

# Лекции по Методам Оптимизации.

Авторы

2022 — 2023

# Содержание

<b>1</b>	<b>Лекция №1. Безусловная оптимизация</b>	<b>4</b>
1.1	Постановка задачи и определения . . . . .	4
1.2	Общий вид оптимизационной задачи . . . . .	4
1.3	Необходимые и достаточные условия экстремума . . . . .	4
1.4	Необходимые условия экстремума I порядка . . . . .	5
1.5	Необходимые и достаточные условия экстремума II порядка . . . . .	5
1.6	Достаточные условия экстремума . . . . .	5
1.7	Теорема Вейерштрасса . . . . .	6
1.8	Критерий Сильвестра . . . . .	7
1.9	Правило решения задачи безусловной оптимизации . . . . .	8
<b>2</b>	<b>Лекция №2. Условная оптимизация</b>	<b>9</b>
2.1	Условная оптимизация с ограничениями-равенствами. Функция Лагранжа . . . . .	9
2.2	Условная оптимизация с ограничениями-равенствами и неравенствами. Математическое программирование . . . . .	11
<b>3</b>	<b>Лекция №3. Выпуклый анализ</b>	<b>13</b>
3.1	Выпуклые множества точек . . . . .	13
3.2	Геометрический смысл решений СЛН и СЛУ . . . . .	14
3.3	Выпуклые множества в n-мерном пространстве. Свойства ЗЛП . . . . .	15
<b>4</b>	<b>Лекция №4. Выпуклые и вогнутые функции</b>	<b>16</b>
4.1	Свойства выпуклых (вогнутых) функций . . . . .	16
4.2	Теорема (Крейна-Мильмана) . . . . .	17
4.3	Теорема . . . . .	17
4.4	Теорема . . . . .	19
4.5	Теорема (дифференциальный признак выпуклых функций) . . . . .	19
<b>5</b>	<b>Лекция №5. Общая постановка выпуклого программирования(ЗВП). Теорема Куна-Таккера</b>	<b>22</b>
5.1	Общая постановка ЗВП . . . . .	22
5.2	Теорема Куна-Таккера. . . . .	23
5.3	Теорема Куна-Таккера (необходимые и достаточные условия решения (глобального минимума) ЗВП) . . . . .	23
<b>6</b>	<b>Лекция №6. Теорема Куна-Таккера(продолжение)</b>	<b>26</b>
6.1	Доказательство теоремы Куна-Таккера . . . . .	26
6.2	Теорема Куна-Таккера (в "седловом" варианте) . . . . .	26
6.3	Линейное программирование . . . . .	28
<b>7</b>	<b>Лекция №7. Название лекции</b>	<b>29</b>
7.1	. . . . .	29
<b>8</b>	<b>Лекция №8. Определение первоначального допустимого базисного решения. Особые случаи симплексного метода. Двойственные задачи линейного программирования.</b>	<b>30</b>
8.1	Определение первоначального допустимого базисного решения . . . . .	30
8.2	Взаимно двойственные задачи линейного программирования(ЗЛП), их свойства. Основное неравенство теории двойственности. Первая(основная) теорема двойственности. Вторая теорема двойственности. . . . .	32
8.2.1	Взаимно двойственные задачи ЛП и их свойства . . . . .	32
8.2.2	Экономико-математическая модель исходной и двойственной задач: . . . . .	33

8.2.3	Свойства взаимно двойственных задач: . . . . .	33
8.2.4	Алгоритм составления двойственной задачи . . . . .	34
8.2.5	Теорема. Основное неравенство теории двойственности . . . . .	35
8.2.6	Теорема. Достаточный признак оптимальности: . . . . .	36
8.2.7	Первая(основная) теорема двойственности. . . . .	36
8.2.8	Экономический смысл первой(основной теоремы) . . . . .	37
8.2.9	Вторая теорема двойственности . . . . .	38
8.2.10	Теорема . . . . .	39
<b>9</b>	<b>Лекция №9. Название лекции</b>	<b>40</b>
9.1	. . . . .	40
<b>10</b>	<b>Словарик выражений</b>	<b>41</b>

# 1 Лекция №1. Безусловная оптимизация

## 1.1 Постановка задачи и определения

**def:** Методы оптимизации — это раздел математики, посвящённый решению (экстремальных) задач, то есть задач на нахождение минимумов и максимумов функций.

**Rem:** Задачу на нахождение максимума функции " $f(x)$ " можно свести к задаче на нахождение минимума функции " $f_1(x) = -f(x)$ " и наоборот.

## 1.2 Общий вид оптимизационной задачи

Найти экстремум (минимум или максимум) функции  $f : X \rightarrow R$  определенной на некотором множестве  $X \in R^n$  при ограничении  $X \in D (D \subset X)$  то есть  $f(x) \rightarrow \text{extr}, X \in D$  (у функции есть экстремум на промежутке  $D$ ).

В большинстве задач область определения функции " $f(x)$ "  $X = R^n$ . Ограничение  $X \in D$  записывается, как правило, в виде уравнений или неравенств. Если множество  $D = X$ , то имеет место задача без ограничений или задача безусловной оптимизации.

При решении оптимизационной задачи находятся как локальные, так и глобальные экстремумы функции.

**def:** Точка " $x^*$ " является точкой локального минимума (максимума) функции " $f(x)$ " if  $\exists \epsilon$  - окрестность  $\mathcal{U}_\epsilon = \{x : |x - x^*| < \epsilon\}$ . Точка  $x^* : f(x^*) \leq f(x) (f(x^*) \geq f(x)) \forall x \in \mathcal{U}_\epsilon$

**Rem:** то что пишется в скобках - для максимума, а то что без - для минимума

**def:** Точка " $x^*$ " является точкой глобального минимума (максимума) функции " $f(x)$ " if  $f(x^*) \leq f(x) (f(x^*) \geq f(x)) \forall x \in D$

## 1.3 Необходимые и достаточные условия экстремума

**def:** Точка  $X^0 = (x_1^0, \dots, x_n^0)$  называется стационарной точкой дифференцируемой функции  $f(x) = f(x_1 \dots x_n)$ , если в ней все частные производные равны нулю, то есть  $f'(x^0) = 0$  или  $\frac{\partial f(x^0)}{\partial x_1} = \dots = \frac{\partial f(x^0)}{\partial x_n} = 0$

## 1.4 Необходимые условия экстремума I порядка

**Теорема:** if точка  $x^* = (x_1^*, \dots, x_n^*)$  - точка локального extr дифференцируемой в точке  $x^*$  функции  $f(x_1, \dots, x_n) \Rightarrow$  then  $\frac{\partial f(x^*)}{\partial x_1} = \dots = \frac{\partial f(x^*)}{\partial x_n} = 0$  (1) (то есть - точка  $x^*$  - точка локального экстремума  $\Rightarrow$  точка  $x^*$  - стационарная точка (обратное утверждение неверно))

**Доказательство:** Рассмотрим функцию одной переменной:

$\varphi(x_i) = f(x_1^*, \dots, x_{i-1}^*, x_i^*, x_{i+1}^*, x_n^*)$  точка  $x^* = (x_1^*, \dots, x_n^*)$  - т. локального extr функции " $f$ "  $\Rightarrow x_i^*$  - т. локального extr функции " $\varphi$ "  $\Rightarrow$  по необходимому условию для функции одной переменной (по т. Ферма) 've:

$$\varphi(x_i^*) = 0 \Leftrightarrow \frac{\partial f(x^*)}{\partial x_i} = 0 \text{ — что и требовалось доказать :)}$$

Для формулировки достаточных условий extr, позволяющих отобрать среди стационарных точек именно точки локального extr (среди стационарных точек могут быть также точки перегиба, седловые точки и т.д.), рассмотрим матрицу вторых производных функции - матрицу Гессе (гессиан):

$$A = f''(x^*) = \left( \frac{\partial^2 f(x^*)}{\partial x_i \partial x_j} \right)_{i,j=\overline{1,n}} = (a_{i,j})_{i,j=\overline{1,n}} \text{ (от 1 до n)}$$

**def:** Матрица "A" называется неотрицательно определённой ( $A \geq 0$ ), если  $\forall h = (h_1, \dots, h_n) \in R^n$  неотрицательной является квадратичная форма:

$$(A * h, h) = \sum_{i,j=1}^n a_{ij} h_i h_j \geq 0$$

**def:** Матрица "A" называется положительно определённой ( $A > 0$ ), если  $(A * h, h) > 0, \forall h \in R^n (h \neq 0)$

## 1.5 Необходимые и достаточные условия экстремума II порядка

**Теорема:** Пусть  $f(x)$  - дважды дифференцируема в точке  $x^*$ . Необходимые условия условия extr:

if точка  $x^*$  - точка локального минимума (максимума) функции  $f(x) \Rightarrow f'(x^*) = 0; (f''(x^*) * h, h) \geq 0 ((f''(x^*) * h, h) \leq 0) \forall h \in R^n$

## 1.6 Достаточные условия экстремума

$f'(x^*) = 0; (f''(x^*) * h, h) > 0 ((f''(x^*) * h, h) < 0) \forall h \in R^n (h \neq 0) \Rightarrow$  точка  $x^*$  - т. локального минимума (максимума) функции  $f(x)$

