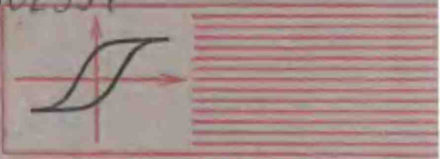

Б. Н. ПШЕНИЧНЫЙ

ВЫПУКЛЫЙ АНАЛИЗ И ЭКСТРЕМАЛЬНЫЕ ЗАДАЧИ

Б. Н. ПШЕНИЧНЫЙ

ВЫПУКЛЫЙ АНАЛИЗ И ЭКСТРЕМАЛЬНЫЕ ЗАДАЧИ

МОСКВА «НАУКА»
ГЛАВНАЯ РЕДАКЦИЯ
ФИЗИКО-МАТЕМАТИЧЕСКОЙ ЛИТЕРАТУРЫ
1980

22.193
П 93
УДК 519.6

Серия «Нелинейный анализ и его приложения»
выпускается под общей редакцией
Н. Н. Боголюбова, М. А. Красносельского, Ю. А. Митропольского

Выпуклый анализ и экстремальные задачи. Пшеничный Б. Н.— М.: Наука. Главная редакция физико-математической литературы, 1980.

Книга написана на основе курса лекций для студентов старших курсов факультета кибернетики Киевского государственного университета. Она посвящена изучению широкого класса экстремальных задач, использующихся при разработке математических моделей экономических и производственных процессов. Отражены хорошо развитые с точки зрения вычислений новые подходы и методы.

П $\frac{20204-010}{053(02)-80}$ 68-80. 1702070000

ОГЛАВЛЕНИЕ

ПРЕДИСЛОВИЕ

Настоящая книга возникла в результате систематизации материала лекций, которые автор читал последние восемь лет на факультете кибернетики и механико-математическом факультете Киевского государственного университета им. Т. Г. Шевченко. Опыт, накопленный при чтении лекций, обусловил структуру книги и ее содержание.

В последние годы теория экстремальных задач приобрела достаточно законченную форму и может быть в систематическом виде изложена для широкого круга читателей. Конечно, это не значит, что книга может читаться совсем легко. Как всякая математическая книга, она требует от читателя активной работы, но в целом овладение излагаемыми ниже понятиями и результатами не требует каких-либо знаний, выходящих за рамки стандартных курсов математического анализа и линейной алгебры. Это достигается за счет того, что все изложение ведется в конечномерных пространствах. Однако этот факт при формулировке теорем и доказательств не используется, если только он не необходим по существу дела. Поэтому читатель, знакомый с функциональным анализом, легко поймет, как тот или иной факт может быть обобщен на случай пространства произвольной размерности.

Первые четыре главы составляют самостоятельное целое и содержат изложение общей теории выпуклого программирования. В основном приводимые здесь факты хорошо известны. Однако при выборе способа изложения был сделан упор на теорию многозначных отображений. Это позволило существенно сократить и унифицировать доказательства и показать, что многие известные теоремы есть в сущности различные формулировки одного и того же факта выпуклого анализа. Кроме того, многозначные отображения в последние годы становятся областью исследований, привлекающей значительное число математиков, и знакомство широкого читателя с ними будет полезным.

Пятая глава содержит основные результаты из теории необходимых условий экстремума. Сюда вошло много результатов, полученных в самое последнее время различными авторами. В качестве объекта для применения в шестой главе рассматривается задача оптимального управления. Это стало уже традиционным в книгах по теории экстремальных задач. Нетрадиционным является постановка задачи оптимального управления в виде экстремальной задачи для дифференциальных включений. Это позволило рассмотреть негладкие задачи оптимального управления и одновременно проиллюстрировать, как работает понятие локально сопряженного отображения к многозначному отображению.

Ограниченный объем книги не позволяет охватить многие интересные приложения общей теории экстремальных задач. Приводимый в конце книги список литературы даст возможность читателю сориентироваться в современной монографической и обзорной литературе и углубить свои знания по отдельным вопросам теории и многообразным ее применениям.

Киев, февраль 1978 г. *Б. Н. Пшеничный*

Глава I

ВЫПУКЛЫЕ МНОЖЕСТВА

В этой главе рассматриваются основные свойства выпуклых множеств.

§ 1. Общие свойства выпуклых множеств

Будем рассматривать векторы x, принадлежащие n-мерному вещественному пространству $X = \mathbf{R}^n$.

Будем рассматривать также его дубликат — пространство $X^* = \mathbf{R}^n$. Обычно в дальнейшем скалярное произведение двух векторов x и x^* ($x \in X$, $x^* \in X^*$) будет обозначаться следующим образом:

$$\langle x, x^* \rangle = \sum_{i=1}^{n} x^i x^{i*},$$

где x^i и x^{i*} — компоненты векторов x и x^* соответственно.

Поясним, почему сделано различие между двумя n-мерными пространствами X и X^*. Дело в том, что элементы из X^* будут играть роль линейных функций, заданных на X. Хорошо известно, да это и несложно показать, что любая линейная функция на X — $l(x)$ — может быть задана некоторым вектором x^* так, что $l(x) = \langle x, x^* \rangle$. Чтобы подчеркнуть тот факт, что вектор x^* задает некоторую линейную функцию, введено такое обозначение.

1. Определение выпуклых множеств.

Определение 1.1. Множество M называется *выпуклым*, если вместе с каждыми своими двумя точками x_1 и x_2 оно содержит и весь отрезок, их соединяющий, т. е. точки вида

$$\lambda_1 x_1 + \lambda_2 x_2 \in M,$$
$$\lambda_1, \lambda_2 \geqslant 0, \quad \lambda_1 + \lambda_2 = 1.$$

Пустое множество $\varnothing$ будем считать выпуклым по определению.

Отметим некоторые простейшие свойства выпуклых множеств.

Лемма 1.1. *Пусть I — произвольное множество индексов i, M_i выпукло для каждого $i \in I$. Тогда множество*

$$M = \bigcap_{i \in I} M_i$$

также выпукло.

Доказательство. Если x_1 и x_2 принадлежат M, то $x_1, x_2 \in M_i$, $i \in I$, и

$$\lambda_1 x_1 + \lambda_2 x_2 \in M_i$$

для любого $i \in I$, т. е. $\lambda_1 x_1 + \lambda_2 x_2 \in M$.

Лемма 1.2. *Если множества M_1 и M_2 выпуклы, c_1, c_2 — вещественные числа, то множество $c_1 M_1 + c_2 M_2$ выпукло.*

Доказательство. Напомним прежде всего, что по определению $c_1 M_1 + c_2 M_2$ есть множество, состоящее из всех точек вида $c_1 x_1 + c_2 x_2$, где $x_1 \in M_1$, $x_2 \in M_2$. Теперь доказательство получается просто из определения выпуклого множества.

Определение 1.2. Будем говорить, что точка x является *выпуклой комбинацией точек* $x_1, \ldots, x_m$, если найдутся числа $\lambda_1, \ldots, \lambda_m$ такие, что

$$\begin{gathered} x = \lambda_1 x_1 + \ldots + \lambda_m x_m, \\ \lambda_i \geqslant 0, \quad i = 1, \ldots, m, \quad \lambda_1 + \ldots + \lambda_m = 1. \end{gathered} \tag{1.1}$$

Лемма 1.3. *Если точки $x_1, \ldots, x_m$ принадлежат выпуклому множеству M, то множество M содержит все их выпуклые комбинации.*

Доказательство удобно провести индукцией по числу точек. Для $m = 2$ утверждение следует из определения. Пусть лемма доказана для $m \leqslant k$. Покажем, что

$$x = \lambda_1 x_1 + \lambda_2 x_2 + \ldots + \lambda_k x_k + \lambda_{k+1} x_{k+1} \in M$$

для любых λ_i таких, что $\lambda_i \geqslant 0$, $\lambda_1 + \ldots + \lambda_{k+1} = 1$. Поскольку можно считать, что $\lambda_i > 0$, то $1 - \lambda_{k+1} = \lambda_1 + \ldots + \lambda_k > 0$. По предположению индукции

$$y = \frac{\lambda_1}{1 - \lambda_{k+1}} x_1 + \ldots + \frac{\lambda_k}{1 - \lambda_{k+1}} x_k \in M,$$

ибо

$$\sum_{i=1}^{k} \frac{\lambda_i}{1-\lambda_{k+1}} = \frac{1}{1-\lambda_{k+1}} \sum_{i=1}^{k} \lambda_i = 1.$$

Но тогда

$$x = (1-\lambda_{k+1})y + \lambda_{k+1}x_{k+1} \in M$$

по определению выпуклого множества.

2. Выпуклые оболочки. Если задано произвольное множество, то по нему можно определить некоторое выпуклое множество, его содержащее.

Определение 1.3. Пересечение всех выпуклых множеств, содержащих данное множество M, называется *выпуклой оболочкой* множества M и обозначается со M.

Согласно лемме 1.1 со M есть выпуклое множество. Так как со M содержит M, то по лемме 1.3 оно должно содержать и все точки вида

$$x = \lambda_1 x_1 + \ldots + \lambda_m x_m, \qquad x_i \in M, \quad i = 1, \ldots, m,$$

$$\lambda_i \geqslant 0, \quad \lambda_1 + \ldots + \lambda_m = 1,$$

т. е. выпуклые комбинации точек из M. Нетрудно убедиться, что множество точек такого вида само выпукло. Итак, со M содержит множество точек вида (1.1), а, с другой стороны, по определению само должно в нем содержаться.

Тем самым доказана

Лемма 1.4. *со M есть множество всех выпуклых комбинаций точек из M.*

Нижеследующая теорема является одной из важнейших в выпуклом анализе.

Теорема 1.1. *В пространстве $X = \mathbf{R}^n$ любую точку* со M *можно представить в виде выпуклой комбинации не более чем $n+1$ точек из M, т. е. для любой $x \in$* со M *найдутся $x_1, \ldots, x_r \in M$ такие, что*

$$\begin{gathered} x = \lambda_1 x_1 + \ldots + \lambda_r x_r, \\ \lambda_1 + \ldots + \lambda_r = 1, \qquad \lambda_1 \geqslant 0, \ldots, \lambda_r \geqslant 0, \end{gathered} \tag{1.2}$$

где $r \leqslant n+1$.

Доказательство. Ясно, что центральным местом в теореме является утверждение о том, что $r \leqslant n+1$.

Возьмем точку вида (1.1) и покажем, что число ненулевых слагаемых в сумме (1.1) можно уменьшить, если $m > n+1$. Достаточно предполагать, что $\lambda_i > 0$. Возьмем $(n+1)$-мерные векторы $(x_i, 1)$, $i = 1, \ldots, m$, в которых первые n компонент совпадают с соответствующими компонентами вектора x_i, а последняя равна 1. Так как число таких векторов $m > n+1$, то они линейно зависимы. Поэтому найдутся не все равные нулю числа α_i, $i = 1, \ldots, m$, такие, что

$$\sum_{i=1}^{m} \alpha_i x_i = 0, \tag{1.3}$$

$$\sum_{i=1}^{m} \alpha_i = 0. \tag{1.4}$$

Среди чисел α_i обязательно есть положительные в силу соотношения (1.4). Положим

$$\varepsilon_0 = \min\left\{\frac{\lambda_i}{\alpha_i} : \alpha_i > 0, \qquad i = 1, \ldots, m\right\}.$$

Пусть минимум достигается при $i = i_0$. Тогда

$$\bar{\lambda}_i = \lambda_i - \varepsilon_0 \alpha_i \geqslant 0$$

для любого $i = 1, \ldots, m$. Это очевидно для $\alpha_i \leqslant 0$, а для $\alpha_i > 0$ следует из выбора ε_0.

Теперь из соотношений

$$\sum_{i=1}^{m} \bar{\lambda}_i x_i = \sum_{i=1}^{m} \lambda_i x_i - \varepsilon_0 \left(\sum_{i=1}^{m} \alpha_i x_i\right) = x,$$

$$\sum_{i=1}^{m} \bar{\lambda}_i = \sum_{i=1}^{m} \lambda_i - \varepsilon_0 \left(\sum_{i=1}^{m} \alpha_i\right) = 1$$

вытекает, что точку x можно представить в виде выпуклой комбинации меньшего числа ненулевых слагаемых. И это уменьшение возможно до тех пор, пока $m > n+1$. Отсюда и следует утверждение теоремы.

3. Топологические свойства.

Определение 1.4. Точка x называется *внутренней точкой* множества M, если существует такое $\varepsilon > 0$, что $x + \varepsilon B \subseteq M$, где B — единичный шар в $\mathbf{R}^n$ с центром в начале координат, т. е.

$$B = \{x\colon \|x\| < 1\}, \qquad \|x\| = \langle x, x\rangle^{1/2}.$$

Множество внутренних точек множества M называется *внутренностью множества* и обозначается $\operatorname{int} M$.

Определение 1.5. Точка x называется *предельной точкой* множества M, если существует последовательность точек $x_k \in M$, сходящаяся к x. Совокупность всех предельных точек множества M называется его *замыканием* и обозначается $\overline{M}$.

Лемма 1.5. *Замыкание и внутренность выпуклого множества выпуклы.*

Доказательство. Если M выпукло, то из $x_1 \in \operatorname{int} M$, $x_2 \in \operatorname{int} M$ следуют включения $x_1 + \varepsilon_1 B \subseteq M$, $x_2 + \varepsilon_2 B \subseteq M$. Пусть $\lambda_1 x_1 + \lambda_2 x_2$ — выпуклая комбинация точек x_1 и x_2. Тогда

$$\lambda_1(x_1 + \varepsilon_1 B) + \lambda_2(x_2 + \varepsilon_2 B) = \\ = \lambda_1 x_1 + \lambda_2 x_2 + (\lambda_1 \varepsilon_1 + \lambda_2 \varepsilon_2) B \subseteq M,$$

т. е. $\lambda_1 x_1 + \lambda_2 x_2$ есть точка $\operatorname{int} M$.

Если $x_1, x_2 \in \overline{M}$, то по определению существуют последовательности точек x_{1k}, $x_{2k} \in M$ такие, что $x_{1k} \to x_1$, $x_{2k} \to x_2$. Пусть $\lambda_1 x_{1k} + \lambda_2 x_{2k}$ — выпуклая комбинация точек x_{1k}, x_{2k}. Тогда

$$\lambda_1 x_1 + \lambda_2 x_2 = \lim_{k \to \infty} (\lambda_1 x_{1k} + \lambda_2 x_{2k}) \in \overline{M},$$

так как $\lambda_1 x_{1k} + \lambda_2 x_{2k} \in M$ в силу выпуклости последнего.

Выпуклые множества обладают тем свойством, что в некотором смысле их всегда можно погрузить в подпространство, относительно которого они уже имеют внутренние точки.

Теорема 1.2. *Выпуклое множество M, лежащее в $\mathbf{R}^n$, либо имеет внутренние точки, либо содержится в подпространстве меньшей размерности, сдвинутом на некоторый вектор.*

Доказательство. Пусть $x_0 \in M$. Рассмотрим все векторы вида $x - x_0$, $x \in M$. Согласно известным теоремам линейной алгебры среди векторов рассматриваемого вида имеется $r \leqslant n$ линейно независимых: $x_1 - x_0, \ldots$ $\ldots, x_r - x_0$. Возможны два случая.

а) $r = n$. Таким образом, имеются n векторов $x_i - x_0$, $x_i \in M$, $i = 1, \ldots, n$, которые линейно независимы. Рассмотрим множество

$$S^n = \{x = \lambda_0 x_0 + \ldots + \lambda_n x_n \colon \lambda_i \geqslant 0,\ \lambda_0 + \ldots + \lambda_n = 1\}.$$

S^n называется *n-мерным симплексом*, натянутым на точки $x_0, x_1, \ldots, x_n$. По лемме 1.3 $S^n \subseteq M$, поэтому, если будет доказано, что S^n имеет внутренние точки, то их будет иметь и M. Докажем, что любая точка $x \in S^n$, коэффициенты λ_i в представлении которой строго положительны, принадлежит $\operatorname{int} S^n$. Рассмотрим систему уравнений относительно λ_i, $i = 1, \ldots, n$:

$$x - x_0 = \sum_{i=1}^{n} \lambda_i (x_i - x_0).$$

Так как векторы $x_i - x_0$ линейно независимы, то эта система имеет единственное решение $\lambda_i(x)$, $i = 1, \ldots, n$, непрерывно зависящее от x (вспомните формулы Крамера для систем уравнений с невырожденным определителем). Поэтому, полагая x равным $\bar{x} = \bar{\lambda}_0 x_0 + \ldots + \bar{\lambda}_n x_n$, $\bar{\lambda}_i > 0$, получим, что $\lambda_i(\bar{x}) = \bar{\lambda}_i > 0$, $i = 1, \ldots, n$, и

$$\lambda_0 = 1 - \sum_{i=1}^{n} \bar{\lambda}_i > 0.$$

Отсюда вытекает, что $\lambda_i(x) > 0$, $i = 1, \ldots, n$, для всех x из некоторой окрестности $\bar{x}$ и

$$\lambda_0(x) = 1 - \sum_{i=1}^{n} \lambda_i(x) > 0.$$

Поэтому для всех точек x из некоторой окрестности точки $\bar{x}$ выполняется включение

$$x = \sum_{i=1}^{n} \lambda_i(x) x_i \in S^n,$$

что доказывает первую часть теоремы.

б) $r < n$. Рассмотрим подпространство X^0, состоящее из векторов

$$y = \sum_{i=1}^{r} \alpha_i (x_i - x_0).$$

По построению $M - x_0 \subseteq X^0$, т. е. $M \subseteq x_0 + X^0$. Теорема доказана.

Построенное в теореме подпространство X^0 r-мерно, и в нем множество $M - x_0$ содержит внутренние точки. Это можно показать так же, как в случае а) доказательства теоремы 1.2. Далее, X^0 не зависит от выбора точки x_0 и векторов $x_i - x_0$, $i = 1, \ldots, r$. В самом деле, любое подпространство, содержащее $M - x_0$, должно содержать векторы $x_i - x_0$, а значит и все X^0. Отсюда следует, что X^0 есть пересечение всех подпространств, содержащих $M - x_0$. Если подпространство X^1 содержит $M - x_0$ для некоторого $x_0 \in M$, то оно содержит и $M - \bar{x}_0$ для любой другой точки $\bar{x}_0 \in M$. Действительно,

$$x - \bar{x}_0 = (x - x_0) - (\bar{x}_0 - x_0),$$

и так как X^1 — подпространство, то оно содержит разность любых двух своих векторов. Таким образом, X^1 содержит $M - x_0$ и $M - \bar{x}_0$ одновременно, т. е. X^0 не зависит от выбора x_0.

Теперь мотивировано следующее определение.

Определение 1.6. Пересечение всех подпространств, содержащих $M - x_0$, где x_0 — любая точка выпуклого множества M, называется *несущим подпространством* множества M и обозначается $\operatorname{Lin} M$; $\operatorname{Aff} M = x_0 + \operatorname{Lin} M$ называется *аффинной оболочкой* M.

Определение 1.7. Точка x называется *относительно внутренней* точкой выпуклого множества M, если $x + \operatorname{Lin} M \cap (\varepsilon B) \subseteq M$, т. е. x содержится в M вместе с шаром радиуса $\varepsilon > 0$, лежащим в $\operatorname{Lin} M$. Множество точек выпуклого множества, внутренних относительно его несущего подпространства, называется *относительной внутренностью* и обозначается $\operatorname{ri} M$.

Лемма 1.6. $\operatorname{ri} \overline{M} = \operatorname{ri} M$.

Доказательство. Так как $\operatorname{Lin} M$ есть замкнутое множество, содержащее M, то $\operatorname{Lin} M \supseteq \overline{M}$. Легко видеть, что $\operatorname{Lin} M = \operatorname{Lin} \overline{M}$. Очевидно, что $\operatorname{ri} \overline{M} \supseteq \operatorname{ri} M$. Докажем обратное включение. Пусть $x \in \operatorname{ri} \overline{M}$, а $e_1, \ldots, e_r$ — базис в $\operatorname{Lin} M$. Тогда при малом ε имеем

$$y_k = x + \varepsilon\left(e_k - \frac{1}{r+1} e\right) \in \overline{M}, \qquad k = 1, \ldots, r,$$

$$y_0 = x - \frac{\varepsilon}{r+1} e \in \overline{M},$$

где $e = e_1 + \ldots + e_r$. Векторы $y_k - y_0 = \varepsilon e_k$ линейно независимы, и

$$x = \frac{1}{r+1} y_0 + \ldots + \frac{1}{r+1} y_r.$$

Последнее равенство означает, что x есть внутренняя точка симплекса, натянутого на точки $y_0, \ldots, y_r$. Если взять точки $\overline{y}_k \in M$, достаточно близкие к y_k, то получится, что x есть внутренняя точка симплекса, натянутого на точки $\overline{y}_k \in M$, а значит, относительно внутренняя точка M; поэтому $\operatorname{ri} \overline{M} \subseteq \operatorname{ri} M$. Лемма доказана.

Теорема 1.3. *Пусть M — выпуклое множество. Если $x_1 \in \overline{M}$, $x_2 \in \operatorname{ri} M$, то при всех $0 < \lambda \leqslant 1$ $(1-\lambda)x_1 + \lambda x_2 \in \operatorname{ri} M$. Кроме того, $\overline{M} = \overline{\operatorname{ri} M}$.*

Доказательство. Если $x_2 \in \operatorname{ri} M$, то $x_2 + \operatorname{Lin} M \cap (\varepsilon B) \subseteq M$. Поэтому и из выпуклости M следует, что

$$(1-\lambda)x_1 + \lambda(x_2 + \operatorname{Lin} M \cap (\varepsilon B)) = (1-\lambda)x_1 + \lambda x_2 + \operatorname{Lin} M \cap (\lambda\varepsilon B) \subseteq \overline{M},$$

т. е. $(1-\lambda)x_1 + \lambda x_2 \in \operatorname{ri} M$, так как $\operatorname{ri} M = \operatorname{ri} \overline{M}$ по лемме 1.6.

Пусть $x_0 \in \overline{M}$, $x_k \in M$, $x_k \to x_0$, $\lambda_k \to 0$. Тогда $(1-\lambda_k)x_k + \lambda_k y \in \operatorname{ri} M$, если $y \in \operatorname{ri} M$; поэтому x_0 есть предельная точка для $\operatorname{ri} M$. Итак, $\overline{M} \subseteq \overline{\operatorname{ri} M}$. Обратное включение очевидно.

Определение 1.8. Размерность пространства $\operatorname{Lin} M$ называется *размерностью выпуклого множества M* и обозначается $\dim M$.

Теорема 1.4. *Если для выпуклых множеств M_1 и M_2 выполняется условие* $\operatorname{ri} M_1 \cap \operatorname{ri} M_2 \neq \varnothing$, то

$$\operatorname{Lin} M_1 \cap \operatorname{Lin} M_2 = \operatorname{Lin}(M_1 \cap M_2),$$

$$\operatorname{ri} M_1 \cap \operatorname{ri} M_2 = \operatorname{ri}(M_1 \cap M_2).$$

Доказательство. Без ограничения общности будем считать, что $0 \in \operatorname{ri} M_1 \cap \operatorname{ri} M_2$. Так как в этом случае $\operatorname{Lin} M_1 \supseteq M_1$, $\operatorname{Lin} M_2 \supseteq M_2$, то

$$\operatorname{Lin} M_1 \cap \operatorname{Lin} M_2 \supseteq M_1 \cap M_2,$$

и поэтому

$$\operatorname{Lin} M_1 \cap \operatorname{Lin} M_2 \supseteq \operatorname{Lin}(M_1 \cap M_2).$$

Обратно, пусть $z \in \operatorname{Lin} M_1 \cap \operatorname{Lin} M_2$. Тогда при достаточно малом $\lambda > 0$ $\lambda z \in M_1$ и $\lambda z \in M_2$, так как $0 \in \operatorname{ri} M_1 \cap \operatorname{ri} M_2$. Отсюда вытекает, что

$$\lambda z \in M_1 \cap M_2,$$

а, значит, $\lambda z \in \operatorname{Lin}(M_1 \cap M_2)$. Так как $\operatorname{Lin}(M_1 \cap M_2)$ — подпространство, то $z \in \operatorname{Lin}(M_1 \cap M_2)$.

Докажем вторую часть утверждения. Если $x \in \operatorname{ri} M_1 \cap \operatorname{ri} M_2$, то $x + \operatorname{Lin} M_1 \cap (\varepsilon B) \subseteq M_1$, $x + \operatorname{Lin} M_2 \cap (\varepsilon B) \subseteq M_2$, так что

$$(x + \operatorname{Lin} M_1 \cap (\varepsilon B)) \cap (x + \operatorname{Lin} M_2 \cap (\varepsilon B)) = \\ = x + \operatorname{Lin}(M_1 \cap M_2) \cap (\varepsilon B) \subseteq M_1 \cap M_2$$

и $x \in \operatorname{ri}(M_1 \cap M_2)$. Поэтому

$$\operatorname{ri}(M_1 \cap M_2) \supseteq \operatorname{ri} M_1 \cap \operatorname{ri} M_2.$$

Пусть теперь $x \in \operatorname{ri}(M_1 \cap M_2)$. Так как $0 \in \operatorname{ri} M_1$, то $(1-\lambda)x \in \operatorname{ri} M_j$, $j = 1, 2$, для $0 < \lambda \leqslant 1$. Поэтому $(1-\lambda)x \in \operatorname{ri} M_1 \cap \operatorname{ri} M_2$. Устремляя λ к нулю, получаем $x \in \overline{\operatorname{ri} M_1 \cap \operatorname{ri} M_2}$. Таким образом,

$$\operatorname{ri} M_1 \cap \operatorname{ri} M_2 \subseteq \operatorname{ri}(M_1 \cap M_2) \subseteq \overline{\operatorname{ri} M_1 \cap \operatorname{ri} M_2}.$$

Пусть $e_1, e_2, \ldots, e_r$ — базис в $\operatorname{Lin}(M_1 \cap M_2)$, $e = e_1 + \ldots + e_r$, $x \in \operatorname{ri}(M_1 \cap M_2)$. Тогда при достаточно малом $\varepsilon > 0$ все точки

$$y_k = x + \varepsilon\left(e_k - \frac{1}{r+1}e\right), \qquad k = 1, \ldots, r,$$

$$y_0 = x - \frac{\varepsilon}{r+1}e$$

принадлежат $\operatorname{ri}(M_1 \cap M_2)$, а значит и $\overline{\operatorname{ri} M_1 \cap \operatorname{ri} M_2}$. В то же время

$$x = \frac{1}{r+1}y_0 + \ldots + \frac{1}{r+1}y_r$$

является внутренней точкой симплекса, натянутого на точки $y_0, \ldots, y_k$, относительно подпространства $\operatorname{Lin}(M_1 \cap \cap M_2)$. Если взять точки $\bar{y}_k \in \operatorname{ri} M_1 \cap \operatorname{ri} M_2$ достаточно близкими к y_k, то ясно, что точка x будет внутренней точкой симплекса, натянутого на $\bar{y}_k$, а потому будет принадлежать $\operatorname{ri} M_1 \cap \operatorname{ri} M_2$. Тем самым доказано, что любая точка x из $\operatorname{ri}(M_1 \cap M_2)$ принадлежит одновременно и $\operatorname{ri} M_1 \cap \operatorname{ri} M_2$. В соответствии с предыдущим это означает, что

$$\operatorname{ri}(M_1 \cap M_2) = \operatorname{ri} M_1 \cap \operatorname{ri} M_2,$$

что и требовалось доказать.

