Введение

1. Безусловный минимум функции конечного числа переменных

Пусть E^n – n-мерное вещественное евклидово пространство, элементами которого являются векторы $x=\left(x_1,...,x_j,...,x_n\right)$, $y=\left(y_1,...,y_j,...,y_n\right)$,... Пусть на E^n задана функция f. Как обычно, через $f\left(x\right)$ обозначим значение функции f в точке x.

<u>*Определение 1.*</u> *Вектор* $x^* \in E^n$ *называется* ***точкой безусловного глобального минимума*** *функции* f *на* E^n, *если для всех* $x \in E^n$ *выполняется неравенство*

$$f\left(x^*\right) \le f\left(x\right). \tag{1}$$

Обозначим $f^* = f\left(x^*\right) = \min\limits_{x \in E^n} f\left(x\right)$.

Для краткости будем использовать термин ***«минимум функции»***, имея в виду точку минимума функции f на E^n. В литературе по математическому анализу глобальный минимум также называют ***абсолютным минимумом***.

Задача поиска f^* и x^* называется ***задачей безусловной минимизации*** функции f.

Наряду с понятием минимума также существует понятие ***максимума***. Задача максимизации функции f легко сводится к задаче минимизации: $\max f(x) = -\min\left(-f(x)\right)$. Понятия «***минимум***» и «***максимум***» объединяются термином «***экстремум***».

*<u>Определение 2.</u> Вектор $x^* \in E^n$ называется **точкой локального безусловного минимума** функции f на E^n, если неравенство (1) выполняется для всех x из некоторой окрестности точки x^*.*

Локальный минимум также называют ***относительным минимумом***.

Пусть функция f дифференцируема в точке $x \in E^n$. ***Градиент функции*** f в точке x (то есть вектор $\left(\frac{\partial f(x)}{\partial x_1}, \ldots, \frac{\partial f(x)}{\partial x_n}\right)$, состоящий из первых частных производных функции f) будем обозначать через $f'(x)$.

<u>Теорема 1.</u> ***(Необходимое условие безусловного минимума первого порядка)*** *Пусть функция f определена на E^n и дифференцируема в точке x^*. Тогда для того, чтобы точка x^* была ло-*

кальным безусловным минимумом функции f *необходимо, чтобы выполнялось равенство*

$$f'\left(x^*\right) = 0. \qquad (2)$$

Условие (2) можно записать покоординатно как систему равенств $\dfrac{\partial f\left(x^*\right)}{\partial x_j} = 0,\ j = 1, ..., n$.

<u>Определение 3.</u> Вектор $x \in E^n$ *называется* ***стационарной точкой*** *функции* f*, дифференцируемой в* x*, если* $f'\left(x\right) = 0$.

Таким образом, стационарная точка – это решение системы уравнений

$$\frac{\partial f\left(x\right)}{\partial x_j} = 0,\ j = 1, ..., n.$$

Точки локального минимума содержатся среди стационарных точек функции. Однако не всякая стационарная точка является точкой минимума. В частности, точки локального максимума и так называемые седловые точки также являются стационарными. Среди стационарных точек могут быть и точки, не являющиеся точками экстремума.

Пусть функция f дважды дифференцируема в точке $x \in E^n$. Матрица

$$f''(x) = \left\{\begin{matrix} \dfrac{\partial^2 f(x)}{\partial x_1^{\,2}}, & \dfrac{\partial^2 f(x)}{\partial x_1 \partial x_2}, & \dots, & \dfrac{\partial^2 f(x)}{\partial x_1 \partial x_n} \\ \dfrac{\partial^2 f(x)}{\partial x_2 \partial x_1}, & \dfrac{\partial^2 f(x)}{\partial x_2^{\,2}}, & \dots, & \dfrac{\partial^2 f(x)}{\partial x_2 \partial x_n} \\ \vdots & \vdots & \vdots & \vdots \\ \dfrac{\partial^2 f(x)}{\partial x_n \partial x_1}, & \dfrac{\partial^2 f(x)}{\partial x_n \partial x_2}, & \dots, & \dfrac{\partial^2 f(x)}{\partial x_n^{\,2}} \end{matrix}\right\},$$

состоящая из вторых частных производных функции f, называется ***гессианом*** (***матрицей Гессе***) функции f в точке x.

Пусть A – квадратная симметричная матрица. Функция $\langle Ax, x\rangle$ называется ***квадратичной формой***.

<u>*Определение 4.*</u> *Говорят, что матрица* A ***неотрицательно определена****, если* $\langle Ax, x\rangle \ge 0$ *при всех* $x \in E^n$, и ***положительно определена****, если* $\langle Ax, x\rangle > 0$, $\forall x \in E^n$, $x \ne 0$.

Аналогично определяются понятия ***неположительно*** и ***отрицательно определенных*** матриц.

<u>*Теорема 2.*</u> (***Необходимое условие безусловного минимума второго порядка***) *Пусть функция* f *определена на* E^n *и дважды дифференцируема в точке* x^*. *Тогда для того, чтобы точка* x^* *была локальным безусловным минимумом функции* f *на*

E^n, необходимо, чтобы матрица $f''(x^*)$ была неотрицательно определена.

Эта теорема позволяет отсеять те из стационарных точек функции f, которые не могут быть точками минимума.

Теорема 3. (***Достаточное условие безусловного минимума второго порядка***) *Пусть функция f определена на E^n и дважды дифференцируема в стационарной точке x^*. Тогда для того, чтобы x^* была локальным безусловным минимумом, достаточно, чтобы матрица $f''(x^*)$ была положительно определена.*

При помощи этого условия можно из стационарных точек, прошедших через «сито» предыдущей теоремы, отобрать точки локального минимума. Оставшиеся точки, а также особые точки функции, которые вместе со стационарными точками объединяются термином ***«критические»*** точки, требуют дополнительного исследования. При этом может потребоваться использование производных более высокого порядка.

2. Условный экстремум

2.1. Основные определения

В предыдущем параграфе рассматривалась задача на безусловный экстремум функции, то есть задача без ограничений на область изменения переменных. Однако во многих проблемах требуется отыскивать экстремум функции с условием, что аргумент может принимать значения только из некоторого множества $D \subset E^n$.

Пусть D – множество из E^n, функция f определена на E^n. Задача минимизации функции f на множестве D называется ***задачей на условный минимум***. При этом множество D принято называть ***допустимой областью***, точки $x \in D$ – ***допустимыми***, функцию f – ***целевой функцией*** (***критерием***) задачи.

Введем некоторые определения.

<u>Определение 1.</u> Точка $x^* \in D$ называется ***точкой условного глобального минимума*** *функции f на множестве D, если для всех $x \in D$ выполняется неравенство $f\left(x^*\right) \le f\left(x\right)$.*

Для краткости будем использовать термин ***«условный минимум функции»***, имея в виду точку минимума функции f на D. Обозначим

$$D^* = \left\{ x^* : \ x^* \in D, \ f\left(x^*\right) = \min_{x \in D} f\left(x\right) \right\}, \quad f^* = f\left(x^*\right).$$

*<u>Определение 2.</u> Точка $x^* \in D$ называется **точкой локального условного минимума** функции f на множестве D, если неравенство $f\left(x^*\right) \le f\left(x\right)$ выполняется для тех $x \in D$, которые принадлежат также некоторой окрестности точки x^*.*

Задача поиска f^* и x^* называется ***задачей условной минимизации*** функции f.

Для анализа и решения этой задачи существенно то, как задано множество D. В частности, далее рассмотрим два варианта:

1) допустимая область задана при помощи системы уравнений;

2) допустимая область задана при помощи системы неравенств.

2.2. Правило множителей Лагранжа

Рассмотрим так называемую классическую задачу на условный минимум или задачу с ограничениями в виде уравнений.

Пусть на E^n заданы функции f, $f_1, ..., f_m$.

Положим $D = \left\{ x : x \in E^n, \ f_i\left(x\right) = 0, \ i = 1, ..., m \right\}$.

Таким образом, множество D представляет собой некоторую поверхность в E^n. Для условной оптимизации наиболее содержательны случаи, когда D – собственное подмножество E^n.

Теперь определим на E^n вектор-функцию $F(x) = \left(f_1(x), ..., f_m(x)\right)$. Тогда, для краткости, запишем $D = \left\{x : x \in E^n,\ F(x) = 0\right\}$.

Введем следующую функцию:

$$L(x, y) = f(x) + \left\langle y, F(x)\right\rangle = f(x) + \sum_{i=1}^{m} y_i f_i(x),$$

где $y = \left(y_1, ..., y_m\right)$. Эта функция $n+m$ переменных называется ***функцией Лагранжа***. Переменные y_i называются ***множителями Лагранжа***.

<u>*Теорема 1.*</u> ***(Правило множителей Лагранжа)*** *Пусть функции $f, f_i,\ i = 1, ..., m$ непрерывно дифференцируемы на E^n, точка $x^* \in D$ такова, что система векторов $\left\{f_i'\left(x^*\right)\right\}_{i=1}^{m}$ линейно независима. Если x^* – локальный минимум функции f на множестве D, то существует вектор y^* такой, что*

$$L_x'\left(x^*, y^*\right) = 0. \qquad (1)$$

Согласно определению функции Лагранжа условие (1) можно записать как систему равенств

$$f'\left(x^*\right) + \sum_{i=1}^{m} y_i^* \, f_i'\left(x^*\right) = 0.$$

Как использовать эту теорему для отыскания условных экстремумов? Составим систему

$$\begin{cases} L_x'\left(x, y\right) = 0, \\ L_y'\left(x, y\right) = 0. \end{cases} \qquad (2)$$

Запишем систему (2) подробнее:

$$\begin{cases} f'\left(x\right) + \sum_{i=1}^{m} y_i \, f_i'\left(x\right) = 0, \\ f_i\left(x\right) = 0, \quad i = 1, \dots, m. \end{cases}$$

Она состоит из $n+m$ уравнений относительно $n+m$ переменных. Пусть вектор $\left(x, y\right)$ – некоторое ее решение. В этом случае вектор x называется ***условно стационарной точкой*** функции f. Таким образом, решая систему (2), мы можем найти все условно стационарные точки. Среди них содержатся все точки условного локального минимума.

Теорема 2. (***Необходимое условие условного минимума второго порядка***) *Пусть выполнены все условия теоремы 1 и, кроме того, функция* f *дважды дифференцируема в точке* x^*. *Тогда для того, чтобы* x^* *была локальным условным минимумом, необходимо, чтобы матрица* $f''\left(x^*\right)$ *была неотрицательно определена на подпространстве*

$$H = \left\{h:\ h \in E^n,\ \left\langle f_i'\left(x^*\right),\ h\right\rangle = 0,\ i = 1, \dots, m\right\}.$$

Так как это условие является только необходимым, оно позволяет отсеять те из условно стационарных точек, которые не могут быть точками условного минимума.

<u>*Теорема 3.*</u> (***Достаточное условие условного минимума второго порядка***) *Пусть выполнены все предположения теоремы 2. Тогда для того, чтобы условно стационарная точка* x^* *была локальным условным минимумом, достаточно, чтобы матрица* $f''\left(x^*\right)$ *была положительно определена на подпространстве* H.

При помощи теорем 2 и 3 можно из условно стационарных точек отобрать точки условного локального минимума. Те же условно стационарные точки, которые не удовлетворяют этим теоремам, требуют дальнейшего исследования.

2.3. Задача нелинейного программирования

Рассмотрим теперь задачу на условный минимум с ограничениями в виде неравенств. Положим $D = \left\{x:\ x \in E^n,\ F(x) \le 0\right\}$ Часто простейшие ограничения системы неравенств выписывают отдельно. Например, такими ограничениями могут быть ограничения на знак переменных. В этом случае $D = \{x : x \in E^n,\ F(x) \le 0,\ x \ge 0\}$.

Целевая функция f и все функции f_i, вообще говоря, нелинейны, поэтому такую задачу отыскания условного минимума называют ***задачей нелинейного программирования***.

Задача нелинейного программирования значительно сложнее классической задачи на условный экстремум. Причиной тому – особенность «конструкции» множества D. Граница его, даже если все функции f_i дифференцируемы, представляет собой, вообще говоря, «негладкое» многообразие. Поэтому приведенные выше теоремы мало чем могут помочь в исследовании и решении задачи нелинейного программирования. Отсутствие в классическом анализе необходимого аппарата исследования потребовало разработки современной теории экстремальных задач. С основами этой теории мы познакомимся в рамках данного пособия.

В классической задаче в любой допустимой точке x равенства $f_i(x) = 0$ выполнены для всех $i = 1, \dots, m$, а в задаче нелинейного программирования часть ограничений может выполняться в виде равенств, а часть – в виде строгих неравенств.

Введем следующие определения.

Определение 3. *Ограничение с номером* i ($i \in \{1, \dots, m\}$) *называется* ***активным*** *(или* ***жестким)*** *в точке* $x \in D$, *если* $f_i(x) = 0$.

<u>*Определение 4.*</u> *Ограничение с номером i ($i \in \{1, ..., m\}$) называется* ***пассивным*** *(или* ***нежестким****) в точке $x \in D$, если $f_i(x) < 0$.*

Множество номеров всех активных в точке x ограничений обозначим $I(x)$. Таким образом,

$$I(x) = \{i: \ i = 1, ..., m; \ f_i(x) = 0\}.$$

<u>*Определение 5.*</u> *Пусть в задаче нелинейного программирования $D = \{x: \ x \in E^n, \ F(x) \le 0, x \ge 0\}$; векторы $\bar{x} \in D$, $\bar{y} \in E^m$, $\bar{y} \ge 0$. Будем говорить, что вектор $(\bar{x}, \bar{y})$ удовлетворяет* ***условию дополняющей нежесткости****, если выполнено равенство*

$$\langle \bar{y}, F(\bar{x}) \rangle = 0. \qquad (3)$$

Условие (3) можно записать как $\sum_{i=1}^{m} \bar{y}_i \ f_i(\bar{x}) = 0$. Отсюда, в силу знакопостоянства слагаемых, $\bar{y}_i \ f_i(\bar{x}) = 0, \ \forall i = 1, ..., m$. Поэтому, если i-тое ограничение задачи нелинейного программирования пассивно в точке $\bar{x}$, то $\bar{y}_i = 0$. Таким образом, ограничение на знак множителя Лагранжа y_i является в точке $\bar{y}$ активным. Если же $\bar{y}_i > 0$ (ограничение на знак y_i в точке $\bar{y}$ пассивно), то $f_i(\bar{x}) = 0$, что означает активность i-того ограничения задачи в точке $\bar{x}$.

