

УСЛОВИЯ ОПТИМАЛЬНОСТИ В ЗАДАЧАХ МАТЕМАТИЧЕСКОГО ПРОГРАММИРОВАНИЯ

По исторически сложившимся причинам *задачей математического программирования* принято называть следующую задачу:

*Найти в E^n максимум функционала $F(x)$,
при условиях :*

$$f_i(x) \leq 0 \quad i = [1, m],$$

или в координатной форме, эта задача записывается:

Найти максимум $F(\xi_1, \xi_2, \dots, \xi_n)$ по $\{\xi_1, \xi_2, \dots, \xi_n\}$,

при условиях:

$$\begin{cases} f_1(\xi_1, \xi_2, \dots, \xi_n) \leq 0, \\ f_2(\xi_1, \xi_2, \dots, \xi_n) \leq 0, \\ \dots \\ f_m(\xi_1, \xi_2, \dots, \xi_n) \leq 0. \end{cases}$$

Если элемент x удовлетворяет всем ограничениям этой задачи, то (для краткости) будем это обозначать как $x \in \Omega$.

Необходимые условия разрешимости задачи математического программирования

Напомним, что ограничение $f_i(x) \leq 0$ называется *активным* на элементе x^* , если $f_i(x^*) \geq 0$.

Пусть

- 1) функции $\{F(x), f_i(x), i = [1, m]\}$ непрерывно дифференцируемы на элементе $x^* \in E^n$,
- 2) $J(x^*)$ – множество индексов активных на элементе x^* ограничений,
- 3) набор элементов $\{\text{grad } f_i(x^*), i \in J(x^*)\}$ линейно независим.

Если ввести обозначение $\delta x = x - x^*$, то оказывается справедливой

Теорема
4.1.1.

Пусть функция $F(x)$ принимает максимальное значение на допустимом элементе x^* . Тогда *необходимое* условие оптимальности имеет вид

$$(\text{grad } F(x^*), \delta x) \leq 0 \quad \forall \delta x \text{ такого, что}$$

$$(\text{grad } f_i(x^*), \delta x) \leq 0, \quad i \in J(x^*).$$

Доказательство.

Вначале покажем, что конус *допустимых вариаций* на элементе x^* описывается системой условий

$$(\text{grad } f_i(x^*), \delta x) \leq 0 \quad \forall i \in J(x^*).$$

Действительно, каждое из активных на допустимом элементе x^* ограничений $f_i(x) \leq 0$ удовлетворяет условию $f_i(x^*) = 0$.

Поэтому в малой окрестности элемента x^* из формулы Тейлора имеем

$$f_i(x) = f_i(x^*) + (\text{grad } f_i(x^*), \delta x) + o(|\delta x|) \leq 0$$

и условие $f_i(x) \leq 0$ аппроксимируется линейным неравенством вида

$$(\text{grad } f_i(x^*), x - x^*) \leq 0.$$

Значит, множество элементов $\delta x = x - x^*$ – допустимых на x^* вариаций – задается системой неравенств

$$(\text{grad } f_i(x^*), \delta x) \leq 0 \quad \forall i \in J(x^*).$$

С другой стороны, множество (конус) *неулучшающих* максимум *вариаций* очевидно определяется условием

$$(\text{grad } F(x^*), \delta x) \leq 0 \quad \forall \delta x$$

А, поскольку из оптимальности элемента x^* следует, что каждая допустимая на этом элементе вариация является *неулучшающей*, то мы приходим к заключению о справедливости утверждения теоремы.

Теорема доказана.

Утверждение теоремы 4.1.1 при замене δx на $x - x^*$ имеет вид

$$(\operatorname{grad} F(x^*), x) \leq (\operatorname{grad} F(x^*), x^*) \quad (4.1.1)$$

$$(\operatorname{grad} f_i(x^*), x) \leq (\operatorname{grad} f_i(x^*), x^*), \quad i \in J(x^*),$$

что, согласно *теореме Фаркаша*, равносильно существованию набора чисел $\{\lambda_1^*, \lambda_2^*, \dots, \lambda_m^*\}$ такого, что

$$\lambda_i^* \geq 0 \quad \text{и} \quad \sum_{i \in J(x^*)} \lambda_i^* \operatorname{grad} f_i(x^*) = \operatorname{grad} F(x^*).$$

