Выпуклые множества

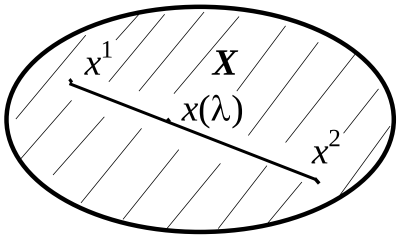
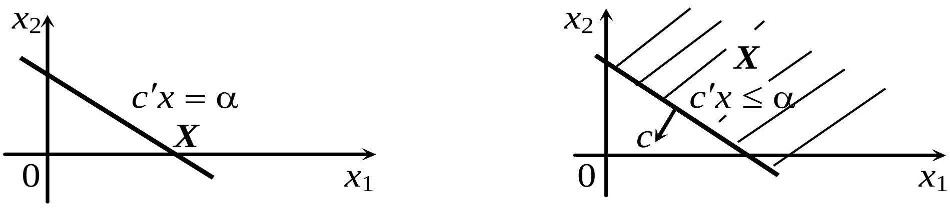


Рис. 5.1

Множество[1] $X \subset \mathbf{R}^n$ называется ***выпуклым***, если $x(\lambda) = \lambda x^1 + (1-\lambda)x^2 \in X$ при всех $x^1,\ x^2 \in X$, $\lambda \in [0;\ 1]$, т. е. если вместе с любыми своими двумя точками множество X содержит и соединяющий их отрезок (рис. 5.1). ***Гиперплоскость*** $X = \{x:\ \ c'x = \ \ = \alpha\}$, $c \in \mathbf{R}^n$, $\alpha \in \mathbf{R}$, ***полупространство*** $X = \{\, x:\ c'x \le \alpha \,\}$ являются выпуклыми множествами (рис. 5.2).

Множество $X \subset \mathbf{R}^n$ называется ***выпуклым конусом*** с вершиной в начале координат, если из того, что $x, y \in X$, следует, что $x + y \in X$, $\lambda x \in X$ для всех $\lambda \ge 0$ (рис. 5.3).



гиперплоскость полупространство

Рис. 5.2

Основные свойства выпуклых множеств:

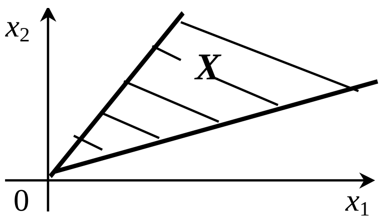


Рис. 5.3

1. Пересечение любого числа выпуклых множеств – выпуклое множество.

2. Если $X_1, \ldots, X_k$ – выпуклые множества в $\mathbf{R}^n$, то множество $X = \sum_{i=1}^{k} \alpha_i X_i = \left\{\, x \in \mathbf{R}^n : x = \sum_{i=1}^{k} \alpha_i x^i,\ \ x^i \in X_i,\ i = \overline{1,k} \,\right\}$ выпукло (α_i, $i = \overline{1,k}$, — любые числа).

3. Выпуклы сумма и разность выпуклых множеств X_1 и X_2, т. е. множества $X_1 \pm X_2 = \{x \in \mathbf{R}^n : x = x^1 \pm x^2,\ x^1 \in \ \in X_1, x^2 \in X_2\}$.

4. Замыкание $\overline{X}$ выпуклого множества X также выпукло.

---

[1] В дальнейшем считаем, если не оговорено противное, что рассматриваемые множества непустые.

Выпуклой оболочкой множества $X \subset \mathbf{R}^n$ называется множество conv$X$ $= \left\{ x : x = \sum_{i=1}^{n+1} \lambda_i x^i,\ x^i \in X,\ \ \sum_{i=1}^{n+1} \lambda_i = 1,\ \lambda_i \ge 0,\ i = \overline{1, n+1} \right\}$ (рис. 5.4).

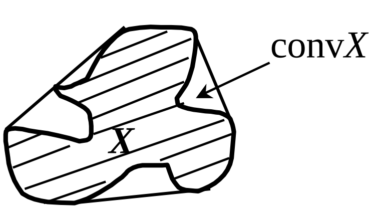


Рис. 5.4

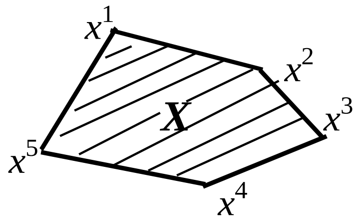


Рис. 5.5

Множество conv$X$ совпадает с пересечением всех выпуклых множеств, содержащих множество X. Выпуклая оболочка множества — множество выпуклое.

Пусть $x^1, \dots, x^k \in \mathbf{R}^n$. Множество $X = \mathrm{conv}\{x^1, \dots, x^k\}$ называется ***выпуклым многогранником*** (рис. 5.5).

При исследовании задач минимизации важную роль играют следующие утверждения.

