ЗАДАЧА МАТЕМАТИЧЕСКОГО ПРОГРАММИРОВАНИЯ

По исторически сложившимся причинам *задачей математического программирования* принято называть задачу поиска в E^n элементов,

доставляющих максимум функции $F(x)$,

на множестве элементов $x \in R$,

т.е. удовлетворяющих условиям вида:

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

или в развернутой форме

$$\begin{cases} f_1(x) \leq 0, \\ f_2(x) \leq 0, \\ \quad \dots \\ f_m(x) \leq 0. \end{cases}$$

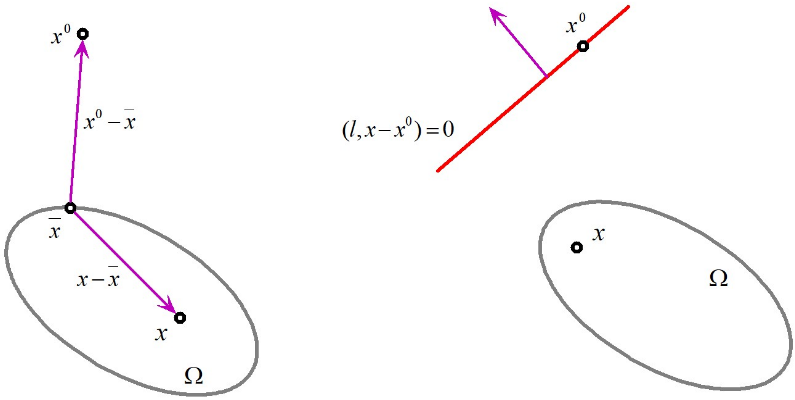
Соответственно в координатной форме эта задача записывается:

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

при условиях:

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

Отметим, что ограничения типа «равенство» могут присутствовать в условии этой задачи в форме двух «встречных» неравенств, например, условие $f_i(x) = 0$ представимо в виде системы $\begin{cases} f_i(x) \le 0 \\ -f_i(x) \le 0 \end{cases}$.

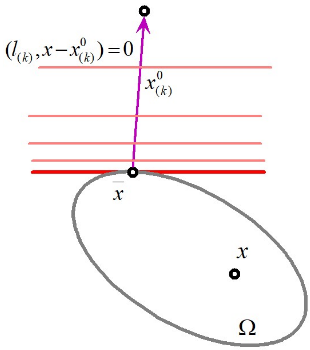


1) 2)

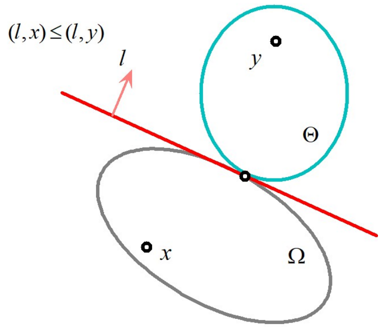
1) Имеет место $\forall x \in \Omega \quad (x - \bar{x}, x^0 - \bar{x}) \le 0$.

2) Существует гиперплоскость $(l, x - x^0) = 0$, такая что

$$(l, x - x^0) < 0 \quad \forall x \in \Omega$$



3)



4)

3) Существует гиперплоскость $(l, x - \bar{x}) = 0$, такая что
$(l, x - \bar{x}) \le 0 \quad \forall x \in \Omega$

4) Существует гиперплоскость $(l, x - x^*) = 0$, такая что
$(l, x) \le (l, y) \quad \forall x \in \Omega \quad \forall y \in \Theta$.

Необходимые условия разрешимости задачи математического программирования

Теорема 3.5.1.1. **Пусть функционал $F(x)$ на множестве R принимает максимальное значение на элементе x^*. Тогда необходимое условие оптимальности имеет вид**

$$(\operatorname{grad} F(x^*), \delta x) \le 0 \quad \text{и} \quad \forall \delta x \in M(x^*) \ :$$

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

Утверждение теоремы 3.5.1.1 при замене δx на $x - x^*$ имеет вид

$$\begin{gathered}(\operatorname{grad} F(x^*), x) \le (\operatorname{grad} F(x^*), x^*) \\ \forall x : (\operatorname{grad} f_i(x^*), x) \le (\operatorname{grad} f_i(x^*), x^*),\ i \in J(x^*)\,,\end{gathered} \tag{3.5.1.1}$$

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

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

Теорема (Фаркаша)

Для того чтобы $\| A \| \| x \| = \| b \|$ – система m линейных уравнений с n неизвестными имела *неотрицательное* частное решение (т.е. $x_k^0 \geq 0 \; \forall k = [1, n]$), необходимо и достаточно, чтобы $\| y \|$ – каждое частное решение системы линейных неравенств вида $\| A \|^{\mathrm{T}} \| y \| \leq \| o \|$ – удовлетворяло условию

$$\| b \|^{\mathrm{T}} \| y \| \leq 0 .$$

Действительно, если систему неравенств (3.5.1.1) привести к виду

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

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

или, в координатной форме

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

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

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 $\lambda^*{}_i \ge 0 \; \forall i \in J(x^*)$.