Действительно, если для любой допустимой вариации δx систему неравенств (4.1.1) можно привести к виду

$$\begin{aligned}(\operatorname{grad} F(x^*), \delta x) &\leq 0, \\ (\operatorname{grad} f_i(x^*), \delta x) &\leq 0 \quad \forall i \in J(x^*)\end{aligned}$$

или, в координатной форме в ортонормированном баазисе

$$\begin{cases} \sum_{j=1}^n \frac{\partial F}{\partial x_j} \delta x_j \leq 0, \\ \sum_{j=1}^n \frac{\partial f_i}{\partial x_j} \delta x_j \leq 0 \quad \forall i \in J(x^*). \end{cases}$$

Если предположить, что *каждое* решение второй системы удовлетворяет первому неравенству, то по теореме Фаркаша должно существовать решение системы линейных уравнений вида

$$\sum_{i \in J(x^*)} \frac{\partial f_i}{\partial x_j} \lambda_i^* = \frac{\partial F}{\partial x_j} \quad \forall j = [1, n],$$

такое, что $\lambda_i^* \geq 0 \quad \forall i \in J(x^*)$.

Иначе говоря, выполняются условия:

$$\lambda_i^* \geq 0 \quad \text{и} \quad \sum_{i \in J(x^*)} \lambda_i^* \text{grad } f_i(x^*) = \text{grad } F(x^*).$$

Заметим, что при использовании обозначений из доказательства теоремы Фаркаша, будут справедливы равенства:

$$\|b\| = \left\| \frac{\partial F}{\partial x_j} \right\| \quad \text{и} \quad \|A\| = \left\| \frac{\partial f_i}{\partial x_j} \right\|.$$

Если доопределить $\lambda_i^* = 0 \quad \forall i \notin J(x^*)$ и ввести в рассмотрение элемент $\Lambda \in E^m$, то последнее равенство принимает вид

$$\begin{aligned} \text{grad } F(x^*) - \sum_{i=1}^m \lambda_i^* \text{grad } f_i(x^*) &= 0, \\ \lambda_i^* f_i(x^*) &= 0, \quad \lambda_i^* \geq 0 \quad \forall i = [1, m]. \end{aligned}$$

Таким образом, оказывается справедливой

Теорема
4.1.2.
(Каруша-
Куна-
Таккера)

Пусть функциз $F(x)$ принимает максимальное значение на допустимом элементе x^* и элементы $\text{grad } f_i(x^*) \quad i \in J(x^*)$ линейно независимы. Тогда существуют числа $\lambda_i^* \quad \forall i = [1, m]$ такие, что

$$\begin{aligned} \text{grad } F(x^*) - \sum_{i=1}^m \lambda_i^* \text{grad } f_i(x^*) &= 0, \\ \lambda_i^* f_i(x^*) &= 0, \quad \lambda_i^* \geq 0 \quad \forall i = [1, m]. \end{aligned}$$

Отметим также, что, равенства

$$\lambda_i^* f_i(x^*) = 0 \quad \forall i = [1, m]$$

принято называть условиями *дополняющей нежесткости*.

Полученные выше необходимые условия экстремальности удобнее формулировать, введя в рассмотрение специальную функцию, зависящую как от x , так и от $\Lambda = \{\lambda_1, \lambda_2, \dots, \lambda_m\}$ и имеющую вид:

$$L(x, \Lambda) = F(x) - \sum_{i=1}^m \lambda_i f_i(x).$$

В этом случае необходимые условия экстремальности формулируются как

Теорема
4.1.3.

Пусть функционал $F(x)$ на множестве Ω принимает максимальное значение на элементе x^* и элементы вектора $\text{grad } f_i(x^*) \quad \forall i \in J(x^*)$ линейно независимы.

Тогда существуют числа $\lambda_i^* \geq 0 \quad \forall i = [1, m]$ такие, что

$$\begin{aligned}\text{grad}_x L(x^*, \Lambda^*) &= 0, \\ \lambda_i^* f_i(x^*) &= 0 \quad \forall i = [1, m].\end{aligned}$$

В теории математического программирования числа $\lambda_i^* \geq 0 \quad i = [1, m]$ принято называть *множителями Лагранжа*, а функцию $L(x, \Lambda)$ – *функцией Лагранжа*.

Функция Лагранжа и ее свойства

При исследовании свойств функции Лагранжа полезными оказываются следующие теоремы.

Теорема 4.2.1. **Справедливо равенство**

$$\max_{x \in \Omega} F(x) = \max_x \min_{\Lambda \geq 0} L(x, \Lambda).$$

Доказательство.