**Доказательство:**

Для случая минимума (для максимума аналогично)

По формуле Тейлора 've:

$$f(x^* + h) = f(x^*) + (f'(x^*), h) + \frac{1}{2}(f''(x^*) * h, h) + r(h),$$

где  $r(h) = o(|h|^2).$ (\*)

### 1) Необходимость:

Пусть точка  $x^*$  - точка локального минимума  $\Rightarrow$  по необходимому условию I порядка  $f'(x^*) = 0$ , а также  $f(x^* + \lambda h) \geq f(x^*)$  (при достаточно малых " $\lambda$ ")  $\Rightarrow$  из (\*) get (g при малых " $\lambda$ " и фиксированном " $h$ "):

$$f(x^* + \lambda h) - f(x^*) = 0 + \frac{\lambda^2}{2}(f''(x^*) * h, h) + r(\lambda * h) \geq 0 \mid : \lambda^2$$

(где  $r(\lambda * h) = o(|\lambda|^2)$ ).  $\frac{1}{2}(f''(x^*) * h, h) + \frac{r(\lambda * h)}{\lambda^2} \geq 0$

При  $\lambda \rightarrow 0$  've :  $(f''(x^*) * h) \geq 0 (\forall h \in R^n) \Rightarrow$  необходимость доказана

### 2) Достаточность:

Можно показать, что в  $R^n$  имеет место эквивалентность условий:

$$(A * h, h) > 0 \forall h \in R^n (h \neq 0) \Leftrightarrow \exists \alpha > 0 : (A * h, h) \geq \alpha * |R|^2 (\forall h \in R^n)$$

Учитывая, что  $f'(x^*) = 0$  и  $(f''(x^*) * h, h) \geq \alpha * |h|^2$

По формуле Тейлора 've:

$f(x^* + h) - f(x^*) = 0 + \frac{1}{2}(f''(x^*) * h, h) + r(h) \geq \frac{\alpha}{2}|h|^2 + r(h) \geq 0$ ), то есть  $f(x^* + h) \geq f(x^*) \Rightarrow$  точка  $x^*$  - точка локального extr функции  $f(x) \Rightarrow$  достаточность доказана.

что и требовалось доказать

**Rem:** Для квадратичной функции

$Q(x) = \sum_{i,j=1}^n a_{ij} x_i x_j$  условие положительной(отрицательной) определённости матрицы  $A = (a_{ij})_{i,j=1}^n > 0$  - это достаточное условие абсолютного минимума(максимума)  $Q(x)$  в стационарной точке.

## 1.7 Теорема Вейерштрасса

### Теорема:

Непрерывная функция  $f : R^n \rightarrow R$  на непустом ограниченном замкнутом подмножестве(компакте) множества  $R^n$  достигает своих абсолютных минимума и максимума[или 1) в стационарной точке внутри; 2) в граничной

точке] - без доказательства

**Следствие:**

if функция " $f(x)$ " непрерывна на  $R^n$  и  $\lim_{|x| \rightarrow \infty} f(x) = \infty$

( $\lim_{|x| \rightarrow \infty} f(x) = -\infty$ )  $\Rightarrow$  then она достигает своего абсолютного минимума (максимума) на  $\forall$  замкнутом подмножестве и  $R^n$ . (без доказательства).

### 1.8 Критерий Сильвестра

**Rem:** В необходимых и достаточных условиях экстремума II порядка используется знакоопределённость матрицы вторых производных (гессиана)  $A = f''(x)$ . Знакоопределённость матрицы устанавливается с помощью критерия Сильвестра.

**Теорема:**

Пусть  $A$  - симметричная матрица

1) Матрица " $A$ " положительно определена ( $A > 0$ )  $\Leftrightarrow$  все её последовательные гл. миноры положительны, т.е.  $A_{1...k} = \begin{vmatrix} a_{11} & a_{1k} \\ a_{k1} & a_{kk} \end{vmatrix} > 0$  ( $k = \overline{1, n}$ )

2) Матрица " $A$ " отрицательно определена ( $A < 0$ )  $\Leftrightarrow$  все её последовательные главные миноры чередуют знак, начиная с отрицательного, т.е.  $(-1)^k * A_k > 0$  ( $k = \overline{1, n}$ )

3) Матрица " $A$ " неотрицательно определена ( $A \geq 0$ )  $\Leftrightarrow$  все её гл. миноры (необязательно только последовательные) неотрицательны, т.е.  $A_{1...k} = \begin{vmatrix} a_{i_1 i_1} & a_{i_1 i_k} \\ a_{i_k i_1} & a_{i_k i_k} \end{vmatrix} \geq 0$  ( $1 \leq i_1 \leq \dots \leq i_k \leq n$ ) ( $k = \overline{1, n}$ )

4) Матрица " $A$ " неположительно определена ( $A \leq 0$ )  $\Leftrightarrow$  все её последовательные главные миноры чередуют знак, начиная с неположительного, т.е.  $(-1)^k * A_{i_1...i_k} \geq 0$  ( $k = \overline{1, n}$ )

(Теорема без доказательства)

**Rem:**

1) Можно показать, что  $A > 0$  ( $A \geq 0$ )  $\Leftrightarrow \forall \lambda_i > 0$  ( $\lambda \geq 0$ ), где  $\lambda_i$  - собственные значения матрицы.

2)

2.1)  $A_{1...k} > 0 \Leftrightarrow A_{i_1...i_k} > 0$

2.2)  $A_{1...k} \geq 0 \not\Leftrightarrow A_{i_1...i_k} \geq 0$  (т.е.  $\not\Leftrightarrow A \geq 0$ )

ex:

$A = \begin{pmatrix} 0 & 0 \\ 0 & -1 \end{pmatrix} \Rightarrow$  последовательные главные миноры  $A_1 = 0$ ;  $A_{12} = \begin{vmatrix} 0 & 0 \\ 0 & -1 \end{vmatrix} = 0$ , но  $A$  не является неотрицательно определённой, так как  $(Ah, h) = ((0; -h), (h, h)) = -h^2 < 0 (\forall h \neq 0)$

### 1.9 Правило решения задачи безусловной оптимизации

1) Найти стационарные точки, то есть точки, удовлетворяющие необходимому условию экстремума I порядка

$$\begin{cases} \frac{\partial f}{\partial x_1} = 0 \\ \frac{\partial f}{\partial x_n} = 0 \end{cases}$$

2) Во всех стационарных точках " $x^0$ " проверяем достаточное условие extr II порядка, то есть проверяем знаки последовательных главных миноров гессиана " $f''(x^0)$ " :

2.1) if  $A_{1..k} > 0$  (k от 1 до n)  $\Rightarrow then x^0 \in locminf$ ;

2.2) if  $(-1)^k * A_{1...k} > 0$  (при k от 1 до n)  $\Rightarrow then x^0 \in locmaxf$ ;

3) Если достаточное условие extr II порядка не выполняется  $\Rightarrow$ , то проверяем в стационарной точке необходимое условие extr II порядка, то есть проверяем знаки главных миноров гессиана " $f''(x^0)$ "

3.1) if гессиан  $f''(x^0) \not\geq 0$  не является неотрицательным отрезком, то есть не выполняется условие  $A_{i1...ik} \geq 0 \Rightarrow then x^0 \notin locminf$ ;

3.2) if гессиан  $f''(x^0) \not\leq 0$  не является неположительным отрезком, то есть не выполняется условие что все  $(-1)^k A_{i1...ik} \geq 0 \Rightarrow then x^0 \notin locmaxf$ ;



## 2 Лекция №2. Условная оптимизация

### 2.1 Условная оптимизация с ограничениями-равенствами. Функция Лагранжа

**def:** Задачей условной оптимизации с ограничениями-равенствами называется следующая задача:

$f_1 \rightarrow \text{extr}$ ;  $f_i(x) = 0 (i = \overline{1, m}), (m < n)$ , где  $f_i(x) : R^n \rightarrow R (i = \overline{0, m})$  и  $\forall f_i(x)$  - дифференцируема.

