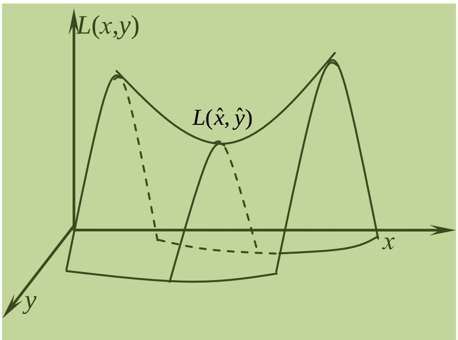
Филиал федерального государственного
бюджетного образовательного учреждения
высшего профессионального образования
«Кемеровский государственный университет»
в г. Анжеро-Судженске

ВЫПУКЛЫЕ ФУНКЦИИ.

ВЫПУКЛОЕ ПРОГРАММИРОВАНИЕ



МИНИСТЕРСТВО ОБРАЗОВАНИЯ И НАУКИ РОССИЙСКОЙ ФЕДЕРАЦИИ

Филиал федерального государственного бюджетного образовательного учреждения высшего профессионального образования «Кемеровский государственный университет» в г. Анжеро-Судженске

ВЫПУКЛЫЕ ФУНКЦИИ. ВЫПУКЛОЕ ПРОГРАММИРОВАНИЕ

Методические указания по решению задач

Анжеро-Судженск
2012

УДК 519.85
ББК 22.1
В92

*Печатается по решению методического совета
филиала Кемеровского государственного университета
в г. Анжеро-Судженске*

Выпуклые функции. Выпуклое программирование: мето-
В92 дические указания по решению задач / Р. Т. Якупов, И. Р. Гарайшина; филиал Кемеровского государственного университета в г. Анжеро-Судженске. — Анжеро-Судженск, 2012. — 37 с.

Данные методические указания содержат необходимые теоретические сведения для усвоения основных понятий, а также методически разобранные примеры и задачи, позволяющие студенту освоить решение задач выпуклого программирования.

Методические указания предназначены для студентов, обучающихся по специальности 080116 «Математические методы в экономике», изучающих курс «Математические методы и модели исследования операций». Могут быть использованы также при самостоятельной работе студентов, обучающихся по другим специальностям и направлениям при изучении дисциплины «Исследование операций».

ББК 22.1

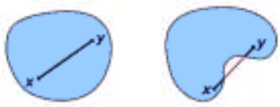
1. Выпуклые множества

Пусть x, y, z — элементы n-мерного действительного евклидова пространства R^n. Будем называть их также векторами или точками пространства R^n.

Определение. *Отрезком*, соединяющим точки x и y, называется множество точек вида

$$z = \lambda x + (1-\lambda)y \ \ (\lambda \in [0,1]).$$

Определение. Множество точек $M \subset R^n$ называется *выпуклым множеством*, если отрезок, соединяющий любые две точки $x, y \in M$, входит в множество M, то есть $\forall \lambda \in [0,1] \ z \in M$ (рис. 1).



а) *б)*

Рис. 1
а) выпуклое множество,
б) невыпуклое множество

Например, выпуклыми множествами являются точка, отрезок, пространство R^n, открытый и замкнутый параллелепипед, открытый и замкнутый шар. Пустое множество не является выпуклым.

Теорема. Непустое пересечение любого числа выпуклых множеств является выпуклым множеством.

Доказательство. Пусть $M_1, M_2, ..., M_k$ — выпуклые множества, и точки x, y принадлежат всем этим множествам одновременно $(x, y \in M_1, ..., x, y \in M_k)$, поэтому

$$x, y \in M = M_1 \bigcap M_2 \bigcap ... \bigcap M_k .$$

Точка $z = \lambda x + (1-\lambda)y$ $(\lambda \in [0,1])$ по определению выпуклого множества принадлежит всем множествам $M_1, M_2, ..., M_k$ одновременно. Таким образом, для любых двух точек $x, y \in M$ точки $z = \lambda x + (1-\lambda)y$ $(\lambda \in [0,1])$ принадлежат множеству M. Поэтому по определению M — выпуклое множество.

Определение. *Гиперплоскостью* в R^n называется множество точек

$$G = \{x : (a, x) = c\},$$

где a — n-мерный *направляющий вектор*, круглые скобки обозначают скалярное произведение $((a, x) = a^T x)$, действительное число c называется свободным членом.

Замечания. 1) Гиперплоскость является выпуклым множеством. Действительно, пусть $x, y \in G$. Тогда для любого $\lambda \in [0,1]$ точка $z = \lambda x + (1-\lambda) y$ принадлежит G, так как

$$(a, z) = (a, \lambda x + (1-\lambda) y) = \lambda (a, x) + (1-\lambda)(a, y) = \lambda c + (1-\lambda) c = c .$$

2) Направляющий вектор a ортогонален гиперплоскости, то есть для любого вектора $z = x - y$, соединяющего две произвольные несовпадающие точки гиперплоскости $(a, z) = 0$. Действительно,

$$(a, z) = (a, x) - (a, y) = c - c = 0.$$

Определение. Множество точек вида

$$S = \{x : (a, x) \le c\}$$

называется *полупространством* в R^n.

Направление неравенства в определении можно взять и противоположным.

Замечание. Полупространство является выпуклым множеством. Действительно, пусть $x, y \in S$. Тогда для любого $\lambda \in [0,1]$ точка $z = \lambda x + (1-\lambda) y$ принадлежит S, так как

$$(a, z) = (a, \lambda x + (1-\lambda) y) = \lambda (a, x) + (1-\lambda)(a, y) \le \lambda c + (1-\lambda) c = c .$$

Определение. Непустое пересечение конечного числа полупространств называется *выпуклым многогранником*.

Применение термина выпуклый многогранник объясняется тем, что полупространство — выпуклое множество, а непустое пересечение конечного числа выпуклых множеств есть выпуклое множество.

Определение. Множество вида

$$S = \{x : x \ge 0\}$$

называется *положительным ортантом*.

Положительный ортант есть выпуклый многогранник. Действительно, неравенство $x \ge 0$ можно интерпретировать как систему неравенств

$$(e^{(1)},\, x) \ge 0,$$
$$\cdots\cdots\cdots\cdots$$
$$(e^{(n)},\, x) \ge 0,$$

где $e^{(1)} = (1,\, 0,\, ...,\, 0),\ \ldots\ ,\ e^{(n)} = (0,\, 0,\, ...,\, n).$

Определение. Пусть выпуклый многогранник G задан системой неравенств

$$(a^{(1)},\, x) \ge c_1,$$
$$\cdots\cdots\cdots\cdots$$
$$(a^{(k)},\, x) \ge c_k,$$

где $a^{(1)},...,\, a^{(k)}$ — направляющие векторы, $k > n$. Если точка $y \in G$ обращает в равенства не менее n неравенств, причем ранг соответствующей системы векторов равен n, то точка y называется *угловой* (или крайней) точкой многогранника.

Отметим, что число угловых точек выпуклого многогранника может быть (в зависимости от n и k) очень большим. Так, при $n = 10$, $k = 20$ это число может быть сравнимо с 10^{11}.

Замечание. Так как равенство вида $(a^{(i)},\, x) = c_i$ можно заменить системой двух неравенств

$$(a^{(i)},\, x) \ge c_i,$$
$$(-a^{(i)},\, x) \ge -c_i,$$

то если в определении часть неравенств (или все неравенства) заменить соответствующими равенствами, то получающаяся система условий также определяет выпуклый многогранник.

Напомним определение часто используемого выпуклого множества.

Определение. ε – *окрестностью* точки $a \in R^n$ называется открытый шар

$$O_\varepsilon(a) = \{x : \|x - a\| < \varepsilon\}.$$

Очевидно, ε – окрестность точки есть выпуклое множество.

Определение. Точка x называется *граничной точкой* множества $X \subset R^n$, если $\forall\ \varepsilon > 0$ ε-окрестность $O_\varepsilon(x)$ содержит точки, принадлежащие множеству X и точки, не принадлежащие множеству X.

Определение. Точка x называется *внутренней точкой* множества $X \subset R^n$, если найдется $\varepsilon > 0$, что ε-окрестность $O_\varepsilon(x)$ целиком лежит внутри множества X.

Замечание. Граничная точка может и не принадлежать множеству X. Например, для множества

$$X = \{x : 0 \le x < 1,\ x \in R^1\}$$

граничная точка $x = 0$ принадлежит X, а граничная точка $x = 1$ не принадлежит X.

Определение. Множество X называется *замкнутым*, если оно содержит все свои граничные точки.

Определение. Множество X называется *открытым*, если оно не содержит свои граничные точки.

Примером замкнутого множества в R^n является замкнутый шар $X = \{x : \|x - a\| \le r\}$ радиуса r с центром в точке a. Примером открытого множества в R^n является открытый шар $X = \{x : \|x - a\| < r\}$.

Определение. *Диаметром* множества X называется число $\operatorname{diam} X = \sup\limits_{x \in X,\, y \in X} \|x - y\|$. В случае замкнутого множества символ sup можно заменить на символ max.

Определение. Множество X называется *ограниченным*, если его диаметр является конечным числом.