В силу неотрицательности чисел λ_i $i = [1, m]$ и структуры функции Лагранжа справедливо равенство

$$\min_{\Lambda \geq 0} L(x, \Lambda) = \begin{cases} F(x) & x \in \Omega, \\ -\infty & x \notin \Omega. \end{cases}$$

Тогда, если x^* – решение исходной задачи, то $x^* \in \Omega$ и

$$F(x^*) = \max_{x \in \Omega} F(x) = \max_x \min_{\Lambda \geq 0} L(x, \Lambda).$$

Обратно: если x^* – оптимальный элемент функционала

$$\max_x \min_{\Lambda \geq 0} L(x, \Lambda),$$

то $x^* \in R$ и справедливо $\max_{x \in \Omega} F(x) = \max_x \min_{\Lambda \geq 0} L(x, \Lambda).$

Теорема доказана.

Следствие 4.2.1. **Задачи поиска $\max_{x \in \Omega} F(x)$ и $\max_x \min_{\Lambda \geq 0} L(x, \Lambda)$ равносильны.**

Задачу поиска

$$\min_{\Lambda \geq 0} \max_x L(x, \Lambda)$$

принято называть *двойственной* задачей

к *прямой* задаче

$$\max_x \min_{\Lambda \geq 0} L(x, \Lambda).$$

Теорема
4.2.2.

Справедливо соотношение

$$\min_{\Lambda \geq 0} \max_x L(x, \Lambda) \geq \max_x \min_{\Lambda \geq 0} L(x, \Lambda).$$

Доказательство.

Очевидно, что $L(x, \Lambda) \geq \min_{\Lambda \geq 0} L(x, \Lambda)$.

Но тогда, как частный случай, верна оценка

$$\forall \Lambda \geq 0 : \quad \max_x L(x, \Lambda) \geq \max_x \min_{\Lambda \geq 0} L(x, \Lambda),$$

из которой следует, что

$$\min_{\Lambda \geq 0} \max_x L(x, \Lambda) \geq \max_x \min_{\Lambda \geq 0} L(x, \Lambda).$$

Теорема доказана.

Теоремы 4.2.1 и 4.2.2 справедливы без каких-либо предположений о выпуклости прямой задачи. Их утверждения позволяют получать в общем случае *верхнюю* оценку ее решения. Наложение же дополнительных условий, приводит к более сильным оценкам, таким как

Теорема
4.2.3.

Пусть $F(x)$ *выпукла вверх* на *выпуклом* множестве Ω , имеющем *внутренние* элементы (*условие регулярности Слейтера*), тогда верны равенства

$$\min_{\Lambda \geq 0} \max_x L(x, \Lambda) = L(x^*, \Lambda^*) = \max_x \min_{\Lambda \geq 0} L(x, \Lambda).$$

Следствие
4.2.2.

Пара элементов $\{x^*, \Lambda^*\}$ есть *седловая точка* функции Лагранжа.

**Замечание о достаточных условиях разрешимости задачи
математического программирования**

Подобно случаю поиска безусловного экстремума, для задачи математического программирования можно сформулировать *достаточные условия второго порядка* существования максимума $F(x)$ на элементе $x^* \in \Omega$.

Эти условия не являются необходимыми.

Примером такого условия может служить

Теорема
 4.3.1.

Пусть существует элемент x^* и набор чисел

$$\lambda_i^* \geq 0 ; i = [1, m]$$

таких, что

$$\begin{aligned} f_i(x^*) &\leq 0 \quad i = [1, m], \\ \lambda_i^* f_i(x^*) &= 0 \quad i = [1, m] \end{aligned} \quad \text{и}$$

$$\text{grad}_x L(x^*, \Lambda^*) = 0$$

и пусть для любого элемента $z \in E^n$ такого, что

$$(\text{grad}_x f_i(x^*), z) = 0 \quad \forall i \in J^*,$$

$$\text{где } J^* : \{i \mid \lambda_i^* > 0\}$$

и

$$(\text{grad}_x f_i(x^*), z) \leq 0 \quad \forall i \in G,$$

$$\text{где } G : \{i \mid \lambda_i^* = 0 \cap f_i(x^*) \geq 0\}$$

выполняется неравенство

$$\left(z, \overset{\wedge}{\text{Hess}_x L(x^*, \Lambda^*)} z \right) < 0$$

тогда x^* — решение исходной задачи математического программирования.