**Теорема 2.1.1.** *необходимое условие экстремума I порядка.*

If  $x^* = (x_1^*, \dots, x_n^*) \in \text{locextr} f_0 \Rightarrow$  then  $\exists$  вектор множителей Лагранжа  $\lambda^* = (\lambda_1^*, \dots, \lambda_m^*) \in R^m (\lambda^* \neq 0)$  такой, что для функции Лагранжа:

$$\mathcal{L}(x_1, \dots, x_n; \lambda_1, \dots, \lambda_m) = f_0(x_1, \dots, x_n) + \sum_{i=1}^m \lambda_i f_i(x_1, \dots, x_n) \quad (1)$$

выполняются условия стационарности:

$$\begin{aligned} \frac{\partial \mathcal{L}(x^*, \lambda^*)}{\partial x_j} &= \frac{\partial f_1(x^*)}{\partial x_j} + \sum_{i=1}^m \lambda_i \frac{\partial f_i(x^*)}{\partial x_j} = 0 \quad (j = \overline{1, n}) \\ \frac{\partial \mathcal{L}(x^*, \lambda^*)}{\partial \lambda_i} &= f_i(x^*) = 0 \quad (i = \overline{1, m}) \end{aligned} \quad (2)$$

т.е. 'ве систему  $n + m$  уравнений для нахождения  $n + m$  неизвестных  $\{x_1^*, \dots, x_n^*; \lambda_1^*, \dots, \lambda_m^*\}$

**Теорема 2.1.2.** *необходимое условия экстремума II порядка.*

If  $x^* = (x_1^*, \dots, x_n^*) \in \text{locmin} f_0$  (условие регулярности) и векторы  $f_1'(x^*), \dots, f_m'(x^*)$  - линейно независимы  $\Rightarrow$  then  $\exists$  вектор множителей Лагранжа  $\lambda^* = (\lambda_1^*, \dots, \lambda_m^*) \in R^m (\lambda^* \neq 0)$  такой, что для функции Лагранжа

$$\mathcal{L}(x, \lambda) = f_0(x) + \sum_{i=1}^m \lambda_i f_i(x)$$

выполняются условия:

1. стационарности (2)

2. неотрицательной определенности матрицы вторых производных:

$$(3) (\mathcal{L}''(x^*, \lambda^*)h, h) \geq 0 \quad \forall h \in \{(f'_i(x^*), h) = 0 (i = \overline{1, m})\}$$

**Rem:** Для т.  $x^* \in \text{locmax} f_0$  need  $\mathcal{L}(x, \lambda) = -f_0(x) + \sum_{i=1}^m \lambda_i f_i(x)$

**Теорема 2.1.3.** достаточное условие экстремума II порядка.

If в т.  $x^* = (x_1^*, \dots, x_n^*)$  векторы  $f'_1(x^*), \dots, f'_m(x^*)$  - линейно независимы и  $\exists$  вектор множителей Лагранжа  $\lambda^* = (\lambda_1^*, \dots, \lambda_m^*) \in R^m (\lambda^* \neq 0)$  такой, что для функции Лагранжа

$$\mathcal{L}(x, \lambda) = f_0(x) + \sum_{i=1}^m \lambda_i f_i(x)$$

в т.  $x^*$  выполняются условия:

1. стационарности (2)

2. положительной определенности матрицы вторых производных:

$$(\mathcal{L}(x^*, \lambda^*)h, h) > 0 \quad \forall h \in \{(f'_i(x^*), h) = 0 (i = \overline{1, m}), (h \neq 0)\} \Rightarrow \\ \Rightarrow \text{then т. } x^* \in \text{locmin} f_0$$

**Rem:** Для т.  $x^* \in \text{locmax} f_0$  need  $\mathcal{L}(x, \lambda) = -f_0(x) + \sum_{i=1}^m \lambda_i f_i$

Частный случай

$z = f(x, y) \rightarrow \text{extr}$ ; при  $\varphi(x, y) = 0 \Rightarrow$  функция Лагранжа 've вид:

$$\mathcal{L}(x, y, \lambda) = f(x, y) + \lambda \varphi(x, y)$$

Условия стационарности(необх. усл. I порядка)(2):

$$\begin{cases} \mathcal{L}'_x(x^*, y^*, \lambda^*) = f'_x(x^*, y^*) + \lambda \varphi'_x(x^*, y^*) = 0 \\ \mathcal{L}'_y(x^*, y^*, \lambda^*) = f'_y(x^*, y^*) + \lambda \varphi'_y(x^*, y^*) = 0 \\ \mathcal{L}'_\lambda(x^*, y^*, \lambda^*) = \varphi'(x^*, y^*) = 0 \end{cases}$$

$\Rightarrow$  система трех уравнений с тремя неизвестными

$\Rightarrow$  находим стационарные точки  $(x^*, y^*, \lambda^*)$

Вычисляется в каждой из get-х стационарных точек  $(x^*, y^*, \lambda^*)$  определитель:

$$\Delta = - \begin{vmatrix} 0 & \varphi'_x & \varphi'_y \\ \varphi'_x & \mathcal{L}''_{xx} & \mathcal{L}''_{xy} \\ \varphi'_y & \mathcal{L}''_{yx} & \mathcal{L}''_{yy} \end{vmatrix}$$

if  $\Delta > 0 \Rightarrow$  then  $\text{т.}(x^*, y^*, \lambda^*) \in \text{locmin } z$

if  $\Delta < 0 \Rightarrow$  then  $\text{т.}(x^*, y^*, \lambda^*) \in \text{locmax } z$

## 2.2 Условная оптимизация с ограничениями-равенствами и неравенствами. Математическое программирование

**def:** Задачей условной оптимизации с ограничениями-равенствами и ограничениями-неравенствами называется следующая задача:

$$(4) f_0(x) \rightarrow \min; (5) f_i(x) \leq 0 \ (i = \overline{1, p}), (6) f_i(x) = 0 \ (i = \overline{p+1, m}),$$

где  $f_i(x) : R^n \rightarrow R (i = \overline{0, m})$

**def:** Эта задача называется задачей математического программирования (ЗМП)

**Rem:**

1) Обычно в ЗМП присутствуют условия неотрицательности переменных  $x_i \geq 0 \ (i = \overline{1, n})$  эти условия записываются в системе неравенств (5) в виде:  $-x_i \leq 0 \ (i = \overline{1, n})$

2) Каждое ограничение-равенство (6) можно заменить двумя неравенствами:

$$f_i(x) = 0 \Leftrightarrow \begin{cases} f_i(x) \leq 0 \\ -f_i(x) \leq 0 \end{cases} \quad (i = \overline{p+1, m})$$

В силу этих замечаний ЗМП можно записать в виде:

$$f_0(x) \rightarrow \min$$

$$f_i(x) \leq 0 \ (i = \overline{1, m}) \quad (8) \text{ где } x = (x_1, \dots, x_n)$$

Для ЗМП составляется функция Лагранжа

$$\mathcal{L}(x_1, \dots, x_n; \lambda_1, \dots, \lambda_m) = f_0(x_1, \dots, x_n) + \sum_{i=1}^m \lambda_i f_i(x_1, \dots, x_n)$$

С помощью функции Лагранжа выписываются необходимые и достаточные условия экстремума тип (2)-(3). Однако, проверка выполнения этих условий становится еще более сложной. При этом требуется решать систему, вообще говоря, нелинейных уравнений и неравенств. Для этого применяются итерационные численные методы, формирующие последовательность точек, сходящуюся к точке экстремума. Однако, эта точка может оказаться точкой локального (а не глобального) экстремума. Это объясняется тем, что ЗМП в такой общей постановке, без каких-либо предположений относительно функций  $f_i(x)$ , является многоэкстремальной задачей. Не существует универсальных методов решения таких задач. Содержательная теория построена лишь для отдельных классов ЗМП, в частности, задач оптимизации выпуклых функций на выпуклом множестве решений систем ограничений-неравенств. Такие задачи, называемые задачами выпуклого программирования (ЗВП), являются, как будет показано ниже, одноэкстремальными задачами.

### 3 Лекция №3. Выпуклый анализ

Ввиду важности задачи выпуклого программирования (ЗВП) рассматриваемые основные понятия раздела математики, называемого выпуклым анализом:

1) Выпуклое множество 2) Выпуклая(вогнутая) функция и ее дифференциальные свойства

#### 3.1 Выпуклые множества точек

**def:** Множество точек называется выпуклым, если оно вместе с  $\forall$  двумя своими точками содержит весь отрезок, соединяющий эти точки **ex:** — Тут нужно иллюстрацию **ex:** круг, сектор, отрезок, прямая, полуплоскость, куб, пирамида

**Теорема 3.1.1.** *Пересечение  $\forall$  числа выпуклых множеств-выпуклое множество, too*

*Доказательство.* Для простоты - пересечение двух выпуклых множеств:  $\forall N, M \in (A \cap B)$ . Множество  $A$  - выпуклое  $\Rightarrow$  отрезок  $MN \in A$ . Аналогично

$MN \in B \Rightarrow MN \in (A \cap B) \Rightarrow (A \cap B)$  - выпуклое — тут нужно иллюстрацию ■

**def:** Внутренняя т. множества -  $\exists$  окрестность этой точки: в ней - только точек  $\in$  множеству.

**def:** Граничная точка множества -  $\forall$  окрестность этой точки содержит как точки  $\in$  множеству, так и точек  $\notin$  множеству

**def:** Угловая точка множества - она не является внутренней ни для какого отрезка, целиком  $\in$ -го множеству.

**ex:** — Тут нужно иллюстрацию точка  $M$  - внутренняя, точка  $N$  граничная, точка  $A$  - угловая ( $AP \in$  множеству целиком, но точка  $A$  - не внутренняя точка для  $AP$ ; точка внутренняя точка для  $KL$ , но  $KL \notin$  множеству

целиком

**def:** Замкнутое множество точек - if оно включает все свои граничные точки.

**def:** Ограниченное множество точек - if  $\exists$  шар конечного радиуса с центром в  $\forall$  точке множества, который полностью содержит в себе данное множество.

Можно показать, что если фигура ограничена only прямыми или их отрезками, то она 've конечное число угловых точек. При криволинейности границ - бесконечное множество угловых точек.

**def:** Выпуклое замкнутое множество точек пространства(плоскости), имеющее конечное число угловых точек, называется выпуклым многогранником(многоугольником), если оно ограниченное, и выпуклой многогранной(многоугольной) областью, если оно неограниченное.