Определение. *Конусом* $K \subset R^n$ называется такое множество, что из того, что $x \in K$ следует, что $\forall\ \lambda \ge 0$ $\lambda x \in K$.

Замечание. Из определения следует, что конус содержит нулевую точку $x = 0$. Конус является неограниченным множеством (за исключением вырожденного случая, когда конус содержит всего лишь одну точку $x = 0$). Конус может быть как замкнутым, так и незамкнутым множеством.

Определение. *Компактом* называется замкнутое ограниченное множество.

Замечание. Замкнутые ограниченные множества представляют особый интерес в связи с теоремой Вейерштрасса, которая утверждает, что непрерывная функция на замкнутом ограниченном множестве (компакте) достигает своего наибольшего и наименьшего значений.

2. Выпуклые функции

2.1. Определение выпуклых функций

Определение. Функция $f(x)$ называется *выпуклой* (выпуклой вниз) на выпуклом множестве $X \subseteq R^n$, если $\forall\, x, y \in X$ и $\forall\, \lambda \in [0,1]$ выполняется неравенство

$$f(\lambda x + (1-\lambda)y) \le \lambda f(x) + (1-\lambda) f(y). \tag{1}$$

Определение. Функция $f(x)$ называется *вогнутой* (выпуклой вверх) на выпуклом множестве $X \subseteq R^n$, если $\forall\, x, y \in X$ и $\forall\, \lambda \in [0,1]$ выполняется неравенство

$$f(\lambda x + (1-\lambda)y) \ge \lambda f(x) + (1-\lambda) f(y). \tag{1$'$}$$

Замечания. 1) Очевидно, функция $f(x)$ вогнута, если функция $-f(x)$ выпукла.

2) Если в (1) неравенство строгое, то функция $f(x)$ называется *строго выпуклой*. Если в (1′) неравенство строгое, то функция $f(x)$ называется *строго вогнутой*.

При доказательстве некоторых утверждений полезным является другое равносильное определение, которое сформулируем в следующей теореме.

Теорема. Для того чтобы функция $f(x)$ была выпукла на выпуклом множестве *X*, необходимо и достаточно, чтобы $\forall\, x, y \in X$ выполнялось неравенство

$$f\left(\tfrac{1}{2}x + \tfrac{1}{2}y\right) \le \tfrac{1}{2}[f(x) + f(y)]. \tag{2}$$

2.2. Свойства выпуклых функций

Теорема. Сумма выпуклых на выпуклом множестве X функций является выпуклой функцией.

Доказательство. Пусть функции $f(x)$ и $g(x)$ — выпуклые функции на множестве *X*. То есть $\forall\, x, y \in X$ и $\forall \lambda \in [0,1]$ выполняются неравенства

$$f(\lambda x + (1-\lambda) y) \le \lambda f(x) + (1-\lambda) f(y),$$
$$g(\lambda x + (1-\lambda) y) \le \lambda g(x) + (1-\lambda) g(y).$$

Сложив эти неравенства, для функции $h(x) = f(x) + g(x)$ получаем:

$$h(\lambda x + (1-\lambda) y) \le \lambda h(x) + (1-\lambda) h(y).$$

Заметим, что теорема будет справедливой и в случае суммы произвольного конечного числа выпуклых функций.

Замечание. Легко видеть, что в n-мерном евклидовом пространстве R^n линейная функция $l(x) = (p, x) + c$, где p — некоторый заданный вектор, c — действительное число, удовлетворяет одновременно неравенствам (1) и ($1'$). То есть линейная функция является одновременно и выпуклой, и вогнутой функцией.

Теорема. Для того чтобы квадратичная функция $f(x) = (x, Bx),$ где B — симметрическая матрица, была выпуклой в R^n, необходимо и достаточно, чтобы матрица B была положительно определенной.

Доказательство. Рассмотрим

$$f(\lambda x + (1-\lambda) y) = (\lambda x + (1-\lambda) y,\ B(\lambda x + (1-\lambda) y)) =$$
$$= \lambda^2 (x, Bx) + 2\lambda(1-\lambda)(x, By) + (1-\lambda)^2 (y, By).$$

Добавим и вычтем в правой части

$$\lambda f(x) + (1-\lambda) f(y) = \lambda(x, Bx) + (1-\lambda)(y, By),$$

получим:

$$f(\lambda x + (1-\lambda) y) = \lambda f(x) + (1-\lambda) f(y) - \lambda(1-\lambda)(x, Bx) +$$
$$+ 2\lambda(1-\lambda)(x, By) + \lambda(1-\lambda)(y, By) =$$
$$= \lambda f(x) + (1-\lambda) f(y) - \lambda(1-\lambda)(x - y, B(x - y)). \qquad (3)$$

1) *Необходимость*. Пусть функция $f(x)$ выпукла. Тогда согласно (1)

$$\lambda(1-\lambda)(x-y, B(x-y)) \geq 0,$$

то есть матрица B положительно определена.

2) *Достаточность*. Пусть матрица B положительно определена. Тогда с учётом (3) будет выполняться неравенство (1), то есть функция $f(x)$ выпукла.

Следствие. Для того чтобы квадратичная функция $f(x)=(x, Bx)+(p, x)+c,$ где B — симметрическая матрица, была выпуклой в R^n, необходимо и достаточно, чтобы матрица B была положительно определенной.

Доказательство. Поскольку для линейной функции $l(x)=(p, x)+c \quad \forall\, x, y \in X$ и $\forall\, \lambda \in [0,1]$ выполняется равенство

$$l(\lambda x+(1-\lambda)y)=\lambda l(x)+(1-\lambda)l(y),$$

то для рассматриваемой квадратичной функции справедливо равенство (3). Поэтому доказательство следствия совпадает с доказательством теоремы.

Замечание. Очевидно, что для строгой выпуклости квадратичной функции необходимо и достаточно, чтобы матрица B была строго положительно определенной.

Теорема. Выпуклая на выпуклом множестве X функция $f(x)$ является непрерывной в каждой внутренней точке множества X.

Теорема. Выпуклая на выпуклом множестве X функция $f(x)$ имеет в каждой внутренней точке $x \in X$ производную по любому направлению a:

$$\frac{\partial f(x)}{\partial a}=\lim_{h\to+0}\frac{f(x+ha)-f(x)}{h}.$$

2.3. Неравенство Йенсена

Теорема. Если функция $f(x)$ выпукла на выпуклом множестве X, и $x_1,...,x_m \in X$, $\lambda_i \geq 0\ (i=\overline{1,m})$, $\sum_{i=1}^{m}\lambda_i = 1$, то

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

Доказательство. По определению выпуклой функции

$$f\left(\sum_{i=1}^{m}\lambda_i x_i\right) = f\left(\lambda_1 x_1 + (1-\lambda_1)\sum_{i=2}^{m}\frac{\lambda_i}{(1-\lambda_1)}x_i\right) \leq$$

$$\leq \lambda_1 f(x_1) + (1-\lambda_1) f\left(\sum_{i=2}^{m}\lambda_i' x_i\right),$$

где $\lambda_i' = \dfrac{\lambda_i}{(1-\lambda_1)}$.

Аналогичным образом получаем:

$$f\left(\sum_{i=2}^{m}\lambda_i' x_i\right) \leq \lambda_2' f(x_2) + (1-\lambda_2') f\left(\sum_{i=3}^{m}\lambda_i^{(2)} x_i\right),$$

где $\lambda_i^{(2)} = \dfrac{\lambda_i'}{(1-\lambda_2')}$,

$$f\left(\sum_{i=3}^{m}\lambda_i^{(2)} x_i\right) \leq \lambda_3^{(2)} f(x_3) + (1-\lambda_3^{(2)}) f\left(\sum_{i=4}^{m}\lambda_i^{(3)} x_i\right),$$

где $\lambda_i^{(3)} = \dfrac{\lambda_i^{(2)}}{(1-\lambda_3^{(2)})}$,

. . . .

$$f\left(\sum_{i=m-1}^{m}\lambda_i^{(m\text{-}2)} x_i\right) \leq \lambda_{m\text{-}1}^{(m\text{-}2)} f(x_{m-1}) + (1-\lambda_{m\text{-}1}^{(m\text{-}2)}) f(x_m),$$

где $\lambda_i^{(m-1)} = \dfrac{\lambda_i^{(m\text{-}2)}}{(1-\lambda_{m-1}^{(m\text{-}2)})}$.

Из этой цепочки неравенств получаем формулу

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

2.4. Условия выпуклости первого порядка

Теорема. Для того чтобы дифференцируемая на выпуклом замкнутом множестве X функция $f(x)$ была выпуклой, необходимо и достаточно, чтобы $\forall x_0\in X$ и $x\in X$ выполнялось неравенство

$$(\nabla f(x_0),\ x-x_0)\le f(x)-f(x_0). \tag{1}$$

Замечание. Запишем (1) в виде

$$f(x_0)+(\nabla f(x_0),\ x-x_0)\le f(x).$$

Слева записано выражение для касательной гиперплоскости к поверхности $y=f(x)$ (или к графику функции $y=f(x)$). Поэтому теорему можно переформулировать следующим образом.

Теорема. Для того чтобы дифференцируемая на выпуклом замкнутом множестве X функция $f(x)$ была выпуклой, необходимо и достаточно, чтобы график функции $f(x)$ на множестве X располагался не ниже касательной гиперплоскости в точке x_0.