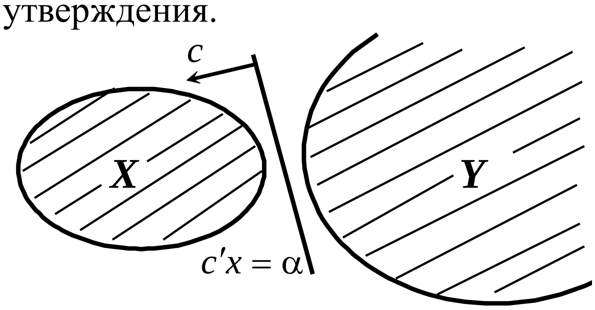


Рис. 5.6

Теорема 5.1 (об отделимости выпуклых множеств). *Если множества $X, Y \subset \mathbf{R}^n$ выпуклы, хотя бы одно из них ограничено и их замыкания не имеют общих точек* ($\overline{X} \cap \overline{Y} = \varnothing$)*, то они строго отделимы, т. е. существуют n-вектор c, $c \ne 0$, и число α такие, что для всех $x \in X, y \in Y$ выполняются неравенства $c'x > \alpha$, $c'y < \alpha$* (*рис.* 5.6). *При этом гиперплоскость $c'x = \alpha$ называется **разделяющей гиперплоскостью**.*

Если X, Y – выпуклые множества, не имеющие общих внутренних точек, то они отделимы, т. е. существуют c, $c \ne 0$, и α такие, что
$c'x \ge \alpha,\quad c'y \le \alpha\ \ \forall\ x \in X, y \in Y$.

Теорема 5.2 (о существовании опорной гиперплоскости). *Если x^* – граничная точка выпуклого множества X, то в этой точке существует опорная гиперплоскость к X, т. е. при некотором $c \ne 0$ выполняется неравенство* $c'x \le c'x^*\ \ \forall\ x \in X$ ($c'x = c'x^*$ – ***опорная гиперплоскость***).

Выпуклые функции

Функция $f(x)$, определенная и конечная на выпуклом множестве $X \subset \mathbf{R}^n$, называется ***выпуклой***, если для любых $x^1,\ x^2 \in X$ выполняется неравенство (рис. 5.7)

$$f(\lambda x^1 + (1-\lambda)x^2) \le \lambda f(x^1) + (1-\lambda) f(x^2) \quad \forall \quad \lambda \in [0;\ 1]. \tag{1}$$

Выпуклая функция $f(x),\ x \in X$, называется ***строго выпуклой***, если для любых $x^1,\ x^2 \in X$, $x^1 \ne x^2$, и $\lambda \in (0;\ 1)$ выполняется строгое неравенство

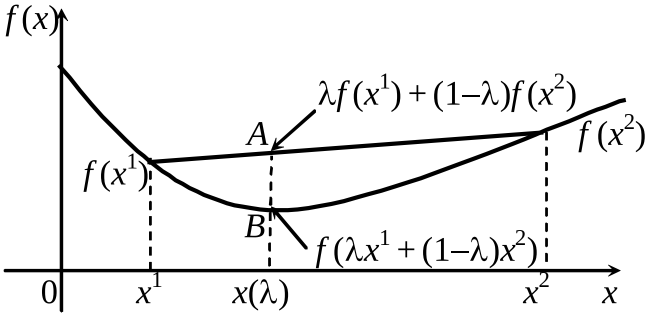


Рис. 5.7

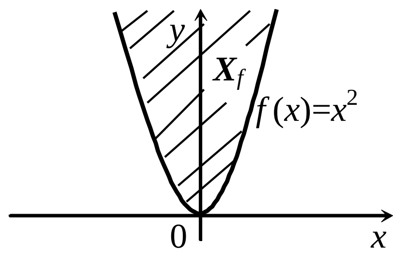


Рис. 5.8

(5.1): $f(\lambda x^1 + (1-\lambda)x^2) < \lambda f(x^1) + (1-\lambda)f(x^2)$. У строго выпуклой функции точки A и B не совпадают (рис. 5.7). Функция $f(x) = x^2,\ \ x \in \mathbf{R}$, строго выпуклая (рис. 5.8).

Приведем несколько ***критериев выпуклости функций***.

а) Для выпуклости функции $f(x)$, *определенной на выпуклом множестве* X, *необходимо и достаточно, чтобы было выпуклым множество* $X_f = \{\{x, y\} : x \in X,\ y \ge f(x)\}$.

Множество X_f называется ***надграфиком*** функции $f(x)$. К примеру, функция $f(x) = x^2$, $x \in \mathbf{R}$, выпуклая, поскольку ее надграфик — выпуклое множество (рис. 5.8).

б) Функция $f(x),\ x \in \mathbf{R}^n$, *выпукла тогда и только тогда, когда выпукла функция* $\varphi(t) = f(x + tl)$ *скалярного аргумента* $t \in \mathbf{R}$ *при любых* $x,\ l \in \mathbf{R}^n$.

в) Функция $f \in C^{(1)}(\mathbf{R}^n)$ *выпукла в том и только в том случае, если выполняется неравенство*

$$f(x) - f(x^*) \ge (x - x^*)' \frac{\partial f(x^*)}{\partial x} \tag{2}$$

для всех $x,\ x^* \in \mathbf{R}^n$. *В случае строгой выпуклости неравенство* (2) *выполняется как строгое для любых* $x,\ x^* \in \mathbf{R}^n$, $x \ne x^*$.