**Rem:** Для выпуклого многогранника(многоугольника) угловые точки  $\equiv$  его вершинам.

Для невыпуклого - не обязательно

**ex:** — Тут нужна иллюстрация точка E - вершина, но не угловая точка, т.к. KL  $\in$  множеству целиком и точка E - внутренняя точка для KL.

**Rem:** В ЗЛП часто число параметров объекта  $n > 3 \Rightarrow$  имеем дело с гипермногогранниками в гиперпространстве с координатами  $x_i (i = \overline{1, n}, n > 3)$

### 3.2 Геометрический смысл решений СЛН и СЛУ

**Теорема 3.2.1.** *Множество всех решений линейного неравенства*

$$a_{11}x_1 + \dots + a_{1n}x_n \leq b_1$$

- это одно из полупространств, на которые  $n$ -мерные гиперпространство делится гиплоскостью

$$a_{11}x_1 + \dots + a_{1n}x_n = b_1$$

включая и эту гиперплоскость

### 3.3 Выпуклые множества в n-мерном пространстве. Свойства ЗЛП

**Выпуклые множества в n-мерном пространстве.** Рассмотрим в n-мерном пространстве k точек (векторов):

$$x_1 = (x_1^{(1)}, \dots, x_n^{(1)}), \dots, x_k = (x_1^{(k)}, \dots, x_n^{(k)})$$

**def:** Точка(вектор)  $X = (x_1, \dots, x_n)$  называется линейной комбинацией точек(векторов)  $X_1, \dots, X_k$ , если справедливо соотношение:

$$X = \alpha_1 X_1 + \dots + \alpha_k X_k \quad (3)$$

где  $\alpha_j - \text{const } j = \overline{1, k}$

**def:** Точка(вектор)  $X$  называется выпуклой линейной комбинацией точек(векторов)  $X = \alpha_1 X_1 + \dots + \alpha_k X_k$ , если:

1.

$$X = \alpha_1 X_1 + \dots + \alpha_k X_k$$

2.

$$\alpha_j \geq 0 \quad j = \overline{1, k} \quad (4)$$

3.

$$\sum_{j=1}^k \alpha_j = 1 \quad (5)$$

Очевидно, что в частном случае при  $k=2$  выпуклой линейной комбинацией двух точек  $X_1$  и  $X_2$  является соединяющий их отрезок, т.к.

$$X = \alpha_1 X_1 + \alpha_2 X_2 = \left\{ \begin{array}{l} \alpha_1 + \alpha_2 = 1 \\ \alpha_2 = \alpha_1 - 1 \end{array} \right\} \alpha X_1 + (1 - \alpha_1) X_2$$

- уравнение точек  $X \in [X_1, X_2]$  (см. аналитич. геометр.) — добавить иллюстрацию

## 4 Лекция №4. Выпуклые и вогнутые функции

Пусть функция " $f(x)$ ", def-на на множестве  $M \subset R^n$  (т.е.  $x = (x_1, \dots, x_n) \in M$ )

**def:** График функции " $f(x)$ " - это множество  $\Gamma(f) \subset R^{n+1}$  состоящее из точек  $(x, f(x)) = (x_1, \dots, x_n, f(x))$ , где  $x \in M$ .

**def:** Надграфик функции " $f(x)$ " - это множество  $\Gamma(f) \subset R^{n+1}$ , состоящее из точек  $(x, x_{n+1}) = (x_1, \dots, x_n, x_{n+1})$ , где  $x \in M$ ,  $x_{n+1} \leq f(x)$

**def:** Функция " $f(x)$ ", заданная на множестве  $m \subset R^n$  наз-ся выпуклой, если ее надграфик  $\Gamma(f)$  является выпуклым множеством в  $R^{n+1}$

**Rem:** Функция " $f(x)$ " вогнута  $\leftrightarrow$  функция " $-f(x)$ " выпукла.

Кроме данного геометрического def-я выпуклых (вогнутых) функций, часто use-ся следующее аналитическое def-e: **def:** Функция " $f(x)$ ", заданная на множестве  $M \subset R^n$  наз-ся выпуклой(вогнутой), если 1)  $M$  - выпуклое, 2)  $\forall x_1 \in M, x_2 \in M$  и числа  $t \in [0, 1]$  've:

$$f((1-t) * x_1 + t x_2) \leq (1-t) * f(x_1) + t * f(x_2) \quad (1)$$

( $\geq$  - для вогнутой функции)

**Rem:** 1) Можно показать, что геометрические и аналитические def-я эквивалентны; (Rem:  $(1-t) * f(x_1) + t * f(x_2)$  - уравнение отрезка  $[x_1, x_2]$  ( $t \in [0, 1]$ ) при  $x_1 \neq x_2$  и  $t \in (0, 1)$ )

2) Если (1) 've строгое неравенство  $\Rightarrow$ , то функция " $f(x)$ " наз-ся строго выпуклой (строго вогнутой).

3)

### 4.1 Свойства выпуклых (вогнутых) функций

1.  $f(x) = const$  и  $f(x) = ax + b$  всюду выпуклы (вогнуты).

2. if функции  $f_i(x)$  ( $i = \overline{1, m}$ ), заданные на  $M \subset R^n$ , выпуклы  $\Rightarrow$



then функция  $f(x) = \sum_{i=1}^m \alpha_i f_i(x)$  выпукла (при  $\forall \alpha_i \geq 0$ )

**3.** if функция " $f(x)$ ", заданная на выпуклом множестве  $M \subset R^n$ , выпукла  
 $\rightarrow$

then  $\forall \alpha$  множество решений неравенства  $f(x) \leq \alpha$ , т.е. множество  $M_\alpha = \{x \in M : f(x) \leq \alpha\}$ , является выпуклым.

**3.1** if функции  $f_1(x), \dots, f_m(x)$ , заданные на выпуклом множестве  $M \subset R^n$ , выпуклы  $\Rightarrow$

then множество решений системы неравенств  $f_i(x) \leq \alpha_i$  ( $i = \overline{1, m}$ ) является выпуклым.

**4.** Выпуклая (вогнутая) функция, заданная на выпуклом множестве  $M \subset R^n$ , непрерывна в  $\forall$  внутренней точке множества.

(вставка на страницу 25)

**def:** Выпуклая оболочка множества - это совокупность всех выпуклых линейных комбинаций его конечных подмножеств  $\{x_1, \dots, x_k$  (где  $x_i \in M$ ).  
(Конечное подмножество - это конечный набор точек  $x_1, \dots, x_k$ )

## 4.2 Теорема (Крейна-Мильмана)

Выпуклый компакт в нормированном пространстве является выпуклой оболочкой своих угловых (крайних) точек.

## 4.3 Теорема

Пусть выпуклая (вогнутая) функция " $f(x)$ " задана на выпуклом множестве  $M \subset R^n \Rightarrow$

Каждый локальный минимум (максимум) функции " $f(x)$ " является её глобальным минимумом (максимумом) на множестве  $M$ .

### Доказательство (для выпуклой функции)

Пусть точка  $x^* \in M$  - точка локального min-а. Пусть точка  $x \in M$  - произвольная точка множества  $M$ . Need доказать:  $f(x) \geq f(x^*)$ .

Отрезок  $[x^*, x] = (1 - t_0)x^* + tx$  ( $t \in [0, 1]$ ) принадлежит " $M$ ". При малом значении  $t_0 \in (0, 1)$  /-щая точка отрезка  $[x^*, x]$  находится в малой окрестности т.  $x^*$ , в котором имеем:

$$f((1 - t_0)x^* + t_0x) \geq f(x^*)$$

Из def-я выпуклой функции  $f(x)$  've:

$$f((1 - t_0)x^* + t_0x) \leq (1 - t_0)f(x^*) + t_0f(x)$$

Т.о. :

$$(1 - t_0) * f(x^*) + t_0 * f(x) \geq f(x^*) \Rightarrow$$

$\Rightarrow f(x) \geq f(x^*)$  ч. и т.д. (для вогнутой " $f(x)$ " доказательство аналогичное.)

### Rem:

1) Задачи выпуклой оптимизации называются одноэкстремальными. В многоэкстремальных задачах может  $\exists$ -ть локальные экстремумы, не совпадающие с глобальными.

2) Одноэкстремальность задач выпуклой оптимизации не означает, что каждая такая задача имеет решение и при том единственное. f.e.: 1) Выпуклая функция одной переменной  $f(x) = x$ , при  $x \in (0, 1)$  не достигает min-а (и max-а) на  $(0, 1)$ ; 2) Множество точек min-а выпуклой функции  $f(x) = C - \text{const}$ ,  $x \in M$ , совпадает со всем  $M$ .

3) Если функция " $f(x)$ " - строго выпуклая (строго вогнутая), то разрешимая задача выпуклой оптимизации (т.е. множество  $M \neq \emptyset$  и ограничено) имеет единственное решение.