<u>*Теорема 4.*</u> *Пусть в задаче нелинейного программирования* $D = \left\{x:\ x \in E^n,\ F(x) \le 0,\ x \ge 0\right\}$. *Пусть неотрицательные векторы* $\bar{x} \in E^n$, $\bar{y} \in E^m$ *таковы, что* $L(\bar{x}, y) \le L(\bar{x}, \bar{y})$ *для всех* $y \ge 0$. *Тогда* $(\bar{x}, \bar{y})$ *удовлетворяет условию дополняющей нежесткости.*

<u>*Доказательство.*</u> По условию теоремы

$$f(\bar{x}) + \langle y, F(\bar{x})\rangle \le f(\bar{x}) + \langle \bar{y}, F(\bar{x})\rangle,\ \forall y \ge 0,$$

то есть

$$\langle y, F(\bar{x})\rangle \le \langle \bar{y}, F(\bar{x})\rangle,\ \forall y \ge 0. \qquad (4)$$

Покажем, что отсюда следует неравенство

$$F(\bar{x}) \le 0. \qquad (5)$$

Действительно, если найдется значение i $(1 \le i \le m)$ такое, что $f_i(\bar{x}) > 0$, то, полагая $y_j = 0$ для всех $j = 1, \ldots, m;$ $j \ne i$ из неравенства (4) имеем $y_i\, f_i(\bar{x}) \le \langle \bar{y}, F(\bar{x})\rangle$ при любых неотрицательных значениях компоненты y_i, что невозможно при достаточно больших y_i.

Итак, (5) имеет место, поэтому с учетом неотрицательности вектора $\bar{x}$ получаем, что $\bar{x} \in D$.

Так как векторы $\bar{y} \ge 0$ и $F(\bar{x}) \le 0$,

$$\langle \bar{y}, F(\bar{x})\rangle \le 0. \qquad (6)$$

С другой стороны, из (4) при $y = 0$ получаем

$$\left\langle \bar{y}, F(\bar{x}) \right\rangle \ge 0. \qquad (7)$$

Из (6) и (7) следует (3). Что и требовалось.

<u>*Теорема 5.*</u> *Пусть в задаче нелинейного программирования* $D = \left\{ x : x \in E^n,\ F(x) \le 0,\ x \ge 0 \right\}$. *Пусть неотрицательные векторы* $x^* \in E^n$, $y^* \in E^m$ *удовлетворяют условию дополняющей нежесткости и неравенству*

$$L\left(x^*, y^*\right) \le L\left(x, y^*\right),\ \forall x \ge 0. \qquad (8)$$

Тогда вектор x^* *– решение задачи нелинейного программирования.*

<u>*Доказательство.*</u> Из (8) и (3) получаем

$$f\left(x^*\right) \le f(x) + \left\langle y^*, F(x) \right\rangle, \quad \forall x \ge 0. \qquad (9)$$

Пусть $x \in D$. Тогда $F(x) \le 0$ и, так как $y^* \ge 0$, имеем $\left\langle y^*, F(x) \right\rangle \le 0$. Отсюда и из (9) получаем $f\left(x^*\right) \le f(x),\ \forall x \in D$. Что и требовалось.

Заметим, что в условиях теоремы 5 точка x^* является условным минимумом функции $L\left(x, y^*\right)$ по переменным x на неотрицательном ортанте $E_+^n = \left\{ x : x \in E^n,\ x_j \ge 0,\ j = 1, ..., n \right\}$ n-мерного евклидова пространства.

<u>*Определение 6.*</u> *Неотрицательный вектор* $\left(x^*, y^*\right)$ *называется* ***седловой точкой*** *функции*

Лагранжа, если для всех $x \geq 0$ *и* $y \geq 0$ *выполняются неравенства*

$$L\left(x^{*}, y\right) \leq L\left(x^{*}, y^{*}\right) \leq L\left(x, y^{*}\right). \qquad (10)$$

<u>*Теорема 6.*</u> ***(О связи седловой точки функции Лагранжа с решением задачи нелинейного программирования)*** *Если* $\left(x^{*}, y^{*}\right)$ *– седловая точка функции Лагранжа, то вектор* x^{*} *– решение задачи нелинейного программирования.*

Справедливость этого утверждения непосредственно вытекает из теорем 4 и 5.

Теорема 6 – достаточное условие оптимальности, но оно, вообще говоря, не является необходимым. Например, седловых точек не существует в случае, если $D^{*} = \varnothing$. Существуют также задачи нелинейного программирования, у которых $D^{*} \neq \varnothing$, но их функция Лагранжа не имеет седловых точек ([2], стр. 238). Тем не менее, для некоторых классов задач нелинейного программирования наличие седловой точки функции Лагранжа является не только достаточным, но и необходимым условием оптимальности (см. ***теоремы Куна-Таккера*** в разделе «Выпуклое программирование»).

Элементы выпуклого анализа

Выпуклый анализ – это раздел современного математического анализа, посвященный изучению выпуклых множеств и выпуклых функций. Изложение основ выпуклого анализа можно найти, например, в [2], [3], [9], [12], [13], [20], [23], [25]. Аппарат выпуклого анализа широко используется во многих математических дисциплинах, в моделировании различных явлений и в численных методах решения разнообразных прикладных задач. В частности, выпуклый анализ лежит в основе теории экстремальных задач.

Приведем основные сведения из выпуклого анализа необходимые для изучения теории и методов решения экстремальных задач.

1. Выпуклые множества

Определение 1. Пусть заданы точки $a, b \in E^n$ *и любое* $t \in [0,1]$. *Линейная комбинация* $ta + (1-t)b$ *называется* ***выпуклой комбинацией*** *точек* a *и* b.

Часто выпуклую комбинацию записывают в виде $b + t\,(a-b)$ или $t_1 a + t_2 b$; $t_1 \geq 0$; $t_2 \geq 0$; $t_1 + t_2 = 1$.

Легко увидеть, что эти формы записи эквивалентны.

<u>*Определение 2.*</u> *Множество* $[a, b]$ *всех* ***выпуклых комбинаций*** *точек* a, b *называется* ***отрезком прямой****, соединяющим эти точки.*

<u>*Определение 3.*</u> *Множество* $D \subset E^n$ *называется* ***выпуклым****, если отрезок* $[x, y]$ *включается в* D *для любых* $x, y \in D$.

<u>*Теорема 1.*</u> *Пусть имеется семейство выпуклых множеств* $\{D_\alpha\}$. *Тогда множество* $D = \bigcap_\alpha D_\alpha$ *является выпуклым.*

<u>*Доказательство.*</u> Пусть $x, y \in D$, тогда $x, y \in D_\alpha$, $\forall \alpha$. Так как все множества D_α – выпуклые, $[x, y] \subset D_\alpha$, $\forall \alpha$, откуда $[x, y] \subset \bigcap_\alpha D_\alpha$. Таким образом, $[x, y] \subset D$. Что и требовалось.

Прежде чем сформулировать следующую теорему, напомним определение операций сложения множеств и умножения множества на число. Для множеств D_i и чисел α_i, $i = 1, ..., m$,

$$\sum_{i=1}^{m} \alpha_i D_i = \left\{ x \in E^n : \ x = \sum_{i=1}^{m} \alpha_i x_i, \ \forall x_i \in D_i, \ i = 1, ..., m \right\}.$$

<u>*Теорема 2.*</u> *Пусть для всех* $i = 1, ..., m$ *множества* $D_i \subset E^n$ *– выпуклые,* $\alpha_i \in R$. *Тогда выпукло и множество* $D = \sum_{i=1}^{m} \alpha_i D_i$.

<u>*Доказательство.*</u> Пусть $x, y \in D$, тогда существуют такие векторы $x_i, y_i \in D_i$, $i = 1, ..., m$, что $x = \sum_{i=1}^{m} \alpha_i x_i$, $y = \sum_{i=1}^{m} \alpha_i y_i$. Пусть $t \in [0, 1]$. Так как все множества D_i являются выпуклыми, то для любого $i = 1, ..., m$ имеем включение $t x_i + (1-t) y_i \in D_i$. Следовательно, $t x + (1-t) y = \sum_{i=1}^{m} \alpha_i \left(t x_i + (1-t) y_i \right) \in D$, что и означает выпуклость множества D.

<u>*Теорема 3.*</u> *Пусть $D \subset E^n$ – выпуклое множество, тогда его замыкание $\bar{D}$ также выпукло.*

<u>*Доказательство.*</u> Пусть $x, y \in \bar{D}$, то есть x, y – предельные точки множества D. Тогда существуют последовательности $\{x_i\}_{i=1,2,...}$, $\{y_i\}_{i=1,2,...} \subset D$ такие, что $\lim_{i \to \infty} x_i = x$, $\lim_{i \to \infty} y_i = y$.

Пусть t – любое из отрезка $[0,1]$. Тогда
$$t x + (1-t) y = t \lim_{i \to \infty} x_i + (1-t) \lim_{i \to \infty} y_i = \lim_{i \to \infty} \left(t x_i + (1-t) y_i \right).$$

В силу выпуклости множества D выполняются включения $t x_i + (1-t) y_i \in D$, $\forall i = 1, 2, ...$ Следовательно, $t x + (1-t) y \in \bar{D}$, что и означает выпуклость множества $\bar{D}$.

<u>*Теорема 4.*</u> *Пусть D – выпуклое множество, тогда его внутренность $\operatorname{int} D$ также выпукла.*

Выше было приведено определение выпуклой комбинации двух векторов. Обобщим это понятие на

случай произвольного конечного числа векторов.

<u>Определение 3.</u> *Линейная комбинация* $\sum_{i=1}^{m} t_i x_i$ *векторов* $\{x_i\} \subset E^n$, $i = 1, ..., m$ *называется* ***выпуклой комбинацией***, *если* $t_i \geq 0$, $i = 1, ..., m$ *и* $\sum_{i=1}^{m} t_i = 1$.

<u>Определение 4.</u> *Множество всевозможных выпуклых комбинаций любого конечного числа векторов из множества* $D \subset E^n$ *называется* ***выпуклой оболочкой*** *множества* D *и обозначается* $convD$.

Очевидно, что для всякого D множество $convD$ является выпуклым. Нетрудно показать, что множество D является выпуклым тогда и только тогда, когда $D = convD$.

Возможен и другой подход к определению выпуклой оболочки множества. Выпуклой оболочкой множества D называется наименьшее выпуклое множество, содержащее D, то есть пересечение всех выпуклых множеств, содержащих D. Эти определения выпуклой оболочки эквивалентны.

<u>Определение 5.</u> *Вектор* x *из выпуклого множества* D *называется* ***крайней точкой*** *множества* D, *если он не является выпуклой комбинацией никаких двух других векторов из* D.

Легко увидеть, что любая крайняя точка выпуклого множества является его граничной точкой, но не всякая граничная точка является крайней.

2. Выпуклые конусы

<u>*Определение 1.*</u> *Множество* $K \subset E^n$ *называется* ***выпуклым конусом***, *если*

1. *для любых* $x \in K$ *и* $t > 0$ *выполняется включение* $t\,x \in K$,
2. *для любых* $x, y \in K$ *выполняется включение* $x + y \in K$.

Легко убедиться в справедливости следующей теоремы.

<u>*Теорема 1.*</u> *Выпуклый конус является выпуклым множеством.*

Следующие 4 теоремы устанавливают некоторые операции допустимые в классе выпуклых конусов. (Рекомендуем доказать теоремы 2 – 4 самостоятельно.)

<u>*Теорема 2.*</u> *Пусть имеется семейство выпуклых конусов* $\{K_\alpha\}$. *Тогда множество* $K = \bigcap_{\alpha} K_\alpha$ *является выпуклым конусом.*

<u>*Теорема 3.*</u> *Пусть* $K_i \subset E^n$, $i = 1, \ldots, m$ *– выпуклые конусы. Тогда множество* $K = \sum_{i=1}^{m} K_i$ *также выпуклый конус.*

<u>*Теорема 4.*</u> *Пусть* K *– выпуклый конус. Тогда* $\bar{K}$ *также выпуклый конус.*

<u>*Теорема 5.*</u> *Пусть* $K \subset E^n$ *– выпуклый конус. Тогда* $\operatorname{int} K$ *также выпуклый конус.*

Легко увидеть, что нулевой вектор пространства E^n является предельной точкой любого выпуклого конуса. Вектор 0 называется ***вершиной*** выпуклого конуса. Выпуклый конус может иметь не более одной крайней точки и этой крайней точкой может быть только вершина конуса.

<u>*Определение 2.*</u> *Линейная комбинация* $\sum_{i=1}^{m} t_i x_i$ *векторов* $\{x_i\}$, $i=1,\dots,m$, *называется* ***конической комбинацией****, если* $t_i \geq 0$, $i=1,\dots,m$.

<u>*Определение 3.*</u> *Множество всевозможных конических комбинаций любого конечного числа векторов из множества* $D \subset E^n$ *называется* ***конической оболочкой*** *множества* D *и обозначается* $coneD$.

Очевидно, что для всякого множества D множество $coneD$ является выпуклым конусом.

