Министерство образования и науки Российской Федерации

САНКТ-ПЕТЕРБУРГСКИЙ
ГОСУДАРСТВЕННЫЙ ПОЛИТЕХНИЧЕСКИЙ УНИВЕРСИТЕТ

Л. В. Петухов   Г. А. Серёгин   Е. А. Родионова

МЕТОДЫ ОПТИМИЗАЦИИ. ЗАДАЧИ ВЫПУКЛОГО ПРОГРАММИРОВАНИЯ

Учебное пособие

ИЗДАТЕЛЬСТВО
ПОЛИТЕХНИЧЕСКОГО УНИВЕРСИТЕТА

Санкт-Петербург
2014

УДК 519.85(075.8)

Петухов Л. В. **Методы оптимизации. Задачи выпуклого программирования :** учеб. пособие / Л. В. Петухов, Г. А. Серёгин, Е. А. Родионова. – СПб. : Изд-во Политехн. ун-та, 2014. – 99 с.

Пособие соответствует государственному образовательному стандарту дисциплины «Методы оптимизации» направления бакалаврской подготовки «Прикладная математика и информатика» и других направлений подготовки бакалавров. В кратком изложении рассматриваются теория аффинных множеств, полиэдров, выпуклых множеств, теория отделимости выпуклых множеств, выпуклые функции. Сформулированы и доказаны необходимые и достаточные условия оптимальности для различных типов задач выпуклого программирования. Приведены элементы теории двойственности. Рассмотрены задачи линейного программирования и симплекс-метод их решения.

Пособие предназначено для студентов, изучающих проблемы минимизации функций. Оно также может быть полезно аспирантам и преподавателям вузов.

Это пособие представляет собой дополненное переиздание учебного пособия «Методы решения задач выпуклого программирования» Л.В. Петухова, Г.А. Серёгина, изданного в 1991 году в ЛГТУ. В настоящее издание включен ряд новых разделов, дополнены разделы, имевшиеся в прежнем, а также изменены доказательства некоторых теорем. В список литературы включены современные книги по теме пособия.

Печатается по решению редакционно-издательского совета
Санкт-Петербургского государственного политехнического университета.

ВВЕДЕНИЕ

Экстремальные задачи встречаются во всех областях науки и техники. Они отражают естественное стремление исследователя найти наилучшее из возможных решений, соответствующее заданному критерию качества. Математически экстремальная задача описывается некоторым множеством Ω элементов любой природы , на котором задана функция φ_0 со значениями в R^1. Ставится задача найти такой элемент $x_* \in \Omega$, для которого выполняется одно из следующих неравенств

$$\varphi_0(x_*) \le \varphi_0(x),\ \forall x \in \Omega \text{ (задача поиска минимума) },$$

$$\varphi_0(x_*) \ge \varphi_0(x),\ \forall x \in \Omega \text{ (задача поиска максимума) }.$$

Так как $\max(\varphi_0) = \min(-\varphi_0)$, то мы будем рассматривать ,как правило, задачу поиска минимума.

В зависимости от множества Ω и вида функции φ_0 возникают разные классы экстремальных задач. В данном учебном пособии рассматриваются экстремальные задачи выпуклого анализа. При этом элементы x являются элементами n-мерного векторного пространства R^n . Множество $\Omega \subset R^n$ называется множеством допустимых точек (или множеством ограничений – в зависимости от типа задачи); x_*– оптимальное решение (оптимальный план); $\varphi_0(x)$– функция цели.

Следует заметить, что теория и методы решения экстремальных задач выпуклого анализа, называемых задачами выпуклого программирования, наиболее развиты. Задачи выпуклого программирования включают в себя сложные задачи, например, в которых функции недифференцируемые.

Приведем обозначения, которые будут использованы ниже.

Везде далее будет использоваться n-мерное векторное пространство R^n.

Элементы R^n обозначим малыми латинскими буквами, кроме букв i, j, k, l, m, n. При необходимости латинские буквы можно снабжать индексами внизу, например, a, b, a_1, b_k, c_l.

Вектор представим столбцом, состоящим из n компонент, каждая из которых имеет индекс вверху в круглых скобках. Нулевой вектор будем обозначать через 0. Таким образом,

$$a = \begin{pmatrix} a^{(1)} \\ a^{(2)} \\ \dots \\ a^{(n)} \end{pmatrix}, \quad c_l = \begin{pmatrix} c_l^{(1)} \\ c_l^{(2)} \\ \dots \\ c_l^{(n)} \end{pmatrix}, \quad 0 = \begin{pmatrix} 0 \\ 0 \\ \dots \\ 0 \end{pmatrix}.$$

Транспонирование вектора обозначим буквой ^{T} справа вверху a^T, b^T, a_1^T, b_k^T, c_l^T, 0^T. Таким образом, транспонированный вектор представляется строкой, состоящей из n компонент

$$a^T = \begin{pmatrix} a^{(1)} & a^{(2)} & \dots & a^{(n)} \end{pmatrix}, \quad c_l^T = \begin{pmatrix} c_l^{(1)} & c_l^{(2)} & \dots & c_l^{(n)} \end{pmatrix}.$$

Числа обозначим малыми греческими буквами, иногда снабженными индексами внизу $\alpha, \beta, \gamma_1, \lambda_k, \varepsilon_0$.

Множества в R^n будем обозначать большими латинскими буквами, иногда снабженными индексами внизу S, T, P_1, S_k.

Для матриц используем большие буквы – A, B и т.д. Как правило, матрицы состоят из векторов строк или векторов столбцов, например, если $a_1, \dots, a_m$ – векторы, то в матрице

$$A = \begin{pmatrix} a_1^{(1)} & \dots & a_1^{(n)} \\ . & . & . \\ a_m^{(1)} & \dots & a_m^{(n)} \end{pmatrix}$$

строки матрицы являются транспонированными векторами $a_1^T, \dots, a_m^T$.

В скалярном произведении векторов и произведении матрицы на вектор будем (без ущерба для наглядности) опускать точку

$$x^T y = \sum_{i=1}^{n} x^{(i)} y^{(i)}, \quad y^T A x = \sum_{i,j=1}^{n} a_j^{(i)} x^{(i)} y^{(j)}.$$

Везде, в основном, используется сферическая норма вектора

$$\|x\| = \sqrt{x^T x}\ .$$

Однако, иногда удобно использовать также кубическую норму

$$\|x\| = \max_{1 \le i \le n} \left|x^{(i)}\right|.$$

Будем писать $x \ge y$, если $x^{(i)} \ge y^{(i)}$, $i=1, \ldots, n$.

Определим множества

$$\lambda S = \left\{ y \,\middle|\, y = \lambda x,\ x \in S, \lambda \in R \right\}, \quad S + T = \left\{ z \,\middle|\, z = x + y,\ x \in S, y \in T \right\}.$$

Для ε-окрестности точки x_0 используем обозначение

$U_\varepsilon(x_0) = U(x_0, \varepsilon) = \left\{ x \in R^n \middle| \|x - x_0\| < \varepsilon \right\}$, в случае ε=1 обозначение ε будем опускать : $U(x_0) = U(x_0, 1)$, а для x_0 =0 обозначать $U(0) = U$.

Замыкание и внутренность множества S обозначаются ,

соответственно, через $\overline{S}$ и $\overset{O}{S}$.

Для дифференцируемой функции $\varphi(x)$ будем обозначать ее градиент в точке x через $\nabla\varphi(x)$, где

$$\nabla^T \varphi\ (x) = \left(\partial\varphi / \partial x^{(1)} \quad \ldots \quad \partial\varphi / \partial x^{(n)} \right).$$

Для дважды дифференцируемой функции $\varphi(x)$ будем обозначать матрицу ее вторых производных через

$$H(x) = \begin{pmatrix} \partial^2\varphi / \partial x^{(1)} \partial x^{(1)} & \ldots & \partial^2\varphi / \partial x^{(1)} \partial x^{(n)} \\ .. & .. & .. \\ \partial^2\varphi / \partial x^{(n)} \partial x^{(1)} & \ldots & \partial^2\varphi / \partial x^{(n)} \partial x^{(n)} \end{pmatrix}.$$

Она носит название матрицы Гессе или гессиана. Гессиан является симметричной матрицей, так как $\partial^2\varphi / \partial x^{(i)} \partial x^{(j)} = \partial^2\varphi / \partial x^{(j)} \partial x^{(i)}$.

Примеры задач математического программирования.

1. Задача о рационе

Пусть имеется n видов продуктов и m видов питательных веществ.

Известно

a_{ij} – содержание $i^{го}$ питательного вещества в единице веса $j^{го}$ продукта, $i=1...m$, $j=1...n$;

b_i – норма потребления $i^{го}$ питательного вещества в соответствующих единицах;

c_j – стоимость единицы веса $j^{го}$ продукта.

Требуется составить рацион так, чтобы удовлетворить потребность во всех питательных веществах и минимизировать стоимость рациона.

Обозначим через x_j – количество $j^{го}$ продукта в рационе. Очевидно, что $x_j \ge 0$. Величина $\sum_{j=1}^{n} a_{ij}x_j$ представляет количество $i^{го}$ питательного вещества в рационе. Тогда множество допустимых точек Ω:

$$x_j \ge 0, j = 1...n$$

$$\sum_{j=1}^{n} a_{ij}x_j \ge b_i, i = 1...m,$$

а функция цели (критерий качества): $f(x) = \sum_{j=1}^{n} c_j x_j \to \min$

2. Транспортная задача

Пусть имеется n пунктов хранения, в которых сосредоточен однотипный груз, и m пунктов потребления, в которые груз должен быть доставлен.

Известны:

матрица тарифов перевозок $\|c_{ij}\|$,

где c_{ij} - стоимость перевозки единицы груза из $i^{го}$ в $j^{й}$ пункт;

a_i – объем хранимого груза в $i^{м}$ пункте отправления, $i = 1...n$;

b_j – потребность в грузе $j^{м}$ пункте назначения , $j = 1...m$.

Требуется составить план перевозок так, чтобы стоимость перевозок была минимальной.

Обозначим x_{ij} – объем перевозок из $i^{го}$ в $j^{й}$ пункт,

$i = 1...n, j = 1...m$.Очевидно, что $x_{ij} \geq 0$.

Имеем тогда

$\sum_{j=1}^{m} x_{ij}$ - количество груза, вывозимого из пункта i;

$\sum_{i=1}^{n} x_{ij}$ - количество груза, ввозимого в пункт j.

Множество Ω будет выглядеть так:

$$
\begin{aligned}
& x_{ij} \geq 0,\ i = 1...n, j = 1...m \\
& \sum_{j=1}^{m} x_{ij} \leq a_i,\ i = 1...n \\
& \sum_{i=1}^{n} x_{ij} = b_j,\ j = 1...m
\end{aligned}
,
$$

а функция цели (стоимость перевозок) : $f(x) = \sum_{i=1}^{n} \sum_{j=1}^{m} c_{ij} x_{ij} \to \min$

Сформулированная транспортная задача называется открытой.В её постановке не предполагается, что весь груз из каждого пункта хранения должен быть вывезен. Поэтому исходные данные задачи должны удовлетворять естественному условию $\sum_{i=1}^{n} a_i \geq \sum_{j=1}^{m} b_j$.

Транспортная задача называется закрытой, если в её постановке присутствует условие обязательного вывоза всего хранимого груза. Тогда должно выполняться условие разрешимости $\sum_{i=1}^{n} a_i = \sum_{j=1}^{m} b_j$.

Отметим, что приведённые постановки называются классическими и могут быть дополнены различными условиями, отражающими практические ситуации (запрет на поставку , обязательность поставки, ограничение по пропускной способности маршрута, возврат порожней тары и т. д.).

3. Обработка экспериментальных данных

Пусть дана таблица наблюдений значений некоторой величины y для ряда известных значений x:

Таблица 1

x	x_1	x_2	x_3	…	x_n
y	y_1	y_2	y_3	…	y_n

Известна аналитическая зависимость между величинами x и y , включающая неизвестные параметры $a,b,c,$ а также ряд условий относительно a,b,c

$$\varphi_0(x,a,b,c) = ae^{-\frac{(x-b)^2}{c}}$$
$$\varphi_1(a,b,c) \le 0$$
$$\varphi_2(a,b,c) = 0$$

Ставится задача: найти параметры $a,\ b,\ c$ из данных таблицы 1.

Воспользуемся принципом наименьших квадратов и выберем в качестве критерия качества

$$f(x) = \sum_{i=1}^{n} [y_i - \varphi_0(x_i,a,b,c)]^2 \to \min$$

Тогда множество допустимых точек Ω для переменных a,b,c :

$$\varphi_1(a,b,c) \le 0$$
$$\varphi_2(a,b,c) = 0$$

I. ЭЛЕМЕНТЫ ВЫПУКЛОГО АНАЛИЗА

Рассматриваются аффинные множества, полиэдры, выпуклые множества, выпуклые функции и их свойства. Основными результатами в этой главе являются разделение выпуклых множеств, теорема Фаркаша, понятие субградиента функции.

1.1. Аффинные множества

Определение 1.1. Пусть $x_1, x_2 \in R^n$. Множество

$$M = \left\{ x \in R^n \middle| x = (1-\lambda)x_1 + \lambda x_2, \lambda \in R \right\}$$

называется прямой, проходящей через x_1, x_2.

Определение 1.2. Множество M называется аффинным множеством, если для $\forall x_1, x_2 \in M$ в M содержится и прямая, проходящая через них.

Теорема 1.1. Подпространства в R^n являются аффинными множествами, содержащими 0 (нулевой элемент).

Доказательство. Пусть M – аффинное множество и $0 \in M$. Тогда $\forall x_1, x_2 \in M$

$$(1-\lambda) \cdot 0 + \lambda x_1 = \lambda x_1 \in M, \tag{1.1}$$

$x_1/2+x_2/2 \in M$, так как M – аффинное множество. С другой стороны, $2(x_1/2+x_2/2) \in M$, что видно из (1.1). Следовательно,

$$x_1+x_2 \in M. \tag{1.2}$$

Из (1.1) и (1.2) следует, что M – подпространство.

Обратное очевидно. Если M – подпространство, то оно аффинное множество, содержащее нулевой элемент. Теорема доказана.

Определение 1.3. Транслянтом множества $M \subset R^n$ для элемента $z \in R^n$ называется множество

$$N = M + z = \{y | y = x + z,\ x \in M\}$$

Определение 1.4. Множества N и M параллельны ($N||M$), если $\exists z \in R^n$ такой, что N есть транслянт M для z: $N=M+z$.

Лемма 1.1. Транслянт аффинного множества есть аффинное множество.

Д о к а з а т е л ь с т в о. Возьмем $\forall y_1, y_2 \in N$. Но $y_1=z+x_1$, $y_2=z+x_2$, где $x_1, x_2 \in M$. Построим прямую, проходящую через y_1 и y_2. Она определяется уравнением $y = (1-\lambda)y_1 + \lambda y_2$. Подставляя сюда y_1 и y_2, получим $y = (1-\lambda)x_1 + \lambda x_2 + (1-\lambda)z + \lambda z = x + z$, где $x = (1-\lambda)x_1 + \lambda x_2 \in M$ (так как M – аффинное множество). Следовательно, $z + x = y \in N$. Лемма доказана.

Лемма 1.2. Пусть $M_1||M$ и $M_2||M$. Тогда $M_2||M_1$.

Д о к а з а т е л ь с т в о. Так как $M_1||M$ и $M_2||M$, то для $\forall x \in M$ получим $x = x_1 + z_1$, $x = x_2 + z_2$ где $x_1 \in M_1$, $x_2 \in M_2$, а z_1, $z_2 \in R^n$. Отсюда $x_2 = x_1 + z$, где $z = z_2 - z_1$, откуда следует, что $M_2||M_1$. Лемма доказана.

Всякое подпространство (как аффинное множество) имеет бесчисленное число параллельных ему аффинных множеств $L=z+M$ для $\forall z \in R^n$. Важно знать и обратное, имеется ли подпространство, которому параллельно каждое аффинное множество.

Теорема 1.2. Каждое аффинное множество параллельно единственному подпространству.

Д о к а з а т е л ь с т в о. Пусть M – аффинное множество. Возьмем $\forall z \in M$. Тогда по теореме 1.1 $L=M-z$ является подпространством, так как $0 \in L$. Возьмем $\forall z_1, z_2 \in M$ и построим два подпространства $L_1=M-z_1$ и

$L_2=M-z_2$. По лемме 1.2 $L_2||L_1$, т.е. $\exists y \in R^n$ такой, что $x_2 = x_1+y$ для $\forall x_1 \in L_1$ *и* $\forall x_2 \in L_2$.

Покажем, что $x_2 \in L_1$. Действительно, так как $0 \in L_2$, то $-y \in L_1$ и $y \in L_1$, а следовательно, и $x_2=x_1+y \in L_1$. Аналогично устанавливается, что $x_1 \in L_2$. Теорема доказана.

Определение 1.5. Размерностью аффинного множества является размерность параллельного ему пространства. Размерность M обозначается $dim\ M$.

Аффинное множество размерности 0 называется точкой.

Аффинное множество размерности 1 называется прямой или линией.

Аффинное множество размерности n-1 называется гиперплоскостью.

Аффинное множество размерности n представляет собой R^n.

Аффинные множества размерностей от 2 до n-2 называются плоскостями.

Определение 1.6. Говорят, что вектор a ортогонален вектору x, если $a^T x = 0$.

Определение 1.7. Подпространство $L^{\perp} = \left\{ a \in R^n \middle| a^T x = 0,\ \forall x \in L \right\}$ называется ортогональным к подпространству L, причем $dim\ L + dim\ L^{\perp} = n$.

Пусть $a_1, \ldots, a_m$ – базис в $L^{\perp}$. Тогда пространство L можно задать в виде

$$L = \left\{ x \in R^n \middle| a_1^T x = 0,\ \ldots,\ a_m^T x = 0 \right\}$$

Пусть $dim\ L = n-1$. Тогда $dim\, L^{\perp} = 1$ и, следовательно, $L = \left\{ x \in R^n \middle| a^T x = 0 \right\}$.

Определение 1.8. Гиперплоскость есть множество $Q=L+z$, где $dim\ L=n-1$.

Возьмем гиперплоскость Q. Она параллельна подпространству L размерности $n-1$. Следовательно, $dim\ L^{\perp} = 1$. Для $\forall y \in Q$ $y=x+z$, где $x \in L$. Но для x имеем $a^T x = 0$, тогда отсюда $a^T y = a^T x + a^T z = a^T z = \alpha$. Таким образом, $Q = \left\{ y \in R^n \middle| a^T y = \alpha \right\}$, где a и α определены с точностью до одинакового ненулевого множителя.

Теорема 1.3. Для того, чтобы M было аффинным множеством размерности n-m, необходимо и достаточно, чтобы существовали матрица A размерности $m \times n$ ранга m и вектор $b \in R^m$ такие, что $M = \left\{ y \middle| Ay = b \right\}$.

Д о к а з а т е л ь с т в о. Н е о б х о д и м о с т ь. Пусть M – аффинное множество размерности n-m. Оно параллельно подпространству L размерности n-m. Выберем в $L^\perp$ m линейно независимых векторов $a_1, \ldots, a_m$. Тогда

$$L = \left\{ x \in R^n \middle| a_1^T x = 0,\ \ldots,\ a_m^T x = 0 \right\}.$$

Элементы $y \in M$ равны: $y=x+z$, где $x \in L$, $z \in R^n$, следовательно,

$$M = \left\{ y \in R^n \middle| a_1^T y = a_1^T z,\ \ldots,\ a_m^T y = a_m^T z \right\}$$

Располагая a_i^T в виде строк матрицы A, получим

$$Ay = b,$$

где $b^{(i)} = a_i^T \cdot z$, $i=1,\ldots,m$. Очевидно, что ранг матрицы A равен m.

Д о с т а т о ч н о с т ь. Пусть имеются матрица A размерности $m \times n$ ранга m и вектор b. Рассмотрим $M = \left\{ y \middle| Ay = b \right\}$. Покажем, что множество M – аффинное. Выберем $y_1, y_2 \in M$, т.е. $Ay_1 = b$, $Ay_2 = b$. Возьмем $y = (1-\lambda)y_1 + \lambda y_2$ и умножим его справа на матрицу A: $Ay = (1-\lambda)Ay_1 + \lambda Ay_2 = (1-\lambda)b + \lambda b = b$. Таким образом, $y \in M$. Возьмем $y_0 \in M$, удовлетворяющий $Ay_0 = b$, тогда $\forall x = y - y_0$ удовлетворяет однородной системе уравнений $Ax = 0$.

Вектор $x \in L \square M$. Размерность L равна n-m, что следует из теорем линейной алгебры [10, 12]. Следовательно, размерность M равна также n-m. Теорема доказана.

Лемма 1.3. Пусть M_1 и M_2 – аффинные множества. Тогда $M = M_1 \cap M_2$ – аффинное множество.

Доказательство. Возьмем $\forall x_1, x_2 \in M$. Так как $M = M_1 \cap M_2$, то $x_1, x_2 \in M_1$ и $x_1, x_2 \in M_2$, следовательно, прямая T, проходящая через x_1 и x_2, содержится в M_1 и M_2 ($T \subset M_1$ и $T \subset M_2$). Но тогда $T \subset M$. Таким образом, M – аффинное. Лемма доказана.

Из теоремы 1.3 и леммы 1.3 вытекает, что любое аффинное множество M можно представить как пересечение гиперплоскостей

$$M = \bigcap_{i=1}^{m} Q_i,\ Q_i = \left\{ x \in R^n \,\middle|\, a_i^T x = \alpha_i \right\}.$$

Определение 1.9. Пусть $S \subset R^n$. Пересечение всех аффинных множеств, содержащих S, называется аффинной оболочкой S и обозначается через *aff(S)*.

Согласно лемме 1.3 множество *aff(S)* – аффинное.

Пример 1. Пусть в R^2 множество S состоит из двух точек x_1 и x_2. Множество *aff(S)* представляет собой прямую, проходящую через точки x_1 и x_2 (рис. 1).

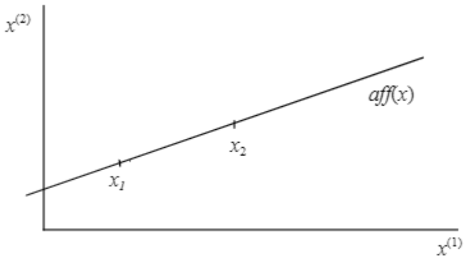

Рис.1

Пример 2. Пусть в R^3 множество S представляет собой треугольник. Множество *aff(S)* представляет собой плоскость, содержащую этот треугольник (рис. 2).

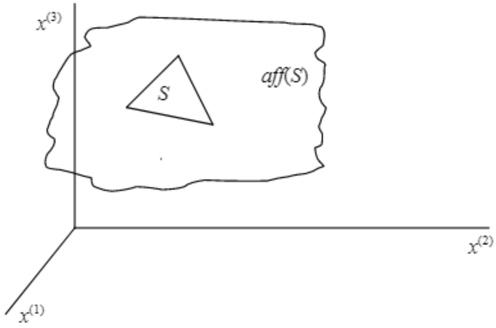

Рис.2

Определение 1.10. Пусть $S \subset R^n$. Множество

$$ac(S) = \left\{ x \middle| \; x = \lambda_1 x_1 + \ldots + \lambda_m x_m, \lambda_1 + \ldots + \lambda_m = 1, x_i \in S, m \geq 1 \right\}$$

называется множеством всех аффинных комбинаций множества *S*.

Лемма 1.4. Множество *ac(S)* – аффинное.

Д о к а з а т е л ь с т в о. Возьмем $\forall x, y \in ac(S)$. Тогда

$x = \lambda_1 x_1 + \ldots + \lambda_m x_m$, $y = \nu_1 y_1 + \ldots + \nu_k y_k$, $x_i, y_j \in S$, $i=1, \ldots, m$, $j=1, \ldots, k$,

$\lambda_1 + \ldots + \lambda_m = 1$, $\nu_1 + \ldots + \nu_k = 1$.

Проведем через *x* и *y* прямую $T = \left\{ z \middle| \; z = (1-\alpha)x + \alpha y \right\}$. Тогда

$z = (1-\alpha)\lambda_1 x_1 + \ldots + (1-\alpha)\lambda_m x_m + \alpha\nu_1 y_1 + \ldots + \alpha\nu_k y_k$. Легко заметить, что

$(1-\alpha)\lambda_1 + \ldots + (1-\alpha)\lambda_m + \alpha\nu_1 + \ldots + \alpha\nu_k = (1-\alpha)(\lambda_1 + \ldots + \lambda_m) + \alpha(\nu_1 + \ldots + \nu_k) = 1 - \alpha + \alpha = 1$,

а $x_i, y_j \in S$. Следовательно, z - аффинная комбинация элементов *S* и $z \in ac(S)$. Лемма доказана.

Теорема 1.4. *aff(S) = ac(S).*

Д о к а з а т е л ь с т в о. Докажем, что $aff(S) \subset ac(S)$. Действительно, *S* является подмножеством *ac(S)* при *m*=1 и *ac(S)* – аффинное. Следовательно, $aff(S) \subset ac(S)$.

Докажем теперь, что $ac(S) \subset aff(S)$. Воспользуемся методом математической индукции. Обозначим через $ac_m(S)$ множество всех аффинных комбинаций, содержащих не более m элементов S.

Случай m=1 тривиален. Рассмотрим m=2. Выберем $\forall y \in ac_2(S)$, тогда $y = \lambda_1 x_1 + \lambda_2 x_2$, $\lambda_1 + \lambda_2 = 1$, $x_1, x_2 \in S$. Очевидно, что $x_1, x_2 \in aff(S)$, следовательно, $y \in aff(S)$, так как $aff(S)$ - аффинное множество. Поэтому $ac_2(S) \subset aff(S)$.

Предположим теперь, что $ac_m(S) \subset aff(S)$ и покажем, что $ac_{m+1}(S) \subset aff(S)$. Возьмем $\forall y \in ac_{m+1}(S)$. Тогда

$$y = \lambda_1 x_1 + \ldots + \lambda_m x_m + \lambda_{m+1} x_{m+1},\ \lambda_1 + \ldots + \lambda_m + \lambda_{m+1} = 1.$$

Очевидно, что хотя бы одно $\lambda_i \neq 1$. Без ограничения общности будем считать, что $\lambda_{m+1} \neq 1$. Представим y в виде

$$y = \left(1 - \lambda_{m+1}\right) z + \lambda_{m+1} x_{m+1}, \qquad z = \frac{\lambda_1}{1 - \lambda_{m+1}} x_1 + \ldots + \frac{\lambda_m}{1 - \lambda_{m+1}} x_m.$$

Но $z \in ac_m(S)$, так как $\dfrac{\lambda_1}{1 - \lambda_{m+1}} + \ldots + \dfrac{\lambda_m}{1 - \lambda_{m+1}} = 1$.

Таким образом, $z \in aff(S)$, $x_{m+1} \in aff(S)$, следовательно, $y \in aff(S)$.

Теорема доказана.

Определение 1.11. Пусть $S \subset R^n$ - произвольное множество. Точка $x \in S$ называется относительно внутренней точкой S, если $\exists \varepsilon > 0$ $U(x, \varepsilon) \cap aff\left(S\right) \subset S$.

Определение 1.12. Совокупность всех относительно внутренних точек множества называется его относительной внутренностью *ri (S)*.

Пример 3. Пусть S представляет собой треугольник в R^3. Тогда $\mathring{S} = \varnothing$. Множество $ri(S) \neq \varnothing$ и представляет собой треугольник S, у которого отброшены граничные отрезки (рис. 3).

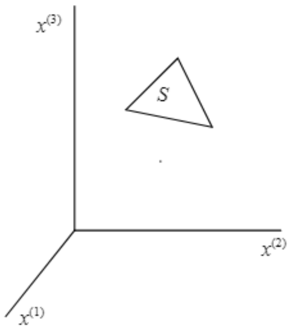

Рис.3

1.2 Полиэдры

Гиперплоскость $Q=\left\{x\in R^n \middle| a^T x=\alpha, a\in R^n, a\neq 0, \alpha\in R^1\right\}$ порождает два замкнутых полупространства

$$Q^- =\left\{x\in R^n \middle| a^T x\le\alpha\right\}, \qquad Q^+ =\left\{x\in R^n \middle| a^T x\ge\alpha\right\}$$

и два открытых полупространства

$$\overset{o}{Q}{}^- =\left\{x\in R^n \middle| a^T x<\alpha\right\}, \qquad \overset{o}{Q}{}^+ =\left\{x\in R^n \middle| a^T x>\alpha\right\}.$$

Вектор нормали к гиперплоскости a направлен в сторону Q^+.