**def:** Производной  $\frac{\partial f(x)}{\partial \ell}$  функции  $f(x) = f(x_1, \dots, x_n)$  по направлению ненулевого единичного вектора  $\ell = (\ell_1, \dots, \ell_n)$  в т.  $x = (x_1, \dots, x_n)$  называется предел

$$\frac{\partial f(x)}{\partial \ell} = \lim_{\lambda \rightarrow +0} \frac{f(x + \lambda \ell) - f(x)}{\lambda}$$

Если функция дифференцируема в т.  $x$ , то она в этой точке производную по  $\forall$  направлению  $\ell = (\ell_1, \dots, \ell_n)$ , которая выражается через частные производные по следующей формуле:

$$\frac{\partial f(x)}{\partial \ell} = \sum_{i=1}^n \frac{\partial f(x)}{\partial x_i} \ell_i$$

Производная по направлению равна скалярному произведению вектора " $\ell$ " и вектора градиента функции " $f(x)$ " в т.  $x$

$$\nabla f(x) = \left( \frac{\partial f(x)}{\partial x_1}, \dots, \frac{\partial f(x)}{\partial x_n} \right) : \frac{\partial f(x)}{\partial \ell} = \nabla f(x) * \ell = \nabla f(x) * |\ell|$$

или

$$\frac{\partial f(x)}{\partial \ell} = |f(x)| * |\ell| * \cos \phi$$

Т.о.:

$\forall$  направления " $\ell$ " производная  $\frac{\partial f}{\partial \ell} \leq |\nabla f(x)|$ .

$\Rightarrow$  Вывод:

1) Производная по направлению " $\frac{\partial f}{\partial \ell}$ " - это скорость изменения функции " $f(x)$ " по направлению " $\ell$ " (знак " $\frac{\partial f}{\partial \ell}$ " - это характер изменения функции (возрастание или убывание)).

2) Направление градиента  $\nabla f(x)$  - это направление наибольшего возрастания функции " $f(x)$ " в т.  $x$ ; длина градиента  $|\nabla f(x)|$  равна наибольшей скорости возрастания функции в этой точке.

#### 4.4 Теорема

Пусть 've дифференцируемую выпуклую функцию " $f(x)$ ", def-ую на выпуклом множестве  $M \subset R^n$ .

1)  $\forall x, y \in M$  've:  $\lambda f(x) * (y - x) \leq f(y) - f(x)$

2) т.  $x^* \in \text{absminf} \Leftrightarrow \nabla f(x^*) = 0$  (2)

#### 4.5 Теорема (дифференциальный признак выпуклых функций)

Дважды дифференцируемая функция " $f(x)$ ", def-ная на выпуклом множестве  $M \subset R^n$  является выпуклой  $\Leftrightarrow \forall x = (x_1, \dots, x_n) \in M$  и  $\forall \ell = (\ell_1, \dots, \ell_n) \in R^n$

've:

$$\sum_{i=1}^n \sum_{j=1}^n \frac{\partial^2 f(x)}{\partial x_i \partial x_j} \ell_i \ell_j \geq 0, (3)$$

т.е. гессиан функции всюду неотрицательно определён:

$$A = f''(x) = \left( \frac{\partial^2 f(x)}{\partial x_i \partial x_j} \right) \geq 0, (\forall x \in M)$$

**Rem:**

1) Функция является строго выпуклой  $\Leftrightarrow$  гессиан положительно определён, т.е.

$$\left( \frac{\partial^2 f(x)}{\partial x_i \partial x_j} \right) > 0, (\forall x \in M)$$

2) Знакоопределённость гессиана устанавливается с помощью критерия Сильвестра.

**ex 1:**

Проверить выпуклость функции:

$$f(x_1, x_2) = 4x_1 + x_2^2 - 2x_1x_2 + 6x_1 - 5x_2 - 2$$

$$f'_{x_1} = 8x_1 = 2x_2 + 6; f''_{x_1x_1} = 8; f''_{x_1x_2} = -2; f''_{x_2x_2} = 2; f'_{x_2} = 2x_2 - 2x_1 - 5$$

$$\text{Гессиан } A = \begin{pmatrix} 8 & -2 \\ -2 & 2 \end{pmatrix}; \Delta_1 = 8 > 0; \Delta_2 = 12 > 0$$

$\Rightarrow$  функция является строго выпуклой.

**ex 2:**

Проверить выпуклость функции:  $f(x) = -\sqrt{x_1x_2}$  на множестве

$$M = \{(x_1, x_2) | x_1 > 0, x_2 > 0\}.$$

$$f'_{x_1} = -\frac{1 \cdot x_2}{2\sqrt{x_1x_2}} = -\frac{1}{2}\sqrt{\frac{x_2}{x_1}};$$

$$f'_{x_2} = -\frac{1}{2}\sqrt{\frac{x_1}{x_2}};$$

$$f''_{x_1x_1} = \left(-\frac{\sqrt{x_2}}{2} * x_1^{-\frac{1}{2}}\right)' = \frac{1}{4} \frac{\sqrt{x_2}}{\sqrt{x_1^3}} = \frac{1}{4x_1} \sqrt{\frac{x_2}{x_1}};$$

$$f''_{x_1x_2} = -\left(\frac{1}{2\sqrt{x_1}} * x_2^{\frac{1}{2}}\right)' = \frac{1}{4\sqrt{x_1x_2}};$$

$$f''_{x_2x_2} = \left(-\frac{1 \cdot \sqrt{x_1}}{2} * x_2^{-\frac{1}{2}}\right)' = \frac{1\sqrt{x_1}}{4\sqrt{x_2^3}} = \frac{1}{4x_2} * \sqrt{\frac{x_1}{x_2}}$$

$$\text{Гессиан } H = \begin{pmatrix} \frac{1}{4x_1} * \sqrt{\frac{x_2}{x_1}} & -\frac{1}{4\sqrt{x_1x_2}} \\ -\frac{1}{4\sqrt{x_1x_2}} & \frac{1}{4x_2} * \sqrt{\frac{x_1}{x_2}} \end{pmatrix}$$

$$\Delta_1 = \frac{1}{4x_1} \sqrt{\frac{x_2}{x_1}} > 0 (\forall x_1, x_2 > 0);$$

$$\Delta_2 = \frac{1}{16x_1x_2} \sqrt{\frac{x_2x_1}{x_1x_2}} - \frac{1}{16x_1x_2} = 0 \Rightarrow$$

Гессиан  $H \geq 0 \Rightarrow$  функция является на множестве  $M = \{(x_1, x_2) | x_1 > 0, x_2 > 0\}$  выпуклой (но не является строго выпуклой).

## 5 Лекция №5. Общая постановка выпуклого программирования(ЗВП). Теорема Куна-Таккера

### 5.1 Общая постановка ЗВП

ЗВП называется следующей задачей математического программирования(ЗМП)

$$f(x_1, \dots, x_n) \rightarrow \min$$

при ограничениях:

$$\begin{cases} g_1(x_1, \dots, x_n) \leq 0 \\ g_m(x_1, \dots, x_n) \leq 0 \end{cases},$$

где  $f, g_1, \dots, g_m$  - выпуклые функции, def-ные на некотором выпуклом множестве  $M \subset R^n$ .

При этом множество  $M$  содержит допустимую область значений, то есть множество, удовлетворяющее системе ограничений. Заметим, что свойству 3.1 выпуклых функций допустимая область ЗВП также выпукла

**Rem:**

1) Аналогично формулируется задача максимизации вогнутой функции "f" при вогнутых функциях " $g_1, \dots, g_m$ ", def-х на некотором выпуклом множестве  $M \subset R^n$ . При этом знак неравенств — " $\geq$ ".

Аналогично допустимая область ЗВП будет также выпуклым множеством(свойство для вогнутых функций)

2) В ??? с теор. на с.??? в ЗВП каждый локальный минимум является глобальным минимумом функции "f" в допустимой области. Причём, если функция "f" строго выпуклая и ограниченная снизу на ограниченном непустом множестве  $M$ , то ЗВП 've единственное решение, то есть минимумом функции "f" достигается в одной точке

$$x^0 = (x_1^0, \dots, x_n^0)$$

(см с.???)

## 5.2 Теорема Куна-Таккера.

**Теорема:** Пусть на ЗВП налагаются следующие требования:

1. Множество, удовлетворяющее системе строгих неравенств:

$$\begin{cases} g_1(x_1, \dots, x_n) < 0 \\ g_m(x_1, \dots, x_n) < 0 \end{cases}$$

и называемое внутренней частью допустимой области, не пусто (т.н. условие Слейтера)

2. Часть допустимой области, в которой некоторые ограничения обращаются в равенства, называется границей допустимой области, а эти ограничения - активными. Пусть градиенты активных ограничений в отвечающих им точках границы линейно независимы (т.н. условие регулярности (см с.???)