<u>*Определение 4.*</u> *Пусть* $h \in E^n$ *– ненулевой вектор. Множество* $L=\{x: \; x=t\,h, \; \forall t \geq 0\}$ *называется* ***лучом****, а вектор* h *называется* ***направляющим вектором*** *этого луча.*

Очевидно, что луч – выпуклый замкнутый конус.

<u>*Определение 5.*</u> *Пусть* K *– выпуклый конус. Луч* $L \subset K$ *называется* ***крайним лучом***, *если он*

не принадлежит конической оболочке двух других лучей этого конуса.

Легко увидеть, что любой крайний луч выпуклого конуса принадлежит его границе, но не всякий луч, принадлежащий границе, является крайним лучом.

3. Выпуклые функции

<u>*Определение 1.*</u> *Функция* f, *определенная на* E^n, *называется* ***выпуклой***, *если для любых* x, y *и любого* $t \in [0, 1]$ *выполняется неравенство*

$$f\left(t\,x + (1-t)\,y\right) \le t\,f(x) + (1-t)\,f(y). \qquad (1)$$

Если при $x \ne y$ *и* $t \in (0,1)$ *неравенство (1) выполняется как строгое, то функция* f *называется* ***строго выпуклой***.

<u>*Определение 2.*</u> *Функция* f, *определенная на* E^n, *называется* ***вогнутой (строго вогнутой)***, *если функция* $-f$ *является выпуклой (строго выпуклой).*

Очевидно, что любая строго выпуклая (строго вогнутая) функция является выпуклой (вогнутой) функцией, но не наоборот.

Приведем некоторые операции допустимые в классе выпуклых функций.

<u>Теорема 1.</u> Пусть все функции f_i, $i=1,...,m$, *выпуклы на* E^n, *числа* $\alpha_i \ge 0$. *Тогда функция* $f=\sum_{i=1}^{m}\alpha_i\, f_i$ *также выпукла.*

<u>Доказательство.</u> Пусть заданы векторы x, y и число $t\in[0,1]$. Так как функции f_i, $i=1,...,m$, выпуклы, то для всех i выполняются неравенства $f_i\big(t\,x+(1-t)\,y\big) \le t\, f_i(x)+(1-t)\, f_i(y)$. Умножая эти неравенства на неотрицательные величины α_i и суммируя их по i, получим неравенство

$$\sum_{i=1}^{m}\alpha_i f_i\big(t\,x+(1-t)\,y\big) \le t\sum_{i=1}^{m}\alpha_i f_i(x)+(1-t)\sum_{i=1}^{m}\alpha_i f_i(y).$$

Следовательно, $f\big(t\,x+(1-t)\,y\big) \le t\, f(x)+(1-t)\, f(y)$. Что и требовалось.

<u>Теорема 2.</u> Пусть на E^n *определены функции* f_i, $i=1,...,m$, $f(x) = \max\limits_{1\le i\le m}\{ f_i(x)\}$. *Если все* f_i *– выпуклые, то функция* f *также выпуклая.*

<u>Доказательство.</u> Пусть заданы векторы x, y и число $t\in[0,1]$. Так как функции f_i выпуклы, то для всех i выполняются неравенства $f_i\big(t\,x+(1-t)\,y\big) \le t\, f_i(x)+(1-t)\, f_i(y)$. Следовательно,

$$f_i\big(t\,x+(1-t)\,y\big) \le t\max_{1\le i\le m}\{ f_i(x)\} + (1-t)\max_{1\le i\le m}\{ f_i(y)\}$$

для всех i. Из полученных неравенств имеем

$$\max_{1\le i\le m}\left\{ f_i\left(t\,x+(1-t)\,y\right)\right\} \le t\max_{1\le i\le m}\left\{ f_i\left(x\right)\right\}+(1-t)\max_{1\le i\le m}\left\{ f_i\left(y\right)\right\},$$

то есть $f\left(t\,x+(1-t)\,y\right)\le t\,f\left(x\right)+(1-t)\,f\left(y\right)$. Что и требовалось.

Приведем теоремы о суперпозициях выпуклых функций.

<u>*Теорема 3.*</u> *Пусть функция φ определена на отрезке $\left[a,b\right]\subset R$ и является на нем выпуклой и неубывающей; функция ψ выпукла на выпуклом множестве $G\subset E^n$, $\psi\left(x\right)\in\left[a,b\right]$, $f\left(x\right)=\varphi\left(\psi\left(x\right)\right)$ для всех $x\in G$. Тогда функция f выпукла на G.*

<u>*Доказательство.*</u> Пусть $x,y\in G$, $t\in\left[0,1\right]$. Тогда $\psi\left(t\,x+(1-t)\,y\right)\le t\psi\left(x\right)+(1-t)\psi\left(y\right)$ в силу выпуклости функции ψ на G. Очевидно, что $t\psi\left(x\right)+\left(1-t\right)\psi\left(y\right)\in\left[a,b\right]$. Поэтому, а также в силу монотонности и выпуклости φ на $\left[a,b\right]$, имеем

$$\varphi\left(\psi\left(t\,x+(1-t)\,y\right)\right)\le\varphi\left(t\psi\left(x\right)+(1-t)\psi\left(y\right)\right)\le$$
$$\le t\varphi\left(\psi\left(x\right)\right)+(1-t)\varphi\left(\psi\left(y\right)\right).$$

Следовательно, $f\left(t\,x+(1-t)\,y\right)\le t\,f\left(x\right)+\left(1-t\right)f\left(y\right)$. Что и требовалось.

<u>*Теорема 4.*</u> *Пусть A – матрица размерности $m\times n$, b – вектор размерности m, ψ – функция, определенная и выпуклая на многообразии*

$\{y : y \in E^m,\ y = Ax + b,\ \forall x \in E^n\}, \quad f(x) = \psi(Ax + b).$ *Тогда функция* f *выпукла на* E^n.

<u>*Доказательство.*</u> Пусть заданы векторы x, y и число $t \in [0,1]$. Тогда имеем

$$f(tx + (1-t)y) = \psi\Big(A\big(tx + (1-t)y\big) + b\Big) =$$

$$= \psi\Big(t(Ax + b) + (1-t)(Ay + b)\Big) \le$$

$$\le t\psi(Ax + b) + (1-t)\psi(Ay + b) = t f(x) + (1-t) f(y).$$

Что и требовалось.

Далее покажем, что выпуклость функции многих переменных можно установить, исследуя на выпуклость ее сужения на всевозможные прямые в E^n. Выпуклость функции одной переменной установить зачастую значительно проще, чем выпуклость функции многих переменных.

Пусть заданы функция f и векторы $x, s \in E^n$. Сужение φ функции f на прямую R определим следующим образом:

$$\varphi(t) = f(x + ts). \tag{2}$$

<u>*Теорема 5.*</u> *Функция* f *является выпуклой тогда и только тогда, когда выпуклой является и функция* φ*, определенная по формуле (2) при любых* $x, s \in E^n$.

<u>*Доказательство.*</u> <u>*Необходимость.*</u> Пусть f –

выпуклая функция, $x, s \in E^n$. Покажем, что функция φ также является выпуклой. Пусть $a, b \in R,\ t \in [0,1]$. Тогда

$$\varphi\big(t\,a + (1-t)b\big) = f\big(t(x+as) + (1-t)(x+bs)\big) \le$$

$$\le t\, f(x+as) + (1-t)\, f(x+bs) = t\,\varphi(a) + (1-t)\varphi(b).$$

<u>*Достаточность.*</u> Предположим, что для произвольных $x, s \in E^n$ функция φ – выпуклая. Пусть $x, y \in E^n$ и $t \in [0,1]$. Тогда

$$f\big(t\,x + (1-t)y\big) = f\big(y + t(x-y)\big) = \varphi(t) =$$

$$= \varphi\big(t \times 1 + (1-t)\times 0\big) \le t\varphi(1) + (1-t)\varphi(0) =$$

$$= t\, f(x) + (1-t)\, f(y).$$

Что и требовалось.

Далее установим связь между выпуклыми множествами и выпуклыми функциями.

Пусть $c \in R$ – некоторая константа. Множество $E_c(f) = \{x:\ x \in E^n,\ f(x) \le c\}$ называется ***лебеговым множеством*** функции f.

<u>*Теорема 6.*</u> *Пусть функция f выпукла на E^n. Тогда любое ее лебегово множество выпукло.*

<u>*Доказательство.*</u> Пусть $x, y \in E_c(f),\ t \in [0,1]$. Тогда из $f(x) \le c,\ f(y) \le c$ и в силу выпуклости f

$f\big(t\,x + (1-t)y\big) \le t\, f(x) + (1-t)\, f(y) \le c$. Таким образом, $[x,y]\subset E_c(f)$, что и означает выпуклость множества $E_c(f)$.

Эта теорема устанавливает одностороннюю связь между выпуклыми множествами и выпуклыми функциями. Утверждение, обратное теореме 6, не имеет места.

4. Критерии выпуклости дифференцируемых функций

В этом параграфе мы рассмотрим свойства выпуклых дифференцируемых функций. Известно, что выпуклая на всем пространстве функция непрерывна ([13], стр. 93 – 94) и дифференцируема в любой точке по любому направлению ([13], стр. 94 – 95). Для дальнейшего изложения материала нам, в первую очередь, понадобятся критерии выпуклости дифференцируемых функций, с которыми мы здесь и познакомимся.

Теорема 1. *Пусть функция* f *определена и дифференцируема на* E^n. *Тогда для того, чтобы* f *была выпуклой, необходимо и достаточно, чтобы для любых* x, x_0 *выполнялось неравенство*

$$f(x)-f(x_0)\geq\langle f'(x_0),x-x_0\rangle. \qquad (1)$$

(Неравенство (1) будем называть ***градиентным***.)

Доказательство. *Необходимость.* Пусть функция f выпукла на E^n, $x_0,\ x \in E^n$ и $t \in (0, 1]$. В силу выпуклости f справедливо неравенство

$$f\left(x_0 + t\left(x - x_0\right)\right) \le f\left(x_0\right) + t\left(f\left(x\right) - f\left(x_0\right)\right).$$

Откуда $f(x) - f(x_0) \ge \frac{1}{t}\left[f\left(x_0 + t\left(x - x_0\right)\right) - f\left(x_0\right)\right]$. Переходя в этом неравенстве к пределу при $t \to +0$, получим (1).

<u>*Достаточность.*</u> Пусть для любой пары точек имеет место (1), $x, y \in E^n$, $t \in [0,1]$. Положим $x_0 = tx + (1-t)y$. Из (1) имеем

$$f(x) - f(x_0) \ge \left\langle f'(x_0), x - x_0 \right\rangle, \tag{2}$$

$$f(y) - f(x_0) \ge \left\langle f'(x_0), y - x_0 \right\rangle. \tag{3}$$

Умножая (2) на t, а (3) – на $(1-t)$ и складывая эти неравенства, получим

$$t\,f(x) + (1-t)\,f(y) - f(x_0) \ge$$

$$\ge \left\langle f'(x_0),\ tx + (1-t)y - x_0 \right\rangle = \left\langle f'(x_0),\ x_0 - x_0 \right\rangle = 0.$$

Откуда $f\left(tx + (1-t)y\right) \le t\,f(x) + (1-t)\,f(y)$, что и означает выпуклость функции f.

<u>*Теорема 2.*</u> *Пусть функция f определена и непрерывно дифференцируема на E^n. Тогда для того, чтобы f была выпуклой, необходимо и достаточно, чтобы для любых x, y выполнялось неравенство*

$$\left\langle f'(x) - f'(y), x - y \right\rangle \geq 0. \qquad (4)$$

(Неравенство (4) называется ***вариационным***.)

<u>*Доказательство.*</u> <u>*Необходимость.*</u> Пусть функция f выпукла на E^n, $x, y \in E^n$. В силу теоремы 1 имеют место неравенства

$$f(x) - f(y) \geq \left\langle f'(y), x - y \right\rangle$$

$$f(y) - f(x) \geq \left\langle f'(x), y - x \right\rangle.$$

Сложив эти неравенства, получим (4).

<u>*Достаточность.*</u> Пусть для любых x, y имеет место (4). Тогда для любого $t \in [0,1]$ имеем $\left\langle f'(y + t(x - y)) - f'(y), t(x - y) \right\rangle \geq 0$. Перепишем полученное неравенство следующим образом: $\left\langle f'(y + t(x - y)), x - y \right\rangle \geq \left\langle f'(y), x - y \right\rangle$. Проинтегрируем это неравенство по t на отрезке $[0,1]$, что возможно в силу непрерывной дифференцируемости функции f. Получим

$$\int_0^1 \left\langle f'(y + t(x - y)), x - y \right\rangle dt \geq \int_0^1 \left\langle f'(y), x - y \right\rangle dt.$$

Откуда $f(x) - f(y) \geq \left\langle f'(y), x - y \right\rangle$. Тогда согласно теореме 1 получаем выпуклость функции f.

Заметим, что в одномерном случае неравенство (4) равносильно монотонности (неубыванию) первой производной функции f. Поэтому это свойство градиента выпуклой функции также принято называть ***монотонностью***.

<u>Теорема 3.</u> Пусть функция f определена и дважды дифференцируема на E^n. Тогда для того, чтобы f была выпуклой, необходимо и достаточно, чтобы матрица $f''(x)$ была неотрицательно определена на E^n.

<u>Доказательство.</u> Пусть $x,\ y \in E^n$. Применяя формулу конечных приращений к неравенству (4), получим

$$\left\langle f''\big(y+\theta(x-y)\big)(x-y),(x-y)\right\rangle \geq 0, \qquad (5)$$

где $\theta \in (0,1)$. Согласно теореме 2 неравенство (4) эквивалентно выпуклости функции f. Тогда неравенство (5) также эквивалентно выпуклости функции f. В силу произвольности векторов x, y (5) и означает неотрицательную определенность матрицы $f''(x)$ на E^n. Что и требовалось.