Двойственные условия оптимальности в линейном программировании

Для задач линейного программирования необходимые и достаточные условия оптимальности (Каруша-Куна–Таккера) могут быть получены путем применения теоремы Фаркаша.

Однако, специфическая форма постановки задачи ЛП позволяет получить эти условия в виде *другой задачи линейного программирования*, сформулированной в двойственном (сопряженном) пространстве.

Напомним, что *прямая* задача формулируется как

$$\max_x \min_{\Lambda \geq 0} L(x, \Lambda),$$

а *двойственная* ей задача имеет вид $\min_{\Lambda \geq 0} \max_x L(x, \Lambda).$

Рассмотрим следующую задачу ЛП:

найти максимум $\sum_{j=1}^n \sigma_j \xi_j$ на $\{\xi_1, \xi_2, \dots, \xi_n\} \in E^n$,

при условиях:

$$\xi_j \geq 0, \quad j = [1, n],$$

$$\sum_{j=1}^n \alpha_{ij} \xi_j \leq \beta_i, \quad i = [1, m]$$

Или же в символическом виде

Найти максимум (c, x) на $x \in E^n$,

при условиях: $x \geq 0, \quad Ax \leq b$.

Функция Лагранжа для этой задачи будет иметь вид

$$L(x, \Lambda) = (c, x) - (\Lambda, Ax - b) = (c, x) + (b, \Lambda) - (\Lambda, Ax)$$

Для того, чтобы получить стандартную формулировку двойственной задачи, преобразуем функцию Лагранжа к виду

$$L(x, \Lambda) = (\Lambda, b) + (c - A^T \Lambda, x).$$

Решение задачи $\max_x L(x, \Lambda)$ в силу условия $x \geq 0$ имеет вид:

$$\max_x L(x, \Lambda) = \begin{cases} (\Lambda, b), & \text{если } c - A^T \Lambda \leq 0, \\ +\infty, & \text{если } c - A^T \Lambda > 0. \end{cases}$$

Поэтому двойственная задача $\min_{\Lambda \geq 0} \max_x L(x, \Lambda)$ будет иметь следующую формулировку:

Найти минимум (Λ, b) по $\Lambda \in E^m$,

при условиях: $A^T \Lambda \geq c, \quad \Lambda \geq 0$.

или в координатной форме:

$$\begin{aligned} \text{найти минимум} \quad & \sum_{i=1}^m \beta_i \lambda_i \quad \text{на} \quad \{\lambda_1, \lambda_2, \dots, \lambda_n\} \in E^m, \\ \text{при условиях:} \quad & \lambda_i \geq 0; \quad i = [1, m], \\ & \sum_{i=1}^m \alpha_{ij} \lambda_i \geq \sigma_j, \quad j = [1, n]. \end{aligned}$$

Преимуществом такой записи прямой задачи линейного программирования является *симметричность* форм записи прямой и двойственной задач.

При этом стоит отметить, что переменные прямой задачи не входят в постановку двойственной только в *линейном* случае.

Проверим, что в линейном случае выполняется

Теорема
4.2.2.

Справедливо соотношение

$$\min_{\Lambda \geq 0} \max_x L(x, \Lambda) \geq \max_x \min_{\Lambda \geq 0} L(x, \Lambda).$$

Имеем прямую задачу

найти максимум $\sum_{j=1}^n \sigma_j \xi_j$ на $\{\xi_1, \xi_2, \dots, \xi_n\} \in E^n$,

при условиях:

$$\xi_j \geq 0, \quad j = [1, n],$$

$$\sum_{j=1}^n \alpha_{ij} \xi_j \leq \beta_i, \quad i = [1, m]$$

и двойственную задачу

найти минимум $\sum_{i=1}^m \beta_i \lambda_i$ на $\{\lambda_1, \lambda_2, \dots, \lambda_n\} \in E^m$,

при условиях: $\lambda_i \geq 0$; $i = [1, m]$,

$$\sum_{i=1}^m \alpha_{ij} \lambda_i \geq \sigma_j \quad j = [1, n].$$

Для любых допустимых $\xi_j \quad j = [1, n]$ и $\lambda_i \quad i = [1, m]$

из $\sum_{j=1}^n \alpha_{ij} \xi_j \leq \beta_i \quad i = [1, m]$ имеем

$$\lambda_i \sum_{j=1}^n \alpha_{ij} \xi_j \leq \beta_i \lambda_i \quad i = [1, m] \quad \Rightarrow \quad \sum_{i=1}^m \sum_{j=1}^n \alpha_{ij} \xi_j \lambda_i \leq \sum_{i=1}^m \beta_i \lambda_i.$$