Теорема 1.5. *Пусть M — выпуклое множество, и пусть $x_0 \in \overline{M}$, но $x_0 \notin M$. Тогда в любой окрестности x_0 найдутся точки, не принадлежащие $\overline{M}$.*

Доказательство. Возьмем точку $y \in \operatorname{ri} M$. Тогда точки луча $y + \lambda(x_0 - y)$, $\lambda \geqslant 0$, при $\lambda > 1$ не принадлежат $\overline{M}$. В самом деле, если при $\lambda > 1$ $x_1 = y + \lambda(x_0 - y) \in \in \overline{M}$, то

$$x_0 = \frac{1}{\lambda} x_1 + \left(1 - \frac{1}{\lambda}\right) y \in \operatorname{ri} M$$

по теореме 1.3, что противоречит тому, что $x_0 \notin M$.

4. Замкнутые выпуклые оболочки. Пересечение любого числа выпуклых множеств есть множество выпуклое. Подобным же свойством обладают и замкнутые множества, поэтому целесообразно ввести следующее определение.

Определение 1.9. Пересечение всех замкнутых выпуклостей множеств, содержащих данное множество M, называется *замкнутой выпуклой оболочкой* M и обозначается $\overline{\operatorname{co}} M$.

Теорема 1.6. $\overline{\operatorname{co}} M = \overline{\operatorname{co} M}$.

Доказательство. Ясно, что $\overline{\operatorname{co}} M \supseteq \operatorname{co} M$, так как в образовании $\operatorname{co} M$ участвуют все выпуклые множества, а не только замкнутые. Отсюда вытекает, что $\overline{\operatorname{co}} M \supseteq \overline{\operatorname{co} M}$.

Обратно, $\overline{\operatorname{co} M}$ есть выпуклое замкнутое множество. Поэтому $\overline{\operatorname{co} M} \supseteq \overline{\operatorname{co}} M$, что завершает доказательство.

Теорема 1.7. *Выпуклая оболочка компакта есть компакт.*

Доказательство. Напомним, что в $\mathbf{R}^n$ компакт есть ограниченное замкнутое множество. Если $x \in \operatorname{co} M$, где M — компакт, то по теореме 1.1

$$x = \sum_{i=1}^{n+1} \lambda_i x_i, \quad x_i \in M, \quad \lambda_i \geqslant 0, \quad \sum_{i=1}^{n+1} \lambda_i = 1,$$

поэтому

$$\|x\| \leqslant \sum_{i=1}^{n+1} \lambda_i \|x_i\| \leqslant c,$$

где c — константа такая, что $\|x\| \leqslant c$ для любого $x \in M$. Итак, $\operatorname{co} M$ — ограниченное множество. Покажем, что оно замкнуто. Пусть

$$x_k = \sum_{i=1}^{n+1} \lambda_{ik} x_{ik}, \quad x_{ik} \in M, \quad \sum_{i=1}^{n+1} \lambda_{ik} = 1. \tag{1.5}$$

Так как последовательности λ_{ik} и x_{ik} ограничены, то из них можно выбрать сходящиеся последовательности. Не ограничивая общности, можно считать, что $\lambda_{ik} \to \lambda_{i0}$, $x_{ik} \to x_{i0} \in M$, поскольку M — компакт. Отсюда, переходя к пределу в формулах (1.5), получаем

$$x_0 = \sum_{i=1}^{n+1} \lambda_{i0} x_{i0}, \quad x_{i0} \in M, \quad \sum_{i=1}^{n+1} \lambda_{i0} = 1.$$

Это означает, что $x_0 \in \operatorname{co} M$. Тем самым замкнутость $\operatorname{co} M$ доказана.

§ 2. Теорема отделимости

Свойство отделимости выпуклых множеств, состоящее в том, что между двумя выпуклыми непересекающимися множествами можно провести гиперплоскость так, чтобы эти множества остались по разные стороны от нее, является одним из основных свойств, широко используемым в приложениях. Здесь будут даны два подхода к построению теорем отделимости. Первый из них существенно связан с конечномерностью рассматриваемого пространства X. Второй основан на известной теореме Хана — Банаха из функционального анализа.

1. Основные теоремы.

Теорема 2.1. *Пусть M — выпуклое множество и точка x_0 не принадлежит его замыканию. Тогда суще-*

ствуют точка x^* и число $\varepsilon > 0$ такие, что

$$\langle x, x^* \rangle \leqslant \langle x_0, x^* \rangle - \varepsilon$$

для всех $x \in M$.

Доказательство. Пусть y — точка из $\overline{M}$, ближайшая к x_0. В силу замкнутости $\overline{M}$ такая точка существует и удовлетворяет неравенству

$$\|x - x_0\| \geqslant \|y - x_0\|, \qquad \forall x \in \overline{M}.$$

Так как $\lambda x + (1-\lambda)y = y + \lambda(x-y) \in \overline{M}$ в силу выпуклости M для всех $\lambda \in [0, 1]$, то

$$\|\lambda x + (1-\lambda)y - x_0\|^2 =$$
$$= \langle y - x_0 + \lambda(x-y),\ y - x_0 + \lambda(x-y) \rangle =$$
$$= \|y - x_0\|^2 + 2\lambda\langle x - y, y - x_0 \rangle + \lambda^2 \|x - y\|^2 \geqslant \|y - x_0\|^2.$$

Отсюда после простых преобразований получаем

$$2\langle x - y, y - x_0 \rangle + \lambda \|x - y\|^2 \geqslant 0, \qquad \lambda \in [0, 1].$$

В частности, при $\lambda = 0$

$$\langle x - y,\ y - x_0 \rangle \geqslant 0. \tag{2.1}$$

Положим $x^* = x_0 - y$, $\varepsilon = \|x^*\|^2$. Так как $x_0 \notin \overline{M}$, то $y \neq x_0$ и $x^* \neq 0$, поэтому $\varepsilon > 0$. Теперь неравенство (2.1) можно переписать в виде

$$\langle x, x^* \rangle \leqslant \langle y, x^* \rangle = \langle x_0, x^* \rangle - \langle x^*, x^* \rangle = \langle x_0, x^* \rangle - \varepsilon.$$

Так как точка $x \in M$ была выбрана произвольно, то доказательство завершено.

Теорема 2.2. *Пусть M — выпуклое множество и $x_0 \notin M$. Тогда существует точка $x^* \neq 0$, удовлетворяющая неравенству*

$$\langle x, x^* \rangle \leqslant \langle x_0, x^* \rangle$$

для всех $x \in M$.

Доказательство. Если $x_0 \notin \overline{M}$, то утверждение следует из предыдущей теоремы. Если же $x_0 \in \overline{M}$, то согласно теореме 1.5 существует последовательность $x_k \to x_0$, $x_k \notin \overline{M}$. Применим предыдущую теорему к M и x_k:

$$\langle x, x_k^* \rangle \leqslant \langle x_k, x_k^* \rangle - \varepsilon_k, \qquad \varepsilon_k > 0.$$

Так как $\varepsilon_k > 0$, то $x_k^* \neq 0$. Отбрасывая ε_k и нормируя x_k^*, получим

$$\left\langle x, \frac{x_k^*}{\|x_k^*\|} \right\rangle \leqslant \left\langle x_k, \frac{x_k^*}{\|x_k^*\|} \right\rangle. \tag{2.2}$$

Из ограниченной по норме последовательности можно выбрать сходящуюся. Без ограничения общности будем считать, что

$$\frac{x_k^*}{\|x_k^*\|} \to x_0^*, \quad \|x_0^*\| = 1.$$

Переходя теперь в неравенстве (2.2) к пределу, получим

$$\langle x, x_0^* \rangle \leqslant \langle x_0, x_0^* \rangle,$$

что и требовалось доказать.

Теорема 2.3. *Если M_1 и M_2 — выпуклые непересекающиеся множества, то существует точка $x^* \neq 0$ такая, что*

$$\langle x_1, x^* \rangle \leqslant \langle x_2, x^* \rangle$$

для всех $x_1 \in M_1$, $x_2 \in M_2$.

Доказательство. Пусть $M = M_1 - M_2$. Так как M_1 и M_2 не пересекаются, то точка $x_0 = 0$ не принадлежит M. Применяя предыдущую теорему к M и $x_0 = 0$, получим, что существует точка $x^* \neq 0$ такая, что

$$\langle x, x^* \rangle \leqslant \langle 0, x^* \rangle \tag{2.3}$$

для всех $x \in M$, т. е. для $x = x_1 - x_2$, $x_1 \in M_1$, $x_2 \in M_2$. Подставляя выражение для x в неравенство (2.3), получим

$$\langle x_1 - x_2, x^* \rangle \leqslant 0, \quad x_1 \in M_1, \quad x_2 \in M_2,$$

что и требовалось доказать.

Теорема 2.4. *Если M_1 и M_2 — выпуклые замкнутые непересекающиеся множества, причем одно из них компактно, то существуют точка x^* и число $\varepsilon > 0$ такие, что*

$$\langle x_1, x^* \rangle \leqslant \langle x_2, x^* \rangle - \varepsilon$$

для всех $x_1 \in M_1$, $x_2 \in M_2$.

Доказательство. Как и в предыдущем доказательстве, точка $x_0 = 0$ не принадлежит $M = M_1 - M_2$. Но из того, что M_1 и M_2 замкнуты, а одно из них компактно, следует, что M — замкнутое множество. Доказательство этого несложного факта предоставляется читателю. Таким образом, точка $x_0 = 0$ не принадлежит замкнутому множеству M. Для завершения доказательства достаточно применить теорему 2.1.

2. Функция Минковского. Эта функция представляет интерес сама по себе, как пример выпуклой функции, которая систематически будет изучаться дальше. В этом параграфе она будет использоваться при доказательстве теоремы отделимости на основе теоремы Хана — Банаха.

Определение 2.1. Пусть M — выпуклое множество и $0 \in \operatorname{int} M$. Положим

$$r_M(x) = \inf\left\{\alpha:\ \alpha > 0,\ \frac{x}{\alpha} \in M\right\}.$$

Функция $r_M(x)$, заданная на всем пространстве X, называется *функцией Минковского*.

Так как $0 \in \operatorname{int} M$, то $\alpha^{-1}x \in M$ при большом α для любого x, поэтому $r_M(x)$ — всегда конечное число. Кроме того, по определению $r_M(x) \geqslant 0$.

Лемма 2.1. *Пусть M — выпуклое множество и $0 \in \operatorname{int} M$. Тогда:*

а) *если $\lambda \geqslant 0$, то $r_M(\lambda x) = \lambda r_M(x)$;*

б) *если $x \in M$, то $r_M(x) \leqslant 1$, а если $x \notin M$, то $r_M(x) \geqslant 1$;*

в) $r_M(x+y) \leqslant r_M(x) + r_M(y)$.

Доказательство. а) Пусть $\lambda\alpha_1 = \alpha$. Имеем

$$r_M(\lambda x) = \inf_{\alpha}\left\{\alpha:\ \alpha > 0,\ \frac{\lambda x}{\alpha} \in M\right\} =$$
$$= \inf_{\alpha_1}\{\lambda\alpha_1:\ \alpha_1 > 0,\ \alpha_1^{-1}x \in M\} =$$
$$= \lambda \inf_{\alpha_1}\{\alpha_1:\ \alpha_1 > 0,\ \alpha_1^{-1}x \in M\} = \lambda r_M(x).$$

б) Если $x \in M$, то $x/1 \in M$, так что $r_M(x) \leqslant 1$. Пусть теперь $x \notin M$. Если $r_M(x) < 1$, то существует $\alpha < 1$ такое, что $\alpha^{-1}x \in M$. Так как $0 \in M$ и M выпукло, то

$$x = (1-\alpha)\cdot 0 + \alpha(\alpha^{-1}x) \in M,$$

что противоречит исходному предположению. Итак, действительно $r_M(x) \geqslant 1$ для $x \notin M$.

в) Пусть $\gamma > r_M(x) + r_M(y)$. Тогда найдутся α и β такие, что $\gamma = \alpha + \beta$, $\alpha > r_M(x)$, $\beta > r_M(y)$. Из неравенства $\alpha > r_M(x)$ вытекает, что $\alpha^{-1}x \in M$. В самом деле, по определению $r_M(x)$ существует число α_1 такое, что $\alpha > \alpha_1 > r_M(x)$ и $\alpha_1^{-1}x \in M$; поэтому $\alpha^{-1}x = (1 - \alpha^{-1}\alpha_1) \cdot 0 + \alpha^{-1}\alpha_1 (\alpha_1^{-1}x) \in M$ в силу выпуклости M и того, что по условию $0 \in M$. Аналогично доказывается, что $\beta^{-1}y \in M$.

Отсюда вытекает справедливость соотношения

$$\frac{x+y}{\gamma} = \frac{x+y}{\alpha+\beta} = \frac{\alpha}{\alpha+\beta}(\alpha^{-1}x) + \frac{\beta}{\alpha+\beta}(\beta^{-1}y) \in M,$$

следовательно, $r_M(x+y) \leqslant \gamma$. Но так как γ было взято произвольным числом, большим $r_M(x) + r_M(y)$, то

$$r_M(x+y) \leqslant r_M(x) + r_M(y).$$

Лемма доказана.

3. Теорема Хана — Банаха. Эта теорема является одной из основных в функциональном анализе. Она практически эквивалентна теореме отделимости.

Теорема 2.5. *Пусть задана функция $p(x)$, определенная на X и такая, что*

$$p(x+y) \leqslant p(x) + p(y), \qquad p(\lambda x) = \lambda p(x)$$

для $\lambda \geqslant 0$. Пусть, кроме того, задана линейная функция $f(x)$ на подпространстве $X^0 \subseteq X$ и $f(x) \leqslant p(x)$ для всех $x \in X^0$. Тогда существует линейная функция $F(x)$, определенная на всем пространстве и такая, что $f(x) = F(x)$ для $x \in X^0$ и $F(x) \leqslant p(x)$ для $x \in X$.

Доказательство. Пусть $x_0 \notin X^0$. По условию для $x_1, x_2 \in X^0$

$$f(x_2) - f(x_1) = f(x_2 - x_1) \leqslant p(x_2 - x_1) =$$
$$= p((x_2 + x_0) + (-x_1 - x_0)) \leqslant p(x_2 + x_0) + p(-x_1 - x_0),$$

откуда

$$-p(-x_1 - x_0) - f(x_1) \leqslant p(x_2 + x_0) - f(x_2). \qquad (2.4)$$

Положим

$$m = \sup_{x_1 \in X^0} [-p(-x_1 - x_0) - f(x_1)], \qquad (2.5)$$

$$M = \inf_{x_2 \in X^0} [p(x_2 + x_0) - f(x_2)]. \qquad (2.6)$$

Очевидно, что числа m и M конечны и $m \leqslant M$. Если взять r_0 между m и M, то для любого $x \in X^0$ будем иметь

$$-p(-x-x_0)-f(x) \leqslant r_0 \leqslant p(x+x_0)-f(x). \qquad (2.7)$$

Рассмотрим теперь множество точек X^1 вида $x+\alpha x_0$, где $x \in X^0$, $\alpha \in \mathbf{R}^1$. Очевидно, что X^1 — подпространство, так как оно замкнуто относительно операций сложения векторов и умножения на число. Далее, если $y = x + \alpha x_0$, то x и α определены однозначно. В самом деле, если существует другое представление $y = x_1 + \alpha_1 x_0$, $x_1 \in X^0$, то

$$x_1 - x = (\alpha - \alpha_1) x_0,$$

и если $\alpha - \alpha_1 \neq 0$, то $x_0 \in X^0$, что противоречит предположению.

Для $y \in X^1$ положим

$$\varphi(y) = f(x) + \alpha r_0.$$

Функция $\varphi(y)$ определена однозначно, линейна и $\varphi(y) = f(y)$ для $y \in X^0$. Покажем, что

$$\varphi(y) \leqslant p(y).$$

Это неравенство можно переписать в виде

$$f(x) + \alpha r_0 \leqslant p(x + \alpha x_0). \qquad (2.8)$$

Если $\alpha > 0$, то (2.8) переходит в неравенство

$$r_0 \leqslant p\left(\frac{x}{\alpha} + x_0\right) - f\left(\frac{x}{\alpha}\right),$$

эквивалентное правому неравенству (2.7), если заменить в нем x на $\alpha^{-1}x$. Аналогично, если $\alpha < 0$, то (2.8) переходит в левое из неравенств (2.7).

Итак, показано, что функция $f(x)$ может быть продолжена с сохранением всех свойств на подпространство X^1, $X^1 \supset X^0$, размерность которого на единицу больше размерности X^0. Повторяя это построение, мы за конечное число шагов построим функцию $f(x)$ во всем пространстве X.

Покажем теперь, как из этой теоремы следует теорема об отделимости.

Теорема 2.6. *Пусть M — выпуклое множество, $\operatorname{int} M \neq \varnothing$ и $x_0 \notin M$. Тогда существует линейная функция*

$F(x)$ такая, что

$$F(x) \leqslant F(x_0), \qquad F(x_0) \geqslant 1,$$

для всех $x \in M$.

Доказательство. Так как $\operatorname{int} M \neq \varnothing$, то без ограничения общности можно считать, что $0 \in \operatorname{int} M$. Иначе можно было бы взять $\bar{x} \in \operatorname{int} M$ и рассмотреть $M - \bar{x}$ и точку $x_0 - \bar{x}$.

Так как $0 \in \operatorname{int} M$, то определена функция Минковского $r_M(x)$. По лемме 2.1 $r_M(x) \leqslant 1$ для $x \in M$ и $r_M(x) \geqslant 1$. Определим теперь на одномерном подпространстве X^0, состоящем из элементов $x = \alpha x_0$, $\alpha \in \mathbf{R}^1$, линейную функцию $f(x)$ следующей формулой:

$$f(x) = \alpha r_M(x_0).$$

Согласно лемме 2.1 функция $r_M(x)$ и линейная функция $f(x)$ на подпространстве X^0 удовлетворяют всем требованиям теоремы 2.5, так как легко проверить справедливость неравенства

$$f(x) \leqslant r_M(x), \qquad x \in X^0.$$

Применяя теорему 2.5, получим линейную функцию $F(x)$ такую, что

$$F(x) \leqslant r_M(x), \qquad x \in X,$$

$$F(x_0) = f(x_0) = r_M(x_0) \geqslant 1.$$

Но так как $r_M(x) \leqslant 1$ для $x \in M$, то

$$F(x) \leqslant r_M(x) \leqslant 1 \leqslant r_M(x_0) = F(x_0),$$

что завершает доказательство.

Теорема 2.6 почти эквивалентна теореме 2.2, так как в конечномерном пространстве $X = \mathbf{R}^n$ любая линейная функция задается скалярным произведением. Остается условие $\operatorname{int} M \neq \varnothing$, не учитывать которое в общем случае нельзя. Однако для конечномерного пространства X это можно проделать, рассмотрев M в $\operatorname{Lin} M$, где его внутренность не пуста.

4. Опорная функция. Выпуклое множество может быть охарактеризовано не только функцией Минковского, но и некоторой другой функцией.

Определение 2.2. Функция

$$W_M(x^*) = \sup_x \{\langle x, x^*\rangle : x \in M\}$$

называется *опорной функцией* множества M.

Теорема 2.7. *Пусть M — замкнутое выпуклое множество. Тогда $x \in M$ в том и только том случае, если*

$$\langle x, x^*\rangle \leqslant W_M(x^*) \tag{2.9}$$

для всех $x^* \in M$.

Доказательство. Если $x \in M$, то неравенство (2.9) выполняется в силу определения опорной функции. Обратно, пусть x_0 удовлетворяет неравенству (2.9), но $x_0 \notin M$. Тогда по теореме 2.1 существуют точка x^* и число $\varepsilon > 0$ такие, что

$$\langle x, x^*\rangle \leqslant \langle x_0, x^*\rangle - \varepsilon.$$

Взяв верхнюю грань в левой части неравенства по всем $x \in M$, получим

$$W_M(x^*) \leqslant \langle x_0, x^*\rangle - \varepsilon,$$

что противоречит предположению.

§ 3. Выпуклые конусы

Выпуклый конус является одним из основных объектов изучения в теории экстремальных задач. Изучение его свойств связано с вычислением сопряженного к нему конуса.

1. Определения и основные свойства.

Определение 3.1. Выпуклое множество K называется *выпуклым конусом*, если из того, что $x \in K$ и $\lambda > 0$, следует, что $\lambda x \in K$.

Лемма 3.1. *Если $x_1, x_2, \ldots, x_m \in K$, $\lambda_1 > 0, \ldots, \lambda_m > 0$, то*

$$\lambda_1 x_1 + \lambda_2 x_2 + \ldots + \lambda_m x_m \in K.$$

Доказательство следует из формулы

$$\lambda_1 x_1 + \ldots + \lambda_m x_m = \lambda\left(\frac{\lambda_1}{\lambda} x_1 + \ldots + \frac{\lambda_m}{\lambda} x_m\right),$$

где $\lambda = \lambda_1 + \ldots + \lambda_m$, леммы 1.3 и определения конуса.

Итак, выпуклый конус содержит любые линейные комбинации своих элементов с положительными коэффициентами.

Так как большинство рассматриваемых далее конусов будут выпуклыми, то будем их называть просто конусами, как правило, опуская слово «выпуклый».

Определение 3.2. Множество векторов $x^* \in X^*$ таких, что

$$\langle x, x^* \rangle \geqslant 0 \quad \text{для всех} \quad x \in K,$$

называется *конусом, сопряженным* к конусу K, и обозначается K^*. Короче это можно записать следующим образом:

$$K^* = \{x^* \in X^* : \ \langle x, x^* \rangle \geqslant 0, \quad x \in K\}.$$

Приведем ряд свойств сопряженных конусов. Очевидно, что K^* — выпуклый конус.

Лемма 3.2. *K^* — замкнутый конус.*

Доказательство. Если $x_k^* \in K^*$, $x_k^* \to x_0^*$, то при переходе к пределу в неравенстве получаем

$$\langle x, x_0^* \rangle \geqslant 0, \qquad x \in K,$$

т. е. $x_0^* \in K^*$.

Лемма 3.3. *K и $\overline{K}$ имеют одинаковые сопряженные конусы, т. е. $K^* = (\overline{K})^*$.*

Доказательство. Если $x^* \in (\overline{K})^*$, то $\langle x, x^* \rangle \geqslant 0$ для всех $x \in \overline{K}$, т. е. и для $x \in K$, поэтому $x^* \in K^*$. Обратно, если $x^* \in K^*$, то $\langle x, x^* \rangle$ — неотрицательная функция на $\overline{K}$ по непрерывности, ибо элементы из $\overline{K}$ являются предельными для элементов из K.

Лемма 3.4. *Если K — замкнутый конус и $\langle x, x^* \rangle \geqslant 0$ для всех $x^* \in K^*$, то $x \in K$.*

Доказательство. Предположим противное. Пусть $x_0 \notin K$, но $\langle x_0, x^* \rangle \geqslant 0$ для всех $x^* \in K^*$. Так как $x_0 \notin K$, то по теореме 2.4 существуют такие $\overline{x}^*$ и $\varepsilon > 0$, что

$$\langle x_0, \overline{x}^* \rangle \leqslant \langle x, \overline{x}^* \rangle - \varepsilon \tag{3.1}$$

для всех $x \in K$.

Покажем, что

$$\inf_{x \in K} \langle x, \bar{x}^* \rangle = 0. \tag{3.2}$$

Действительно, так как вместе с x конус содержит и точки λx, $\lambda > 0$, то, устремляя λ к нулю, убеждаемся, что любой конус содержит элементы, как угодно близкие к нулю. В частности, замкнутый конус всегда содержит точку нуль, поэтому точная нижняя грань произведения $\langle x, \bar{x}^* \rangle$ не может быть больше нуля. С другой стороны, произведение $\langle x, \bar{x}^* \rangle$ ограничено снизу согласно оценке (3.1). Покажем, что $\langle x, \bar{x}^* \rangle \geqslant 0$. Если бы существовала точка $x_1 \in K$, для которой $\langle x_1, \bar{x}^* \rangle < 0$, то, взяв $x = \lambda x_1$ и устремив λ к $+\infty$, получили бы, что произведение $\langle x, \bar{x}^* \rangle$ стремится к $-\infty$, что противоречит ограниченности снизу. Тем самым соотношение (3.2) доказано. Из него следует, что $\bar{x}^* \in K^*$.

