного набора выпуклых функций — выпуклая функция;
- для любой выпуклой функции f и любого $\alpha > 0$ функция αf — выпукла;
- супремум любого семейства выпуклых функций — выпуклая функция.

Нам понадобятся следующие свойства выпуклых множеств.

Лемма 9.10. *Пусть $A \subset \mathcal{X}$ — выпуклое подмножество линейного топологического пространства $\mathcal{X}$, $a \in \operatorname{Int} A$, $b \in \overline{A}$. Тогда $[a, b) \subset \operatorname{Int} A$.*

Доказательство. Пусть $c = ta + (1-t)b \in [a, b)$, $0 \le t < 1$, нам нужно показать, что $c \in \operatorname{Int} A$. Так как $a \in \operatorname{Int} A$, то существует такая окрестность нуля $0 \in V \subset \mathcal{X}$, что $a + V \subset \operatorname{Int} A$. Пусть $b' \in A$. Так как A выпукло, то $t(a+V) + (1-t)b' \subset A$. Поэтому, прибавляя и вычитая $(1-t)b$, получаем

$$A \supset ta+tV+(1-t)b'+(1-t)b-(1-t)b = ta+(1-t)b+tV-(1-t)(b-b') = c+tV-(1-t)(b-b').$$

Для завершения доказательства леммы достаточно найти такое $b' \in A$, что $tV-(1-t)(b-b')$ является окрестностью нуля в $\mathcal{X}$ (тогда $c+tV-(1-t)(b-b'$ — окрестность c, лежащая в A). Так как окрестности нуля переходят друг в друга при гомотетии, последнее равносильно тому, что

$$\frac{t}{1-t}V - (b-b') \quad \text{является окрестностью нуля в } \mathcal{X},$$

или, что все равно,

$$W = \frac{t}{1-t}V \quad \text{является окрестностью } (b-b') \text{ в } \mathcal{X}.$$

Однако любая окрестность нуля является окрестностью каждой своей точки, поэтому достаточно выбрать b' так, чтобы $b' \in b - W$. Так как кривая $c(t) = ta + (1-t)b$ непрерывна при всех t, в том числе при $t = 1$, то для любой окрестности точки $c(1) = b$, в том числе и для $b - W$, найдется такое $\delta > 0$, что $c\big((1-\delta, 1+\delta)\big) \subset b - W$. Выбрав $b' = c(1-\delta/2) \in A$, получим требуемое. □

Лемма 9.11. *Пусть $C \subset E$ — выпуклое подмножество, тогда множество* $\operatorname{Int} C$ *также выпукло. Кроме того, если* $\operatorname{Int} C \neq \emptyset$, *то* $\overline{C} = \overline{\operatorname{Int} C}$.

Доказательство. Первое утверждение леммы сразу следует из леммы 9.10. Действительно, если $x, y \in \operatorname{Int} C$, то $[x, y) \subset \operatorname{Int} C$ и $y \in \operatorname{Int} C$, поэтому $[x, y] \subset \operatorname{Int} C$.

Далее, включение $\overline{C} \supset \overline{\operatorname{Int} C}$ очевидно, докажем обратное. Возьмем произвольный $x \in \overline{C}$ и $y \in \operatorname{Int} C$. Тогда, в силу леммы 9.10, $[y, x) \subset \operatorname{Int} C$, поэтому $x \in \overline{\operatorname{Int} C}$. □

Упражнение 9.12. Докажите, что замыкание выпуклого множества также выпукло.

Через E^* будем обозначать топологически двойственное пространство, т.е. пространство непрерывных линейных функций на E (со слабой топологией).

Пример 9.13. Пусть $A \subset E$ — непустое подмножество. Определим функцию $S_A \colon E^* \to \mathbb{R} \cup \{+\infty\}$, положив $S_A(y^*) = \sup_{x \in A} \langle x, y^* \rangle$. Функция S_A выпукла как точная верхняя грань выпуклых (даже линейных) функций. Она называется *опорной функцией множества A*.

Упражнение 9.14. Проверьте, что если B — единичный шар в нормированном пространстве E, то $S_B(y^*) = \|y^*\|$, где справа стоит двойственная норма, т.е. $\|y^*\| = \sup_{\|x\| \le 1} \langle x, y^* \rangle$.