Иначе говоря, выполняются условия:

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

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 3.5.1.2. (Каруша-Куна–Таккера) **Пусть функционал $F(x)$ на множестве R принимает максимальное значение на 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**

$$\operatorname{grad} F(x^*) - \sum_{i=1}^{m} \lambda_i^* \operatorname{grad} f_i(x^*) = o;$$

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

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 3.5.1.3. **Пусть функционал $F(x)$ на множестве R принимает максимальное значение на элементе x^* и элементы**

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

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 3.5.2.1. **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 $\max\limits_{x\in R} F(x) = \max\limits_{x} \min\limits_{\Lambda\geq o} L(x,\Lambda)$.

Теорема доказана.

Следствие 3.5.2.1. **Задачи поиска $\max\limits_{x\in R} F(x)$ и $\max\limits_{x} \min\limits_{\Lambda\geq o} L(x,\Lambda)$ равносильны.**

Задачу поиска

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

принято назвать *двойственной* к *прямой* задаче

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

Теорема 3.5.2.2. **Справедливо соотношение**

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

Доказательство.

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

Но тогда, как частный случай, верна оценка

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

из которой следует, что

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

Теорема доказана.

Теоремы 3.5.2.1 и 3.5.2.2 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 3.5.2.3. **Пусть $F(x)$ *выпукла вверх* на *выпуклом* множестве R, имеющем внутренние элементы (условие регулярности Слейтера), тогда верны равенства**

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

Следствие 3.5.2.2. **Пара элементов $\{x^*,\Lambda^*\}$ есть *седловая точка* функции Лагранжа.**

ПОСТАНОВКА ЗАДАЧИ ЛИНЕЙНОГО ПРОГРАММИРОВАНИЯ

Единственный класс задач математического программирования, для которого разработаны практически эффективные методы решения, составляют так называемые задачи *линейного программирования* (ЛП), то есть задачи отыскания в E^n экстремумов целевого *линейного* функционала $F(x)$ при условиях типа "нестрогое неравенство", накладываемых на значения некоторого конечного набора *линейных* функционалов $\{ f_j(x), j = [1, m]\}$, который задает допустимое множество задачи ЛП.

Рассмотрим конкретные формы постановки задач линейного программирования.

Пусть некоторый элемент $x \in E^n$ имеет в ортонормированном базисе $\{e_1, e_2, \ldots, e_n\}$ координатное представление

$$\| x \| = \left\| \begin{matrix} \xi_1 \\ \xi_2 \\ \ldots \\ \xi_n \end{matrix} \right\| .$$

Известно, что каждый линейный функционал в E^n (в том числе и целевой функционал $F(x)$) полностью и однозначно задается в базисе $\{e_1, e_2, \ldots, e_n\}$ n-компонентной строкой, получающейся при транспонировании столбца

$$\|c\| = \begin{Vmatrix} \sigma_1 \\ \sigma_2 \\ \cdots \\ \sigma_n \end{Vmatrix},$$

где $\sigma_j = F(e_j), \forall j = [1, n]$, а его значение на элементе x находится по одной из следующих равносильных формул:

$$F(x) = (c, x) \qquad \text{– символическая форма;}$$

$$F(x) = \|c\|^{\mathrm{T}} \|x\| \qquad \text{– матричная форма;}$$

$$F(x) = \sum_{j=1}^{n} \sigma_j \xi_j \qquad \text{– координатная форма.}$$

В задаче ЛП допустимое множество R задается системой неравенств вида

$$\begin{cases} f_1(x) \le \beta_1, \\ f_2(x) \le \beta_2, \\ \quad \dots \\ f_m(x) \le \beta_m, \end{cases}$$

где $f_i(x) = (a_i, x)$ – i-ый ограничивающий линейный функционал, для которого, кроме символического, также допустимы матричное и координатное представления

$$f_i(x) = \| a_i \| \| x \|,$$

$$f_i(x) = \sum_{j=1}^{n} \alpha_{ij} \xi_j ,$$

где n-компонентная строка $\| a_i \|$ есть координатное представление линейного функционала $f_i(x)$ в E^n, то есть

$$\| a_i \| = \| \alpha_{i1} \quad \alpha_{i2} \quad \dots \quad \alpha_{in} \|, \text{ где } i = [1, m].$$