Полагая в неравенстве (3.1) $x = 0$, получаем, что $\langle x_0, \bar{x}^* \rangle < -\varepsilon$ в противоречии с допущением. Итак, $x_0 \in K$, что и требовалось доказать.

Так как K^* — выпуклый конус в X^* — дубликате $X = \mathbf{R}^n$,— то можно поставить вопрос о вычислении сопряженного к нему конуса $(K^*)^*$, т. е. K^{**}. Этот последний будем брать в исходном пространстве X.

Лемма 3.5. *Если K — замкнутый конус, то $K^{**} = K$.*

Доказательство. В соответствии с определением 3.2

$$K^{**} = \{x\colon \langle x, x^* \rangle \geqslant 0, \quad x^* \in K^*\}.$$

Но если $x \in K$, то $\langle x, x^* \rangle \geqslant 0$ для $x^* \in K^*$. Обратно, если K замкнут, то согласно предыдущей лемме из неравенства $\langle x, x^* \rangle \geqslant 0$ для всех $x^* \in K^*$ следует, что $x \in K$.

В общем случае $K^{**} = \bar{K}$. Это следует из леммы 3.3, ибо $K^{**} = (\bar{K})^*$ и поэтому $K^{**} = (\bar{K})^{**} = \bar{K}$.

Лемма 3.6. *Пусть K_1 и K_2 — конусы. Тогда $K_1 + K_2$ — также выпуклый конус и*

$$(K_1 + K_2)^* = K_1^* \cap K_2^*.$$

Доказательство. То, что $K_1 + K_2$ — выпуклый конус, следует из леммы 1.2 и легко проверяемого факта, что K_1+K_2 вместе с $x=x_1+x_2$, $x_1 \in K_1$, $x_2 \in K_2$ содержит и $\lambda x=\lambda x_1+\lambda x_2$, $\lambda>0$.

Далее, $x^* \in (K_1+K_2)^*$ тогда и только тогда, когда

$$\langle x_1+x_2,\ x^*\rangle \geqslant 0, \quad x_1 \in K_1, \quad x_2 \in K_2.$$

Поскольку x_1 и x_2 меняются независимо и при необходимости одно из них можно устремить к нулю, то последнее неравенство эквивалентно следующим двум:

$$\langle x_1,\ x^*\rangle \geqslant 0, \qquad x_1 \in K_1,$$
$$\langle x_2,\ x^*\rangle \geqslant 0, \qquad x_2 \in K_2,$$

т. е. $x^* \in K_1^*$ и $x^* \in K_2^*$. Итак, доказано, что $x^* \in$ $\in (K_1+K_2)^*$ тогда и только тогда, когда $x^* \in K_1^* \cap K_2^*$.

Следующая лемма очень важна.

Лемма 3.7. *Для замкнутых конусов K_1 и K_2*

$$(K_1 \cap K_2)^* = \overline{K_1^*+K_2^*}. \tag{3.3}$$

Замечание. Напомним, что черта над множеством обозначает замыкание множества. Эту черту в соотношении (3.3) в общем случае убрать нельзя, так как можно привести примеры конусов — выпуклых и замкнутых,— для которых сумма не замкнута, в то время как сопряженный конус в левой части неравенства (3.3) всегда замкнут.

Доказательство леммы несколько формально и опирается на предыдущие результаты:

$$(K_1 \cap K_2)^* = \left(K_1^{**} \cap K_2^{**}\right)^* = \left((K_1^*+K_2^*)^*\right)^* =$$
$$= \left(K_1^*+K_2^*\right)^{**} = \overline{K_1^*+K_2^*}.$$

Закончим этот раздел простой, но полезной леммой.

Лемма 3.8. *Если произведение $\langle x,\ x^*\rangle$ ограничено снизу для всех $x \in K$, то $x^* \in K^*$. Если $x \in \operatorname{int} K$, то*

$$\langle x,\ x^*\rangle > 0$$

для всех $x^* \in K^*$, $x^* \neq 0$.

Доказательство первого утверждения было дано при выводе формулы (3.2). Докажем второе. Если $x \in \operatorname{int} K$, то существует такое $\varepsilon>0$, что $x+\varepsilon B \subseteq K$, где B — единичный шар с центром в нуле. Поэтому

$$\langle x+\varepsilon z,\ x^*\rangle \geqslant 0$$

для $x^* \in K^*$ при всех $z \in B$. Отсюда следует, что

$$\langle x, x^* \rangle \geqslant \varepsilon \sup_{z \in B} \langle -z, x^* \rangle \geqslant \varepsilon \left\langle \frac{x^*}{\|x^*\|}, x^* \right\rangle = \varepsilon \|x^*\| > 0,$$

что и требовалось доказать.

2. Отделимость выпуклых конусов. Для дальнейшего необходимо уточнить теоремы, связанные с отделимостью выпуклых конусов.

Теорема 3.1. *Пусть $K_1, \ldots, K_m$ — выпуклые конусы. Для того чтобы их пересечение было пусто, необходимо, чтобы нашлись такие $x_i^* \in K_i^*$, $i = 1, \ldots, m$, не все равные нулю, что*

$$x_1^* + \ldots + x_m^* = 0.$$

Доказательство. Рассмотрим пространство $X^m = \overbrace{X \times X \times \ldots \times X}^{m}$, т. е. пространство $\mathbf{R}^{mn}$, элементы которого имеют вид $(x_1, x_2, \ldots, x_m)$, где каждая компонента x_i принадлежит X. Очевидно, что скалярное произведение векторов $(x_1, \ldots, x_m)$ и $(x_1^*, \ldots, x_m^*)$ в таком пространстве можно записать в виде суммы

$$\langle x_1, x_1^* \rangle + \langle x_2, x_2^* \rangle + \ldots + \langle x_m, x_m^* \rangle.$$

Рассмотрим в X^m два конуса:

$$\widetilde{K} = K_1 \times \ldots \times K_m = \{(x_1, \ldots, x_m) : x_1 \in K_1, \ldots, x_m \in K_m\},$$

$$\widetilde{P} = \{(x, \ldots, x) : x \in X\}.$$

Второй конус состоит из векторов, все n-мерные компоненты которых равны между собой. Так как пересечение конусов $K_1, \ldots, K_m$ пусто, то нетрудно видеть, что $\widetilde{K}$ и $\widetilde{P}$ не пересекаются. Применяя теорему 2.3, получим, что существует такой вектор $(x_1^*, \ldots, x_m^*)$, что

$$\langle x, x_1^* \rangle + \ldots + \langle x, x_m^* \rangle \leqslant \langle x_1, x_1^* \rangle + \ldots + \langle x_m, x_m^* \rangle, \tag{3.4}$$

$$x \in X, \quad x_1 \in K_1, \ldots, x_m \in K_m.$$

Из неравенства (3.4) следует, что $\langle x_i, x_i^* \rangle$ ограничено снизу на конусе K_i, $i = 1, 2, \ldots, m$. По лемме 3.8 отсюда следует, что

$$x_i^* \in K_i^*, \qquad i = 1, \ldots, m. \tag{3.5}$$

Левая часть неравенства (3.4) в свою очередь показывает, что произведение $\langle x, x_1^* + \ldots + x_m^* \rangle$ ограничено сверху для всех $x \in X$. Но это может быть только в том случае, если

$$x_1^* + \ldots + x_m^* = 0, \tag{3.6}$$

так как ненулевая линейная функция может принимать как угодно большие значения. Из соотношений (3.5) и (3.6) следует утверждение теоремы.

Теорема 3.2. *Пусть*

$$K = K_1 \cap K_2 \cap \ldots \cap K_m,$$

$$K_1 \cap \operatorname{int} K_2 \cap \ldots \cap \operatorname{int} K_m \neq \varnothing.$$

Тогда

$$K^* = K_1^* + \ldots + K_m^*.$$

Доказательство. Справедливость включения

$$K^* \supseteq K_1^* + \ldots + K_m^*$$

проверяется непосредственно, исходя из определения сопряженного конуса.

Докажем обратное включение. Пусть $x^* \in K^*$, $x^* \neq 0$. Положим

$$K_0 = \{x\colon \langle x, x^* \rangle < 0\}.$$

Конусы K_0 и K не пересекаются, так как в противном случае из существования элемента $x_1 \in K_0 \cap K \neq \varnothing$ следует

$$\langle x_1, x^* \rangle < 0,$$

так как $x_1 \in K_0$, и

$$\langle x_1, x^* \rangle \geqslant 0,$$

так как $x_1 \in K$, $x^* \in K^*$. Рассмотрим конус, сопряженный к K_0. Пусть из $\langle x, x^* \rangle < 0$ следует, что $\langle x, y^* \rangle \geqslant 0$, т. е. $y^* \in K_0^*$, $y^* \neq 0$. Тогда x^* и y^* линейно зависимы. Значит, при некоторых не равных одновременно нулю α_1 и α_2 выполняется соотношение

$$\alpha_1 x^* - \alpha_2 y^* = 0.$$

Так как $x^* \neq 0$ и $y^* \neq 0$, то $\alpha_2 \neq 0$ и

$$y^* = \lambda x^*, \qquad \lambda = \alpha_1/\alpha_2.$$

Далее, для $x \in K_0$

$$0 \leqslant \langle x, y^* \rangle = \lambda \langle x, x^* \rangle.$$

Так как $\langle x, x^* \rangle < 0$, то $\lambda < 0$. В случае, если $y^* = 0$, полагаем $y^* = 0 \cdot x^*$.

Итак, доказано, что

$$K_0^* = \{y^*\colon\ y^* = \lambda x^*,\ \lambda \leqslant 0\}.$$

Поскольку конусы K_0 и K (т. е. $K_0, K_1, \ldots, K_m$) не пересекаются, то согласно предыдущей теореме существуют такие $y^* \in K_0^*$, $x_i^* \in K_i^*$, $i = 1, \ldots, m$, что

$$y^* + x_1^* + \ldots + x_m^* = 0, \tag{3.7}$$

причем не все слагаемые в этой сумме равны нулю. Перепишем равенство (3.7) в виде

$$-\lambda x^* = x_1^* + \ldots + x_m^*, \tag{3.8}$$

используя то, что $y^* = \lambda x^*$, $\lambda \leqslant 0$. Если $\lambda < 0$, то

$$x^* = \left(-\frac{1}{\lambda}\right) x_1^* + \left(-\frac{1}{\lambda}\right) x_2^* + \ldots$$
$$\ldots + \left(-\frac{1}{\lambda}\right) x_m^* \in K_1^* + \ldots + K_m^*,$$

так как x_i^* принадлежат конусам K_i^*, $i = 1, \ldots, m$.

Покажем, что λ не может быть равным нулю. Допустим противное: $\lambda = 0$. Тогда соотношение (3.8) переходит в равенство

$$x_1^* + \ldots + x_m^* = 0. \tag{3.9}$$

Здесь не все x_i^* равны нулю, значит, не равны нулю по крайней мере два слагаемых, например x_1^* и x_2^*. По предположению теоремы существует точка $x_0 \in K_1 \cap \operatorname{int} K_2 \cap \ldots$ $\ldots \cap \operatorname{int} K_m$ такая, что по лемме 3.8

$$\langle x_0, x_2^* \rangle > 0,$$

а $\langle x_0, x_i^* \rangle \geqslant 0$ для $i \neq 2$.

Умножая теперь обе части соотношения (3.9) на x_0 скалярно, получаем противоречие:

$$0 = \langle x_0, x_1^* + x_2^* + \ldots + x_m^* \rangle = \\ = \langle x_0, x_1^* \rangle + \ldots + \langle x_0, x_m^* \rangle > 0.$$

Теорема доказана.

Теорема 3.3. *Пусть $K_1, \ldots, K_m$ — выпуклые конусы, $K = K_1 \cap \ldots \cap K_m$. Тогда либо*

$$K^* = K_1^* + \ldots + K_m^*, \tag{3.10}$$

либо существуют не все равные нулю векторы $x_i^* \in K_i^*$ такие, что

$$x_1^* + \ldots + x_m^* = 0.$$

Доказательство. Вновь обратимся к соотношению (3.8). Если для всех $x^* \in K^*$ в (3.8) оказывается, что $\lambda < 0$, то x^* можно представить в виде суммы $x_i^* \in K_i^*$. Это доказывает равенство (3.10). Если же для некоторого x^* $\lambda = 0$, то выполнено равенство (3.9). Теорема доказана.

Как показывает теорема 3.3, если пересечение конусов пусто, то выполнено равенство (3.9). Поставим вопрос: при каких наиболее общих предположениях может быть выполнено это равенство и когда нет?

Определение 3.3. Конусы $K_1, \ldots, K_m$ *отделимы*, если существуют такие, не все равные нулю $x_i^* \in K_i^*$, что

$$x_1^* + \ldots + x_m^* = 0.$$

Теорема 3.4. *Для того чтобы конусы K_1 и K_2 были неотделимы, необходимо и достаточно выполнения условий:*

а) $\operatorname{ri} K_1 \cap \operatorname{ri} K_2 \neq \varnothing$;

б) $\operatorname{Lin} K_1 + \operatorname{Lin} K_2 = \mathbf{R}^n$.

Доказательство. Если конусы отделимы, то найдутся $x_1^* \in K_1^*$, $x_2^* \in K_2^*$, $x_1^* \neq 0$ такие, что $x_1^* + x_2^* = 0$. Тогда

$$\begin{gathered} \langle x_1 - x_2, x_1^* \rangle = \langle x_1, x_1^* \rangle + \langle x_2, x_2^* \rangle \geqslant 0, \\ x_1 \in K_1, \quad x_2 \in K_2. \end{gathered} \tag{3.11}$$

Из (3.11) следует, что если конусы K_1 и K_2 отделимы, то нуль не может быть внутренней точкой $K_1 - K_2$. Действительно, если $0 \in \operatorname{int}(K_1 - K_2)$, то $-\varepsilon x_1^* \in K_1 - K_2$ при достаточно малом $\varepsilon > 0$, т. е. $-\varepsilon x_1^* = \bar{x}_1 - \bar{x}_2$, $\bar{x}_1 \in K_1$, $\bar{x}_2 \in K_2$. Тогда

$$\langle \bar{x}_1 - \bar{x}_2, x_1^* \rangle = -\varepsilon \| x_1^* \|^2 < 0,$$

что противоречит неравенству (3.11).

Обратно, если $0 \notin \operatorname{int}(K_1 - K_2)$, то существует точка $x^* \neq 0$ такая, что

$$\langle x_1 - x_2, x^* \rangle \geqslant 0, \quad x_1 - x_2 \in \operatorname{int}(K_1 - K_2).$$

Полагая $x_1^* = x^*$, $x_2^* = -x^*$, получаем, что $x_1^* \in K_1^*$, $x_2^* \in K_2^*$ и $x_1^* + x_2^* = 0$, т. е. конусы K_1 и K_2 отделимы.

Итак, необходимым и достаточным условием неотделимости двух конусов является условие $0 \in \operatorname{int}(K_1 - K_2)$.

Если $\operatorname{ri} K_1 \cap \operatorname{ri} K_2 = \varnothing$, то K_1 и K_2 отделимы. Для этого достаточно отделить $\operatorname{ri} K_1$ и $\operatorname{ri} K_2$:

$$\langle x_1, x^* \rangle \geqslant \langle x_2, x^* \rangle, \quad x_1 \in \operatorname{ri} K_1, \quad x_2 \in \operatorname{ri} K_2, \tag{3.12}$$

и обозначить $x_1^* = x^*$, $x_2^* = -x^*$. Остается заметить, что $\bar{K}_i = \overline{\operatorname{ri} K_i}$ и $K_i^* = (\bar{K}_i)^* = (\operatorname{ri} K_i)^*$.

Допустим теперь, что $\operatorname{ri} K_1 \cap \operatorname{ri} K_2 \neq \varnothing$, но $\operatorname{Lin} K_1 + \operatorname{Lin} K_2 \neq \mathbf{R}^n$. Так как $0 \in \bar{K}$ и $\operatorname{Lin} \bar{K} = \operatorname{Lin} K$, то $K \subseteq \operatorname{Lin} K$ для конуса. Далее, $K_1 - K_2 \subseteq \operatorname{Lin} K_1 - \operatorname{Lin} K_2$. Так как $\operatorname{Lin} K_2$ — подпространство, которое с элементом x содержит и $-x$, то $-\operatorname{Lin} K_2 = \operatorname{Lin} K_2$, поэтому окончательно имеем

$$K_1 - K_2 \subseteq \operatorname{Lin} K_1 + \operatorname{Lin} K_2.$$

Но $\operatorname{Lin} K_1 + \operatorname{Lin} K_2 \neq \mathbf{R}^n$ и поэтому является собственным подпространством $\mathbf{R}^n$. Значит, существует вектор $x^* \neq 0$, ортогональный всем векторам из $\operatorname{Lin} K_1 + \operatorname{Lin} K_2$, в частности,

$$\langle x_1 - x_2, x^* \rangle = 0, \qquad x_1 \in K_1, \quad x_2 \in K_2.$$

Последнее равенство фактически показывает, что K_1 и K_2 отделимы.

Осталось рассмотреть случай, когда $\operatorname{ri} K_1 \cap \operatorname{ri} K_2 \neq \varnothing$ и $\operatorname{Lin} K_1 + \operatorname{Lin} K_2 = \mathbf{R}^n$. Пусть $e_1, e_2, \ldots, e_n$ — орты в $\mathbf{R}^n$. Так как $e_k \in \operatorname{Lin} K_1 + \operatorname{Lin} K_2$, то существуют такие $g_k \in \operatorname{Lin} K_1$ и $f_k \in \operatorname{Lin} K_2$, что $e_k = g_k - f_k$.

Пусть $x_0 \in \operatorname{ri} K_1 \cap \operatorname{ri} K_2$. Обозначим через $g = g_1 + \ldots + g_n$, $f = f_1 + \ldots + f_n$. Очевидно, что $g \in \operatorname{Lin} K_1$, $f \in \operatorname{Lin} K_2$. Выберем теперь $\varepsilon > 0$ настолько малым, чтобы выполнялись включения

$$x_0 + \varepsilon\left(g_k - \frac{1}{n+1} g\right) \in K_1, \quad k = 1, \ldots, n, \quad x_0 - \frac{\varepsilon}{n+1} g \in K_1,$$

$$x_0 + \varepsilon\left(f_k - \frac{1}{n+1} f\right) \in K_2, \quad k = 1, \ldots, n, \quad x_0 - \frac{\varepsilon}{n+1} f \in K_2.$$

Тогда точки

$$y_k = \left(x_0 + \varepsilon\left(g_k - \frac{1}{n+1} g\right)\right) - \left(x_0 + \varepsilon\left(f_k - \frac{1}{n+1}\right)\right) =$$

$$= \varepsilon\left(e_k - \frac{1}{n+1} e\right) \in K_1 - K_2, \quad k = 1, \ldots, n,$$

$$y_0 = \left(x_0 - \frac{\varepsilon}{n+1} g\right) - \left(x_0 - \frac{\varepsilon}{n+1} f\right) = -\frac{\varepsilon}{n+1} e \in K_1 - K_2,$$

где $e = g - f = e_1 + \ldots + e_k$, являются вершинами симплекса в $\mathbf{R}^n$, для которого точка нуль — внутренняя. Этот симплекс целиком лежит в $K_1 - K_2$, ибо его вершины лежат в $K_1 - K_2$. Значит, $0 \in \operatorname{int}(K_1 - K_2)$ и K_1 и K_2 неотделимы.

Обратимся к общему случаю.

Теорема 3.5. *Для того чтобы конусы $K_1, \ldots, K_m$ были неотделимы, необходимо и достаточно выполнения условий:*

а) $\operatorname{ri} K_1 \cap \operatorname{ri} K_2 \cap \ldots \cap \operatorname{ri} K_m \neq \varnothing$;

б) $\operatorname{Lin} K_1 \cap \ldots \cap \operatorname{Lin} K_{j-1} + \operatorname{Lin} K_j = \mathbf{R}^n$, *для всех* $j = 2, \ldots, m$.

Доказательство. Конусы $K_1, \ldots, K_m$ отделимы тогда и только тогда, когда при некотором $j = 2, \ldots, m$ отделимы конусы $K_1 \cap K_2 \cap \ldots \cap K_{j-1}$ и K_j. Действительно, если последние отделимы, то существуют такие $x^* \in (K_1 \cap \ldots \cap K_{j-1})^*$ и $x_j^* \in K_j^*$, что

$$x^* + x_j^* = 0 \tag{3.13}$$

Но согласно теореме 3.3 либо $x^* = x_1^* + \ldots + x_{j-1}^*$, $x_i^* \in K_i^*$, либо

$$x_1^* + \ldots + x_{j-1}^* = 0. \tag{3.14}$$

В первом случае, подставляя выражение для x^* в равенство (3.13) и полагая $x_{j+1}^* = \ldots = x_m^* = 0$, добьемся выполнения условия (3.9). Во втором случае — дополним недостающие в формуле (3.14) слагаемые до (3.9) нулями. Получим, что $K_1, \ldots, K_m$ отделимы.

Обратно, если $K_1, \ldots, K_m$ отделимы, то выберем в соотношении (3.9) наибольший номер j, для которого $x_j^* \neq 0$, $j \leqslant m$. Тогда соотношение (3.9) перейдет в равенство (3.13), если положить $x^* = x_1^* + \ldots + x_{j-1}^*$. Но $x^* \in (K_1 \cap \ldots \cap K_{j-1})^*$ всегда, так что конусы $K_1 \cap \ldots \cap K_{j-1}$ и K_j отделимы.

Рассмотрим конусы K_j и $K_1 \cap \ldots \cap K_{j-1}$, $j = 2, \ldots, m$, и найдем необходимые и достаточные условия их неотделимости индукцией по числу j.

При $j = 2$ применение теоремы 3.4 дает $\operatorname{ri} K_1 \cap \operatorname{ri} K_2 \neq \varnothing$ и $\operatorname{Lin} K_1 + \operatorname{Lin} K_2 = \mathbf{R}^n$. Пусть уже показано, что для неотделимости $K_1 \cap \ldots \cap K_{j-2}$ и K_{j-1} необходимо и достаточно выполнения условий

$$\operatorname{ri} K_1 \cap \ldots \cap \operatorname{ri} K_{j-1} \neq \varnothing, \tag{3.15}$$

$$\operatorname{Lin} K_1 \cap \ldots \cap \operatorname{Lin} K_{j-2} + \operatorname{Lin} K_{j-1} = \mathbf{R}^n. \tag{3.16}$$

Покажем их справедливость для неотделимости $K_1 \cap \ldots \cap K_{j-1}$ и K_j. Применение теоремы 3.4 к этим конусам дает

$$\operatorname{ri}(K_1 \cap \ldots \cap K_{j-1}) \cap \operatorname{ri} K_j \neq \varnothing, \tag{3.17}$$

$$\operatorname{Lin}(K_1 \cap \ldots \cap K_{j-1}) + \operatorname{Lin} K_j = \mathbf{R}^n. \tag{3.18}$$

Но в силу условия (3.15) и теоремы 1.4

$$\operatorname{ri}(K_1 \cap \ldots \cap K_{j-1}) = \operatorname{ri} K_1 \cap \ldots \cap \operatorname{ri} K_{j-1},$$

$$\operatorname{Lin}(K_1 \cap \ldots \cap K_{j-1}) = \operatorname{Lin} K_1 \cap \ldots \cap \operatorname{Lin} K_{j-1}.$$

Соотношения (3.17) и (3.18) переходят соответственно в соотношения

$$\operatorname{ri} K_1 \cap \ldots \cap \operatorname{ri} K_j \neq \varnothing, \tag{3.19}$$

$$\operatorname{Lin} K_1 \cap \ldots \cap \operatorname{Lin} K_{j-1} + \operatorname{Lin} K_j = \mathbf{R}^n. \tag{3.20}$$

Теорема доказана.

3. Выпуклые конусы и частичная упорядоченность. Если в пространстве X задан выпуклый замкнутый конус K, то в нем можно задать частичный порядок, т. е. определить для некоторых его элементов отношение «больше».

Определение 3.4. Будем говорить, что x *больше нуля*, если $x \in K$, и писать $x \overset{K}{\geqslant} 0$. Будем говорить, что x *больше* y, если $x - y \in K$, и писать $x \overset{K}{\geqslant} y$.

Если ясно, о каком конусе K идет речь, то будем иногда вместо $x \overset{K}{\geqslant} 0$ писать просто $x \geqslant 0$. Кроме того, удобно договориться, что соотношение $x \leqslant y$ эквивалентно $y \geqslant x$.

Вообще говоря, не для всяких x и y можно написать соотношения $x \overset{K}{\geqslant} y$ или $y \overset{K}{\geqslant} x$, поэтому упорядоченность лишь частичная. С другой стороны, введенное отношение обладает целым рядом свойств, присущих обычному неравенству. Перечислим некоторые из них.

1. Из $x \leqslant y$ следует, что $\alpha x \leqslant \alpha y$ для $\alpha \geqslant 0$ и $\alpha y \leqslant \alpha x$ для $\alpha < 0$.
2. Если $x_1 \leqslant y_1$ и $x_2 \leqslant y_2$, то $x_1 + x_2 \leqslant y_1 + y_2$.
3. Если $x \leqslant y$ и $y \leqslant z$, то $x \leqslant z$.
4. Если $x_k \to x_0$, $y_k \to y_0$, $x_k \leqslant y_k$, то $x_0 \leqslant y_0$, т. е. в неравенстве можно переходить к пределу.

Наконец, отметим еще одно полезное понятие. Будем писать $x \overset{K}{>} 0$ или просто $x > 0$, если $x \in \operatorname{int} K$. Аналогично, $x \underset{K}{>} y$ или $x > y$, если $x - y \in \operatorname{int} K$.

Пример. Пусть $X = \mathbf{R}^n$ и

$$K = \{x : \ x^i \geqslant 0, \quad i = 1, \ldots, n\}.$$

В этом случае соотношение $x \overset{K}{\geqslant} y$ эквивалентно тому, что $x^i \geqslant y^i$ для всех $i = 1, \ldots, n$, а соотношение $x \overset{K}{>} y$ эквивалентно $x^i > y^i$, $i = 1, \ldots, n$.