Определение 2.1. Множество $S\subset R^n$ называется полиэдром (или многогранным множеством), если его можно представить как пересечение конечного числа замкнутых полупространств и гиперплоскостей.

$$S=\left\{x\in R^n \middle| a_i^T x\le b^{(i)}, i=\overline{1,k}, a_i^T x=b^{(i)}, i=\overline{k+1,m_1}, a_i\in R^n, b^{(i)}\in R^1, a_i\neq 0\right\} \qquad (2.1)$$

Поскольку каждое из равенств в представлении (2.1) можно заменить системой неравенств

$$a_i^T x\le b^{(i)}, i=\overline{1,k}$$

$-a_i^T x \le -b^{(i)}$, $i = \overline{k+1, m_1}$,

будем рассматривать основное представление полиэдра

$$S_{\text{осн}} = \left\{ x \in R^n \middle| a_i^T x \le b^{(i)}, i = \overline{1,m} \right\} \text{ или } S_{\text{осн}} = \left\{ x \in R^n \middle| Ax \le b \right\},$$

где a_i^T расположены по строкам A, вектор b составлен из величин $b^{(i)}$.

Обозначим $M=\{1,2, \ldots,m\}$ – отрезок чисел натурального ряда от 1 до m. Для $\forall x \in S$ введем в рассмотрение множество номеров активных ограничений $I(x) = \left\{ i \in M \middle| a_i^T x = b^{(i)} \right\}$.

Определение 2.2. Точка $x \in S$ называется вершиной полиэдра, если $rang\, A_{I(x)}=n$, где $A_{I(x)}$ – матрица, составленная из векторов a_i, $i \in I(x)$.

Определение 2.3. Вершина x называется невырожденной, если $|I(x)|=n$ и вырожденной в противном случае ($|I(x)|>n$).

Определение 2.4. Пусть $S \subset R^n$, $S \ne \varnothing$, S – полиэдр. Ограничение с номером i называется жестким, если оно выполняется как равенство в S и нежестким иначе.

Теорема 2.1. Аффинная оболочка непустого полиэдра S порождается системой его жестких ограничений: $aff(S) = \left\{ x \middle| A_{I(S)} x = b_{I(S)} \right\}$, $I(S)$ –множество индексов жестких ограничений.

Д о к а з а т е л ь с т в о. Рассмотрим множество $W = \left\{ x \middle| A_{I(S)} x = b_{I(S)} \right\}$. W – аффинное множество (как пересечение гиперплоскостей), $S \subset W$ (очевидно). Следовательно, $aff(S) \subset W$.

Покажем, что $W \subset aff(S)$. Выберем $\forall y \in W$ и установим, что $y \in aff(S)$. Возможны следующие случаи

1) $I(S)=M$, следовательно, все ограничения жесткие, поэтому $y \in S$. Значит, $y \in aff(S)$.

2) $I(S) \ne M$. Обозначим $K(S)=M \setminus I(S)$ – множество индексов нежестких ограничений. Пусть точки $z_i \in S$, в которых $a_i^T z_i < b^{(i)}$. Построим

$z = \sum_{i \in K} \lambda_i z_i$, $\lambda_i > 0$, $\sum_{i \in K} \lambda_i = 1$. Умножим это равенство слева на A:

$Az = \sum_{i \in K} \lambda_i A z_i \le \sum_{i \in K} \lambda_i b = b$. Отсюда $Az \le b$, следовательно, $z \in S$.

В точке z выполняются условия $A_{I(S)}x = b_{I(S)}$ и $A_{K(S)}x < b_{K(S)}$, для точки y справедливо $A_{I(S)}y = b_{I(S)}$. Проведем прямую через точки z и y и выберем на ней точку $\bar{x}$:

$\bar{x} = (1 - \gamma)z + \gamma y = z + \gamma(y - z)$. Подберем γ так, чтобы $\bar{x} \in S$. Запишем для точки y: $y = \frac{\gamma - 1}{\gamma} z + \frac{1}{\gamma} \bar{x}$.

Имеем $z \in S$, следовательно, $z \in aff(S)$. $\bar{x} \in S$, значит, $\bar{x} \in aff(S)$. Тогда $y \in aff(S)$ и , следовательно, $W \subset aff(S)$. Теорема доказана.

Теорема 2.2. Пусть $S \subset R^n$, $S \ne \varnothing$ - полиэдр. Тогда $dim\, S = n\text{-}rang\, A_{I(S)}$.

Д о к а з а т е л ь с т в о. Из предыдущей теоремы $aff(S) = \{x \mid A_{I(S)}x = b_{I(S)}\}$. Тогда по теореме о представлении аффинного множества $dim\, S = n\text{-}rang\, A_{I(S)}$. Теорема доказана.

Теорема 2.3. Пусть $S \subset R^n$, $S \ne \varnothing$ - полиэдр (2.1). Если система

$$y^T A = 0 \qquad (2.2)$$

$$y^T b = 0, \quad y^{(i)} \ge 0, i = 1, 2, \ldots k \qquad (2.3)$$

имеет решение y , для которого $y^{(j)} > 0$ при некотором $j \in \{1, 2, \ldots k\}$, то $j \in I(S)$.

Д о к а з а т е л ь с т в о . Пусть $j \notin I(S)$. Тогда существует $z \in S$, для которого $a_j^T z < b_j$. Из (2.2) следует $y^T Az = 0$, но $y^T Az < y^T b = 0$ из (2.3). Полученное противоречие показывает, что $j \in I(S)$.Теорема доказана.

Пример 4. Пусть полиэдр S задан системой

$x_1 + 2x_2 \le 4, \quad x_1 + x_2 \ge 2, \quad x_2 \ge 2$. Найдём $I(S)$.

Точка $(0,2) \in S$, поэтому $S \ne \varnothing$. Запишем систему условий (2.2) и (2.3):

$$
\begin{aligned}
y_1 - \ \ y_2 \qquad &= 0 \\
2y_1 - \ \ y_2 - \ y_3 &= 0 \\
4y_1 - 2y_2 - 2y_3 &= 0
\end{aligned}
$$

Приводя матрицу системы к треугольному виду, получим общее решение системы уравнений

$y_1 = \beta, y_2 = y_1, y_3 = y_1, \beta > 0$, откуда следует, что $I(S) = \{1,2,3\}$ (рис. 4)

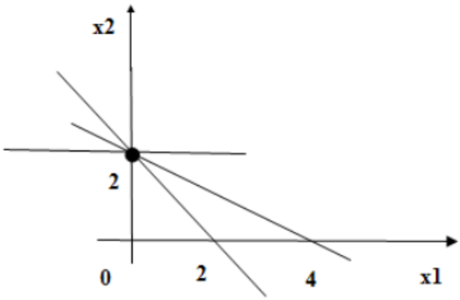

Рис.4

Определение 2.5. Выпуклая оболочка конечного числа точек $\{x_i\}, i = \overline{1,m}$, называется выпуклым многогранником.

Если векторы x_i - x_1, $i = \overline{2,m}$, линейно независимы, то выпуклая оболочка точек $\{x_i\}, i = \overline{1,m}$ называется симплексом с вершинами в точках $\{x_i\}, i = \overline{1,m}$.

Максимальное число линейно независимых векторов в R^n равно n. Поэтому в R^n не может быть симплекса, у которого более, чем n+1 вершина.

1.3. Выпуклые множества

Определение 3.1. Пусть $x_1, x_2 \in R^n$. Отрезком, соединяющим x_1 и x_2, называется множество

$[x_1, x_2] = \left\{ x \in R^n \middle| \; x = (1-\lambda)x_1 + \lambda x_2,\ 0 \le \lambda \le 1 \right\}$.

Определение 3.2. Пусть $x_1, x_2 \in R^n$. Открытым отрезком, соединяющим x_1 и x_2, называется множество

$(x_1, x_2) = \left\{ x \in R^n \middle| \; x = (1-\lambda)x_1 + \lambda x_2,\ 0 < \lambda < 1 \right\}$.

Аналогичным образом можно определить множества (x_1, x_2] и [x_1, x_2).

Определение 3.3. Множество S называется выпуклым, если для $\forall x_1, x_2 \in S \quad [x_1, x_2] \subset S$.

Пример 1. Множество $Q^- = \left\{ x \middle| a^T x \le \alpha, a \ne 0 \right\}$ - выпуклое. Возьмем $\forall x_1, x_2 \in Q^-$ и рассмотрим точку $x = (1-\lambda)x_1 + \lambda x_2,\ 0 \le \lambda \le 1$. Для нее $a^T x = (1-\lambda)a^T x_1 + \lambda a^T x_2 \le (1-\lambda)\alpha + \lambda\alpha = \alpha$. Следовательно, $x \in Q^-$ и множество Q^- - выпуклое.

Пример 2. Множества $Q^+ = \left\{ x \middle| a^T x \ge \alpha, a \ne 0 \right\}$, $Q = \left\{ x \middle| a^T x = \alpha, a \ne 0 \right\}$,

$\overset{o}{Q}{}^- = \left\{ x \middle| a^T x < \alpha, a \ne 0 \right\}$, $\overset{o}{Q}{}^+ = \left\{ x \middle| a^T x > \alpha, a \ne 0 \right\}$ - выпуклые.

Доказательство проводится так же, как и в примере 1.

Пример 3. Множество $U(x_0,\varepsilon)$ – выпуклое. Возьмем $\forall x_1, x_2 \in U(x_0,\varepsilon)$ и построим точку $x = (1-\lambda)x_1 + \lambda x_2$, $0 \le \lambda \le 1$. Для нее

$$\|x - x_0\| = \|(1-\lambda)x_1 + \lambda x_2 - x_0\| = \|(1-\lambda)(x - x_0) + \lambda(x_2 - x_0)\| \le (1-\lambda)\|x_1 - x_0\| + \lambda\|x_2 - x_0\| <$$

$$< (1-\lambda)\varepsilon + \lambda\varepsilon = \varepsilon.$$

Таким образом, $x \in U(x_0,\varepsilon)$ и $U(x_0,\varepsilon)$ – выпуклое.

Теорема 3.1. Пусть S_1 и S_2 – выпуклые множества. Тогда $S_1 \cap S_2$, μS_i, $S_1 + S_2$ – выпуклые множества.

Д о к а з а т е л ь с т в о. Возьмем $\forall x_1, x_2 \in S_1 \cap S_2$. Тогда $x_1, x_2 \in S_1$ и $x_1, x_2 \in S_2$. Так как S_1 и S_2 – выпуклые множества, то $[x_1, x_2] \subset S_1$ и $[x_1, x_2] \subset S_2$. Тогда отсюда $[x_1, x_2] \subset S_1 \cap S_2$.

Возьмем теперь $\forall y_1, y_2 \in \mu S_i$ и построим $y = (1-\lambda) y_1 + \lambda y_2$, $0 \le \lambda \le 1$. Но $y_1 = \mu x_1$, $y_2 = \mu x_2$, $x_1, x_2 \in S$, S выпуклое. Следовательно, $(1-\lambda) x_1 + \lambda x_2 \in S_i$. Но $y = \mu[(1-\lambda) x_1 + \lambda x_2]$, поэтому $y \in \mu S_i$.

Возьмем теперь $\forall y_1, y_2 \in S_1 + S_2$ и построим $y = (1-\lambda) y_1 + \lambda y_2$, $0 \le \lambda \le 1$. Но $y_1 = x_1 + z_1$, $y_2 = x_2 + z_2$, где $x_1, x_2 \in S_1$, а $z_1, z_2 \in S_2$, причем S_1, S_2 - выпуклые. Следовательно, $(1-\lambda) x_1 + \lambda x_2 \in S_1$, $(1-\lambda) z_1 + \lambda z_2 \in S_2$. Тогда $y = (1-\lambda) x_1 + \lambda x_2 + (1-\lambda) z_1 + \lambda z_2 \in S_1 + S_2$. Следовательно, $S_1 + S_2$ - выпуклое. Теорема доказана.

Теорема 3.2. Пусть $S \subset R^n$ - произвольное множество. Для того, чтобы S было выпуклым, необходимо и достаточно, чтобы для $\forall \lambda_1 \ge 0$, $\forall \lambda_2 \ge 0$ $\lambda_1 S + \lambda_2 S = (\lambda_1 + \lambda_2) S$.

Д о к а з а т е л ь с т в о. Н е о б х о д и м о с т ь. Пусть S – выпуклое множество. Покажем , что $\lambda_1 S + \lambda_2 S \subset (\lambda_1 + \lambda_2) S$. Выберем $\forall y \in \lambda_1 S + \lambda_2 S$. Тогда $y = y_1 + y_2$, $y_1 = \lambda_1 x_1$, $x_1 \in S$, $y_2 = \lambda_2 x_2$, $x_2 \in S$.

Получим $y = y_1 + y_2 = \lambda_1 x_1 + \lambda_2 x_2 = (\lambda_1 + \lambda_2)\left(\frac{\lambda_1}{\lambda_1 + \lambda_2} x_1 + \frac{\lambda_2}{\lambda_1 + \lambda_2} x_2\right) \in (\lambda_1 + \lambda_2) S$, так как S – выпуклое.

Покажем теперь, что $(\lambda_1 + \lambda_2) S \subset \lambda_1 S + \lambda_2 S$. Выберем $\forall y \in (\lambda_1 + \lambda_2) S$. Тогда $y = (\lambda_1 + \lambda_2) x = \lambda_1 x + \lambda_2 x$, $x \in S$. Значит, $y \in \lambda_1 S + \lambda_2 S$.

Д о с т а т о ч н о с т ь. Выберем $\forall x_1 \in S$, $\forall x_2 \in S$ и рассмотрим $y = (1-\lambda) x_1 + \lambda x_2$ для некоторого $\lambda \in [0,1]$. По условию $y \in ((1-\lambda) + \lambda) S = S$, следовательно, S – выпуклое. Теорема доказана.

Определение 3.4. Пусть S – множество. Выпуклой оболочкой $co(S)$ называется пересечение всех выпуклых множеств, содержащих S.

Определение 3.5. Пусть S – множество. Множеством всех выпуклых комбинаций S является множество

$$cc(S) = \left\{ x \middle| x = \sum_{i=1}^{m} \lambda_i x_i, x_i \in S, \quad \lambda_i \geq 0, \sum_{i=1}^{m} \lambda_i = 1, m \geq 1 \right\}$$

Лемма 3.1. $cc(S)$ – выпуклое множество.

Д о к а з а т е л ь с т в о. Возьмем $\forall x, y \in cc(S)$. Тогда

$$x = \sum_{i=1}^{m} \nu_i x_i, x_i \in S, \quad \nu_i \geq 0, \ \sum_{i=1}^{m} \nu_i = 1, \quad y = \sum_{j=1}^{l} \mu_j y_j, y_j \in S, \quad \mu_j \geq 0, \ \sum_{j=1}^{l} \mu_j = 1.$$

Составим выпуклую комбинацию

$z = (1-\lambda)x + \lambda y = (1-\lambda)\nu_1 x_1 + \ldots + (1-\lambda)\nu_m x_m + \lambda\mu_1 y_1 + \ldots + \lambda\mu_l y_l$. Коэффициенты $(1-\lambda)\nu_i \geq 0, \ \lambda\mu_j \geq 0$, а $(1-\lambda)\nu_1 + \ldots + (1-\lambda)\nu_m + \lambda\mu_1 + \ldots + \lambda\mu_l = 1$. Следовательно, $z \in cc(S)$. Лемма доказана.

Теорема 3.3. $co(S)=cc(S)$.

Д о к а з а т е л ь с т в о. 1.Множество $cc(S)$ – выпуклое и $S \subset cc(S)$, следовательно,

$co(S) \subset cc(S)$.

2. Обозначим через $cc_m(S)$ множество выпуклых комбинаций не более m элементов множества S. Очевидно, что $cc_1(S) \subset co(S)$ и $cc_2(S) \subset co(S)$, так как $S \subset co(S)$ и $co(S)$ – выпуклое множество.

Предположим теперь, что $cc_m(S) \subset co(S)$ и покажем, что $cc_{m+1}(S) \subset co(S)$. Возьмем $\forall y \in cc_{m+1}(S)$

$$y = \lambda_1 x_1 + \ldots + \lambda_m x_m + \lambda_{m+1} x_{m+1}, \ x_i \in S, \ \lambda_i \geq 0, \ \lambda_1 + \ldots + \lambda_m + \lambda_{m+1} = 1.$$

Так как $\lambda_{m+1} \neq 1$, представим y в виде $y = (1-\lambda_{m+1})z + \lambda_{m+1} x_{m+1}$,

где $$z = \frac{\lambda_1}{1-\lambda_{m+1}} x_1 + \ldots + \frac{\lambda_m}{1-\lambda_{m+1}} x_m.$$

Вектор $z \in cc_m(S) \subset co(S)$, так как $\dfrac{\lambda_i}{1-\lambda_{m+1}} \geq 0, \quad \dfrac{\lambda_1}{1-\lambda_{m+1}} + \ldots + \dfrac{\lambda_m}{1-\lambda_{m+1}} = 1.$

Следовательно, $y \in co(S)$. Теорема доказана.

Следствие 3.1 . Если S – выпуклое множество, то $co(S)=S$.

Следствие 3.2. Пусть S состоит из m точек x_i, $i=1,\dots,m$. Тогда

$$co(S)=\left\{x \middle| x=\sum_{i=1}^{m}\lambda_i x_i,\ x_i \in S,\quad \lambda_i \ge 0, \sum_{i=1}^{m}\lambda_i = 1\right\}.$$

Пример 4. Пусть S состоит из 2^n точек x_i, координаты которых равны ± 1. Тогда $co(S)=\bar{U}_\infty$,

где $$\bar{U}_\infty=\left\{x \in R^n \middle| \ \|x\|_\infty = \max_{1 \le i \le n} \left|x^{(i)}\right| \le 1\right\}.$$

1. Покажем, что $co(S) \subset \bar{U}_\infty$. Возьмем $\forall x \in co(S)$. Тогда по следствию 3.2

$$x=\sum_{i=1}^{2^n}\lambda_i x_i,\quad \sum_{i=1}^{2^n}\lambda_i = 1,\ \lambda_i \ge 0.$$

Рассмотрим $x^{(j)}=\sum_{i=1}^{2^n}\lambda_i x_i^{(j)}$ Координаты $x_i^{(j)}=\pm 1$, причем 2^{n-1} координат равны +1, а 2^{n-1} координат равны -1. Без ограничения общности можно считать, что значения j-х координат первых 2^{n-1} векторов равны +1, а последних 2^{n-1} векторов -1. Тогда

$$x^{(j)}=\sum_{i=1}^{2^{n-1}}\lambda_i - \sum_{i=2^{n-1}+1}^{2^n}\lambda_i .$$

Так как $\lambda_1+\dots+\lambda_{2^n}=1$, все $\lambda_i \ge 0$, то $\min x^{(j)}=-1,\quad \max x^{(j)}=1$,

т.е. $-1 \le x^{(j)} \le 1$ и $x \in \bar{U}_\infty$. Включение $co(S) \subset \bar{U}_\infty$ установлено.

2. Покажем теперь, что $\bar{U}_\infty \subset co(S)$. Возьмем $\forall x \in \bar{U}_\infty$. Координаты $x^{(j)}$ удовлетворяют неравенству $-1 \le x^{(j)} \le 1$.

Представим вектор x в виде выпуклой комбинации двух векторов

$$u_{11}^{\mathrm{T}}=\begin{pmatrix}1 & x^{(2)} & \dots & x^{(n)}\end{pmatrix} \text{ и } \quad v_{11}^{T}=\begin{pmatrix}-1 & x^{(2)} & \dots & x^{(n)}\end{pmatrix} \in \bar{U}_2$$

$$x=(1-\mu_1)u_{11}+\mu_1 v_{11},\ \mu_1=\left(1-x^{(1)}\right)\Big/2 .$$

Для $-1 \le x^{(1)} \le 1$ имеем $0 \le \mu_1 \le 1$. Введем обозначения $x_{11} = u_{11}$, $x_{12} = v_{11}$, $\gamma_{11} = 1 - \mu_1$, $\gamma_{12} = \mu_1$. Тогда

$$x = \sum_{i=1}^{2} \gamma_{1i} x_{1i}, \quad \sum_{i=1}^{2} \gamma_{1i} = 1, \gamma_{1i} \ge 0. \qquad (3.1)$$

Представим теперь каждый вектор x_{1i} в виде выпуклой комбинации двух векторов $u_{2i}^{\mathrm{T}} = \begin{pmatrix} \pm 1 & +1 & x^{(3)} & \dots & x^{(n)} \end{pmatrix}$, $v_{2i}^{\mathrm{T}} = \begin{pmatrix} \pm 1 & -1 & x^{(3)} & \dots & x^{(n)} \end{pmatrix}$. Тогда

$$x = \sum_{i=1}^{2} \gamma_{1i} \left[(1-\mu_2) u_{2i} + \mu_2 v_{2i} \right], \quad \mu_2 = \left(1 - x^{(2)}\right) / 2.$$

Для $-1 \le x^{(2)} \le 1$ будет $0 \le \mu_2 \le 1$. Введем обозначения $x_{21} = u_{21}$, $x_{22} = u_{22}$, $x_{23} = v_{21}$, $x_{24} = v_{22}$, $\gamma_{21} = (1-\mu_2)\gamma_{11}, \gamma_{22} = (1-\mu_2)\gamma_{12}$, $\gamma_{23} = \mu_2 \gamma_{11}$, $\gamma_{24} = \mu_2 \gamma_{12}$. Тогда

$$x = \sum_{i=1}^{2^2} \gamma_{2i} x_{2i}, \quad \sum_{i=1}^{2^2} \gamma_{2i} = 1, \gamma_{2i} \ge 0. \qquad (3.2)$$

Продолжая такой процесс, предположим, что на k-м шаге x представлен в виде выпуклой комбинации векторов $x_{ki}^{\mathrm{T}} = \begin{pmatrix} \pm 1 & \dots & \pm 1 & x^{(k+1)} & \dots & x^{(n)} \end{pmatrix}$

$$x = \sum_{i=1}^{2^k} \gamma_{ki} x_{ki}, \quad \sum_{i=1}^{2^k} \gamma_{ki} = 1, \gamma_{ki} \ge 0. \qquad (3.3)$$

Представим каждый вектор x_{ki} в виде выпуклой комбинации двух векторов $u_{k+1,i}^{\mathrm{T}} = \begin{pmatrix} \pm 1 & \dots & \pm 1 & +1 & x^{(k+2)} & \dots & x^{(n)} \end{pmatrix}$,

$v_{k+1,i}^{\mathrm{T}} = \begin{pmatrix} \pm 1 & \dots & \pm 1 & -1 & x^{(k+2)} & \dots & x^{(n)} \end{pmatrix}$

$$x = \sum_{i=1}^{2^k} \gamma_{ki} \left((1-\mu_{k+1}) u_{k+1,i} + \mu_{k+1} v_{k+1,i} \right), \; \mu_{k+1,i} = \left(1 - x^{(k+1)}\right) / 2.$$

Для $-1 \le x^{(k+1)} \le 1$ будет $0 \le \mu_{k+1} \le 1$. Введем, как и выше, обозначения $x_{k+1,i} = u_{k+1,i}$, $\gamma_{k+1,i} = (1-\mu_{k+1})\gamma_{ki}$, $i = 1, \dots, 2^k$, $x_{k+1,i} = v_{k+1,i}$, $\gamma_{k+1,i} = \mu_{k+1}\gamma_{ki}$, $i = 2^k + 1, \dots, 2^{k+1}$. Тогда

$$x = \sum_{i=1}^{2^{k+1}} \gamma_{k+1,i} x_{k+1,i}, \quad \sum_{i=1}^{2^{k+1}} \gamma_{k+1,i} = 1, \gamma_{k+1,i} \ge 0. \qquad (3.4)$$

Таким образом, предположив (3.3), установили (3.4). Соотношения (3.3) и (3.4) отличаются друг от друга только номерами k и $k+1$. Для k=1 и k=2 соотношения (3.3) установлены (см. (3.1),(3.2)), поэтому (3.3) будут справедливы и для $k=n$

$$x=\sum_{i=1}^{2^n}\lambda_i x_i,\quad \sum_{i=1}^{2^n}\lambda_i=1,\ \lambda_i\ge 0,\tag{3.5}$$

где $x_i=x_{ni}=\begin{pmatrix}\pm1 & \pm1 & \dots & \pm1\end{pmatrix}$, $\lambda_i=\gamma_{ni}$. Включение $\bar{U}_\infty\subset co(S)$ установлено, а следовательно, установлено и $co(S)=\bar{U}_\infty$.

По определению выпуклой оболочки *co(S)* множества S $\forall x\in co(S)$ может быть представлена в виде выпуклой комбинации конечного числа точек из S. Теорема Каратеодори показывает, что минимальное число таких точек не превосходит $n+1$.

Теорема 3.4 Каратеодори. Пусть S – произвольное множество из R^n. Тогда $\forall x\in co(S)$ можно представить в виде выпуклой комбинации не более n+1 точки множества S.

Д о к а з а т е л ь с т в о. Так как *co(S)=cc(S),* то для $\forall x\in co(S)$

$$x=\sum_{i=1}^{m}\lambda_i x_i,\quad \lambda_i\ge 0,\ \sum_{i=1}^{m}\lambda_i=1,\ x_i\in S.$$

Пусть $m>n+1$, тогда векторы $x_2-x_1,\dots,x_m-x_1$ - линейно зависимы, т.е. $\exists\mu_i$, $i=2,\dots,m,$ не все равные нулю, такие, что

$$\sum_{i=2}^{m}\mu_i\left(x_i-x_1\right)=0 \text{ или } \sum_{i=1}^{m}\mu_i x_i=0,\tag{3.6}$$

где $\mu_1=-\mu_2-\ldots-\mu_m$.

Будем считать, что существует хотя бы одно $\mu_j>0$, в противном случае последнее равенство (3.6) можно умножить на -1 и среди $-\mu_i$ окажется положительное число.

Домножим равенство (3.6) на $\varepsilon>0$ и представим $x=\sum_{i=1}^{m}(\lambda_i-\varepsilon\mu_i)x_i$.

Положим $\varepsilon_0 = \min_{\mu_i > 0}\{\lambda_i / \mu_i\} = \lambda_j / \mu_j > 0$. Обозначим $\lambda_i - \varepsilon_0 \mu_i = \nu_i, i = 1,...,m$.

Тогда

$$x = \sum_{i=1}^{m} \nu_i x_i , \tag{3.7}$$

$$\sum_{i=1}^{m} \nu_i = \sum_{i=1}^{m} \lambda_i - \varepsilon_0 \sum_{i=1}^{m} \mu_i = 1 , \quad \nu_i \ge 0, \; i = 1,...,m \; .$$

В равенстве (3.7) $\nu_j = 0$, поэтому вектор x является выпуклой комбинацией m-1 вектора x_i. Перенумеруем x_i из (3.7), тогда получим

$$x = \sum_{i=1}^{m-1} \nu_i x_i \, .$$

Продолжая такой процесс, можно уменьшить количество векторов x_i до n+1. Теорема доказана.

Теорема 3.5. Пусть $S_1, \, S_2 \subset R^n$ - произвольные множества. Тогда $co(S_1+S_2) = co(S_1)+co(S_2)$.

Д о к а з а т е л ь с т в о. Покажем, что $co(S_1+S_2) \subset co(S_1)+co(S_2)$. Имеем $S_1 \subset co(S_1)$, $S_2 \subset co(S_2)$, следовательно, $S_1+S_2 \subset co(S_1)+co(S_2)$.

co(S₁) и co(S₂) – выпуклые множества, значит, $co(S_1)+co(S_2)$ – выпуклое множество. Отсюда получаем, что $co(S_1+S_2) \subset co(S_1)+co(S_2)$.

Покажем теперь, что $co(S_1)+co(S_2) \subset co(S_1+S_2)$. Выберем $\forall x \in co(S_1)+co(S_2)$. Имеем

$$x = \sum_{i=1}^{l} \lambda_i y_i + \sum_{j=1}^{k} \nu_j z_j \, , \; y_i \in S_1, \; z_j \in S_2, \; \lambda_i \ge 0 \, , \; \sum_{i=1}^{l} \lambda_i = 1 \, , \; \nu_j \ge 0 \, , \; \sum_{j=1}^{k} \nu_j = 1 \, .$$

Представим $\lambda_i = \lambda_i \cdot 1 = \lambda_i \sum_{j=1}^{k} \nu_j = \sum_{j=1}^{k} \lambda_i \nu_j$,

$$\nu_j = \nu_j \cdot 1 = \nu_j \sum_{i=1}^{l} \lambda_i = \sum_{i=1}^{l} \nu_j \lambda_i \, .$$

Тогда $x=\sum_{i=1}^{l}\sum_{j=1}^{k}\mu_{ij}\left(y_i+z_j\right)$, где $\mu_{ij}=\lambda_i\nu_j$. Значит, $x=\sum_{i=1}^{l}\sum_{j=1}^{k}\mu_{ij}\left(y_i+z_j\right)\in$ *co(S₁+S₂)*,

так как $y_i+z_j\in S_1+S_2$, $\mu_{ij}\geq 0$, $\sum_{i=1}^{l}\sum_{j=1}^{k}\mu_{ij}=1$. Теорема доказана .