С другой стороны из $\sum_{i=1}^m \alpha_{ij} \lambda_i \geq \sigma_j \quad j = [1, n]$ получаем

$$\xi_j \sum_{i=1}^m \alpha_{ij} \lambda_i \geq \sigma_j \xi_j \quad j = [1, n] \quad \Rightarrow \quad \sum_{j=1}^n \sum_{i=1}^m \alpha_{ij} \xi_j \lambda_i \geq \sum_{j=1}^n \sigma_j \xi_j.$$

Откуда $\sum_{i=1}^m \beta_i \lambda_i \geq \sum_{j=1}^n \sigma_j \xi_j$ при всех допустимых $\xi_j \quad j = [1, n]$ и

$$\lambda_i \quad i = [1, m]. \text{ Тогда и } \sum_{i=1}^m \beta_i \lambda_i^* \geq \sum_{j=1}^n \sigma_j \xi_j^*.$$

**Пояснения к доказательству
теоремы Каруша-Куна-Таккера**

Теорема Фаркаша Для того чтобы $\ A\ \ x\ = \ b\$ – система m линейных уравнений с n неизвестными имела неотрицательное частное решение $\ x^0\ \geq \ o\$, необходимо и достаточно, чтобы $\ y\$ – каждое частное решение сопряженной системы линейных неравенств $\ A\ ^T \ y\ \leq \ o\ $ – удовлетворяло условию $\ b\ ^T \ y\ \leq 0.$ Пусть $\ x\ = \ \lambda\$ и $\ y\ = \ \delta x\$. $\ b\ = \left\ \frac{\partial F}{\partial x_j} \right\ \quad \text{и}$ $\ A\ = \left\ \frac{\partial f_i}{\partial x_j} \right\ .$	Задача математического программирования Найти в E^n максимум функции $F(x)$, при условии: $f_i(x) \leq 0 \quad i = [1, m]$, Необходимое условие оптимальности $(\text{grad } F(x^*), \delta x) \leq 0$ $\forall \delta x : (\text{grad } f_i(x^*), \delta x) \leq 0, \quad \forall i \in J(x^*)$ Тогда, по теореме Фаркаша $\sum_{i \in J(x^*)} \frac{\partial f_i}{\partial x_j} \lambda_i^* = \frac{\partial F}{\partial x_j} \quad \forall j = [1, n],$ с $\lambda_i^* \geq 0$ или $\sum_{i \in J(x^*)} \lambda_i^* \text{grad } f_i(x^*) = \text{grad } F(x^*)$
--	---

Теперь покажем, что справедлива

Теорема 4.2.3. Пусть $F(x)$ выпукла вверх на выпуклом множестве Ω , имеющем внутренние элементы (условие регулярности Слейтера), тогда верны равенства

$$\min_{\Lambda \geq 0} \max_x L(x, \Lambda) = L(x^*, \Lambda^*) = \max_x \min_{\Lambda \geq 0} L(x, \Lambda).$$

Доказательство.

Поскольку утверждение теоремы 3.5.2.2 справедливо и в рассматриваемом случае, для доказательства достаточно убедиться в справедливости оценки

$$\min_{\Lambda \geq 0} \max_x L(x, \Lambda) \leq \max_x \min_{\Lambda \geq 0} L(x, \Lambda).$$

Пусть x^* решение прямой задачи. Рассмотрим евклидово простран-

ство E^{m+1} с элементами y такими, что $\|y\| = \begin{pmatrix} \eta_0 \\ \eta_1 \\ \dots \\ \eta_m \end{pmatrix}$. В этом простран-

стве выделим множество Ω с элементами, удовлетворяющими системе неравенств

$$\begin{cases} \eta_0 \leq F(x), \\ \eta_1 \leq -f_1(x), \\ \dots \\ \eta_m \leq -f_m(x), \end{cases}$$

для некоторого фиксированного допустимого элемента x , а также множество Θ с элементами, удовлетворяющими системе условий вида

$$\begin{cases} F(x^*) < \eta_0, \\ \eta_1 = 0, \\ \dots \\ \eta_m = 0. \end{cases}.$$

Множества Ω и Θ по условию теоремы *выпуклы* и, что очевидно, по построению не имеют общих элементов в E^{m+1} .