Доказательство. *Необходимость*. Пусть функция $f(x)$ выпукла на множестве X. Запишем неравенство из определения выпуклости в несколько ином виде. Для этого рассмотрим

$$f(x_0+h(x-x_0))=f(hx+(1-h)x_0)\le hf(x)+(1-h)f(x_0)=$$
$$=f(x_0)+h[f(x)-f(x_0)].$$

Таким образом, имеем:

$$f(x_0+h(x-x_0))\le f(x_0)+h[f(x)-f(x_0)],\ h\in[0,1].$$

Перенесем $f(x_0)$ в левую часть и разделим полученное неравенство на $h\ne 0$:

$$\frac{f(x_0+h(x-x_0))-f(x_0)}{h}\le f(x)-f(x_0). \tag{2}$$

После предельного перехода при $h \to +0$ в левой части получим производную по направлению вектора $x - x_0$:

$$\frac{\partial f(x_0)}{\partial (x - x_0)} = \frac{\partial f(x_0)}{\partial x_1}(x - x_{01}) + \ldots + \frac{\partial f(x_0)}{\partial x_n}(x - x_{0n}) =$$

$$= (\nabla f(x_0),\ x - x_0).$$

Поэтому неравенство (2) примет вид (1).

Достаточность. Пусть имеет место неравенство (1). Построим функцию одной переменной

$$u(h) = f(x_0 + h(x - x_0)) = f(hx + (1-h)x_0).$$

Производная этой функции

$$u'(h) = \sum_{i=1}^{n} \frac{\partial f(x_0 + h(x - x_0))}{\partial x_i}(x_i - x_{0i}) = (\nabla f(x_0 + h(x - x_0)),\ x - x_0). \quad (3)$$

Возьмем $h_2 > h_1$. Рассмотрим две точки:

$$y_0 = x_0 + h_1(x - x_0); \quad y = x_0 + h_2(x - x_0).$$

Из (1) имеем:

$$(\nabla f(y_0),\ y - y_0) = ((\nabla f(x_0 + h_1(x - x_0)),\ (h_2 - h_1)(x - x_0)) \le$$

$$\le f(y) - f(y_0) = u(h_2) - u(h_1),$$

откуда с учётом (3) получаем:

$$u'(h_1)(h_2 - h_1) \le u(h_2) - u(h_1).$$

Аналогично после замены $h_1 \to h_2,\ h_2 \to h_1$ получаем:

$$u'(h_2)(h_1 - h_2) \le u(h_1) - u(h_2).$$

Складывая эти два неравенства, имеем:

$$[u'(h_1) - u'(h_2)](h_2 - h_1) \le 0.$$

Так как $h_2 > h_1$, то $u'(h_2) \ge u'(h_1)$, то есть производная $u'(h)$ монотонно возрастает. А это означает выпуклость функции $u(h)$.

Для завершения доказательства рассмотрим

$$f(hx + (1-h)x_0) = f(x_0 + h(x - x_0)) = u(h) =$$

$$= u(h \cdot 1 + (1-h) \cdot 0) \le hu(1) + (1-h)u(0) = hf(x) + (1-h) f(x_0).$$

Так как это неравенство верно для всех $x,\ x_0 \in X$ и $h \in [0,1]$, то функция $f(x)$ является выпуклой на множестве X.

2.5. Примеры выпуклых функций

1. Пусть функция $f(x)$ выпукла на выпуклом множестве X, и $f(x) \geq 0 \ \forall x \in X$. Тогда функция $f^2(x)$ также выпукла на множестве X.

Для доказательства этого утверждения рассмотрим

$$f^2(\tfrac{1}{2}x + \tfrac{1}{2}y) \leq [\tfrac{1}{2}f(x) + \tfrac{1}{2}f(y)]^2 =$$
$$= \tfrac{1}{2}f^2(x) + \tfrac{1}{2}f^2(y) - \tfrac{1}{4}[f^2(x) - 2f(x)f(y) + f^2(y)].$$

Так как выражение внутри квадратной скобки неотрицательно, то

$$f^2(\tfrac{1}{2}x + \tfrac{1}{2}y) \leq \tfrac{1}{2}f^2(x) + \tfrac{1}{2}f^2(y),$$

что и означает выпуклость функции $f^2(x)$.

2. Пусть функции $f_1(x), ..., f_m(x)$ выпуклы на выпуклом множестве X, тогда функция $h(x) = \max(f_1(x), ..., f_m(x))$ также выпукла на множестве X.

Достаточно доказать утверждение для случая $m = 2$. Для общего случая доказательство затем легко проводится методом математической индукции.

Рассмотрим

$$h(\tfrac{1}{2}x + \tfrac{1}{2}y) = \max[f_1(\tfrac{1}{2}x + \tfrac{1}{2}y),\ f_2(\tfrac{1}{2}x + \tfrac{1}{2}y)] \leq$$
$$\leq \max[\tfrac{1}{2}f_1(x) + \tfrac{1}{2}f_1(y),\ \tfrac{1}{2}f_2(x) + \tfrac{1}{2}f_2(y)] \leq$$
$$\leq \tfrac{1}{2}\max[f_1(x),\ f_2(x)] + \tfrac{1}{2}\max[f_1(y),\ f_2(y)] = \tfrac{1}{2}h(x) + \tfrac{1}{2}h(y).$$

Таким образом, $h(\tfrac{1}{2}x + \tfrac{1}{2}y) \leq \tfrac{1}{2}h(x) + \tfrac{1}{2}h(y)$, то есть функция $h(x)$ выпукла.

2.6. Условия выпуклости второго порядка

Теорема. Пусть функция $f(x)$ дважды дифференцируема на выпуклом множестве X. Тогда для того чтобы функция $f(x)$ была выпуклой на множестве X необходимо и достаточно, чтобы матрица вторых производных $H(x)$ функции $f(x)$ (матрица Гессе) была положительно определенной для всех $x \in X$.

Доказательство. *Необходимость*. Пусть функция $f(x)$ выпукла на множестве X, и точки $x_0,\ x \in X,\ \lambda \in [0,1]$. Обозначим $\Delta x = x - x_0$. Тогда функцию $f(x_0 + \lambda x)$ одной переменной λ можно разложить по формуле Тейлора (Маклорена)

$$f(x_0 + \lambda x) - f(x_0) = \lambda \nabla^T f(x_0)\Delta x + \tfrac{1}{2}\lambda^2 (\Delta x)^T H(x_0)\Delta x + o(\lambda^2). \quad (1)$$

По необходимому и достаточному условию выпуклости первого порядка

$$\lambda \nabla^T f(x_0)\Delta x \le f(x_0 + \lambda x) - f(x_0).$$

Поэтому $\frac{1}{2}\lambda^2 (\Delta x)^T H(x_0)\Delta x + o(\lambda^2) \ge 0.$ Поделим это неравенство на $\lambda^2 \ne 0$ и перейдем к пределу при $\lambda \to +0$, после чего получим неравенство

$$(\Delta x)^T H(x_0)\Delta x \ge 0,$$

верное для всех $x_0,\ x \in X$. Отсюда следует, что матрица H положительно определена.

Достаточность. Пусть матрица H положительно определена для всех $x \in X$. Тогда по формуле Тейлора с остаточным членом в форме Лагранжа для всех $x_0,\ x \in X$

$$f(x) - f(x_0) = \nabla^T f(x_0)\Delta x + \tfrac{1}{2}(\Delta x)^T H(x_0 + \theta \Delta x)\Delta x, \text{ где } \theta \in [0,1].$$

Поэтому $f(x) - f(x_0) \ge \nabla^T f(x_0)(x - x_0)$. Отсюда по необходимому и достаточному условию выпуклости первого порядка следует, что функция $f(x)$ выпукла на множестве X.

2.7. Множества, определяемые неравенствами с выпуклыми функциями

Теорема. Множество вида

$$D = \{x :\ f(x) \le b\},$$

где $f(x)$ — выпуклая функция, $b \in R,$ является выпуклым.

Доказательство. Пусть $x \in D,\ y \in D,$ то есть

$$f(x) \le b,\ f(y) \le b.$$

В силу выпуклости $f(x)$ для $z = \lambda x + (1-\lambda)y$, где $\lambda \in [0,1]$, имеем неравенство

$$f(z) \le \lambda f(x) + (1-\lambda) f(y) \le \lambda b + (1-\lambda)b = b.$$

Полученное неравенство $f(z) \le b$ означает, что $z \in D$. Отсюда по определению выпуклого множества следует, что D — выпуклое множество.

Следствие 1. Множество вида

$$D = \{x: \ f_1(x) \le b_1, \ ..., \ f_m(x) \le b_m\},$$

где $f_1(x), ..., f_m(x)$ — выпуклые функции, является выпуклым.

Доказательство следует из того, что множества $D_i = \{x: \ f_i(x) \le b_i\} \ (i = \overline{1,m})$ являются выпуклыми, а D есть их пересечение: $D = D_1 \cap D_2 ... \cap D_m$.