Замыкание и внутренность выпуклого множества.

Определение 3.6. Пусть $S\subset R^n$ - произвольное множество. Точка x принадлежит замыканию множества ($x\in cl(S)$ или $x\in\overline{S}$), если $S\cap U_\varepsilon(x)\neq\varnothing$ для $\forall\varepsilon>0$. Если $S=\overline{S}$, то множество S называется замкнутым.

Определение 3.7. Пусть $S\subset R^n$ - произвольное множество. Точка x принадлежит внутренности множества S ($x\in\operatorname{int}(S)$ или $x\in\overset{\circ}{S}$), если $U_\varepsilon(x)\subset S$ для некоторого $\varepsilon>0$.

Если $S=\overset{\circ}{S}$, то множество S называется открытым.

Теорема 3.6. Пусть S – выпуклое множество. Тогда

1) $\overline{S}$ – выпуклое множество;
2) Если $\overset{\circ}{S}\neq\varnothing$, то $\overset{\circ}{S}$ - выпуклое множество.

Д о к а з а т е л ь с т в о. Выберем $\forall y_1,y_2\in\overline{S}$. Пусть $y=(1-\lambda)y_1+\lambda y_2, 0\leq\lambda\leq 1$ - произвольная точка отрезка $[y_1,y_2]$, $y\neq y_1$, $y\neq y_2$. Построим $U_\varepsilon(y_1)$, $U_\varepsilon(y_2)$ для некоторого $\varepsilon>0$. Очевидно, $\exists x_1\in S\cap U_\varepsilon(y_1)$ и $x_2\in S\cap U_\varepsilon(y_2)$. Положим $x=(1-\lambda)x_1+\lambda x_2, x\in S$, так как S – выпуклое. Рассмотрим

$\|y-x\|\leq(1-\lambda)\|y_1-x_1\|+\lambda\|y_2-x_2\|<(1-\lambda)\varepsilon+\lambda\varepsilon=\varepsilon$, следовательно, $\exists x\in S\cap U(y,\varepsilon)$.

Значит, $y\in\overline{S}$ и $\overline{S}$ - выпуклое.

Докажем второе утверждение теоремы.

Возьмем $\forall x_1, x_2 \in \overset{\circ}{S}$. Тогда $\exists \varepsilon_1 > 0$, $U_{\varepsilon_1}(x_1) \subset S$ и $\exists \varepsilon_2 > 0$, $U_{\varepsilon_2}(x_2) \subset S$.

Рассмотрим $x = (1-\lambda)x_1 + \lambda x_2 \in S$ для некоторого $\lambda \in (0,1)$. Покажем, что $x \in \overset{\circ}{S}$, то есть $\exists \varepsilon > 0$, $U_\varepsilon(x) \subset S$. Для этого установим, что $\forall y \in U_\varepsilon(x)$ следует $y \in S$.

Выберем $\varepsilon = \min\{\varepsilon_1, \varepsilon_2\}$ и рассмотрим $U_\varepsilon(x)$. Пусть $\forall y \in U_\varepsilon(x)$. Построим $y_1 = x_1 + y - x \in U_\varepsilon(x) \subset S$, $y_2 = x_2 + y - x \in U_\varepsilon(x) \subset S$. $(1-\lambda)y_1 + \lambda y_2 \in S$, так как S – выпуклое. Подставим выражения для y_1 и y_2.

$(1-\lambda)y_1 + \lambda y_2 = (1-\lambda)x_1 + \lambda x_2 + y - x = y \in S$. Следовательно, $x \in \overset{\circ}{S}$. Теорема доказана полностью.

Теорема 3.7 Об отрезке, соединяющем точки замыкания и внутренности.

Пусть S – выпуклое множество, $\overset{\circ}{S} \neq \varnothing$. Тогда для $\forall x_1 \in \overline{S}$ и $\forall x_2 \in \overset{\circ}{S}$ $\lambda x_1 + (1-\lambda)x_2 \in \overset{\circ}{S}$ для всех $\lambda \in [0,1)$.

Д о к а з а т е л ь с т в о. Так как $x_2 \in \overset{\circ}{S}$, то $\exists \varepsilon > 0$ такое, что $U_\varepsilon(x_2) \subset S$. Выберем $y = \lambda x_1 + (1-\lambda)x_2$, $\lambda \in (0,1)$ и построим $U_{(1-\lambda)\varepsilon}(y) = \{\tilde{z} \mid \|\tilde{z} - y\| < (1-\lambda)\varepsilon\}$. Покажем, что $U_{(1-\lambda)\varepsilon}(y) \subset S$. Для этого выберем $\forall \tilde{z} \in U_{(1-\lambda)\varepsilon}(y)$ и докажем, что $\tilde{z} \in S$.

Так как $x_1 \in \overline{S}$, то $\forall \tilde{\varepsilon} > 0$ множество $U_{\tilde{\varepsilon}}(x_1) \cap S \neq \varnothing$. Выберем $\tilde{\varepsilon} = \dfrac{(1-\lambda)\varepsilon - \|\tilde{z} - y\|}{\lambda}$, $\lambda \neq 0$, и рассмотрим $z_1 \in S$ и $\|z_1 - x_1\| < \tilde{\varepsilon}$. Пусть $z_2 = \dfrac{\tilde{z} - \lambda z_1}{1-\lambda}$. Тогда

$$\|z_2 - x_2\| = \left\| \frac{\tilde{z} - \lambda z_1}{1-\lambda} - x_2 \right\| = \left\| \frac{\tilde{z} - \lambda z_1 - (y - \lambda x_1)}{1-\lambda} \right\| = \frac{1}{1-\lambda} \|\tilde{z} - y + \lambda(x_1 - z_1)\| \leq$$

$$\leq \frac{1}{1-\lambda}\left(\|\tilde{z} - y\| + \lambda\|x_1 - z_1\|\right) < \frac{1}{1-\lambda}\left((1-\lambda)\varepsilon - \|\tilde{z} - y\| + \|\tilde{z} - y\|\right) = \varepsilon$$

Следовательно, $z_2 \in S$. Получили, что $\tilde{z} = \lambda z_1 + (1-\lambda)z_2$, $z_1 \in S$, $z_2 \in S$, поэтому $\tilde{z} \in S$. Значит, $y \in \overset{\circ}{S}$. Теорема доказана.

Следствие 3.3. Пусть S – выпуклое и $\overset{\circ}{S} \neq \varnothing$. Тогда $\overline{\overset{\circ}{S}} = \overline{S}$, $\overset{\circ}{S} = \overset{\circ}{\overline{S}}$.

Д о к а з а т е л ь с т в о. $\overset{\circ}{S} \subset S$, следовательно, $\overline{\overset{\circ}{S}} = \overline{S}$. Возьмем $\forall x \in \overset{\circ}{S}$ и $\forall y \in \overline{S}$. Тогда по теореме 3.7 $[x, y) \subset \overset{\circ}{S}$ и следовательно $\overline{[x,y)} \subset \overline{\overset{\circ}{S}}$, т.е. $y \in \overline{\overset{\circ}{S}}$.

Имеем $S \subset \overline{S}$, следовательно, $\overset{\circ}{S} \subset \overset{\circ}{\overline{S}}$. Покажем, что $\overset{\circ}{\overline{S}} \subset \overset{\circ}{S}$. Выберем $\forall x_1 \in \overset{\circ}{\overline{S}}$. Тогда $\exists \varepsilon > 0$ такое, что $U_\varepsilon(x_1) \subset \overline{S}$. Пусть $x_2 \in \overset{\circ}{S}, x_2 \neq x_1$.

Возьмем $y = (1+\lambda)x_1 - \lambda x_2$ и $\lambda = \dfrac{\varepsilon}{2\|x_2 - x_1\|}$. Тогда $x_1 = \dfrac{1}{1+\lambda} y + \dfrac{\lambda}{1+\lambda} x_2$.

Найдем $\|y - x_1\| = \|(1+\lambda)x_1 - \lambda x_2 - x_1\| = \lambda\|x_1 - x_2\| = \dfrac{\varepsilon}{2}$. Следовательно, $y \in \overline{S}$.

Таким образом, по теореме 3.7 $x_1 \in \overset{\circ}{S}$. Следствие установлено.

Следствие 3.4. Пусть S_i – выпуклые множества, $i \in I$ и $\bigcap\limits_{i\in I} \overset{\circ}{S}_i \neq \varnothing$. Тогда $\overline{\bigcap\limits_{i\in I} S_i} = \bigcap\limits_{i\in I} \overline{S_i}$.

Д о к а з а т е л ь с т в о. $S_i \subset \overline{S}_i$, следовательно, $\bigcap\limits_{i\in I} S_i \subset \bigcap\limits_{i\in I} \overline{S_i}$ и поэтому $\overline{\bigcap\limits_{i\in I} S_i} \subset \overline{\bigcap\limits_{i\in I} \overline{S_i}} = \bigcap\limits_{i\in I} \overline{S_i}$.

Возьмем $x \in \bigcap\limits_{i\in I} \overset{\circ}{S}_i$, $y \in \bigcap\limits_{i\in I} \overline{S_i}$. Тогда $x \in \overset{\circ}{S}_i$, $y \in \overline{S_i}$, $i \in I$. По теореме 3.7 $[x,y) \subset \overset{\circ}{S}_i$, следовательно, $[x,y) \subset \bigcap\limits_{i\in I} \overset{\circ}{S}_i \subset \bigcap\limits_{i\in I} S_i$ и поэтому $\overline{[x,y)} \subset \overline{\bigcap\limits_{i\in I} S_i}$. Таким образом $y \in \overline{\bigcap\limits_{i\in I} S_i}$. Следствие установлено.

Теорема 3.8. Пусть $S \subset R^n$, $S = \overline{S}$ и для $\forall x, y \in S$, $x \neq y$ существует $\alpha \in (0,1)$ такое, что $(1-\alpha)x + \alpha y \in S$. Тогда множество S – выпуклое.

Д о к а з а т е л ь с т в о. Предположим противное – множество S невыпуклое. Следовательно, существуют точки $x \in S$ и $y \in S$ такие, что $[x, y] \not\subset S$. Обозначим $z(\lambda) = (1-\lambda)x + \lambda y$ и рассмотрим $z(\lambda_0) \notin S$, $\lambda_0 \in (0,1)$.

Пусть $\lambda_1 = \sup\{\lambda | z(\lambda) \in S, 0 < \lambda \le \lambda_0\}$, $\lambda_2 = \inf\{\lambda | z(\lambda) \in S, \lambda_0 \le \lambda < 1\}$. В силу замкнутости множества S, имеем $z(\lambda_1) \in S, z(\lambda_2) \in S$, причем, по построению $\lambda_1 < \lambda_0 < \lambda_2$.

Рассмотрим точки $z(\lambda_1)$ и $z(\lambda_2)$. Это точки множества S, между которыми нет точек этого множества. Получили противоречие предположению, что S – невыпуклое множество. Теорема доказана.

1.4. Разделение выпуклых множеств

Определение 4.1. Точка z называется граничной точкой множества S, если для $\forall \varepsilon > 0$ $\exists x \in S \cap U_\varepsilon(z)$ и $\exists y \notin S$, $y \in U_\varepsilon(z)$.

Определение 4.2. Говорят, что гиперплоскость $Q = \{z | a^T z = \alpha, a \ne 0\}$ разделяет множества S_1 и S_2, если $a^T x \le \alpha$ для $\forall x \in S_1$ и $a^T y \ge \alpha$ для $\forall y \in S_2$.

Определение 4.3. Разделение называется собственным, если $S_1 \cup S_2 \not\subset Q$.

Определение 4.4. Говорят, что гиперплоскость $Q = \{z | a^T z = \alpha, a \ne 0\}$ строго разделяет множества S_1 и S_2, если $a^T x < \alpha$ для $\forall x \in S_1$ и $a^T y > \alpha$ для $\forall y \in S_2$.

Определение 4.5. Говорят, что гиперплоскость $Q = \{z | a^T z = \alpha, a \ne 0\}$ сильно разделяет множества S_1 и S_2, если $\exists \varepsilon > 0$ такое, что $a^T x \le \alpha$ для $\forall x \in S_1$ и $a^T y \ge \alpha + \varepsilon$ для $\forall y \in S_2$.

Определение 4.6. Гиперплоскость Q называется опорной к множеству S в граничной точке z, если $a^T z = \alpha$ и для $\forall x \in S$ $a^T x \le \alpha$ или $a^T x \ge \alpha$.

Определение 4.7. Проекцией точки y на множество S называется такая точка $p \in S$, $p = \pi_S(y)$, что $\|p - y\| = \inf\limits_{x \in S} \|x - y\| = \rho(y, S)$.

$\rho(y, S)$ называется расстоянием от точки y до множества S.

Теорема 4.1. Пусть S выпуклое, $S = \overline{S}$, $y \notin S$.

Тогда существует единственная точка $p \in S$, $p = \pi_S(y)$.

Д о к а з а т е л ь с т в о. Обозначим $\inf\limits_{x \in S} \|y - x\| = \gamma > 0$.

Тогда существует последовательность $\{x_k\} \subset S$, для которой

$$\lim_{k \to \infty} \|x_k - y\| = \gamma . \tag{4.1}$$

Покажем, что $\lim\limits_{k \to \infty} x_k = p$, $p \in S$.

Запишем для точек последовательности

$$\|x_k - x_m\|^2 = 2\|x_k - y\|^2 + 2\|x_m - y\|^2 - \|x_k + x_m - 2y\|^2 =$$

$$= 2\|x_k - y\|^2 + 2\|x_m - y\|^2 - 4\left\|\frac{x_k + x_m}{2} - y\right\|^2 \tag{4.2}$$

В силу выпуклости S, $\frac{x_k + x_m}{2} \in S$, тогда $\left\|\frac{x_k + x_m}{2} - y\right\|^2 \geq \gamma^2$

и из (4.2) вытекает, что

$$\|x_k - x_m\|^2 \leq 2\|x_k - y\|^2 + 2\|x_m - y\|^2 - 4\gamma^2 \tag{4.3}$$

Из (4.1) следует, что для $\forall \varepsilon > 0$ $\exists N$ такой, что для $k, m > N$ справедливо

$\|x_k - y\|^2 - \gamma^2 < \varepsilon/4$ и $\|x_m - y\|^2 - \gamma^2 < \varepsilon/4$.

Тогда получим из (4.3) $\|x_k - x_m\|^2 < \varepsilon$. Отсюда заключаем, что $\lim\limits_{k \to \infty} x_k = p$, $p \in S$, так как S – замкнутое множество.

Покажем теперь, что p – единственная.

Предположим противное, существует $q \neq p$ такая, что $\|q - y\| = \|p - y\| = \gamma$.

В силу выпуклости S, $\frac{p + q}{2} \in S$. Тогда справедливо неравенство

$\left\| y-\frac{p+q}{2} \right\| \le \frac{1}{2}\|y-p\|+\frac{1}{2}\|y-q\|=\gamma$. Но $\left\| y-\frac{p+q}{2} \right\| \ge \gamma$, поскольку $\gamma=\inf_{x\in S}\|y-x\|$.

Следовательно, $\left\| y-\frac{p+q}{2} \right\| = \gamma$, отсюда $y-p=\lambda(y-q)$, $\lambda \in R^1$. Значит,

$\|y-p\|=|\lambda|\cdot\|y-q\|$ и $|\lambda|=1$.

Положим $\lambda=1$, тогда $y-p=y-q$ и $p=q$, а это не так.

Положим теперь $\lambda=-1$, тогда $y-p=-(y-q)$. Отсюда $y=\frac{p+q}{2}\in S$, а по условию y – внешняя точка. Теорема доказана.

Теорема 4.2. Пусть S – выпуклое множество, $S=\bar{S}$. Для того, чтобы $p=\pi_S(y)$, необходимо и достаточно, чтобы для $\forall x\in S$ выполнялось неравенство

$(x-p)^T(y-p)\le 0$.

Доказательство. Достаточность. Возьмем $\forall x\in S$. Рассмотрим выражение

$$\|y-x\|^2=\|y-p+p-x\|^2=\|y-p\|^2+\|p-x\|^2+2(p-x)^T(y-p)\ge\|y-p\|^2.$$

Следовательно, p – ближайшая к y точка из S.

Необходимость. Пусть $p=\pi_S(y)$. Возьмем $\forall x\in S$ и рассмотрим точку $z=\lambda x+(1-\lambda)p$, $\lambda\in(0,1)$. В силу выпуклости S, $z\in S$. Запишем

$$\|z-y\|^2=\|\lambda x+(1-\lambda)p-y\|^2=\|\lambda(x-p)+(p-y)\|^2=$$

$$=\lambda^2\|x-p\|^2+\|p-y\|^2+2\lambda(x-p)^T(p-y)\ge\|p-y\|^2$$

Отсюда $\lambda^2\|x-p\|^2+\|p-y\|^2+2\lambda(x-p)^T(p-y)\ge 0$.

Это неравенство должно быть справедливо для всех $\lambda\in(0,1)$, значит,

$(x-p)^T(y-p)\le 0$.

Теорема доказана .

Теорема 4.3. Пусть S - выпуклое, $S=\overline{S}$, точки $y_1, y_2 \notin S$, $y_1 \neq y_2$. Тогда оператор проектирования $\pi_S(y)$ обладает свойством нерастяжения расстояний, то есть
$\left\|\pi_S(y_1)-\pi_S(y_2)\right\| \le \left\|y_1-y_2\right\|$.

Д о к а з а т е л ь с т в о. Пусть $p_1=\pi_S(y_1)$, $p_2=\pi_S(y_2)$. По теореме 4.2 будем иметь
$(x-p_1)^T(y_1-p_1)\le 0$ и $(x-p_2)^T(y_2-p_2)\le 0$ для $\forall x\in S$.

Запишем первое условие для $x=p_2$, второе для $x=p_1$. Получим
$(p_2-p_1)^T(y_1-p_1)\le 0$ и $(p_1-p_2)^T(y_2-p_2)\le 0$.
Отсюда $(p_2-p_1)^T(y_1-p_1)+(p_1-p_2)^T(y_2-p_2)\le 0$,
следовательно, $(p_1-p_2)^T(y_2-p_2-y_1+p_1)\le 0$.
Отсюда $\left\|p_1-p_2\right\|^2 \le (p_1-p_2)^T(y_1-y_2)\le \left\|p_1-p_2\right\|\cdot\left\|y_1-y_2\right\|$. Значит,
$\left\|p_1-p_2\right\|\le\left\|y_1-y_2\right\|$. Теорема доказана.

Теорема 4.4 (теорема отделимости 1).
Пусть $S\neq\varnothing$, S - выпуклое, $S=\overline{S}$ и $y\notin S$. Тогда существует гиперплоскость $Q=\left\{z\middle|a^Tz=\alpha, a\neq 0, \alpha\in R^1\right\}$ такая, что $a^Ty>\alpha$ и $a^Tx\le\alpha$ для $\forall x\in S$.

Д о к а з а т е л ь с т в о. По теоремам 4.1 и 4.2 существует единственная точка $p\in S$, $p=\pi_S(y)$ и $(x-p)^T(y-p)\le 0$, $\forall x\in S$.

Рассмотрим
$\left\|y-p\right\|^2=(y-p)^T(y-p)=y^T(y-p)-p^T(y-p)\le y^T(y-p)-x^T(y-p)=(y-p)^T(y-x)$.

Обозначим через $a=y-p\neq 0$. Тогда $\left\|y-p\right\|^2\le a^T(y-x)$. Отсюда
$a^Ty\ge a^Tx+\left\|y-p\right\|^2$, $\forall x\in S$.

Пусть $\alpha=\sup a^Tx$, $\forall x\in S$. Тогда $a^Ty>\alpha$ и $a^Tx\le\alpha$, $\forall x\in S$. Теорема доказана.

Теорема 4.5 (теорема отделимости 2).

Пусть $S \neq \varnothing$, S - выпуклое, $S = \overline{S}$ и $y \notin S$. Тогда существует гиперплоскость $Q = \left\{ z \middle| a^T z = \alpha, a \neq 0, \alpha \in R^1 \right\}$ такая, что $a^T y = \alpha$ и $a^T x < \alpha$ для $\forall x \in S$, то есть существует проходящая через y гиперплоскость Q такая, что множество S лежит в одном из полупространств, определяемых Q.

Доказать самостоятельно .

Теорема 4.6 (теорема об опорной гиперлоскости).

Пусть $S \neq \varnothing$, S - выпуклое и $\overline{x} \in \partial S$ ($\overline{x}$ – граничная точка S) . Тогда существует гиперплоскость , опорная к S в $\overline{x}$, то есть $\exists a \neq 0$ и число α такие, что $Q = \left\{ x \middle| a^T x = \alpha, a \neq 0, \alpha \in R^1, \alpha = a^T \overline{x} \right\}$ и $a^T x \leq \alpha$ для всех $x \in \overline{S}$.

Д о к а з а т е л ь с т в о. Поскольку $\overline{x} \in \partial S$, существует последовательность точек $\{y_k\}$, внешних относительно замыкания, $y_k \notin \overline{S}$, таких, что $\lim\limits_{k \to \infty} y_k = \overline{x}$. По теореме 4.5 для каждой точки y_k существует гиперплоскость $Q_k = \left\{ z \middle| {a_k}^T x = \alpha_k, a \neq 0, \alpha \in R^1 \right\}$ такая, что ${a_k}^T y_k = \alpha_k$ и ${a_k}^T x < \alpha_k$ для $\forall x \in S$. Без потери общности можно предположить(см. доказательство теоремы 4.4), что $\|a_k\| = 1$. Из ограниченной последовательности $\{a_k\}$ выделим сходящуюся подпоследовательность $\{a_{k_i}\} \xrightarrow[k_i \to \infty]{} a$. Для элементов этой подпоследовательности справедливо $a^T_{k_i} \cdot y_{k_i} > a^T_{k_i} \cdot x$ при $\forall x \in S$. Фиксируем x и перейдём к пределу при $k_i \to \infty$. Получим $a^T \cdot \overline{x} \geq a^T \cdot x$ при $\forall x \in S$. Выберем $\alpha = a^T \cdot \overline{x}$.Тогда гиперплоскость $Q = \left\{ x \middle| a^T x = \alpha, a \neq 0 \right\}$-искомая опорная гиперплоскость. Теорема доказана.

Теорема 4.7 (обобщенная теорема отделимости).

Пусть $S \neq \varnothing$, S - выпуклое, $S = \overline{S}$ и $y \notin S$. Тогда $\exists a \in R^n$, $a \neq 0$ такой, что $a^T(x - y) \leq 0$ для $\forall x \in \overline{S}$.

Д о к а з а т е л ь с т в о. Пусть $y \notin \overline{S}$, тогда утверждение следует из теоремы отделимости 4.4 для $\overline{S}$ и точки y.

Пусть $y \in \overline{S}$, тогда утверждение следует из теоремы 4.6 об опорной гиперплоскости для $\alpha = a^T y$.

Теорема 4.8 (об отделимости выпуклых множеств).

Пусть S_1 и S_2 выпуклые множества, $\overset{\circ}{S}_2 \neq \varnothing$ и $S_1 \cap \overset{\circ}{S}_2 = \varnothing$. Тогда существует гиперплоскость Q,, разделяющая S_1 и S_2.

Д о к а з а т е л ь с т в о. Рассмотрим множество

$$S = S_1 - \overset{\circ}{S}_2 = \left\{ z \middle| z = x_1 - x_2, x_1 \in S_1, x_2 \in \overset{\circ}{S}_2 \right\}.$$

Имеем $0 \notin S$, тогда по теореме 4.7 $\exists a \in R^n$, $a \neq 0$ такой, что $a^T(z-0) \leq 0$. Следовательно, $a^T x_1 \leq a^T x_2$ для $\forall x_1 \in S_1$ и $\forall x_2 \in S_2$. Теорема доказана.

Теорема 4.9 (о сильной отделимости выпуклых множеств).

Пусть S_1 и S_2 выпуклые множества, $S_1 = \overline{S}_1$, $S_2 = \overline{S}_2$ и S_1-ограниченное. Если $S_1 \cap S_2 = \varnothing$, то существуют $a \in R^n$, $a \neq 0$ и $\varepsilon > 0$ такие, что

$$\inf\left\{a^T y \middle| y \in S_1\right\} \geq \varepsilon + \sup\left\{a^T z \middle| z \in S_2\right\}.$$

Д о к а з а т е л ь с т в о. Рассмотрим множество $S = S_1 - S_2$. Покажем, что $S = \overline{S}$.

Пусть последовательность $\{x_k\}$, $\{x_k\} \subset S$, $\{x_k\} \to x$ при $k \to \infty$. Тогда $x_k = y_k - z_k, y_k \in S_1, z_k \in S_2$. Из ограниченной последовательности $\{y_k\}$ выделим сходящуюся подпоследовательность $\{y_{k_i}\} \xrightarrow[k_i \to \infty]{} y \in S_1$. Тогда $x_{k_j} = y_{k_j} - z_{k_j}, x_{k_j} \to x, y_{k_j} \to y$, следовательно, $z_{k_j} \to z$ при $k_j \to \infty$, причём, $z \in S_2$. Значит, $x \in S$ и $S = \overline{S}$.

По теореме 4.4 существуют $a \in R^n$, $a \neq 0$ и $\varepsilon > 0$ такие, что $a^T x \geq \varepsilon$ $\forall x \in S$ и $a^T 0 < \varepsilon$. Отсюда $a^T (y - z) \geq \varepsilon, \forall y \in S_1, \forall z \in S_2$ и $\inf\left\{a^T y \middle| y \in S_1\right\} \geq \varepsilon + \sup\left\{a^T z \middle| z \in S_2\right\}$.

Пример 1. Множества S_1 и S_2 не имеют общих точек , и $\overline{S_1}$ и $\overline{S_2}$ также не имеют общих точек. Кроме того, S_1 и S_2 ограничены, поэтому их можно сильно разделить гиперплоскостью Q (рис.5)

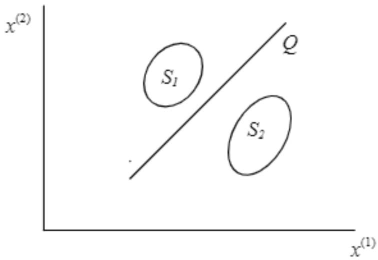

Рис. 5

Пример 2. Рассмотрим два множества $S_1 = \left\{x \in R^2 \middle| x^{(2)} \leq 0\right\}$ и $S_2 = \left\{x \in R^2 \middle| x^{(2)} \geq \frac{1}{x^{(1)}}, x^{(1)} > 0\right\}$.Эти два множества выпуклые замкнутые , не пересекаются. Однако, их можно только разделить гиперплоскостью Q (рис.6).

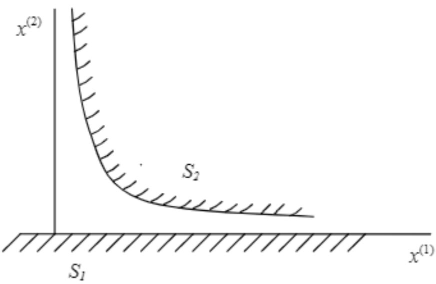

Рис. 6

Определение 4.8. Множество $S_m = \left\{ x \in R^n \middle| x = A^T y, y \in R^m, y \ge 0 \right\}$, где A – матрица размером $m \times n$, называется многогранным конусом.

Лемма 4.1 (о замкнутости многогранного конуса). Многогранный конус S_m - выпуклое замкнутое множество.

Д о к а з а т е л ь с т в о. Выпуклость множества S_m очевидна. Представим $x = A^T y$ в виде

$$x = \sum_{j=1}^{m} y^{(j)} a_j \,,\ y^{(j)} \ge 0 \,,$$

где a_j – строки матрицы A. Очевидно, что S_1 – замкнутое множество. Предположим, что S_k – замкнутое множество и покажем, что тогда и S_{k+1} – замкнутое множество.

Возьмем $\forall z \in \overline{S}_{k+1}$ и покажем, что $z \in S_{k+1}$. Вектор $z \in \overline{S}_{k+1}$, следовательно,

$$z = \lim_{i \to \infty} x_i \,,\ x_i = \sum_{j=1}^{k+1} y_i^{(j)} a_j \,.$$

Обозначим через $\lambda_i = \max y_i^{(j)}$, $1 \le j \le k+1$.

Возможны два случая: 1) $\lambda_i \le const$; 2) λ_i - не ограничено.