4. Важный пример выпуклого конуса. Пусть M — непустое множество. Обозначим

$$\operatorname{con} M = \{x : \ x = \lambda x_1, \quad x_1 \in M, \quad \lambda > 0\}.$$

Теорема 3.6. *Если M — выпуклое множество, то $\operatorname{con} M$ — выпуклый конус, причем $x \in \operatorname{int}(\operatorname{con} M)$, если $x = \lambda x_1$, $x_1 \in \operatorname{int} M$, $\lambda > 0$.*

Доказательство. Возьмем $x_1 \in \operatorname{con} M$, $x_2 \in \operatorname{con} M$. Тогда

$$x_1 = \lambda_1 \bar{x}_1, \quad x_2 = \lambda_2 \bar{x}_2, \quad \bar{x}_1, \ \bar{x}_2 \in M, \quad \lambda_1, \lambda_2 > 0.$$

Пусть $\gamma_1 > 0$, $\gamma_2 > 0$ произвольны. Имеем

$$\gamma_1 x_1 + \gamma_2 x_2 = \gamma_1\lambda_1\bar{x}_1 + \gamma_2\lambda_2\bar{x}_2 =$$
$$= (\gamma_1\lambda_1 + \gamma_2\lambda_2)\left[\frac{\gamma_1\lambda_1}{\gamma_1\lambda_1+\gamma_2\lambda_2}\bar{x}_1 + \frac{\gamma_2\lambda_2}{\gamma_1\lambda_1+\gamma_2\lambda_2}\bar{x}_2\right] \in \operatorname{con} M,$$

так как в силу выпуклости M выражение в квадратных скобках принадлежит M.

Далее, если $x = \lambda x_1$, $x_1 \in \operatorname{int} M$, $\lambda > 0$, то $x_1 + \varepsilon B \subseteq M$ при достаточно малом $\varepsilon > 0$. Поэтому

$$x + \lambda\varepsilon B = \lambda(x_1 + \varepsilon B) \subseteq \operatorname{con} M,$$

т. е. $x \in \operatorname{int}\operatorname{con} M$.

Теорема 3.7. *Если $M_1, \ldots, M_k$ — выпуклые множества и $0 \in M_i$, то* $\bigcap\limits_{i=1}^{k} \operatorname{con} M_i = \operatorname{con}\left(\bigcap\limits_{i=1}^{k} M_i\right)$.

Доказательство. Если $x = \lambda x_1$, где $x_1 \in \bigcap\limits_{i=1}^{k} M_i$, то ясно, что $x \in \operatorname{con} M_i$ для всех $i = 1, \ldots, k$. Обратно, пусть $x \in \operatorname{con} M_i$, $i = 1, \ldots, k$, т. е. $x = \lambda_i x_i$, $\lambda_i > 0$, $x_i \in M_i$, $i = 1, \ldots, k$. Тогда $\lambda_i^{-1}x \in M_i$ и

$$\lambda\left(\lambda_i^{-1}x\right) = (1 - \lambda)\cdot 0 + \lambda\left(\lambda_i^{-1}x\right) \in M_i$$

при $0 \leqslant \lambda < 1$. Выбирая μ из условия $0 < \mu \leqslant \min \lambda_i^{-1}$, получим

$$\mu x = (\mu\lambda_i)\left(\lambda_i^{-1}x\right) \in M_i, \qquad i = 1, \ldots, k,$$

т. е.

$$\mu x \in \bigcap_{i=1}^{k} M_i, \qquad x = \left(\frac{1}{\mu}\right)(\mu x) \in \operatorname{con}\left(\bigcap_{i=1}^{k} M_i\right),$$

что и требовалось доказать.

Теорема 3.8. *$x^* \in (\operatorname{con} M)^*$ тогда и только тогда, когда $\langle x, x^*\rangle \geqslant 0$ для всех $x \in M$.*

Доказательство. В самом деле, $x^* \in (\operatorname{con} M)^*$, если и только если

$$\langle \lambda x, x^*\rangle \geqslant 0$$

при всех $\lambda \geqslant 0$ и любых $x \in M$. Так как $\lambda \geqslant 0$, то из этого неравенства следует утверждение.

Теорема 3.9. *Пусть* $0 \in M$, *тогда* $x \in \operatorname{con} M$ *в том, и только том случае, когда* $\lambda x \in M$ *при достаточно малых* $\lambda > 0$.

Доказательство. Если $x \in \operatorname{con} M$, то $x = \lambda_1 x_1$, где $\lambda_1 > 0$, $x_1 \in M$. Поэтому

$$\lambda x = \lambda\lambda_1 x_1 = (1 - \lambda\lambda_1) \cdot 0 + \lambda\lambda_1 x_1 \in M,$$

как только $\lambda\lambda_1 \leqslant 1$. Обратно, если $\lambda x \in M$, $\lambda > 0$, то $x = \frac{1}{\lambda} x_1$, $x_1 \in M$ и $x \in \operatorname{con} M$.

§ 4. Крайние точки и многогранные множества

В § 1 показано (см. лемму 1.3), что в выпуклом множестве любая выпуклая комбинация его точек снова принадлежит этому множеству. Возникает вопрос: какие точки выпуклого множества можно представить такими комбинациями других точек из этого множества и какие нельзя? Оказывается, что именно последние являются характерными и полностью задают данное выпуклое множество.

1. Крайние точки.

Определение 4.1. Точка x выпуклого множества M называется *крайней* точкой, если не существует таких точек $x_1, x_2 \in M$, $x_1 \neq x_2$, что

$$\frac{1}{2} x_1 + \frac{1}{2} x_2 = x.$$

Геометрически это означает, что крайняя точка выпуклого множества не может быть серединой отрезка, концы которого лежат в этом множестве. Вообще говоря, не во всяком выпуклом множестве существуют крайние точки. Например, точки, принадлежащие некоторой гиперплоскости, не могут быть крайними для выпуклого множества, совпадающего с этой гиперплоскостью, ибо через каждую точку этого множества проходит целая прямая. С другой стороны, компактные выпуклые множества полностью характеризуются своими крайними точками.

Теорема 4.1. *Компактное выпуклое множество в* $\mathbf{R}^n$ *совпадает с совокупностью выпуклых комбинаций своих крайних точек.*

Доказательство. Удобнее всего провести доказательство индукцией по числу измерений n. Для $n = 1$ единственным видом компактных выпуклых множеств являются замкнутые отрезки; поэтому теорема верна, ибо для всякого отрезка его концы являются крайними точками, а любая другая точка отрезка есть выпуклая комбинация его концов.

Предположим теперь, что теорема верна для $n \leqslant k$, и докажем ее для $n = k+1$. Можно считать, что рассматриваемое выпуклое множество имеет в $\mathbf{R}^{k+1}$ внутренние точки. В противном случае его можно было бы погрузить в пространство меньшего числа измерений, где теорема верна по предположению индукции. Итак, $M \subseteq \mathbf{R}^{k+1}$, $\operatorname{int} M \neq \varnothing$, $x_0 \in \operatorname{int} M$. Возьмем произвольный вектор $p \neq 0$ и рассмотрим прямую $x_0 + \lambda p$, $\lambda \in \mathbf{R}^1$. Так как x_0 — внутренняя точка M, а M выпукло и компактно, то существуют $\lambda_1 > 0$ и $\lambda_2 < 0$ такие, что $x_0 + \lambda p \in M$ для всех $\lambda \leqslant \lambda_1$ и $x_0 + \lambda p \in M$ для всех $\lambda \geqslant \lambda_2$. Ясно, что

$$x_0 = \frac{(-\lambda_2)}{\lambda_1 - \lambda_2}(x_0 + \lambda_1 p) + \frac{\lambda_1}{\lambda_1 - \lambda_2}(x_0 + \lambda_2 p),$$

т. е. x_0 есть точка отрезка с концами

$$x_1 = x_0 + \lambda_1 p, \qquad x_2 = x_0 + \lambda_2 p.$$

Рассмотрим точку x_1. По определению точки $x_0 + \lambda p$ при $\lambda > \lambda_1$ не принадлежат M. Поэтому множества

$$L = \{x = x_0 + \lambda p \colon \lambda > \lambda_1\}$$

и M можно разделить, т. е. существует такой вектор x^*, что

$$\langle x, x^* \rangle \leqslant \langle x_0 + \lambda p, x^* \rangle, \quad x \in M, \ \lambda > \lambda_1. \tag{4.1}$$

Отсюда следует, что $\langle p, x^* \rangle \geqslant 0$. Действительно, предположив, что $\langle p, x^* \rangle < 0$, из (4.1) получим

$$\langle x, x^* \rangle \leqslant \langle x_0, x^* \rangle + \lambda \langle p, x^* \rangle.$$

Правая часть этого неравенства стремится к $-\infty$, если $\lambda \to +\infty$, что приводит к противоречию. Поэтому из неравенства (4.1) получаем, что

$$\langle x, x^* \rangle \leqslant \langle x_0 + \lambda_1 p, x^* \rangle = \langle x_1, x^* \rangle, \ x \in M. \tag{4.2}$$

Положим

$$P_1 = \{x \in M \colon \langle x, x^* \rangle = \langle x_1, x^* \rangle\}. \tag{4.3}$$

Это множество не пусто, ибо содержит x_1. Кроме того, оно представляет собой пересечение компакта M с замкнутой гиперплоскостью

$$\langle x, x^*\rangle = \langle x_1, x^*\rangle$$

и поэтому является ограниченным замкнутым выпуклым множеством размерности не более чем k. Согласно предположению индукции точка x_1 есть выпуклая комбинация крайних точек P_1.

Покажем, что если y — крайняя точка P_1, то y — крайняя точка M. Предположим противное: y не является крайней точкой множества M, т. е.

$$y = \frac{1}{2} y_1 + \frac{1}{2} y_2, \quad y_1 \in M, \quad y_2 \in M, \quad y_1 \neq y_2.$$

В силу условий (4.2) и (4.3) выполняются соотношения

$$\langle y_1, x^*\rangle \leqslant \langle x_1, x^*\rangle,$$

$$\langle y_2, x^*\rangle \leqslant \langle x_1, x^*\rangle,$$

$$\langle y, x^*\rangle = \langle x_1, x^*\rangle,$$

$$\langle x_1, x^*\rangle = \langle y, x^*\rangle = \frac{1}{2}\langle y_1, x^*\rangle + \frac{1}{2}\langle y_2, x^*\rangle \leqslant \langle x_1, x^*\rangle,$$

поэтому

$$\langle y_1, x^*\rangle = \langle x_1, x^*\rangle, \qquad \langle y_2, x^*\rangle = \langle x_1, x^*\rangle,$$

т. е. $y_1, y_2 \in P_1$. Таким образом, y не есть крайняя точка P_1, что противоречит допущению.

Итак, если y — крайняя точка P_1, то y — крайняя точка M. Аналогичными рассуждениями доказываются свойства крайних точек множества

$$P_2 = \{x \in M \colon \langle x, x^*\rangle = \langle x_2, x^*\rangle\}.$$

Таким образом, согласно предположению теоремы и проведенным рассуждениям

$$x_1 = \sum_{i=1}^{q_1} \alpha_i y_i, \quad y_i \in P_1, \quad \alpha_i \geqslant 0, \quad \sum_{i=1}^{q_1} \alpha_i = 1,$$

$$x_2 = \sum_{j=1}^{q_2} \beta_j z_j, \quad z_j \in P_2, \quad \beta_j \geqslant 0, \quad \sum_{j=1}^{q_2} \beta_j = 1,$$

где y_i, z_j — крайние точки множеств P_1, P_2 и, следовательно, M. Но, как было установлено ранее,

$$x_0 = \gamma_1 x_1 + \gamma_2 x_2 = \sum_{i=1}^{q_1} \gamma_i \alpha_i y_i + \sum_{j=1}^{q_2} \gamma_2 \beta_j z_j,$$

где

$$\gamma_1 = \frac{-\lambda_2}{\lambda_1 - \lambda_2}, \quad \gamma_2 = \frac{\lambda_1}{\lambda_1 - \lambda_2},$$

и, как легко проверить,

$$\sum_{i=1}^{q_1} \gamma_1 \alpha_i + \sum_{j=1}^{q_2} \gamma_2 \beta_j = 1.$$

Это означает, что точка x_0 есть выпуклая комбинация крайних точек из M. Если же $x_0 \in M$, $x_0 \notin \operatorname{int} M$, то либо x_0 — крайняя точка, либо x_0 и $\operatorname{int} M$ можно отделить и повторить рассуждения.

Теорема 4.2. *Для того чтобы точка x_0 была крайней точкой множества $M \subseteq \mathbf{R}^n$, заданного системой линейных неравенств*

$$\langle x, x_i^* \rangle \leqslant \alpha_i, \qquad i \in I = \{1, 2, \ldots, m\},$$

необходимо и достаточно, чтобы множество

$$I(x_0) = \{i\colon \langle x_0, x_i^* \rangle = \alpha_i, \quad i \in I\} \tag{4.4}$$

содержало подмножество I_0, мощности n и чтобы точки x_i^*, $i \in I_0$, были линейно независимы.

Доказательство. Необходимость. Пусть множество $\{x_i^*, \ i \in I(x_0)\}$ содержит меньше чем n линейно независимых элементов. Тогда на основании известных теорем линейной алгебры система линейных относительно x уравнений

$$\langle x, x_i^* \rangle = 0, \qquad i \in I(x_0),$$

имеет ненулевое решение $\bar{x}$. Из определения (4.4) множества $I(x_0)$ вытекает, что

$$\langle x_0 \pm \varepsilon \bar{x}, x_i^* \rangle = \alpha_i, \qquad i \in I(x_0),$$

$$\langle x_0 \pm \varepsilon \bar{x}, x_i^* \rangle < \alpha_i, \qquad i \notin I \setminus I(x_0),$$

при достаточно малом $\varepsilon > 0$, так что

$$x_0 + \varepsilon \bar{x} \in M, \ x_0 - \varepsilon \bar{x} \in M,$$

$$x_0 = \frac{1}{2}(x_0 + \varepsilon \bar{x}) + \frac{1}{2}(x_0 - \varepsilon \bar{x}) \in M,$$

т. е. x_0 не есть крайняя точка M.

Достаточность. Пусть точка $x_0 \in M$ такова, что мощность I_0 равна n и для $i \in I_0$ векторы x_i^* линейно независимы. Тогда система неравенств, описывающих множество M, может быть записана в следующем виде:

$$\langle x_0, x_i^* \rangle = \alpha_i, \qquad i \in I_0, \tag{4.5}$$

$$\langle x_0, x_i^* \rangle \leqslant \alpha_i, \qquad i \in I \setminus I_0. \tag{4.6}$$

Допустим, что

$$x_0 = \frac{1}{2} x_1 + \frac{1}{2} x_2, \qquad x_1 \in M, \quad x_2 \in M, \quad x_1 \neq x_2. \tag{4.7}$$

Так как x_1, $x_2 \in M$, то по определению M справедливы неравенства

$$\langle x_k, x_i^* \rangle \leqslant \alpha_i, \qquad i \in I, \quad k = 1, 2. \tag{4.8}$$

В силу условий (4.5), (4.8) соотношение (4.7) выполняется, только если

$$\langle x_k, x_i^* \rangle = \alpha_i, \qquad i \in I_0, \quad k = 1, 2. \tag{4.9}$$

Итак, две различные точки удовлетворяют системе уравнений (4.9), в которой уравнения линейно независимы и число их, совпадающее с мощностью множества I_0, равно n — размерности пространства. Это невозможно в силу известных теорем о разрешимости систем линейных уравнений. Теорема доказана.

Из двух вышеприведенных теорем следует

Теорема 4.3. *Пусть M есть ограниченное множество, заданное конечной системой линейных неравенств*

$$\langle x, x_i^* \rangle \leqslant \alpha_i, \qquad i \in I. \tag{4.10}$$

Тогда M есть выпуклая оболочка своих крайних точек, число которых конечно.

Доказательство. В силу условия (4.10) множество M компактно. На основании теорем 4.1, 4.2 множество M есть выпуклая оболочка множества своих крайних точек, причем каждая крайняя точка множества M есть единственное решение невырожденной системы n уравнений с n неизвестными (4.5) и поэтому однозначно определяется множеством I_0. Так как из конечной системы индексов I подмножество из n индексов можно выбрать лишь конечным числом способов, то число крайних точек конечно.

2. Выпуклые многогранники. Предыдущая теорема мотивирует следующее

Определение 4.2. *Выпуклым многогранником* называется выпуклая оболочка конечного числа точек.

Из теоремы 4.3 теперь сразу вытекает

Теорема 4.4. *Ограниченное множество, заданное конечной системой линейных неравенств, есть многогранник.*

Итак, если M — многогранник, то любая его точка x есть выпуклая комбинация конечного числа точек $x_1, \ldots, x_m$:

$$x = \sum_{i=1}^{m} \lambda_i x_i, \qquad \lambda_i \geqslant 0, \qquad \sum_{i=1}^{m} \lambda_i = 1.$$

Легко вычислить опорную функцию многогранника:

$$W_M(x^*) = \sup_{x \in M} \langle x, x^* \rangle = \max_{1 \leqslant i \leqslant m} \langle x_i, x^* \rangle.$$

Пусть

$$I(x^*) = \{i\colon \langle x_i, x^* \rangle = W_M(x^*),\ i = 1, \ldots, m\}$$

есть множество индексов i, для которых скалярное произведение $\langle x_i, x^* \rangle$ достигает максимума.

Определение 4.3. Гиперплоскость

$$\langle x, x^* \rangle = W_M(x^*)$$

называется *опорой*. Если $j \in I(x^*)$ и среди векторов $x_i - x_j$, $i \in I(x^*)$, имеется $n-1$ линейно независимых, то опора называется *крайней*.

Теорема 4.5. *Пусть многогранник M содержит внутренние точки. Тогда любую точку $x_0 \notin M$ можно отделить при помощи крайней опоры, т. е. найдется вектор x^*, определяющий крайнюю опору и такой, что*

$$\langle x, x^* \rangle < \langle x_0, x^* \rangle, \qquad x \in M.$$

Доказательство. Без ограничения общности можно считать, что $0 \in \operatorname{int} M$. По теореме 2.1 существует такой вектор x^*, что

$$\langle x, x^* \rangle \leqslant \langle x_0, x^* \rangle - \varepsilon \tag{4.11}$$

для всех $x \in M$. Ясно, что

$$W_M(x^*) < \langle x_0, x^* \rangle. \tag{4.12}$$

Предположим, что x^* не определяет крайнюю опору. Тогда среди векторов $x_i - x_j$ ($i \in I(x^*)$, j — фиксированный элемент $I(x^*)$) имеются не более $n-2$ линейно независимых. Выберем теперь вектор $\bar{x}^*$ так, чтобы выполнялись соотношения

$$\begin{aligned} &\langle x_i - x_j, \bar{x}^* \rangle = 0, \quad i \in I(x^*), \\ &\langle x_0 - x_j, \bar{x}^* \rangle = 0, \\ &\langle x_0, \bar{x}^* \rangle \leqslant 0. \end{aligned} \tag{4.13}$$

Это всегда можно сделать, так как среди векторов $x_0 - x_j$, $x_i - x_j$, $i \in I(x^*)$, имеются не более $n-1$ линейно независимых.

Покажем, что найдется такой индекс $q \in I$, что

$$\langle x_q, \bar{x}^* \rangle > 0.$$

В самом деле, так как по предположению $0 \in \operatorname{int} M$, то

$$0 < W_M(\bar{x}^*) = \max_{1 \leqslant i \leqslant m} \langle x_i, \bar{x}^* \rangle,$$

откуда следует существование q. В силу условий (4.13) $\langle x_i, \bar{x}^* \rangle \leqslant 0$ для $i \in I(x^*)$, поэтому $q \notin I(x^*)$. Для $i \in I(x^*)$ из этих же соотношений получаем

$$\langle x_i, x^* + \alpha \bar{x}^* \rangle = \langle x_i, x^* \rangle + \alpha \langle x_i, \bar{x}^* \rangle = \\ = W_M(x^*) + \alpha \langle x_0, \bar{x}^* \rangle. \tag{4.14}$$

Определим α_i для $i \notin I(x^*)$ равенством

$$\langle x_i, x^* + \alpha \bar{x}^* \rangle = W_M(x^*) + \alpha \langle x_0, \bar{x}^* \rangle,$$

т. е.

$$\alpha_i = \frac{W_M(x^*) - \langle x_i, x^* \rangle}{\langle x_i, \bar{x}^* \rangle - \langle x_0, \bar{x}^* \rangle}, \qquad i \notin I(x^*). \tag{4.15}$$

Числитель в выражении для α_i всегда положителен, поэтому $\alpha_i \neq 0$. Так как $\langle x_0, \bar{x}^* \rangle \leqslant 0$, а $\langle x_q, \bar{x}^* \rangle > 0$, то $\alpha_q > 0$. Выберем среди положительных α_i, $i \notin I(x^*)$, наименьшее: α_p. Так как

$$\langle x_i, x^* \rangle < W_M(x^*), \quad i \notin I(x^*),$$

то

$$\langle x_i,\ x^* + \alpha_p \bar{x}^*\rangle \leqslant W_M(x^*) + \alpha_p \langle x_0,\ \bar{x}^*\rangle, \qquad i \notin I(x^*),$$

причем

$$\langle x_p,\ x^* + \alpha_p \bar{x}^*\rangle = W_M(x^*) + \alpha_p\langle x_0,\ \bar{x}^*\rangle.$$

Сравнивая последнее равенство с соотношением (4.14), получаем

$$W_M(x^* + \alpha_p \bar{x}^*) = W_M(x^*) + \alpha_p\langle x_0,\ \bar{x}^*\rangle,$$

$$I(x^* + \alpha_p \bar{x}^*) \supseteq I(x^*) \cup \{p\}.$$

Таким образом, множество $I(x^* + \alpha_p \bar{x}^*)$ содержит больше индексов, чем $I(x^*)$. Более того, векторы $x_i - x_j$, $i \in I(x^*)$, и $x_p - x_j$ линейно независимы, так как

$$\langle x_i - x_j,\ \bar{x}^*\rangle = 0, \qquad i \in J(x^*),$$

а

$$\langle x_p - x_j,\ \bar{x}^*\rangle = \langle x_p,\ \bar{x}^*\rangle - \langle x_0,\ \bar{x}^*\rangle > 0,$$

что вытекает из формулы (4.15) и положительности α_p.

Итак, показано, что среди векторов $x_i - x_j$, $i \in I(x^* + \alpha_p \bar{x}^*)$, число линейно независимых по крайней мере на один больше, чем среди векторов $x_i - x_j$, $i \in I(x^*)$. Кроме того, для всех $x \in M$

$$\langle x,\ x^* + \alpha_p \bar{x}^*\rangle \leqslant W_M(x^* + \alpha_p \bar{x}^*) =$$
$$= W_M(x^*) + \alpha_p\langle x_0,\ \bar{x}^*\rangle < \langle x_0,\ x^*\rangle + \alpha_p\langle x_0,\ \bar{x}^*\rangle =$$
$$= \langle x_0,\ x^* + \alpha_p \bar{x}^*\rangle,$$

т. е. вектор $x^* + \alpha_p \bar{x}^*$ отделяет точку x_0 от M. Если среди векторов $x_i - x_j$, $i \in I(x^* + \alpha_p \bar{x}^*)$, число линейно независимых меньше, чем $n - 1$, то описанную выше процедуру можно повторить.

Очевидно, что крайняя опора будет построена за конечное число шагов.

Нетрудно видеть, что число крайних опор конечно. Действительно, если вектор x^* определяет крайнюю опору, то он ортогонален $n - 1$ векторам $x_i - x_j$, $i \in I(x^*)$, и, следовательно, определяется однозначно с точностью до скалярного множителя. Так как $I(x^*)$ есть подмножество конечного множества I, то число возможных подмножеств $I(x^*)$ конечно, а значит, конечно и число крайних опор.

Теорема 4.6. *Выпуклый многогранник может быть задан конечной системой линейных неравенств.*

Доказательство. Пусть $\operatorname{int} M \neq \varnothing$ и векторы x_k^*, $k = 1, \ldots, l$, определяют все крайние опоры к M. Тогда система линейных неравенств

$$\langle x, x_k^* \rangle \leqslant W_M(x_k^*), \qquad k = 1, \ldots, l,$$

определяет многогранник M. Действительно, любая точка $x \in M$ удовлетворяет этой системе, а любая точка $x \notin M$ должна нарушать хотя бы одно из этих неравенств согласно теореме 4.5.

Если $\operatorname{int} M = \varnothing$, то можно рассмотреть множество $M - x_0$, $x_0 \in M$, относительно подпространства $\operatorname{Lin} M$, в котором $M - x_0$ лежит и имеет внутренние точки. В подпространстве $\operatorname{Lin} M$ многогранник M может быть задан конечной системой линейных неравенств. Само подпространство $\operatorname{Lin} M$ можно задать конечным числом линейных уравнений. Совмещая все эти соотношения, получаем утверждение теоремы.

Из теорем 4.4 и 4.6 непосредственно следует, что определения многогранника как выпуклой оболочки конечного числа вершин и как ограниченного множества, точки которого описываются конечной системой линейных неравенств, эквивалентны. Поэтому в зависимости от обстоятельств можно пользоваться либо одним, либо другим определением.

3. Многогранные конусы. Многогранники представляют собой ограниченные множества, которые можно задать конечной системой линейных неравенств. Рассмотрим важный класс неограниченных множеств, которые можно задать системой однородных линейных неравенств.

