ЗАДАЧА МАТЕМАТИЧЕСКОГО ПРОГРАММИРОВАНИЯ

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

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

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

или в координатной форме эта задача записывается:

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

при условиях:
$$\begin{cases} f_1(\xi_1, \xi_2, \ldots, \xi_n) \le 0, \\ f_2(\xi_1, \xi_2, \ldots, \xi_n) \le 0, \\ \qquad \ldots \\ f_m(\xi_1, \xi_2, \ldots, \xi_n) \le 0. \end{cases}$$

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

Необходимые условия разрешимости задачи математического программирования

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

Пусть

1) функционалы $\{ F(x), f_i(x), i = [1, m] \}$ 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 функционал $F(x)$ принимает максимальное значение на допустимом элементе x^*. Тогда *необходимое условие оптимальности* имеет вид**

$$(\operatorname{grad} F(x^*), \delta x) \le 0 \qquad \forall \delta x \quad \textbf{такого, что}$$

$$(\operatorname{grad} f_i(x^*), \delta x) \le 0\,, \;\; i \in J(x^*).$$

Доказательство.

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

$$(\operatorname{grad} f_i(x^*), \delta x) \le 0 \qquad \forall i \in J(x^*)\,.$$

Действительно, каждое из активных на допустимом элементе x^* ограничений $f_i(x) \le 0$ удовлетворяет условию $f_i(x^*) = 0$. Поэтому в малой окрестности элемента x^* из формулы Тейлора имеем

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

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

$$(\operatorname{grad} f_i(x^*), x - x^*) \le 0\,.$$

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

$$(\operatorname{grad} f_i(x^*), \delta x) \le 0 \qquad \forall i \in J(x^*)\,.$$

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

$$(\operatorname{grad} F(x^*), \delta x) \le 0 \qquad \forall \delta x$$

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

$$(\operatorname{grad} F(x^*), x) \le (\operatorname{grad} F(x^*), x^*) \tag{4.1.1}$$

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

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

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

Действительно, если систему неравенств (4.1.1) привести к виду

$$(\operatorname{grad} F(x^*), \delta x) \le 0$$

$$\forall \delta x : (\operatorname{grad} f_i(x^*), \delta x) \le 0, \qquad \forall i \in J(x^*)$$

или, в координатной форме

$$\sum_{j=1}^{n} \frac{\partial F}{\partial x_j} \delta x_j \le 0$$

$$\forall \delta x : \quad \sum_{j=1}^{n} \frac{\partial f_i}{\partial x_j} \delta x_j \le 0 \quad \forall i \in J(x^*)$$

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

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

такое, что $\lambda^*{}_i \geq 0 \; \forall i \in J(x^*)$.

Иначе говоря, выполняются условия:

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

Заметим, что, если использовать обозначения из теоремы Фаркаша, то 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 $\lambda_i^* = 0$, $\forall i \notin J(x^*)$ и ввести в рассмотрение элемент $\Lambda \in E^m$, то последнее равенство принимает вид

$$\operatorname{grad} F(x^*) - \sum_{i=1}^{m} \lambda_i^* \operatorname{grad} f_i(x^*) = o;$$
$$\lambda_i^* f_i(x^*) = 0; \quad \lambda_i^* \geq 0; \quad \forall i = [1, m].$$

Таким образом, оказывается справедливой

Теорема 4.1.2. (Каруша-Куна–Таккера)

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

$$\operatorname{grad} F(x^*) - \sum_{i=1}^{m} \lambda_i^* \operatorname{grad} f_i(x^*) = o;$$
$$\lambda_i^* f_i(x^*) = 0; \quad \lambda_i^* \geq 0; \quad \forall i = [1, m].$$

Отметим также, что, равенства

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

принято называть условиями *дополняющей нежесткости*.

Полученные выше необходимые условия экстремальности удобнее формулировать, введя в рассмотрение специальный функционал, зависящий как от x, так и от $\Lambda = \{\lambda_1, \lambda_2, \ldots, \lambda_m\}$ и имеющий 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 R принимает максимальное значение на элементе x^* и элементы**

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

линейно независимы.

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

$$\operatorname{grad}_x L(x^*, \Lambda^*) = o\;;$$
$$\lambda_i^* f_i(x^*) = 0 \quad \forall i = [1, m].$$

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

Функция Лагранжа и ее свойства

При исследовании свойств функции Лагранжа полезными оказываются следующие теоремы.

Теорема 4.2.1. **Справедливо равенство**

$$\max_{x\in R} F(x) = \max_{x} \min_{\Lambda\geq o} L(x,\Lambda).$$

Доказательство.

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

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

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

$$F(x^*) = \max_{x\in R} F(x) = \max_{x} \min_{\Lambda\geq o} L(x,\Lambda).$$

Обратно: если x^* – оптимальный элемент функционала

$$\max_{x} \min_{\Lambda\geq o} L(x,\Lambda),$$

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

Теорема доказана.

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

Задачу поиска

$$\min_{\Lambda \geq o} \max_{x} L(x, \Lambda)$$

принято назвать *двойственной*

к *прямой* задаче

$$\max_{x} \min_{\Lambda \geq o} L(x, \Lambda).$$

Теорема 4.2.2. **Справедливо соотношение**

$$\min_{\Lambda \geq o} \max_{x} L(x, \Lambda) \geq \max_{x} \min_{\Lambda \geq o} L(x, \Lambda).$$

Доказательство.