Заметим, что совокупность строк $\left\| \alpha_{j1} \quad \alpha_{j2} \quad \dots \quad \alpha_{jn} \right\|$ при необходимости может быть также записана и в виде матрицы размера $m \times n$

$$\| A \| = \begin{Vmatrix} \alpha_{11} & \alpha_{12} & \dots & \alpha_{1n} \\ \alpha_{21} & \alpha_{22} & \dots & \alpha_{2n} \\ \dots & \dots & \dots & \dots \\ \alpha_{m1} & \alpha_{m2} & \dots & \alpha_{mn} \end{Vmatrix}.$$

Введем обозначение: $\| b \| = \begin{Vmatrix} \beta_1 \\ \beta_2 \\ \dots \\ \beta_m \end{Vmatrix}$. Тогда множество R – допустимых состояний – может быть задано системой условий в одной из следующих форм

$$\begin{cases} (a_1, x) \le \beta_1, \\ (a_2, x) \le \beta_2, \\ \quad \dots \\ (a_m, x) \le \beta_m, \end{cases} \qquad \begin{cases} \| a_1 \| \| x \| \le \beta_1, \\ \| a_2 \| \| x \| \le \beta_2, \\ \quad \dots \\ \| a_m \| \| x \| \le \beta_m, \end{cases}$$

$$\begin{cases} \sum_{j=1}^{n} \alpha_{1j} x_j \le \beta_1, \\ \sum_{j=1}^{n} \alpha_{2j} x_j \le \beta_2, \\ \quad \dots \\ \sum_{j=1}^{n} \alpha_{mj} x_j \le \beta_m \end{cases} \quad \text{или же} \quad \| A \| \| x \| \le \| b \|..$$

Формы постановки задачи линейного программирования при необходимости могут быть изменены при помощи правил равносильности неравенств и очевидных свойств операций нахождения максимума или минимума. Наиболее часто используются отношения следующего вида

$$\begin{gathered}
\alpha \geq \beta \quad \Leftrightarrow \quad -\alpha \leq -\beta \\
\alpha = \beta \quad \Leftrightarrow \quad \begin{cases} \alpha \geq \beta, \\ \alpha \leq \beta, \end{cases} \\
\alpha \geq \beta \quad \Leftrightarrow \quad \begin{cases} \alpha + \gamma = \beta, \\ \gamma \leq 0, \end{cases} \\
\max_x (c, x) = \min_x (-c, x).
\end{gathered} \tag{4.1.1}$$

Рассмотрим теперь стандартные постановки задач линейного программирования.

Канонической формой задачи ЛП, к которой может быть сведена любая задача линейного программирования, принято называть задачу:

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

$$\textit{при условиях: } \xi_j \geq 0, \quad j = [1, n],$$

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

Достаточно часто используются *матричные формы* представления задач ЛП, например для первой канонической формы условие задачи ЛП можно записать в виде

Найти максимум $\left\| c \right\|^{\mathrm{T}} \left\| x \right\|$ *на* $x \in E^n$,

при условиях: $\left\| x \right\| \geq \left\| o \right\|, \quad \left\| A \right\| \left\| x \right\| \leq \left\| b \right\|$

где каждое матричное неравенство определяется как система соответствующих покомпонентных числовых неравенств.

Будем обозначать решение задачи ЛП (если оно существует) как x^* с координатным представлением $\left\| x^* \right\| = \left\| \xi_1^* \quad \xi_2^* \quad \dots \quad \xi_n^* \right\|^{\mathrm{T}}$.

Двойственные условия оптимальности

Для задач линейного программирования как частного случая задач, рассмотренных в главе 3, необходимые и достаточные условия оптимальности могут быть получены также и в альтернативной, двойственной форме (Куна–Таккера), путем применения теоремы 1.1.4.4 (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 в этом случае формулируется как

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

а *двойственная* ей задача имеет вид

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

В качестве прямой задачи рассмотрим следующую задачу ЛП:

$$\begin{array}{l} \textit{найти максимум } \sum_{j=1}^{n} \sigma_j \xi_j \quad \textit{на} \quad \{\xi_1, \xi_2, \ldots, \xi_n\} \in E^n , \\ \textit{при условиях}: \\ \xi_j \ge 0 , \quad j = [1, n], \\ \sum_{j=1}^{n} \alpha_{ij} \xi_j \le \beta_i , \quad i = [1, m] \end{array} \tag{4.2.2.2}$$

Или же в символическом виде

$$\begin{array}{ll} \textit{Найти максимум} & (c, x) \quad \textit{на} \quad x \in E^n , \\ \textit{при условиях:} & x \ge o , \quad Ax \le b . \end{array}$$