Определение 4.4. Конус K называется *многогранным*, если существует конечный набор векторов $x_1, \ldots$ $\ldots, x_m$ таких, что для векторов $x \in K$ и только для них выполняется равенство

$$x = \sum_{i=1}^{m} \lambda_i x_i, \qquad \lambda_i \geqslant 0, \quad i = 1, \ldots, m. \tag{4.16}$$

Теорема 4.7. *Многогранный конус может быть задан конечной системой линейных однородных неравенств*

$$\langle x, x_k^* \rangle \geqslant 0, \qquad k = 1, \ldots, l. \tag{4.17}$$

Доказательство. Рассмотрим множество M точек вида (4.16), для которых дополнительно выполнено неравенство

$$\sum_{i=1}^{m} \lambda_i \leqslant 1.$$

Нетрудно видеть, что M есть многогранник, натянутый на точки $0, x_1, \ldots, x_m$, поэтому он может быть задан конечной системой линейных неравенств

$$\langle x, x_k^* \rangle \geqslant \alpha_k, \qquad k = 1, \ldots, l_1. \tag{4.18}$$

Подставляя $0 \in M$ в систему (4.18), получим, что $\alpha_k \leqslant 0$, $k = 1, \ldots, l_1$. Пусть $\alpha_k = 0$ для $k = 1, \ldots, l \leqslant l_1$, т. е. при $1 \leqslant k \leqslant l$ неравенства (4.18) переходят в (4.17).

Покажем, что точки x, удовлетворяющие условиям вида (4.17), и только они принадлежат конусу K. В самом деле, если x определяется формулой (4.16), то при достаточно малом $\alpha > 0$

$$\alpha x = \sum_{i=1}^{m} \alpha \lambda_i x_i, \qquad \sum_{i=1}^{m} \alpha \lambda_i \leqslant 1,$$

т. е. $\alpha x \in M$, αx удовлетворяет неравенствам (4.18), и, в частности, (4.17). Но тогда точка x удовлетворяет однородным неравенствам (4.17), так как $\alpha > 0$.

Обратно, если x удовлетворяет неравенствам (4.17), то при достаточно малом $\alpha > 0$ точка αx удовлетворяет условиям (4.18), т. е. принадлежит многограннику M. Это означает, что

$$\alpha x = \sum_{i=1}^{m} \lambda_i x_i, \quad \sum_{i=1}^{m} \lambda_i \leqslant 1, \qquad \lambda_i \geqslant 0;$$

поэтому

$$x = \sum_{i=1}^{m} \frac{\lambda_i}{\alpha} x_i \in K.$$

Теорема 4.8. *Многогранный конус замкнут.*

Доказательство. На основании предыдущей теоремы многогранный конус задается системой нестрогих линейных неравенств. Отсюда сразу следует, что он замкнут.

Теорема 4.9. *Если конус K задан системой линейных неравенств (4.17), то сопряженный ему конус K^* является многогранным и состоит из элементов*

$$x^* = \sum_{k=1}^{l} \gamma_k x_k^*, \qquad \gamma_k \geqslant 0.$$

Доказательство. Рассмотрим многогранный конус

$$\widetilde{K} = \left\{x^*:\ x^* = \sum_{k=1}^{l} \gamma_k x_k^*, \quad \gamma_k \geqslant 0\right\}.$$

Он замкнут согласно предыдущей теореме. По определению $x \in (\widetilde{K})^*$, если

$$\left\langle x, \sum_{k=1}^{l} \gamma_k x_k^* \right\rangle \geqslant 0, \qquad \gamma_k \geqslant 0,\ k = 1, \ldots\ l.$$

Но последнее возможно лишь в том случае, когда

$$\langle x, x_k^* \rangle \geqslant 0, \qquad k = 1, \ldots, l,$$

т. е. $x \in K$. Итак, $K = (\widetilde{K})^*$. Так как $\widetilde{K}$ замкнут, то

$$K^* = (\widetilde{K})^{**} = \widetilde{K},$$

что и требовалось доказать.

Теорема 4.10. *Конус, заданный системой линейных неравенств, является многогранным.*

Доказательство. Пусть конус K задается системой линейных неравенств (4.17). Так как сопряженный ему конус K^* является многогранным, то существуют такие точки x_i, $i = 1, \ldots, m$, что точки $x^* \in K^*$ и только они удовлетворяют неравенствам

$$\langle x_i, x^* \rangle \geqslant 0, \quad i = 1, \ldots, m.$$

Так как конус K замкнут, то $K = (K^*)^*$ и на основании теоремы 4.9

$$x = \sum_{i=1}^{m} \lambda_i x_i, \qquad \lambda_i \geqslant 0,$$

т. е. K — многогранный конус.

На основании теорем 4.7 и 4.10 можно заключить, что многогранный конус можно определить двумя способами. либо определением 4.4, либо при помощи конечной однородной системы линейных неравенств. Из существования двух способов задания многогранных конусов можно извлечь важные следствия.

Теорема 4.11. *Сумма многогранных конусов есть многогранный конус.*

Доказательство. Пусть K_1 и K_2 — многогранные конусы. Тогда их элементы можно представить соответственно в виде

$$x_1 = \sum_{i=1}^{m} \lambda_i x_{i1}, \quad \lambda_i \geqslant 0, \qquad x_2 = \sum_{j=1}^{l} \gamma_j x_{j2}, \quad \gamma_j \geqslant 0.$$

Если теперь $x \in K_1 + K_2$, то

$$x = x_1 + x_2 = \sum_{i=1}^{m} \lambda_i x_{i1} + \sum_{j=1}^{l} \gamma_j x_{j2}, \quad \lambda_i \geqslant 0, \quad \gamma_j \geqslant 0,$$

т. е. $K_1 + K_2$ есть снова многогранный конус.

Теорема 4.12. *Пересечение многогранных конусов есть снова многогранный конус.*

Доказательство. По теореме 4.7 многогранные конусы K_1 и K_2 могут быть заданы конечными линейными системами неравенств вида (4.17). Ясно, что пересечение $K_1 \cap K_2$ удовлетворяет линейной системе неравенств, образованной путем объединения систем, задающих K_1 и K_2. По теореме 4.10 $K_1 \cap K_2$ — многогранный конус. Прямым следствием теорем 4.7 и 4.9 является следующее утверждение.

Теорема 4.13. *Конус, сопряженный многогранному, является многогранным.*

Теорема 4.14. *Для многогранных конусов* $K_1, \ldots, K_m$ *выполняется соотношение*

$$(K_1 \cap \ldots \cap K_m)^* = K_1^* + \ldots + K_m^*.$$

Доказательство. Согласно лемме 3.7 имеет место формула

$$(K_1 \cap \ldots \cap K_m)^* = \overline{K_1^* + \ldots + K_m^*}.$$

Но из предыдущего следует, что K_i^*, $i = 1, \ldots, m$, —

многогранные конусы (теорема 4.13), $K_1^* + \ldots + K_m^*$ — многогранный конус (теорема 4.11), а многогранный конус замкнут (теорема 4.8). Поэтому черту замыкания над суммой в последней формуле можно убрать.

4. Многогранные множества. В предыдущих разделах были изучены два типа множеств — многогранники и многогранные конусы. Оба эти типа множеств являются частным случаем более общего многогранного множества.

Определение 4.5. Множество точек, удовлетворяющих системе линейных неравенств

$$\langle x, x_i^* \rangle \geqslant \alpha_i, \qquad i = 1, \ldots, l, \tag{4.19}$$

называется *многогранным.*

Теорема 4.15. *Многогранное множество есть сумма многогранника и многогранного конуса, и, наоборот, сумма многогранника и многогранного конуса есть многогранное множество.*

Доказательство. Введем дополнительную координату x^0 и запишем систему неравенств

$$\langle x, x_i^* \rangle - x^0\alpha_i \geqslant 0, \qquad i = 1, \ldots, l, \quad x^0 \geqslant 0. \tag{4.20}$$

Это однородная система линейных неравенств в $\mathbf{R}^{n+1}$; согласно теореме 4.10 она задает многогранный конус, элементы которого можно представить в виде

$$\begin{pmatrix} x^0 \\ x \end{pmatrix} = \sum_{j=1}^{m} \lambda_j \begin{pmatrix} x_j^0 \\ x_j \end{pmatrix}, \qquad \lambda_j \geqslant 0, \tag{4.21}$$

где через $\begin{pmatrix} x^0 \\ x \end{pmatrix}$ обозначен $(n+1)$-мерный вектор из $\mathbf{R}^{n+1}$ с компонентами $x^0, x^1, \ldots, x^n$. При любых $\lambda_j \geqslant 0$ выражение в правой части формулы (4.21) должно удовлетворять неравенствам (4.20). В частности, если $\lambda_j = 1$, а $\lambda_i = 0$ при $i \neq j$, то из неравенств вытекает, что $x_j^0 \geqslant 0$. Пусть

$$I^0 = \{j \colon x_j^0 = 0, \quad j = 1, \ldots, m\},$$

$$I^+ = \{j \colon x_j^0 > 0, \quad j = 1, \ldots, m\}.$$

Обозначим $y_j = x_j/x_j^0$, $\gamma_j = \lambda_j x_j^0$, $j \in I^+$, и перепишем формулы (4.21) в следующем виде:

$$x^0 = \sum_{j\in I^+} \lambda_j x_j^0 = \sum_{j\in I^+} \gamma_j, \qquad \gamma_j \geqslant 0, \tag{4.22}$$

$$x = \sum_{j\in I^0} \lambda_j x_j + \sum_{j\in I^+} \lambda_j x_j^0 \left(\frac{x_j}{x_j^0}\right) =$$

$$= \sum_{j\in I^0} \lambda_j x_j + \sum_{j\in I^+} \gamma_j y_j, \qquad \gamma_j \geqslant 0. \tag{4.23}$$

Итак, каждое решение системы (4.20) можно представить в виде (4.22), (4.23). Но решение системы (4.19) получается из решения системы (4.20), если положить $x^0 = 1$. Таким образом, любое решение системы (4.19) можно представить в виде

$$x = \sum_{j\in I^0} \lambda_j x_j + \sum_{j\in I^+} \gamma_j y_j, \qquad \lambda_j \geqslant 0, \tag{4.24}$$

$$\sum_{j\in I^+} \gamma_j = 1, \qquad \gamma_j \geqslant 0, \quad j \in I^+. \tag{4.25}$$

Из этих формул видно, что первая сумма в правой части формулы (4.24) представляет собой точки некоторого многогранного конуса, вторая — многогранника.

То, что сумма многогранника и многогранного конуса есть многогранное множество, легко доказывается обратным ходом рассуждений от формул (4.24), (4.25) через соотношения (4.22), (4.23) к системе (4.20) и от нее ($x^0 = 1$) к системе (4.19). Теорема доказана.

5. Пересечение многогранного и произвольного конусов.

Теорема 4.16. *Если K — выпуклый, а K_0 — многогранный конусы и* $\operatorname{ri} K \cap K_0 \neq \varnothing$, *то*

$$(K \cap K_0)^* = K^* + K_0^*.$$

Доказательство. Так как $K \subseteq \operatorname{Lin} K$, то $K^* \supseteq (\operatorname{Lin} K)^\perp$, где $(\operatorname{Lin} K)^\perp$ — ортогональное дополнение $\operatorname{Lin} K$ до всего пространства X. Очевидно, что $K \cap K_0 \subseteq \operatorname{Lin} K$. Возьмем в качестве основного пространства $\operatorname{Lin} K$ и найдем сопряженный конус $(K \cap K_0)^*$ относительно этого пространства, т. е. рассматривая векторы x^* только из $\operatorname{Lin} K$. Так как $\operatorname{ri} K \cap K_0 \neq \varnothing$, то в пространстве $\operatorname{Lin} K$ применима теорема 3.2, т. е.

$$(K \cap K_0)^*_{\mathrm{Lin}K} = (K \cap (K_0 \cap \mathrm{Lin}\, K))^*_{\mathrm{Lin}K} = \\ = K^*_{\mathrm{Lin}K} + (K_0 \cap \mathrm{Lin}\, K)^*_{\mathrm{Lin}K}.$$

Нетрудно убедиться, что для любого конуса K_1, лежащего в $\mathrm{Lin}\, K$,

$$K_1^* = (K_1)^*_{\mathrm{Lin}K} + (\mathrm{Lin}\, K)^{\perp},$$

поэтому

$$(K \cap K_0)^* = K^*_{\mathrm{Lin}K} + (K_0 \cap \mathrm{Lin}\, K)^*_{\mathrm{Lin}K} + (\mathrm{Lin}\, K)^{\perp} = \\ = \left(K^*_{\mathrm{Lin}K} + (\mathrm{Lin}\, K)^{\perp}\right) + \left((K_0 \cap \mathrm{Lin}\, K)_{\mathrm{Lin}K} + (\mathrm{Lin}\, K)^{\perp}\right) = \\ = K^* + (K_0 \cap \mathrm{Lin}\, K)^*.$$

Но $\mathrm{Lin}\, K$ — многогранный конус, причем $(\mathrm{Lin}\, K)^* = (\mathrm{Lin}\, K)^{\perp}$, поэтому по теореме 4.14

$$(K_0 \cap \mathrm{Lin}\, K)^* = K_0^* + (\mathrm{Lin}\, K)^{\perp}.$$

Учитывая, что $K^* \supseteq (\mathrm{Lin}\, K)^{\perp}$, получаем

$$(K \cap K_0)^* = K^* + K_0^* + (\mathrm{Lin}\, K)^{\perp} = K^* + K_0^*.$$

Глава II

ВЫПУКЛЫЕ ФУНКЦИИ

§ 1. Основные свойства выпуклых функций

Выпуклые функции, как и выпуклые множества, представляют собой основной предмет изучения в выпуклом анализе и теории экстремальных задач. Оказывается, что из свойства выпуклости сразу следует целый ряд других свойств таких, как непрерывность, дифференцируемость по направлению и многие другие.

Будем рассматривать заданные на пространстве X $(=\mathbf{R}^n)$ функции $f(x)$, значения которых принадлежат расширенной действительной оси $[-\infty, +\infty]$. Таким образом, допускается, что функции принимают значения, равные $-\infty$ и $+\infty$.

С каждой функцией f свяжем два множества. Первое из них — *эффективная область* — множество точек, в которых $f(x)$ может принимать лишь конечное значение или значение $-\infty$:

$$\operatorname{dom} f = \{x\colon\ f(x) < +\infty\}.$$

Второе множество — это *надграфик функции* f — множество пар $x^0 \in R^1$ и $x \in X$ таких, что $x^0 \geqslant f(x)$:

$$\operatorname{epi} f = \{(x^0, x)\colon\ x^0 \geqslant f(x)\}.$$

Важно заметить, что точка $(x^0, x) \in \mathbf{R}^{n+1}$ только тогда может принадлежать $\operatorname{epi} f$, когда $x \in \operatorname{dom} f$, так как если $f(x) = +\infty$, то не существует вещественного числа x^0 такого, что $x^0 \geqslant f(x)$.

Надграфик функции f ее определяет. Действительно, легко проверить, что

$$f(x) = \inf_{x^0} \{x^0\colon\ (x^0, x) \in \operatorname{epi} f\}. \qquad (1.1)$$

Вообще, если взять в пространстве $\mathbf{R}^{n+1}$ произвольное множество, которое вместе с точкой (x^0, x) содержит и точки (y^0, x), где $y^0 \geqslant x^0$, и объявить его надграфиком некоторой функции, то формула (1.1) позволяет вычислить эту функцию. Таким образом, существует тесная связь между множествами в $\mathbf{R}^{n+1}$ и функциями на $\mathbf{R}^n$, которая позволяет свойства одних переносить на свойства других. Например, для функции имеется понятие производной, которого нет для множества, хотя это понятие может быть для множеств соответствующим образом проинтерпретировано.

1. Определения и основные свойства. Перейдем теперь к определению выпуклой функции.

Определение 1.1. Функция f называется *выпуклой*, если ее надграфик $\operatorname{epi} f$ есть выпуклое множество.

Определение 1.2. Функция называется *собственной*, если она не принимает значения $-\infty$ и не равна тождественно $+\infty$.

Из определения следует, что для собственной функции f $\operatorname{dom} f \neq \varnothing$ и $f(x)$ принимает конечные значения при $x \in \operatorname{dom} f$. Для собственной функции f определение выпуклости через надграфик может быть заменено на другое, более привычное и более часто употребляемое.

Лемма 1.1. *Если f — собственная функция, то, для того чтобы она была выпуклой, необходимо и достаточно, чтобы при всех x_1, x_2 выполнялось соотношение*

$$f(\lambda_1 x_1 + \lambda_2 x_2) \leqslant \lambda_1 f(x_1) + \lambda_2 f(x_2),$$
$$\lambda_1 \geqslant 0, \quad \lambda_2 \geqslant 0, \quad \lambda_1 + \lambda_2 = 1. \tag{1.2}$$

Доказательство. Если f — выпуклая функция, то по определению для всех $\lambda_1, \lambda_2 \geqslant 0$, $\lambda_1 + \lambda_2 = 1$, выполнено включение

$$\lambda_1 (x_1^0, x_1) + \lambda_2 (x_2^0, x_2) = (\lambda_1 x_1^0 + \lambda_2 x_2^0, \lambda_1 x_1 + \lambda_2 x_2) \in \operatorname{epi} f,$$

если $(x_1^0, x_1) \in \operatorname{epi} f$, $(x_2^0, x_2) \in \operatorname{epi} f$, или, иначе,

$$f(\lambda_1 x_1 + \lambda_2 x_2) \leqslant \lambda_1 x_1^0 + \lambda_2 x_2^0.$$

В частности, если $x_1^0 = f(x_1)$, $x_2^0 = f(x_2)$, то

$$f(\lambda_1 x_1 + \lambda_2 x_2) \leqslant \lambda_1 f(x_1) + \lambda_2 f(x_2),$$
$$\lambda_1 + \lambda_2 = 1, \qquad \lambda_1, \lambda_2 \geqslant 0.$$

Нетрудно видеть, что верно и обратное: если собственная функция удовлетворяет соотношению (1.2), то она выпукла. При этом в неравенстве (1.2) может встречаться значение, равное $+\infty$, однако, если обращаться с этим несобственным числом естественным образом, то недоразумений не возникает; просто всегда надо считать, что $+\infty$ — это «очень большое число», большее любого другого.

Лемма 1.2. $\operatorname{dom} f$ — *выпуклое множество, если функция* f *выпукла.*

Доказательство. Пусть $x_1, x_2 \in \operatorname{dom} f$. Тогда существуют такие конечные числа x_1^0, x_2^0, что $f(x_1) \leqslant x_1^0$, $f(x_2) \leqslant x_2^0$, так что $(x_1^0, x_1) \in \operatorname{epi} f$, $(x_2^0, x_2) \in \operatorname{epi} f$. Поэтому

$$\lambda_1 (x_1^0, x_1) + \lambda_2 (x_2^0, x_2) \in \operatorname{epi} f,$$

$$\lambda_1 + \lambda_2 = 1, \quad \lambda_1, \lambda_2 \geqslant 0,$$

т. е.

$$f(\lambda_1 x_1 + \lambda_2 x_2) \leqslant \lambda_1 x_1^0 + \lambda_2 x_2^0.$$

Значит, $f(\lambda_1 x_1 + \lambda_2 x_2)$ не равно $+\infty$, а поэтому $\lambda_1 x_1 + \lambda_2 x_2 \in \operatorname{dom} f$, что доказывает лемму.

Как показывает лемма, $\operatorname{dom} f$ — выпуклое множество, даже если f — несобственная выпуклая функция. Рассмотрим такую функцию. Пусть $y \in \operatorname{dom} f$ — точка, в которой $f(y) = -\infty$, а $x \in \operatorname{ri} \operatorname{dom} f$. Тогда $(x - y) \in \operatorname{Lin} \operatorname{dom} f$, и поэтому при достаточно малом $\varepsilon > 0$ $x_1 = x + \varepsilon(x - y) \in \operatorname{dom} f$. Легко видеть, что

$$x = \frac{1}{1+\varepsilon} x_1 + \frac{\varepsilon}{1+\varepsilon} y.$$

Если y^0 — любое число, $x_1^0 \geqslant f(x_1)$, то $(y^0, y) \in \operatorname{epi} f$ (так как $f(y) = -\infty$), $(x_1^0, x_1) \in \operatorname{epi} f$ и поэтому в силу выпуклости $\operatorname{epi} f$

$$\left(\frac{1}{1+\varepsilon} x_1^0 + \frac{\varepsilon}{1+\varepsilon} y^0, \frac{1}{1+\varepsilon} x_1 + \frac{\varepsilon}{1+\varepsilon} y\right) \in \operatorname{epi} f,$$

$$f(x) = f\left(\frac{1}{1+\varepsilon} x_1 + \frac{\varepsilon}{1+\varepsilon} y\right) \leqslant \frac{1}{1+\varepsilon} x_1^0 + \frac{\varepsilon}{1+\varepsilon} y^0.$$

Если в последнем неравенстве $y^0 \to -\infty$, то получим, что $f(x) = -\infty$.

Таким образом, если f — несобственная выпуклая функция, то $f(x) = -\infty$ для всех $x \in \operatorname{ri} \operatorname{dom} f$. Отсюда следует, что несобственная выпуклая функция может принимать конечные значения лишь на относительной границе своей области определения.

Приведем несколько свойств выпуклых функций.

Лемма 1.3. *Если $f_i(x)$, $i \in I$,— семейство выпуклых функций, где I — произвольное множество индексов, то функция*

$$f(x) = \sup_{i \in I} f_i(x)$$

выпуклая.

Доказательство. Нетрудно видеть, что

$$\operatorname{epi} f = \bigcap_{i \in I} \operatorname{epi} f_i.$$

Но пересечение выпуклых множеств выпукло по лемме I.1.1. Значит, $\operatorname{epi} f$ — выпуклое множество и функция f выпукла.

Лемма 1.4. *Пусть f — собственная выпуклая функция. Тогда*

$$f(\lambda_1 x_1 + \ldots + \lambda_m x_m) \leqslant \lambda_1 f(x_1) + \ldots + \lambda_m f(x_m),$$
$$\lambda_i \geqslant 0, \quad i = 1, \ldots, m, \quad \lambda_1 + \ldots + \lambda_m = 1.$$

Доказательство. Можно считать, что все λ_i больше нуля, ибо иначе число слагаемых можно было бы уменьшить. Если $x_i \notin \operatorname{dom} f$, то $f(x_i) = +\infty$, $\lambda_i f(x_i) = +\infty$ и неравенство тривиально выполняется, так как правая его часть равна $+\infty$.

Пусть теперь $x_i \in \operatorname{dom} f$, $i = 1, \ldots, m$. Так как $\operatorname{epi} f$ есть выпуклое множество, то из включения $(f(x_i), x_i) \in \operatorname{epi} f$ в силу леммы I.1.3 следует, что

$$(\lambda_1 f(x_1) + \ldots + \lambda_m f(x_m), \lambda_1 x_1 + \ldots + \lambda_m x_m) \in \operatorname{epi} f.$$

Поэтому

$$f(\lambda_1 x_1 + \ldots + \lambda_m x_m) \leqslant \lambda_1 f(x_1) + \ldots + \lambda_m f(x_m),$$

что и требовалось доказать.

Лемма 1.5. *Сумма собственных выпуклых функций с неотрицательными коэффициентами есть снова выпуклая функция.*

Доказательство. Если f_i, $i = 1, \ldots, m$,— собственные выпуклые функции, то в силу леммы 1.1

$$f_i(\lambda_1 x_1 + \lambda_2 x_2) \leqslant \lambda_1 f_i(x_1) + \lambda_2 f_i(x_2),$$

где $\lambda_1 \geqslant 0$, $\lambda_2 \geqslant 0$, $\lambda_1 + \lambda_2 = 1$. Умножая эти неравенства на неотрицательные коэффициенты $\alpha_i \geqslant 0$ и складывая, получим неравенство

$$\sum_{i=1}^{m} \alpha_i f_i(\lambda_1 x_1 + \lambda_2 x_2) \leqslant \lambda_1 \sum_{i=1}^{m} \alpha_i f_i(x_1) + \lambda_2 \sum_{i=1}^{m} \alpha_i f_i(x_2),$$

из которого следует справедливость утверждения леммы.

2. Критерии выпуклости. Леммы 1.3—1.5 дают возможность конструировать новые выпуклые функции из уже известных. Приведем несколько простых критериев, позволяющих распознавать выпуклость функции.

Пусть $f(x)$ — собственная функция. Выберем точку $p \in X$ и построим функцию

$$g_{x,p}(\alpha) = f(x + \alpha p).$$

Лемма 1.6. *Функция f выпукла тогда и только тогда, когда функция $g_{x,p}(\alpha)$ выпукла по α для любых x и p.*

Доказательство. Если $f(x)$ — выпуклая функция, то для любых $\lambda_1 \geqslant 0$, $\lambda_2 \geqslant 0$, $\lambda_1 + \lambda_2 = 1$, имеем

$$g_{x,p}(\lambda_1\alpha_1 + \lambda_2\alpha_2) = f(x + (\lambda_1\alpha_1 + \lambda_2\alpha_2)p) =$$
$$= f(\lambda_1(x + \alpha_1 p) + \lambda_2(x + \alpha_2 p)) \leqslant \lambda_1 f(x + \alpha_1 p) +$$
$$+ \lambda_2 f(x + \alpha_2 p) = \lambda_1 g_{x,p}(\alpha_1) + \lambda_2 g_{x,p}(\alpha_2).$$

Обратно, пусть функция $g_{x,p}(\alpha)$ выпукла по α при любых x и p. Тогда

$$f(\lambda_1 x_1 + \lambda_2 x_2) = f(x_2 + (\lambda_1 \cdot 1 + \lambda_2 \cdot 0)(x_1 - x_2)) =$$
$$= g_{x_2, x_1 - x_2}(\lambda_1 \cdot 1 + \lambda_2 \cdot 0) \leqslant \lambda_1 g_{x_2, x_1 - x_2}(1) +$$
$$+ \lambda_2 g_{x_2, x_1 - x_2}(0) = \lambda_1 f(x_1) + \lambda_2 f(x_2).$$

Эта простая лемма часто бывает полезной, так как позволяет сводить исследование выпуклости функции многих переменных к исследованию выпуклости функций одной переменной.