Легко убедиться, что в одномерном случае выпуклость дважды дифференцируемой функции эквивалентна неотрицательности ее второй производной.

Заметим, что критерием вогнутости функции f является неположительная определенность матрицы $f''(x)$, критерием строгой выпуклости является ее положительная определенность, а критерием строгой вогнутости – отрицательная определенность.

Теорема 3 удобна для установления выпуклости функций. Рассмотрим пример. Выясним, при каких условиях квадратичная функция

$f(x) = \frac{1}{2}\langle Ax, x\rangle + \langle b, x\rangle + c$ выпукла на E^n. Для этой функции $f'(x) = Ax+b$, $f''(x) = A$. На основании теоремы 3 получаем, что неотрицательная определенность матрицы A эквивалентна выпуклости квадратичной функции.

5. Конусы релаксационных и возможных направлений

Определение 1. *Пусть функция f определена на E^n. Вектор $s \in E^n$ называется* ***релаксационным направлением (направлением убывания)*** *функции f в точке x, если существует число $\alpha > 0$ такое, что для любого $t \in (0, \alpha)$ выполняется неравенство $f(x+ts) < f(x)$.*

Обозначим множество релаксационных направлений функции f в точке x через $R(x)$.

Теорема 1. *Пусть функция f выпукла на E^n. Тогда для любого $x \in E^n$ множество $R(x)$ – выпуклый конус.*

Доказательство. Пусть вектор $s \in R(x)$, число $\lambda > 0$. Тогда согласно определению 1 имеем $f(x+t\lambda s) < f(x)$ для любого $t \in \left(0, \frac{\alpha}{\lambda}\right)$, то есть вектор $\lambda s \in R(x)$.

Проверим теперь выполнение второго требования определения выпуклого конуса. Пусть векторы $s, h \in R(x)$. Согласно определению 1 найдутся $\alpha, \beta > 0$ такие, что $f(x + ts) < f(x)$ при всех $t \in (0, \alpha)$ и $f(x + th) < f(x)$ при всех $t \in (0, \beta)$. Таким образом, оба неравенства справедливы при всех $t \in (0, \gamma)$, где $\gamma = \min\{\alpha, \beta\}$. В силу выпуклости функции f имеем

$$f\left(\frac{1}{2}(x + ts) + \frac{1}{2}(x + th)\right) \le \frac{1}{2} f(x + ts) + \frac{1}{2} f(x + th).$$

Следовательно, $f\left(x + \dfrac{t}{2}(s + h)\right) < f(x)$, то есть $f(x + \tau(s + h)) < f(x)$ при всех $\tau \in (0, \gamma/2)$, где $\tau = \dfrac{t}{2}$. Итак, $s + h \in R(x)$. Что и требовалось.

Релаксационные направления часто используются как при исследовании задач на минимум, так и в различных методах решения оптимизационных задач. В случае, когда решается задача максимизации, используются ***направления возрастания*** функции в точке, удовлетворяющие неравенству $f(x + ts) > f(x)$ при $t \in (0, \alpha)$.

Определение 1 не всегда позволяет непосредственно отыскивать релаксационные направления функции или устанавливать их отсутствие. Для выпуклых дифференцируемых функций в этом может помочь следующая теорема.

<u>*Теорема 2.*</u> *Пусть* f *– выпуклая дифференцируемая в точке* $x \in E^n$ *функция. Тогда*

$$R(x) = \left\{ s:\ s \in E^n,\ \left\langle f'(x), s \right\rangle < 0 \right\}. \qquad (1)$$

<u>*Доказательство.*</u> Докажем сначала включение $R(x)$ во множество $\left\{ s:\ s \in E^n,\ \left\langle f'(x), s \right\rangle < 0 \right\}$. Пусть $s \in R(x)$. Тогда существует $\alpha > 0$ такое, что $f(x+ts) < f(x),\ \forall t \in (0, \alpha)$. Из теоремы 4.1 получаем $f(x+ts) - f(x) \geq t\left\langle f'(x), s \right\rangle,\ \forall t \in (0, \alpha)$. Из этих двух неравенств и следует $\left\langle f'(x), s \right\rangle < 0$. Что и требовалось.

Докажем обратное включение. Пусть имеет место неравенство $\left\langle f'(x),\ s \right\rangle < 0$. Так как по условию функция f дифференцируема в точке x, имеем $f(x+ts) - f(x) = t\left\langle f'(x),\ s \right\rangle + o(x, s, t)$, где $\lim\limits_{t \to 0} \dfrac{o(x, s, t)}{t} = 0$. Поэтому для достаточно малых t знак приращения функции f совпадает со знаком произведения $\left\langle f'(x), s \right\rangle$. Тогда существует $\alpha > 0$ такое, что $f(x+ts) < f(x),\ \forall t \in (0, \alpha)$, то есть $s \in R(x)$. Что и требовалось.

Заметим, что при доказательстве второго включения выпуклость функции не использовалась.

Заметим также, что в условиях теоремы 2 при $f'(x) \neq 0$ конус $R(x)$ является открытым полупространством.

Наконец, легко увидеть, что если функция f вогнута и дифференцируема в точке $x \in E^n$, то вектор s является направлением возрастания функции f в точке x тогда и только тогда, когда выполняется неравенство $\langle f'(x),\ s \rangle > 0$.

В случае, когда функция $f(x)$ линейна ($f(x) = \langle c, x \rangle$), а значит, выпукла и вогнута одновременно, неравенство $\langle c, s \rangle < 0$ задает конус направлений убывания, а $\langle c, s \rangle > 0$ – конус направлений возрастания в любой точке x.

<u>*Определение 2.*</u> *Пусть D – множество из E^n, точка $x \in D$. Вектор $s \in E^n$ называется* ***возможным направлением*** *в точке x для множества D, если существует число $\alpha > 0$ такое, что $x + ts \in D$ для любого $t \in [0, \alpha]$.*

Обозначим множество возможных направлений в точке x для множества D через $K(x)$.

<u>*Теорема 3.*</u> *Пусть $D \subset E^n$ – выпуклое множество, $x \in D$. Тогда $K(x)$ – выпуклый конус.*

<u>*Доказательство.*</u> Пусть вектор $s \in K(x)$, число $\lambda > 0$. Тогда согласно определению 2 имеем $x + t\lambda s \in D$ для любого $t \in \left[0, \alpha/\lambda\right]$, то есть вектор $\lambda s \in K(x)$.

Проверим теперь выполнение второго требования определения выпуклого конуса. Пусть векторы $s, h \in K(x)$. Согласно определению 2 найдутся $\alpha, \beta > 0$ такие, что $x + ts \in D$ при всех $t \in [0, \alpha]$ и $x + th \in D$ при всех $t \in [0, \beta]$. Таким образом, эти включения справедливы при всех $t \in [0, \gamma]$, где $\gamma = \min\{\alpha, \beta\}$. В силу выпуклости множества D имеем $\frac{1}{2}(x + ts) + \frac{1}{2}(x + th) \in D$, то есть $x + \tau(s + h) \in D$ при всех $\tau \in [0, \gamma/2]$, где $\tau = \frac{t}{2}$. Таким образом, $s + h \in K(x)$. Что и требовалось.

Заметим, что если $x \in \operatorname{int} D$, то $K(x) = E^n$.

<u>*Теорема 4*</u>. *Если $D \subset E^n$ – выпуклое множество, точки $x, y \in D$, то вектор $s = y - x \in K(x)$.*

Справедливость этого утверждения непосредственно следует из определений выпуклого множества и возможного направления.

При исследовании задач на условный экстремум нам понадобятся так называемые ***условно релаксационные направления***.

<u>*Определение 3.*</u> *Пусть функция f определена на множестве $D \subset E^n$, точка $x \in D$. Вектор $s \in E^n$ называется* ***условно релаксационным направлением*** *функции f в точке x относительно множества D, если в этой точке направление s является возможным для D и релаксационным*

для функции f.

Обозначим множество условно релаксационных направлений функции f в точке x через $R(x, D)$. Итак, $R(x, D) = K(x) \bigcap R(x)$, а значит, в условиях теорем 1 и 3 множество $R(x, D)$ является выпуклым конусом.

6. Экстремальные свойства выпуклых функций

Данный параграф посвящен изучению экстремумов выпуклых функций на выпуклых множествах. Из теорем, с которыми мы здесь познакомимся, следует, что этот класс задач на экстремум удобен для исследования и решения.

Теорема 1. Пусть D *– выпуклое множество из* E^n, *функция* f *– выпукла на* D. *Тогда всякий локальный условный минимум функции* f *на множестве* D *является и глобальным.*

Доказательство. Пусть x^* – точка локального минимума функции f на множестве D. Тогда существует такое число $\varepsilon > 0$, что для всех $x \in D \bigcap U(x^*, \varepsilon)$ выполняется неравенство

$$f(x^*) \le f(x). \tag{1}$$

(Здесь $U(x^*, \varepsilon) = \{x : \|x - x^*\| < \varepsilon\}$.) Предположим

противное, то есть, что существует точка $y \in D$

такая, что
$$f\left(x^*\right) > f\left(y\right). \tag{2}$$

В силу выпуклости множества D имеем $x = x^* + t\left(y - x^*\right) \in D,\ \forall t \in [0,1]$. Следовательно, $x = x^* + t\left(y - x^*\right) \in D \bigcap U\left(x^*, \varepsilon\right)$ для достаточно малых значений $t > 0$. Из неравенств (1), (2) и в силу выпуклости функции f на множестве D для таких t имеем
$$f\left(x^*\right) \le f\left(x^* + t\left(y - x^*\right)\right) \le$$
$$\le f\left(x^*\right) + t\left(f\left(y\right) - f\left(x^*\right)\right) < f\left(x^*\right).$$
Полученное противоречие доказывает теорему.

Следующие две теоремы устанавливают свойства множества D^* точек условного минимума.

<u>*Теорема 2.*</u> *Пусть D – выпуклое множество из E^n, функция f – выпукла на D. Тогда D^* – выпуклое множество.*

<u>*Доказательство.*</u> Очевидно, что
$$D^* = D \bigcap \left\{x \in E^n : f\left(x\right) \le f\left(x^*\right)\right\}.$$
Поэтому выпуклость множества D^* следует из теорем 3.6 и 1.1.

<u>*Теорема 3.*</u> *Пусть D – выпуклое множество*

из E^n*, функция* f *– строго выпукла на* D*. Тогда множество* D^* *содержит не более одной точки.*

<u>*Доказательство.*</u> Пусть $D^* \neq \varnothing$. Докажем, что оно состоит только из одной точки. Пусть это не так, то есть существуют два различных вектора $x^*, y^* \in D^*$. Тогда в силу выпуклости множества D^* (см. предыдущую теорему) $t\,x^* + (1-t)\,y^* \in D^*$ при любом $t \in (0,1)$, и в силу строгой выпуклости функции f имеем

$$f^* = f\left(x^*\right) = f\left(y^*\right) = f\left(t\,x^* + (1-t)\,y^*\right) <$$
$$< t\,f\left(x^*\right) + (1-t)\,f\left(y^*\right) = f^*.$$

Полученное противоречие доказывает теорему.

Оставшаяся часть параграфа посвящена необходимым и достаточным условиям экстремума.

<u>*Теорема 4.*</u> ***(Критерий условного экстремума в терминах конусов условно релаксационных направлений)*** *Пусть* D *– выпуклое множество из* E^n*, функция* f *выпукла на* D*. Тогда для того, чтобы точка* x^* *была минимумом функции* f *на множестве* D*, необходимо и достаточно, чтобы* $R\left(x^*, D\right) = \varnothing$.

<u>*Доказательство.*</u> <u>*Необходимость.*</u> Пусть x^* – минимум функции f на D. Убедимся, что конус $R\left(x^*, D\right) = \varnothing$. Предположим противное, то есть

существует вектор $s \in R\left(x^*,\, D\right)$. Поскольку по определению $R\left(x^*,\, D\right) = K\left(x^*\right) \bigcap R\left(x^*\right)$, найдется число $\alpha > 0$ такое, что при всех $t \in \left(0,\, \alpha\right)$ имеем $x^* + t\,s \in D$ и $f\left(x^* + t\,s\right) < f\left(x^*\right)$. Таким образом, получено противоречие с тем, что x^* – условный минимум.

<u>*Достаточность*</u>. Пусть $R\left(x^*,\, D\right) = \varnothing$. Докажем, что $x^* \in D^*$. Предположим противное. Пусть существует точка $y \in D$ такая, что

$$f\left(y\right) < f\left(x^*\right). \qquad (3).$$

Обозначим $s = y - x^*$. Согласно теореме 5.4 имеем $s \in K\left(x^*\right)$. Учитывая неравенство (3) и выпуклость функции f, получаем неравенства $f\left(x^* + t\,s\right) \le f\left(x^*\right) + t\left(f\left(y\right) - f\left(x^*\right)\right) < f\left(x^*\right)$ справедливые при всех $t \in \left(0,\, 1\right)$. Значит, $s \in R\left(x^*\right)$. Таким образом, $s \in R\left(x^*,\, D\right)$. Полученное противоречие доказывает теорему.

<u>*Следствие.*</u> *Пусть функция f выпукла на E^n. Тогда для того, чтобы точка x^* была безусловным минимумом функции f, необходимо и достаточно, чтобы $R\left(x^*\right) = \varnothing$.*

Справедливость этого утверждения следует из того, что здесь $D = E^n$ и $K\left(x^*\right) = E^n$ (см. параграф 5).

<u>*Теорема 5.*</u> (***Критерий условного экстремума первого порядка***) *Пусть* $D \subset E^n$ *– выпуклое множество, дифференцируемая функция* f *выпукла на* D. *Тогда для того, чтобы точка* x^* *была минимумом функции* f *на множестве* D,

необходимо и достаточно, чтобы выполнялось условие

$$\left\langle f'\left(x^*\right),\ x - x^*\right\rangle \geq 0, \quad \forall x \in D. \qquad (4)$$

Справедливость этого утверждения следует из теорем 4, 5.2 и 5.4.