Упражнение 9.15. Проверьте, что опорная функция произвольного подмножества $A \subset E$ равна функционалу Минковского его поляры (где поляра рассматривается как подмножество в E^*).

9.4 Двойственность Фенхеля–Рокафеллера

Пусть $f \colon E \to \mathbb{R}$ — выпуклая функция на нормированном пространстве E.

Преобразованием Лежандра или *преобразованием Юнга–Фенхеля* функции f или просто *двойственной к f функцией* называется функция f^* на топологически двойственном пространстве E^*, заданная формулой

$$f^*(z^*) = \sup_{z \in E} \big(\langle z^*, z \rangle - f(z) \big).$$

Пример 9.16. Пусть $E = E^* = \mathbb{R}$, $\langle z^*, z \rangle = z^* z$, $f(z)$ — произвольная дифференцируемая выпуклая вниз функция. Для каждого фиксированного z^* требуется найти $\sup_z \big(z^* z - f(z)\big)$. Дифференцируя по z, получаем $z^* - f'(z) = 0$. Из выпуклости функции f вытекает, что это уравнение имеет не более одного решения, и если $z = z(z^*)$ — решение, то $f^*(z^*) = z^* z(z^*) - f\big(z(z^*)\big)$. Геометрически это можно себе представить так. На плоскости (z, y) рассматриваем график функции $y = f(z)$ и прямую $y = z^* z$ (здесь z^* — фиксировано и задает угол наклона прямой). Нужно найти точную верхнюю грань значений разности $z^* z - f(z)$. Он очевидно достигается в такой точке z, в которой касательная к графику $y = f(z)$ параллельна прямой $y = z^* z$. Значение $f^*(z^*)$ по модулю равно y-координате точки пересечения этой касательной с осью Oy.

Например, если $f(z) = e^z$, то уравнение принимает вид $z^* = e^z$, что при положительных z^* дает $z = \log z^*$, откуда $f^*(z^*) = z^* \log z^* - z^*$. В этом случае при $z^* = 0$ имеем $f^*(0) = \sup_z(-e^z) = 0$, а при отрицательных z^* величина $z^* z - e^z$ неограниченно растет при $z \to \infty$, поэтому $f^*(z^*) = +\infty$, $z^* < 0$.

Пример 9.17. Пусть $A \subset E$ — произвольное непустое подмножество. Тогда $I_A^*(z^*) = S_A(z^*)$.

Следующий минимакс принцип называют теоремой Фенхеля–Рокафеллера.

Теорема 9.18. *Пусть E — линейное нормированное пространство, E^* — топологически двойственное к нему, и пусть F и G — выпуклые функции на E со значениями в $\mathbb{R} \cup \{+\infty\}$. Пусть F^* и G^* — двойственные к ним функции на E^*. Предположим, что существует точка $z_0 \in E$, в которой обе функции F и G принимают конечные значения и F непрерывна в z_0. Тогда*

$$\inf_{z \in E} \big(F(z) + G(z)\big) = \sup_{z^* \in E^*} \big(- F^*(-z^*) - G^*(z^*)\big) = \max_{z^* \in E^*} \big(- F^*(-z^*) - G^*(z^*)\big),$$

в частности, супремум в правой части равенства достигается.

Доказательство. Прежде всего, заметим, что

$$G^*(z^*) = \sup_{y \in E} \big(\langle z^*, y\rangle - G(y)\big) = -\inf_y \big(G(y) - \langle z^*, y\rangle\big),$$

откуда

$$-G^*(z^*) = \inf_y \big(G(y) - \langle z^*, y\rangle\big) \qquad \text{и} \qquad - F^*(-z^*) = \inf_x \big(F(x) + \langle z^*, x\rangle\big),$$

поэтому нужно проверить следующее равенство:

$$\sup_{z^* \in E^*} \inf_{x,y \in E} \Big[F(x) + G(y) + \langle z^*, x - y\rangle\Big] = \inf_x \Big[F(x) + G(x)\Big],$$

и показать, что супремум в левой части достигается (и равен максимуму).

При $x = y$ инфимум в левой части равен правой и не зависит от z^*, поэтому левая часть не превосходит правой. Остается проверить, что существует такой функционал $z^* \in E^*$, что

$$F(x) + G(y) + \langle z^*, x - y\rangle \geq \inf_x \big(F(x) + G(x)\big) \qquad \text{при всех } x, y \in E.$$

Тогда на нем достигается супремум (и превращается в максимум), и выполняется искомое равенство.