Будем обозначать градиент дифференцируемой функции в точке x через $f'(x)$, а матрицу вторых производных — через $f''(x)$. Таким образом, $f'(x)$ это вектор-столбец с компонентами $\partial f(x)/\partial x^i$, $i = 1, \ldots, n$, а

$$f''(x) = \left\{\frac{\partial^2 f(x)}{\partial x^i \partial x^j}\right\}, \qquad i = 1, \ldots, n, \quad j = 1, \ldots, n.$$

Если $g(\alpha) = f(x + \alpha p)$, то, пользуясь правилами вычисления производной от сложной функции, легко установить, что

$$g'(\alpha)=\langle p,\ f'(x+\alpha p)\rangle,$$

$$g''(\alpha)=\langle p,\ f''(x+\alpha p)p\rangle.$$

Теорема 1.1. *Пусть $f(x)$ — дифференцируемая функция. Тогда следующие утверждения эквивалентны:*

1) $f(x)$ — выпуклая функция;

2) $f(x_2)-f(x_1)\geqslant\langle x_2-x_1,\ f'(x_1)\rangle$ для любых x и p;

3) $\langle p,\ f'(x+\alpha p)\rangle$ — неубывающая функция α для любых x и p.

Если f дважды дифференцируема, то

4) $f''(x)$ есть неотрицательно определенная матрица.

Доказательство. Покажем, что из 1) следует 2) или, коротко, 1) $\to$ 2). Действительно поскольку

$$f((1-\lambda)x_1+\lambda x_2)\leqslant(1-\lambda)f(x_1)+\lambda f(x_2),\qquad 0<\lambda<1,$$

то

$$\frac{f(x_1+\lambda(x_2-x_1))-f(x_1)}{\lambda}\leqslant f(x_2)-f(x_1).$$

Переходя к пределу при $\lambda\to 0$, получим

$$\langle x_2-x_1,\ f'(x_1)\rangle\leqslant f(x_2)-f(x_1). \tag{1.3}$$

Покажем теперь, что 2) $\to$ 3). Положим из неравенства (1.3) при $x_1=x+\alpha_1 p$, $x_2=x+\alpha_2 p$ следует, что

$$g_{x,p}(\alpha_2)-g_{x,p}(\alpha_1)\geqslant(\alpha_2-\alpha_1)\langle p,\ f'(x+\alpha_1 p)\rangle,$$

а при $x_1=x+\alpha_2 p$, $x_2=x+\alpha_1 p$ —

$$g_{x,p}(\alpha_1)-g_{x,p}(\alpha_2)\geqslant(\alpha_1-\alpha_2)\langle p,\ f'(x+\alpha_2 p)\rangle.$$

Из двух последних неравенств (для определенности $\alpha_2>\alpha_1$) следует, что

$$\langle p,\ f'(x+\alpha_1 p)\rangle\leqslant\frac{g_{x,p}(\alpha_2)-g_{x,p}(\alpha_1)}{\alpha_2-\alpha_1}\leqslant\langle p,\ f'(x+\alpha_2 p)\rangle$$

Покажем, что 3) $\to$ 1). Пусть $g'_{x,p}(\alpha)=\langle p,\ f'(x+\alpha p)\rangle$ — неубывающая функция α. Тогда $g'_{x,p}(\alpha_1)\leqslant g'_{x,p}(\alpha_2)$ при $\alpha_2\geqslant\alpha_1$.

Если $0 < \mu < 1$, то

$$0 \leqslant \mu(\alpha_2 - \alpha_1) \int_0^1 \left[g'_{x,p}(\alpha_1 + \tau(\alpha_2 - \alpha_1)) - \right.$$

$$\left. - g'_{x,p}(\alpha_1 + \tau\mu(\alpha_2 - \alpha_1)) \right] d\tau = (1-\mu) g_{x,p}(\alpha_1) + \mu g_{x,p}(\alpha_2) -$$

$$- g(x, p)((1-\mu)\alpha_1 + \mu\alpha_2),$$

т. е. $g_{x,p}(\alpha)$ — выпуклая функция. На основании леммы 1.6 отсюда можно заключить, что $f(x)$ выпукла.

Пусть теперь f дважды дифференцируема. Покажем, что 4) $\rightleftarrows$ 3). Так как $g'_{x,p}(\alpha)$ — неубывающая функция, то

$$g''_{x,p}(\alpha) = \langle p, f''(x + \alpha p) \rangle \geqslant 0, \tag{1.4}$$

откуда следует, что матрица $f''(x)$ неотрицательно определена. Обратно, если условие (1.4) выполнено, то $g''_{x,p}(\alpha)$ неотрицательна, и, значит, $g'_{x,p}(\alpha) = \langle p, f'(x + \alpha p) \rangle$ — неубывающая функция.

Так как мы показали, что 1) $\to$ 2) $\to$ 3) $\to$ 1) и 4) $\rightleftarrows$ 3), то тем самым эквивалентность всех четырех утверждений доказана.

Лемма 1.7. *Если функция $f_0(x)$ определена на выпуклом множестве D и удовлетворяет на нем соотношению* (1.2), *то*

$$f(x) = \begin{cases} f_0(x), & x \in D, \\ +\infty, & x \notin D, \end{cases}$$

— выпуклая функция.

Доказательство леммы предоставляется читателю.

Эта лемма часто бывает полезной. Так, в частности, если G — выпуклое множество, то функция

$$\delta(x|G) = \begin{cases} 0, & x \in G, \\ +\infty, & x \notin G, \end{cases}$$

называемая *индикаторной функцией множества*, является выпуклой.

Приведем несколько примеров выпуклых функций (их выпуклость легко проверяется на основании теоремы 1.1 и леммы 1.7):

1) $f(x) = e^{\alpha x}$, где $-\infty < \alpha < +\infty$;

2) $f(x)=x^p$, если $x\geqslant 0$, $1\leqslant p<\infty$,
$f(x)=+\infty$, если $x<0$;
3) $f(x)=-x^p$, если $x\geqslant 0$, $0\leqslant p\leqslant 1$;
$f(x)=+\infty$, если $x<0$,
4) $f(x)=x^p$, если $x>0$, $-\infty<p\leqslant 0$,
$f(x)=+\infty$, если $x\leqslant 0$;
5) $f(x)=-\ln x$, если $x>0$,
$f(x)=+\infty$, если $x\leqslant 0$;
6) $f(x)=\frac{1}{2}\langle x, Ax\rangle+\langle d, x\rangle$, если A — неотрицательно определенная матрица.

3. Вспомогательная лемма. Приводимая ниже лемма часто используется при доказательствах последующих утверждений.

Лемма 1.8. *Пусть $g(\alpha)$ — выпуклая функция аргумента α; $\alpha_0<\alpha_1<\alpha_2$, α_0, α_1, $\alpha_2\in\operatorname{dom} g$. Тогда*

$$\frac{g(\alpha_2)-g(\alpha_0)}{\alpha_2-\alpha_0}\geqslant\frac{g(\alpha_1)-g(\alpha_0)}{\alpha_1-\alpha_0},$$

$$\frac{g(\alpha_1)-g(\alpha_0)}{\alpha_1-\alpha_0}\leqslant\frac{g(\alpha_2)-g(\alpha_1)}{\alpha_2-\alpha_1}.$$

Доказательство. Положим

$$\lambda_1=\frac{\alpha_1-\alpha_0}{\alpha_2-\alpha_0},\quad \lambda_2=1-\lambda_1=\frac{\alpha_2-\alpha_1}{\alpha_2-\alpha_0}.$$

Тогда

$$\lambda_1\alpha_2+\lambda_2\alpha_0=\frac{\alpha_1-\alpha_0}{\alpha_2-\alpha_0}\alpha_2+\frac{\alpha_2-\alpha_1}{\alpha_2-\alpha_0}\alpha_0=\alpha_1.$$

Поэтому

$$g(\alpha_1)=g(\lambda_1\alpha_2+\lambda_2\alpha_0)\leqslant\frac{\alpha_1-\alpha_0}{\alpha_2-\alpha_0}\,g(\alpha_2)+\frac{\alpha_2-\alpha_1}{\alpha_2-\alpha_0}g(\alpha_0).\tag{1.5}$$

Полученное неравенство можно преобразовать двумя способами.

1. Вычитая из обеих частей $g(\alpha_0)$, получим

$$g(\alpha_1)-g(\alpha_0)\leqslant\frac{\alpha_1-\alpha_0}{\alpha_2-\alpha_0}(g(\alpha_2)-g(\alpha_0)),$$

откуда, после деления на $\alpha_1-\alpha_0$, вытекает первое неравенство леммы.

2. Так как $\lambda_1 + \lambda_2 = 1$, то

$$\frac{\alpha_1 - \alpha_0}{\alpha_2 - \alpha_0} g(\alpha_1) + \frac{\alpha_2 - \alpha_1}{\alpha_2 - \alpha_0} g(\alpha_1) \leqslant \frac{\alpha_1 - \alpha_0}{\alpha_2 - \alpha_0} g(\alpha_2) + \frac{\alpha_2 - \alpha_1}{\alpha_2 - \alpha_0} g(\alpha_0),$$

или

$$\frac{\alpha_2 - \alpha_1}{\alpha_2 - \alpha_0} [g(\alpha_1) - g(\alpha_0)] \leqslant \frac{\alpha_1 - \alpha_0}{\alpha_2 - \alpha_0} (g(\alpha_2) - g(\alpha_1)),$$

откуда получаем второе неравенство леммы.

Заметим, что первое из неравенств, доказанных в лемме, показывает, что функция

$$\frac{g(\alpha) - g(\alpha_0)}{\alpha - \alpha_0}$$

монотонно не убывает при возрастании α, $\alpha > \alpha_0$.

4. Непрерывность выпуклых функций. Свойство выпуклости оказывается тесно связанным со свойством непрерывности.

Теорема 1.2. *Пусть собственная выпуклая функция ограничена сверху в окрестности некоторой точки $x_0 \in \operatorname{dom} f$. Тогда она непрерывна в этой точке.*

Доказательство. Без ограничения общности предположим, что выбранной точкой является точка $x_0 = 0$. Пусть Ω — некоторый открытый шар с центром в нуле и $f(x) \leqslant c_1$ для всех $x \in \Omega$. Рассмотрим функцию

$$g(\alpha) = f(\alpha x)$$

при фиксированном $x \in \Omega$. Положив в первом неравенстве леммы 1.8 $\alpha_0 = 0$, $\alpha_1 = \alpha > 0$, $\alpha_2 = 1$, получим

$$\frac{g(\alpha) - g(0)}{\alpha} \leqslant \frac{g(1) - g(0)}{1}.$$

Так как

$$g(1) = f(x) \leqslant c_1, \quad g(0) = f(0) \leqslant c_1,$$

то

$$f(\alpha x) - f(0) \leqslant 2c_1\alpha. \tag{1.6}$$

Далее, положив во втором неравенстве леммы 1.8 $\alpha_0 = -1$, $\alpha_1 = 0$, $\alpha_2 = \alpha$, получим

$$\frac{g(0) - g(-1)}{0 - (-1)} \leqslant \frac{g(\alpha) - g(0)}{\alpha},$$

откуда

$$-2c_1\alpha \leqslant f(\alpha x) - f(0). \tag{1.7}$$

Итак, согласно неравенствам (1.6) и (1.7) имеем

$$|f(\alpha x) - f(0)| \leqslant 2c_1\alpha. \tag{1.8}$$

Возьмем теперь $\varepsilon > 0$ и положим $\delta = \varepsilon/(2c_1) < 1$, $\Omega_\delta = \delta\Omega$. Пусть $y \in \Omega_\delta$. Тогда существует такой вектор $x \in \Omega$, что $y = \delta x$. Согласно оценке (1.8) получим

$$|f(y) - f(0)| = |f(\delta x) - f(0)| \leqslant 2\delta c_1 = \varepsilon,$$

что доказывает непрерывность функции в точке нуль.

Теорема 1.3. *Если $f(x)$ — выпуклая функция, непрерывная в точке x_0, то она удовлетворяет условию Липшица в этой точке, т. е.*

$$|f(x) - f(x_0)| \leqslant L\|x - x_0\|$$

при $L = \mathrm{const}$ для всех x из некоторой окрестности точки x_0.

Доказательство. Будем считать, что $x_0 = 0$. Пусть r — радиус шара Ω с центром в нуле. Возьмем $y \in \Omega$, удовлетворяющее неравенству $\|y\| < r/2$, и положим

$$x = \frac{r}{2}\frac{y}{\|y\|}.$$

Тогда, используя неравенство (1.8), получим

$$|f(y) - f(0)| = \left|f\left(\frac{2\|y\|}{r}x\right) - f(0)\right| \leqslant L\|y\|,$$

где $L = 4c/r$.

Теорема 1.4. *Пусть f — выпуклая собственная функция. Тогда она непрерывна на множестве* $\operatorname{ri}\operatorname{dom} f$.

Доказательство. Любую точку $x_0 \in \operatorname{ri}\operatorname{dom} f$ можно сделать внутренней точкой некоторого симплекса с вершинами $y_0, \ldots, y_k \in \operatorname{dom} f$, где $k = \dim\operatorname{dom} f$. Любая точка этого симплекса имеет вид

$$x = \lambda_0 y_0 + \ldots + \lambda_k y_k, \quad \lambda_i \geqslant 0, \quad i = 0, 1, \ldots, k,$$
$$\lambda_0 + \lambda_1 + \ldots + \lambda_k = 1.$$

Поэтому

$$f(x) \leqslant \lambda_0 f(y_0) + \ldots + \lambda_k f(y_k) \leqslant \max_i f(y_i),$$

т. е. функция $f(x)$ ограничена в некоторой окрестности точки x_0. Рассмотрим выпуклую функцию от $y = x - x_0$:

$$g(y) = f(y + x_0),$$

относительно подпространства $\operatorname{Lin} \operatorname{dom} f$. Применив теорему 1.2, получим требуемый результат.

Как видно из доказанной теоремы, выпуклая функция непрерывна внутри области своего определения и может иметь разрывы только на ее границе. Для того чтобы охарактеризовать случаи, когда таких разрывов нет, удобно ввести понятие замкнутости функции.

Определение 1.3. Функция f называется *замкнутой*, если ее надграфик $\operatorname{epi} f$ есть замкнутое множество в $\mathbf{R}^{n+1}$.

Теорема 1.5. *Следующие три утверждения эквивалентны:*

1) *функция f замкнута;*
2) *множества уровня $C_\alpha = \{x\colon f(x) \leqslant \alpha\}$ замкнуты;*
3) *функция f полунепрерывна снизу.*

Напомним, что функция f называется *полунепрерывной снизу* в точке x_0, если для любой последовательности $x_k \to x_0$.

$$\liminf_k f(x_k) \geqslant f(x_0). \tag{1.9}$$

Доказательство. Покажем, что из 1) следует 2). Пусть $x_k \to x_0$ и $f(x_k) \leqslant \alpha$. Без ограничения общности можно считать, что $f(x_k) \to \mu \leqslant \alpha$, где μ либо конечно, либо равно $-\infty$.

Покажем, что если $\mu = -\infty$, то $f(x_0) = -\infty$. В самом деле, если $f(x_0) = \mu_0$ конечно, то для больших k $f(x_k) < \mu_0 - \varepsilon$. Рассмотрим последовательность точек $(\mu_0 - \varepsilon, x_k)$, которые принадлежат $\operatorname{epi} f$. Так как $(\mu_0 - \varepsilon, x_k) \to (\mu_0 - \varepsilon, x_0)$ и $\operatorname{epi} f$ — замкнутое множество, то $(\mu_0 - \varepsilon, x_0) \in \operatorname{epi} f$, т. е. $\mu_0 - \varepsilon \geqslant f(x_0) = \mu_0$; получено противоречие. Значит, если $\mu = -\infty$, то $f(x_0) = -\infty$, т. е. $x_0 \in C_\alpha$.

Если μ конечно, то $(f(x_k), x_k) \to (\mu, x_0) \in \operatorname{epi} f$ и, значит, $\alpha \geqslant \mu \geqslant f(x_0)$, т. е. опять $x_0 \in C_\alpha$. Значит, C_α замкнуто.

Пусть теперь 2) выполнено. Если $x_k \to x_0$ и $f(x_k) \to \alpha$, то для любого $\varepsilon > 0$ и достаточно больших k выполняется неравенство $f(x_k) \leqslant \alpha + \varepsilon$, и, значит, $f(x_0) \leqslant \alpha + \varepsilon$, ибо

$C_{\alpha+\varepsilon}$ замкнуто. Так как ε произвольно мало, то $f(x_0) \leqslant \alpha$, что показывает справедливость утверждения 3).

Наконец, покажем, что из 3) следует 1). Действительно, если $x_k^0 \geqslant f(x_k)$, $(x_k^0, x_k) \to (x^0, x_0)$, то согласно неравенству (1.9)

$$x^0 \geqslant \liminf_k f(x_k) \geqslant f(x_0),$$

т. е. $(x^0, x_0) \in \operatorname{epi} f$, что доказывает замкнутость $\operatorname{epi} f$.

Замечание. На основании утверждения 3) теоремы легко проверить, что сумма замкнутых функций также замкнута (на основании утверждения 1) это сделать гораздо сложнее).

Теорема 1.6. *Если $f(x)$ — выпуклая замкнутая функция, принимающая в некоторой точке x_0 конечное значение, то $f(x) > -\infty$.*

Доказательство. Выше было показано, что если $f(x)$ принимает значение $-\infty$, то $f(x) = -\infty$ для всех $x \in \operatorname{ri}\operatorname{dom} f$. Но по теореме I.1.3, если $x \in \operatorname{dom} f$, то $(1-\lambda)x + \lambda x_1 \in \operatorname{ri}\operatorname{dom} f$ для всех λ и x_1 таких, что $x_1 \in \in \operatorname{ri}\operatorname{dom} f$. Так как функция f замкнута, то

$$f(x) \leqslant \lim_{\lambda \downarrow 0} f((1-\lambda)x + \lambda x_1) = -\infty,$$

т. е. $f(x) = -\infty$ для всех $x \in \operatorname{dom} f$.

Итак, если $f(x)$ принимает конечное значение хотя бы в одной точке, то $f(x) > -\infty$ при всех x.

В заключение этого параграфа введем

Определение 1.4. Пусть f — выпуклая функция; тогда функцию $\bar{f}$, надграфик которой есть замыкание надграфика f, называют *замыканием функции* f и пишут

$$\operatorname{epi} \bar{f} = \overline{\operatorname{epi} f},$$

$$\bar{f}(x) = \inf_{x^0} \{x^0 \colon (x^0, x) \in \overline{\operatorname{epi} f}\}.$$

Так как замыкание выпуклого множества снова выпукло (лемма I.1.5), то $\bar{f}$ — выпуклая функция.

§ 2. Сопряженные функции

Применение теорем отделимости к выпуклым функциям приводит естественным образом к понятию сопряженной функции. Напомним, что согласно § 2 главы I замкнутое выпуклое множество полностью характеризуется своей опорной функцией

$$W_M(x^*) = \sup_x \{\langle x, x^*\rangle : x \in M\}.$$

Вычислим опорную функцию надграфика замкнутой выпуклой функции f. Учитывая, что $\operatorname{epi} f$ лежит в пространстве $\mathbf{R}^{n+1}$, получим

$$W_{\operatorname{epi} f}(x^{0*}, x^*) = \sup_{(x^0, x)} \{\langle x, x^*\rangle + x^0 x^{0*} : (x^0, x) \in \operatorname{epi} f\}.$$

Если $x^{0*} > 0$, то при данном x число x^0 может быть любым, бо́льшим $f(x)$; поэтому $W_{\operatorname{epi} f}(x^{0*}, x^*) = +\infty$. Остается вычислить $W_{\operatorname{epi} f}$ при $x^{0*} \leqslant 0$. Ниже будет показано, что для вычисления $W_{\operatorname{epi} f}$ при $x^{0*} \leqslant 0$ в случае, когда f — замкнутая выпуклая функция, достаточно найти $W_{\operatorname{epi} f}(-1, x^*)$. Если же $x^{0*} = -1$, то

$$W_{\operatorname{epi} f}(-1, x^*) = \sup_{(x^0, x)} \{\langle x, x^*\rangle - x^0 : x^0 \geqslant f(x)\} = \sup_x \{\langle x, x^*\rangle - f(x)\}.$$

1. Определение и основные свойства.

Определение 2.1. Функция

$$f^*(x^*) = \sup_x \{\langle x, x^*\rangle - f(x)\}$$

называется *сопряженной к* f.

Лемма 2.1. *Сопряженная функция замкнута и выпукла.*

Доказательство. При фиксированном x функция $\langle x, x^*\rangle - f(x)$ есть линейная функция x^*, а значит, она замкнута и выпукла по этому аргументу. Надграфик $f^*(x^*)$ есть пересечение надграфиков функций $\langle x, x^*\rangle - f(x)$, т. е. пересечение замкнутых выпуклых множеств, и поэтому $\operatorname{epi} f^*$ — также замкнутое и выпуклое множество.

Лемма 2.2. *Справедливо неравенство*

$$f(x) + f^*(x^*) \geqslant \langle x, x^*\rangle.$$

Это неравенство, называющееся *неравенством Юнга*, вытекает непосредственно из определения сопряженной функции.

Следующая теорема, являющаяся в некотором смысле основной, описывает связь между функцией f и сопряженной к ней функцией f^*.

Теорема 2.1. *Пусть f — собственная выпуклая замкнутая функция. Тогда*

$$f(x) = f^{**}(x),$$

где согласно определению 2.1

$$f^{**}(x) = \sup_{x^*} \{\langle x, x^* \rangle - f^*(x^*)\}. \tag{2.1}$$

Доказательство. Так как согласно предыдущей лемме

$$f(x) \geqslant \langle x, x^* \rangle - f^*(x^*),$$

то

$$f(x) \geqslant f^{**}(x).$$

Отсюда, в частности, следует, что $\operatorname{dom} f \subseteq \operatorname{dom} f^{**}$.

Пусть $x_0 \in \operatorname{dom} f$. Выберем $\gamma < f(x_0)$ и рассмотрим в $\mathbf{R}^{n+1}$ множество P_ε:

$$P_\varepsilon = \{(x^0, x):\ x - x_0 \in \varepsilon B,\ x^0 \leqslant \gamma\},$$

где B — единичный шар в $\mathbf{R}^n$. Так как epi f — замкнутое множество, то P_ε и epi f при достаточно малом ε не пересекаются. Предположим противное. Тогда найдется последовательность $(x_k^0, x_k) \in P_{\varepsilon_k}$, $\varepsilon_k \to 0$, $x_k^0 \leqslant \gamma$, такая, что $(x_k^0, x_k) \in \operatorname{epi} f$. Так как $\varepsilon_k \to 0$, то $x_k \to x_0$, $\lim\limits_{k \to \infty} x_k^0 \leqslant \gamma$ и в силу замкнутости epi f $\gamma \geqslant \lim\limits_{k \to \infty} x_k^0 \geqslant f(x_0)$, что противоречит выбору γ.

Разделим epi f и P_ε, т. е. найдем такой ненулевой вектор (x^{0*}, x^*), что

$$\begin{gathered} x^{0*}y^0 + \langle y, x^* \rangle \geqslant x^{0*}x^0 + \langle x, x^* \rangle, \\ (y^0, y) \in P_\varepsilon,\ (x^0, x) \in \operatorname{epi} f. \end{gathered} \tag{2.2}$$

Так как y^0 можно взять любым, меньшим γ, то из неравенства (2.2) сразу следует, что $x^{0*} \leqslant 0$. Покажем, что на самом деле $x^{0*} < 0$. В предположении $x^{0*} = 0$ формула

(2.2) переходит в формулу

$$\langle y, x^*\rangle \geqslant \langle x, x^*\rangle, \quad y - x_0 \in \varepsilon B, \quad x \in \operatorname{dom} f.$$

Положив $x = x_0$, получим, что

$$\langle y - x_0, x^*\rangle \geqslant 0, \qquad y - x_0 \in \varepsilon B.$$

В силу произвольности $y - x_0 \in \varepsilon B$ это неравенство выполняется только в случае, если $x^* = 0$, что находится в противоречии с тем, что вектор (x^{0*}, x^*) не равен нулю.

Положим теперь $x^{0*} = -1$. Тогда (2.2) приобретает следующий вид:

$$\begin{gathered} -y^0 + \langle y, x^*\rangle \geqslant -x^0 + \langle x, x^*\rangle, \\ y^0 \leqslant \gamma, \quad y - x_0 \in \varepsilon B, \quad x^0 \geqslant f(x). \end{gathered} \tag{2.3}$$

Пусть $y^0 = \gamma$, $x^0 = f(x)$, $y = x_0$. Тогда из (2.3) получаем, что

$$\langle x_0, x^*\rangle - \gamma \geqslant \sup_x \{\langle x, x^*\rangle - f(x)\} = f^*(x^*)$$

или

$$\gamma \leqslant \langle x_0, x^*\rangle - f^*(x^*),$$

откуда

$$\gamma \leqslant \sup_{x^*} \{\langle x_0, x^*\rangle - f(x^*)\} = f^{**}(x_0). \tag{2.4}$$

Так как γ — произвольное число, меньшее $f(x_0)$, то из соотношения (2.4) получаем, что

$$f(x_0) \leqslant f^{**}(x_0).$$

Сопоставляя это неравенство с доказанным ранее ($f \geqslant f^{**}$), получаем требуемый результат: $f(x) = f^{**}(x_0)$.

Пусть теперь $x_0 \notin \operatorname{dom} f$, т. е. $f(x_0) = +\infty$. Тогда точно так же, как и ранее, взяв произвольное γ, можно убедиться, что epi f и P_ε при достаточно малом ε не пересекаются и справедливо неравенство (2.2).

Если для произвольно больших γ имеем $x^{0*} < 0$, то для этих γ выполняется неравенство (2.4), и, значит, $f^{**}(x_0) = +\infty$.