## 5.3 Теорема Куна-Таккера (необходимые и достаточные условия решения (глобального минимума) ЗВП)

Точка  $x^0 = (x_1^0, \dots, x_n^0)$  является решением ЗВП (то есть точкой глобального минимума функции "f")  $\Leftrightarrow$  (Н. и Д.) в ней выполнены следующие условия (Условия Куна-Таккера):

1) Условие стационарности функции Лагранжа (см.с.???)

$$\lambda(x_1, \dots, x_n; \lambda_1, \dots, \lambda_m) = f(x_1, \dots, x_n) + \sum_{i=1}^m \lambda_i * g_i(x_1, \dots, x_n)$$

$$\frac{\partial L}{\partial x_1} = \dots = \frac{\partial L}{\partial x_n} = 0 \text{ (то есть } \nabla(x^0) = 0);$$

2) Условия "дополняющей нежёсткости":

$$\lambda_i * g_i(x_1^0, \dots, x_m^0) = 0 \text{ (} i = \overline{1, m});$$

3) Условия неотрицательности:

$$\lambda_i \geq 0 \text{ (} i = \overline{1, m})$$

### Доказательство:

Заметим, что в условиях К.-Т. множители Лагранжа  $\lambda_1, \dots, \lambda_m$  являются "выключателями" делящих ограничений. Если, например,  $\lambda_k = 0$ , то ограничение  $g_k \leq 0$  исключается из условий К.-Т., так как : 1) это ограничение

не входит в функцию Лагранжа (слагаемое  $\lambda_k * g_k = 0$ ) и,??? не входит в условие 1; 2) условия 2 и 3 для этого ограничения выполняются автоматически

## 1. Необходимость

Пусть точка  $x^0 = (x_1^0, \dots, x_n^0)$  - решение ЗВП, то есть  $\min f(x_1, \dots, x_n) = f(x_1^0, \dots, x_n^0)$ . Целевая функция "f" может достигать минимума внутри допустимой области или на её границе.

Пусть минимум достигается внутри области  $\Rightarrow$  в точке  $x^0$  выполняются необходимые условия локального(безусловного экстремума функции "f"):

$$\frac{\partial f}{\partial x_1} = \dots = \frac{\partial f}{\partial x_n}$$

Это частный случай условия 1 К.-Т.(условия стационарности функции Лагранжа) при  $\lambda_1 = \dots = \lambda_m = 0$  все ограничения выключены, так как 've безусловную оптимизацию)

Условия 2 и 3 при этом (при  $\lambda_i = 0 (i = \overline{1, m})$ ) выполняются автоматически.

Пусть минимум достигается на границе допустимой области, то есть является условным extr функции "f" с активными ограничениями задачи для точки минимума точки  $x^0$ . Необходимые условия условного extr (см. с. ???) ( $\frac{\partial L}{\partial x_j} = 0; (\frac{\partial L}{\partial \lambda_j} = 0 (j = \overline{1, n}; i = \overline{1, m}))$ ) влекут за собой выполнение условий 1 и 2 К.-Т., если положительно равными нулю множит. Лагранжа для неактивных ограничений (так как реально они не входят в функцию Лагранжа).

Остаётся доказать, что выполняется 3 условие К.-Т. :  $\lambda_i \geq 0 (i = \overline{1, m})$

Для наглядности ограничимся случаем двух переменных (n=2).

Уравнение  $g(x_1, x_2)$  означает плоскую кривую, а неравенство  $g(x_1, x_2) \leq 0$  - одну из частей плоскости, ограниченную этой кривой

Система ограничений

$$\begin{cases} g_1(x_1, x_2) \leq 0 \\ \dots \\ g_m(x_1, x_2) \leq 0 \end{cases}$$

задает криволинейный многоугольник, в каждой вершине которого пересекаются две кривые.

Минимум на границе допустимой области достигается либо на строке этого многоугольника, либо в его вершине<sub>24</sub>



Пусть минимум - на стороне многоугольника  $\Rightarrow$  активным является только одно ограничение и только отвечающий ему множитель Лагранжа  $\lambda_k \neq 0$ .  
Функция Лагранжа 've вид:

$$L = f(x_1, \dots, x_n) + \lambda_k * g_k(x_1, \dots, x_n),$$

где  $g_k = 0$  - активное ограничение.

В точке минимума выполняются необходимые условия. Условия условного extr(см. с. ???):

$$\frac{\partial L}{\partial x_1} = \dots = \frac{\partial L}{\partial x_n} = \frac{\partial L}{\partial \lambda_k}$$

или

$$\begin{cases} \nabla L = \nabla f + \lambda_k * \nabla g_k = 0 \\ g_k = 0, \end{cases}$$

$$\Rightarrow \nabla f = -\lambda_k * \nabla g_k, (*)$$

то есть векторы  $\nabla f$  и  $\nabla g_k$  — коллинеарны

Функция  $g_k < 0$  внутри допустимой области и  $g_k > 0$  вне её  $\Rightarrow$  вектор  $\nabla g_k$  направлен из допустимой области (в сторону возрастания функции "  $g_k$  ")

Значение функции "  $f$  " внутри допустимой области больше, чем в точке минимума  $\Rightarrow \nabla f$  направлен внутрь допустимой области.

Таким образом, векторы  $\nabla f$  и  $\nabla g_k$  противоположны и в равенстве (\*) коэффициенты коллинеарности  $\lambda_k \leq 0 \Rightarrow \lambda_k \geq 0$ , то есть условие 3 К.-Т. доказано для случая, когда минимум достигается на стороне допустимого многоугольника.

Можно показать, что условие 3 К.-Т. ( $\lambda_i \geq 0 (i = \overline{1, m})$ ) выполняется и для случая, когда минимум достигается в вершине допустимого многоугольника.

Таким образом, необходимость К.-Т. доказана.

## 6 Лекция №6. Теорема Куна-Таккера(продолжение)

### 6.1 Доказательство теоремы Куна-Таккера

#### 2. Достаточность

Пусть в точке  $x^0 = (x_1^0, \dots, x_n^0)$  выполняются условия 1-3 теоремы Куна-Таккера.

Имеем выпуклые функции " $f, g_1, \dots, g_m$ " и выпуклую допустимую область  $\Rightarrow$

$\Rightarrow$  Функция Лагранжа

$L = f + \lambda_1 g_1 + \dots + \lambda_m g_m$  ( $\lambda_1 \geq 0, \dots, \lambda_m \geq 0$ ) также выпукла (см. св-во 2 выпуклых ф-ий)

Если в точке  $x^0 = (x_1^0, \dots, x_n^0)$  выполняется условие 1 теоремы Куна-Таккера  $\frac{\partial L}{\partial x_1} = \dots = \frac{\partial L}{\partial x_n} = 0$ , (т.е.  $\nabla L(x^0) = 0$ ) то эта точка - точка  $\min$  выпуклой функции Лагранжа " $L$ " в допустимой области.

Обозначим  $\min L(x) = L(x^0) = L^0 \Rightarrow L^0 \leq L$

Но вследствие условия 2 теоремы Куна-Таккера

$$\lambda_1 g_1^0 = \dots = \lambda_m g_m^0 = 0 \Rightarrow L^0 = f^0, f^0 \leq L$$

Кроме того, в допустимой области  $g_1 \leq 0, \dots, g_m \leq 0$  и по условию 3 теоремы Куна-Таккера,  $\lambda_1 \geq 0, \dots, \lambda_m \leq 0 \Rightarrow \lambda_1 g_1 = \dots = \lambda_m g_m \leq 0 \Rightarrow$

в допустимой области

$L = f + \lambda_1 g_1 + \dots + \lambda_m g_m \leq f \Rightarrow$  в допустимой области:

$f^0 \leq L \leq f$ , т.е.  $f^0 \leq f \Rightarrow \min f = f^0 = f(x^0)$ , т.е. точка  $x^0 = (x_1^0, \dots, x_n^0)$  - точка минимума (глобального) целевой функции в допустимой области. Т.

О. достаточность теоремы Куна-Таккера доказана.

ч и т.д.

### 6.2 Теорема Куна-Таккера (в "седловом" варианте)

Точка  $x^0 = (x_1^0, \dots, x_n^0)$  является решением ЗВП (т.е. точкой глобального  $\min$ -а функции " $f$ ")  $\Leftrightarrow \exists$  неотрицательный вектор множителей Лагранжа

$\lambda^0 = (\lambda_1^0, \dots, \lambda_m^0)$  ( $\lambda_i^0 \geq 0; i = \overline{1, m}$ ) такой, что для функции Лагранжа

$L = (x_1, \dots, x_n; \lambda_1, \dots, \lambda_m) = f(x_1, \dots, x_n) + \sum_{i=1}^m \lambda_i g_i(x_1, \dots, x_n)$  точка  $(x^0; \lambda^0)$  является седловой точкой, т.е.

$$L(x^0, \lambda) \leq (x^0, \lambda^0) \leq L(x, \lambda^0) (*)$$

$\forall x \in$  допустимой области,  $\lambda \geq 0$   
(без доказательства)

**Rem:**

1. Из неравенства (\*) следует, что точка - точка минимума функции, а точка - точка максимума функции по. Существование седловой точки означает равенство минимакса максимуму.
2. Можно показать, что необходимые и достаточные условия того, чтобы являлась седловой точкой функции Лагранжа, имеют вид:

**Rem:** 1. При доказательстве необходимости условий теоремы Куна-Таккера выпуклость функций и допустимой области не требовалась  $\Rightarrow$  условия Куна-Таккера являются необходимыми для произвольной (необязательно выпуклой) ЗМП.

2. Множители Лагранжа " $\lambda_k$ " представляют собой "цену" ограничения " $g_k \leq 0$ ", т.е. чувствительность оптимального значения целевой функции " $f^0$ " к нарушению этого ограничения. В функции Лагранжа слагаемые " $\lambda_k g_k$ " представляет собой как бы "штраф" за нарушение ограничения " $g_k \leq 0$ ". Чем больше значение множителя Лагранжа " $\lambda_k$ " и чем больше нарушено /-щее ограничение (т.е. стало " $g_k > 0$ "), тем больше величина этого "штрафа". На этой идее основан один из методов решения ЗМП - метод штрафных функций, который был рассмотрен ниже.