Положим $m = \inf_x \big(F(x) + G(x)\big)$. Рассмотрим множества

$$C = \big\{(x, \lambda) \in E \times \mathbb{R} : F(x) < \lambda\big\}, \qquad D = \big\{(y, \mu) \in E \times \mathbb{R} : \mu \leq m - G(y)\big\}.$$

Из выпуклости функций F и G следует, что множества C и D —выпуклы, кроме того, если $\lambda_0 = F(z_0) + 1$, то $(z_0, \lambda_0) \in C$, и из непрерывности отображения $(z, t) \mapsto \big(F(z), t\big)$ в точке (z_0, λ_0) вытекает, что существует такая окрестность $U \subset E \times \mathbb{R}$ точки (z_0, λ_0), что $F(u) < \lambda$ для всех $(u, \lambda) \in U$, т.е. $U \subset C$, и внутренность C не пуста. Кроме того, если $(x, \lambda) \in C \cap D$, то $F(x) < \lambda \leq m - G(x)$, откуда $F(x) + G(x) < m$, что противоречит выбору m, поэтому $C \cap D = \emptyset$. Заметим также, что $(z_0, \mu) \in D$ при всех $\mu \leq m - G(z_0)$.

В силу леммы 9.11 множества $\operatorname{Int} C$ и D выпуклы, не пусты, $\operatorname{Int} C$ открыто, поэтому применима теорема Хана–Банаха об отделимости (теорема 8.15), согласно которой существует ненулевой непрерывный линейный функционал $\phi \in (E \times \mathbb{R})^*$ и число α такие, что $\phi(x) < \alpha \le \phi(y)$ для любых $x \in \operatorname{Int} C$ и $y \in D$, поэтому (здесь мы снова используем тот факт, что $\overline{\operatorname{Int} C} = \overline{C}$, см. лемму 9.11, и непрерывность ϕ)

$$\sup_{x \in C} \phi(x) = \sup_{x \in \overline{C}} \phi(x) = \sup_{x \in \overline{\operatorname{Int} C}} \phi(x) = \sup_{x \in \operatorname{Int} C} \phi(x) \le \alpha \le \inf_{y \in D} \phi(y).$$

Другими словами, существуют $w^* \in E^*$ и $a \in \mathbb{R}$ (не равные одновременно нулю), для которых

$$\langle w^*, x\rangle + a\lambda \le \alpha \le \langle w*, y\rangle + a\mu$$

при всех (x, λ) и (y, μ), для которых $F(x) < \lambda$ и $\mu \le m - G(y)$.

Лемма 9.19. *В сделанных предположениях, $a < 0$.*

Доказательство. Возьмем $x = z_0$, тогда $F(z_0)$ конечно, и пара $(z_0, \lambda) \in C$ при всех достаточно больших λ, поэтому $a \le 0$ (иначе левая часть неравенства неограниченно растет при фиксированной правой).

Покажем, что $a \ne 0$. Предположим противное, тогда $\langle w^*, x\rangle \le \alpha \le \langle w*, y\rangle$ при всех $(x, \lambda) \in C$, $(y, \mu) \in D$. В частности, это верно для всех (x, λ) из окрестности точки $(z_0, \lambda_0) \subset C$, поэтому при малом $\varepsilon > 0$ неравенство $\langle w^*, z_0 + \varepsilon z\rangle \le \alpha$ выполнено при всех z из подходящей окрестности нуля. С другой стороны, $(z_0, \mu) \in D$ при подходящих μ, поэтому

$$\langle w^*, z_0 + \varepsilon z\rangle \le \alpha \le \langle w^*, z_0\rangle,$$

откуда $\varepsilon\langle w^*, z\rangle \le 0$ при всех z из окрестности нуля, что возможно только при $w^* = 0$. Последнее невозможно, так как функционал ϕ — ненулевой. □

Положим $z^* = w^*/a$. Тогда

$$\langle z^*, x\rangle + \lambda \ge \langle z^*, y\rangle + \mu$$

в частности, это верно при $\lambda = F(x)$ и $\mu = m - G(y)$,

$$\langle z^*, x\rangle + F(x) \ge \langle z^*, y\rangle + m - G(y),$$

откуда

$$F(x) + G(y) + \langle z^*, x - y\rangle \ge m$$

при всех $x, y \in E$, что и требовалось. □

Идея доказательства теоремы 9.1. Основная идея состоит в использовании принципа двойственности или так называемого минимакс-принципа (теоремы 9.18), который позволяет при определенных условиях заменить $\inf\sup$ на $\sup\inf$. Мы займемся обоснованием применимости этой теоремы позже, а здесь покажем только как она может работает.