Следствие 2. Множество вида

$$D = \{x: \ f_i(x) \le b_i \ (i = \overline{1,m}); \ (a_i, x) = b_i \ (i = \overline{m+1, m+k})\},$$

где $f_1(x), ..., f_m(x)$ — выпуклые функции, $b_i \in R \ (i = \overline{1, m+k})$, $a_i \in R^n$, является выпуклым.

Доказательство следует из того, что множества $D_i \ (i = \overline{1,m})$ — выпуклые, множество

$$\Delta = \{x: (a_i, x) = b_i \ (i = \overline{m+1, m+k})\} \ —$$

выпуклый многогранник, а D есть пересечение этих выпуклых множеств: $D = D_1 \cap D_2 ... \cap D_m \cap \Delta$.

2.8. Экстремальные свойства функций на выпуклых множествах

Определение. Направление, определяемое вектором $a \ne 0$ в точке $x \in X$, где X — выпуклое множество, называется *возможным* (или допустимым), если существует такое число $\lambda^* > 0$, что $\forall \lambda \in [0, \lambda^*]$ точка $x + \lambda a \in X$.

Замечание. Очевидно, если x — внутренняя точка множества X, то в этой точке любое направление является возможным (см. определение внутренней точки).

Теорема. Для того чтобы точка $\hat{x}$ выпуклого множества X была точкой локального минимума дифференцируемой на X функции $f(x)$, необходимо, чтобы существовала окрестность $O_\delta(\hat{x})$ точки $\hat{x}$ такая, что $\forall x \in O_\delta(x) \cap X$ выполнялось неравенство $\nabla^T f(\hat{x})(x-\hat{x}) \geq 0$.

Доказательство. Пусть $\hat{x}$ — точка локального минимума, то есть существует окрестность $O_\delta(\hat{x})$ точки $\hat{x}$ такая, что $\forall x \in O_\delta(x) \cap X$ выполняется неравенство $f(x) \geq f(\hat{x})$. Для краткости введем обозначение $U = O_\delta(x) \cap X$.

Так как множество U выпуклое, то отрезок, соединяющий точки x и $\hat{x}$ из U, будет принадлежать множеству U.

Вектор $a = \dfrac{1}{\|x-\hat{x}\|}(x-\hat{x})$ имеет единичную норму $\|a\| = 1$ и задает допустимое направление. Положим $\lambda^* = \dfrac{\delta}{\|a\|}$, где возьмем $0 < \delta < 1$. Тогда $\forall \lambda \in (0, \lambda^*)$

$$\hat{x} + \lambda a = \lambda x + (1-\lambda)\hat{x} \in U.$$

Поэтому $f(\hat{x}+\lambda a) - f(\hat{x}) \geq 0$, и

$$\lim_{\lambda \to +0} \frac{1}{\lambda}[f(\hat{x}+\lambda a) - f(\hat{x})] = \frac{\partial f(\hat{x})}{\partial a} = \nabla^T f(\hat{x})(x-\hat{x}) \geq 0.$$

Замечание. Таким образом, для того чтобы точка $\hat{x}$ выпуклого множества X была точкой локального минимума дифференцируемой на X функции $f(x)$, необходимо, чтобы в точке $\hat{x}$ производные по всем возможным направлениям были неотрицательными.

Теорема. Локальный минимум выпуклой функции $f(x)$ на выпуклом множестве X является её глобальным минимумом.

Доказательство. Воспользуемся методом математической индукции. Предположим, что $\hat{x}$ — точка локального минимума, а $\tilde{x}$ — не совпадающая с $\hat{x}$ точка глобального минимума, то есть

$$f(\hat{x}) > f(\tilde{x}). \qquad (1)$$

По определению выпуклой функции

$$f(\lambda\hat{x} + (1-\lambda)\tilde{x}) \le \lambda f(\hat{x}) + (1-\lambda) f(\tilde{x}),\ \lambda \in [0,1]. \qquad (2)$$

Заменим в (2) в правой части $f(\tilde{x})$ на $f(\hat{x})$. Тогда с учётом (1) получим строгое неравенство

$$f(\lambda\hat{x} + (1-\lambda)\tilde{x}) < \lambda f(\hat{x}) + (1-\lambda) f(\hat{x}) = f(\hat{x}),\ \lambda \in [0,1).$$

Так как это неравенство выполняется для λ, сколь угодно близких к 1, то в сколь угодно малой окрестности $O_\delta(\hat{x})$ точки $\hat{x}$ существуют точки $x' = \lambda\hat{x} + (1-\lambda)\tilde{x} \in O_\delta(\hat{x})$, в которых $f(x') < f(\hat{x})$, что противоречит тому, что $\hat{x}$ — точка локального минимума.

Теорема. Множество точек глобального минимума выпуклой на выпуклом множестве X функции $f(x)$ является выпуклым.

Доказательство. Пусть x_1, x_2 — две точки, в которых достигается глобальный минимум: $f(x_1) = f(x_2)$. По определению выпуклой функции

$$f(\lambda x_1 + (1-\lambda)x_2) \le \lambda f(x_1) + (1-\lambda) f(x_2) = f(x_1).$$

Так как x_1 — точка глобального минимума, то не может быть точки из X, для которой

$$f(\lambda x_1 + (1-\lambda)x_2) < f(x_1).$$

Поэтому

$$f(\lambda x_1 + (1-\lambda)x_2) = f(x_1).$$

Таким образом, точки $\lambda x_1 + (1-\lambda)x_2$ принадлежат множеству точек глобального минимума. Поэтому это множество выпукло.

Теорема. Строго выпуклая на выпуклом множестве X функция $f(x)$ имеет единственную точку минимума.

Доказательство. Докажем утверждение методом от противного. Предположим, что есть две точки глобального минимума x_1 и x_2 из множества X, в которых значение функции равно f^*, то есть

$$f(x_1) = f(x_2) = f^*.$$

По определению строго выпуклой функции $\forall\ \lambda \in (0,1)$

$$f(\lambda x_1 + (1-\lambda)x_2) < \lambda f(x_1) + (1-\lambda) f(x_2) =$$
$$= \lambda f(x_1) + (1-\lambda) f(x_1) = f^*.$$

Таким образом, есть точки из X, в которых значение функции $f(x)$ меньше f^*. Это противоречит определению глобального минимума.

Теорема. Если для выпуклой на выпуклом множестве X функции $f(x)$ в точке $\hat{x}$ все частные производные равны нулю, то $\hat{x}$ — точка глобального минимума.

Доказательство. Из условия выпуклости 1-го порядка

$$\forall x \in X \quad f(x) - f(\hat{x}) \geq \nabla^T f(\hat{x})(x - \hat{x}) = 0,$$

то есть $\forall x \in X \quad f(x) \geq f(\hat{x}),$ а это означает, что $\hat{x}$ — точка глобального минимума.

2.9. Свойства максимумов выпуклой функции

Свойства максимума выпуклой функции очень сильно отличаются от свойств её минимума. Отметим одно из важных свойств максимума выпуклой функции.

Теорема. Если $f(x)$ — выпуклая на выпуклом замкнутом ограниченном множестве X функция, то

1) множество точек глобального максимума функции $f(x)$ на множестве X содержит хотя бы одну крайнюю точку множества X;

2) если существует хотя бы один локальный максимум во внутренней точке множества X, то функция $f(x)$ постоянна на множестве X.

Замечание. Так как в случае вогнутой функции $f(x)$ функция $(-f(x))$ является выпуклой, то для точек минимума вогнутой функции $f(x)$ на выпуклом замкнутом ограниченном множестве X справедлива аналогичная теорема.

2.10. Свойства максимумов и минимумов линейной функции

Так как линейная функция является одновременно выпуклой и вогнутой функцией, то из приведенных выше экстремальных свойств выпуклых и вогнутых функций вытекает следующее утверждение.

Теорема. Если $f(x)$ — линейная функция (но не постоянная), заданная на выпуклом замкнутом ограниченном многограннике X, то минимумы (максимумы) $f(x)$ достигаются в угловых точках. Если при этом точка минимума (максимума) не единственна, то множество точек минимума (максимума) является выпуклым. Все минимумы (максимумы) являются глобальными.

Таким образом, при поиске минимума (максимума) линейной функции на выпуклом замкнутом ограниченном многограннике достаточно рассмотреть угловые точки многогранника.

3. Выпуклое программирование

3.1. Задачи выпуклого программирования

Пусть функция $f(x)$ выпукла на выпуклом замкнутом множестве X, которое определяется следующим образом:

$$X = \{x : g(x) \le b,\ x \in X_0\}, \qquad (1)$$

где $x \in R^n$, X_0 — выпуклое замкнутое множество из R^n,

$$g(x) = (g_1(x), \dots, g_m(x)^T),\ b = (b_1, \dots, b_m)^T,$$

функции $g_i(x)$ $(i = \overline{1, m})$ выпуклы на множестве X_0.

Определение. Задача

$$\min_{x \in X} f(x), \qquad (2)$$

где X — выпуклое замкнутое множество вида (1), $f(x)$ выпуклая на множестве X функция, называется *задачей выпуклого программирования*.

Замечание. Прежде чем приступать к решению задачи (2), нужно убедиться, что

1) множество X не пусто;

2) функция $f(x)$ не является неограниченной снизу на множестве X.