В матричной форме эта задача будет

$$\begin{array}{ll} \textit{Найти максимум} & \| c \|^{\mathrm{T}} \| x \| \quad \textit{на} \quad x \in E^n , \\ \textit{при условиях:} & \| x \| \ge \| o \| , \quad \| A \| \| x \| \le \| b \| \end{array}$$

Составим функцию Лагранжа для прямой задачи

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

Используя распределительное свойство скалярного произведения, эту формулу можно записать в виде

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

Введя в рассмотрение линейный оператор A^*, сопряженный к A, получим

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

и, после группировки слагаемых, приходим к формуле

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

Тогда решение задачи $\max\limits_{x} L(x, \Lambda)$ в силу условия $x \geq 0$ имеет иной вид:

$$\max_{x} L(x,\Lambda) = \begin{bmatrix} (\Lambda, b), \; \textit{если } c - A^{*}\Lambda \leq o, \\ +\infty, \quad \textit{если } c - A^{*}\Lambda > o. \end{bmatrix}$$

При этом двойственная задача $\min\limits_{\Lambda \geq o} \max\limits_{x} L(x, \Lambda)$ будет иметь следующую формулировку:

$$\textit{Найти минимум} \qquad (\Lambda, b) \quad \textit{по} \quad \Lambda \in E^{m},$$
$$\textit{при условиях:} \qquad A^{\mathrm{T}}\Lambda \geq c\,, \quad \Lambda \geq o\,.$$

или же, в координатной форме:

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

$$\text{\textit{при условиях}:} \qquad \lambda_i \geq 0\,; \quad i = [1, m], \tag{4.2.2.3}$$

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

и, наконец, в матричном виде

$$\text{\textit{Найти минимум}} \qquad \| b \|^{\mathrm{T}} \| \Lambda \| \ \text{\textit{на}} \quad \Lambda \in E^m,$$

$$\text{\textit{при условиях:}} \qquad \| \Lambda \| \geq \| o \|, \quad \| A \|^{T} \| \Lambda \| \geq \| c \|.$$

Связь между условиями и решениями двойственной пары задач

Рассмотрим следующую пару задач ЛП:

задачу (***P***):

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

при условиях:

$$\xi_j \geq 0\ ,\ j = [1, n], \qquad \sum_{j=1}^{n} \alpha_{ij} \xi_j \leq \beta_i\ ,\ j = [1, m]$$

и ***задачу*** (***D***)**:**

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

при условиях:

$$\lambda_i \geq 0\ ,\ i = [1, m], \qquad \sum_{i=1}^{m} \alpha_{ij} \lambda_i \geq \sigma_i\ ,\ i = [1, n].$$

Основные правила, связывающие условия прямой и двойственной задач таковы:

Если в условии прямой задачи,	**то в условии двойственной задачи**
Найти максимум	Найти минимум
Найти минимум	Найти максимум
Число переменных	Число ограничений
Число ограничений	Число переменных
j -й коэффициент целевого функционала	правая часть j -го неравенства
j -я неотрицательная переменная	j -е неравенство типа $\geq$
j -я неограниченная переменная	j -е равенство
j -й столбец в матрице ограничений	j -я строка в матрице ограничений
правая часть i -го неравенства	i -й коэффициент целевого функционала
i -я строка в матрице ограничений	i -й столбец в матрице ограничений
i -е неравенство типа $\leq$	i -я неотрицательная переменная
i -е равенство	i -я неограниченная переменная

Теорема двойственности в линейном программировании

Для каждой пары взаимодвойственных задач верна

Теорема 4.3.2.1

Если $\left\| x^* \right\| = \left\| \xi_1^* \quad \xi_2^* \quad \dots \quad \xi_n^* \right\|^{\mathrm{T}}$ – оптимальное решение прямой задачи, а $\left\| \Lambda^* \right\| = \left\| \lambda_1^* \quad \lambda_2^* \quad \dots \quad \lambda_n^* \right\|^{\mathrm{T}}$ – оптимальное решение двойственной задачи, то справедливы равенства:

1°. ***основное соотношение двойственности***

$$\sum_{j=1}^{n} \sigma_j \xi_j^* = \sum_{i=1}^{m} \beta_i \lambda_i^* ;$$

2°. ***соотношения дополняющей нежесткости***

$$\lambda_i^* (-\beta_i + \sum_{j=1}^{n} \alpha_{ij} \xi_j^*) = 0 \; ; \forall i = [1, m],$$

$$\xi_j^* (-\sigma_j + \sum_{i=1}^{m} \alpha_{ij} \lambda_i^*) = 0 \; ; \forall j = [1, n].$$