3. Условия теоремы Куна-Таккера представляют собой систему нелинейных (в общем случае) уравнений и неравенств, которая допускает лишь приближенное решение численными методами. Для решения ЗВП разработали специальные итерационные численные методы, которые заключаются в построении последовательных приближений к точке целевой функции в допустимой области. Критерием окончания итерационного процесса является достижение заданной точности вычислений. В результате определяются приближенные значения.

4. Точное решение ЗВП допускает в случае, когда и целевая функция и

все ограничения являются линейными функциями переменных (линейные функции являются выпуклыми в /-ии со свойством 1 выпуклых функций). Такая ЗВП называется задачей линейного программирования (ЗЛП).

### **6.3    Линейное программирование**

## 7 Лекция №7. Название лекции

### 7.1

## 8 Лекция №8. Определение первоначального допустимого базисного решения. Особые случаи симплексного метода. Двойственные задачи линейного программирования.

### 8.1 Определение первоначального допустимого базисного решения

Базисные решения, получаемые на I шаге, не всегда являются допустимыми.

Рассмотрим один из алгоритмов получения допустимого базисного решения (ДБР):

1. Если в каждом уравнении дополнительная переменная и свободный член, стоящий в правой части, имеют одинаковые знаки, то дополнительные переменные берём в качестве базисных и при этом получаем ДБР.
2. Если хотя бы в одном уравнении дополнительная переменная и свободный член имеют противоположные знаки (и дополнительные переменные — в качестве базисных, то есть 1 базисное решение получается недопустимым), то в системе (1) (в которой базисные выражаются через свободные) рассмотрим  $\forall$  ур-е с отрицательным свободным членом и переводим в базисные  $\forall$  из свободных переменных, входящих в это уравнение с положительным коэффициентом.

Процедура повторяется до достижения ДБР. При этом (при возможности выбора) следует переводить из свободных в базисные ту свободную переменную, которая def-т в качестве разрешающего уравнения с отрицательным свободным членом. Только в этом случае новые базисные решения будут иметь меньше отрицательных компонент.

3. Если базисное решение недопустимое и в уравнениях с отрицательным свободным членом  $\nexists$  свободных переменных с положительным коэффициентом, то  $\nexists$  ДБР (то есть условия ЗЛП противоречивы).

ex:

$F = x_1 + x_2 \rightarrow \max$  при

$$\begin{cases} x_1 - x_2 \leq -1 \\ x_1 - x_2 \geq -3 \\ x_1 \leq 3 \\ x_1, x_2 \geq 0 \end{cases}$$

В каноническом виде

$$\begin{cases} x_1 - x_2 + x_3 & = -1 \\ x_1 - x_2 & - x_4 = -3 \\ x_1 & + x_5 = 3 \\ x_i \geq 0 \ (i = \overline{1,5}) \end{cases}$$

### I шаг

В соответствии с правилом со стр. ??? в качестве базисных берём дополнительные переменные  $\{x_3, x_4, x_5\}$ ; свободные -  $\{x_1, x_2\}$ . Базисные — через свободные:

$$\begin{cases} x_3 = \boxed{-1} - x_1 + x_2 \\ x_4 = \boxed{3} + x_1 - x_2 \\ x_5 = \boxed{3} - x_1 \end{cases}$$

$\Rightarrow$  базисное решение  $X_1 = (0; 0; -1; 3; 3)$  - не является допустимым, то есть оно  $\notin$  многоугольнику решений  $\Rightarrow \nexists F(X_1)$

Уравнение с отрицательным свободным членом — 1-ое, в нём с положительным коэффициентом — " $x_2''$ "  $\Rightarrow$  её в базисные (так как при этом будет увеличиваться " $x_3''$ ")

Оценочные отношения  $\Rightarrow$  Наибольшее возможное значение:  $x_2 = \min\{1; 3; \infty\} \Rightarrow$  1-ое уравнение  $\boxed{x_3 = -1 - x_1 + x_2}$  — разрешающее

### II шаг

Базисные —  $\{x_2; x_4; x_5\}$ ; свободные —  $\{x_1, x_3\}$

$$\begin{cases} x_2 = \boxed{1} + x_1 + x_3 \\ x_4 = \boxed{3} + x_1 - (1 + x_1 + x_3) = 2 \\ x_5 = \boxed{3} - x_1 \end{cases}$$

$\Rightarrow$  get ДБР:  $X_2 = (0; 1; 0; 2; 3);$

$$F = x_1 + (1 + x_1 + x_3) = 1 + 2x_1 + x_3;$$

$$F(X_2) = 1 \neq F_{max}$$

$\Rightarrow$  следующий шаг симплексного метода (так как в F есть положительные коэффициенты, а значит он не max)...<sub>31</sub>

## 8.2 Взаимно двойственные задачи линейного программирования(ЗЛП), их свойства. Основное неравенство теории двойственности. Первая(основная) теорема двойственности. Вторая теорема двойственности.

### 8.2.1 Взаимно двойственные задачи ЛП и их свойства

∀ ЗЛП соответствует двойственная (сопряжённая) ей задача ЗЛП.

**ex:** (см стр. ??? об use-нии ресурсов):

Из ресурсов  $S_i$  ( $i = \overline{1, m}$ ) с запасами  $b_i$  ( $i = \overline{1, m}$ ) изготавливаются виды продукции  $P_j$  ( $j = \overline{1, n}$ ) в количестве  $x_j$  ( $j = \overline{1, n}$ ). Технологические коэффициенты —  $a_{ij}$  ( $i = \overline{1, m}, j = \overline{1, n}$ ) — число единиц ресурса  $S_i$  затративших на изготовление одной единицы продукции  $p_j$ . Прибыль(выручка) от реализации единицы продукции  $P_j$  —  $c_j$ , ( $j = \overline{1, n}$ ) (то есть цена единицы продукции  $P_j$ ).

Целевая функция:

$F = \sum_{j=1}^n c_j * x_j \rightarrow \max$  — прибыль от реализации всей продукции.

При ограничениях:

$$\sum_{j=1}^n a_{ij} * x_j \leq b_i, (i = \overline{1, m})$$

Пусть другая организация хочет купить у этого предприятия все ресурсы  $S_i$  ( $i = \overline{1, m}$ ).

Целевая функция(для другой организации, так как они хотят купить как можно дешевле):

$Z = \sum_{i=1}^m b_i * y_i \rightarrow \min$  — затраты на покупку всех ресурсов по ценам " $y_i$ ";

При ограничениях:

$Z = \sum_{i=1}^m a_{ij} * y_i \geq c_j, (j = \overline{1, n})$  — выручка продавца (предприятия) должна быть не менее той суммы, которую предприятие может получить при переработке ресурсов  $S_i$  в готовую продукцию  $P_j$ .



### 8.2.2 Экономико-математическая модель исходной и двойственной задач:

**Задача I (исходная):**

$$F = c_1x_1 + \dots + c_nx_n \rightarrow \max \quad (1)$$

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n \leq b_1 \\ \dots \\ a_{m1}x_1 + a_{m2}x_2 + \dots + a_{mn}x_n \leq b_m \end{cases} \quad (2)$$

$$x_j \geq 0 \quad (j = \overline{1, n}) \quad (3)$$

Найти: план выпуска продукции  $X = (x_1, \dots, x_n)$ , прибыль от её реализации  $F = F_{\max}$ , при условии, что потребление ресурсов  $S_i$  не превзойдёт запасов  $b_i$

**Задача II (двойственная):**

$$Z = b_1y_1 + \dots + b_ny_n \rightarrow \min \quad (4)$$

$$\begin{cases} a_{11}y_1 + a_{12}y_2 + \dots + a_{1n}y_n \geq c_1 \\ \dots \\ a_{m1}y_1 + a_{m2}y_2 + \dots + a_{mn}y_n \geq c_m \end{cases} \quad (5)$$

$$y_j \geq 0 \quad (j = \overline{1, m}) \quad (6)$$

Найти: набор цен(оценок ресурсов)  $Y = (y_1, \dots, y_n)$ , общие затраты на ресурсы  $Z = Z_{\min}$ , при условии, что затраты на ресурсы при производстве продукции  $P_j$  не менее прибыли от её реализации  $c_j$

### 8.2.3 Свойства взаимно двойственных задач:

1. В одной задаче  $\rightarrow \max$ , в другой —  $\rightarrow \min$
2. Коэффициенты целевой функции одной задачи — свободные члены системы ограничений другой задачи
3. Обе ЗЛП — в стандартной форме; в задаче  $\max$ -ции неравенства вида  $\leq$  — в  $\min$ -ции —  $\geq$
4. Матрицы системы ограничений транспонированные друг к другу
5. Число неравенств системы ограничений одной задачи — число переменных другой задачи
6. В обеих задачах неотрицательные переменные

#### 8.2.4 Алгоритм составления двойственной задачи

1. Все неравенства к одному виду: if  $\max$  —  $\leq$ , if  $\min$  —  $\geq$
  2. Составить расширенную матрицу исходной системы: дописываем  $n+1$  столбец свободных членов,  $m+1$  строка — строка коэффициентов целевой функции.
  3. Транспонируем её
  4. Формулируем двойственную задачу на основе этой матрицы.
- Будет  $n$  ограничений,  $m$  переменных.