3.2. Условия регулярности

Определение 1. Если $\forall\ i = \overline{1,m}$ существуют точки $x_i \in X$ такие, что

$$g_i(b_i) < b_i, \tag{1}$$

то говорят, что множество X удовлетворяет *условию регулярности*.

Определение 2. Если существует точка $x \in X$ такая, что

$$g(b) < b, \tag{2}$$

то говорят, что множество X удовлетворяет *условию регулярности Слейтера*.

Теорема. Условия регулярности в смысле определения 1 и условия регулярности по Слейтеру из определения 2 равносильны.

Доказательство. 1) Пусть $\forall\ i = \overline{1,m}$ существуют точки $x_i \in X$ такие, что $g_i(b_i) < b_i$. Построим выпуклую линейную комбинацию этих точек:

$$x = \sum_{i=1}^{m} \lambda_i x_i, \quad \sum_{i=1}^{m} \lambda_i = 1, \quad \lambda_i \ge 0 \ \ (i = \overline{1,m}).$$

Воспользовавшись неравенством Йенсена, получим:

$$g_i(x) = g_i\left(\sum_{j=1}^{m} \lambda_j x_j\right) \le \sum_{j=1}^{m} \lambda_j g_i(x_j) < \sum_{j=1}^{m} \lambda_j b_i = b_i.$$

Таким образом, существует точка $x \in X$ такая, что $g(b) < b$.

2) Пусть теперь существует точка $x \in X$ такая, что $g(b) < b.$ Положим $x_1 = x_2 = ... = x_m = x.$ Тогда из предыдущего неравенства имеем:

$$g_i(b_i) < b_i \ \ (i = \overline{1, m}).$$

3.3. Условия минимума в задаче выпуклого программирования. Функция Лагранжа

Определение. Функция

$$L(x, y) = f(x) + (y,\ g(x) - b), \tag{1}$$

где $x \in X_0,\ y \geq 0,$ называется *функцией Лагранжа* для задачи выпуклого программирования.

Определение. Пара $(\hat{x},\ \hat{y})$ называется *седловой точкой* функции Лагранжа $L(x, y)$ (рис. 2) на множестве $x \in X_0,\ y \geq 0,$ если

1) $\hat{x} \in X_0,\ \hat{y} \geq 0;$

2) $L(\hat{x}, y) \leq L(\hat{x}, \hat{y}) \leq L(x, \hat{y})$ для всех $x \in X_0,\ y \geq 0.$ (2)

Замечание. Условие (2) можно записать также в виде:

$$L(\hat{x},\ \hat{y}) = \min_{x \in X_0} \max_{y \geq 0} L(x, y) = \max_{y \geq 0} \min_{x \in X_0} L(x, y). \tag{3}$$

Теорема. Если пара $(\hat{x},\ \hat{y})$ является седловой точкой функции Лагранжа на множестве $x \in X_0,\ y \geq 0,$ то $\hat{x}$ — оптимальная точка задачи выпуклого программирования.

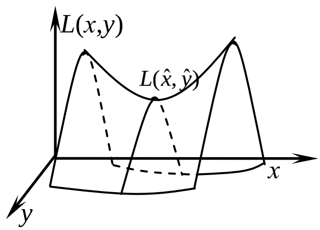


Рис. 2. Седловая точка функции Лагранжа

Доказательство. Пусть пара $(\hat{x},\ \hat{y})$ — седловая точка. Обозначим

$$r(x) = g(x) - b,\ \ r(x) \leq 0\ \ \forall\, x \in X.$$

С учётом этого и вида функции Лагранжа неравенство (2) запишем в виде:

$$f(\hat{x}) + (y,\ r(\hat{x})) \leq f(\hat{x}) + (\hat{y},\ r(\hat{x})) \leq f(x) + (\hat{y},\ r(x)). \tag{4}$$

1) Из левого неравенства в (4) имеем:

$$(y,\, r(\hat{x})) \le (\hat{y},\, r(\hat{x}))\ \forall\, y \ge 0,$$

откуда при $y = 0$ получаем:

$$(\hat{y},\, r(\hat{x})) \ge 0. \tag{5}$$

С другой стороны, так как в скалярном произведении $(\hat{y},\, r(\hat{x}))$ $\hat{y} \ge 0,\ r(\hat{x}) \le 0,$ то скалярное произведение неположительное:

$$(\hat{y},\, r(\hat{x})) \le 0. \tag{$5'$}$$

Из сравнения (5) и ($5'$) следует, что

$$(\hat{y},\, r(\hat{x})) = 0. \tag{6}$$

2) Из правого неравенства в (4) с учётом (6) и неравенства $(\hat{y},\, r(x)) \le 0,$ которое получается из тех же соображений, что и ($5'$), имеем:

$$f(\hat{x}) \le f(x) + (\hat{y},\, r(x)) \le f(x)\ \ \forall\, x \in X.$$

Это означает, что $\hat{x}$ — точка глобального минимума.

Замечание. Условие (6), записанное в виде

$$(\hat{y},\, g(\hat{x}) - b) = 0\ ,$$

называют обычно условием дополняющей *нежёсткости*. Смысл его состоит в следующем:

1) если $\hat{y}_i > 0,$ то $g_i(\hat{x}) - b_i = 0;$

2) если $g_i(\hat{x}) - b_i < 0,$ то $\hat{y}_i = 0.$

Замечание. Из вида функции Лагранжа (1) и условия (6) получаем равенство $f(\hat{x}) = L(\hat{x}, \hat{y}),$ то есть значение целевой функции $f(x)$ в оптимальной точке $\hat{x}$ совпадает со значением функции Лагранжа в седловой точке.

Приведём *необходимое и достаточное условие* оптимальности в задаче выпуклого программирования.

Теорема. Пусть в задаче выпуклого программирования множество X обладает свойством регулярности (или свойством регулярности Слейтера). Для того чтобы $\hat{x}$ была оптимальной точкой, необходимо и достаточно, чтобы существовала точка

$\hat{y} \geq 0$ такая, чтобы пара $(\hat{x},\ \hat{y})$ была седловой точкой функции Лагранжа $L(x, y)$ на множестве $x \in X_0,\ y \geq 0$.

Достаточность условий теоремы следует из предыдущей теоремы.

Необходимость условий теоремы фактически была доказана при рассмотрении общей задачи нелинейного программирования, где множество X_0 было положительным ортантом. Условие Слейтера необходимо для того, чтобы в допустимой области были внутренние точки.

3.4. Условия Куна-Таккера для задачи выпуклого программирования

Пусть множество X_0 есть положительный ортант, то есть $X_0 = \{x : x \geq 0\}$.

Тогда рассматриваемая задача выпуклого программирования примет вид

$$\min_{x \in X} f(x);\ \ X = \{x : g(x) \geq 0,\ x \geq 0\}. \tag{*}$$

При рассмотрении такой задачи в случае произвольных функций $f(x)$ и $g(x)$ было получено необходимое условие оптимальности в форме условий Куна-Таккера. Для задачи выпуклого программирования при выполнении условий регулярности условия Куна-Таккера будут и необходимыми и достаточными условиями оптимальности.

Теорема. Пусть функции $f(x)$ и $g(x)$ непрерывно дифференцируемы на множестве $X_0 = \{x : x \geq 0\}$, множество X удовлетворяет условию регулярности (или условию регулярности Слейтера). Для того чтобы $\hat{x}$ была оптимальной точкой, необходимо и достаточно, чтобы существовала точка $\hat{y} \geq 0$ такая, чтобы пара $(\hat{x},\ \hat{y})$ была седловой точкой функции Лагранжа $L(x, y)$ на множестве $x \geq 0,\ y \geq 0$ или чтобы выполнялись условия Куна-Таккера:

$$\frac{\partial \hat{L}}{\partial x} \geq 0; \qquad (1) \qquad\qquad \frac{\partial \hat{L}}{\partial y} \leq 0; \qquad (4)$$

$$\left(\hat{x}, \frac{\partial \hat{L}}{\partial x}\right) = 0; \qquad (2) \qquad\qquad \left(\hat{y}, \frac{\partial \hat{L}}{\partial y}\right) = 0; \qquad (5)$$

$$\hat{x} \geq 0; \qquad (3) \qquad\qquad \hat{y} \geq 0; \qquad (6)$$

где использованы обозначения:

$$\frac{\partial L}{\partial x} = \left(\frac{\partial L}{\partial x_1}, \ldots, \frac{\partial L}{\partial x_n}\right)^T; \ \frac{\partial L}{\partial y} = \left(\frac{\partial L}{\partial y_1}, \ldots, \frac{\partial L}{\partial y_m}\right)^T;$$

$$\frac{\partial \hat{L}}{\partial x} = \left.\frac{\partial L(x, y)}{\partial x}\right|_{\substack{x=\hat{x} \\ y=\hat{y}}}; \ \frac{\partial \hat{L}}{\partial y} = \left.\frac{\partial L(x, y)}{\partial y}\right|_{\substack{x=\hat{x} \\ y=\hat{y}}}.$$

Замечание. Условия (1) – (6) можно записать в эквивалентной покомпонентной форме:

$$\frac{\partial \hat{L}}{\partial x_i} \geq 0 \ \ (i = \overline{1, n}); \qquad (1а) \qquad\qquad \frac{\partial \hat{L}}{\partial y_j} \leq 0 \ \ (j = \overline{1, m}); \qquad (4а)$$

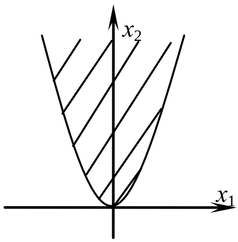
$$\hat{x}_i \frac{\partial \hat{L}}{\partial x_i} = 0 \ \ (i = \overline{1, n}); \qquad (2а) \qquad\qquad \hat{y}_j \frac{\partial \hat{L}}{\partial y_j} = 0 \ \ (j = \overline{1, m}); \qquad (5а)$$

$$\hat{x}_i \geq 0 \ \ (i = \overline{1, n}); \qquad (3а) \qquad\qquad \hat{y}_j \geq 0 \ \ (j = \overline{1, m}). \qquad (6а)$$

Замечание. Если в задаче выпуклого программирования (*) положить $X_0 = R^n$, то в (1) – (6) условия (1) – (3) нужно заменить одним условием $\frac{\partial \hat{L}}{\partial x} = 0$.