<u>*Следствие.*</u> *Пусть* f *– выпуклая и дифференцируемая на* E^n *функция. Тогда для того, чтобы точка* x^* *была безусловным минимумом функции* f, *необходимо и достаточно, чтобы* $f'\left(x^*\right) = 0$.

Справедливость этого утверждения очевидным образом следует из (4) при $D = E^n$.

7. Проекция точки на множество

<u>*Определение 1.*</u> **Проекцией** *точки* $x \in E^n$ *на*

множество $D \subset E^n$ *называется вектор* $P(x) \in D$, *удовлетворяющий условию*

$$\| x - P(x) \| = \min_{y \in D} \| x - y \|. \qquad (1)$$

Иногда проекцию точки x на множество D будем обозначать $P(x, D)$.

Легко увидеть, что $P(x) = x$ тогда и только тогда, когда $x \in D$.

Очевидно также, что имеет место равенство

$$\| x - P(x) \|^2 = \min_{y \in D} \| x - y \|^2. \qquad (2)$$

Таким образом, согласно (2) отыскание проекции является задачей минимизации функции $\| x - y \|^2$ на множестве D. Заметим при этом, что функция $\| x - y \|^2$ является строго выпуклой (в этом нетрудно убедиться, пользуясь критериями, изученными в параграфе 4).

Существование и единственность проекции зависят от свойств множества D. В связи с этим познакомимся со следующими двумя теоремами.

<u>*Теорема 1.*</u> *Пусть* D *– замкнутое множество из* E^n, *тогда* $P(x)$ *существует для любого* $x \in E^n$.

<u>*Доказательство.*</u> Для произвольной точки $z \in D$ обозначим $D(z) = \{ y : y \in D, \| x - y \| \le \| x - z \| \}$.

Очевидно, что $P(x, D) = P(x, D(z))$. Поэтому задача (1) эквивалентна задаче

$$\min_{y \in D(z)} \| x - y \|. \tag{3}$$

Решение задачи (3) существует, так как множество $D(z)$ компактно, а функция $\| x - y \|$ непрерывна.

Нарушение условия теоремы 1 может привести к отсутствию проекции. Например, проекция не существует, если множество D открыто, а вектор x не принадлежит множеству D.

<u>*Теорема 2.*</u> *Пусть* D *– выпуклое замкнутое множество из* E^n, *тогда всякая точка* $x \in E^n$ *имеет единственную проекцию на множество* D.

<u>*Доказательство.*</u> Существование проекции доказано в предыдущей теореме. Докажем ее единственность. Поскольку, как отмечено выше, функция $\| x - y \|^2$ строго выпукла по y, единственность ее минимума на выпуклом множестве D вытекает из теоремы 6.3. Что и требовалось.

Нарушение условия выпуклости множества D в теореме 2 может привести к неединственности проекции.

Далее приведем критерий, который может быть полезен для нахождения проекции.

<u>*Теорема 3.*</u> *Для того, чтобы точка* $P(x) \in D$ *была проекцией* $x \in E^n$ *на выпуклое замкнутое*

множество D*, необходимо и достаточно, чтобы для всех* $y \in D$ *выполнялось неравенство*

$$\langle P(x) - x,\ y - P(x) \rangle \geq 0. \qquad (4)$$

<u>*Доказательство.*</u> Для того, чтобы вектор $P(x) \in D$ удовлетворял равенству (2), необходимо и достаточно, как следует из теоремы 6.5, чтобы в точке $x^* = P(x)$ выполнялось неравенство

$$\langle f'(P(x)),\ y - P(x) \rangle \geq 0, \quad \forall y \in D, \qquad (5)$$

где $f(y) = \|y - x\|^2$. Так как $f'(y) = 2(y - x)$, а $f'(P(x)) = 2(P(x) - x)$, то условия (5) и (4) эквивалентны. Что и требовалось.

Отыскание проекции – это задача минимизации выпуклой функции на выпуклом множестве. Решать такие задачи мы только лишь учимся. Необходимо заметить, что, как правило, отыскать точно проекцию невозможно. Однако при достаточно «простых» множествах D проекцию можно вычислить по явным формулам. Приведем несколько примеров таких множеств.

Неотрицательный ортант:

$$D = E_+^n.$$

Обозначим j-ую координату вектора $P(x)$ через p_j. Тогда

$$p_j = \begin{cases} 0, & x_j < 0, \\ x_j, & x_j \ge 0, \end{cases} \quad j = 1, \ldots, n.$$

Гиперпараллелепипед:

$$D = \left\{ y : y \in E^n,\ \alpha_j \le y_j \le \beta_j,\ j = 1, \ldots, n \right\}.$$

$$p_j = \begin{cases} \alpha_j, & x_j < \alpha_j, \\ x_j, & \alpha_j \le x_j \le \beta_j, \quad j = 1, \ldots, n, \\ \beta_j, & x_j > \beta_j. \end{cases}$$

Шар:

$$D = S[0, r] = \left\{ y : y \in E^n,\ \| y \| \le r \right\}.$$

$$P(x) = \begin{cases} x, & x \in S[0, r], \\ \dfrac{r}{\| x \|} x, & x \notin S[0, r]. \end{cases}$$

Гиперплоскость:

$$D = \left\{ y : y \in E^n,\ \langle c, y \rangle = b \right\},$$

где $c \in E^n,\ b \in R.$

$$P(x) = x + \frac{b - \langle c, x \rangle}{\| c \|^2} c.$$

Полупространство:

$$D = \left\{ y :\, y \in E^n,\ \langle c,\, y \rangle \le b \right\}.$$

$$P(x) = \begin{cases} x, & x \in D, \\ x + \dfrac{b - \langle c, x \rangle}{\| c \|^2} c, & x \notin D. \end{cases}$$

Многообразие:

$$D = \left\{ y :\, y \in E^n,\ \ Ay = b \right\},$$

где A – матрица размерности $m \times n$, $b \in E^m$.

$$P(x) = x - A^{\mathrm{T}} \left(AA^{\mathrm{T}} \right)^{-1} \left(Ax - b \right).$$

В тех же случаях, когда множество D является более «сложным», проекцию находят приближенно с помощью итерационных процедур.

8. Опорные векторы

<u>Определение 1.</u> Вектор $g \in E^n$ *называется* ***опорным*** *в точке* $x \in E^n$ *ко множеству* $D \subset E^n$, *если выполняется неравенство* $\langle g, y - x \rangle \le 0 \quad \forall y \in D.$

При этом для $g \ne 0$ *гиперплоскость* $\left\{ y :\, y \in E^n, \langle g,\, y - x \rangle = 0 \right\}$ *называется* ***опорной*** *в точке* x *ко множеству* D.

Легко увидеть, что опорные векторы определяются не единственным образом. Обозначим через $T(x)$ множество опорных векторов в точке x ко множеству D. Иногда множество векторов опорных в x к D будем обозначать $T(x, D)$.

Очевидно, что нулевой вектор всегда включается во множество $T(x)$, причем если $x \in \operatorname{int} D$, то $T(x) = \{0\}$. Далее в этом параграфе мы изучим условия существования ненулевых опорных векторов. Но прежде приведем следующее определение.

Определение 2. *Вектор g называется* ***строго опорным*** *в точке x ко множеству D, если выполняется неравенство $\langle g, y-x\rangle < 0,\ \forall y \in D,\ y \neq x$.*

Очевидно, что строго опорный вектор является опорным. Обратное утверждение, вообще говоря, неверно.

Теорема 1. *Пусть D – выпуклое множество из E^n и $x \notin \bar{D}$. Тогда существует вектор g строго опорный в точке x ко множеству D.*

Доказательство. Согласно теореме 7.2 существует проекция $P(x, \bar{D})$. Обозначим для краткости $P(x, \bar{D}) = P$ и положим $g = x - P$. Так как $x \notin \bar{D}$, вектор $g \neq 0$. Убедимся, что вектор g – строго опорный к D в точке x. Согласно теореме 1.3 из выпуклости множества D следует выпуклость $\bar{D}$. Тогда из теоремы 7.3 следует, что неравенство $\langle g, y-P\rangle \le 0$ справедливо для всех $y \in \bar{D}$, а значит, и для всех $y \in D$. Преобразуем это неравенство следующим образом:

$$\langle g, (y-x)+(x-P)\rangle = \langle g, y-x\rangle + \langle g, x-P\rangle \le 0,$$

откуда $\langle g, y-x\rangle \le \langle g, P-x\rangle = -\| g \|^2 < 0, \quad \forall y \in D$. Что и требовалось.

<u>*Теорема 2.*</u> *Пусть D – выпуклое множество и $x \notin \operatorname{int} D$. Тогда существует ненулевой опорный вектор g в точке x ко множеству D.*

<u>*Доказательство.*</u> Если $x \notin \bar{D}$, то этот факт следует из теоремы 1. Пусть $x \in \bar{D}$. Тогда из условия теоремы следует, что x – граничная точка множества D. Поэтому существует последовательность $\{x_k\}_{k=1,2,\dots} \not\subset \bar{D}$ такая, что $\lim\limits_{k\to\infty} x_k = x$. Согласно теореме 1 для любого k существует ненулевой вектор g_k строго опорный в точке x_k ко множеству D. Следовательно, для всех $k = 1, 2, \dots$ имеем

$$\langle g_k, y-x_k\rangle < 0, \quad \forall y \in D. \qquad (1)$$

Не нарушая общности, можно считать, что $\| g_k \| = 1$ для всех $k = 1, 2, \dots$ Поэтому последовательность

$\{g_k\}_{k=1,2,\dots}$ имеет предельную точку. Так же без ограничения общности будем считать, что эта последовательность сходится. Положим $g=\lim_{k\to\infty} g_k$. Очевидно, что $g\neq 0$. Перейдем в (1) к пределу по $k\to\infty$. Получим $\langle g,\, y-x\rangle \le 0,\ \ \forall y\in D$. Таким образом, g – опорный вектор в точке x ко множеству D. Что и требовалось.

Замечание. *Ненулевой опорный вектор g в точке x ко множеству D является строго опорным вектором в x ко множеству* $\operatorname{int} D$.

9. Выпуклые многогранные множества

Раздел выпуклого анализа, посвященный выпуклым многогранным множествам, тесно связан с теорией систем линейных неравенств. Достаточно полное и детальное изложение теории систем линейных неравенств и выпуклых многогранных множеств можно найти в ([16], [23], [25]).

Определение 1. *Множество $D\subset E^n$ называется* ***выпуклым многогранным****, если оно представимо как пересечение конечного числа замкнутых полупространств.*

Таким образом, например, множества решений системы линейных алгебраических уравнений $Ax=b$ либо неравенств $Ax\le b$, где A – матрица размерности $m \times n$, $b \in E^m$, являются выпуклыми многогранными множествами.

<u>*Определение 2.*</u> *Ограниченное выпуклое многогранное множество называется* ***выпуклым многогранником***.

Легко увидеть, что справедливы следующие утверждения.

<u>*Теорема 1.*</u> *Выпуклое многогранное множество является выпуклым множеством.*

<u>*Теорема 2.*</u> *Пусть* $D_i \subset E^n$, $i = 1, \dots, m$ *– выпуклые многогранные множества. Тогда* $D = \bigcap_{i=1}^{m} D_i$ *– также выпуклое многогранное множество.*

Особую роль в «устройстве» многогранного множества играет его граница. Следуя ([15], стр. 112 – 114), приведем следующие понятия.

Для каждого выпуклого многогранного множества $D \subset E^n$ существует целое число r $(0 \le r \le n)$ такое, что множество D содержится в некотором линейном многообразии размерности r (r-мерной плоскости), но не содержится целиком ни в какой $(r-1)$-мерной плоскости. При этом существует только одна такая r-мерная плоскость, содержащая множество D. Она называется ***несущей плоскостью*** многогранного множества D, а число r называется ***размерностью*** этого множества. В частности, нульмерный многогранник представляет собой точку n-мерного пространства. Несущей плоскостью одномерного многогранного множества является прямая. В случае $(n-1)$-мерного многогранного множества его несущей плоскостью является некоторая гиперплоскость. В случае же n-мерного многогранного множества несущая плоскость совпадает со всем пространством.

Граница любого r-мерного многогранного множества (при $r>0$) состоит из конечного числа $(r-1)$-мерных многогранных множеств, причем все они имеют различные несущие плоскости. Эти многогранные множества называются $(r-1)$-мерными гранями рассматриваемого r-мерного многогранного множества. Каждая из этих $(r-1)$-мерных граней, в свою очередь, имеет $(r-2)$-мерные грани. Они также являются гранями исходного r-мерного многогранного множества. Аналогично определяются $(r-3)$-мерные грани и грани меньших размерностей. Итак, у r-мерного многогранного множества могут быть грани размерностей $r-1$, $r-2$, …, 2, 1, 0. Одномерные грани называются ***ребрами***. Ребро может быть отрезком, лучом, либо прямой. Нульмерные грани называются ***вершинами***.

Приведем следующую теорему без доказательства.

<u>Теорема 3.</u> Для того, чтобы вектор v из выпуклого многогранного множества D был его вершиной, необходимо и достаточно, чтобы он

был крайней точкой множества D.

<u>*Определение 3.*</u> *Множество* $K \subset E^n$ *называется* ***выпуклым многогранным конусом****, если оно одновременно является выпуклым многогранным множеством и выпуклым конусом.*

Легко увидеть, что справедлива следующая теорема.