Очевидно, что $L(x, \Lambda) \geq \min\limits_{\Lambda \geq o} L(x, \Lambda)$.

Но тогда, как частный случай, верна оценка

$$\forall \Lambda \geq o: \max_{x} L(x, \Lambda) \geq \max_{x} \min_{\Lambda \geq o} L(x, \Lambda),$$

из которой следует, что

$$\min_{\Lambda \geq o} \max_{x} L(x, \Lambda) \geq \max_{x} \min_{\Lambda \geq o} L(x, \Lambda).$$

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 R, имеющем внутренние элементы (условие регулярности Слейтера), тогда верны равенства**

$$\min_{\Lambda \ge o} \max_{x} L(x,\Lambda) = L(x^*,\Lambda^*) = \max_{x} \min_{\Lambda \ge o} L(x,\Lambda).$$

Следствие 4.2.2. **Пара элементов $\{x^*,\Lambda^*\}$ есть *седловая точка* функции Лагранжа.**

Достаточные условия 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 $x^* \in R$.

Теорема 4.3.1. **Пусть существует элемент x^* и набор чисел**

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

таких, что

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

и

$$\operatorname{grad}_x L(x^*, \Lambda^*) = o$$

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

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

где $J^* : \left\{ i \,\middle|\, \lambda_i^* > 0 \right\}$

и

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

где $G : \left\{ i \,\middle|\, \lambda_i^* = 0 \bigcap f_i(x^*) \geq 0 \right\}$

выполняется неравенство

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

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

$$\max_{x} \min_{\Lambda \geq o} L(x, \Lambda),$$

а *двойственная* ей задача имеет вид $\min\limits_{\Lambda \geq o} \max\limits_{x} L(x, \Lambda)$.

Рассмотрим следующую задачу ЛП:

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

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: $x \geq o$, $Ax \leq b$.

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

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

Решение задачи $\max\limits_x L(x,\Lambda)$ в силу условия $x \ge 0$ имеет вид:

$$\max_x L(x,\Lambda) = \left[\begin{array}{ll} (\Lambda,b), & \text{если } c - A^{\mathrm{T}}\Lambda \le o, \\ +\infty, & \text{если } c - A^{\mathrm{T}}\Lambda > o. \end{array}\right.$$

Поэтому двойственная задача $\min\limits_{\Lambda \ge o} \max\limits_x L(x,\Lambda)$ будет иметь следующую формулировку:

Найти минимум (Λ,b) *по* $\Lambda \in E^m$,
при условиях: $A^{\mathrm{T}}\Lambda \ge c$, $\Lambda \ge o$.

или в координатной форме:

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

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

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

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**

$$\min_{\Lambda \geq o} \max_{x} L(x, \Lambda) \geq \max_{x} \min_{\Lambda \geq o} L(x, \Lambda).$$

Имеем прямую задачу

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

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* $\{\lambda_1, \lambda_2, \ldots, \lambda_n\} \in E^m$,

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

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

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

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

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

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

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

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

Пояснения к доказательству теоремы Каруша-Куна-Таккера

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

Теперь покажем, что справедлива

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

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

Доказательство.

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

$$\min_{\Lambda \geq o} \max_{x} L(x,\Lambda) \leq \max_{x} \min_{\Lambda \geq o} L(x,\Lambda).$$

Пусть x^* решение прямой задачи. Рассмотрим евклидово пространство E^{m+1} с элементами y такими, что $\| y \| = \left\| \begin{matrix} \eta_0 \\ \eta_1 \\ \cdots \\ \eta_m \end{matrix} \right\|$. В этом пространстве выделим множество Ω с элементами, удовлетворяющими системе неравенств

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

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, у которого $\| y^* \| = \left\| \begin{matrix} \eta_0^* \\ \eta_1^* \\ \dots \\ \eta_m^* \end{matrix} \right\| = \left\| \begin{matrix} F(x^*) \\ 0 \\ \dots \\ 0 \end{matrix} \right\|$, а в качестве *ненулевого* (нормального) элемента $l^* \neq o$ взять какой-то элемент, $\| l^* \| = \left\| \begin{matrix} \lambda_0^* \\ \lambda_1^* \\ \dots \\ \lambda_m^* \end{matrix} \right\|$ с *неотрицательными* 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 $(l^*, y) \le (l^*, y^*)$, или, что то же самое

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

причем $l^* \ge o$,

Последнее неравенство верно $\forall y \in \Omega$, поэтому оно будет верным и для элемента $\| y \| = \begin{Vmatrix} \eta_0 \\ \eta_1 \\ \cdots \\ \eta_m \end{Vmatrix} = \begin{Vmatrix} F(x) \\ -f_1(x) \\ \cdots \\ -f_m(x) \end{Vmatrix}$, что дает оценку

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

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

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

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

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

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

$$\max_{x} L(x,\Lambda^{*}) \geq \min_{\Lambda \geq o} \max_{x} L(x,\Lambda),$$

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

$$\max_{x} \min_{\Lambda \geq o} L(x,\Lambda) \geq \min_{\Lambda \geq o} \max_{x} L(x,\Lambda).$$

И в сочетании с ранее полученным (теорема 4.2.2) неравенством

$$\min_{\Lambda \geq o} \max_{x} L(x,\Lambda) \geq \max_{x} \min_{\Lambda \geq o} L(x,\Lambda)$$

приходим к

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

Теорема доказана.