4. Задачи для самостоятельного решения

4.1. Изобразить графически и определить, является ли выпуклым множество



<u>Пример</u>. Рассмотрим множество $\{x \in R^2 : x_2 \geq x_1^2\}$. Для его графического изображения сначала построим границу множества — параболу $x_2 = x_1^2$. Неравенству $x_2 \geq x_1^2$ удовлетворяют точки параболы и точки, лежащие выше неё. Это легко проверить, взяв, например, точку (0; 1).

Данное множество, очевидно, является выпуклым.

4.1.1. $a;b$, где $a = (-1;3)$, $b = (1;5)$.

4.1.2. $\{x \in R^2 : x_1 x_2 \geq 1\}$.

4.1.3. $\{x \in R^2 : x_1 x_2 \geq 1, x_1 > 0\}$.

4.1.4. $\{x \in R^2 : x_1^2 + x_2^2 \leq 1\}$.

4.1.5. $\{x \in R^2 : x_1^2 + x_2^2 = 1\}$.

4.1.6. $\{x \in R^2 : |x_i|, i = 1, 2\}$.

4.1.7. $\{x \in R^2 : -1 \leq x_1 \leq 3, 0 \leq x_2 \leq 4\}$.

4.1.8. $\{x \in R^2 : x_1 = 1 + 2t, x_2 = 3 - t, t \in R\}$.

4.2. Определить, при каких значениях параметра *a* множество *M* будет выпуклым.

<u>Пример</u>. Рассмотрим множество

$$M = \{x \in R^2 : x_1^2(a^2 + 3a + 2) - x_2 \geq 0\}.$$

При $(a^2 + 3a + 2) \geq 0$ множество M представляет собой множество точек параболы $x_1^2(a^2 + 3a + 2) - x_2 = 0$, ветви которой направлены вверх и точки, лежащие выше неё, и является выпуклым.

При $(a^2+3a+2)<0$ ветви параболы направлены вниз и множество M точек, лежащих выше параболы выпуклым не является.

Таким образом, искомые значения параметра a являются решениями неравенства $(a^2+3a+2)\ge 0$, а именно, $a\in(-\infty;-2]\cup[-1;+\infty)$.

4.2.1. $M=\{x\in R^2 : a(x_1-x_2^2)=0, x_1+x_2=1\}$.

4.2.2. $M=\{x\in R^2 : a(x_1-x_2^2)=0, x_1+x_2=a\}$.

4.2.3. $M=\{x\in R^2 : a(x_1-x_2^2)\le 0, x_1+x_2=a\}$.

4.2.4. $M=\{x\in R^2 : x_1+x_2\ge 1, x_2\le ax_1^2\}$.

4.2.5. $M=\{x\in R^2 : e^{x_1}(a^2-5a+6)-x_2(a^2+2)\le 0\}$.

4.2.6. $M=\{x\in R^2 : x_2^2\le e^{x_1}, x_2\le ax_1\}$.

4.3. Исследовать на выпуклость функции многих переменных

<u>Пример</u>. Рассмотрим функцию $z=x^2+x^2y+xy^2+y^2$ в области $0\le x<1;\ 0\le y<1$. Составим матрицу вторых производных (матрицу Гессе):

$$H=\begin{pmatrix} z''_{xx} & z''_{xy} \\ z''_{yx} & z''_{yy}\end{pmatrix}=\begin{pmatrix} 2+2y & 2x+2y \\ 2x+2y & 2+2x\end{pmatrix}.$$

Главные угловые диагональные миноры

$$M^{(1)}=2+2y>0,$$

$$M^{(2)}=\begin{vmatrix} 2+2y & 2x+2y \\ 2x+2y & 2+2x\end{vmatrix}=4+4x+4y+4xy-4x^2-8xy-4y^2>0.$$

Так как в области $0\le x<1;\ 0\le y<1$ главные угловые миноры положительны, то исследуемая функция выпукла в этой области.

4.3.1. $z=4-x^2+x-2y^2;\ (x,y)\in R^2$.

4.3.2. $z = x^2 + y^2 + 3x + 1;\ \ (x, y) \in R^2$.

4.3.3. $z = 2x - x^2 - 2y^2;\ \ (x, y) \in R^2$.

4.3.4. $z = (2 - x)^2 + 3y^2;\ \ (x, y) \in R^2$.

4.3.5. $z = x^2 + 2xy + 3y^2;\ \ (x, y) \in R^2$.

4.3.6. $z = 1/x + 1/y;\ \ x > 0, y > 0$.

4.3.7. $z = x^2 - xy + y^2 + 3;\ \ x \ge 1, y \ge 1$.

4.3.8. $z = \ln(x + 1) - xy - y^2;\ \ 0 < x < \sqrt{2} - 1, -\infty < y < +\infty$.

4.3.9. $z = \sqrt[3]{xy};\ \ (x, y) \in R^2$.

4.3.10. $z = y^2 + ye^x;\ \ x \le 0, -1 \le y \le 1$.

4.3.11. $z = \sqrt{x} + \sqrt{x + y} + \sqrt{y};\ \ x > 0,\ y > 0$.

4.4. Проверить с использованием условий Куна-Таккера, является ли заданная точка решением задачи выпуклого программирования

<u>Пример</u>. Проверим, что точка (1; 1) является решением задачи выпуклого программирования:

$$z = (x - 1)^2 + (y - 1)^2 \Rightarrow \min;$$

$$x^2 + y^2 \le 4;\ x - 2y \le 0;\ \ y - 2x \le 0;\ \ x \ge 0,\ \ y \ge 0.$$

Составим функцию Лагранжа

$$L = (x - 1)^2 + (y - 1)^2 + \lambda(x^2 + y^2 - 4) + \mu(x - 2y) + \nu(y - 2x).$$

Составим условия Куна-Таккера:

$$L'_x = 2(x - 1) + 2\lambda x + \mu - 2\nu \ge 0,$$

$$L'_y = 2(y - 1) + 2\lambda y - 2\mu + \nu \ge 0,$$

$$xL'_x = x(2(x - 1) + 2\lambda x + \mu - 2\nu) = 0,$$

$$yL'_y = y(2(y - 1) + 2\lambda y - 2\mu + \nu) = 0,$$

$$L'_\lambda = x^2 + y^2 - 4 \le 0,$$

$$L'_\mu = x - 2y \le 0,$$

$$L'_\nu = y - 2x \le 0,$$

$$\lambda L'_\lambda = \lambda(x^2 + y^2 - 4) = 0,$$
$$\mu L'_\mu = \mu(x - 2y) = 0,$$
$$\nu L'_\nu = \nu(y - 2x) = 0,$$
$$\lambda \ge 0,\ \ \mu \ge 0,\ \ \nu \ge 0.$$

Из последних трёх равенств находим $\lambda^* = 0,\ \mu^* = 0,\ \nu^* = 0.$ Подставим $x^* = 1,\quad y^* = 1;\ \lambda^* = 0,\ \mu^* = 0,\ \nu^* = 0$ в условия Куна-Таккера:

$$L'_x = 2(1-1) + 2\cdot 0\cdot 1 + 0 - 2\cdot 0 = 0,$$
$$L'_y = 2(1-1) + 2\cdot 0\cdot 1 - 2\cdot 0 + 0 = 0,$$
$$xL'_x = 1\cdot(2(1-1) + 2\cdot 0\cdot 1 + 0 - 2\cdot 0) = 0,$$
$$yL'_y = 1\cdot(2(1-1) + 2\cdot 0\cdot 1 - 2\cdot 0 + 0) = 0,$$
$$L'_\lambda = 1^2 + 1^2 - 4 = -2 < 0,$$
$$L'_\mu = 1 - 2\cdot 1 = -1 < 0,$$
$$L'_\nu = 1 - 2\cdot 1 = -1 < 0,$$
$$\lambda L'_\lambda = 0\cdot(1^2 + 1^2 - 4) = 0,$$
$$\mu L'_\mu = 0\cdot(1 - 2\cdot 1) = 0,$$
$$\nu L'_\nu = 0\cdot(1 - 2\cdot 1) = 0,$$
$$\lambda = 0,\ \ \mu = 0,\ \ \nu = 0.$$

Все условия Куна-Таккера выполнены. Поэтому точка (1; 1) является решением рассматриваемой задачи выпуклого программирования.