В первом случае множества $\left\{ y_i^{(j)} \right\}$, $j=1,\ldots,k+1$ ограничены, поэтому из них можно выбрать сходящиеся последовательности $\left\{ y_{i_1}^{(j)} \right\}$ такие, что $y_{i_1}^{(j)} \to \overline{y}^{(j)} \ge 0$. Тогда

$$z = \lim_{i \to \infty} \sum_{j=1}^{k+1} y_i^{(j)} a_j = \lim_{i_1 \to \infty} \sum_{j=1}^{k+1} y_{i_1}^{(j)} a_j = \sum_{j=1}^{k+1} \overline{y}^{(j)} a_j \,,$$

откуда следует, что $z \in S_{k+1}$.

Во втором случае λ_i не ограничены сверху. Представим z в виде

$$z = \lim_{i \to \infty} \lambda_i \sum_{j=1}^{k+1} \left(y_i^{(j)} / \lambda_i \right) a_j \,.$$

Последовательности $\left\{y_i^{(j)}/\lambda_i\right\}$, $j=1,\ldots,k+1$ ограничены единицей, поэтому из них можно выбрать подпоследовательности $\left\{y_{i_l}^{(j)}/\lambda_{i_l}\right\}$, сходящиеся к α_j.

Преобразуем z к следующему виду

$$z=\lim_{i_l\to\infty}\lambda_{i_l}\left[\sum_{j=1}^{k+1}\alpha_j a_j+\sum_{j=1}^{k+1}\left(y_{i_l}^{(j)}/\lambda_{i_l}-\alpha_j\right)a_j\right]. \qquad (4.4)$$

В правой части (4.4) $\lambda_{i_l}\to\infty$, поэтому $\sum_{j=1}^{k+1}\left(y_{i_l}^{(j)}/\lambda_{i_l}-\alpha_j\right)a_j\to 0$ и, следовательно,

$$\sum_{j=1}^{k+1}\alpha_j a_j=0, \qquad (4.5)$$

причем в (4.5) не все α_j равны нулю.

Воспользуемся равенством (4.5) и представим

$$x_i=\sum_{j=1}^{k+1}\left(y_i^{(j)}-\varepsilon_i\alpha_j\right)a_j, \qquad (4.6)$$

где ε_i определяется соотношением $\varepsilon_i=\min\limits_{1\le j\le k+1}\dfrac{y_i^{(j)}}{\alpha_j}=\dfrac{y_i^{(j_i)}}{\alpha_{j_i}}$.

В представлении (4.6) $y_i^{(j)}-\varepsilon_i\alpha_j\ge 0$, $j=1,\ldots,k+1$, причем по крайней мере один из коэффициентов $y_i^{(j_i)}-\varepsilon_i\alpha_{j_i}=0$.

Элементов x_i – бесконечное число, поэтому найдется такой индекс j_k, для которого $y_i^{(j_k)}-\varepsilon_i\alpha_{j_k}=0$ бесконечное число раз. Выберем из последовательности $\{x_i\}$ подпоследовательность $\{x_{i_k}\}$, для которой $y_{i_k}^{(j_k)}-\varepsilon_{i_k}\alpha_{j_k}=0$. Тогда

$$z=\lim_{i_k\to\infty}x_{i_k}=\lim_{i_k\to\infty}\sum_{j=1,j\ne j_k}^{k+1}\left(y_{i_k}^{(j_k)}-\varepsilon_{i_k}\alpha_{j_k}\right)\cdot a_j \qquad (4.7)$$

В (4.7) $x_{i_k}\in S_k$, S_k- замкнутое множество. Следовательно, $z\in S_k\subset S_{k+1}$. Лемма доказана.

Теорема 4.10 Фаркаша. Пусть A – матрица размером $m\times n$ и $c\in R^n$. Для того, чтобы задача

$$A^T y = c,\ y \ge 0,\ y \in R^m \tag{4.8}$$

имела решение необходимо и достаточно, чтобы

$$\forall x \in S = \left\{ x \in R^n \middle| Ax \le 0 \right\} \text{ было } c^T x \le 0. \tag{4.9}$$

Д о к а з а т е л ь с т в о. Н е о б х о д и м о с т ь. Пусть задача (4.8) имеет решение $\bar{y} \ge 0$. Тогда для $\forall x \in S$

$$\bar{y}^T Ax = x^T A^T \bar{y} = x^T c = c^T x.$$

Так как $\bar{y} \ge 0$, а $Ax \le 0$, то $c^T x \le 0$.

Д о с т а т о ч н о с т ь. Пусть выполнены условия (4.9). Рассмотрим множество

$$V = \left\{ x \in R^n \middle| x = A^T y, y \ge 0 \right\}$$

По лемме 4.1 это множество является выпуклым замкнутым множеством. Если $c \in V$, то теорема доказана. Предположим, что $c \notin V$. Тогда по теореме отделимости 1 $\exists u \ne 0$ такой, что

$$u^T c > \alpha,\ u^T x \le \alpha,\ \forall x \in V.$$

Так как $0 \in V$, то

$$\alpha \ge 0,\ u^T c > 0,\ u^T A^T y = y^T Au \le \alpha. \tag{4.10}$$

Но $y \ge 0$, следовательно, $Au \le 0$. Если бы какой-нибудь элемент $(Au)^{(j)} > 0$, то, положив $y^{(i)} = 0$, $i \ne j$ и устремив $y^{(j)} \to \infty$, можно было получить неравенство

$$y^T Au = y^{(j)} \cdot (Au)^{(j)} > \alpha$$

и нарушить (4.10).

Вектор $u \in S$, и поэтому должно быть $c^T u \le 0$. Но в (4.10) $c^T u > 0$. Это противоречие и доказывает теорему.

Следствие 4.1. Пусть A – матрица размером $m \times n$ и $c \in R^n$. Для того, чтобы задача

$$A^T y \ge c, y \ge 0,\ y \in R^m \tag{4.11}$$

имела решение необходимо и достаточно, чтобы

$$\forall x \in S = \left\{ x \in R^n \middle| x \geq 0, Ax \leq 0 \right\} \text{ было } c^T x \leq 0. \qquad (4.12)$$

Д о к а з а т е л ь с т в о. Введем матрицу $\overline{A}^T = \left(A^T \middle| -E \right)$. Теперь множество S можно представить так: $S = \left\{ x \in R^n \middle| \overline{A}x \leq 0 \right\}$, и тогда по теореме Фаркаша

$$A^T y - Ez = c,\ y \geq 0,\ z \geq 0,\ y \in R^m,\ z \in R^n,$$

откуда вытекает $A^T y \geq c, y \geq 0,\ y \in R^m$.Следствие установлено.

Следствие 4.2. Пусть A и B – матрицы размером $m \times n$ и $k \times n$, соответственно, и $c \in R^n$. Для того, чтобы задача

$$A^T y + B^T z = c, y \geq 0,\ y \in R^m,\ z \in R^n \qquad (4.13)$$

имела решение необходимо и достаточно, чтобы

$$\forall x \in S = \left\{ x \in R^n \middle| Ax \leq 0, Bx = 0 \right\} \text{ следовало } c^T x \leq 0. \qquad (4.14)$$

Д о к а з а т е л ь с т в о. Введем матрицу $\overline{A}^T = \left(A^T \middle| B^T \middle| -B^T \right)$. Тогда множество S можно представить так: $S = \left\{ x \in R^n \middle| \overline{A}x \leq 0 \right\}$, и по теореме Фаркаша

$$A^T y + B^T u - B^T v = c,\ y \geq 0,\ u \geq 0,\ v \geq 0,\ y \in R^m,\ u, v \in R^k,$$

откуда, вводя $z = u - v$, получим

$A^T \cdot y + B^T \cdot z = c, y \geq 0,\ y \in R^m,\ z \in R^k$.Следствие установлено.

1.5. Выпуклые функции

Пусть на непустом выпуклом множестве $S \subset R^n$ задана функция $\varphi : S \to R^1$.

Определение 5.1. Функция $\varphi(x)$ называется выпуклой, если для $\forall x_1, x_2 \in S$ и $0 \leq \lambda \leq 1$ выполняется неравенство

$$\varphi\left[(1-\lambda) x_1 + \lambda x_2 \right] \leq (1-\lambda)\varphi(x_1) + \lambda \varphi(x_2).$$

Определение 5.2. Функция $\varphi(x)$ называется строго выпуклой, если для $\forall x_1, x_2 \in S$, $x_1 \neq x_2$ и $0 < \lambda < 1$ выполняется неравенство

$$\varphi\left[(1-\lambda)x_1 + \lambda x_2\right] < (1-\lambda)\varphi(x_1) + \lambda\varphi(x_2).$$

Определение 5.3. Функция $\varphi(x)$ называется сильно выпуклой, если для $\forall x_1, x_2 \in S$, $0 \leq \lambda \leq 1$ и $\gamma > 0$ выполняется неравенство

$$\varphi\left[(1-\lambda)x_1 + \lambda x_2\right] \leq (1-\lambda)\varphi(x_1) + \lambda\varphi(x_2) - \gamma\lambda(1-\lambda)\|x_1 - x_2\|^2.$$

Определение 5.4. Функция $\varphi(x)$ называется вогнутой, строго вогнутой, сильно вогнутой, если $-\varphi(x)$ является ,соответственно, выпуклой, строго выпуклой, сильно выпуклой функцией.

Дадим геометрическую интерпретацию определения выпуклой и сильно выпуклой функции на примере функции одной переменной $y = \varphi(x)$. График выпуклой функции $y = \varphi(x)$ расположен ниже секущей, соединяющей две точки графика. Для сильно выпуклой функци в области, ограниченной секущей и графиком функции, можно построить параболу, соединяющую эти точки.

Пример 1. Пусть $\varphi(x) = a^T x + \alpha$ - линейная функция. Тогда

$$\varphi\left[(1-\lambda)x_1 + \lambda x_2\right] = a^T\left[(1-\lambda)x_1 + \lambda x_2\right] + \alpha =$$

$$= (1-\lambda)\left(a^T x_1 + \alpha\right) + \lambda\left(a^T x_2 + \alpha\right) = (1-\lambda)\varphi(x_1) + \lambda\varphi(x_2),$$

откуда следует, что линейная функция является одновременно и выпуклой, и вогнутой.

Пример 2. Пусть $\varphi(x) = x^T Ax$ - квадратичная форма, где A – симметричная матрица. Тогда

$$\varphi\left[(1-\lambda)x_1 + \lambda x_2\right] = \left[(1-\lambda)x_1 + \lambda x_2\right]^T A\left[(1-\lambda)x_1 + \lambda x_2\right] =$$

$$= (1-\lambda)^2 x_1^T Ax_1 + 2\lambda(1-\lambda)x_1^T Ax_2 + \lambda^2 x_2^T Ax_2 =$$

$$= (1-\lambda)x_1^T Ax_1 + \lambda x_2^T Ax_2 + \left[(1-\lambda)^2 - (1-\lambda)\right]x_1^T Ax_1 +$$

$$+\left(\lambda^2-\lambda\right)x_2^T Ax_2+2\lambda\left(1-\lambda\right)x_1^T Ax_2=\left(1-\lambda\right)\varphi\left(x_1\right)+\lambda\varphi\left(x_2\right)+$$

$$+2\lambda\left(1-\lambda\right)x_1^T Ax_2-\left(1-\lambda\right)\lambda x_1^T Ax_1-\left(1-\lambda\right)\lambda x_2^T Ax_2=$$

$$=\left(1-\lambda\right)\varphi\left(x_1\right)+\lambda\varphi\left(x_2\right)-\lambda\left(1-\lambda\right)\left(x_1-x_2\right)^T A\left(x_1-x_2\right) \qquad (5.1)$$

Если A – положительно-полуопределеннная матрица, т.е. $y^T Ay\geq 0$ для $\forall y\in R^n$, то из (5.1) следует $\varphi\left[\left(1-\lambda\right)x_1+\lambda x_2\right]\leq\left(1-\lambda\right)\varphi\left(x_1\right)+\lambda\varphi\left(x_2\right)$, т.е. $\varphi(x)$ - выпуклая функция.

Если матрица A – положительно-определеннная, т.е. $y^T Ay\geq\gamma\|y\|^2$, $\forall y\in R^n$, $\gamma>0$, то из (5.1) следует $\varphi\left[\left(1-\lambda\right)x_1+\lambda x_2\right]\leq\left(1-\lambda\right)\varphi\left(x_1\right)+\lambda\varphi\left(x_2\right)-\lambda\left(1-\lambda\right)\gamma\left\|x_1-x_2\right\|^2$ и, следовательно, $\varphi\left(x\right)$ - сильно выпуклая функция.

Теорема 5.1. Пусть $\varphi_1(x)$ и $\varphi_2(x)$ - выпуклые функции, $\alpha_1\geq 0$, $\alpha_2\geq 0$. Тогда $\varphi(x)=\alpha_1\varphi_1(x)+\alpha_2\varphi_2(x)$ - выпуклая функция.

Д о к а з а т е л ь с т в о. Приведем очевидные неравенства

$$\varphi\left[\left(1-\lambda\right)x_1+\lambda x_2\right]=\alpha_1\varphi_1\left[\left(1-\lambda\right)x_1+\lambda x_2\right]+\alpha_2\varphi_2\left[\left(1-\lambda\right)x_1+\lambda x_2\right]\leq$$

$$\leq\alpha_1\left[\left(1-\lambda\right)\varphi_1\left(x_1\right)+\lambda\varphi_1\left(x_2\right)\right]+\alpha_2\left[\left(1-\lambda\right)\varphi_2\left(x_1\right)+\lambda\varphi_2\left(x_2\right)\right]=\left(1-\lambda\right)\varphi\left(x_1\right)+\lambda\varphi\left(x_2\right)$$

Теорема доказана.

Пример 3. Квадратичная функция $\varphi\left(x\right)=\dfrac{1}{2}x^T Ax+b^T x+\alpha$ является выпуклой (сильно выпуклой), если A – положительная (положительно-определеннная) матрица, и вогнутой (сильно вогнутой), если A – отрицательная (отрицательно-определеннная) матрица.

Пример 4 . Недифференцируемая функция. Пусть $\varphi(x)$ - выпуклая функция,заданная на выпуклом множестве S. .Рассмотрим функцию $f(x)=\max\left\{0,\varphi(x)\right\}$. Возьмем $x_1,x_2\in S$ и запишем

$$f\left(\left(1-\lambda\right)x_1+\lambda x_2\right)=\max\left\{0,\varphi\left(\left(1-\lambda\right)x_1+\lambda x_2\right)\right\}\leq\max\left\{0,\left(1-\lambda\right)\varphi\left(x_1\right)+\lambda\varphi\left(x_2\right)\right\}=$$

$$= \frac{\left|(1-\lambda)\varphi(x_1)+\lambda\varphi(x_2)\right| + (1-\lambda)\varphi(x_1)+\lambda\varphi(x_2)}{2} \le \frac{(1-\lambda)(\varphi(x_1)+|\varphi(x_1)|)}{2} +$$

$$+\frac{\lambda(\varphi(x_2)+|\varphi(x_2)|)}{2} = (1-\lambda)\max\{0,\varphi(x_1)\} + \lambda\max\{0,\varphi(x_2)\} = (1-\lambda)f(x_1)+\lambda f(x_2).$$

Следовательно, $f(x)$ - выпуклая функция.

Теорема 5.2. (Неравенство Йенсена). Пусть $\varphi(x)$ - выпуклая функция. Тогда для $\forall x_1,\dots,x_m \in S$ и $\lambda_1 \ge 0,\dots,\lambda_m \ge 0$, $\lambda_1+\dots+\lambda_m = 1$

$$\varphi\left(\sum_{i=1}^{m}\lambda_i x_i\right) \le \sum_{i=1}^{m}\lambda_i\varphi(x_i).$$

Д о к а з а т е л ь с т в о. Для m=1 и m=2 утверждение теоремы очевидно. Предположим, что теорема справедлива для $m=k$ и покажем, что в этом случае утверждение теоремы справедливо и для $m=k+1$. Не умаляя общности можно предположить, что $\lambda_{k+1} \ne 0$.Имеем

$$\varphi\left(\sum_{i=1}^{k+1}\lambda_i x_i\right) = \varphi\left[\left(1-\lambda_{k+1}\right)x + \lambda_{k+1}x_{k+1}\right], \quad x = \sum_{i=1}^{k}\frac{\lambda_i x_i}{1-\lambda_{k+1}}.$$

Но $x \in S$, так как

$$\sum_{i=1}^{k}\frac{\lambda_i}{1-\lambda_{k+1}} = 1, \quad \frac{\lambda_i}{1-\lambda_{k+1}} \ge 0,$$

поэтому

$$\varphi\left[(1-\lambda_{k+1})x + \lambda_{k+1}x_{k+1}\right] \le (1-\lambda_{k+1})\varphi(x) + \lambda_{k+1}\varphi(x_{k+1}) \le$$

$$\le \left(1-\lambda_{k+1}\right)\sum_{i=1}^{k}\frac{\lambda_i\varphi(x_i)}{1-\lambda_{k+1}} + \lambda_{k+1}\varphi\left(x_{k+1}\right) = \sum_{i=1}^{k+1}\lambda_i\varphi\left(x_i\right).$$ Теорема доказана.

Определение 5.5. Множество $\left\{(x,\alpha) \middle| x \in S, \alpha = \varphi(x)\right\}$ называется графиком функции $\varphi(x)$.

Определение 5.6. Множество $epi(\varphi) = \left\{(x,\alpha) \middle| x \in S, \alpha \ge \varphi(x)\right\}$ называется надграфиком функции $\varphi(x)$.

Теорема 5.3. Для того, чтобы функция $\varphi(x)$ была выпуклой, необходимо и достаточно, чтобы множество $epi(\varphi)$ было выпуклым множеством.

Д о к а з а т е л ь с т в о. Н е о б х о д и м о с т ь. Пусть $\varphi(x)$ - выпуклая функция. Возьмем две точки $(x_1,\alpha_1),(x_2,\alpha_2)\in epi(\varphi)$. Для них $\alpha_1 \geq \varphi(x_1)$, $\alpha_2 \geq \varphi(x_2)$ и поэтому

$(1-\lambda)\alpha_1+\lambda\alpha_2 \geq (1-\lambda)\varphi(x_1)+\lambda\,\varphi(x_2)\geq\varphi\left[(1-\lambda)x_1+\lambda x_2\right]$, откуда следует

$((1-\lambda)x_1+\lambda x_2,(1-\lambda)\alpha_1+\lambda\alpha_2)\in epi(\varphi)$.

Д о с т а т о ч н о с т ь. Пусть $epi(\varphi)$ - выпуклое множество. Возьмем $(x_1,\varphi(x_1)),(x_2,\varphi(x_2))\in epi(\varphi)$. Тогда $((1-\lambda)x_1+\lambda x_2,(1-\lambda)\varphi(x_1)+\lambda\varphi(x_2))\in epi(\varphi)$, откуда следует, что $(1-\lambda)\varphi(x_1)+\lambda\varphi(x_2)\geq\varphi((1-\lambda)x_1+\lambda x_2)$. Теорема доказана.

Пример 5. Пусть $\varphi(x)$ - выпуклая функция. Тогда $S=\{x\in R^n \,|\, \varphi(x)\leq\alpha\}$ - выпуклое множество.

Действительно, возьмем $x_1,x_2\in S$. Тогда из очевидных неравенств $\varphi((1-\lambda)x_1+\lambda x_2)\leq(1-\lambda)\varphi(x_1)+\lambda\varphi(x_2)\leq(1-\lambda)\alpha+\lambda\alpha=\alpha$ следует $(1-\lambda)x_1+\lambda x_2\in S$, значит, S-выпуклое множество.

Определение 5.7. Пусть $\varphi(x)$ - функция, заданная на множестве S и $x\in S$. Пусть $z\in R^n$, z- ненулевой элемент, такой, что $x+\lambda z\in S$ при всех $0\leq\lambda\leq\lambda_0>0$, где λ_0 -некоторое число. Производной функции $\varphi(x)$ в точке x по направлению z называется

$$\lim_{\lambda\to+0}\frac{\varphi(x+\lambda z)-\varphi(x)}{\lambda}=\frac{\partial\varphi(x)}{\partial z},\quad \text{если он существует.}$$

Предположим, что функция $\varphi(x)$ - дифференцируема, то есть существует вектор $\nabla\varphi(x)=\left(\varphi_{x^{(1)}}(x)\ \ \dots\ \ \varphi_{x^{(n)}}(x)\right)^T$ - градиент функции. В этом случае $\dfrac{\partial\varphi(x)}{\partial z}=\nabla^T\varphi(x)\cdot z$, где $\nabla\varphi(x)$ - градиент функции $\varphi(x)$, вычисленный в точке x.

Теорема 5.4. Пусть S – выпуклое множество, $\overset{\circ}{S} \neq \varnothing$, а $\varphi(x)$ - выпуклая функция, заданная на S. Тогда в каждой $x \in \overset{\circ}{S}$ существует $\dfrac{\partial\varphi(x)}{\partial z}$ для $\forall z \in R^n$.

Д о к а з а т е л ь с т в о. Пусть $x \in \overset{\circ}{S}$ и $z \in R^n$. Возьмем $-\lambda_0 < 0 < \lambda_1 \leq \lambda_2 \leq \lambda_0$ такие, что $x + \lambda z \in S$ для $\forall \lambda \in [-\lambda_0, \lambda_0]$.

Представим $x + \lambda_1 z$ в виде

$$x + \lambda_1 z = \frac{\lambda_1}{\lambda_2} x + \left(1 - \frac{\lambda_1}{\lambda_2}\right) x + \frac{\lambda_1}{\lambda_2} \lambda_2 z = \frac{\lambda_1}{\lambda_2}(x + \lambda_2 z) + \left(1 - \frac{\lambda_1}{\lambda_2}\right) x.$$

Тогда

$$\varphi(x + \lambda_1 z) = \varphi\left(\frac{\lambda_1}{\lambda_2}(x + \lambda_2 z) + \left(1 - \frac{\lambda_1}{\lambda_2}\right) x\right) \leq \frac{\lambda_1}{\lambda_2}\varphi(x + \lambda_2 z) + \left(1 - \frac{\lambda_1}{\lambda_2}\right)\varphi(x),$$

откуда для $0 < \lambda_1 \leq \lambda_2$ следует неравенство

$$\frac{\varphi(x + \lambda_1 z) - \varphi(x)}{\lambda_1} \leq \frac{\varphi(x + \lambda_2 z) - \varphi(x)}{\lambda_2}, \tag{5.2}$$

то есть $\dfrac{\varphi(x + \lambda z) - \varphi(x)}{\lambda}$ - неубывающая функция λ.

Представим x в виде

$$x = \frac{\lambda_0}{\lambda_0 + \lambda}(x + \lambda z) + \frac{\lambda}{\lambda_0 + \lambda}(x - \lambda_0 z).$$

Тогда

$$\varphi(x) = \varphi\left[\frac{\lambda_0}{\lambda_0 + \lambda}(x + \lambda z) + \frac{\lambda}{\lambda_0 + \lambda}(x - \lambda_0 z)\right] \leq \frac{\lambda_0}{\lambda_0 + \lambda}\varphi(x + \lambda z) + \frac{\lambda}{\lambda_0 + \lambda}\varphi(x - \lambda_0 z),$$

откуда при $\lambda > 0$ следует неравенство

$$\frac{\varphi(x) - \varphi(x - \lambda_0 z)}{\lambda_0} \leq \frac{\varphi(x + \lambda z) - \varphi(x)}{\lambda}, \tag{5.3}$$

то есть функция $\dfrac{\varphi(x + \lambda z) - \varphi(x)}{\lambda}$ ограничена снизу.

Из (5.2) и (5.3) получили,что функция

$$\frac{\varphi(x+\lambda z)-\varphi(x)}{\lambda}$$

является неубывающей функцией λ и ограничена снизу при всех $\lambda > 0$,

поэтому предел $\lim\limits_{\lambda\to+0}\dfrac{\varphi(x+\lambda z)-\varphi(x)}{\lambda}$ существует. Теорема доказана.

Теорема 5.5. Пусть S – выпуклое множество, $\overset{\circ}{S} \neq \varnothing$, а $\varphi(x)$ - выпуклая функция. Тогда в $\forall x_0 \in \overset{\circ}{S}$ функция $\varphi(x)$ непрерывна.

Д о к а з а т е л ь с т в о. Возьмем $\forall x_0 \in \overset{\circ}{S}$. Тогда существует такое $\delta > 0$, что n-мерный куб $\delta\bar{U}_\infty(x_0) \subset S$. (Здесь $\delta\bar{U}_\infty(x_0) = \{x \mid \|x - x_0\|_\infty \le \delta\}$).

Любую точку $x \in \delta\bar{U}_\infty(x_0)$ можно представить в виде выпуклой комбинации его вершин (см. пример 4 п. 1.3)

$$x = x_0 + \delta\sum_{i=1}^{2^n}\lambda_i p_i,$$

где p_i - векторы, компоненты которых равны ± 1.

По теореме 5.4 в точке x_0 существуют производные $\dfrac{\partial\varphi(x_0)}{\partial p_i}$, $i = 1,\dots,2^n$ по всем направлениям p_i поэтому для $\forall\varepsilon > 0$ существуют $\delta_i > 0$ такие, что

$\varphi(x_0 + \delta_i p_i) - \varphi(x_0) \le \varepsilon$.Пусть $\delta = \min\delta_i, i = 1\dots2^n$

Рассмотрим цепочку очевидных равенств и неравенств

$$\varphi(x)-\varphi(x_0) = \varphi\left(x_0 + \delta\sum_{i=1}^{2^n}\lambda_i p_i\right) - \varphi(x_0) = \varphi\left(\sum_{i=1}^{2^n}\lambda_i(x_0 + \delta p_i)\right) - \varphi(x_0) \le$$

$$\le \sum_{i=1}^{2^n}\lambda_i\left(\varphi(x_0 + \delta p_i) - \varphi(x_0)\right) \le \sum_{i=1}^{2^n}\lambda_i\varepsilon = \varepsilon \qquad (5.4)$$

Покажем, что $\varphi(x_0) - \varphi(x) \le \varepsilon$. Возьмем теперь точку $y = x_0 - \delta\sum\limits_{i=1}^{2^n}\lambda_i p_i$, очевидно, $y \in \delta\bar{U}_\infty(x_0)$. Тогда $x_0 = \dfrac{x+y}{2}$ и $\varphi(x_0) \le \dfrac{\varphi(x)+\varphi(y)}{2}$ в силу выпуклости функции $\varphi(x)$.

Отсюда вытекает

$$\varphi(x_0)-\varphi(x)\le\varphi(y)-\varphi(x_0)\le\varepsilon \qquad (5.5)$$

Из неравенств (5.4) и (5.5) следует неравенство $|\varphi(x)-\varphi(x_0)|\le\varepsilon$.

Теорема доказана.

Пример 6. Рассмотрим выпуклую разрывную функцию $\varphi(x)$, Заданную на отрезке $[a,b]$. По теореме 5.5 она может быть разрывной только в точках a и b (рис. 7).

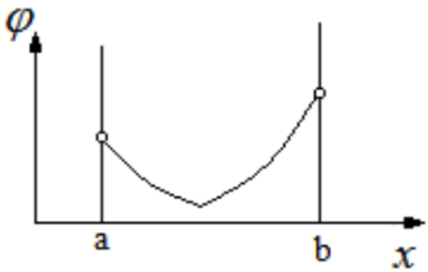

Рис.7

Мы показали, что надграфик выпуклой функции есть выпуклое множество. Поэтому в его граничных точках существуют опорные гиперплоскости, которые приводят к понятию субградиента функции.

Определение 5.8. Пусть $\varphi(x)$ - функция, заданная на множестве S. Вектор $y\in R^n$ называется субградиентом функции $\varphi(x)$ в точке $x_0\in S$, если

$$\varphi(x)\ge\varphi(x_0)+y^T(x-x_0),\ \forall x\in S. \qquad (5.6)$$

Заметим, что $y=y(x_0)$.

Теорема 5.6. Множество субградиентов функции $\varphi(x)$ в каждой точке множества S – выпуклое множество.

Д о к а з а т е л ь с т в о. Пусть y_1 и y_2 - субградиенты $\varphi(x)$ в точке x_0. Тогда для $\forall x\in S$

$$\varphi(x)\ge\varphi(x_0)+y_1^T(x-x_0),\ \varphi(x)\ge\varphi(x_0)+y_2^T(x-x_0),$$

откуда, домножая первое на $1-\lambda$, а второе на λ $(0 \le \lambda \le 1)$ и складывая, получим

$$\varphi(x) \ge \varphi(x_0) + \left[(1-\lambda)y_1 + \lambda y_2\right]^T \cdot (x - x_0), \ \forall x \in S.$$

Следовательно, $(1-\lambda)y_1 + \lambda y_2$ - субградиент функции $\varphi(x)$ в точке x_0.

Теорема 5.7. Пусть S – выпуклое множество, $\overset{\circ}{S} \neq \varnothing$. Для того, чтобы $\varphi(x)$ была выпуклой на $\overset{\circ}{S}$ необходимо и достаточно, чтобы для $\forall x_0 \in \overset{0}{S}$ существовал субградиент.