<u>*Теорема 4.*</u> *Пусть* $K_i \subset E^n$, $i=1,\dots,m$ *– выпуклые многогранные конусы. Тогда* $K = \bigcap_{i=1}^{m} K_i$ *– также выпуклый многогранный конус.*

Нулевой вектор пространства принадлежит любому выпуклому многогранному конусу и является его единственной вершиной и как вершина

выпуклого конуса и, если он является крайней точкой, то и как вершина многогранного множества. Ребрами выпуклого многогранного конуса могут быть только лучи либо прямые.

Нетрудно увидеть, что луч является выпуклым многогранным конусом.

Приведем ряд утверждений о свойствах выпуклых многогранных множеств.

<u>*Теорема 5.*</u> *Для того чтобы луч* L *из выпуклого многогранного конуса* K *был его ребром, необходимо и достаточно, чтобы он был крайним лучом конуса* K.

<u>*Теорема 6.*</u> ***(О представлении выпуклого многогранного множества)*** *Пусть выпуклое мно-*

гогранное множество D задано системой линейных неравенств $Ax \le b$, где A – матрица размерности $m \times n$, $b \in E^m$, ранг матрицы $r(A) = n$. Тогда множество D представимо в виде $D = M + K$, где выпуклый многогранник $M = convV$, V – совокупность вершин множества D, выпуклый многогранный конус $K = coneS$, S – множество направляющих векторов неограниченных ребер (лучей) множества D.

<u>Теорема 7.</u> Пусть выполнены условия теоремы 6. Тогда для $K = coneS$ в любой точке $x \in D$ справедливо включение $K \subset K(x)$.

<u>Доказательство.</u> Пусть точка $x \in D$, вектор $h \in K$. Покажем, что $h \in K(x)$. По теореме 6 найдутся векторы $z \in M$ и $s \in K$ такие, что $x = z + s$. Пусть t – произвольное положительное число. Рассмотрим точку $x + th = z + (s + th)$. Так как K – выпуклый конус, вектор $s + th \in K$. Поэтому согласно теореме 6 $x + th \in D$, что и доказывает теорему.

10. Конусы опорных векторов

Ранее в параграфе 8 было введено множество $T(x) = \left\{g : g \in E^n, \langle g, y - x \rangle \le 0, \forall y \in D\right\}$ векторов, опорных в точке x ко множеству $D \subset E^n$. Изучим свойства этих множеств.

Теорема 1. *В любой точке $x \in E^n$ множество $T(x)$ является выпуклым замкнутым конусом.*

Доказательство. Пусть $g, q \in T(x)$, $t > 0$, $u > 0$. Для любого $y \in D$ имеем $\langle g, y - x \rangle \le 0$ и $\langle q, y - x \rangle \le 0$, а значит, $\langle tg + uq, y - x \rangle \le 0$, то есть $tg + uq \in T(x)$. Итак, $T(x)$ – выпуклый конус.

Проверим его замкнутость. Пусть g – предельная точка конуса $T(x)$. Это означает, что существует последовательность $\{g_k\}_{k=1,2,\ldots} \subset T(x)$ такая, что $\lim\limits_{k\to\infty} g_k = g$. Для любой точки $y \in D$ справедливы неравенства $\langle g_k, y - x \rangle \le 0, k = 1,2,\ldots$ Следовательно, $\lim\limits_{k\to\infty} \langle g_k, y - x \rangle = \langle g, y - x \rangle \le 0$, то есть $g \in T(x)$. Что и означает замкнутость $T(x)$.

Легко увидеть, что справедливо следующее утверждение.

Теорема 2. *Пусть B и D – множества из E^n, $B \subset D$. Тогда $T(x, D) \subset T(x, B)$.*

Следующая теорема непосредственно вытекает из теорем 2 и 1.

<u>*Теорема 3.*</u> *Пусть* $D_i \subset E^n, i = 1, \ldots, m, D = \bigcap_{i=1}^{m} D_i$.

Тогда $\sum_{i=1}^{m} T(x, D_i) \subset T(x, D)$.

Задача построения конуса векторов опорных для данного множества в данной точке, вообще говоря, является достаточно сложной. Не существует явных формул или конечных алгоритмов, решающих эту задачу в общем случае. Однако для некоторых классов множеств эта задача решается сравнительно просто. Рассмотрим далее несколько случаев таких множеств, которые нам

10. Конусы опорных векторов

Ранее в параграфе 8 было введено множество $T(x) = \left\{ g : g \in E^n, \langle g, y - x \rangle \le 0, \forall y \in D \right\}$ векторов, опорных в точке x ко множеству $D \subset E^n$. Изучим свойства этих множеств.

<u>*Теорема 1.*</u> *В любой точке* $x \in E^n$ *множество* $T(x)$ *является выпуклым замкнутым конусом.*

<u>*Доказательство.*</u> Пусть $g, q \in T(x)$, $t > 0$, $u > 0$. Для любого $y \in D$ имеем $\langle g, y - x \rangle \le 0$ и $\langle q, y - x \rangle \le 0$, а значит, $\langle tg + uq, y - x \rangle \le 0$, то есть $t\,g + u\,q \in T(x)$. Итак, $T(x)$ – выпуклый конус.

Проверим его замкнутость. Пусть g — предельная точка конуса $T(x)$. Это означает, что существует последовательность $\{g_k\}_{k=1,2,\ldots} \subset T(x)$ такая, что $\lim\limits_{k\to\infty} g_k = g$. Для любой точки $y \in D$ справедливы неравенства $\langle g_k, y-x \rangle \le 0, k = 1,2,\ldots$ Следовательно, $\lim\limits_{k\to\infty} \langle g_k, y-x \rangle = \langle g, y-x \rangle \le 0$, то есть $g \in T(x)$. Что и означает замкнутость $T(x)$.

Легко увидеть, что справедливо следующее утверждение.

<u>*Теорема 2.*</u> *Пусть* B *и* D *– множества из* E^n, $B \subset D$. *Тогда* $T(x, D) \subset T(x, B)$.

Следующая теорема непосредственно вытекает из теорем 2 и 1.

<u>*Теорема 3.*</u> *Пусть* $D_i \subset E^n, i = 1, \ldots, m, D = \bigcap\limits_{i=1}^{m} D_i$.

Тогда $\sum\limits_{i=1}^{m} T(x, D_i) \subset T(x, D)$.

Задача построения конуса векторов опорных для данного множества в данной точке, вообще говоря, является достаточно сложной. Не существует явных формул или конечных алгоритмов, решающих эту задачу в общем случае. Однако для некоторых классов множеств эта задача решается сравнительно просто. Рассмотрим далее несколько случаев таких множеств, которые нам потребуются в дальнейшем.

<u>*Теорема 4.*</u> *Пусть* $D=\{y: y\in E^n,\ f(y)\le 0\}$, *где функция f определена, является выпуклой и непрерывно дифференцируемой на* E^n, *точка x такова, что* $f(x)\ge 0$. *Тогда* $f'(x)\in T(x)$.

<u>*Доказательство*</u>. Пусть точка $y\in D$, то есть $f(y)\le 0$. Учитывая, что по условиям $f(x)\ge 0$, имеем $f(y)-f(x)\le 0$. Отсюда и из теоремы 4.1 следует $\langle f'(x), y-x\rangle\le 0$. Что и требовалось.

Получим теперь правило построения конуса опорных векторов для класса множеств, образованных системами выпуклых неравенств. Нам понадобится следующее условие.

<u>*Условие Слейтера*</u>. *Пусть на* E^n *определены функции* f_i, $i=1,\dots,m$. *Говорят, что система неравенств* $f_i(x)\le 0,\ i=1,\dots,m$, *удовлетворяет* ***условию Слейтера*** *относительно некоторого множества* B *из* E^n, *если существует точка* $z\in B$ *такая, что* $f_i(z)<0$ *для всех* $i=1,\dots,m$.

Если данная система неравенств удовлетворяет условию Слейтера относительно $B=E^n$, то будем просто говорить, что она удовлетворяет условию Слейтера.

Для системы выпуклых неравенств выполнение условия Слейтера обеспечивает непустоту внутренности множества D решений системы.

То есть $\operatorname{int} D = \left\{ y : y \in E^n,\ f_i(y) < 0,\ i = 1, \dots, m \right\}$.

<u>*Теорема 5.*</u> *Пусть $D \subset E^n$ – множество решений системы $f_i(y) \le 0,\ i = 1, \dots, m$, удовлетворяющей условию Слейтера, все функции f_i выпуклы и непрерывно дифференцируемы. Пусть точка $x \in D$ такова, что $I(x) \ne \varnothing$. Тогда $T(x) = coneS(x)$, где $S(x) = \left\{ f_i'(x),\ \forall i \in I(x) \right\}$.*

<u>*Доказательство*</u>. Так как

$$coneS(x) = \left\{ g : g = \sum_{i \in I(x)} t_i f_i'(x), \forall t_i \ge 0,\ i \in I(x) \right\},$$

то из теорем 4, 3 и 1 следует, что

$$coneS(x) \subset T(x). \qquad (1)$$

Докажем теперь включение обратное (1). Предположим противное. Это означает, что существует вектор $h \in T(x)$ такой, что $h \notin coneS(x)$. (Очевидно, что $h \ne 0$.) Так как по теореме 9.6 множество $coneS(x)$ является выпуклым и замкнутым, то в силу теоремы 8.1 найдется вектор s строго опорный ко множеству $coneS(x)$ в точке h. Тогда для всех $g \in coneS(x)$

$$\langle s, g - h \rangle < 0. \qquad (2)$$

Отсюда при $g = 0$ получаем

$$\langle s, h \rangle > 0. \qquad (3)$$

Выберем произвольно $i \in I(x)$ и $t > 0$. Положим $g_i = t\, f_i'(x)$. Из неравенства (2) при $g = g_i$ получим $\langle s, t\, f_i'(x) - h \rangle < 0$. Отсюда $\left\langle s,\, f_i'(x) - \frac{1}{t}h \right\rangle < 0$. Устремляя в этом неравенстве t к бесконечности, получим

$$\left\langle f_i'(x),\, s \right\rangle \le 0. \tag{4}$$

Так как $i \in I(x)$, то $f_i(x) = 0$. Из условия Слейтера следует, что нулевое значение функции f_i не является минимальным. Поэтому $f_i'(x) \ne 0$.

Пусть вектор $z \in \operatorname{int} D$. Положим $u = z - x$. Согласно включению (1) $f_i'(x) \in T(x)$. Из замечания к теореме 8.2 следует, что вектор $f_i'(x)$ является строго опорным в точке x ко множеству $\operatorname{int} D$. Поэтому $\langle f_i'(x), z - x \rangle < 0$. Таким образом,

$$\langle f_i'(x), u \rangle < 0. \tag{5}$$

Положим $s(\theta) = s + \theta u$ для произвольного $\theta > 0$. Тогда $\langle f_i'(x), s(\theta) \rangle = \langle f_i'(x), s \rangle + \theta \langle f_i'(x), u \rangle$. Отсюда и из (4), (5) получаем $\langle f_i'(x),\, s(\theta) \rangle < 0$. Таким образом, по теореме 5.2 вектор $s(\theta)$ является релаксационным направлением функции f_i в точке x. Поэтому найдется такое $\alpha_i > 0$, что

$$f_i\big(x + t\, s(\theta)\big) < f_i(x) = 0 \tag{6}$$

для всех $t \in (0, \alpha_i)$. Так как i – номер произвольного активного ограничения, то (6) выполняется для всех $i \in I(x)$.

Пусть теперь $i \notin I(x)$, то есть $f_i(x) < 0$. Тогда в силу непрерывности функции f_i найдется такое $\alpha_i > 0$, что для всех $t \in (0, \alpha_i)$

$$f_i(x + t\,s(\theta)) < 0. \tag{7}$$

Положим $\alpha = \min\limits_{i=1,\dots,m} \{\alpha_i\}$. Тогда с учетом неравенства (6) получаем, что (7) справедливо для всех $t \in (0, \alpha)$ и всех $i = 1,\dots,m$.

Таким образом, согласно (7) справедливо включение $x + t\,s(\theta) \in \operatorname{int} D$ для $t \in (0, \alpha)$.

Из замечания к теореме 8.2 следует, что вектор h (опорный к D в точке x) является строго опорным ко множеству $\operatorname{int} D$ в точке x. Поэтому $\langle h,\, x + t\,s(\theta) - x \rangle < 0$, откуда $\langle h, s(\theta) \rangle < 0$, то есть $\langle h, s + \theta u \rangle < 0,\ \forall \theta > 0$. При $\theta \to 0$ получим $\langle h, s \rangle \le 0$, что противоречит (3). Полученное противоречие доказывает теорему.

Приведем теперь без доказательства правило построения конуса опорных векторов для множества, образованного системой линейных неравенств.

<u>Теорема 6.</u> Пусть $D = \{x : x \in E^n, Ax \le b\}$, *где* A – *матрица размерности* $m \times n$, *вектор* $b \in E^m$,

точка $x \in D$ *такова, что* $I(x) \neq \varnothing$. *Тогда* $T(x)$ *совпадает с конической оболочкой системы векторов* $\{a_i\}$, $i \in I(x)$, *где* a_i – i*-тая строка матрицы* A.

Заметим, что эта теорема, вообще говоря, не является частным случаем теоремы 5, так как в ней не предполагается выполнение условия Слейтера.

11. Выпуклое программирование

Определение 1. ***Задачей выпуклого программирования*** *называется задача нелинейного программирования, где все функции* f, $f_1, \ldots, f_m$ *являются выпуклыми.*

Таким образом, задача выпуклого программирования является задачей условной минимизации, где целевая функция выпукла и допустимая область представляет собой выпуклое множество, образованное системой выпуклых неравенств. Поэтому утверждения, полученные ранее в параграфе 6, справедливы для задачи выпуклого программирования. В данном параграфе мы конкретизируем эти общие результаты и приведем их в форму более удобную для исследования и решения следующей задачи выпуклого программирования:

$$\min f(x) \tag{1}$$

$$f_i(x) \le 0,\ i=1,\dots,m, \tag{2}$$

$$x_j \ge 0,\ j=1,\dots,n. \tag{3}$$

Нам понадобятся некоторые вспомогательные построения в пространстве E^{m+1} векторов $\bar{u} = (u_1,\dots,u_m,u_{m+1})$. Вектор из первых m компонент точки $\bar{u}$ будем обозначать через u. Итак, $\bar{u} = (u, u_{m+1})$.