4.4.1. $(x^*,\ y^*) = (4;\ 0)$; $3x^2 + 4xy + 5y^2 \Rightarrow \min;$
$x + y \ge 4;\ x \ge 0,\ y \ge 0.$

4.4.2. $(x^*,\ y^*) = (2{,}5;\ 1{,}5)$; $x^2 - 2x + y^2 \Rightarrow \min;$
$x^2 + 4y^2 - 4x - 4y \le 0;\ x + y \ge 4,\ x \ge 0,\ y \ge 0.$

4.4.3. $(x^*,\ y^*) = (0{,}4;\ 1{,}8)$; $x^2 - 2x - 8y + y^2 \Rightarrow \min;$
$x + 2y^2 \le 4;\ 3x + y \ge 3,\ x \ge 0,\ y \ge 0.$

4.5. Решить задачи выпуклого программирования с использованием условий Куна-Таккера

Пример. Решим задачу выпуклого программирования

$$z = (x-4)^2 + (y-3)^2 \Rightarrow \min;$$

$$x + 2y \le 6,$$

$$2x + y \le 6.$$

Составим функцию Лагранжа

$$L(x, y; \lambda, \mu) = (x-4)^2 + (y-3)^2 + \lambda(x + 2y - 6) + \mu(2x + y - 6).$$

Вычислим её частные производные и запишем условия Куна-Таккера:

$$L'_x = 2(x-4) + \lambda + 2\mu = 0, \quad \text{(а)}$$

$$L'_y = 2(y-3) + 2\lambda + \mu = 0, \quad \text{(б)}$$

$$L'_\lambda = x + 2y - 6 \le 0, \quad \text{(в)}$$

$$L'_\mu = 2x + y - 6 \le 0, \quad \text{(г)}$$

$$\lambda L'_\lambda = \lambda(x + 2y - 6) = 0, \quad \text{(д)}$$

$$\mu L'_\mu = \mu(2x + y - 6) = 0, \quad \text{(е)}$$

$$\lambda \ge 0, \ \ \mu \ge 0. \quad \text{(ж)}$$

Рассмотрим 4 возможных варианта для множителей Лагранжа:

$(\lambda > 0, \mu > 0)$, $(\lambda > 0, \mu = 0)$, $(\lambda = 0, \mu > 0)$ и $(\lambda = 0, \mu = 0)$.

а) Пусть $\lambda > 0, \mu > 0$. Тогда, разделив, равенство (д) на λ, а равенство (е) — на μ, получим

$$x + 2y - 6 = 0, \quad \text{(д}'\text{)}$$

$$2x + y - 6 = 0. \quad \text{(е}'\text{)}$$

В этом случае неравенства (в) и (г) выполняются автоматически и их можно пока из рассмотрения исключить. Таким образом, из системы (а) – (ж) получаем систему вида:

$$2(x-4) + \lambda + 2\mu = 0, \quad \text{(а)}$$

$$2(y-3) + 2\lambda + \mu = 0, \quad \text{(б)}$$

$$x + 2y - 6 = 0, \quad \text{(д}'\text{)}$$

$$2x + y - 6 = 0 \quad \text{(е}'\text{)}$$

$$\lambda > 0, \quad \mu > 0. \qquad (ж')$$

Из уравнений (д′) и (е′) находим: $x^* = 2,\ y^* = 2.$

Подставляя эти значения в первое и второе уравнения, получаем систему

$$\lambda + 2\mu = 4,$$

$$2\lambda + \mu = 2,$$

решая которую находим множители Лагранжа

$$\lambda^* = 0,\ \mu^* = 2.$$

Так как это противоречит предположению $\lambda > 0,$ то данный вариант отбрасываем.

б) Пусть $\lambda > 0,\ \mu = 0$. В этом случае условия Куна-Таккера можно записать так:

$$2(x-4) + \lambda = 0, \qquad (а)$$

$$2(y-3) + 2\lambda = 0, \qquad (б)$$

$$2x + y - 6 \le 0, \qquad (г)$$

$$x + 2y - 6 = 0, \qquad (д')$$

$$\lambda > 0, \quad \mu = 0.$$

Из первого и второго уравнений получаем:

$$x = 4 - 0{,}5\lambda,$$

$$y = 3 - \lambda.$$

Подставляя эти выражения в уравнение (д′), получим

$$\lambda^* = 1{,}6 > 0.$$

Проверим выполнение условия (г):

$$2x + y - 6 = 2(4 - 0{,}5 \cdot \lambda^*) + (3 - \lambda^*) - 6 = 1{,}8 > 0.$$

Таким образом, это условие Куна-Таккера не выполняется, поэтому переходим к рассмотрению следующего варианта.

в) Пусть $\lambda = 0,\ \mu > 0$. В этом случае условия Куна-Таккера можно записать так:

$$2(x-4) + 2\mu = 0, \qquad (а)$$

$$2(y-3) + \mu = 0, \qquad (б)$$

$$x + 2y - 6 \le 0, \qquad (в)$$

$$2x + y - 6 = 0, \qquad (e')$$
$$\lambda = 0, \ \ \mu > 0.$$

Из первого и второго уравнений получаем:

$$x = 4 - \mu,$$
$$y = 3 - 0{,}5\mu.$$

Подставляя эти выражения в условие (е′), получим $\mu^* = 2 > 0.$

Проверим выполнение условия (в):

$$x + 2y - 6 = 4 - \mu^* + 2(3 - 0{,}5\mu^*) - 6 = 4 - 2\mu^* = 0\,.$$

Остальные условия также выполняются. Поэтому получено решение задачи:

$$x^* = 4 - \mu^* = 2; \ \ y^* = 3 - 0{,}5\mu^* = 2\,.$$

Значение функции в этой точке $z^* = z_{\min} = 5$.

Задача решена, поэтому последний вариант можно не рассматривать. Однако в учебных целях мы его также рассмотрим.

г) Пусть $\lambda = 0, \mu = 0$. В этом случае условия Куна-Таккера можно записать так:

$$2(x - 4) = 0, \qquad (а)$$
$$2(y - 3) = 0, \qquad (б)$$
$$x + 2y - 6 \le 0, \qquad (в)$$
$$2x + y - 6 \le 0, \qquad (г)$$
$$\lambda = 0, \ \ \mu = 0.$$

Из первых двух условий получаем: $x^* = 4; \ y^* = 3.$

Легко видеть, что при этих значениях x и y третье и четвёртое условия не выполняются.

<u>Пример</u>. Решим задачу выпуклого программирования

$$z = (x - 3)^2 + (y - 2)^2 \Rightarrow \min;$$
$$x + y \le 2,$$
$$y \le 1,$$
$$x \ge 0, \ \ y \ge 0.$$

Составим функцию Лагранжа

$$L(x, y; \lambda, \mu) = (x-3)^2 + (y-2)^2 + \lambda(x+y-2) + \mu(y-1).$$

Вычислим её частные производные и запишем условия Куна-Таккера:

$$L'_x = 2(x-3) + \lambda \geq 0, \qquad \text{(а)}$$

$$L'_y = 2(y-2) + \lambda + \mu \geq 0, \qquad \text{(б)}$$

$$xL'_x = 2x(x-3) + \lambda x = 0, \qquad \text{(в)}$$

$$yL'_y = 2y(y-2) + (\lambda + \mu)y = 0, \qquad \text{(г)}$$

$$L'_\lambda = x + y - 2 \leq 0, \qquad \text{(д)}$$

$$L'_\mu = y - 1 \leq 0, \qquad \text{(е)}$$

$$\lambda L'_\lambda = \lambda(x+y-2) = 0, \qquad \text{(ж)}$$

$$\mu L'_\mu = \mu(y-1) = 0, \qquad \text{(з)}$$

$$x \geq 0,\ \ y \geq 0,\ \ \lambda \geq 0,\ \ \mu \geq 0. \qquad \text{(и)}$$

Рассмотрим 4 возможных варианта для множителей Лагранжа: $(\lambda > 0,\ \mu > 0)$, $(\lambda > 0,\ \mu = 0)$, $(\lambda = 0,\ \mu > 0)$ и $(\lambda = 0,\ \mu = 0)$.

а) Пусть $\lambda > 0, \mu > 0$. В этом случае условия (д) – (з), как это было показано в предыдущем примере, можно заменить следующими:

$$x + y - 2 = 0, \qquad \text{(ж}'\text{)}$$

$$y - 1 = 0. \qquad \text{(з}'\text{)}$$

Тогда условия Куна-Таккера можно записать так:

$$2(x-3) + \lambda \geq 0, \qquad \text{(а)}$$

$$2(y-2) + \lambda + \mu \geq 0, \qquad \text{(б)}$$

$$2x(x-3) + \lambda x = 0, \qquad \text{(в)}$$

$$2y(y-2) + (\lambda + \mu)y = 0, \qquad \text{(г)}$$

$$x + y - 2 = 0, \qquad \text{(ж}'\text{)}$$

$$y - 1 = 0, \qquad \text{(з}'\text{)}$$

$$x \geq 0,\ \ y \geq 0,\ \ \lambda > 0,\ \ \mu > 0. \qquad \text{(и}'\text{)}$$

Решая систему уравнений

$$x+y-2=0,$$
$$y-1=0,$$

находим $x^*=1,\ \ y^*=1.$

Условия (в) и (г) запишутся так:

$$2(1-3)+\lambda=-4+\lambda=0,$$
$$yL'_y=2(1-2)+(\lambda+\mu)=-2+\lambda+\mu=0.$$

Откуда находим $\lambda^*=4,\ \mu^*=-2$. Это противоречит условию $\mu>0$.