Д о к а з а т е л ь с т в о. Н е о б х о д и м о с т ь. Пусть $\varphi(x)$ - выпуклая функция. Тогда $epi(\varphi)$ – выпуклое множество. Возьмем $\forall x_0 \in \overset{0}{S}$. Тогда $(x_0, \varphi(x_0)) \in epi(\varphi)$. Эта точка является граничной $epi(\varphi)$, поэтому в ней существует опорная гиперплоскость, такая, что

$$(u, \beta)\begin{pmatrix} x - x_0 \\ \alpha - \varphi(x_0) \end{pmatrix} \le 0$$, причем $(u, \beta) \neq 0$.Тогда будем иметь

$$u^T x + \beta\alpha \le u^T x_0 + \beta\varphi(x_0), \ \forall (x, \alpha) \in epi(\varphi) \tag{5.7}$$

Рассмотрим, какие значения может принимать β.

Пусть $\beta > 0$.Тогда при достаточно больших $\alpha \to \infty$ неравенство (5.7) нарушится.

Предположим, что $\beta = 0$, тогда неравенство (5.7) упрощается

$$u^T x \le u^T x_0 \tag{5.8}$$

Поскольку $x_0 \in \overset{0}{S}$, существует $\lambda > 0$ такое, что $x = x_0 + \lambda u \in S$. Тогда из (5.8)

$\lambda \|u\|^2 \le 0$, откуда следует $u = 0$, что противоречит условию $(u, \beta) \neq 0$.

Следовательно, $\beta < 0$.Разделим (5.7) на $|\beta|$, обозначим через $y = \dfrac{u}{|\beta|}$ и положим $\alpha = \varphi(x)$. Тогда из (5.7), получим

$\varphi(x) \ge \varphi(x_0) + y^T (x - x_0)$. $\forall x_0 \in \overset{0}{S}$.

Д о с т а т о ч н о с т ь. Пусть в каждой точке $x \in \overset{0}{S}$ существует субградиент. Возьмем $\forall x_1, x_2 \in \overset{0}{S}$. Тогда $(1-\lambda)x_1 + \lambda x_2 \in \overset{0}{S}$. В точке $(1-\lambda)x_1 + \lambda x_2$ существует субградиент *y*, так что

$$\varphi(x_1) \ge \varphi\left[(1-\lambda)x_1 + \lambda x_2\right] + y^T \left[x_1 - (1-\lambda)x_1 - \lambda x_2\right] \tag{5.9}$$

$$\varphi(x_2) \ge \varphi\left[(1-\lambda)x_1 + \lambda x_2\right] + y^T \left[x_2 - (1-\lambda)x_1 - \lambda x_2\right] \tag{5.10}$$

Домножая неравенство (5.9) на $1-\lambda$, а (5.10) на λ и складывая, получим

$$(1-\lambda)\varphi(x_1) + \lambda\varphi(x_2) \ge \varphi\left[(1-\lambda)x_1 + \lambda x_2\right],$$

откуда следует, что $\varphi(x)$ - выпуклая функция. Теорема доказана.

Определение 5.9. Множество субградиентов в точке x_0 называется субдифференциалом в этой точке и обозначается $\partial\varphi(x_0)$.

Пример 7. Пусть $\varphi(x) = \|x\|$. Для построения субдифференциала $\varphi(x)$ в точке 0 применим неравенство (5.6) для $\varphi(x)$ в точке $x_0 = 0$. Тогда получим

$$\|x\| \ge y^T \cdot x \tag{5.11}$$

Преобразуем (5.11) к виду

$$\|y\| u^T \cdot v \le 1, \tag{5.12}$$

где $u = y/\|y\|$, $v = x/\|x\|$. Неравенство (5.12) должно выполняться для $\forall v \in R^n$, для которых $\|v\| = 1$. Максимум левой части (5.12) достигается для *v*=*u*. Таким образом, субдифференциал представляет собой множество $\|y\| \le 1$.

1.6. Выпукло-дифференцируемые функции

Определение 6.1. Выпуклая функция $\varphi(x)$ называется выпуклодифференцируемой, если в каждой точке $x \in S$ существует градиент $\nabla\varphi(x)$.

Теорема 6.1. Пусть $\varphi(x)$ – выпукло-дифференцируемая функция, заданная на выпуклом множестве S, причём $\overset{\circ}{S} \neq 0$. Тогда в каждой точке $x_0 \in \overset{\circ}{S}$ субградиент $y = \nabla\varphi(x_0)$.

Д о к а з а т е л ь с т в о. Так как $x_0 \in \overset{\circ}{S}$, то для $\forall z \in R^n$ $\exists \lambda_0 > 0$ такое, что $x_0 + \lambda z \in S,\ 0 \le \lambda \le \lambda_0$.

По теореме 5.7 в точке x_0 существует субградиент $y = y(x_0)$, для которого справедливо неравенство

$$\varphi(x_0 + \lambda z) \ge \varphi(x_0) + \lambda y^T z, \quad 0 \le \lambda \le \lambda_0 \tag{6.1}$$

Разложим в ряд Тейлора в точке x_0 функцию $\varphi(x_0 + \lambda z)$

$$\varphi(x_0 + \lambda z) = \varphi(x_0) + \lambda\nabla^T\varphi(x_0)z + \lambda[\nabla^T\varphi(\bar{x}) - \nabla^T\varphi(x_0)]\cdot z\ , \tag{6.2}$$

где $\bar{x} = x_0 + \nu z, 0 \le \nu \le \lambda$. Подставляя правую часть равенства (6.2) в левую часть неравенства (6.1), получим

$$[\nabla\varphi(x_0) - y]^T \cdot z + [\nabla\varphi(\bar{x}) - \nabla\varphi(x_0)]^T \cdot z \ge 0 \tag{6.3}$$

Из неравенства (6.3) следует, что

$$[\nabla\varphi(x_0) - y]^T \cdot z \ge 0 \tag{6.4}$$

Действительно, предположим, что $[\nabla\varphi(x_0) - y]^T \cdot z < 0$. Тогда устремляя $\lambda \to 0$, получим $\bar{x} \to x_0$. Следовательно, $\nabla\varphi(\bar{x}) \to \nabla\varphi(x_0)$ и неравенство (6.3) будет нарушено. Таким образом, неравенство (6.4) установлено.

Возьмём вектор $z = -(\nabla\varphi(x_0) - y)$ и подставим его в (6.4). Тогда

$$-\|\nabla\varphi(x_0) - y\|^2 \ge 0$$

и, следовательно, $y = \nabla\varphi(x_0)$. Теорема доказана.

Следствие 6.1. Пусть $\varphi(x)$ - дифференцируемая функция, заданная на выпуклом множестве *S*, где $\overset{\circ}{S} \neq \varnothing$. Тогда для того, чтобы $\varphi(x)$ была выпуклой, необходимо и достаточно, чтобы в каждой точке $x_0 \in S$ выполнялось неравенство

$$\varphi(x) \geq \varphi(x_0) + \nabla^T \varphi(x_0)(x - x_0), \forall x \in S \tag{6.5}$$

Д о к а з а т е л ь с т в о. Из теорем 5.7 и 6.1 вытекает, что следствие справедливо для $x_0 \in \overset{\circ}{S}$.

Если $x_0 \in S$, то можно построить $\{x_k\} \subset \overset{\circ}{S}, x_k \to x_0$ и, переходя в неравенстве

$$\varphi(x) \geq \varphi(x_k) + \nabla^T \varphi(x_k)(x - x_k), \forall x \in S$$

к пределу при $k \to \infty$, получить (6.5). Следствие установлено.

Теорема 6.2. Пусть $S \neq \varnothing$, $S = \overset{o}{S}$, S – выпуклое, $\varphi(x)$ - дифференцируемая функция, заданная на *S*. Для того, чтобы $\varphi(x)$ была выпуклой, необходимо и достаточно, чтобы при $\forall x_1, x_2 \in S$ выполнялось неравенство

$$\left(\nabla\varphi(x_2) - \nabla\varphi(x_1)\right)^T (x_2 - x_1) \geq 0.$$

Д о к а з а т е л ь с т в о . Н е о б х о д и м о с т ь. Пусть $\varphi(x)$ - выпуклая функция. Выберем $x_1, x_2 \in S$ и запишем по (6.5)

$$\varphi(x_1) \geq \varphi(x_2) + \nabla^T \varphi(x_1)(x_1 - x_2)$$

$$\varphi(x_2) \geq \varphi(x_1) + \nabla^T \varphi(x_2)(x_2 - x_1).$$

Отсюда

$$\left(\nabla\varphi(x_2) - \nabla\varphi(x_1)\right)^T (x_2 - x_1) \geq 0.$$

Д о с т а т о ч н о с т ь. Выберем $\forall x_1, x_2 \in S$ и запишем по теореме о среднем

$$\varphi(x_2)-\varphi(x_1)=\nabla^T\varphi(x)(x_2-x_1),\, x\in(x_1,x_2), \tag{6.6}$$

иными словами, $x=(1-\lambda)x_1+\lambda x_2$, $\lambda\in(0,1)$. Отсюда $x-x_1=\lambda(x_2-x_1)$.

По условию $\left(\nabla\varphi(x)-\nabla\varphi(x_1)\right)^T\cdot(x-x_1)\ge 0$, значит, $\lambda\left(\nabla\varphi(x)-\nabla\varphi(x_1)\right)^T\cdot(x_2-x_1)\ge 0$. Отсюда $\nabla^T\varphi(x)(x_2-x_1)\ge\nabla^T\varphi(x_1)(x_2-x_1)$. Тогда из (6.6) получим

$\varphi(x_2)-\varphi(x_1)\ge\nabla^T\varphi(x_1)(x_2-x_1)$. По следствию 6.1 заключаем, что $\varphi(x)$ - выпуклая функция.

Определение 6.2. Функция $\varphi(x)$ называется выпуклой дважды дифференцируемой на выпуклом множестве *S*, если она выпукла и имеет непрерывный гессиан $H(x)$.

Теорема 6.3. Пусть S – непустое открытое выпуклое множество, $\varphi(x)$ - дважды дифференцируемая функция. Тогда для того, чтобы $\varphi(x)$ была выпуклой функцией на *S*, необходимо и достаточно, чтобы её гессиан $H(x)$ был положительно-полуопределённой матрицей.

Д о к а з а т е л ь с т в о . Н е о б х о д и м о с т ь . Пусть $\varphi(x)$ - выпуклая функция. Тогда по следствию 6.1 для $\forall x\in S$ и $\forall z\in R^n$ $\exists\lambda_0>0$ такое, что при $0\le\lambda\le\lambda_0$

$$\varphi(x+\lambda z)\ge\varphi(x)+\lambda\nabla\varphi^T(x)\cdot z \tag{6.7}$$

Разложим $\varphi(x+\lambda z)$ в ряд Тейлора (6.8)

$$\varphi(x+\lambda z)=\varphi(x)+\lambda\nabla^T\varphi(x)\cdot z+\lambda^2 z^T H(\bar{x})z/2$$

где $\bar{x}=x_0+\nu z, 0\le\nu\le\lambda$. Подставляя правую часть равенства (6.8) в левую часть (6.7), получим

$z^T H(\bar{x})z\ge 0, \forall z\in R^n$. Здесь $\bar{x}\in S$ так как $\bar{x}=x_0+\nu z,\ 0\le\nu\le\lambda\le\lambda_0$.

Д о с т а т о ч н о с т ь. Пусть $H(x)$ – положительно-полуопределённая матрица Гессе для $\varphi(x)$. По условию теоремы S – непустое открытое множество , поэтому для $\forall x \in S$ и $\forall z \in R^n$ $\exists \lambda_0 > 0$ такое, что $x + \lambda z \in S$ при $0 \le \lambda \le \lambda_0$.

Разложим $\varphi(x + \lambda z)$ в ряд Тейлора

$$\varphi(x + \lambda z) \ge \varphi(x) + \lambda \nabla^T \varphi(x) \cdot z + \lambda^2 z^T H(\bar{x}) z / 2$$

где $\bar{x} = x_0 + \nu z, 0 \le \nu \le \lambda$, $\bar{x} \in S$. Так как матрица $H(x)$ - положительно-полуопределенная, то

$z^T \cdot H(\bar{x}) \cdot z \ge 0$. Таким образом, $\varphi(x + \lambda z) \ge \varphi(x) + \lambda \nabla^T \varphi(x) \cdot z, \forall x \in S$, откуда по теореме 5.7 следует выпуклость $\varphi(x)$. Теорема доказана.

Определение 6.3. Пусть S –выпуклое множество в R^n, $\mathring{S} \ne \varnothing$, $\varphi(x): S \to R^1$, $\varphi(x)$ - дифференцируемая на S. Функция $\varphi(x)$ называется псевдовыпуклой, если для $\forall\, x_1, x_2 \in S$ таких, что $\nabla^T \varphi(x_1)(x_2 - x_1) \ge 0$ выполнено $\varphi(x_2) \ge \varphi(x_1)$.

По доказанному ранее, непрерывно дифференцируемая выпуклая функция является псевдовыпуклой.

Определение 6.4. Пусть S – выпуклое множество в R^n, $\mathring{S} \ne \varnothing$, $\varphi(x): S \to R^1$. $\varphi(x)$ называется квазивыпуклой, если для $\forall\, x_1, x_2 \in S$ и $\forall \lambda \in [0,1]$ справедливо $\varphi[(1-\lambda)x_1 + \lambda x_2] \le \max\{\varphi(x_1), \varphi(x_2)\}$.

Теорема 6.4. Пусть S – выпуклое множество в R^n, $\mathring{S} \ne \varnothing$, $\varphi(x): S \to R^1$. Для того, чтобы $\varphi(x)$ была квазивыпуклой, необходимо и достаточно, чтобы для $\forall\, \gamma \in R^1$ множество $S_\gamma = \{x \in S \,|\, \varphi(x) \le \gamma\}$ было выпуклым.

Д о к а з а т е л ь с т в о . Н е о б х о д и м о с т ь . Пусть $\varphi(x)$ - квазивыпуклая функция.Тогда для $\forall\, x_1, x_2 \in S_\gamma$

$\varphi(x_1) \le \gamma,$
$\varphi(x_2) \le \gamma$.

Отсюда для $\forall \lambda \in [0,1]$ будет выполнено

$\varphi[(1-\lambda)x_1 + \lambda x_2] \le \max\{\varphi(x_1), \varphi(x_2)\} \le \gamma$. Следовательно, $(1-\lambda)x_1 + \lambda x_2 \in S_\gamma$ и S_γ -выпуклое множество.

Д о с т а т о ч н о с т ь. Выберем $\forall\, x_1, x_2 \in S_\gamma$ и пусть , для определённости, $\varphi(x_1) \ge \varphi(x_2)$. Обозначим $\gamma = \varphi(x_1)$.

S_γ - выпуклое , поэтому для $\forall \lambda \in [0,1]$ $(1-\lambda)x_1 + \lambda x_2] \in S_\gamma$. Отсюда $\varphi[(1-\lambda)x_1 + \lambda x_2] \le \gamma = \varphi(x_1) = \max\{\varphi(x_1), \varphi(x_2)\}$. Теорема доказана.

II. ЗАДАЧИ ВЫПУКЛОГО ПРОГРАММИРОВАНИЯ

В этой главе рассматриваются постановки задач выпуклого программирования, теория необходимых и достаточных условий, теория двойственности. Особое внимание уделено применению теории выпуклого программирования к задачам линейного программирования. Приведен симплекс- метод решения задач линейного программирования.

2.1. Постановки задач выпуклого программирования

Пусть определено множество $S \subset R^n$, на котором задана целевая функция $\varphi_0(x)$. Решение задачи минимизация целевой функции $\varphi_0(x), \forall x \in S$ заключается в следующем:

1) найти $x_* \in S$ такую, что $\varphi_0(x_*) \le \varphi_0(x), \forall x \in S$;

2) либо найти $\varphi_{0*} = \inf \varphi_0(x), \forall x \in S, x_* \notin S$;

3) либо доказать, что $\varphi_0(x)$ неограничена на S снизу;

4) либо показать, что $S = \varnothing$.

Определение 1.1. Функция $\varphi_0(x)$ называется целевой функцией или функцией цели, а множество S, на котором разыскивается минимум целевой функции –допустимым или множеством допустимых точек.

Определение 1.2. Точка $x_* \in S$ называется оптимальной точкой (оптимальным решением) или точкой глобального минимума.

Рассмотрим множество $S = \{x \in R^n \mid \varphi_i(x) \ge 0, i = 1,...,m\}$, где $\varphi_i(x)$ - заданные скалярные функции. Задачу минимизации функции $\varphi_0(x)$ на множестве S будем называть основной задачей математического программирования.

Определение 1.3.Неравенства $\varphi_i(x) \ge 0$, определяющие допустимое множество, называются ограничениями задачи.

Если множество S – выпуклое, $\varphi_0(x)$ - выпуклая, то задачу называют задачей выпуклого программирования.

Из теоремы 3.1 и примера 5 (гл.I) вытекает, что для выпуклости множества S достаточно, чтобы функции $\varphi_i(x)$были вогнутыми.

Если в задаче выпуклого программирования все функции $\varphi_i(x)$ вогнуты, а функция $\varphi_0(x)$ - выпуклая, то задачу будем называть основной задачей выпуклого программирования.

Далее будут рассматриваться только задачи минимизации.

Вопрос существования оптимального решения для задач математического программирования базируется на известной теореме Вейерштрасса.

Теорема 1.1.(Вейерштрасса). Непрерывная функция достигает минимума и максимума на компактном множестве.

Доказательство можно найти, например, в [4].

Определение 1.4. Точка $x_* \in S$ сообщает локальный минимум функции $\varphi_0(x)$, если $\exists \varepsilon > 0$ такое, что $\varphi_0(x_*) \le \varphi_0(x), \forall x \in U(x_*, \varepsilon) \bigcap S$.

Рассмотрим задачу минимизации

$$\min \varphi_0(x), \forall x \in S, \tag{1.1}$$

где S – выпуклое множество, а $\varphi_0(x)$ - выпуклая функция.

Теорема 1.2. Точка x_*, в которой достигается локальный минимум в задаче (1.1), является точкой глобального минимума.

Д о к а з а т е л ь с т в о. Пусть x_* - точка локального минимума. Тогда $\exists \varepsilon > 0$ такое, что для $\forall x \in U(x_*, \varepsilon) \bigcap S$ справедливо неравенство $\varphi_0(x_*) \le \varphi_0(x)$.

Предположим, что точка x_* не является точкой глобального минимума, значит, $\exists y_* \in S$, для которого $\varphi_0(y_*) < \varphi_0(x_*)$. Выберем такое λ $(0 < \lambda < 1)$, чтобы $x = (1-\lambda)x_* + \lambda y_* \in U(x_*, \varepsilon)$.

Запишем $x - x_* = \lambda(y_* - x_*)$, отсюда $\|x - x_*\| = \lambda \|y_* - x_*\| < \varepsilon$.Положим $\lambda < \dfrac{\varepsilon}{\|y_* - x_*\|}$.Очевидно, что $\lambda < 1$, поскольку x_* - точка локального минимума и , значит, $y_* \notin U(x_*, \varepsilon)$.

Тогда $\varphi_0[(1-\lambda)x_* + \lambda y_*] \le (1-\lambda)\varphi_0(x_*) + \lambda\varphi_0(y_*) < (1-\lambda)\varphi_0(x_*) + \lambda\varphi_0(x_*) = \varphi_0(x_*)$, откуда следует, что $\varphi_0(x_*)$ не локальный минимум $\varphi_0(x)$. Теорема доказана.

Из этой теоремы следует, что все локальные минимумы явдяются глобальными в задаче выпуклого программирования, поэтому для решения задачи достаточно найти какую-либо точку локального минимума.

Теорема 1.3. Множество оптимальных точек $S_* \subset S$ в задаче выпуклого программирования является выпуклым.

Д о к а з а т е л ь с т в о. Пусть $x_*, y_* \in S_*$.Рассмотрим

$x = (1-\lambda)x_* + \lambda y_*, 0 \le \lambda \le 1$ и запишем для $\varphi_0(x)$:

$\varphi_0(x) \le (1-\lambda)\varphi_0(x_*) + \lambda\varphi_0(y_*) = \varphi_0(x_*),$ так как $\varphi_0(x_*) = \varphi_0(y_*)$. Теорема доказана.

Пример 1.

На рисунке 8 приведена выпуклая функция $\varphi(x)$, заданная на отрезке. Видно, что оптимальная точка не единственная, а множество S_* - выпуклое

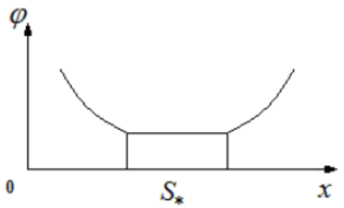

Рис.8

Теорема 1.4. Для строго выпуклой функции цели множество S_* не может иметь более одной точки.

Д о к а з а т е л ь с т в о . Предположим, что $\exists x_*, y_* \in S_*$. Тогда

$\varphi_0(\frac{x_*}{2} + \frac{y_*}{2}) < \frac{1}{2}\varphi_0(x_*) + \frac{1}{2}\varphi_0(y_*) = \varphi_0(x_*)$, что противоречит оптимальности x_* и y_*. Теорема доказана.

Обобщим доказанные утверждения .

Теорема 1.5. Пусть S – непустое выпуклое множество в R^n. Рассматривается задача $\min \varphi_0(x), \forall x \in S$. Пусть $\bar{x}$ - локальное оптимальное решение (точка локального минимума). Тогда

1) если $\varphi_0(x)$ - выпуклая функция, то $\bar{x}$ - глобальное оптимальное решение;

2) если $\varphi_0(x)$ - строго выпуклая функция, то $\bar{x}$ - единственное глобальное оптимальное решение.

Теорема 1.6. Пусть в задаче (1.1) $\varphi_0(x)$ - псевдовыпуклая функция. Тогда точка x_*, в которой достигается локальный минимум , является точкой глобального минимума.

Д о к а з а т е л ь с т в о. Пусть x_* - точка локального минимума. Тогда $\exists\varepsilon > 0$ и для

$\forall x \in U(x_*, \varepsilon) \bigcap S$ будет $\varphi_0(x_*) \le \varphi_0(x)$.

Предположим, что точка x_* не является точкой глобального минимума, значит, $\exists y_* \in S$, для которого $\varphi_0(y_*) < \varphi_0(x_*)$. Выберем такое λ $(0 < \lambda < 1)$, чтобы $x = (1-\lambda)x_* + \lambda y_* \in U(x_*, \varepsilon)$.

Запишем $x - x_* = \lambda(y_* - x_*)$, отсюда $\|x - x_*\| = \lambda\|y_* - x_*\| < \varepsilon$.Положим $\lambda < \dfrac{\varepsilon}{\|y_* - x_*\|}$.Очевидно, что $\lambda < 1$, поскольку x_* - точка локального минимума и , значит, $y_* \notin U(x_*, \varepsilon)$.

Так как $\varphi_0(y_*) < \varphi_0(x_*)$, то $\nabla^T\varphi_0(x_*)(y_* - x_*) < 0$ (в противном случае $\varphi_0(y_*) \ge \varphi_0(x_*)$).

Функция $\varphi_0(x)$ - псевдовыпуклая, следовательно, дифференцируемая, и можно записать $\varphi_0(x) = \varphi_0(x_*) + \nabla^T\varphi_0(x_*)(x - x_*) + \|x - x_*\|\alpha(x_*; x - x_*)$,

где $\alpha(x_*; x - x_*) \to 0$ при $x \to x_*$. Отсюда

$\varphi_0(x) - \varphi_0(x_*) = \lambda\nabla^T\varphi_0(x_*)(y_* - x_*) + \lambda\|y_* - x_*\|\alpha(x_*; \lambda(y_* - x_*))$ и при малых λ получим $\varphi_0(x) < \varphi_0(x_*)$, откуда следует, что $\varphi_0(x_*)$ не локальный минимум $\varphi_0(x)$. Теорема доказана.

Для квазивыпуклых функций аналогичная теорема неверна.Рассмотрим следующий пример.

Пример 2. Пусть $\varphi_0(x) = sign(x), x \in [-2, 2]$. Точка $x_* = 1$ является точкой локального, но не глобального минимума.

Определение 1.5. Пусть S – непустое выпуклое множество в R^n, $\varphi(x): S \to R^1$. $\varphi(x)$ называется строго квазивыпуклой, если для $\forall\, x_1, x_2 \in S$ таких, что $\varphi(x_1) \neq \varphi(x_2)$ для $\forall \lambda \in (0,1)$ справедливо $\varphi[(1-\lambda)x_1 + \lambda x_2] < \max\{\varphi(x_1), \varphi(x_2)\}$.

Очевидно, что выпуклая функция является строго квазивыпуклой.

Теорема 1.7. Пусть в задаче (1.1) $\varphi_0(x)$ - строго квазивыпуклая функция. Тогда точка x_*, в которой достигается локальный минимум , является точкой глобального минимума.

Д о к а з а т е л ь с т в о. Пусть x_* - точка локального минимума. Тогда $\exists \varepsilon > 0$ и для
$\forall x \in U(x_*, \varepsilon) \bigcap S$ будет $\varphi_0(x_*) \le \varphi_0(x)$.

Предположим, что точка x_* не является точкой глобального минимума, значит, $\exists y_* \in S$, для которого $\varphi_0(y_*) < \varphi_0(x_*)$. Выберем такое λ $(0 < \lambda < 1)$, чтобы $x = (1-\lambda)x_* + \lambda y_* \in U(x_*, \varepsilon)$.

Запишем $x - x_* = \lambda(y_* - x_*)$, отсюда $\|x - x_*\| = \lambda \|y_* - x_*\| < \varepsilon$. Положим $\lambda < \dfrac{\varepsilon}{\|y_* - x_*\|}$.Очевидно, что $\lambda < 1$, поскольку x_* - точка локального минимума и , значит, $y_* \notin U(x_*, \varepsilon)$. Тогда $\varphi_0[(1-\lambda)x_* + \lambda y_*] < \max\{\varphi_0(x_*), \varphi_0(y_*)\} = \varphi_0(x_*)$, откуда следует, что $\varphi_0(x_*)$ не локальный минимум $\varphi_0(x)$. Теорема доказана.

2.2. Необходимые и достаточные условия

Следующая теорема является наиболее общей теоремой о необходимых и достаточных условиях минимума в общей задаче выпуклого программирования (1.1) . Назовём её задачей I.

Теорема 2.1. В задаче I выпуклого программирования $x_* \in S$ будет оптимальной точкой тогда и только тогда, когда в ней существует субградиент $y = y(x_*)$, для которого выполняется условие

$$y^T(x-x_*)\ge 0, \forall x\in S \tag{2.1}$$

Д о к а з а т е л ь с т в о . Д о с т а т о ч н о с т ь. Пусть в точке x_* существует субградиент y, удовлетворяющий условию (2.1.) Функция $\varphi_0(x)$ - выпуклая, поэтому для нее $\varphi_0(x)\ge\varphi_0(x_*)+y^T(x-x_*), \forall x\in S.$ Отсюда, с учётом (2.1), получаем

$$\varphi_0(x)\ge\varphi_0(x_*)\,, \forall x\in S\,.$$

Н е о б х о д и м о с т ь. Пусть x_* - оптимальное решение задачи I. Построим два множества

$$V=\{(x-x_*,\alpha)^T\in R^{n+1}\mid x\in R^n, \alpha>\varphi_0(x)-\varphi_0(x_*)\},$$
$$W=\{(x-x_*,\alpha)^T\in R^{n+1}\mid x\in S, \alpha\le 0\}$$

Множества V и W – выпуклые, так как S – выпуклое, а $\alpha>\varphi_0(x)-\varphi_0(x_*)$ и $\alpha\le 0$ - полупространства. Очевидно, что $V\bigcap W=\varnothing$, иначе нашлась бы точка $(x-x_*,\alpha)^T$ такая, что $x\in S$ и $0\ge\alpha>\varphi_0(x)-\varphi_0(x_*)$.Поэтому V и W можно разделить гиперплоскостью, определяемой вектором $(u,\beta)^T\ne 0$, $u\in R^n, \beta\in R^1$ и $\gamma\in R^1$

$$(u,\beta)^T(x-x_*,\alpha)=u^T(x-x_*)+\beta\alpha\le\gamma, x\in R^n, \alpha>\varphi_0(x)-\varphi_0(x_*) \tag{2.2}$$

$$(u,\beta)^T(x-x_*,\alpha)=u^T(x-x_*)+\beta\alpha\ge\gamma, x\in S, \alpha\le 0 \tag{2.3}$$

Подставляя в неравенство (2.2) $x=x_*, \alpha=\varepsilon>0$, в (2.3) $x=x_*, \alpha=0$, найдем

$$\beta\varepsilon\le\gamma\le 0\,. \tag{2.4}$$

Из неравенства (2.4) следует $\beta\le 0$ и $\gamma=0$, так как в нем ε может быть сколь угодно мало.