Для задачи (1) – (3) определим множество

$$A = \left\{\bar{u}:\ \ \bar{u} \in E^{m+1};\ u \ge F(x);\ u_{m+1} \ge f(x);\ x \ge 0\right\},$$

где $F = (f_1,\dots,f_m)$.

Лемма. *Для задачи выпуклого программирования (1) – (3) множество A выпукло.*

Доказательство. Выберем произвольные векторы $\bar{u}, \bar{v}$ из множества A и число $t \in [0,1]$. Тогда существуют точки x и q из E_+^n такие, что $u \ge F(x),\ u_{m+1} \ge f(x)$ и $v \ge F(q),\ v_{m+1} \ge f(q)$. Умножим эти неравенства на t и $(1-t)$, соответственно, и сложим их. В силу выпуклости всех функций получаем

$$tu + (1-t)v \ge tF(x) + (1-t)F(q) \ge F(tx + (1-t)q),$$

$$tu_{m+1} + (1-t)v_{m+1} \ge t\,f(x) + (1-t)\,f(q) \ \ge f(tx + (1-t)q).$$

Из полученных неравенств и вытекает выпуклость множества A.

<u>*Теорема 1.*</u> ***(Теорема Куна-Таккера в форме о седловой точке функции Лагранжа задачи выпуклого программирования)*** *Пусть в задаче выпуклого программирования (1) – (3) система (2) удовлетворяет условию Слейтера относительно E_+^n. Тогда для того, чтобы неотрицательный вектор x^* был решением задачи (1) – (3), необходимо и достаточно, чтобы существовал неотрицательный вектор y^* такой, что $\left(x^*, y^*\right)$ – седловая точка функции Лагранжа.*

<u>*Доказательство.*</u> Поскольку достаточность этого условия уже доказана для произвольной задачи нелинейного программирования (см. теорему 2.6 введения), осталось доказать только необходимость.

<u>*Необходимость.*</u> Пусть x^* – решение задачи (1) – (3). Положим $\bar{w} = \left(0, ..., 0, f^*\right) \in E^{m+1}$. Очевидно, что $\bar{w} \in A$, так как $x^* \geq 0$, $f\left(x^*\right) = f^*$ и $F\left(x^*\right) \leq 0$.

Убедимся, что $\bar{w} \notin \operatorname{int} A$. Предположим противное. Это означает, что найдется точка $\tilde{x} \geq 0$ такая, что $F\left(\tilde{x}\right) < 0$, $f\left(\tilde{x}\right) < f^*$. Следовательно, $\tilde{x}$ – такая допустимая точка, значение целевой функции в которой меньше минимального. Получаем противоречие с тем, что x^* – решение задачи выпуклого программирования.

Итак, $\bar{w} \notin \operatorname{int} A$. Согласно лемме множество A выпукло. Следовательно, выполняются все требования теоремы 8.2. Поэтому существует ненулевой вектор $\bar{g} = (g, g_{m+1}) \in E^{m+1}$ опорный в точке $\bar{w}$ ко множеству A:

$$\langle \bar{g}, \bar{u} - \bar{w} \rangle \leq 0, \quad \forall \bar{u} \in A. \tag{4}$$

Убедимся далее, что все координаты вектора $\bar{g}$ неположительны. Предположим противное. Пусть существует координата $g_j > 0$. Зафиксируем у вектора $\bar{u}$ все компоненты, кроме j-ой. Тогда, учитывая, что произведение $g_j u_j$ может принимать сколь угодно большие значения (в силу неограниченности сверху координаты u_j), получаем противоречие с неравенством (4).

Легко увидеть, что при любом $x \geq 0$ векторы $\bar{u} = (F(x), f(x))$ включаются во множество A. Тогда из (4) имеем:

$$\langle g, F(x) \rangle + g_{m+1}\left(f(x) - f^* \right) \leq 0, \quad \forall x \geq 0. \tag{5}$$

Покажем, что $g_{m+1} < 0$. Пусть это не так. Тогда $g_{m+1} = 0$. Следовательно, $\langle g, F(x) \rangle \leq 0, \quad \forall x \geq 0$. По условию Слейтера существует вектор $z \geq 0$ такой, что $F(z) < 0$. Поэтому $\langle g, F(z) \rangle > 0$. Полученное противоречие и означает, что $g_{m+1} < 0$.

Обозначим $y^* = \dfrac{1}{g_{m+1}} g$. Покажем, что построенный вектор y^* представляет собой искомый вектор множителей Лагранжа. Очевидно, что $y^* \geq 0$ и из (5) получаем

$$f(x) + \left\langle y^*,\ F(x)\right\rangle \geq f^*, \quad \forall x \geq 0. \qquad (6)$$

Отсюда при $x = x^*$ следует

$$\left\langle y^*,\ F\left(x^*\right)\right\rangle \geq 0. \qquad (7)$$

С другой стороны, так как $F\left(x^*\right) \leq 0$ (поскольку $x^* \in D$) и $y^* \geq 0$, получаем неравенство $\left\langle y^*,\ F\left(x^*\right)\right\rangle \leq 0$. Отсюда и из (7) следует, что в точке $\left(x^*,\ y^*\right)$ выполняется условие дополняющей нежесткости:

$$\left\langle y^*,\ F\left(x^*\right)\right\rangle = 0. \qquad (8)$$

Из (6) и (8) имеем

$$f\left(x^*\right) + \left\langle y^*,\ F\left(x^*\right)\right\rangle \leq f(x) + \left\langle y^*,\ F(x)\right\rangle, \ \forall x \geq 0,$$

или, что то же самое,

$$L\left(x^*,\ y^*\right) \leq L\left(x,\ y^*\right), \quad \forall x \geq 0. \qquad (9)$$

Далее, пусть $y \geq 0$. Тогда $\left\langle y,\ F\left(x^*\right)\right\rangle \leq 0$. Отсюда и из (8) получаем неравенство

$$f\left(x^*\right) + \left\langle y,\ F\left(x^*\right)\right\rangle \leq f\left(x^*\right) + \left\langle y^*,\ F\left(x^*\right)\right\rangle.$$

Итак,

$$L\left(x^*,\ y\right) \leq L\left(x^*,\ y^*\right), \quad \forall y \geq 0. \qquad (10)$$

Неравенства (9), (10) и означают, что $\left(x^*, y^*\right)$ – седловая точка функции Лагранжа задачи выпуклого программирования. Что и требовалось.

Прежде, чем познакомиться с еще одним вариантом теоремы Куна-Таккера, приведем следующую теорему, которая представляет собой критерий условного минимума в терминах конусов опорных векторов.

<u>*Теорема 2.*</u> *Пусть f – выпуклая и дифференцируемая на E^n функция, множество D выпукло. Тогда для того, чтобы точка $x^* \in D$ была условным минимумом функции f на множестве D, необходимо и достаточно, чтобы выполнялось включение*

$$-f'\left(x^*\right) \in T\left(x^*\right). \tag{11}$$

Доказательство следует непосредственно из теоремы 6.5 и определения конуса $T(x)$ опорных векторов в точке x ко множеству D.

<u>*Теорема 3.*</u> ***(Теорема Куна-Таккера в дифференциальной форме для задачи выпуклого программирования)*** *Пусть дана задача выпуклого программирования в виде (1), (2), где все функции f, $f_1, ..., f_m$ непрерывно дифференцируемы, система (2) удовлетворяет условию Слейтера. Тогда для того, чтобы вектор $x^* \in D$ был решением задачи (1), (2), необходимо и достаточно, чтобы существовал неотрицательный вектор y^* такой, что выполняются условия*

$$f'\left(x^*\right) + \sum_{i=1}^{m} y_i^* f_i'\left(x^*\right) = 0, \qquad (12)$$

$$\left\langle y,\ F\left(x^*\right)\right\rangle = 0. \qquad (13)$$

Доказательство. Покажем, что условия (12) и (13) эквивалентны включению (11). Пусть точка $x^* \in D$ такова, что $I\left(x^*\right) = \varnothing$. Тогда $f'\left(x^*\right) = 0$ и $y^* = 0$.

Пусть теперь $I\left(x^*\right) \neq \varnothing$. Тогда из теорем 2 и 10.5 следует, что необходимым и достаточным условием экстремума является существование таких множителей $y_i^* \geq 0$, $i \in I\left(x^*\right)$, для которых $-f'\left(x^*\right) = \sum_{i \in I\left(x^*\right)} y_i^* f_i'\left(x^*\right)$. Положим $y_i^* = 0$ для всех $i \notin I\left(x^*\right)$ и получим из последнего равенства условия (12) и (13). Что и требовалось.

В заключение параграфа приведем формулировки двух теорем Куна-Таккера для задачи выпуклого программирования с линейными ограничениями.

Теорема 4. *Пусть в задаче выпуклого программирования (1) – (3) система ограничений (2) имеет вид* $Ax \leq b$, *где* A *– матрица размерности* $m \times n$, b *– вектор размерности* m. *Тогда для того, чтобы неотрицательный вектор* x^* *был решением задачи, необходимо и достаточно, чтобы существовал неотрицательный вектор* y^* *такой,* *что* $\left(x^*, y^*\right)$ *– седловая точка функции Лагранжа данной задачи.*

Отметим, что в этом случае функция Лагранжа имеет вид $L(x, y) = f(x) + \langle y, Ax - b\rangle$.

<u>*Теорема*</u> <u>5.</u> *Пусть в задаче выпуклого программирования (1), (2) целевая функция* f *непрерывно дифференцируема, система ограничений (2) имеет вид* $Ax \le b$, *где* A *– матрица размерности* $m \times n$, b *– вектор размерности* m. *Тогда для того, чтобы вектор* $x^* \in D$ *был решением задачи, необходимо и достаточно, чтобы существовал неотрицательный вектор* y^* *такой, что выполняются условия* $f'\left(x^*\right) + y^* A = 0$, $\left\langle y^*, Ax^* - b\right\rangle = 0$.

Заметим, что в теоремах 4 и 5 не требуется выполнения условия Слейтера, поэтому они не являются частными случаями теорем 1 и 3 и требуют самостоятельного доказательства.

12. Линейное программирование

<u>*Определение 1.*</u> ***Задачей линейного программирования*** *называется задача выпуклого программирования, у которой все функции* $f, f_1, \ldots, f_m$ *являются линейными.*

Таким образом, задача линейного программирования является задачей нахождения условного экстремума линейной функции на выпуклом многогранном множестве, образованном системой линейных неравенств. Целью данного параграфа является изучение тех свойств задачи линейного программирования, которые присущи только этому классу задач и, вообще говоря, не имеют места для задач выпуклого программирования.

В предыдущих разделах пособия мы изучали задачи минимизации. Задача линейного программирования чаще формулируется как задача максимизации. Это объясняется тем, что ее истоки связаны с экономическими приложениями, например, с проблемой планирования производства с целью максимизации дохода.

Итак, ***общая*** форма записи задачи линейного программирования имеет вид:

$$\max \sum_{j=1}^{n} c_j x_j$$

при условиях

$$\sum_{j=1}^{n} a_{ij} x_j = b_i, \quad i = 1, \ldots, k,$$

$$\sum_{j=1}^{n} a_{ij} x_j \le b_i, \quad i = k+1, \ldots, m,$$

$$x_j \ge 0, \quad j = 1, \ldots, s, \ (s \le n).$$

Введем следующие обозначения: $c = (c_1, \ldots, c_n)$; $b = (b_1, \ldots, b_m)$; $A = \{a_{ij}\}$ – матрица размерности

$m \times n$; $a_i = \left(a_{i\,1}, \ldots, a_{i\,n}\right)$ – вектор-строка матрицы A, $i = 1, \ldots, m$; $A_j = \left(a_{1j}, \ldots, a_{m\,j}\right)^T$ – вектор-столбец матрицы A, $j = 1, \ldots, n$.

В дальнейшем будем использовать следующие две формы записи задачи линейного программирования.

Каноническая форма:

$$\max \langle c, x \rangle$$

при условиях

$$Ax = b,$$
$$x \geq 0.$$

Симметричная форма:

$$\max \langle c, x \rangle$$

при условиях

$$Ax \leq b,$$
$$x \geq 0.$$

Задача линейного программирования легко переводится из одной формы записи в другую при помощи простых формальных преобразований.

Легко увидеть, что задача линейного программирования имеет решение, если допустимая область D, определяемая системой ограничений, непуста и целевая функция $\langle c, x \rangle$ ограничена сверху на D. Если же либо множество D пусто, либо функция $\langle c,x\rangle$ неограниченна сверху на D, то задача линейного программирования не имеет решения.

Все свойства задачи выпуклого программирования, естественно, имеют место и для задачи линейного программирования.

Укажем теперь ***основное свойство задачи линейного программирования,*** которое отличает ее от задачи выпуклого программирования.

Теорема 1. Пусть задача линейного программирования в симметричной форме имеет решение. Тогда существует вершина множества D, которая является решением задачи.

Доказательство. Пусть вектор x^* – решение задачи. Очевидно, что ранг системы ограничений задачи равен n, тогда по теореме 9.6 $D=M+K$. Поэтому существуют векторы $z\in M$ и $h\in K$ такие, что $x^*=z+h$. Отсюда

$$\langle c,x^*\rangle=\langle c,z\rangle+\langle c,h\rangle. \tag{2}$$

Как следует из теоремы 9.7, $h\in K\left(x^*\right)$. В то же время, так как x^* – условный максимум, вектор h не является направлением возрастания функции $\langle c,\ x\rangle$ в точке x^*. Тогда согласно замечанию к теореме 5.2 $\langle c,\ h\rangle\le 0$.