Последняя  $n+1$  строка — бывшие коэффициенты свободных членов в исходной матрице. Они теперь являются коэффициентами целевой функции двойственной задачи ( $Z$ )

Столбец свободных членов  $n+1$  в исходной задаче — матрица коэффициентов целевой функции  $F$ . Теперь это вектор свободных членов в системе ограничений для двойственной задачи.

**ex:**

Исходная ЗЛП:

$$F = -x_1 + 2x_2 \rightarrow \max$$
$$\begin{cases} 2x_1 - x_2 \geq 1 \\ -x_1 + 4x_2 \leq 24 \\ x_1 - x_2 \leq 3 \\ x_1 + x_2 \geq 5 \\ x_1 \geq 0; x_2 \geq 0 \end{cases} \Rightarrow m = 4; n = 2$$

**Пункт №1.** Приводим неравенства данной задачи  $\max$ -ции к виду " $\leq$ "

$$\begin{cases} -2x_1 + x_2 \leq 1 \\ -x_1 + 4x_2 \leq 24 \\ x_1 - x_2 \leq 3 \\ -x_1 - x_2 \leq 5 \\ x_1 \geq 0; x_2 \geq 0 \end{cases}$$

**Пункт №2.** Составляем расширенную матрицу системы.

$$A = \left( \begin{array}{cc|c} -2 & 1 & -1 \\ -1 & 4 & 24 \\ 1 & -1 & 3 \\ -1 & -1 & -5 \\ \hline -1 & 2 & F \end{array} \right)$$

**Пункт №3.** Составляем расширенную матрицу двойственной задачи (Транспонировали предыдущую)

$$A' = \left( \begin{array}{cccc|c} -2 & -1 & 1 & -1 & -1 \\ 1 & 4 & -1 & -1 & 2 \\ \hline -1 & 24 & 3 & -5 & Z \end{array} \right)$$

**Пункт №4.** Двойственная ЗЛП:

$$z = -y_1 + 24y_2 + 3y_3 - 5y_4 \rightarrow \min;$$

при

$$\begin{cases} -2y_1 - y_2 + y_3 - y_4 \geq -1 \\ y_1 + 4y_2 - y_3 - y_4 \geq 2 \\ y_i, (i = \overline{1, 4}) \end{cases}$$

### 8.2.5 Теорема. Основное неравенство теории двойственности

**Основное неравенство:**

$\forall$  пары допустимых решений исходной  $X = (x_1, \dots, x_n)$  и двойственной  $Y = (y_1, \dots, y_m)$  имеем:

$$F(X) \leq Z(Y), \text{ то есть } \sum_{j=1}^n c_j x_j \leq \sum_{i=1}^m b_i y_i$$

**Доказательство:**

Умножим каждое  $i$ -ое неравенство системы ограничений задачи I (2) на неотрицательную переменную  $y_i (i = \overline{1, m})$  и сложим все неравенства, то есть:

$$\sum_{j=1}^n a_{ij} x_j \leq b_i \mid * y_i \Rightarrow \sum_{i=1}^m y_i * (\sum_{j=1}^n a_{ij} x_j) \leq \sum_{i=1}^m b_i y_i \boxed{=Z(Y)}$$

Аналогично каждое "j" неравенство задачи II (5) — на  $x_j (j = \overline{1, n})$  и сложим, то есть:

$$\sum_{i=1}^m a_{ij} y_i \geq c_j \mid * x_j \Rightarrow \sum_{j=1}^n x_j * (\sum_{i=1}^m a_{ij} y_i) \geq \sum_{j=1}^n c_j x_j \boxed{=F(X)}$$

Таким образом get-уем:

$$F(X) \leq \sum_{i=1}^m \sum_{j=1}^n a_{ij} x_j y_i \leq Z(Y)$$

то, что по центру это как раз и есть то общее, что было в предыдущих двух неравенствах ■

### 8.2.6 Теорема. Достаточный признак оптимальности:

if  $X^* = (x_1^*, \dots, x_n^*), Y^* = (y_1^*)$  — допустимые решения взаимно-двойственных задач и

$$F(X^*) = Z(Y^*) (7) \Rightarrow$$

$\Rightarrow$  then  $X^* = X_{max}, Y^* = Y_{min}$  (то есть это оптимальные решения своих задач)

### 8.2.7 Первая(основная) теорема двойственности.

1. Если существует оптимальное решение одной из взаимно-двойственных задач  $\Rightarrow$  then  $\exists$  решение другой задачи и они равны:

$$F_{max} = Z_{min} \text{ or } F(X^*) = Z(Y^*)$$

2. if целевая функция одной из задач неограничена  $\Rightarrow$  then условия другой задачи противоречивы.

#### Доказательство:

1) Первая часть — без доказательства. Но, таким образом, равенство(7) — не только достаточный, но и необходимый признак оптимальности решений взаимно двойственных задач, то есть

$$F(X^*) = Z(Y^*) \Leftrightarrow X^* = X_{max}, Y^* = Y_{min}.$$

2) Вторая часть доказывается от противного.

Пусть  $F_{max} = \infty$ , а условия другой задачи не является противоречивыми, то есть  $\exists$  допустимое решение  $Y = (y_1, \dots, y_m) \Rightarrow \forall X : F(X) \leq Z(Y) \Rightarrow \nexists$  допустимых решений  $Y$ . ■

ex:

I

$$F = 2x_1 + 3x_2 \rightarrow \max$$

$$\begin{cases} x_1 + 3x_2 \leq 18 \\ 2x_1 + x_2 \leq 16 \\ x_2 \leq 5 \\ 3x_1 \leq 21 \\ x_i \geq 0 (i = \overline{1, 2}) \end{cases}$$

II

$$Z = 18y_1 + 16y_2 + 5y_3 + 21y_4 \rightarrow \min$$

$$\begin{cases} y_1 + 2y_2 + 3y_4 \geq 2 \\ 3y_1 + y_2 + y_3 \geq 3 \\ y_i \geq 0 (i = \overline{1, 4}) \end{cases}$$

$F_{max} = 24$ , значит можно сделать вывод о том, что  $Z_{min} = 24$  и наоборот.

### 8.2.8 Экономический смысл первой(основной теоремы)

План производства  $X^* = (x_1^*, \dots, x_n^*)$  и набор цен(оценок) ресурсов  $Y^* = (y_1^*, \dots, y_m^*)$  оптимальны  $\Leftrightarrow$  прибыль(выручка) от продукции  $F(X)$  равная затратам на ресурсы  $Z(Y)$ .

Для других планов  $X$  и  $Y$  прибыль (выручка) от продукции всегда меньше затрат на ресурсы.

Таким образом, предприятию безразлично либо

1) производить продукцию по оптимальному плану  $X^*$ , либо

2) продавать ресурсы по оптимальному плану  $Y^*$  и возместить от продажи равные  $F_{max}$  затраты на ресурсы  $Z_{min}$ .

### 8.2.9 Вторая теорема двойственности

Две взаимно двойственные ЗЛП характеризуются не только равенством оптимальных значений их целевых функций, а также взаимосвязью между первоначальными переменными одной из задач и дополнительными переменными другой задачи в их оптимальных ДБР  $X^*$  и  $Y^*$

Рассмотрим взаимно двойственные задачи со стр. ??? . Для решения их симплексным методом они должны быть приведены к каноническому виду путем введения в систему ограничений дополнительных неотрицательных переменных:

Было  $\dots \leq b_i$ :

$$\sum_{j=1}^n a_{ij} * x_j + x_{n+i} = b_i, (i = \overline{1, m}) \quad (8)$$

Было  $\dots \geq c_j$ :

$$\sum_{i=1}^m a_{ij} * y_i - y_{m+j} = c_j, (j = \overline{1, n}) \quad (9)$$

Установим соответствие между переменными:

Первоначальные(исходной) — дополнительные(двойственной)

$$x_1 \leftrightarrow y_{m+1}$$

...

$$x_j \leftrightarrow y_{m+j}$$

...

$$x_n \leftrightarrow y_{m+n}$$

Теперь наоборот:

Дополнительные(исходной) — первоначальные(двойственной):

$$x_{n+1} \leftrightarrow y_1$$

...

$$x_{n+i} \leftrightarrow y_j$$

...

$$x_{n+m} \leftrightarrow y_m$$

### 8.2.10 Теорема

Положительным компонентам оптимального решения одной из задач соответствуют нулевые компоненты оптимального решения

## 9 Лекция №9. Название лекции

### 9.1



## 10 Словарик выражений

1. get-уем — получаем
2.  $\div$  — соответствует
3. 've — имеем
4. зча — задача
5. зчи — задачи
6. зч — задач
7. Мцы — матрицы
8. if — если
9. Т.О. — таким образом
10. opt-ные — оптимальные
11. then — тогда
12. смы — системы
13. М. — можно
14. max-я — максимизация
15. min-я — минимизация
16. реш-я — решения
17. def-уем — определяем
18. def-ся — определяется
19. def-ть — определить
20. the end — конец