Покажем, что $\beta<0$. Предположим, что $\beta=0$. Тогда неравенство (2.2) будет иметь вид

$$u^T(x-x_*)\le 0, \forall x\in R^n \tag{2.5}$$

Положим в (2.5) $x=x_*+u$. Тогда получим неравенство $\|u\|^2\le 0$, откуда следует $u=0$, что противоречит условию $(u,\beta)^T\ne 0$.

Поделив неравенство (2.2) на $-\beta>0$ и учитывая $\gamma=0$, получим

$\varphi_0(x) \ge \varphi_0(x_*) + y^T(x - x_*), \forall x \in S,$ где $y = -u / \beta$ - является субградиентом функции $\varphi_0(x)$ в точке x_*.

Поделив неравенство (2.3) на $-\beta > 0$ и учитывая $\gamma = 0$, получим для $\alpha = 0$ неравенство (2.1). Теорема доказана.

Следствие 2.1. Пусть в задаче I $S = \overset{o}{S}$. Точка x_* будет оптимальной тогда и только тогда, когда в ней существует субградиент $y = y(x_*) = 0$.

Д о к а з а т е ль с т в о. Н е о б х о д и м о с т ь. Повторяя ход доказательства теоремы 2.1, получим условие $y^T(x - x_*) \ge 0, \forall x \in S$. Положим $x = x_* - \varepsilon y,$ где $\varepsilon > 0$ выберем таким, чтобы $x \in S$. Тогда $-\varepsilon \| y \|^2 \ge 0$, откуда получим $y = 0$.

Д о с т а т о ч н о с т ь . Доказательство очевидно. Следствие установлено.

Следствие 2.2. Пусть в задаче I функция $\varphi_0(x)$ - дифференцируемая. Точка $x_* \in S$ будет оптимальной тогда и только тогда, когда выполняется неравенство

$$\nabla^T \varphi(x_*)(x - x_*) \ge 0, \forall x \in S. \qquad (2.6)$$

Доказательство следует из теоремы 6.1 гл.I и теоремы 2.1.

Следствие 2.3. Пусть в задаче I функция $\varphi_0(x)$ - дифференцируемая, $S = \overset{o}{S}$. Точка $x_* \in S$ будет оптимальной тогда и только тогда, когда $\nabla \varphi(x_*) = 0$.

Доказательство следует из теоремы 6.1 гл.I следствия 2.1.

Пример 1. Пусть в задаче I множество $S = \{x \mid x \ge 0\}$ и $\varphi_0(x)$ - дифференцируемая.

Необходимым и достаточным условием минимума является неравенство (2.6), которое в развернутом виде будет иметь вид

$$\sum_{i=1}^{n} \frac{\partial \varphi_0(x_*)}{\partial x^{(i)}}(x^{(i)} - x_*^{(i)}) \ge 0, \forall x^{(i)} \ge 0, \qquad (2.7)$$

причём $x_*^{(i)}$ - может быть равно нулю или принимать положительное значение.

Если $x_*^{(j)} = 0$, то $x^{(j)} - x_*^{(j)} \geq 0$ и эта разность может быть сколь угодно большой. Положим $x^{(i)} - x_*^{(i)} = 0, i \neq j$. Тогда из неравенства (2.7) следует, что $\frac{\partial \varphi_0(x_*)}{\partial x^{(j)}} \geq 0$.

Если $x_*^{(j)} > 0$, то $x^{(j)} - x_*^{(j)}$ может быть как отрицательным, так и положительным. Опять, полагая $x^{(i)} - x_*^{(i)} = 0, i \neq j$, получим $\frac{\partial \varphi_0(x_*)}{\partial x^{(j)}} = 0$.

Таким образом, необходимыми и достаточными условиями минимума являются

$$\frac{\partial \varphi_0(x_*)}{\partial x^{(i)}} = \begin{cases} \geq 0, x_*^{(i)} = 0, \\ = 0, x_*^{(i)} > 0 \end{cases}$$

Рассмотрим задачу минимизации

$$\min \varphi_0(x), \;\; \varphi_i(x) \leq 0, \;\; i = 1, ..., m, \;\; x \in V_0 \;, \tag{2.8}$$

где V_0 - выпуклое множество, а $\varphi_0(x), \varphi_1(x), ..., \varphi_m(x)$ -выпуклые функции. Назовём эту задачу задачей II.

Множество допустимых точек $S = \{x \in V_0 \mid \varphi_i(x) \leq 0, i = 1, ..., m\}$ является выпуклым, как пересечение выпуклых множеств $V_0, V_1, ..., V_m$,где

$$V_i = \{x \in R^n \mid \varphi_i(x) \leq 0\}, i = 1, ..., m .$$

Обозначим через M множество чисел $\{1, ..., m\}$. Возьмем $\forall x \in S$ и построим два множества

$$M_0(x) = \{i \in M \mid \varphi_i(x) = 0\}, \;\; M_1(x) = \{i \in M \mid \varphi_i(x) < 0\} .$$

Говорят, что для $i \in M_0(x)$ ограничения $\varphi_i(x)$ выполняются активно, а для $i \in M_1(x)$ - пассивно.

Определение 2.1. Функция $\psi(x, y) = \varphi_0(x) + \sum_{i=1}^{m} y^{(i)} \varphi_i(x), \forall x \in V_0, \forall y \geq 0$ называется функцией Лагранжа, а y – вектором множителей Лагранжа.

Если ввести вектор $\varphi^T = (\varphi_1, ..., \varphi_m)$, то функцию Лагранжа можно записать в виде

$$\psi(x, y) = \varphi_0(x) + y^T \varphi(x), \forall x \in V_0, \forall y \geq 0 .$$

Функция Лагранжа является выпуклой по x, так как $y \geq 0$ (см. теорему 5.1 гл I.).

Определение 2.2. Говорят, что множество S удовлетворяет условию Слейтера, если $\exists \bar{x} \in S$ такая, что $\varphi(\bar{x}) < 0$.

Определение 2.3. Говорят, что функция $\psi(x, y)$ имеет седловую точку (x_*, y_*), $x_* \in V_0, y_* \geq 0$, если

$$\psi(x_*, y) \leq \psi(x_*, y_*) \leq \psi(x, y_*) \quad \forall x \in V_0, \forall y \geq 0 \tag{2.9}$$

Теорема 2.2 Куна-Таккера. Пусть в задаче II (2.8) множество S удовлетворяет условию Слейтера. Тогда для того, чтобы x_* была оптимальной точкой, необходимо и достаточно, чтобы $\exists y_* \geq 0$ такой, чтобы (x_*, y_*) была седловой точкой функции Лагранжа $\psi(x, y)$.

Д о к а з а т е л ь с т в о. Д о с т а т о ч н о с т ь . Пусть выполнено условие Слейтера и $\exists y_* \geq 0$ такой, что (x_*, y_*) является седловой точкой функции $\psi(x, y)$.

Рассмотрим левое неравенство (2.9)

$$\varphi_0(x_*) + y^T \varphi(x_*) \leq \varphi_0(x_*) + y_*^T \varphi(x_*), \forall y \geq 0 , \tag{2.10}$$

откуда следует, что $\varphi(x_*) \leq 0$. Действительно, предположим, что $\varphi_j(x_*) > 0$ для некоторого $j \in M$. Устремляя $y^{(j)} \to \infty$, придем к нарушению неравенства (2.10).

Полагая $y = 0$, получим $y_*^T \varphi(x_*) \geq 0$. С другой стороны, $\varphi(x_*) \leq 0$, а $y_* \geq 0$, поэтому $y_*^T \varphi(x_*) \leq 0$, откуда и из предыдущего неравенства следует, что $y_*^T \varphi(x_*) = 0$.

Рассмотрим правое неравенство (2.9)

$$\varphi_0(x_*) + y_*^T \varphi(x_*) \le \varphi_0(x) + y_*^T \varphi(x), \forall x \in V_0 \qquad (2.11)$$

Так как $y_*^T \cdot \varphi(x_*) = 0$, то $\varphi_0(x_*) \le \varphi_0(x) + y_*^T \varphi(x) \le \varphi_0(x)$ для $\forall x \in S$, т.е. x_* - оптимальное решение задачи II.

Н е о б х о д и м о с т ь. Пусть x_*- оптимальное решение задачи II. Покажем выполнение неравенства седловой точки. Построим два множества

$$Z = \bigcup_{x \in V_0} Z_x, Z_x = \{(z,\alpha)^T \in R^{m+1} \mid z \ge \varphi(x), \alpha \ge \varphi_0(x)\}$$

$$W = \{(v,\beta)^T \in R^{m+1} \mid v \le 0, \beta < \varphi_0(x_*)\}.$$

Покажем, что множество Z – выпуклое. Возьмём $\forall (z_1,\alpha_1)^T, (z_2,\alpha_2)^T \in Z$. Тогда $(z_1,\alpha_1)^T \in Z_{x_1}, (z_2,\alpha_2)^T \in Z_{x_2}$ и, следовательно,

$z_1 \ge \varphi(x_1), \alpha_1 \ge \varphi_0(x_1), z_2 \ge \varphi(x_2), \alpha_2 \ge \varphi_0(x_2)$.

Рассматривая выпуклые комбинации

$$\begin{aligned} &(1-\lambda)z_1 + \lambda z_2 \ge (1-\lambda)\varphi(x_1) + \lambda\varphi(x_2) \ge \varphi[(1-\lambda)x_1 + \lambda x_2], \\ &(1-\lambda)\alpha_1 + \lambda\alpha_2 \ge (1-\lambda)\varphi_0(x_1) + \lambda\varphi_0(x_2) \ge \varphi_0[(1-\lambda)x_1 + \lambda x_2], 0 \le \lambda \le 1, \end{aligned}$$

получим, что

$$(1-\lambda)(z_1,\alpha_1)^T + \lambda(z_2,\alpha_2)^T \in Z_{(1-\lambda)x_1+\lambda x_2} \subset Z$$

Очевидно, что множество W - выпуклое.

Покажем, что $Z \cap W = \varnothing$.

Для $\forall x \in S, \varphi(x) \le 0$ и $\varphi_0(x_*) \le \varphi_0(x)$, т.е. $\alpha > \beta$.

Для $\forall x \in V_0 \setminus S$ существует индекс $j \in M$ такой, что $\varphi_j(x) > 0$ и , следовательно, $z^{(j)} > 0$. Но соответствующее $v^{(j)} \le 0$, значит, $v^{(j)} < z^{(j)}$.

Таким образом, $Z \cap W = \varnothing$ и существует вектор $(u,\gamma)^T \ne 0$, определяющий гиперплоскость, разделяющую множества Z и W, т.е.

$$u^T z + \gamma\alpha \ge u^T v + \gamma\beta, \ \forall (z,\alpha)^T \in Z, \ \forall (v,\beta)^T \in W \qquad (2.12)$$

Покажем, что $u \ge 0, \gamma \ge 0$.   Так как $v \le 0, \beta < \varphi_0(x_*)$, то из (2.12) следует, что $u \ge 0, \gamma \ge 0$. Действительно, если $u^{(j)} < 0$ или $\gamma < 0$, то устремляя $v^{(j)} \to -\infty$ или $\beta \to -\infty$, получим нарушение (2.12).

Покажем, что $\gamma > 0$. Предположим, что $\gamma = 0$. Подставляя в неравенство (2.12) $\gamma = 0, v = 0$, $z = \varphi(\bar{x})$, где $\bar{x}$ – точка Слейтера (из условия Слейтера), получим $u^T \varphi(\bar{x}) \geq 0$. Это неравенство может выполняться только при $u = 0$ и тогда $(u, \gamma)^T = 0$. Значит, предположение $\gamma = 0$ неверно. Таким образом, $\gamma > 0$.

Обозначая $u / \gamma = y_* \geq 0$ и полагая в (2.12) $v = 0, z = \varphi(x), \alpha = \varphi_0(x)$, получим $\beta \leq \varphi_0(x) + y_*^T \varphi(x), \forall x \in V_0$. Но по определению $\beta < \varphi_0(x_*)$, значит, должно быть

$$\varphi_0(x_*) \leq \varphi_0(x) + y_*^T \varphi(x), \forall x \in V_0 \tag{2.13}$$

Для $x = x_*$ неравенство (2.13) принимает вид $y_*^T \varphi(x_*) \geq 0$. С другой стороны, так как $y_* \geq 0$, а $\varphi_0(x_*) \leq 0$ $\quad y_*^T \varphi(x_*) \leq 0$ и, следовательно, $y_*^T \varphi(x_*) = 0$. Добавляя $y_*^T \varphi(x_*)$ в левую часть неравенства (2.13), получим правое неравенство (2.9).

Левое неравенство (2.9) очевидно, так как $y \geq 0$, а $\varphi(x_*) \leq 0$. Теорема доказана.

Следствие 2.4. Для $i \in M_0(x_*)$ множители Лагранжа $y_*^{(i)} \geq 0$, а для $i \in M_1(x_*), \quad y_*^{(i)} = 0$.

Д о к а з а т е л ь с т в о. В теореме Куна-Таккера установлено равенство $y_*^T \varphi(x_*) = 0$, где $y_* \geq 0$.

Пусть $\varphi_i(x_*) < 0$, $i \in M_1(x_*)$. Тогда соответствующий множитель Лагранжа $y_*^{(i)} = 0$, в противном случае равенство $y_*^T \varphi(x_*) = 0$ невозможно. Следствие установлено.

Рассмотрим задачу минимизации

$$\min \varphi_0(x),\ \varphi_i(x) \leq 0, i = 1, ..., m, x \in \overset{\circ}{V}, \tag{2.14}$$

где $\overset{\circ}{V}$ - выпуклое множество, а $\varphi_0(x), \varphi_1(x), ..., \varphi_m(x)$ – выпукло-дифференцируемые функции.

Назовём эту задачу задачей III. Она отличается от задачи (2.8) тем, что, во-первых, функции $\varphi_0(x), \varphi_1(x), ..., \varphi_m(x)$ ещё и дифференцируемы и, во-вторых, множество $\overset{\circ}{V}$ открытое.

Множество допустимых точек $S = \{x \in \overset{\circ}{V} \mid \varphi_i(x) \le 0, i = 1,..., m\}$ – выпуклое.

Теорема 2.3. Пусть в задаче III (2.14) множество S удовлетворяет условию Слейтера. Тогда для того чтобы x_* была оптимальной точкой, необходимо и достаточно, чтобы $\exists y_* \ge 0$ такой, что

$$y_*^T \varphi(x_*) = 0 \tag{2.15}$$

$$\nabla \varphi_0(x_*) + \sum_{i=1}^{m} y_*^{(i)} \nabla \varphi_i(x_*) = 0 \tag{2.16}$$

Д о к а з а т е л ь с т в о. Д о с т а т о ч н о с т ь. Так как выполнены условия теоремы (2.15) и (2.16), то функция $\psi(x, y)$ по следствию 2.3 удовлетворяет неравенству $\psi(x_*, y_*) \le \psi(x, y_*)$; неравенство $\psi(x_*, y) \le \psi(x_*, y_*)$ очевидно. Поэтому по теореме Куна-Таккера x_* - оптимальное решение задачи III.

Н е о б х о д и м о с т ь. Пусть x_* - оптимальное решение. По теореме Куна-Таккера (x_*, y_*) является седловой точкой функции Лагранжа и выполняется условие (2.15) , т.е. $\psi(x_*, y_*) \le \psi(x, y_*)$ и $y_*^T \varphi(x_*) = 0$. Отсюда и из следствия 2.3 заключаем, что

$$\frac{\partial \psi(x_*, y_*)}{\partial x^{(i)}} = 0, i = 1,..., m. \tag{2.17}$$

Левые части равенств (2.17) и (2.16) совпадают, что и доказывает теорему.

Рассмотрим задачу минимизации IV

$$\min \varphi_0(x),\ \varphi_i(x) \le 0,\ i = 1,..., m,\ x \ge 0,\ x \in \overset{\circ}{V}, \tag{2.18}$$

где $\overset{\circ}{V}$ - выпуклое множество, а $\varphi_0(x), \varphi_1(x), ..., \varphi_m(x)$ – выпукло-дифференцируемые функции.

Эта задача отличается от задачи III только наличием ограничений $x \geq 0$. Множество допустимых точек $S = \{x \in \overset{\circ}{V} \mid \varphi_i(x) \leq 0, i = 1, ..., m; x \geq 0\}$.

Теорема 2.4. Пусть в задаче IV множество S удовлетворяет условию Слейтера. Тогда для того чтобы x_* была оптимальной точкой, необходимо и достаточно, чтобы $\exists y_* \geq 0$ такой, что

$$y_*^T \varphi(x_*) = 0 \tag{2.19}$$

$$\nabla \varphi_0(x_*) + \sum_{i=1}^{m} y_*^{(i)} \nabla \varphi_i(x_*) \geq 0 \ , \tag{2.20}$$

причём

$$\frac{\partial \varphi_0(x_*)}{\partial x^{(j)}} + \sum_{i=1}^{m} y_*^{(i)} \frac{\partial \varphi_i(x_*)}{\partial x^{(j)}} = 0 \text{, если } x_*^{(j)} > 0 \, . \tag{2.21}$$

Д о к а з а т е л ь с т в о. Задачу (2.18) можно свести к задаче (2.14), если ввести функции

$\varphi_i(x) = -x^{(i-m)}, i = m+1, ..., m+n$. Эти функции выпуклые, поэтому, применяя теорему 2.3, получим, что необходимыми и достаточными условиями оптимальности x_* будут: $\exists y_* \in R^{m+n}$, $y_* \geq 0$ такого, что

$$\sum_{i=1}^{m+n} y_*^{(i)} \varphi_i(x_*) = 0 \tag{2.22}$$

$$\nabla \varphi_0(x_*) + \sum_{i=1}^{m+n} y_*^{(i)} \nabla \varphi_i(x_*) = 0 \tag{2.23}$$

Производные функций $\varphi_i(x)$ для $i = m+1, ..., m+n$ могут быть легко вычислены и $\nabla \varphi_i(x) = -e_i$, ($e_i$ – единичные орты $e_i^T = (0 \ldots 1 \ldots 0)$, единица стоит в (*i-m)*-й позиции), $i = m+1, \ldots, m+n$. Поэтому равенство (2.23) можно записать в виде

$$\nabla \varphi_0(x_*) + \sum_{i=1}^{m} y_*^{(i)} \nabla \varphi_i(x_*) = \sum_{i=m+1}^{m+n} y_*^{(i)} e_i \geq 0$$

Из (2.22) видно, что $y_*^{(m+j)} = 0$, если $x_*^{(j)} > 0$ и справедливо (2.21). И наоборот, если $x_*^{(j)} > 0$, то из (2.22) следует, что $y_*^{(m+j)} = 0$. Теорема доказана.

В формулировках теорем 2.2, 2.3, 2.4 присутствует не очень удобное условие Слейтера, которое используется при доказательстве необходимости. Это условие существенно.

Пример 2. Пусть $m=1, \varphi_0(x)=-x, \varphi_1(x)=x^2, \varphi_2(x)=x$. Допустимое множество состоит из одной точки $x=0$, которая и является оптимальной точкой.

Построим функцию Лагранжа $\psi(x,y)=-x+y^{(1)}x^2+y^{(2)}x$. Очевидно, что функция Лагранжа не имеет минимума по x в точке $x_*=0$.

Однако имеется класс задач выпуклого программирования, в которых условие Слейтера может быть опущено.

Рассмотрим задачу минимизации V

$$\min \varphi_0(x),\ a_i^T x - b^{(i)} \le 0,\ i \in M = \{1,...m\}. \tag{2.24}$$

Теорема 2.5. Для того, чтобы x_* была оптимальной в задаче V, необходимо и достаточно, чтобы $\exists y_* \ge 0$ такой, что

$$\sum_{i\in M} y_*^{(i)}(a_i^T x_* - b^{(i)}) = 0 \tag{2.25}$$

$$\nabla \varphi_0(x_*) + \sum_{i\in M} y_*^{(i)} a_i = 0 \tag{2.26}$$

Д о к а з а т е л ь с т в о. Необходимым и достаточным условием минимума в точке x_* является неравенство

$$\nabla^T \varphi_0(x_*)(x-x_*) \ge 0, \forall x \in S, \tag{2.27}$$

где $S=\{x \mid a_i^T x - b^{(i)} \le 0, i\in M\}$.

Обозначим $M_0(x_*)=\{i\in M \mid a_i^T x_* - b^{(i)} = 0\}$. Тогда

$$a_i^T(x-x_*) \le 0,\ i \in M_0(x_*), \forall x \in S. \tag{2.28}$$

Обозначим через A - матрицу, строки которой образуют векторы $a_i^T, i\in M_0(x_*)$. Тогда (2.27) и (2.28) можно представить в виде

$$\nabla^T \varphi_0(x_*) z \ge 0,\ Az \le 0,\ z = x - x_*. \tag{2.29}$$

По теореме Фаркаша задача (2.29) эквивалентна задаче

$\exists y_* \ge 0,\ A^T y_* = -\nabla\varphi_0(x_*)$

или в другой форме

$$\exists y_* \ge 0,\ \nabla\varphi_0(x_*) + \sum_{i\in M_0(x_*)} y_*^{(j)} a_i = 0\ . \tag{2.30}$$

Положим $y_*^{(j)} = 0, i \in M_1(x_*)$, тогда (2.30) можно записать в виде

$$\exists y_* \ge 0,\ \nabla\varphi_0(x_*) + \sum_{i\in M} y_*^{(i)} a_i = 0\ , \tag{2.31}$$

причем

$\sum_{i\in M} y_*^{(i)}(a_i^T x_* - b^{(i)}) = 0$, так как $\forall i \in M_1(x_*),\ y_*^{(i)} = 0$ и

$\forall i \in M_0(x_*),\ a_i^T x_* - b^{(i)} = 0$. Теорема доказана.

Следствие 2.5 Для того, что бы x_* была оптимальной в задаче V, необходимо и достаточно $\exists y_* \ge 0$, такой чтобы (x_*, y_*) была седловой точкой функции Лагранжа для задачи V .

Д о к а з а т е л ь с т в о. Д о с т а т о ч н о с т ь доказана для более общей задачи в теореме Куна-Таккера.

Н е о б х о д и м о с т ь. Пусть x_* - оптимальная точка. Тогда справедливы условия (2.25), (2.26).

Для функции Лагранжа $\psi(x,y) = \varphi_0(x) + \sum_{i\in M} y^{(i)}(a_i^T x - b^{(i)})$

условие (2.26) является необходимым и достаточным условием минимума функции $\psi(x,y)$ в точке x_* при $y = y_*$, т.е. $\psi(x_*, y_*) \le \psi(x, y_*)$.

Неравенство $\psi(x_*, y) \le \psi(x_*, y_*), \forall y \ge 0$ следует из (2.25). Следствие установлено.

Рассмотрим задачу VI :

$\min \varphi_0(x), \forall x \in S$,

$$S = \{x \in V \mid \varphi_i(x) \le 0, i = 1,...,m,\ a_j^T x - b^{(j)} = 0,\ j = 1,...,l\} \tag{2.32}$$

Теорема 2.6. Пусть множество S удовлетворяет условию Слейтера, т.е. $\exists \bar{x} \in S,\ \varphi(\bar{x}) < 0$. Тогда для того, что бы x_* была оптимальной в задаче VI , необходимо и достаточно, что бы $\exists y_* \in R^m$, $y_* \geq 0$ и $z_* \in R^l$, такие, что функция Лагранжа

$$\psi(x,y,z) = \varphi_0(x) + y^T\varphi(x) + z^T(Ax-b), \text{ где}$$

$$y \in R^m, z \in R^l, A = \begin{pmatrix} a_1^{(1)} & \dots & a_1^{(n)} \\ \vdots & \ddots & \vdots \\ a_l^{(1)} & \dots & a_l^{(n)} \end{pmatrix}, b = \begin{pmatrix} b^{(1)} \\ \vdots \\ b^{(l)} \end{pmatrix}, \varphi(x) = \begin{pmatrix} \varphi_1(x) \\ \vdots \\ \varphi_m(x) \end{pmatrix}$$

имела седловую точку, т.е.

$$\psi(x_*, y, z) \leq \psi(x_*, y_*, z_*) \leq \psi(x, y_*, z_*)$$

Доказательство этой теоремы не приводится.

2.3. Двойственные задачи

Рассмотрим выпуклые функции $\varphi_0(x), \varphi_1(x), ..., \varphi_m(x)$, заданные на выпуклом множестве *V*, и функцию Лагранжа

$\psi(x,y) = \varphi_0(x) + y^T\varphi(x),\ x \in V, y \geq 0$.

Построим две функции:

$$\psi_*(y) = \inf \psi(x,y), \forall x \in V, \quad \psi^*(x) = \sup \psi(x,y), \forall y \geq 0$$

Теорема 3.1. Функция $\psi^*(x)$ – выпуклая на множестве V, а функция $\psi_*(y)$-вогнутая на множестве, задаваемом неравенством $y \geq 0$.

Д о к а з а т е л ь с т в о. Докажем выпуклость $\psi^*(x)$.Возьмем $\forall x_1, x_2 \in V, 0 \leq \lambda \leq 1$, тогда

$$(1-\lambda)\psi^*(x_1) + \lambda\psi^*(x_2) \geq (1-\lambda)\psi(x_1, y) + \lambda\psi(x_2, y) \geq \varphi_0[(1-\lambda)x_1 + \lambda x_2] + y^T\varphi[(1-\lambda)x_1 + \lambda x_2] =$$
$$= \psi[(1-\lambda)x_1 + \lambda x_2, y]$$

Переходя в правой части полученного неравенства к супремуму по *y*, получим

$(1-\lambda)\psi^*(x_1) + \lambda\psi^*(x_2) \geq \psi^*[(1-\lambda)x_1 + \lambda x_2]$, что и доказывает выпуклость $\psi^*(x)$.

Возьмем $y_1 \geq 0$ и $y_2 \geq 0$, тогда

$(1-\lambda)\psi_*(y_1)+\lambda\psi_*(y_2)\le(1-\lambda)\psi(x,y_1)+\lambda\psi(x,y_2)=\varphi_0(x)+[(1-\lambda)y_1+\lambda y_2]^T\varphi(x)=$
$=\psi[x,(1-\lambda)y_1+\lambda y_2]$

Переходя в правой части к инфимуму по x, получим
$(1-\lambda)\psi_*(y_1)+\lambda\psi_*(y_2)\le\psi_*[(1-\lambda)y_1+\lambda y_2]$,

что и доказывает вогнутость $\psi_*(y)$. Теорема доказана.

Для функций $\psi^*(x)$ и $\psi_*(y)$ можно поставить две задачи выпуклого программирования

$$\min\psi^*(x), \forall x\in V \tag{3.1}$$

$$\max\psi_*(y), \forall y\ge 0 \tag{3.2}$$

Задачу (3.1) будем называть прямой задачей, а (3.2) – двойственной задачей.

Функцию $\psi^*(x)$ легко построить. Действительно, так как

$$\psi^*(x)=\sup[\varphi_0(x)+y^T\varphi(x)],\ y\ge 0,$$

то

$$\psi^*(x)=\begin{cases}\varphi_0(x),\ если\ \varphi(x)\le 0\\ \infty,\ если\ \exists j\in M, \varphi_j(x)>0\end{cases}$$

Таким образом, задача (3.1) совпадает с задачей II. Прямая и двойственная задачи связаны следующим образом

. **Теорема 3.2.** Пусть $x\in S$, а $y\ge 0$. Тогда $\psi_*(y)\le\varphi_0(x)$. Если еще выполнены

условия теоремы Куна-Таккера, то $\max\limits_{y\ge o}\psi_*(y)=\min\limits_{x\in S}\varphi_0(x)$

Д о к а з а т е л ь с т в о . Пусть $x\in V$. Тогда

$$\psi_*(y)\le\varphi_0(x)+y^T\varphi(x)\,. \tag{3.3}$$

Но $\varphi(x)\le 0$ для $x\in S$ и $y\ge 0$, поэтому из неравенства (3.3) следует

$\psi_*(y)\le\varphi_0(x)$ для $x\in S$

Пусть еще выполнены условия теоремы Куна-Таккера. Тогда необходимым и достаточным условием оптимальности x_* в задаче II является $\exists y_* \geq 0$ такого, что функция $\psi(x,y)$ имеет седловую точку (x_*, y_*) ,то есть $\psi(x_*,y) \leq \psi(x_*,y_*) \leq \psi(x,y_*)$.Отсюда $\sup\limits_{y\geq 0} \psi(x_*,y) \leq \psi(x_*,y_*) \leq \inf\limits_{x\in V} \psi(x,y_*)$.Кроме того, $\psi_*(y) \leq \psi^*(x)$, поэтому $\psi_*(y_*) = \psi^*(x_*) = \min \varphi_0(x), \forall x \in S$. Теорема доказана.