Перепишем левую часть искомого равенства так:

$$\inf_{\Pi(\mu,\nu)} K(\pi) = \inf_{\pi \in M_+(X \times Y)} \big(K(\pi) + I_{\Pi(\mu,\nu)}(\pi)\big),$$

где $M_+(X \times Y)$ — множество всех неотрицательных борелевских мер на $X \times Y$, а $I_{\Pi(\mu,\nu)}$ — характеристическая функция множества $\Pi(\mu,\nu)$ на $M_+(X \times Y)$.

Далее, сама эта функция может быть записана как супремум так:

$$I_{\Pi(\mu,\nu)}(\pi) = \sup_{(\varphi,\psi)} \Big[\int_X \varphi \, d\mu + \int_Y \psi \, d\nu - \int_{X \times Y} \big(\varphi(x) + \psi(y)\big) \, d\pi \Big],$$

где супремум берется по всем парам непрерывных ограниченных функций (φ, ψ). Таким образом, внося под супремум $K(\pi)$, не зависящий от φ и ψ, получаем:

$$\inf_{\Pi(\mu,\nu)} K(\pi) = \inf_{\pi \in M_+(X \times Y)} \sup_{(\varphi,\psi)} \Big[\int_{X \times Y} c(x,y) \, d\pi + \\ + \int_X \varphi \, d\mu + \int_Y \psi \, d\nu - \int_{X \times Y} \big(\varphi(x) + \psi(y)\big) \, d\pi \Big].$$

Воспользовавшись минимакс-принципом, поменяем супремум и инфимум местами, вынесем за инфимум по π, ставший внутренним, те слагаемые, которые не зависят от π, и заменим инфимум на супремум по формуле $\inf f = -\sup(-f)$. Получим:

$$\inf_{\Pi(\mu,\nu)} K(\pi) = \sup_{(\varphi,\psi)} \inf_{\pi \in M_+(X \times Y)} \Big[\int_{X \times Y} c(x,y) \, d\pi + \\ + \int_X \varphi \, d\mu + \int_Y \psi \, d\nu - \int_{X \times Y} \big(\varphi(x) + \psi(y)\big) \, d\pi \Big] = \\ = \sup_{(\varphi,\psi)} \Big[\int_X \varphi \, d\mu + \int_Y \psi \, d\nu - \sup_{\pi \in M_+(X \times Y)} \int_{X \times Y} \big(\varphi(x) + \psi(y) - c(x,y)\big) \, d\pi \Big].$$

Вычислим внутренний супремум. Заметим, что если функция $\zeta(x,y) = \varphi(x) + \psi(y) - c(x,y)$ принимает положительное значение в некоторой точке (x_0, y_0), то взяв $\pi = \lambda \delta_{(x_0,y_0)}$, где $\lambda \to +\infty$, заключаем, что этот супремум равен $+\infty$. Если же функция $\zeta(x,y)$ принимает $(\mu \otimes \nu)$-почти всюду неположительные значения, т.е. $(\varphi, \psi) \in \Phi_c$, то супремум достигается на нулевой мере π. Поэтому

$$\sup_{\pi \in M_+(X \times Y)} \int_{X \times Y} \big(\varphi(x) + \psi(y) - c(x,y)\big) \, d\pi = I_{\Phi_c}(\varphi, \psi),$$

и, подставив это равенство в предыдущее, заключаем, что

$$\inf_{\Pi(\mu,\nu)} K(\pi) = \sup_{(\varphi,\psi)} \Big[\int_X \varphi \, d\mu + \int_Y \psi \, d\nu - I_{\Phi_c}(\varphi, \psi) \Big] = \\ = \sup_{(\varphi,\psi) \in \Phi_c} \Big[\int_X \varphi \, d\mu + \int_Y \psi \, d\nu \Big] = \sup_{(\varphi,\psi) \in \Phi_c} J(\varphi, \psi),$$

что и требовалось. □