Следовательно, из (2) получаем неравенство $\langle c,x^*\rangle\le\langle c,z\rangle$. Так как x^* – точка максимума, это неравенство может выполниться лишь как равенство

$$\langle c, x^* \rangle = \langle c, z \rangle. \quad (3)$$

Так как $z \in M$, а $M = convV$ (см. теорему 9.6), то $z = \sum_{i=1}^{q} t_i v_i$, $t_i \geq 0$, $i = 1, ..., q$, $\sum_{i=1}^{q} t_i = 1$, где векторы $v_i, i = 1, ..., q$ – все вершины выпуклого многогранного множества D. Тогда из (3)

$$\langle c, x^* \rangle = \sum_{i=1}^{q} t_i \langle c, v_i \rangle. \quad (4)$$

Пусть номер $k \in \{1, ..., q\}$ вершины множества D таков, что $\langle c, v_k \rangle = \max_{1 \leq i \leq q} \{\langle c, v_i \rangle\}$. Тогда из равенства (4), учитывая неотрицательность всех t_i, имеем соотношения $\langle c, x^* \rangle \leq \sum_{i=1}^{q} t_i \langle c, v_k \rangle = \langle c, v_k \rangle \sum_{i=1}^{q} t_i = \langle c, v_k \rangle$. Следовательно, так как вектор x^* – точка максимума, $\langle c, x^* \rangle = \langle c, v_k \rangle$. Что и требовалось.

13. Элементы теории двойственности в линейном программировании

Пусть дана задача линейного программирования в симметричной форме записи

$$\max \langle c, x \rangle$$

$$Ax \le b \qquad (1)$$
$$x \ge 0.$$

Приведем определение еще одной задачи линейного программирования, тесно связанной с задачей (1).

Определение 1. Задача линейного программирования

$$\min \langle b, y \rangle$$

при условиях

$$yA \ge c,$$
$$y \ge 0,$$

где вектор $y \in E^m$, *называется* ***двойственной*** *к задаче (1).*

В теории двойственности задачу (1) принято называть ***прямой***. Легко убедиться, что задача, двойственная к двойственной, является прямой задачей линейного программирования. В связи с этим часто пару задач линейного программирования (прямую и двойственную к ней) называют парой ***взаимосопряженных*** задач линейного программирования.

Пусть теперь прямая задача линейного программирования записана в канонической форме

$$\max \langle c, x \rangle$$

$$Ax = b \qquad (2)$$
$$x \ge 0.$$

Тогда задача двойственная к (2) определяется следующим образом:

$$\min\langle b, y\rangle$$

при условиях

$$yA \geq c\,.$$

Многогранную область пространства E^m, определяемую ограничениями двойственной задачи, будем обозначать D .

Следующие три теоремы устанавливают свойства взаимосопряженных задач линейного программирования, вытекающие непосредственно из их определения.

Теорема 1. *Пусть дана пара взаимосопряженных задач линейного программирования в симметричной форме, векторы $x \in D$ и $y \in D$. Тогда*

$$\langle c, x\rangle \leq \langle b, y\rangle\,. \tag{3}$$

Доказательство. Легко увидеть, что из ограничений прямой и двойственной задач линейного программирования следуют неравенства:

$$\langle c, x\rangle \leq \langle yA, x\rangle \leq \langle b, y\rangle.$$

Из них и получаем неравенство (3).

Теорема 2. *Пусть дана пара взаимосопряженных задач линейного программирования в симметричной форме, векторы $\overline{x} \in D$ и $\overline{y} \in D$ таковы, что*

$$\langle c, \overline{x}\rangle = \langle b, \overline{y}\rangle\,. \tag{4}$$

Тогда векторы $\bar{x}, \bar{y}$ *являются решениями, соответственно, прямой и двойственной задач линейного программирования.*

<u>Доказательство.</u> Из теоремы 1 следует, что $\langle c, x\rangle \le \langle b, \bar{y}\rangle$ для любого $x \in D$. Отсюда и из равенства (4) имеем $\langle c, x\rangle \le \langle c, \bar{x}\rangle$ для любого $x \in D$. Таким образом, $\bar{x}$ – решение прямой задачи линейного программирования. Второе утверждение теоремы доказывается аналогично.

<u>Теорема 3.</u> Для того, чтобы множество D *было пустым, достаточно, чтобы функция* $\langle b, y\rangle$ *была неограниченна снизу на множестве* D.

Справедливость этого утверждения следует непосредственно из теоремы 1.

Очевидно, что утверждение теоремы 3 в применении к двойственной задаче принимает вид:

Для того, чтобы множество D *было пустым, достаточно, чтобы функция* $\langle c, x\rangle$ *была неограниченна сверху на множестве* D.

<u>Теорема 4.</u> ***(Теорема двойственности)*** *Пусть одна из пары взаимосопряженных задач линейного программирования в симметричной форме имеет решение. Тогда и двойственная к ней задача тоже имеет решение и для любых векторов* $x^* \in D^*$ *и* $y^* \in D^*$ *выполняется равенство*

$$\left\langle c, x^* \right\rangle = \left\langle b, y^* \right\rangle. \tag{5}$$

<u>*Доказательство.*</u> Пусть имеет решение прямая задача линейного программирования. Тогда из теоремы 11.4 следует, что найдется такой вектор $y^* \geq 0$, что вектор $\left(x^*, y^*\right)$ – седловая точка функции Лагранжа прямой задачи. Таким образом, $L\left(x^*, y\right) \leq L\left(x^*, y^*\right) \leq L\left(x, y^*\right)$ при всех $x \geq 0,\ y \geq 0$. Отсюда, учитывая то, что для задачи (1) функция Лагранжа имеет вид $L(x, y) = -\langle c, x\rangle + \langle y, Ax - b\rangle$, получаем неравенства $-\left\langle c, x^*\right\rangle + \left\langle y, Ax^* - b\right\rangle \leq$ $\leq -\left\langle c, x^*\right\rangle + \left\langle y^*, Ax^* - b\right\rangle \leq -\langle c, x\rangle + \left\langle y^*, Ax - b\right\rangle$ при всех $x \geq 0,\ y \geq 0$. Перепишем их в виде $\left\langle b, y^*\right\rangle + \left\langle x, c - y^*A\right\rangle \leq \left\langle b, y^*\right\rangle + \left\langle x^*, c - y^*A\right\rangle \leq \langle b, y\rangle +$ $+ \left\langle x^*, c - yA\right\rangle$ при всех $x \geq 0,\ y \geq 0$. Отсюда, так как функция Лагранжа для двойственной задачи имеет вид $L\ (y,\ x) = \langle b, y\rangle + \langle x, c - yA\rangle$, получаем

$L\ \left(y^*, x\right) \leq L\ \left(y^*, x^*\right) \leq L\ \left(y, x^*\right),\ \forall y \geq 0,\ x \geq 0$. Последние неравенства означают, что вектор $\left(y^*, x^*\right)$ является седловой точкой функции Лагранжа для двойственной задачи. Тогда из теоремы 11.4 следует, что вектор y^* – решение двойственной задачи линейного программирования.

Для доказательства второго утверждения теоремы выберем точки $x^* \in D^*$ и $y^* \in D^*$. Так как x^* – решение прямой задачи, то по теореме 11.4 существует вектор $\tilde{y} \ge 0$ такой, что $\left(x^*, \tilde{y}\right)$ – седловая точка функции Лагранжа прямой задачи. Следовательно, как доказано выше, $\tilde{y} \in D^*$ и выполняется условие $\left\langle \tilde{y}, Ax^* - b \right\rangle = 0$, то есть $\left\langle \tilde{y}, Ax^* \right\rangle = \left\langle b, \tilde{y} \right\rangle$. С другой стороны, аналогично, вектор $\left(\tilde{y}, x^*\right)$ – седловая точка функции Лагранжа двойственной задачи, а значит, выполняется условие $\left\langle x^*, c - \tilde{y}A \right\rangle = 0$, то есть $\left\langle \tilde{y}A, x^* \right\rangle = \left\langle c, x^* \right\rangle$. Таким образом, $\left\langle c, x^* \right\rangle = \left\langle b, \tilde{y} \right\rangle$. Откуда, учитывая, что $\left\langle b, \tilde{y} \right\rangle = \left\langle b, y^* \right\rangle$, получаем равенство (5). Что и требовалось.

Заметим, что все результаты этого параграфа были изложены применительно к задачам линейного программирования в симметричной форме. Аналогичные утверждения справедливы и для задач в других формах записи с небольшими изменениями, связанными только лишь с конкретной формой.

Подводя итоги полученным в этом параграфе результатам, можно сделать вывод, что имеет место одна из следующих ситуаций:

– обе взаимосопряженные задачи линейного программирования имеют решения;

– обе задачи не имеют решений, так как целевая функция одной из задач неограниченна на допустимой области, а допустимая область двойственной к ней задачи пуста;

– обе задачи не имеют решений, так как их допустимые области пусты.

Литература

Основная литература

1. Ашманов С.А. Линейное программирование [Текст] / С.А. Ашманов. – М.: Наука, 1981. – 304 с.
2. Васильев Ф.П. Численные методы решения экстремальных задач [Текст] / Ф.П. Васильев. – М.: Наука, 1988. – 550 с.
3. Васильев Ф.П. Методы оптимизации [Текст] / Ф.П. Васильев. – М.: Факториал, 2002. – 823 с.
4. Васильев Ф.П. Линейное программирование [Текст] / Ф.П. Васильев, А.Ю. Иваницкий. – М.: Факториал, 1998. – 176 с.
5. Габасов Р.Ф. Методы оптимизации [Текст] / Р.Ф. Габасов, Ф.М. Кириллова. – Минск: изд-во БГУ, 1981. – 352 с.
6. Заботин Я.И. Учебные задания по курсу математические методы в исследовании операций

[Текст] / Я.И. Заботин, А.И. Кораблев, М.И. Крейнин. – Казань: изд-во КГУ, 1984. – 48 с.

7. Заботин Я.И. Лекции по линейному программированию [Текст] / Я.И. Заботин. – Казань: изд-во КГУ, 1985. – 98 с.
8. Заславский Ю.Л. Сборник задач по линейному программированию [Текст] / Ю.Л. Заславский. – М.: Наука, 1969. – 256 с.
9. Карманов В.Г. Математическое программирование [Текст] / В.Г. Карманов. – М.: Физматлит, 2004. – 264 с.
10. Моисеев Н.Н. Методы оптимизации [Текст] / Н.Н. Моисеев, Ю.П. Иванилов, Е.М. Столярова. – М.: Наука, 1978. – 352 с.
11. Мухачева Э.А. Математическое программирование [Текст] / Э.А. Мухачева, Г.Ш. Рубинштейн. – Новосибирск: Наука, 1987. – 272 с.
12. Поляк Б.Т. Введение в оптимизацию [Текст] / Б.Т. Поляк. – М.: Наука, 1983. – 385 с.
13. Сухарев А.Г. Курс методов оптимизации [Текст] / А.Г. Сухарев, А.В. Тимохов, В.В. Федоров. – М.: Физматлит, 2005. – 368 с.

Дополнительная литература

14. Акулич И.Л. Математическое программирование в примерах и задачах [Текст] / И.Л. Акулич. – М.: Высшая школа, 1986. – 319 с.
15. Болтянский В.Г. Математические методы оптимального управления [Текст] / В.Г. Болтянский. – М.: Наука, 1969. – 408 с.
16. Гавурин М.К. Экстремальные задачи с линейными ограничениями [Текст] / М.К. Гавурин, В.Н. Малоземов. – Л.: изд-во ЛГУ, 1984. – 175 с.
17. Гасс С. Линейное программирование [Текст] / С. Гасс. – М.: Физматгиз, 1961. – 304 с.
18. Данциг Дж. Линейное программирование, его обобщения и применения [Текст] / Дж. Данциг. – М.: Прогресс, 1966. – 600 с.
19. Гейл Д. Теория линейных экономических моделей [Текст] / Д. Гейл. – М.: ИЛ, 1963. – 418 с.
20. Заботин Я.И. Выпуклые функции и выпуклые множества [Текст] / Я.И. Заботин. – Казань, изд-во КГУ, 1968. – 40 с.
21. Заботин Я.И. Методы возможных направлений в математическом программировании [Текст] / Я.И. Заботин. – Казань, изд-во КГУ, 1968. – 56 с.
22. Канторович Л.В. Математические методы организации и планирования производства [Текст] / Л.В. Канторович. – Л.: изд-во ЛГУ, 1939. –

68 с.

23. Кун Г. Линейные неравенства и смежные вопросы [Текст] / Г. Кун, А. Таккер. – М.: ИЛ, 1959. – 470 с.
24. Кюнци Г.П. Нелинейное программирование [Текст] / Г.П. Кюнци, В Крелле. – М.: Сов. рад., 1965. – 303 с.
25. Рокафеллар Р. Выпуклый анализ [Текст] / Р. Рокафеллар. – М.: Мир, 1973. 469 с.
26. Юдин Д.Б. Линейное программирование (теория, методы и приложения) [Текст] / Д.Б. Юдин, Е.Г. Гольштейн. – М.: Наука, 1969. – 424 с.

Сетевые ресурсы

27. http://kek.ksu.ru/EOS/MO/ASP/links.asp
28. http://emics.ksu.ru/Link_r.phtml
29. Кораблев А.И. Электронная образовательная система «Исследование операций и методы оптимизации (ОРТ)» [Интернет-ресурс] / А.И. Кораблев, О.А. Кашина, Н.М. Яковлева, В.А. Галеева. – ОФАП 2379, ВНТИЦ 50200300154, 2003. http://kek.ksu.ru/EOS/MO/uchebnik.asp