г) *Если* $f \in C^{(2)}(\mathbf{R}^n)$, *то для ее выпуклости необходимо и достаточно, чтобы* $\partial^2 f(x)/\partial x^2 \geq 0$, $x \in \mathbf{R}^n$ (*если* $\partial^2 f(x)/\partial x^2 > 0$, $x \in \mathbf{R}^n$, *то функция f строго выпукла*).

Напомним, что симметричная *n*×*n*-***матрица*** $A = (a_{ij}, i, j = \overline{1,n})$ называется ***неотрицательной*** (***положительной***) и обозначается $A \geq 0$ ($A > 0$), если ***знакоположительна*** (***определенно положительна***) ***квадратичная форма***: $x'Ax \geq 0$ ($x'Ax > 0$, $x \neq 0$) при любых $x \in \mathbf{R}^n$. Справедливы следующие утверждения (***критерии Сильвестра***):

а) *Для неотрицательности матрицы A необходимо и достаточно, чтобы все ее* ***главные миноры*** *были неотрицательны:*

$$\det A\begin{pmatrix} i_1, \ldots, i_s \\ i_1, \ldots, i_s \end{pmatrix} \geq 0,\ 1 \leq i_1 < i_2 < \ldots < i_s \leq n;\ s = \overline{1,n}.$$

Минор называется ***главным***, если он составлен из строк и столбцов с одинаковыми номерами.

б) *Для положительности матрицы A необходимо и достаточно, чтобы все ее* ***последовательные главные миноры*** D_s *были положительны, т. е.* $D_s = \det(a_{ij},\ i = \overline{1,s};\ j = \overline{1,s}) > 0$, $s = \overline{1,n}$.

Аналогично можно записать и следующие критерии:

а) ***отрицательности матрицы***: $(-1)^s D_s > 0$, $s = \overline{1,n}$;

б) ***неположительности матрицы***: $(-1)^s \det A\begin{pmatrix} i_1, \ldots, i_s \\ i_1, \ldots, i_s \end{pmatrix} \geq 0$, $s = \overline{1,n}$.

Функция *f*(*x*), определенная на выпуклом множестве *X*, называется ***вогнутой*** (***строго вогнутой***), если функция –*f*(*x*), $x \in X$, выпукла (строго выпукла).

Основные свойства выпуклых функций:

1. Выпуклая функция *f*(*x*), $x \in X$, непрерывна в точках относительной внутренности множества *X*.

2. Если функции $f_i(x)$, $i = \overline{1,m}$, $x \in X$, выпуклы, то при любых $\alpha_i \geq 0$, $i = \overline{1,m}$, выпуклы и функции $f(x) = \sum_{i=1}^{m} \alpha_i f_i(x)$, $f(x) = \max\{f_i(x),\ i = \overline{1,m}\}$.

3. У выпуклой функции локальный минимум является и глобальным. У строго выпуклой функции минимум может достигаться в единственной точке.

4. Множество уровня $\{x : f(x) \leq c\}$ выпуклой функции *f*(*x*), $x \in \mathbf{R}^n$, или пусто, или выпукло.

5. В каждой точке $x \in \mathbf{R}^n$ выпуклая функция *f*(*x*) имеет производную по любому направлению $l \in \mathbf{R}^n$:

$$\frac{\partial f(x)}{\partial l} = \lim_{t \to +0} \frac{f(x+tl) - f(x)}{t}.$$

6. Вектор $c \in \mathbf{R}^n$ называется ***субградиентом*** выпуклой функции $f(x)$, $x \in \mathbf{R}^n$, в точке x^*, если для любых $x \in \mathbf{R}^n$ выполняется неравенство $f(x) - f(x^*) \ge c'(x - x^*)$. Множество субградиентов в точке x^* называется ***субдифференциалом*** и обозначается $\partial f(x^*)$. В каждой точке субдифференциал – непустой выпуклый компакт (состоит из единственного элемента $\partial f(x)/\partial x$, если $f(x)$ дифференцируема в x). Справедлива формула

$$\frac{\partial f(x)}{\partial l} = \max c'l, \quad c \in \partial f(x).$$

7. Пусть X – выпуклое множество. Вектор $l \in \mathbf{R}^n$, $l \ne 0$, назовем ***допустимым направлением*** в точке $x \in X$ относительно множества X, если найдется число ε_0, $\varepsilon_0 > 0$, такое, что $x + \varepsilon l \in X$ при любом $\varepsilon \in [0;\ \varepsilon_0]$.

Теорема 5.3 (критерий оптимальности). *Для того чтобы точка $x^0 \in X$ была точкой минимума выпуклой функции $f(x)$, $x \in X$, необходимо и достаточно, чтобы для любого допустимого направления l в точке x^0 выполнялось неравенство*

$$\frac{\partial f(x^0)}{\partial l} \ge 0.$$

Из этого критерия следует, что каждая стационарная точка x^* $\left(\dfrac{\partial f(x^*)}{\partial x} = 0\right)$ выпуклой гладкой функции $f(x)$ является ее точкой минимума.