Поэтому, в силу теоремы о разделяющей гиперплоскости, можно утверждать, что существует разделяющая множества Ω и Θ гиперплоскость. Например, вида

$$(l^*, y - y^*) = 0,$$

где в качестве y^* (т.е. точки, через которую эта гиперплоскость проходит) можно принять элемент, y которого

$$\|y^*\| = \left\| \begin{pmatrix} \eta_0^* \\ \eta_1^* \\ \dots \\ \eta_m^* \end{pmatrix} \right\| = \left\| \begin{pmatrix} F(x^*) \\ 0 \\ \dots \\ 0 \end{pmatrix} \right\|, \text{ а в качестве ненулевого (нормального) эле-}$$

$$\text{мента } l^* \neq 0 \text{ взять какой-то элемент, } \|l^*\| = \left\| \begin{pmatrix} \lambda_0^* \\ \lambda_1^* \\ \dots \\ \lambda_m^* \end{pmatrix} \right\| \text{ с неотрицатель-}$$

ными компонентами, поскольку множеству Ω принадлежат элементы со сколь угодно большими по модулю и отрицательными по знаку координатами.

Опять-таки, в силу теоремы о разделяющей гиперплоскости, в этом случае $\forall y \in \Omega$ будет справедлива оценка $(l^*, y) \leq (l^*, y^*)$, или, что то же самое

$$\lambda_0^* \eta_0^* + \sum_{i=1}^m \lambda_i^* \eta_i^* \leq \lambda_0^* F(x^*) \quad \forall x \in R,$$

причем $l^* \geq 0$,

Последнее неравенство верно $\forall y \in \Omega$, поэтому оно будет верным и

для элемента $\|y\| = \left\| \begin{pmatrix} \eta_0 \\ \eta_1 \\ \dots \\ \eta_m \end{pmatrix} \right\| = \left\| \begin{pmatrix} F(x) \\ -f_1(x) \\ \dots \\ -f_m(x) \end{pmatrix} \right\|$, что дает оценку

$$\lambda_0^* F(x) - \sum_{i=1}^m \lambda_i^* f_i(x) \leq \lambda_0^* F(x^*); \quad \forall x.$$

Заметим также, что $\lambda_0^* > 0$. Действительно, из предположения $\lambda_0^* = 0$ в силу $\lambda_j^* \geq 0 \quad \forall j = [1, m]$ получаем, что

$$\sum_{j=1}^m \lambda_j^* f_j(x) \geq 0 \quad \forall x,$$

а это противоречит условию *регулярности Слейтера*, то есть предположению, что у множества R есть строго внутренние точки.

Если $\lambda_0^* > 0$, то можно пронормировать это неравенство, поделив обе его части на λ_0^* . Тогда, сохранив (после нормировки) прежние обозначения для чисел $\lambda_i \quad i = [1, m]$, получаем оценку

$$F(x) - \sum_{i=1}^m \lambda_i^* f_i(x) \leq F(x^*) \quad \forall x.$$

При этом в силу произвольности x будет справедливо и неравенство

$$\max_x \left[F(x) - \sum_{i=1}^m \lambda_i^* f_i(x) \right] \leq F(x^*),$$

то есть, что

$$\max_x L(x, \Lambda^*) \leq F(x^*).$$

Наконец, из очевидного

$$\max_x L(x, \Lambda^*) \geq \min_{\Lambda \geq 0} \max_x L(x, \Lambda),$$

учитывая, что $F(x^*) = \max_{x \in R} F(x) = \max_x \min_{\Lambda \geq 0} L(x, \Lambda^*)$, получаем требуемое неравенство

$$\max_x \min_{\Lambda \geq 0} L(x, \Lambda) \geq \min_{\Lambda \geq 0} \max_x L(x, \Lambda).$$

И в сочетании с ранее полученным (теорема 4.2.2) неравенством

$$\min_{\Lambda \geq 0} \max_x L(x, \Lambda) \geq \max_x \min_{\Lambda \geq 0} L(x, \Lambda)$$

приходим к

$$\min_{\Lambda \geq 0} \max_x L(x, \Lambda) = L(x^*, \Lambda^*) = \max_x \min_{\Lambda \geq 0} L(x, \Lambda).$$

Теорема доказана.