Пусть теперь $x^{0*} = 0$ при некотором γ. Тогда из условия (2.2) получаем, что

$$\langle y, x^*\rangle \geqslant \langle x, x^*\rangle, \quad y - x_0 \in \varepsilon B, \quad x \in \operatorname{dom} f.$$

Вычтем из обеих частей этого неравенства произведение $\langle x_0, x^* \rangle$ и, взяв в левой части минимум по $y - x \in \varepsilon B$, а в правой — точную верхнюю грань по $x \in \operatorname{dom} f$, получим

$$-\varepsilon \| x^* \| \geqslant \sup_{x \in \operatorname{dom} f} \langle x, x^* \rangle - \langle x_0, x^* \rangle. \tag{2.5}$$

Вспомним теперь, что существует точка x_1, в которой $f(x_1)$ — конечное число. Если применить к этой точке предыдущие рассуждения, то, подставляя в неравенство (2.3) $\gamma < f(x_1)$, $y = x_1$, получим

$$-\gamma + \langle x_1, x^* \rangle \geqslant -f(x) + \langle x, x^* \rangle,$$

т. е. $f^*(x^*) \leqslant -\gamma + \langle x_1, x^* \rangle$. Значит, существует такой вектор x^* (обозначим его через x_1^*), что $f\left(x_1^*\right)$ конечно. Отсюда следует, что

$$f^*\left(x_1^* + \alpha x^*\right) = \sup_x \{\langle x, x_1^* + \alpha x^* \rangle - f(x)\} \leqslant$$
$$\leqslant f^*\left(x_1^*\right) + \alpha \sup_{x \in \operatorname{dom} f} \langle x, x^* \rangle,$$

а поэтому с учетом неравенства (2.5) имеем

$$f^{**}(x_0) \geqslant \langle x_0, x_1^* + \alpha x^* \rangle - f^*\left(x_1^* + \alpha x^*\right) \geqslant$$
$$\geqslant \langle x_0\, x_1^* \rangle - f^*\left(x_1^*\right) + \alpha\left[\langle x_0, x^* \rangle - \sup_{x \in \operatorname{dom} f} \langle x, x^* \rangle\right] \geqslant$$
$$\geqslant \langle x_0, x_1^* \rangle - f^*\left(x_1^*\right) + \alpha\varepsilon \| x^* \| \to +\infty$$

при $\alpha \to +\infty$. Теорема доказана.

Теорема 2.2. *Если $f(x)$ — собственная замкнутая выпуклая функция, то $f^*(x^*)$ — также собственная замкнутая выпуклая функция.*

Доказательство. Согласно лемме 2.1 функция $f^*(x^*)$ — замкнутая и выпуклая. Остается доказать, что $f^*(x^*)$ — собственная функция. Так как $f(x_1)$ для некоторого $x_1 \in \operatorname{dom} f$ конечно, то

$$f^*(x^*) \geqslant \langle x_1, x^* \rangle - f(x_1),$$

и поэтому $f^*(x^*)$ не принимает значения $-\infty$. В ходе доказательства предыдущей теоремы показано существование вектора x_1^*, для которого $f^*(x_1^*)$ конечно. Отсюда следует утверждение теоремы.

Заметим теперь, что доказательство теоремы 2.1 в случае $x_0 \in \operatorname{dom} f$ основывалось на том, что $\operatorname{epi} f$ и P_ε не пересекаются при достаточно малом $\varepsilon > 0$. При этом использовалась замкнутость $\operatorname{epi} f$. Нетрудно видеть, что утверждение остается справедливым, если потребовать, чтобы функция $f(x)$ была полунепрерывна снизу в точке x_0. Поэтому справедлива

Теорема 2.3. *Если $x_0 \in \operatorname{dom} f$ и функция f полунепрерывна снизу в точке x_0, то*

$$f(x_0) = f^{**}(x_0).$$

2. Положительно однородные выпуклые функции. Функция называется *положительно однородной*, если $f(\lambda x) = \lambda f(x)$ для $\lambda > 0$. Если к тому же $f(x)$ выпукла, то

$$f(x_1 + x_2) = f\left(2\left(\frac{1}{2} x_1 + \frac{1}{2} x_2\right)\right) = 2f\left(\frac{1}{2} x_1 + \frac{1}{2} x_2\right) \leqslant \leqslant f(x_1) + f(x_2),$$

или, в более общем виде,

$$f(x_1 + \ldots + x_m) \leqslant f(x_1) + \ldots + f(x_m).$$

Полагая в предыдущем неравенстве $x_1 = 0$, $x_2 = x$, получим $f(x) \leqslant f(0) + f(x)$, т. е. $f(0) \geqslant 0$.

Пусть $f(x)$ положительно однородна, выпукла и замкнута. В силу замкнутости f

$$\lim_{\lambda \downarrow 0} f(\lambda x) = \lim_{\lambda \downarrow 0} \lambda f(x) = 0 \geqslant f(0).$$

Таким образом, $f(0) = 0$ для замкнутой выпуклой положительно однородной функции.

Вычислим сопряженную функцию для такой функции:

$$f^*(x^*) = \sup_{x^*} \{\langle x, x^* \rangle - f(x)\} \geqslant -f(0) = 0.$$

Пусть существует такое x_1, что

$$\langle x_1, x^* \rangle - f(x_1) > 0.$$

Тогда

$$f^*(x^*) \geqslant \sup_{\lambda > 0} \{\langle \lambda x_1, x^* \rangle - f(\lambda x_1)\} = = \sup_{\lambda > 0} \lambda [\langle x_1, x^* \rangle - f(x_1)] = +\infty.$$

Таким образом, $f^*(x^*)$ принимает только два значения: 0 и $+\infty$. Поэтому

$$f^*(x^*)=\delta(x^*\,|\,\operatorname{dom} f^*)=\begin{cases} \quad 0, & x^*\in\operatorname{dom} f^*,\\ +\infty, & x^*\notin\operatorname{dom} f^*.\end{cases}$$

Теорема 2.4. *Если f — положительно однородная выпуклая замкнутая функция, то*

$$f^*(x^*)=\delta(x^*|\operatorname{dom} f^*)$$

и $\operatorname{dom} f^*$ *— замкнутое множество.*

Осталось доказать только последнее утверждение теоремы. Но $f^*(x^*)$ — замкнутая функция, а легко видеть, что индикаторная функция множества замкнута тогда и только тогда, когда это множество замкнуто.

Теорема 2.5. *Если f — положительно однородная выпуклая замкнутая функция, то*

$$f(x)=\sup_{x^*}\{\langle x, x^*\rangle\colon\ x^*\in\operatorname{dom} f^*\}.$$

Доказательство получается прямым применением теоремы 2.1. Действительно, $f(x)=f^{**}(x)$. Но по теореме 2.4 $f^*=\delta(\,\cdot\,|\operatorname{dom} f^*)$ и

$$f^{**}(x)=\sup_{x^*}\{\langle x, x^*\rangle-\delta(x^*\,|\operatorname{dom} f^*)\}=$$
$$=\sup_{x^*}\{\langle x, x^*\rangle\colon\ x^*\in\operatorname{dom} f^*\}.$$

Теорема 2.6. *Функция, сопряженная к опорной функции замкнутого выпуклого множества, есть индикаторная функция этого множества.*

Доказательство. Пусть C^* — выпуклое замкнутое множество в X^* и

$$f(x)=\sup_{x^*}\{\langle x, x^*\rangle\colon\ x^*\in C^*\}. \tag{2.6}$$

Очевидно, что $f(x)$ выпукла, замкнута и положительно однородна. Кроме того, $f(x)$ есть опорная функция множества C^*.

Рассмотрим индикаторную функцию множества C^*: $\delta(x^*|C^*)$. Тогда сопряженной к ней функцией будет

$$\sup_{x^*}\{\langle x, x^*\rangle-\delta(x^*\,|\,C^*)\}=\sup_{x^*}\{\langle x, x^*\rangle\colon\ x^*\in C^*\}=f(x).$$

Итак, $f(x)$ есть функция, сопряженная к $\delta(x^*|C^*)$. Применяя теорему 2.1, получим

$$f^*(x^*) = \delta(x^*|C^*), \qquad \operatorname{dom} f^* = C^*.$$

§ 3. Производные по направлениям и субдифференциалы

Выпуклые функции, вообще говоря, не дифференцируемы в обычном смысле. Тем не менее они обладают производными по направлению. Кроме того, для выпуклых функций можно определить понятие субградиента, которое заменяет обычное понятие градиента гладкой функции в задачах на экстремум. Введению этих понятий и их свойствам посвящен данный параграф.

Напомним, что *производной по направлению* $p \in X$ от функции f в точке x называется величина

$$f'(x, p) = \lim_{\lambda \downarrow 0} \frac{f(x+\lambda p) - f(x)}{\lambda},$$

если предел в правой части равенства существует.

1. Определения и основные свойства.

Лемма 3.1. *Пусть $f(x)$ — выпуклая собственная функция, $x_0 \in \operatorname{dom} f$. Тогда величина $f'(x_0, p)$ — конечная или бесконечная — существует при любом p.*

Доказательство. Если $x_0 + \lambda p \notin \operatorname{dom} f$ при всех $\lambda > 0$, то $f(x_0+\lambda p) = +\infty$ и $f'(x_0, p) = +\infty$. Если $x_0 + \lambda p \in \operatorname{dom} f$ для малых λ, то согласно лемме 1.8 отношение

$$\frac{f(x_0+\lambda p) - f(x_0)}{\lambda}$$

есть невозрастающая функция λ, когда λ стремится к нулю убывая. Поэтому предел

$$f'(x, p) = \lim_{\lambda \downarrow 0} \frac{f(x_0+\lambda p) - f(x_0)}{\lambda}$$

существует всегда, что доказывает лемму.

Лемма 3.2. *Если $x_0 + \lambda p \in \operatorname{dom} f$ для $\lambda \in [-\varepsilon, +\varepsilon]$, $\varepsilon > 0$, то $f'(x_0, p)$ есть конечная величина.*

Доказательство. Согласно второму из неравенств леммы 1.8 при $\alpha_0 = -\varepsilon$, $\alpha_1 = 0$, $\alpha_2 = \lambda$ имеем

$$\frac{f(x_0) - f(x_0 - \varepsilon p)}{\varepsilon} \leqslant \frac{f(x_0+\lambda p) - f(x_0)}{\lambda},$$

откуда, переходя к пределу по $\lambda \downarrow 0$, получаем

$$\frac{f(x_0 + \lambda p) - f(x_0)}{\lambda} \geqslant f'(x_0, p) \geqslant \frac{f(x_0) - f(x_0 - \varepsilon p)}{\varepsilon}. \quad (3.1)$$

Лемма 3.3. *$f'(x_0, p)$ есть положительно однородная выпуклая функция p.*

Доказательство. По определению для $\alpha > 0$

$$f'(x_0, \alpha p) = \lim_{\lambda \downarrow 0} \frac{f(x_0 + \lambda\alpha p) - f(x_0)}{\lambda} =$$

$$= \alpha \lim_{\lambda \downarrow 0} \frac{f(x_0 + (\lambda\alpha) p) - f(x_0)}{(\lambda\alpha)} = \alpha f'(x_0, p).$$

Далее, для $\lambda_1, \lambda_2 \geqslant 0$, $\lambda_1 + \lambda_2 = 1$, имеем

$$\frac{f(x_0 + \lambda(\lambda_1 p_1 + \lambda_2 p_2)) - f(x_0)}{\lambda} =$$

$$= \frac{f(\lambda_1 (x_0 + \lambda p_1) + \lambda_2 (x_0 + \lambda p_2)) - \lambda_1 f(x_0) - \lambda_2 f(x_0)}{\lambda} \leqslant$$

$$\leqslant \lambda_1 \frac{f(x_0 + \lambda p_1) - f(x_0)}{\lambda} + \lambda_2 \frac{f(x_0 + \lambda p_2) - f(x_0)}{\lambda}.$$

Переходя к пределу по $\lambda \downarrow 0$, получаем

$$f'(x_0, \lambda_1 p_1 + \lambda_2 p_2) \leqslant \lambda_1 f'(x_0, p_1) + \lambda_2 f'(x_0, p_2),$$

что доказывает лемму.

Определение 3.1. Вектор x^* называется *субградиентом* выпуклой собственной функции $f(x)$ в точке $x_0 \in \operatorname{dom} f$, если

$$f(x) - f(x_0) \geqslant \langle x - x_0, x^* \rangle$$

для всех x.

Данное определение субградиента нелокально, так как использует все x, а не только окрестность точки x_0. Следующая лемма показывает, что на самом деле субградиент есть локальное понятие.

Лемма 3.4. *Вектор x^* является субградиентом функции $f(x)$ в точке x_0 тогда и только тогда, когда*

$$f'(x_0, p) \geqslant \langle p, x^* \rangle \quad (3.2)$$

для всех p.

Доказательство. Если x^* — субградиент, то

$$f(x_0+\lambda p)-f(x_0)\geqslant\lambda\langle p,\ x^*\rangle,$$

откуда

$$\frac{f(x_0+\lambda p)-f(x_0)}{\lambda}\geqslant\langle p,\ x^*\rangle$$

и $f'(x_0,\ p)\geqslant\langle p,\ x^*\rangle$.

Обратно, если неравенство (3.2) выполнено, то для $0<\lambda<1$ (см. доказательство леммы 3.1)

$$f(x)-f(x_0)=\frac{f(x_0+1(x-x_0))-f(x_0)}{1}\geqslant$$

$$\geqslant\frac{f(x_0+\lambda(x-x_0))-f(x_0)}{\lambda}\geqslant f'(x_0,\ x-x_0)\geqslant$$

$$\geqslant\langle x-x_0,\ x^*\rangle,$$

т. е. x^* — субградиент.

Определение 3.2. Множество субградиентов в точке x_0 называется *субдифференциалом* и обозначается $\partial f(x_0)$. Это записывается

$$\partial f(x_0)=\{x^*:\ f(x)-f(x_0)\geqslant\langle x-x_0,\ x^*\rangle,\ \forall x\}.$$

Из леммы 3.4 следует, что

$$\partial f(x_0)=\partial_p f'(x_0,\ 0), \tag{3.3}$$

где ∂_p означает субдифференциал от $f'(x_0,\ p)$ по аргументу p. Действительно, $f'(x_0,\ 0)=0$ по определению, поэтому неравенство (3.2) может быть переписано в виде

$$f'(x_0,\ p)-f'(x_0,\ 0)\geqslant\langle p,\ x^*\rangle,$$

а это означает, что $x^*\in\partial_p f'(x_0,\ p)$.

Теорема 3.1. *Если $f'(x_0,\ p)$ есть замкнутая функция p, то $\partial f(x_0)$ не пусто и*

$$f'(x_0,\ p)=\sup_{x^*}\{\langle p,\ x^*\rangle:\ x^*\in\partial f(x_0)\}.$$

Доказательство. Так как $f'(x_0,\ p)$ положительно однородна и выпукла, то при условии замкнутости к ней может быть применена теорема 2.5. Согласно этой теореме

$$f'(x_0,\ p)=\sup_{x^*}\{\langle p,\ x^*\rangle:\ x^*\in\operatorname{dom}(f'(x_0,\ \cdot))^*\}, \tag{3.4}$$

где $(f'(x_0, \cdot))^*$ означает функцию, сопряженную к $f'(x_0, p)$ относительно p, т. е.

$$(f'(x_0, \cdot))^*(x^*) = \sup_p \{\langle p, x^*\rangle - f'(x_0, p)\}. \tag{3.5}$$

Но, как показано в § 2,

$$(f'(x_0, \cdot))^*(x^*) = \begin{cases} 0, & x^* \in \operatorname{dom}(f'(x_0, \cdot))^*, \\ +\infty, & x^* \notin \operatorname{dom}(f'(x_0, \cdot))^*; \end{cases}$$

поэтому с учетом соотношения (3.5) получаем, что $x^* \in \operatorname{dom}(f'(x_0, \cdot))^*$ тогда и только тогда, когда

$$0 \geqslant \langle p, x^*\rangle - f'(x_0, p),$$

что эквивалентно неравенству (3.2). Отсюда следует, что

$$\partial f(x_0) = \partial_p f'(x_0, 0) = \operatorname{dom}(f'(x_0, \cdot))^*.$$

Последняя формула совместно с формулой (3.4) доказывает теорему.

Теорема 3.2. *$\partial f(x_0)$ есть выпуклое замкнутое множество.*

Доказательство получается непосредственной проверкой указанных в утверждении свойств, исходя из определения 3.2.

Теорема 3.3. *$x^* \in \partial f(x_0)$ тогда и только тогда, когда*

$$\langle x_0, x^*\rangle - f(x_0) = f^*(x^*).$$

Доказательство. Действительно, по определению $x^* \in \partial f(x_0)$, если

$$\langle x_0, x^*\rangle - f(x_0) \geqslant \langle x, x^*\rangle - f(x).$$

Беря в правой части верхнюю грань, получаем

$$\langle x_0, x^*\rangle - f(x_0) \geqslant f^*(x^*).$$

Но $\langle x_0, x^*\rangle - f(x_0) \leqslant f^*(x^*)$ по лемме 2.2, поэтому

$$\langle x_0, x^*\rangle - f(x_0) = f^*(x^*).$$

Теорема 3.4. *Если выпуклая функция дифференцируема в точке x_0, то*

$$\partial f(x_0) = \{f'(x_0)\},$$

т. е. $\partial f(x_0)$ состоит из единственного вектора $f'(x_0)$ — градиента функции f.

Доказательство. Если $f(x)$ дифференцируема в точке x_0, то

$$f'(x_0, p) = \langle p, f'(x_0)\rangle,$$

и поэтому $f'(x_0, p)$ есть замкнутая функция p. По теореме 3.1 справедлива формула

$$\langle p, f'(x_0)\rangle = \sup_{x^*} \{\langle p, x^*\rangle : x^* \in \partial f(x_0)\}.$$

Легко видеть, что это возможно лишь в том случае, если множество $\partial f(x_0)$ состоит из одной точки $f'(x_0)$.

Теорема 3.5. *Пусть выпуклая функция $f(x)$ непрерывна в точке x_0. Тогда $\partial f(x_0)$ — непустое ограниченное множество и*

$$f'(x_0, p) = \max_{x^*} \{\langle p, x^*\rangle : x^* \in \partial f(x_0)\}.$$

Доказательство. Так как $f(x_0)$ непрерывна в точке x_0, то по лемме 3.2 $f'(x_0, p)$ — конечная величина. Рассмотрим в $\mathbf{R}^{n+1}$ луч

$$\mathscr{L} = \{(x^0, x) : x^0 = f(x_0) + \lambda f'(x_0, p), \quad x = x_0 + \lambda p, \quad \lambda \geqslant 0\}$$

и множество

$$A = \{(x^0, x) : x^0 > f(x)\}.$$

Множества A и $\mathscr{L}$ выпуклы и не пересекаются. Докажем это. Выпуклость A сразу следует из выпуклости $f(x)$. Из неравенства

$$\frac{f(x_0 + \lambda p) - f(x_0)}{\lambda} \geqslant f'(x_0, p)$$

вытекает, что

$$f(x_0 + \lambda p) \geqslant f(x_0) + \lambda f'(x_0, p).$$

Если же точка $(f(x_0) + \lambda f'(x_0, p),\ x_0 + \lambda p)$ луча $\mathscr{L}$ принадлежит A, то

$$f(x_0) + \lambda f'(x_0, p) > f(x_0 + \lambda p)$$

в противоречии с предыдущим неравенством.

По теореме отделимости I.2.3 существует такой не равный нулю вектор $(x^{0*}, x^*) \in \mathbf{R}^{n+1}$ что

$$\begin{gathered} y^0 \cdot x^{0*} + \langle y, x^*\rangle \leqslant x^0 \cdot x^{0*} + \langle x, x^*\rangle, \\ (y^0, y) \in \mathscr{L}, \quad (x^0, x) \in A. \end{gathered} \tag{3.6}$$

Так как $x^0 > f(x)$ для $(x^0, x) \in A$, то x^0 можно устремлять к $+\infty$. Неравенство (3.6) тогда нарушалось бы в случае $x^{0*} < 0$; поэтому $x^{0*} \geqslant 0$. Покажем, что $x^{0*} > 0$. Если $x^{0*} = 0$, то (3.6) переходит в неравенство

$$\langle y, x^* \rangle \leqslant \langle x, x^* \rangle, \qquad (y^0, y) \in \mathscr{L}, \quad x \in \operatorname{dom} f. \tag{3.7}$$

Точка $(f(x_0), x_0) \in \mathscr{L}$. С другой стороны, $x_0 + \varepsilon z \in \operatorname{dom} f$, $z \in B$, при достаточно малом $\varepsilon > 0$, так как x_0 — точка непрерывности f, и, значит, $x_0 \in \operatorname{int} \operatorname{dom} f$. Поэтому (3.7) дает

$$\langle x_0, \ x^* \rangle \leqslant \langle x_0, \ x^* \rangle + \varepsilon \langle z, \ x^* \rangle,$$

т. е.

$$\langle z, x^* \rangle \geqslant 0, \qquad z \in B,$$

где B — единичный шар с центром в нуле. Последнее возможно лишь в том случае, если $x^* = 0$, т. е. $(x^{0*}, x^*) = (0, 0)$. Получено противоречие с тем, что этот вектор ненулевой.

Итак, $x^{0*} > 0$ и можно положить $x^{0*} = 1$ (иначе следовало бы разделить обе части неравенства (3.6) на x^{0*}).

Подставим в неравенство (3.6) вместо $(y^0, y) \in \mathscr{L}$ их выражение для y_0, y, а в правой части положим $x^0 = f(x)$. Последнее возможно по непрерывности, так как неравенство (3.6) справедливо для всех $x^0 > f(x)$. Теперь оно приобретает вид

$$f(x_0) + \lambda f'(x_0, \ p) + \langle x_0 + \lambda p, \ x^* \rangle \leqslant f(x) + \langle x, \ x^* \rangle, \qquad \lambda \geqslant 0. \tag{3.8}$$

Положив $\lambda = 0$, получаем

$$\langle x - x_0, \ -x^* \rangle \leqslant f(x) - f(x_0),$$

т. е. $x_0^* = -x^* \in \partial f(x_0)$. Таким образом, $\partial f(x_0) \neq \varnothing$. С учетом этого можно переписать неравенство (3.8) в следующем виде:

$$f(x) - f(x_0) \geqslant \langle x - x_0, x_0^* \rangle + \lambda \left[f'(x_0, p) - \langle p, x_0^* \rangle \right].$$

Полагая $x = x_0$, получаем

$$f'(x_0, p) \leqslant \langle p, x_0^* \rangle, \qquad x_0^* \in \partial f(x_0). \tag{3.9}$$

По лемме 3.4

$$f'(x_0, \ p) \geqslant \langle p, \ x^* \rangle, \qquad x^* \in \partial f(x_0). \tag{3.10}$$

Совмещение неравенств (3.9) и (3.10) дает

$$f'(x_0, p) = \langle p, x_0^* \rangle = \max_{x^*} \{\langle p, x^* \rangle : x^* \in \partial f(x_0)\}. \tag{3.11}$$

Осталось доказать, что $\partial f(x_0)$ есть ограниченное множество. Согласно теореме 1.3 $f(x)$ удовлетворяет условию Липшица в точке x_0, поэтому

$$f'(x_0, p) \leqslant \frac{f(x_0 + \lambda p) - f(x_0)}{\lambda} \leqslant L \|p\|.$$

Отсюда с учетом соотношения (3.11) получаем

$$\langle p, x^* \rangle \leqslant L\|p\|$$

для всех $x^* \in \partial f(x_0)$ и любых p. Полагая $p = x^*$, получим, что

$$\|x^*\| \leqslant L.$$

Теорема полностью доказана.

Теорема 3.6. *Если* $x_0 \in \operatorname{ri} \operatorname{dom} f$, *то* $\partial f(x_0) \neq \varnothing$ *и*

$$f'(x_0, p) = \sup_{x^*} \{\langle p, x^* \rangle : x^* \in \partial f(x_0)\}.$$

Доказательство. Пусть $\mathscr{L} = \operatorname{Lin} \operatorname{dom} f$; тогда

$$f_{\mathscr{L}}(y) = f(x_0 + y), \quad y \in \mathscr{L}, \tag{3.12}$$

т. е. $f_{\mathscr{L}}(y)$ определена на подпространстве $\mathscr{L}$ (будем ее рассматривать только на этом подпространстве). Тогда точка $y = 0$ есть внутренняя точка $\operatorname{dom} f_{\mathscr{L}}$ в этом подпространстве, так как $x_0 \in \operatorname{ri} \operatorname{dom} f$.

Согласно теореме 1.4 $f_{\mathscr{L}}(y)$ непрерывна в точке $y = 0$, и поэтому по предыдущей теореме

$$f'_{\mathscr{L}}(0, p) = \max_{y^*} \{\langle p, y^* \rangle : y^* \in \partial f_{\mathscr{L}}(0)\} \tag{3.13}$$

для $p \in \mathscr{L}$. При этом $\partial f_{\mathscr{L}}(0)$ также вычисляется только относительно подпространства $\mathscr{L}$, т. е. $\partial f_{\mathscr{L}}(0) \subseteq \mathscr{L}$. Обозначим через $\mathscr{L}^{\perp}$ ортогональное дополнение $\mathscr{L}$ до всего пространства $X = \mathbf{R}^n$, т. е. $z \in \mathscr{L}^{\perp}$ тогда и только тогда, когда $\langle p, z \rangle = 0$, $p \in \mathscr{L}$. Любой вектор $p' \in \mathbf{R}^n$ можно однозначно представить в виде $p' = p + z$, $p \in \mathscr{L}$, $z \in \mathscr{L}^{\perp}$.