2.4. Линейное программирование

Рассмотрим три задачи линейного программирования. Первую из них назовем общей задачей линейного программирования

$$\min c^T x, \forall x \in S \tag{4.1}$$

$$S = \left\{ x \mid a_i^T x - b^{(i)} \geq 0,\ i \in M_1\ , a_i^T x - b^{(i)} = 0,\ i \in M_2, x^{(i)} \geq 0, i \in N_1, x^{(i)} - \forall\ \textit{знака}, i \in N_2 \right\}$$

$$M_1 \cup M_2 = M = \{1,\ldots,m\}, \qquad N_1 \cup N_2 = N = \{1,\ldots,n\}$$

Вторая называется симметричной задачей линейного программирования

$$\min c^T x, \forall x \in S \tag{4.2}$$

$$S = \left\{ x \mid a_i^T x - b^{(i)} \geq 0,\ i \in M,\ x \geq 0 \right\}.$$

Третья задача называется канонической задачей линейного программирования

$$\min c^T x, \forall x \in S \tag{4.3}$$

$$S = \left\{ x \mid a_i^T x - b^{(i)} = 0,\ i \in M,\ x \geq 0 \right\}.$$

Прежде всего введем удобные обозначения для матриц и векторов, которые более наглядно позволяют представить задачи линейного программирования. Введем

$x = x[N],\ b = b[M],\ A[M,N],\ c = c[N],\ b[M_1],\ b[M_2],\ A[M_1,N],\ A[M_2,N],\ x[N_1],\ x[N_2]$.

Здесь в квадратных скобках приведены множества значений, которые могут пробегать индексы соответствующих переменных с индексами

$x^{(i)}, b^{(i)}, c^{(i)}$. Строками матриц $A[M,N]$, $A[M_1,N]$, $A[M_2,N]$ являются векторы a_i^T, $i \in M$, a_i^T, $i \in M_1$. a_i^T, $i \in M_2$, соответственно.

Такие обозначения позволяют выделить из векторов и матриц различные части векторов и матриц. Задачи (4.1) – (4.3) могут быть представлены в виде

$$\min c^T[N] \cdot x[N],\ x[N] \in S \tag{4.4}$$

$$S = \{x[N] \mid A[M_1,N] \cdot x[N] \geq b[M_1],\ A[M_2,N] \cdot x[N] = b[M_2],\ x[N_1] \geq 0\}.$$

$$\min c^T[N] \cdot x[N],\ x[N] \in S \tag{4.5}$$

$$S = \{x[N] \mid A[M,N] \cdot x[N] \geq b[M],\ x[N] \geq 0\}.$$

$$\min c^T[N] \cdot x[N],\ x[N] \in S \tag{4.6}$$

$$S = \{x[N] \mid A[M,N] \cdot x[N] = b[M],\ x[N] \geq 0\}.$$

Покажем, что все три задачи линейного программирования эквивалентны. Задачи (4.5) и (4.6) являются частным случаем задачи (4.4), первая при $M_1 = M,\ M_2 = \varnothing,\ N_1 = N,\ N_2 = \varnothing$, вторая при $M_1 = \varnothing,\ M_2 = M,\ N_1 = N,\ N_2 = \varnothing$.

Для сведения задачи (4.4) к (4.5) представим равенство

$A[M_2,N] \cdot x[N] = b[M_2]$

в виде двух неравенств

$$A[M_2,N] \cdot x[N] \geq b[M_2],\ -A[M_2,N] \cdot x[N] \geq -b[M_2].$$

Вектор $x[N_2]$ представим в виде разности двух векторов $x[N_2] = u[N_2] - v[N_2],\ \text{где}\ u[N_2] \geq 0,\ v[N_2] \geq 0$.

Введем новые векторы $\bar{x}[\bar{N}],\ \bar{b}[\bar{M}]$ и матрицу $\bar{A}[\bar{M},\bar{N}]$

$$\bar{x}^T[\bar{N}] = \begin{pmatrix} x^T[N_1] & u^T[N_2] & v^T[N_2] \end{pmatrix}$$

$$\bar{b}^T[\bar{M}] = \begin{pmatrix} b^T[M_1] & b^T[M_2] & -b^T[M_2] \end{pmatrix}$$

$$\bar{c}^T[\bar{N}] = \begin{pmatrix} c^T[N_1] & c^T[N_2] & -c^T[N_2] \end{pmatrix}$$

$$\bar{A}[\bar{M},\bar{N}] = \begin{pmatrix} A[M_1,N_1] & A[M_1,N_2] & -A[M_1,N_2] \\ A[M_2,N_1] & A[M_2,N_2] & -A[M_2,N_2] \\ -A[M_2,N_1] & -A[M_2,N_2] & A[M_2,N_2] \end{pmatrix}$$

Тогда задача (4.4) может быть представлена в виде

$$\min \bar{c}^T[\bar{N}] \cdot \bar{x}[\bar{N}],\ \bar{x}[\bar{N}] \in S \qquad (4.7)$$

$$S = \left\{ \bar{x}[\bar{N}] \mid \bar{A}[\bar{M},\bar{N}] \cdot \bar{x}[\bar{N}] \ge \bar{b}[\bar{M}],\ \bar{x}[\bar{N}] \ge 0 \right\}.$$

Для сведения задачи (4.4) к (4.6) запишем ограничение

$A[M_1,N] \cdot x[N] \ge b[M_1]$

в виде

$$A[M_1,N] \cdot x[N] - w[M_1] = b[M_1],$$

где $w[M_1]$ - дополнительный неотрицательный вектор дополнительных неизвестных. Введем новые векторы и матрицу

$$\bar{x}[\bar{N}] = \left(x^T[N_1] \quad u^T[N_2] \quad v^T[N_2] \quad w^T[M_1] \right)$$

$$\bar{b}^T[\bar{M}] = \left(b^T[M_1] \quad b^T[M_2] \right)$$

$$\bar{c}^T[\bar{N}] = \left(c^T[N_1] \quad c^T[N_2] \quad -c^T[N_2] \quad \mathrm{O}[M_1] \right)$$

$$\bar{A}[\bar{M},\bar{N}] = \begin{pmatrix} A[M_1,N_1] & A[M_1,N_2] & -A[M_1,N_2] & -E[M_1,N_1] \\ A[M_2,N_1] & A[M_2,N_2] & -A[M_2,N_2] & \mathrm{O}[M_2,M_1] \end{pmatrix}$$

Тогда задача (4.4) может быть представлена в виде (4.7).

Сведение задач (4.5), (4.6) друг к другу делается аналогично.

Теорема 4.1. Для того, чтобы $x_*[N]$ была оптимальной точкой в задаче линейного программирования (4.4), необходимо и достаточно, чтобы $\exists y_*[M]$ такой, что:

$$y_*[M_1] \ge 0 \qquad (4.8)$$

$$c^T[N_1] - y_*^T[M] \cdot A[M,N_1] \ge 0 \qquad (4.9)$$

$$c^T[N_2] - y_*^T[M] \cdot A[M,N_2] = 0 \qquad (4.10)$$

$$y_*^T[M_1] \cdot (A[M_1,N] \cdot x_*[N] - b[M_1]) = 0 \qquad (4.11)$$

$$(c^T[N_1]-y_*^T[M]\cdot A[M,N_1])\cdot x_*[N_1]=0 \tag{4.12}$$

Д о к а з а т е л ь с т в о. Запишем ограничения задачи (4.4) в виде неравенств

$$-A[M_1,N]\cdot x[N]+b[M_1]\le 0,\ A[M_2,N]\cdot x[N]-b[M_2]\le 0$$

$$-A[M_2,N]\cdot x[N]+b[M_2]\le 0,\ -x[N_1]\le 0$$

и воспользуемся теоремой 2.5, из которой следует, что необходимым и достаточным условием оптимальности $x_*[N]$ будет

$\exists y_*[M_1]\ge 0,\quad u_*[M_2]\ge 0,\quad v_*[M_2]\ge 0,\quad w_*[N_1]\ge 0$

таких, что

$$c^T[N]-y_*^T[M_1]\cdot A[M_1,N]+u_*^T[M_2]\cdot A[M_2,N]-v_*^T[M_2]\cdot A[M_2,N]-\sum_{i\in N_1} w_*[i]\cdot e_i[N]=0, \tag{4.13}$$

где

$$y_*[i]=0,\ если\ \ -A[i,N]\cdot x_*[N]+b[i]<0,\ i\in M_1 \tag{4.14}$$

$$w_*[i]=0,\ если\ \ -x_*[i]<0,\ i\in N_1 \tag{4.15}$$

Обозначим $y_*[M_2]=v_*[M_2]-u_*[M_2]$. Тогда получаем (4.8) и из (4.13) следуют условия

$$c^T[N_1]-y_*^T[M]\cdot A[M,N_1]\ge 0$$

$$c^T[N_2]-y_*^T[M]\cdot A[M,N_2]=0,$$

так как $w_*[N_1]\ge 0$.

Из (4.14) и (4.15) следуют равенства (4.11) и (4.12). Теорема доказана.

Следствие 4.1. Для того, чтобы $x_*[N]$ была оптимальной точкой в задаче линейного программирования (4.5) , необходимо и достаточно, чтобы $\exists y_*[M]\ge 0$ такой, что

$$c^T[N]-y_*^T[M]\cdot A[M,N]\ge 0. \tag{4.16}$$

$$y_*^T[M]\cdot(A[M,N]\cdot x_*[N]-b[M])=0. \tag{4.17}$$

$$(c^T[N]-y_*^T[M]\cdot A[M,N])\cdot x_*[N]=0 \tag{4.18}$$

Доказательство следует непосредственно из теоремы 4.1.

Следствие 4.2. Для того, чтобы $x_*[N]$ была оптимальной точкой в задаче линейного программирования (4.6) , необходимо и достаточно, чтобы $\exists y_*[M]$ такой, что

$$c^T[N] - y_*^T[M]\cdot A[M,N] \geq 0 \tag{4.19}$$

$$(c^T[N] - y_*^T[M]\cdot A[M,N])\cdot x_*[N] = 0 \tag{4.20}$$

Доказательство следует непосредственно из теоремы 4.1.

Теорема 4.2. Для того, чтобы $x_*[N]$ была оптимальной точкой в задаче линейного программирования (4.4), необходимо и достаточно, чтобы $\exists y_*[M]$ такой, что $y_*[M_1] \geq 0$ и функция Лагранжа

$$\psi(x,y) = c^T[N]\cdot x[N] + y^T[M]\cdot(b[M] - A[M,N]\cdot x[N]) \tag{4.21}$$

на множествах $x[N_1] \geq 0,\ y[M_1] \geq 0$ имела седловую точку $(x_*[N],\ y_*[M])$.

Д о к а з а т е л ь с т в о. Эта теорема является частным случаем теоремы 2.6.

Применим теорию двойственности к общей задаче линейного программирования.

Построим функцию $\psi^*(x)$ на множестве $x[N_1] \geq 0$

$\psi^*(x) = \sup\psi(x,y),\ \forall y[M_1] \geq 0$,

где $\psi(x,y)$ задана правой частью (4.21).

Для определения $\psi^*(x)$ представим $\psi(x,y)$ в виде

$\psi(x,y) = c^T[N]\cdot x[N] + y^T[M_1]\cdot(b[M_1] - A[M_1,N]\cdot x[N]) + y^T[M_2]\cdot(b[M_2] - A[M_2,N]\cdot x[N])$.

Если $\exists i \in M_1:\ \ b[i] > A[i,N]\cdot x[N],$ то $\psi^*(x) = \infty$,

если $\exists i \in M_2:\ \ b[i] \neq A[i,N]\cdot x[N],$, то $\psi^*(x) = \infty$.

Если же $b[M_1] - A[M_1,N]\cdot x[N] \leq 0$ и $b[M_2] - A[M_2,N]\cdot x[N] = 0$, то $\psi^*(x) = c^T[N]\cdot x[N]$.

Тогда запишем функцию

$$\psi^*(x) = \begin{cases} \infty, \ b[i] > A[i,N]\cdot x[N], \ i\in M_1 \\ \infty, \ b[i] \neq A[i,N]\cdot x[N], \ i\in M_2 \\ c^T[N]\cdot x[N], \ \begin{matrix} b[M_1] - A[M_1,N]\cdot x[N] \le 0, \\ b[M_2] - A[M_2,N]\cdot x[N] = 0 \end{matrix} \end{cases}$$

Построим функцию $\psi_*(y)$ на множестве $y[M_1]\ge 0$

$$\psi_*(y) = \inf \psi(x,y), \quad \forall x[N_1]\ge 0,$$

где $\psi(x,y)$ определена правой частью (4.21).

Для определения $\psi_*(y)$ представим $\psi(x,y)$ в виде

$$\psi(x,y) = y^T[M]\cdot b[M] + (c^T[N_1] - y^T[M]\cdot A[M,N_1])\cdot x[N_1] + (c^T[N_2] - y^T[M]\cdot A[M,N_2])\cdot x[N_2]$$

Если $\exists i\in N_1: \ \ c[i] < y^T[M]\cdot A[M,i]$, то $\psi_*(y) = -\infty$,

если $\exists i\in N_2: \ \ c[i] \neq y^T[M]\cdot A[M,i]$, то $\psi_*(y) = -\infty$.

Если же $c^T[N_1] \ge y^T[M]\cdot A[M,N_1]$ и $c^T[N_2] = y^T[M]\cdot A[M,N_2]$, то $\psi_*(y) = y^T[M]\cdot b[M]$.

Тогда

$$\psi_*(y) = \begin{cases} -\infty, \ c[i] < y^T[M]\cdot A[M,i], \ i\in N_1 \\ -\infty, \ c[i] \neq y^T[M]\cdot A[M,i], \ i\in N_2 \\ y^T[M]\cdot b[M], \ \begin{matrix} c^T[N_1] \ge y^T[M]\cdot A[M,N_1], \\ c^T[N_2] = y^T[M]\cdot A[M,N_2] \end{matrix} \end{cases}$$

Задача минимизации $\psi^*(x), \ \forall x[N_1]\ge 0$ эквивалентна общей задаче линейного программирования (4.4).

Задача максимизации $\psi_*(y)$, $\forall y[M_1]\ge 0$ эквивалентна задаче

$$\begin{aligned} &\max \ b^T[M]\cdot y[M], \ \ \forall y[M]\in S_{дв.} \\ &S_{дв.} = \{y[M] \mid A^T[N_1,M]\cdot y[M] \le c[N_1], \\ &\qquad A^T[N_2,M]\cdot y[M] = c[N_2], \ y[M_1]\ge 0\} \end{aligned} \tag{4.22}$$

Эта задача является двойственной задачей линейного программирования к основной задаче линейного программирования.

Из (4.22) легко получить двойственную задачу к симметричной задаче линейного программирования

$$\max \ b^T[M]\cdot y[M]\ ,\ \forall y[M]\in S_{дв.}$$
$$S_{дв.}=\{y[M]\mid A^T[N,M]\cdot y[M]\le c[N],\, y[M]\ge 0\}, \qquad (4.23)$$

а также к канонической задаче линейного программирования

$$\max \ b^T[M]\cdot y[M]\ ,\ \forall y[M]\in S_{дв.}$$
$$S_{дв.}=\{y[M]\mid A^T[N,M]\cdot y[M]\le c[N]\} \qquad (4.24)$$

Применяя к двойственным задачам (4.22), (4.23) и (4.24) опять теорию двойственности, можно увидеть, что двойственными к ним будут, соответственно,
задачи (4.4), (4.5) и (4.6).

Для основной задачи линейного программирования возможны следующие решения:

1) $S\ne\varnothing, \min c^T[N]\cdot c[N]\ -\ существует;$

2) $S\ne\varnothing, \min c^T[N]\cdot x[N]\ -\ не\ \ существует;$

3) $S=\varnothing.$

Для двойственной к основной задаче линейного программирования возможны следующие решения:

1) $S_{дв.}\ne\varnothing, \max b^T[M]\cdot y[M]\ \ -\ существует;$

2) $S_{дв.}\ne\varnothing, \max b^T[M]\cdot y[M]\ \ -\ не\ \ существует;$

3) $S_{дв.}=\varnothing.$

Рассмотрим, как соотносятся решения основной (О) и двойственной (Д) задач линейного программирования. Возможные варианты приведены в таблице 2.

Таблица 2.

О \ Д	1	2	3
1		X	X
2	X	X	
3	X		

Теорема 4.3. Для основной и двойственной задач линейного программирования возможными являются только комбинации 1-1, 2-3, 3-2, 3-3 таблицы.

Д о к а з а т е л ь с т в о . Пусть $\exists x_*[N] \in S$. Тогда по теореме 4.2 $\exists y_*[M]$ такой, что

$y_*[M_1] \geq 0$ и $\psi(x_*, y) \leq \psi(x_*, y_*) \leq \psi(x, y_*)$,

откуда и из теоремы 3.2 вытекает, что $y_*[M] \in S_{дв.}$ является оптимальным решением двойственной задачи. Верно и обратное, так как двойственной задачей к (4.22) является задача (4.4).

Заметим также, что невозможна комбинация 2 – 2, так как $\psi_*(y) \leq \psi^*(x)$ и невозможно, чтобы $\psi_*(y) \to \infty$, а $\psi^*(x) \to -\infty$. Теорема доказана.

Определение 4.1. Точка $x_0 \in S$ - выпуклому множеству называется крайней точкой *S*, если не существует $x_1, x_2 \in S, x_1 \neq x_2$, таких, что $x_0 = (1-\lambda)x_1 + \lambda x_2$ для некоторого $0 < \lambda < 1$.

Определение 4.2. Множество S называется многогранным множеством, если оно является пересечением конечного числа замкнутых полупространств и гиперплоскостей (рис.9), (рис.10).

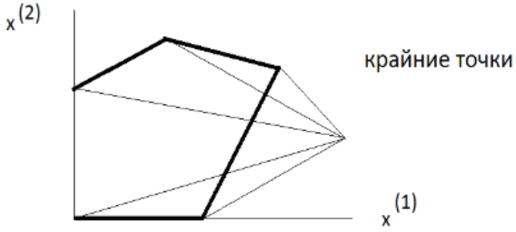

Рис.9.

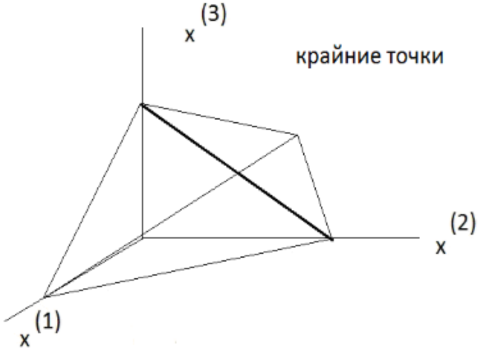

Рис.10.

В дальнейшем будем рассматривать каноническую задачу линейного программирования.

$$\min c^T[N] \cdot x[N] \tag{4.25}$$

$S = \{x[N] \mid A[M,N] \cdot x[N] = b[M], x[N] \geq 0\}$.

Будем считать, что количество строк матрицы A равно $m = |M|$, а количество столбцов - $n = |N|$, причем $m < n$ и ранг матрицы A равен m.

Возьмем любой допустимый вектор $x_0[N] \in S$. Его компоненты или равны нулю или положительные. Введем множества

$N_0^+ = \{i \in N \mid x_0[i] > 0\}, \quad N_0^0 = \{i \in N \mid x_0[i] = 0\}$,

Тогда ограничение типа равенства можно записать в виде

$$A[M, N_0^+] \cdot x[N_0^+] = b[M], \tag{4.26}$$

так как $x_0[N_0^0]=0$.

Равенство (4.26) можно представить в виде

$$\sum_{j\in N_0^+} x_0[j]\cdot \bar{a}_j[M]=b[M], \tag{4.27}$$

где $\bar{a}_j[M]$ - столбцы матрицы $A[M,N]$.

Определение 4.3. Вектор $x_0[N]\in S$ называется опорным, если векторы $\bar{a}_j[M],\ j\in N_0^+$ линейно независимы.

Для опорного вектора $x_0[N]$ $|N_0^+|\le m$, так как любая система $m+1$ вектора $\bar{a}_j[M]$ всегда линейно зависима.

Определение 4.4. Опорный вектор $x_0[N]$ называется невырожденным, если $|N_0^+|=m$, и вырожденным, если $|N_0^+|<m$.

Для невырожденного опорного вектора система линейных уравнений (4.26) имеет квадратную матрицу, причем $\det A[M,N_0^+]\ne 0$; поэтому (4.26) имеет единственное решение.

Для вырожденного опорного вектора система уравнений (4.26) имеет меньшее число неизвестных $|N_0^+|<m$. Так как матрица $A[M,N]$ имеет ранг m, то систему векторов $\bar{a}_j[M],\ j\in N_0^+$ можно пополнить до линейно независимой [10,12] $\bar{a}_j[M],\ j\in N_0,\ |N_0|=m$, где $x_0[N_0\setminus N_0^+]=0$.

В этом случае опять получим систему m линейных уравнений

$$A[M,N_0]\cdot x_0[N_0]=b[M], \tag{4.28}$$

для которой $\det A[M,N_0]\ne 0$, поэтому она имеет единственное решение

$$x_0[N_0^+],\quad x_0[N_0\setminus N_0^+]=0. \tag{4.29}$$

Пополнение матрицы $A[M,N_0^+]$ до $A[M,N_0]$ можно сделать не единственным образом, однако во всех случаях система (4.28) будет иметь решение (4.29). Отличие будет только в изменении множества N_0/N_0^+ в (4.29). Однако, поскольку $x_0[N_0^0]=0$, то в зависимости от пополнения матрицы $A[M,N_0^+]$ будут изменяться индексы нулевых компонент вектора $x_0[N]$ в множестве N_0.

Определение 4.5. Векторы $\overline{a}_j[M],\ j\in N_0$ называются базисом опорного вектора $x_0[N]$.

Из этого определения и предыдущих рассуждений следует, что каждому базису отвечает один опорный вектор. Но вырожденному опорному вектору могут отвечать несколько базисов.

Лемма 4.1. Число опорных векторов множества S не больше C_n^m.

Д о к а з а т е л ь с т в о. Число всевозможных сочетаний из n векторов $\overline{a}_j[M],\ j\in N$ по m равно C_n^m, поэтому можно составить матрицы $A[M,N_k], k=1,...,C_n^m$.

Некоторые матрицы $A[M,N_k]$ будут иметь $\det A[M,N_k]=0$. Тогда их можно отбросить. Для $\det A[M,N_k]\neq 0$ составим систему линейных уравнений

$$A[M,N_k]\cdot x[N_k]=b[M],$$

которая имеет единственное решение $x[N_k]$.

Если вектор $x[N_k]\geq 0$, то , полагая $x[N\setminus N_k]=0$, получим опорный вектор $x[N]\geq 0$.

Если $\exists x[i]<0,\ i\in N_k$, то матрица $A[M,N_k]$ отбрасывается. Лемма доказана.

Пример 1. Пусть множество S определяется уравнениями

$$2x^{(1)}+2x^{(2)}+3x^{(3)}+x^{(4)}=2 \qquad (4.30)$$
$$x^{(1)}+2x^{(2)}-2x^{(3)}-x^{(4)}=1$$

и неравенством $x\geq 0$.

Полагая любые две компоненты x равными нулю, получим шесть систем уравнений

$$1)\; 2x^{(1)}+2x^{(2)}=2 \qquad 2)\; 2x^{(1)}+3x^{(3)}=2 \qquad 3)\; 2x^{(1)}+x^{(4)}=2$$

$$x^{(1)}+2x^{(2)}=1 \qquad x^{(1)}-2x^{(3)}=1 \qquad x^{(1)}-x^{(4)}=1$$

$$4)\; 2x^{(2)}+3x^{(3)}=2 \qquad 5)\; 2x^{(2)}+x^{(4)}=2 \qquad 6)\; 3x^{(3)}+x^{(4)}=2$$

$$2x^{(2)}-2x^{(3)}=1 \qquad 2x^{(2)}-x^{(4)}=1 \qquad -2x^{(3)}-x^{(4)}=1,$$

из которых найдем шесть решений

$$x_1^T=(1 \quad 0 \quad 0 \quad 0), \quad x_2^T=(1 \quad 0 \quad 0 \quad 0), \quad x_3^T=(1 \quad 0 \quad 0 \quad 0),$$

$$x_4^T=(0 \quad 7/10 \quad 1/5 \quad 0), \quad x_5^T=(0 \quad 3/4 \quad 0 \quad 1/2), \quad x_6^T=(0 \quad 0 \quad 3 \quad -7).$$

Первые три вектора одинаковые, однако, им соответствуют разные базисы:

$$1)\; \overline{a}_1^T=(2 \quad 1), \quad \overline{a}_2^T=(2 \quad 2); \qquad 2)\; \overline{a}_1^T=(2 \quad 1), \quad \overline{a}_3^T=(3 \quad -2);$$

$$3)\; \overline{a}_1^T=(2 \quad 1), \quad \overline{a}_4^T=(1 \quad -1).$$

Вектор x_6 не является допустимым, x_4 и x_5 - являются невырожденными опорными векторами, которым соответствуют базисы $\overline{a}_2^T=(2 \quad 2), \quad \overline{a}_3^T=(3 \quad -2)$ и $\overline{a}_2^T=(2 \quad 2), \quad \overline{a}_4^T=(1 \quad -1)$.

Лемма 4.2. Вектор $x[N]$ является опорным тогда и только тогда, когда $x[N]$ является крайней точкой.

Д о к а з а т е л ь с т в о. Пусть $x[N]$ - не опорный вектор. Тогда для $x[N^+]>0,\ \overline{a}_i[M],\ i\in N^+$ линейно зависимы, т.е. $\exists u[N^+]\neq 0$ такой, что

$$A[M,N^+]\cdot u[N^+]=0. \tag{4.31}$$

Будем считать, что $\exists u[i]>0, i\in N^+$, в противном случае можно $u[N^+]$ умножить на -1.

Домножим (4.31) на θ и вычтем его из $A[M,N^+]\cdot x[N^+]=b[M]$.

Тогда $A[M,N^+]\cdot(x[N^+]-\theta u[N^+])=b[M]$.

Выберем $\theta_1=\min\limits_{i\in N^+,u[i]>0}(x[i]/u[i])$. Тогда $x[i]-\theta_1 u[i]\geq 0,\ i\in N^+$.

Рассмотрим $A[M,N^+]\cdot(x[N^+]+\theta u[N^+])=b[M]$.

Если $\exists u[i]<0,\ i\in N^+$, то выберем $\theta_2 = \min\limits_{i\in N^+, u[i]<0} (-x[i]/u[i])$.

Тогда $x[i]+\theta_2 u[i]\ge 0, i\in N^+$.

Если $u[N^+]\ge 0$, то положим $\theta_2=\theta_1$. Выберем $\theta_0=\min\{\theta_1,\theta_2\}$,положим $u[N\setminus N^+]=0$ и постоим два вектора

$$x_1[N]=x[N]+\theta_0 u[N],\ x_2[N]=x[N]-\theta_0 u[N].$$

Тогда $x[N]$ – не крайняя точка, так как

$$x[N]=(x_1[N]+x_2[N])/2.$$

Пусть $x[N]$ - не крайняя точка, т.е. $\exists x_1[N], x_2[N]\in S$, $x_1[N]\ne x_2[N]$ такие, что

$$x[N]=(1-\lambda)x_1[N]+\lambda x_2[N]), 0<\lambda<1.$$

Так как

$x[N^0]=0$, то $x_1[N^0]=x_2[N^0]=0$.

Из $A[M,N^+]\cdot x_1[N^+]=b[M]$ и $A[M,N^+]\cdot x_2[N^+]=b[M]$ следует $A[M,N^+]\cdot(x_1[N^+]-x_2[N^+])=0$, $x_1[N^+]\ne x_2[N^+]$. Таким образом, $\bar{a}_i,\ i\in N^+$ линейно зависимы и, следовательно, $x[N]$ – не опорный вектор. Лемма доказана.

Теорема 4.4. Пусть $x_*[N]$ - оптимальное решение задачи линейного программирования (4.25). Тогда $x_*[N]$ - опорный вектор или $\exists\, \bar{x}_*[N]\in S_*$, где $\bar{x}_*[N]$ - опорный вектор.

Д о к а з а т е л ь с т в о. Выберем из множества оптимальных векторов такой, который имеет минимальное количество положительных компонент. Обозначим его через $\bar{x}_*[N]$.

Предположим, что $\bar{x}_*[N]$ - не опорный вектор. Тогда $\exists u[\bar{N}^+]\ne 0$ такой, что $A[M,\bar{N}^+]\cdot u[\overline{N}^+]=0$. Положим $u[\bar{N}^0]=0$ и построим вектор $x_\theta[N]=x_*[N]+\theta u[N]$.