б) Пусть $\lambda>0, \mu=0$. В этом случае условия Куна-Таккера можно записать так:

$$L'_x=2(x-3)+\lambda\geq 0, \qquad \text{(а)}$$
$$L'_y=2(y-2)+\lambda\geq 0, \qquad \text{(б)}$$
$$xL'_x=2x(x-3)+\lambda x=0, \qquad \text{(в)}$$
$$yL'_y=2y(y-2)+\lambda y=0, \qquad \text{(г)}$$
$$L'_\mu=y-1\leq 0, \qquad \text{(е)}$$
$$L'_\lambda=x+y-2=0, \qquad \text{(ж}'\text{)}$$
$$x\geq 0,\ \ y\geq 0,\ \ \lambda>0,\ \ \mu=0.$$

Предположим, что $x>0,\ \ y>0$. Третье уравнение разделим на *x*, четвёртое разделим на *y* и получим систему уравнений

$$2(x-3)+\lambda=0,$$
$$2(y-2)+\lambda=0,$$
$$x+y-2=0.$$

Решение этой системы: $x^*=1{,}5;\ \ y^*=0{,}5;\ \ \lambda^*=3.$

Проверяем выполнение условий Куна-Таккера:

$$L'_x=2(1{,}5-3)+3=0, \qquad \text{(а)}$$
$$L'_y=2(0{,}5-2)+3=0, \qquad \text{(б)}$$
$$xL'_x=0, \qquad \text{(в)}$$
$$yL'_y=0, \qquad \text{(г)}$$

$$L'_\mu = 0{,}5 - 1 = -0{,}5 < 0, \qquad \text{(е)}$$

$$L'_\lambda = 1{,}5 + 0{,}5 - 2 = 0, \qquad \text{(ж}'\text{)}$$

$$x^* > 0,\ \ y^* > 0,\ \ \lambda^* > 0,\ \ \mu^* = 0.$$

Все условия Куна-Таккера выполнены, поэтому получено оптимальное решение

$$x^* = 1{,}5;\ \ y^* = 0{,}5.$$

Значение целевой функции в этой точке $z^* = z_{\min} = 4{,}5$.

г) Вариант $\lambda = 0,\ \mu = 0$ можно не рассматривать, так как задача выпуклого программирования имеет единственное решение. Однако для тренировки рекомендуется исследовать все возможные варианты. В данном случае условия Куна-Таккера выглядят так:

$$L'_x = 2(x-3) \ge 0,$$

$$L'_y = 2(y-2) \ge 0,$$

$$xL'_x = 2x(x-3) = 0,$$

$$yL'_y = 2y(y-2) = 0,$$

$$L'_\lambda = x + y - 2 \le 0,$$

$$L'_\mu = y - 1 \le 0,$$

$$\lambda L'_\lambda = 0,$$

$$\mu L'_\mu = 0,$$

$$x \ge 0,\ \ y \ge 0,\ \ \lambda = 0,\ \ \mu = 0.$$

Система уравнений, составленная из третьего и четвёртого равенств, имеет четыре решения: (0; 0), (0; 2), (3; 0), (3; 2).

Первое решение противоречит первому и второму условию Куна-Таккера.

Второе решение противоречит первому и шестому условию Куна-Таккера.

Третье решение противоречит второму и пятому условию Куна-Таккера.

Четвёртое решение противоречит пятому и шестому условию Куна-Таккера.

Таким образом, вариант г) должен быть отброшен.

4.5.1 $z = x^2 + 2y^2 - 2x - 4y \Rightarrow \min;$

$x + 2y \le 8,\ 2x - y \le 12$.

4.5.2. $z = (x-4)^2 + (y-2)^2 \Rightarrow \min;$

$x + y \le 3,\ x + 2y \le 4$.

4.5.3. $z = x + 4y + xy - 2x^2 - 2y^2 \Rightarrow \max;$

$x + 2y \le 12,\ 3x + y \le 15$.

4.5.4. $z = 2y^2 - 2x - 3y \Rightarrow \min;$

$x + 4y \le 4,\ x + y \le 2,\ \ x \ge 0,\ y \ge 0$.

4.5.5. $z = 8y - 2x - x^2 - y^2 \Rightarrow \max;$

$x + 2y \le 12,\ -x + y \ge -8,\ \ x \ge 0,\ y \ge 0$.

4.5.6. $z = 20x - y^2 \Rightarrow \max;$

$x - y \ge 0,\ -x + 2y \le 2,\ \ x \ge 0,\ y \ge 0$.

4.5.7. $z = x + 8y - x^2 - y^2 \Rightarrow \max;$

$x + y \le 7,\ y \le 5,\ \ x \ge 0,\ y \ge 0$.

4.5.8. $z = x^2 + y^2 - 2y \Rightarrow \min;$

$x + y \le 18,\ y \le 12,\ \ x \ge 0,\ y \ge 0$.

Литература

1. Акулич И. Л. Математическое программирование в примерах и задачах / И. Л. Акулич. — М.: Высшая школа, 1986. — 319 с.

2. Алексеев В. М. Сборник задач по оптимизации. Теория. Примеры. Задачи. / В. М. Алексеев, Э. М. Галеев, В. М. Тихонов. — М.: Наука, 1984. — 288 с.

3. Васильев О. В. Методы оптимизации в задачах и упражнениях / О. В. Васильев, А. В. Аргучинцев. — М.: Физматлит, 1999. — 208 с.

4. Ильченко А. Н. Практикум по экономико-математическим методам /А. Н. Ильченко, О. Л. Ксенофонтова, Г. В. Канакина. — М.: Финансы и статистика, ИНФРА-М, 2009. — 288 с.

5. Кузнецов А. В. Сборник задач и упражнений по высшей математике: математическое программирование / А. В. Кузнецов, В. А. Сакович, Н. И. Холод, Р. А. Рутковский, Н. М. Слукин, Л. Ф. Дежурко, М. А. Хотомцева. — Минск: Вышэйшая школа, 2002. — 447 с.

6. Кузнецов Б. Т. Математические методы и модели исследования операций / Б. Т. Кузнецов. — М.: ЮНИТИ-ДАНА, 2005. — 390 с.

7. Лугинин О. Е. Экономико-математические методы и модели: теория и практика с решением задач / О. Ю. Лугинин, В. Н. Фомишина. — Ростов н/Д: Феникс, 2009. — 440 с.

8. Математические методы и модели исследования операций / под ред. В. А. Колемаева. — М.: ЮНИТИ-ДАНА, 2008. — 592 с.

9. Пантелеев А. В. Методы оптимизации в примерах и задачах / А. В. Пантелеев, Т. А. Летова. — М.: Высшая школа, 2005. — 544 с.

10. Шапкин А. С. Математические методы и модели исследования операций / А. С. Шапкин, В. А. Шапкин. — М.: Дашков и К, 2009. — 400 с.

Содержание

1. Выпуклые множества 0
2. Выпуклые функции 7
2.1. Определение выпуклых функций 7
2.2. Свойства выпуклых функций 8
2.3. Неравенство Йенсена 10
2.4. Условия выпуклости первого порядка 11
2.5. Примеры выпуклых функций 13
2.6. Условия выпуклости второго порядка...............................13
2.7. Множества, определяемые неравенствами с выпуклыми функциями 14
2.8. Экстремальные свойства функций на выпуклых множествах 15
2.9. Свойства максимумов выпуклой функции 18
2.10. Свойства максимумов и минимумов линейной функции 19
3. Выпуклое программирование...19
3.1. Понятие задачи выпуклого программирования.................19
3.2. Условия регулярности 20
3.3. Условия минимума в задаче выпуклого программирования. Функция Лагранжа...................................21
3.4. Условия Куна-Таккера для задачи выпуклого программирования 23
4. Задачи для самостоятельного решения 25
Литература...36

Учебное издание

ВЫПУКЛЫЕ ФУНКЦИИ. ВЫПУКЛОЕ ПРОГРАММИРОВАНИЕ

Методические указания по решению задач

Составители:
Якупов Рафаэль Тимирович, **Гарайшина** Ирина Рашитовна

Подписано к печати 28.04.2012. Формат 60x84 1/16. Усл. печ. л. 2,1. Ризография. Тираж 100 экз. Заказ № 1008 .

Филиал федерального государственного бюджетного образовательного учреждения высшего профессионального образования «Кемеровский государственный университет» в г. Анжеро-Судженске.
652470, Кемеровская область, г. Анжеро-Судженск, ул. Ленина, 8.

Отпечатано на участке оперативной полиграфии филиала федерального государственного бюджетного образовательного учреждения высшего профессионального образования «Кемеровский государственный университет» в г. Анжеро-Судженске.
652470, Кемеровская область, г. Анжеро-Судженск, ул. Ленина, 8.