Так как $x_*[\overline{N}^+] > 0$, а $u[\overline{N}^0] = 0$, то $\exists \theta_0 > 0$, что $x_\theta[N] \ge 0$ для $\theta \in [-\theta_0, \theta_0]$. Кроме того, вектор $x_\theta[N]$ удовлетворяет уравнению:

$$A[M,N]\cdot x_\theta[N] = A[M,\overline{N}^+]\cdot x_\theta[\overline{N}^+] = A[M,\overline{N}^+]\cdot \overline{x_*}[\overline{N}^+] + \theta A[M,\overline{N}^+]\cdot u[\overline{N}^+] = b[M]$$

Подсчитаем $c^T[N]\cdot x_\theta[N] = c^T[N]\cdot \overline{x}_*[N] + \theta\cdot c^T[N]\cdot u[N]$. Скалярное произведение $c^T[N]\cdot u[N] = 0$ в противном случае при $\theta > 0$или $\theta < 0$можно было бы получить меньшее значение целевой функции, чем оптимальное.

В зависимости от знака $u[\overline{N}^+]$ выберем положительное или отрицательное значение θ_0 такое, чтобы по крайней мере одна компонента вектора $x_\theta[\overline{N}^+]$обратилась в нуль. Это всегда можно сделать (см. доказательство леммы 4.2).

Обозначим $\overline{\overline{x}}_*[N] = \overline{x}[N] + \theta_0 u[N]$.Вектор $\overline{\overline{x}}_*[N]$ - оптимальный, причем $\left|\overline{\overline{N}}^+\right| < \left|\overline{N}^+\right|$Но это противоречит предположению, что $\left|\overline{N}^+\right|$ - минимальное значение. Следовательно, $\overline{x_*}[N]$ - опорный вектор. Теорема доказана.

Приведём теоремы, показывающие связь оптимального решения задачи линейного программирования и крайней точки множества допустимых точек.

Теорема 4.5. Пусть $S \in R^n$ - выпуклое, замкнутое, ограниченное множество. Тогда $\forall x_0 \in S$ может быть представлена в виде выпуклой комбинации конечного числа крайних точек множества S.

Д о к а з а т е л ь с т в о. Если n=1, то S- отрезок $[x_1, x_2]$ и $\forall x_0 \in S$ $x_0 = \alpha_1 x_1 + \alpha_2 x_1, \alpha_1 \ge 0, \alpha_2 \ge 0, \alpha_1 + \alpha_2 = 1$.

Пусть утверждение справедливо для n=k. Для n=k+1 рассмотрим два случая, когда точка x_0- граничная и внутренняя.

Если $x_0 \in \partial S$, то построим гиперплоскость, опорную к S в точке x_0

$Q=\{x \mid a^T x = a^T x_0, a \neq 0\}$. Построим $S_0 = S \cap Q$ - выпуклое замкнутое ограниченное множество. Очевидно, что $S_0 \in Q$, поэтому $\exists x_1, x_2 ... x_m$ - крайние точки S_0 такие, что

$$x_0 = \sum_{i=1}^{m} \alpha_i x_i,\ \alpha_i \geq 0,\ \sum_{i=1}^{m} \alpha_i = 1.$$

Покажем, что $x_1, x_2 ... x_m$ - крайние точки S.

Действительно, если x_i не является крайней точкой S, то

$\exists\, x_i^1, x_i^2 \in S, x_i^1 \neq x_i^2, \lambda \in (0,1): \quad x_i = \lambda\, x_i^1 + (1-\lambda) x_i^2$.

Имеем $x_i \in S_0, S_0 \subset Q$, поэтому $a^T x_i = a^T x_0$.

Q-гиперплоскость, опорная к S, значит $a^T x_i^1 \leq a^T x_0, a^T x_i^2 \leq a^T x_0$. С другой стороны,

$a^T x_i^1 = \dfrac{1}{\lambda}(a^T x_i - (1-\lambda) a^T x_i^2) \geq \dfrac{1}{\lambda}(a^T x_0 - (1-\lambda) a^T x_0) = a^T x_0$, поэтому $a^T x_i^1 = a^T x_0$.

Следовательно, $x_i^1 \in Q$. Но $x_i^1 \in S$, значит, $x_i^1 \in S_0$. Рассуждая аналогично, заключаем, что $x_i^2 \in S_0$. Тогда x_i не может быть крайней точкой S_0.

Если x_0 - внутренняя точка *S*, то построим прямую, проходящую через x_0, и представим $x_0 = \lambda\, y_1 + (1-\lambda) y_2, \lambda \in (0,1),\ y_1, y_2 \in \partial S$. Согласно доказанному выше, для граничных точек y_1, y_2 можно записать

$$y_1 = \sum_{i=1}^{m1} \beta_i x_i,\ \beta_i \geq 0,\ \sum_{i=1}^{m1} \beta_i = 1,$$
$$y_2 = \sum_{i=1}^{m2} \delta_i z_i,\ \delta_i \geq 0,\ \sum_{i=1}^{m2} \delta_i = 1,$$

где x_i, z_i - крайние точки *S*. Тогда $x_0 = \lambda \sum_{i=1}^{m1} \beta_i x_i + (1-\lambda) \sum_{i=1}^{m2} \delta_i z_i$,

$\lambda\beta_i \geq 0, (1-\lambda)\delta_i \geq 0, \sum_{i=1}^{m1} \lambda\beta_i + \sum_{i=1}^{m2} (1-\lambda)\delta_i = 1$. Теорема доказана.

Теорема 4.6. Если x_* - оптимальное решение задачи линейного программирования. то существует крайняя точка $\bar{x}$ такая, что $c^T \bar{x} = c^T x_*$.

Д о к а з а т е л ь с т в о. Пусть S – ограниченное множество. Тогда по теореме 4.5 $\exists x_1, x_2 ... x_m$ - крайние точки S такие, что $x_* = \sum_{i=1}^{m} \alpha_i x_i,\ \alpha_i \geq 0,\ \sum_{i=1}^{m} \alpha_i = 1$.

Если $\bar{\exists} x_i,\ c^T x_i = c^T x_*$, то $c^T x_* = \sum_{i=1}^{m} \alpha_i c^T x_i > \sum_{i=1}^{m} \alpha_i c^T x_* = c^T x_*$.

Пусть теперь S – неограниченное множество. Построим множество $S_0 = S \cap D,\quad D = \left\{ x \middle| \sum_{j=1}^{n} x^{(j)} \leq \gamma \right\},\ \gamma > 0,\ x_* \in S_0$, причём,

$x_* \notin d = \left\{ x \middle| \sum_{i=1}^{n} x^{(j)} = \gamma \right\}$.

S_0 -ограниченное множество, поэтому $\exists x_1, x_2 ... x_m$ - крайние точки S_0 и

$x_* = \sum_{i=1}^{m} \alpha_i x_i,\ \alpha_i \geq 0,\ \sum_{i=1}^{m} \alpha_i = 1$. Выберем здесь точки, для которых $\alpha_i > 0$,

$x_* = \sum_{i=1}^{m1} \alpha_i x_i,\ \alpha_i > 0,\ \sum_{i=1}^{m1} \alpha_i = 1$ и перенумеруем точки по неубыванию значений функции цели

$c^T x_* \leq c^T x_1 \leq ... \leq c^T x_{m1}$.

Тогда $c^T x_* = \sum_{i=1}^{m1} \alpha_i c^T x_i \geq \sum_{i=1}^{m1} \alpha_i c^T x_1 = c^T x_1$, следовательно, $c^T x_* = c^T x_1$.

Воспользуется полученным условием и оценим $c^T x_*$ так

$c^T x_* = \sum_{i=1}^{m1} \alpha_i c^T x_i = \alpha_1 c^T x_1 + \sum_{i=2}^{m1} \alpha_i c^T x_i \geq \alpha_1 c^T x_* + \sum_{i=2}^{m1} \alpha_i c^T x_2 = \alpha_1 c^T x_* + (1-\alpha_1) c^T x_2$. Тогда $c^T x_* = c^T x_2$. Рассматривая аналогично оставшиеся точки, получим

$c^T x_* = c^T x_1 = ... = c^T x_{m1}$.

Если среди $x_1, x_2 ... x_{m1}$ имеется хотя бы одна точка $x_i \notin d$, то x_i - крайняя точка S, и теорема доказана.

Предположим теперь, что $x_i \in d, i = 1,...m1$. Тогда

$$\sum_{j=1}^{n} x^{(j)} = \sum_{j=1}^{n}\sum_{i=1}^{m1}\alpha_i x_i^{(j)} = \gamma\sum_{i=1}^{m1}\alpha_i = \gamma$$. значит, $x_* \in d$. Но $x_* \overline{\notin} d$ по построению множества S_0 .Теорема доказана.

2.5. Симплекс-метод

В п.4 показано, что если оптимальное решение задачи (4.25) достигается, то оно достигается в опорном векторе. Количество опорных векторов не быть больше, чем C_n^m , однако, это число может быть достаточно большим при больших n и $m = [n/2]$.

Симплекс-метод позволяет переходить от одного опорного вектора к другому так, что значение целевой функции не увеличивается.

Рассмотрим один шаг симплекс -метода. Пусть имеется $x_k[N]$ – опорный вектор. Построим опорный вектор $x_{k+1}[N]$, для которого $c^T[N]\cdot x_{k+1}[N] \le c^T[N]\cdot x_k[N]$.

Пусть, как обычно, $x_{k+1}[N_k^+] > 0, x_k[N_k^0] = 0$. Столбцы матрицы $A[M,N_k^+]$ – линейно независимы, так как $x_k[N]$ – опорный вектор. Пополним столбцы матрицы $A[M,N_k^+]$ столбцами из $A[M,N_k^0]$ так, чтобы матрица $A[M,N_k]$ была квадратной и $\det A[M,N_k^+] \ne 0$. Обозначим через $L_k = N \setminus N_k$.

Для построения алгоритма воспользуемся необходимыми и достаточными условиями (4.19), (4.20) оптимальности для канонической задачи линейного программирования . Воспользуемся условием (4.19) и составим уравнение относительно $y_k[M]$

$$c^T[N_k] - y_k^T[M]\cdot A[M,N_k] = 0 \tag{5.1}$$

Так как $\det A[M,N_k] \ne 0$ можно разрешить (5.1) относительно $y_k[M]$

$$y_k^T[M] = c^T[N_k]\cdot B[N_k,M], \tag{5.2}$$

где $B[N_k,M] = A^{-1}[M,N_k]$.

Для таким образом определённого $y_k[M]$ условие (4.20) следствия 4.2 выполняется автоматически, так как $x_k[L_k]=0$

Проверим теперь условие (4.19). Подставляя $y_k[M]$ из (5.2) в левую часть (4.19), получим

$$c^T[N]-c^T[N_k]\cdot B[N_k,M]\cdot A[M,N]=d_k^T[N] \tag{5.3}$$

Разобьем вектор (5.3) на два вектора

$$c^T[N_k]-c^T[N_k]\cdot B[N_k,M]\cdot A[M,N_k]=d_k^T[N_k] \tag{5.4}$$

$$c^T[L_k]-c^T[N_k]\cdot B[N_k,M]\cdot A[M,L_k]=d_k^T[L_k] \tag{5.5}$$

Очевидно, что в (5.4) $d_k^T[N_k]=0$, так как $B[N_k,M]=A^{-1}[M,N_k]$.

Рассмотрим теперь вектор $d_k[L_k]$.

Если $d_k[L_k]\geq 0$, то $x_k[N]=x_*[N]$ является оптимальным, так как для него выполнены необходимые и достаточные условия (4.19), (4.20) следствия 4.2. В этом случае алгоритм симплекс-метода заканчивается.

Если $\exists j_k\in L_k$ для которого $d_k[j_k]<0$, то нужно искать следующий опорный вектор $x_{k+1}[N]$.

Построим вектор $u_k[N_k]=B[N_k,M]\cdot A[M,j_k],\ u_k[j_k]=-1, u_k[L_k\setminus j_k]=0$ и при помощи него получим вектор $x_\theta[N]=x_k[N]-\theta u_k[N]$.

Покажем, что вектор $x_\theta[N]$ является допустимым для ограничений типа равенства задачи (4.25) при любом θ. Действительно

$$A[M,N]\cdot x_\theta[N]=A[M,N]\cdot x_k[N]-\theta A[M,N]\cdot u_k[N]=$$
$$=b[M]-\theta A[M,N_k]\cdot B[N_k,M]\cdot A[M,j_k]+\theta A[M,j_k]=b[M]$$

Оставим пока без внимания условие $x_\theta[N]\geq 0$ и рассмотрим значения целевой функции для $x_\theta[N]$

$$c^T[N]\cdot x_\theta[N]=c^T[N]\cdot x_k[N]-\theta c^T[N]\cdot u_k[N]=$$
$$=c^T[N]\cdot x_k[N]-\theta c^T[N_k]\cdot B[N_k,M]\cdot A[M,j_k]+\theta c[j_k]=$$
$$=c^T[N]\cdot x_k[N]+\theta d_k[j_k]$$

Но $d_k[j_k]<0$, поэтому при $\theta>0$значение целевой функции уменьшается, причем желательно, чтобы θ было как можно больше.

Рассмотрим теперь вопрос выбора $\theta>0$.

Если $u_k[N_k]\le 0$ то, так как $u_k[j_k]=-1$, а $u_k[L_k \setminus j_k]=0$, можно $\theta\to\infty$ и следовательно, $c^T[N]\cdot x_\theta[N]\to-\infty$. В этом случае алгоритм симплекс – метода заканчивается, так как задача решена – показано, что целевая функция не ограничена снизу на множестве *S*.

Если $\exists i\in N_k$, для которых $u_k[i]>0$, то θ ограничено сверху условием $x_\theta[N]\ge 0$. Рассмотрим два случая.

Пусть $x_k[N]$ - невырожденный опорный вектор. Тогда $N_k = N_k^+$. В это случае можно найти $\theta_k>0$. Найдем

$$\theta_k = \min_{i\in N_k, u_k[i]>0}\frac{x_k[i]}{u_k[i]}=\frac{x_k[i_k]}{u_k[i_k]} \tag{5.6}$$

и построим вектор

$$x_{k+1}[N]=x_k[N]-\theta_k u_k[N] \tag{5.7}$$

Из определения θ_k ясно, что по крайней мере одна компонента вектора $x_{k+1}[N]$ равна нулю: $x_{k+1}[i_k]$. Одна компонента вектора $x_{k+1}[N]$ становится положительной: $x_{k+1}[j_k]=\theta_k$. Осталось только показать, что $x_{k+1}[N]$ – опорный вектор. Это будет сделано ниже.

Пусть $x_k[N]$ – вырожденный опорный вектор.

Если $u_k[N_k \setminus N_k^+]\le 0$, то можно найти $\theta_k>0$ по формуле (5.6) и построить вектор $x_{k+1}[N]$ по формуле (5.7).

Если $\exists i\in N_k \setminus N_k^+$, для которого $u_k[i]>0$, то нельзя найти $\theta_k>0$. Для продолжения алгоритма не меняют вектор $x_k[N]$, а пытаются изменить базис вектора $x_k[N]$

Это делают следующим образом. Берется столбец $\overline{a_i}, i \in N_k \setminus N_k^+$ и заменяется вектором $\overline{a_i}, i \in L_k$. Далее проверяется, является ли новая система векторов $\overline{a_i}$ линейно независимой.

Если система векторов $\overline{a_i}$ - линейно независима, то происходит возврат к началу алгоритма. Если система векторов $\overline{a_i}$ – линейно зависима, то следует снова поменять столбцы $\overline{a_i}$ из $N_k \setminus N_k^+$ и L_k

В результате таких манипуляций может оказаться, что не удастся сменить вектор $x_k[N]$. Это явление носит название явление зацикливания.

Явление зацикливания встречается только в специально построенных задачах. При реализации алгоритма симплекс-метода на компьютере даже в специально построенных задачах явления зацикливания не встречается. Мы не будем здесь разбирать способы борьбы с зацикливанием. Некоторые правила и примеры можно найти , например, в [5]

Покажем теперь, что новый вектор $x_{k+1}[N]$- опорный. Для этого вектора определим множество N_{k+1}. Оно равно: $N_{k+1} = N_k \setminus \{i_k\} \cup \{j_k\}$. Это означает, что в матрице $A[M, N_k]$ удаляется столбец $\overline{a_{i_k}}$ и на его место ставится столбец $\overline{a_{j_k}}$. В результате получится новая матрица $A[M, N_{k+1}]$. Если $\det A[M, N_{k+1}] \neq 0$, то $x_{k+1}[N]$ – опорный вектор.

Будем строить матрицу $B[N_{k+1}, M]$. , которая нужна на следующем шаге алгоритма. Матрица $A[M, N_k]$ отличается от матрицы $A[M, N_{k+1}]$ только одним столбцом, поэтому положим

$$B[N_{k+1}, M] = F[N_{k+1}, N_k] \cdot B[N_k, M] \tag{5.8}$$

Так как $B[N_{k+1}, M]$ - обратная для $A[M, N_{k+1}]$, то

$B[N_{k+1}, M] \cdot A[M, N_{k+1}] = F[N_{k+1}, N_k] \cdot B[N_k, M] \cdot A[M, N_{k+1}] = E[N_{k+1}, N_{k+1}]$, где E - единичная матрица.

Обозначим через $G[N_k, N_{k+1}] = B[N_k, M] \cdot A[M, N_{k+1}]$. Матрица $A[M, N_{k+1}]$ отличается от матрицы $A[M, N_k]$ столбцом, стоящим на i_k-*м* месте, поэтому матрица $G[N_k, N_{k+1}]$ будет отличаться от единичной только столбцом, стоящим на i_k-*м* месте. Этот столбец получается умножением матрицы $B[N_k, M]$ на j_k–*й* столбец $A[M, N_{k+1}]$ т.е.

$$B[N_k, M] \cdot A[M, j_k] = u_k[N_k] \tag{5.9}$$

Таким образом, матрица

$$G[N_k, N_{k+1}] = \begin{pmatrix} 1 & \dots & u_k[1] & \dots & 0 \\ \vdots & \ddots & . & . & \vdots \\ 0 & . & u_k[i_k] & . & 0 \\ \vdots & . & & \ddots & \vdots \\ 0 & \dots & u_k[m] & \dots & 1 \end{pmatrix}$$

Очевидно, что

$$\det G[N_k, N_{k+1}] = u_k[i_k] > 0, \tag{5.10}$$

так как именно $u_k[i_k]$ стоит в знаменателе выражения, определяющего θ_k. Определитель (5.10) отличен от нуля, поэтому матрица $G[N_k, N_{k+1}]$ имеет обратную, которую легко определить

$$F[N_{k+1}, N_k] = \begin{pmatrix} 1 & \dots & -u_k[1]/u_k[i_k] & \dots & 0 \\ \vdots & \ddots & . & . & \vdots \\ 0 & . & 1/u_k[i_k] & . & 0 \\ \vdots & . & . & \ddots & \vdots \\ 0 & \dots & -u_k[m]/u_k[i_k] & \dots & 1 \end{pmatrix}$$

Теперь легко определить матрицу $B[N_{k+1}, M]$, используя равенство (5.8). Таким образом, матрица $B[N_{k+1}, M]$ – существует, и найденный вектор $x_{k+1}[N]$ является опорным.

Точно так же определяется матрица $B[N_{k+1}, M]$ при замене базиса вырожденного опорного вектора $x_k[N]$, когда $\exists i \in N_k \setminus N_k^+$, $u_k[i] > 0$. В результате смены столбцов матрица $A[M, N_{k+1}]$ отличается от $A[M, N_k]$ только одним столбцом, поэтому можно определить вектор (5.9) и в случае, если $u_k[i_k] \neq 0$, определить новый базис вектора $x_k[N]$.

Приведем блок-схему алгоритма симплекс-метода (рис. 11).

Заметим, что в алгоритме отсутствуют учет явления зацикливания, а также алгоритм построения начального опорного вектора.

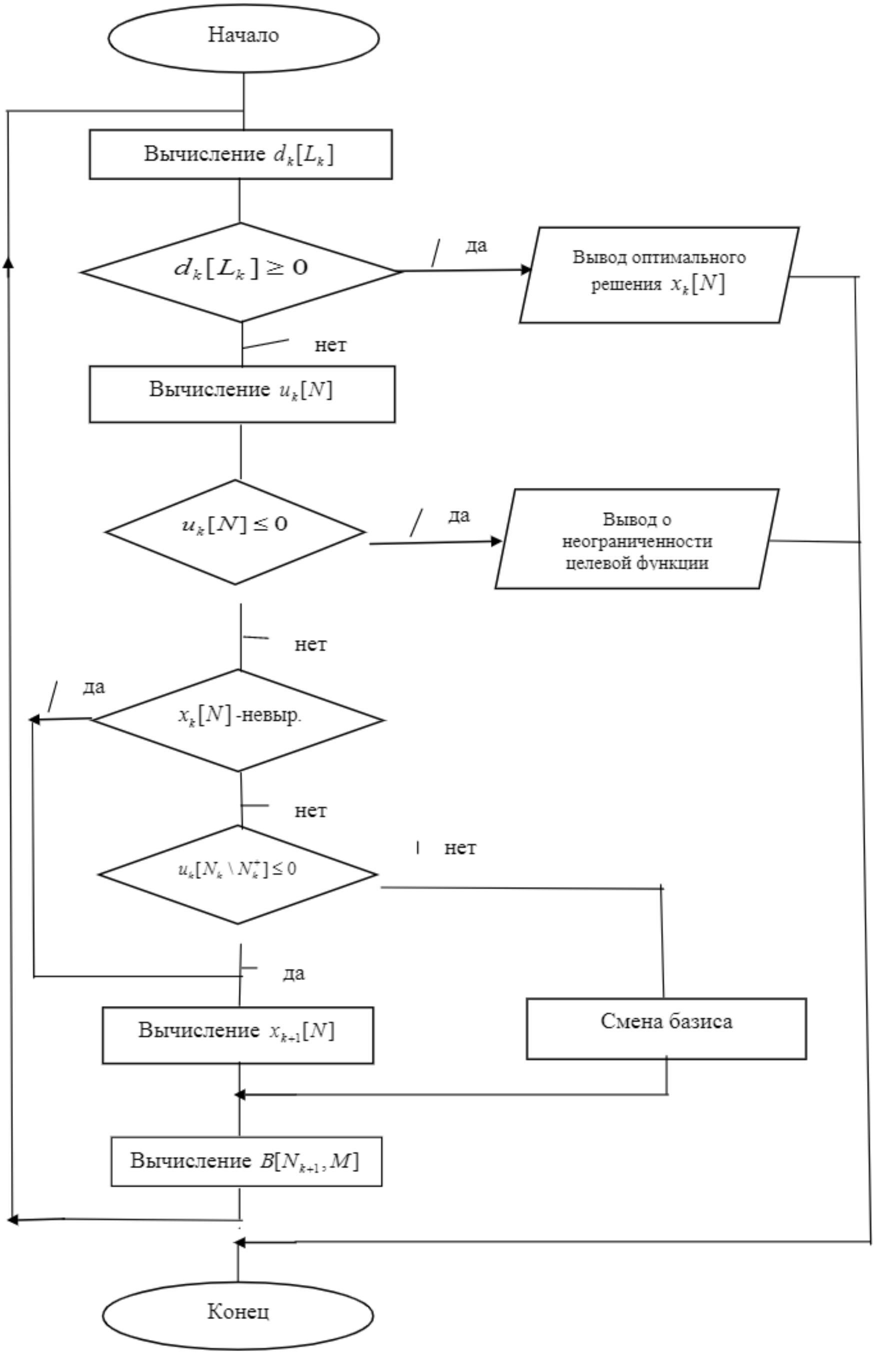
Начало
Вычисление $d_k[L_k]$
$d_k[L_k] \geq 0$
да
Вывод оптимального решения $x_k[N]$
нет
Вычисление $u_k[N]$
$u_k[N] \leq 0$
да
Вывод о неограниченности целевой функции
нет
да
$x_k[N]$ -невыр.
нет
$u_k[N_k \setminus N_k^+] \leq 0$
нет
да
Вычисление $x_{k+1}[N]$
Смена базиса
Вычисление $B[N_{k+1}, M]$
Конец

Рис.11

Рассмотрим построение начального опорного вектора. Можно рекомендовать два способа.

Первый основывается на лемме 4.1. Пользуясь её доказательством, разыскивается первый опорный вектор, с которого и начинается алгоритм симплекс-метода.

Второй заключается в рассмотрении вспомогательной канонической задачи линейного программирования

$$\min \sum_{i \in M} y[i] \tag{5.11}$$

$$A[M,N] \cdot x[N] + E[M,M] \cdot y[M] = b[M]$$

$$x[N] \geq 0, y[M] \geq 0$$

Без ограничения общности можно считать, что $b[M] \geq 0$. Если какой-либо коэффициент $b[i] < 0$, то можно левую и правую части $i-го$ уравнения умножить на -1.

Для задачи (5.11) начальный опорный вектор легко найти : $x[N] = 0, y[M] = b[M] \geq 0$, поэтому можно использовать для нее алгоритм симплекс-метода. Возможны два варианта решения задачи (5.11).

1. Для полученного решения $\overline{x_*}[N], \overline{y_*}[M]$ $\exists i \in M$, для которого $\overline{y_*}[i] > 0$. В этом случае, очевидно, исходная задача (4.25) не имеет ни одной допустимой точки, т.е. $S = \varnothing$. Это также является решением задачи (4.25).

2. В полученном решении $\overline{y_*}[M] = 0$. В этом случае $\overline{x_*}[N]$ является опорным вектором.

Если $\overline{x_*}[N]$ – невырожденный опорный вектор, то его можно использовать в качестве исходного непосредственно.

Если $\overline{x_*}[N]$ – вырожденный опорный вектор, то может оказаться, что в базисе присутствуют единичные столбцы матрицы $E[M,M]$, соответствующие элементам $\bar{x}_*[N]$. В этом случае их необходимо заменить по одному на столбцы матрицы $A[M, N \setminus \overline{N}^+], (\bar{x}_*[\overline{N}^+] > 0)$.

После каждой замены необходимо, конечно, вычислять и матрицу B, пользуясь вышеизложенным алгоритмом.

Литература

1. Акулич И.П. Математическое программирование в примерах и задачах.-Лань, 2011.-352 с.
2. Базара М., Шетти К. Нелинейное программирование.-М.: Мир, 1982.-583 с.
3. Булавский В.А., Звягина .Р.А., Яковлева М.А. Численные методы линейного программирования.-М.:Наука, 1977.-368 с.
4. Васильев Ф.П. Численные методы решения экстремальных задач.-М.:Наука, 1980.
5. Васильев Ф.П., Иваницкий А.Ю. Линейное программирование.-М.: Факториал, 1998.- 176 с.
6. Галеев Э.М. Оптимизация.Теория, примеры, задачи.-Либроком, 2012.-336 с.
7. Гольдштейн Е.Г. Выпуклое программирование. Элементы теории.-Красанд,2009. -72 с.
8. Грешилов А.А. Прикладные задачи математического программирования: Учебное пособие. - 2-е изд. - М.: Логос, 2006. – 288 с.
9. Зангвилл У. Нелигейное программирование.-М.:Советское радио, 1973.-312 с.
10. Ильин В. А., Позняк Э.Г. Линейная алгебра: Учеб.: Для вузов. — 6-е изд., стер. - М.: Физматлит, 2005. - 280 с.
11. Карманов В.Г. Математическое программирование .-М.:Физматлит, 2008.-264 с.
12. Курош А.Г. Курс высшей алгебры.- М.:Наука, 1965.-432 с.

13. Пшеничный Б.Н., Данилин Ю.М.Численные методы в экстремальных задачах. -М.:Наука, 1975.-320 с.
14. Рокафеллар Р. Выпуклый анализ.- М.: Мир, 1973. -471 с.
15. Сухарев А.Г., Тимохов А.В., Фёдоров В.В. Курс методов оптимизации. - М.: Физматлит, 2011. - 368 с.

ОГЛАВЛЕНИЕ

Петухов Л. В.
Серёгин Г. А.
Родионова Е. А.

МЕТОДЫ ОПТИМИЗАЦИИ. ЗАДАЧИ ВЫПУКЛОГО ПРОГРАММИРОВАНИЯ

Учебное пособие

Налоговая льгота – Общероссийский классификатор продукции
ОК 005-93, т. 2; 95 3005 – учебная литература

Подписано в печать 18.06.2014. Формат 60×84/16. Печать цифровая.
Усл. печ. л. 6,25. Тираж 52. Заказ 11995b.

Отпечатано с готового оригинал-макета,
предоставленного Издательством Политехнического университета,
в Типографии Политехнического университета.
195251, Санкт-Петербург, Политехническая ул., 29.
Тел.: (812) 552-77-17; 550-40-14.