Пусть теперь $p \notin \mathscr{L} = \mathrm{Lin}\,\mathrm{dom}\, f$. Тогда $x_0 + \lambda p \notin \mathrm{dom}\, f$ для $\lambda > 0$, и поэтому

$$f'(x_0, p) = +\infty, \quad p \notin \mathscr{L}. \tag{3.14}$$

Если $p \in \mathscr{L}$, то

$$f'(x_0, p) = f'_{\mathscr{X}}(0, p). \tag{3.15}$$

Положим

$$A = \partial f_{\mathscr{X}}(0) + \mathscr{L}^{\perp}. \tag{3.16}$$

Тогда для любого вектора p имеем

$$p = p_0 + z, \quad p_0 \in \mathscr{L}, \quad z \in \mathscr{L}^{\perp}, \langle p_0, z\rangle = 0,$$

$$\max_{y^* \in A} \langle p, y^* \rangle =$$
$$= \max_{y_0^*, z^*} \{\langle p_0 + z, y_0^* + z^* \rangle : y_0^* \in \partial f_{\mathscr{X}}(0), z^* \in \mathscr{L}^{\perp}\} =$$
$$= \max_{y_0^*} \{\langle p_0, y_0^* \rangle : y_0^* \in \partial f_{\mathscr{X}}(0)\} + \max_{z^*} \{\langle z, z^* \rangle : z^* \in \mathscr{L}^{\perp}\} =$$
$$= \begin{cases} +\infty, & z \neq 0, \\ f'_{\mathscr{X}}(0, p_0), & z = 0, \end{cases}$$

поскольку, как легко видеть, при $z \in \mathscr{L}^{\perp}$

$$\max_{z^*} \{\langle z, z^* \rangle : z^* \in \mathscr{L}^{\perp}\} = \begin{cases} 0, & z = 0, \\ +\infty, & z \neq 0. \end{cases}$$

На основании соотношений (3.14) и (3.15), получаем

$$f'(x_0, p) = \max_{x^*} \{\langle p, x^* \rangle : x^* \in A\}. \tag{3.17}$$

Так как по теореме 3.5 $\partial f_{\mathscr{X}}(0)$ — ограниченное, а по теореме 3.2 и замкнутое выпуклое множество, то нетрудно показать, исходя из формулы (3.16), что A также выпукло и замкнуто.

Далее, из ограниченности $\partial f_{\mathscr{X}}(0)$ и формулы (3.13) следует, что $f'_{\mathscr{X}}(0, p)$ ограничена при всех $p \in \mathscr{L}$, а, значит, по теореме 1.4 она непрерывна на $\mathscr{L}$. Формула (3.15) показывает теперь, что $f'(x_0, p)$ непрерывна на $\mathscr{L}$, а в силу (3.14) $f'(x_0, p) = +\infty$ вне $\mathscr{L}$. Отсюда следует, что $f(x_0, p)$ — замкнутая функция p и можно применить теорему 3.1. Из теоремы и соотношения (3.17) следует, что выпуклые замкнутые множества A и $\partial f(x_0)$ имеют одну и ту же опорную функцию $f'(x_0, p)$. Но согласно теореме I.2.7 опорная функция полностью характеризует выпуклое замкнутое множество, так что $A = \partial f(x_0)$. Поэтому из (3.17) следует утверждение теоремы.

2. Субдифференциал суммы выпуклых функций. Операции умножения на положительную константу, суммирования и взятия максимума не выводят из класса выпуклых функций. Поэтому естественно поставить вопрос о том, как вычислить субдифференциал вновь полученной выпуклой функции, если известны субдифференциалы исходных. Нижеследующие теоремы дают ответы на этот вопрос.

Теорема 3.7. *Пусть $f(x) = \alpha f_0(x)$, где $f_0(x)$ — выпуклая функция, $\alpha > 0$. Тогда*

$$\partial f(x_0) = \alpha \partial f_0(x_0).$$

Доказательство непосредственно следует из определения субдифференциала.

Следующая теорема имеет многочисленные и важные приложения.

Теорема 3.8. *Пусть $f = f_1 + f_2$, где f_1 и f_2 — собственные выпуклые функции, и существует точка $x_1 \in \operatorname{dom} f_1 \cap \operatorname{dom} f_2$, в которой f_1 непрерывна. Тогда*

$$\partial f(x_0) = \partial f_1(x_0) + \partial f_2(x_0).$$

Доказательство. Если $x_1^* \in \partial f_1(x_0)$, $x_2^* \in \partial f_2(x_0)$, то, суммируя неравенства

$$f_1(x) - f_1(x_0) \geqslant \langle x - x_0, x_1^* \rangle,$$

$$f_2(x) - f_2(x_0) \geqslant \langle x - x_0, x_2^* \rangle,$$

получаем

$$f(x) - f(x_0) \geqslant \langle x - x_0, x_1^* + x_2^* \rangle,$$

т. е. $x_1^* + x_2^* \in \partial f(x_0)$. Таким образом, $\partial f \supseteq \partial f_1 + \partial f_2$.

Докажем обратное включение. Для упрощения выкладок предположим, что $x_0 = 0$, $f_1(0) = f_2(0) = f(0) = 0$ (этого всегда можно добиться сдвигом начала координат и вычитанием констант из функций f_1 и f_2). Пусть $x^* \in$

$\in \partial f(0)$. По определению это означает

$$f_1(x) + f_2(x) \geqslant \langle x, x^* \rangle, \tag{3.18}$$

где учтено, что $x_0 = 0$, $f_1(x_0) = f_2(x_0) = 0$.

Перепишем неравенство (3.18) в виде

$$f_1(x) - \langle x, x^* \rangle \geqslant -f_2(x). \tag{3.19}$$

Введем в $\mathbf{R}^{n+1}$ выпуклые множества

$$A = \{(x^0, x) \colon x^0 > f_1(x) - \langle x, x^* \rangle\},$$
$$B = \{(y^0, y) \colon y^0 < -f_2(y)\}.$$

Выпуклость A и B вытекает из выпуклости f_1 и f_2.

Из неравенства (3.19) следует, что A и B не пересекаются. Поэтому их можно разделить, т. е. существует такой ненулевой вектор (x^{0*}, x_0^*), что

$$y^0 x^{0*} + \langle y, x_0^* \rangle \leqslant x^0 x^{0*} + \langle x, x_0^* \rangle,$$
$$(y^0, y) \in B, \qquad (x^0, x) \in A. \tag{3.20}$$

Так как x^0 может увеличиваться неограниченно, то из (3.20) следует, что $x^{0*} \geqslant 0$. Покажем, что $x^{0*} > 0$. Действительно, если $x^{0*} = 0$, то неравенство (3.20) можно переписать в виде

$$\langle y, x_0^* \rangle \leqslant \langle x, x_0^* \rangle, \qquad y \in \operatorname{dom} f_2, \quad x \in \operatorname{dom} f_1. \tag{3.21}$$

Положим $y = x_1$. Далее, так как $x_1 \in \operatorname{dom} f_1$ — точка непрерывности f_1, то $x_1 + \varepsilon z \in \operatorname{dom} f_1$, $\|z\| \leqslant 1$, при достаточно малом $\varepsilon > 0$. Подставляя $y = x_1$, $x = x_1 + \varepsilon z$ в неравенство (3.21), получаем

$$0 \leqslant \varepsilon \langle z, x_0^* \rangle, \qquad \|z\| \leqslant 1,$$

что невозможно, если $x_0^* \neq 0$. Итак, $x^{0*} = 0$, $x_0^* = 0$ в противоречии с тем, что вектор (x^{0*}, x_0^*)не равен нулю; поэтому $x^{0*} > 0$ и можно положить $x^{0*} = 1$.

Теперь с учетом определения множеств A и B из соотношения (3.20) получаем

$$-f_2(y) + \langle y, x_0^* \rangle \leqslant f_1(x) - \langle x, x^* \rangle + \langle x, x_0^* \rangle, \tag{3.22}$$
$$y \in \operatorname{dom} f_2, \qquad x \in \operatorname{dom} f_1.$$

Так как вне $\operatorname{dom} f_2$ и $\operatorname{dom} f_1$ значения функций f_2 и f_1 бесконечны, то можно считать, что неравенство (3.22) справедливо при всех y и x. Положив $y = x_0 = 0$, получаем

$$\langle x, x^* - x_0^* \rangle \leqslant f_1(x),$$

т. е. $(x^* - x_0^*) \in \partial f_1(0)$. Положив $x = x_0 = 0$, получаем

$$\langle y, x_0^* \rangle \leqslant f_2(y),$$

т. е. $x_0^* \in \partial f_2(0)$. Поэтому

$$x^* = (x^* - x_0^*) + x_0^* \in \partial f_1(0) + \partial f_2(0).$$

Так как x^* — произвольный элемент $\partial f(0)$, то

$$\partial f(0) \subseteq \partial f_1(0) + \partial f_2(0).$$

Учитывая ранее доказанное противоположное включение, окончательно получаем

$$\partial f(0) = \partial f_1(0) + \partial f_2(0),$$

что и требовалось доказать.

Теорема 3.9. *Пусть f_1 и f_2 — собственные выпуклые функции и* $\mathrm{ri\,dom}\, f_1 \cap \mathrm{ri\,dom}\, f_2 \neq \varnothing$. *Тогда*

$$\partial f(x_0) = \partial f_1(x_0) + \partial f_2(x_0),$$

где $f = f_1 + f_2$, $x_0 \in \mathrm{dom}\, f_1 \cap \mathrm{dom}\, f_2$.

Доказательство. Как и ранее, без ограничения общности предполагаем, что $x_0 = 0$, $f_1(0) = f_2(0) = 0$. Тогда

$$\mathrm{dom}\, f_1 \subseteq \mathrm{Lin\,dom}\, f_1, \quad \mathrm{dom}\, f_2 \subseteq \mathrm{Lin\,dom}\, f_2.$$

Пусть $x_1 \in \mathrm{ri\,dom}\, f_1 \cap \mathrm{ri\,dom}\, f_2$. По теореме I.1.4

$$\mathrm{Lin\,dom}\, f = \mathrm{Lin\,dom}\, f_1 \cap \mathrm{Lin\,dom}\, f_2.$$

Обозначим $\mathscr{L} = \mathrm{Lin\,dom}\, f_1 + \mathrm{Lin\,dom}\, f_2$. Тогда $\mathrm{dom}\, f_1$, $\mathrm{dom}\, f_2$ и $\mathrm{dom}\, f$ лежат в $\mathscr{L}$, а надграфики всех рассматриваемых функций лежат в $\mathbf{R}^1 \times \mathscr{L}$. Поэтому можно все рассмотрения ввести в $\mathscr{L}$ или $\mathbf{R}^1 \times \mathscr{L}$.

Повторяя рассуждения предыдущей теоремы с учетом того, что множества A и B лежат в $\mathbf{R}^1 \times \mathscr{L}$, получим, что A и B можно разделить в $\mathbf{R}^1 \times \mathscr{L}$, т. е. найдется такой ненулевой вектор (x^{0*}, x_0^*), что $x_0^* \in \mathscr{L}$ и выполнено неравенство (3.20).

Покажем, что $x^{0*} > 0$. То, что x^{0*} не может быть меньше нуля, доказывается, как в предыдущей теореме. Пусть $x^{0*} = 0$. Тогда справедливо неравенство (3.21). По предположению

$$x_1 + \varepsilon z_1 \in \operatorname{dom} f_1, \quad z_1 \in \operatorname{Lin} \operatorname{dom} f_1, \quad \|z_1\| \leqslant 1,$$

$$x_1 - \varepsilon z_2 \in \operatorname{dom} f_2, \quad z_2 \in \operatorname{Lin} \operatorname{dom} f_2, \quad \|z_2\| \leqslant 1,$$

при некотором $\varepsilon > 0$. Поэтому из соотношения (3.21) получаем

$$\langle x_1 - \varepsilon z_2, x_0^* \rangle \leqslant \langle x_1 + \varepsilon z_1, x_0^* \rangle,$$

$$z_1 \in \operatorname{Lin} \operatorname{dom} f_1, \quad z_2 \in \operatorname{Lin} \operatorname{dom} f_2, \quad \|z_1\| \leqslant 1, \quad \|z_2\| \leqslant 1,$$

или

$$\begin{gathered} \langle z_1 + z_2, x_0^* \rangle \geqslant 0, \\ z_i \in \operatorname{Lin} \operatorname{dom} f_i, \qquad \| z_i \| \leqslant 1, \quad i = 1, 2. \end{gathered} \tag{3.23}$$

Так как для любого $z \in \mathscr{L}$, $z = \bar{z}_1 + \bar{z}_2$, $\bar{z}_i \in \operatorname{Lin} \operatorname{dom} f_i$, $i = 1, 2$, и при достаточно малом $\lambda > 0$ $\|\lambda \bar{z}_1\| \leqslant 1$, $\|\lambda \bar{z}_2\| \leqslant 1$, то, подставляя $\lambda z = \lambda \bar{z}_1 + \lambda \bar{z}_2$ в неравенство (3.23), получаем

$$\langle z, x_0^* \rangle \geqslant 0$$

при всех $z \in \mathscr{L}$. Но так как $x_0^* \in \mathscr{L}$, то это возможно лишь, если $x_0^* = 0$. А это противоречит тому, что $(x^{0*}, x_0^*) \neq 0$. Итак, как и в предыдущей теореме, можно положить $x_0^* = 1$. Повторение доказательства предыдущей теоремы завершает доказательство теоремы 3.9.

Проиллюстрируем, насколько содержательны теоремы 3.8 и 3.9, на важном примере индикаторной функции множества. Пусть M — выпуклое множество, а $\delta(x|M)$ — его индикаторная функция. Вычислим субдифференциал $\partial\delta(x_0|M)$, $x_0 \in M$. По определению $x^* \in \partial\delta(x_0|M)$ тогда и только тогда, когда

$$\delta(x|M) - \delta(x_0|M) \geqslant \langle x - x_0, \quad x^* \rangle.$$

Но $\delta(x_0|M) = 0$. Если $x \notin M$, то $\delta(xM) = +\infty$ и неравенство выполняется всегда. Для $x \in M$ получаем

$$0 \geqslant \langle x - x_0, x^* \rangle. \tag{3.24}$$

Пусть

$$\operatorname{con} (M - x_0) = \{p\colon \ p = \lambda(x - x_0), \quad \lambda > 0, \quad x \in M\}.$$

Тогда неравенство (3.24) эквивалентно тому, что

$$-x^* \in (\mathrm{con}\,(M - x_0))^*.$$

Таким образом,

$$\partial\delta(x_0|M) = -(\mathrm{con}\,(M - x_0))^*. \tag{3.25}$$

Если теперь K_i, $i = 1, \ldots, m$,— выпуклые конусы и $0 \in K_i$, то $\mathrm{con}\,(K_i - 0) = K_i$, и поэтому

$$\partial\delta(0\,|\,K_i) = -K_i^*. \tag{3.26}$$

Заметим теперь, что

$$\delta\left(x\,\middle|\,\bigcap_{i=1}^{m} K_i\right) = \delta(x\,|\,K_1) + \ldots + \delta(x\,|\,K_m). \tag{3.27}$$

Если теперь применить к $\delta\left(x\,\middle|\,\bigcap_{i=1}^{m} K_i\right)$ теорему 3.8, то получится результат, соответствующий теореме I.3.2.

Теорема 3.10. *Пусть* $\mathrm{ri}\,K_1 \cap \ldots \cap \mathrm{ri}\,K_m \neq \varnothing$. *Тогда*

$$(K_1 \cap \ldots \cap K_m)^* = K_1^* + \ldots + K_m^*.$$

Доказательство. Очевидно, что $\mathrm{dom}\,\delta(\cdot|K_i) = K_i$, поэтому условия теоремы показывают, что

$$\mathrm{ri\,dom}\,\delta(\cdot|K_1) \cap \ldots \cap \mathrm{ri\,dom}\,\delta(\cdot|K_m) \neq \varnothing,$$

и можно применить теорему 3.9 к функции (3.27). С учетом равенства (3.26) получаем

$$-\left(\bigcap_{i=1}^{m} K_i\right)^* = \partial\delta\left(0\,\middle|\,\bigcap_{i=1}^{m} K_i\right) =$$

$$= \partial\delta(0\,|\,K_1) + \ldots + \partial\delta(0\,|\,K_m) = -K_1^* - \ldots - K_m^*,$$

что и требовалось доказать.

3. Субдифференциал максимума выпуклых функций. Перейдем теперь к изучению выпуклых функций, полученных в результате взятия максимума по параметру.

Теорема 3.11. *Пусть M^* — выпуклое замкнутое множество и*

$$f(x) = \sup_{x^*}\{\langle x, x^*\rangle:\ x^* \in M^*\}.$$

Тогда

$$\partial f(x_0) = \{x^* \in M^*:\ \langle x_0, x^*\rangle = f(x_0)\}.$$

В частности, если $x_0 = 0$, то $\partial f(0) = M^*$.

Доказательство. Если $x^* \in M^*$, $\langle x_0, x^*\rangle = f(x_0)$, то

$$f(x) - f(x_0) \geqslant \langle x, x^*\rangle - \langle x_0, x^*\rangle = \langle x - x_0, x^*\rangle,$$

т. е. $x^* \in \partial f(x_0)$.

Обратно, пусть $x_0^* \in \partial f(x_0)$. Допустим, что $x_0^* \notin M^*$. Тогда по теореме I.2.1 существует такой вектор p, что

$$\sup_{x^*} \{\langle p, x^*\rangle :\ x^* \in M^*\} < \langle p, x_0^*\rangle. \tag{3.28}$$

С другой стороны, так как максимум разности не меньше, чем разность максимумов, то

$$\sup_{x^*} \{\langle x - x_0, x^*\rangle :\ x^* \in M^*\} \geqslant f(x) - f(x_0) \geqslant \langle x - x_0, x_0^*\rangle.$$

Полагая в этом неравенстве $x = x_0 + p$, получаем

$$\sup_{x^*} \{\langle p, x^*\rangle :\ x^* \in M^*\} \geqslant \langle p, x_0^*\rangle$$

в противоречии с неравенством (3.28). Значит, $x_0^* \in M^*$. Так как по предположению $x_0^* \in \partial f(x_0)$, то

$$f(x) - \langle x, x_0^*\rangle \geqslant f(x_0) - \langle x_0, x_0^*\rangle.$$

Полагая $x = 0$, получаем $\langle x_0, x_0^*\rangle \geqslant f(x_0)$. Но $x_0^* \in M^*$, поэтому $f(x_0) \geqslant \langle x_0, x_0^*\rangle$. Из двух полученных неравенств следует

$$\langle x_0, x_0^*\rangle = f(x_0),$$

что и требовалось доказать.

Если $x_0 = 0$, то $f(0) = \langle 0, x^*\rangle = 0$ для всех $x^* \in M^*$, и поэтому $\partial f(0) = M^*$.

Теорема 3.12. *Пусть $f(x)$ — положительно однородная выпуклая замкнутая функция. Тогда*

$$\partial f(x_0) = \{x^* \in \operatorname{dom} f^* :\ \langle x_0, x^*\rangle = f(x_0)\}.$$

Доказательство следует из теорем 2.5 и 3.11.

Пусть теперь A — некоторое множество индексов и $f(x, \alpha)$ — выпуклые при каждом α функции от x. Как показано ранее,

$$f(x) = \sup_{\alpha} \{f(x, \alpha) :\ \alpha \in A\}$$

есть снова выпуклая функция. Чтобы ответить на вопрос, как выражается субдифференциал $f(x)$ через субдифференциалы $f(x, \alpha)$, предварительно докажем две леммы.

Лемма 3.5. *Пусть A — компакт и функция $f(x, \alpha)$ непрерывна по x и α в окрестности точки x_0 и $\alpha \in A$. Тогда функция*

$$f(x) = \max_{\alpha} \{f(x, \alpha):\ \alpha \in A\} \tag{3.29}$$

непрерывна в точке x_0.

Доказательство. Обозначим

$$A(x) = \{\alpha \in A:\ f(x, \alpha) = f(x)\}.$$

Тогда

$$f(x, \alpha) = f(x) \geqslant f(x, \alpha_0),$$
$$f(x_0, \alpha) \leqslant f(x_0) = f(x_0, \alpha_0)$$

при $\alpha \in A(x)$, $\alpha_0 \in A(x_0)$. Вычитая из первого неравенства второе, получим

$$f(x, \alpha) - f(x_0, \alpha) \geqslant f(x) - f(x_0) \geqslant f(x, \alpha_0) - f(x_0, \alpha_0). \tag{3.30}$$

При $x \to x_0$ выражение в правой части стремится к нулю. Покажем, что и левая часть стремится к нулю. Действительно, пусть для некоторой последовательности $x_k \to x_0$ и $\alpha_k \in A(x_k) \subseteq A$

$$|f(x_k, \alpha_k) - f(x_0, \alpha_k)| \geqslant \varepsilon > 0.$$

Так как A — компакт, то можно считать, что $\alpha_k \to \alpha \in A$. Но тогда

$$f(x_k, \alpha_k) - f(x_0, \alpha_k) \to f(x_0, \alpha) - f(x_0, \alpha) = 0,$$

что противоречит допущению.

Итак, при $x \to x_0$ левая и правая части неравенства (3.30) стремятся к нулю, а поэтому $f(x) \to f(x_0)$. Лемма доказана.

Лемма 3.6. *Пусть выполнены предположения леммы 3.5. Тогда для любого открытого множества C такого, что $A\ (x_0) \subseteq C \subseteq A$, существует окрестность точки x_0, для любой точки x из которой выполняется включение*

$$A(x) \subseteq C.$$

Доказательство. Пусть для некоторого открытого множества C_0, $A(x_0) \subseteq C_0$, существует такая последовательность $x_k \to x_0$, что найдутся точки $\alpha_k \in A(x_k)$, $\alpha_k \notin C_0$. Так как A — компакт, то можно считать, что $\alpha_k \to \alpha_0$.

По определению

$$f(x_k) = f(x_k, \alpha_k).$$

Переходя к пределу, согласно лемме 3.5 получаем

$$f(x_0) = f(x_0, \alpha_0),$$

т. е. $\alpha_0 \in A(x_0) \subseteq C_0$. Тогда α_k при достаточно больших k должно принадлежать открытому множеству C_0 — окрестности α_0 — по определению предела. Последнее противоречит тому, что $\alpha_k \notin C_0$. Лемма доказана.

Теорема 3.13. *Пусть A — компакт, $f(x, \alpha)$ — выпуклая функция при каждом $\alpha \in A$. Пусть, кроме того, $f(x + \lambda p, \alpha)$ непрерывна по λ и α для $\lambda \in [-\delta, \delta]$, $\delta > 0$, $\alpha \in A$. Тогда функция $f(x)$, определяемая* (3.29), *дифференцируема по направлению p и*

$$f'(x_0, p) = \max_{\alpha} \{f'(x_0, p, \alpha) : \alpha \in A(x_0)\},$$

где $f'(x_0, p, \alpha)$ есть производная $f(x, \alpha)$ по направлению p в точке x_0.

Доказательство. Положим для простоты

$$g(\lambda) = f(x_0 + \lambda p), \qquad g(\lambda, \alpha) = f(x_0 + \lambda p, \alpha),$$
$$A(\lambda) = A(x_0 + \lambda p).$$

Ясно, что

$$f'(x_0, p) = g'(0), \qquad f'(x_0, p, \alpha) = g'(0, \alpha),$$

где

$$g'(0) = \lim_{\lambda \downarrow 0} \frac{g(\lambda) - g(0)}{\lambda},$$

$$g'(0, \alpha) = \lim_{\lambda \downarrow 0} \frac{g(\lambda, \alpha) - g(0, \alpha)}{\lambda}.$$

Функция

$$g(\lambda) = \max_{\alpha} \{g(\lambda, \alpha) : \alpha \in A\}$$

конечна при $\lambda \in [-\delta, \delta]$, ибо $g(\lambda, \alpha)$ непрерывна по α, и A — компакт. Поэтому производная $g'(0)$ существует и ограничена. Положив в неравенства (3.30) $x=x_0+\lambda p$, $\lambda>0$, получим

$$\frac{g(\lambda,\alpha)-g(0,\alpha)}{\lambda}\geqslant\frac{g(\lambda)-g(0)}{\lambda}\geqslant\frac{g(\lambda,\alpha_0)-g(0,\alpha_0)}{\lambda},\qquad(3.31)$$
$$\alpha\in A(\lambda),\quad \alpha_0\in A(0).$$

Так как разностное отношение для выпуклой функции стремится к нулю, монотонно убывая, то из правого неравенства (3.31) получаем

$$\frac{g(\lambda)-g(0)}{\lambda}\geqslant g'(0,\alpha_0),\qquad \alpha_0\in A(0).$$

Поэтому

$$g'(0)\geqslant\sup_{\alpha}\{g'(0,\alpha):\ \alpha\in A(0)\}.$$

Обратимся теперь к левому неравенству (3.31). По лемме 3.6 для любого открытого множества $C\supseteq A(x_0)=$ $=A(0)$ найдется такое $\varepsilon>0$, что $A(\lambda)\subseteq C$ при $0\leqslant\lambda<\varepsilon$. Поэтому для $\lambda<\varepsilon$ из неравенства (3.31) получаем

$$g'(0)\leqslant\frac{g(\lambda)-g(0)}{\lambda}\leqslant\sup_{\alpha}\left\{\frac{g(\lambda,\alpha)-g(0,\alpha)}{\lambda}:\ \alpha\in C\right\}.\qquad(3.32)$$

Возьмем точную нижнюю грань выражения, стоящего в правой части соотношения (3.32), по всем $C\supseteq A(0)$. Так как $g(\lambda,\alpha)$ непрерывна по α и $A(0)$ — замкнутое подмножество компактного множества, то нетрудно убедиться, что

$$\inf_{C\supseteq A(0)}\sup_{\alpha}\left\{\frac{g(\lambda,\alpha)-g(0,\alpha)}{\lambda}:\ \alpha\in C\right\}=$$
$$=\max_{\alpha}\left\{\frac{g(\lambda,\alpha)-g(0,\alpha)}{\lambda}:\ \alpha\in A(0)\right\}.$$

Поэтому

$$g'(0)\leqslant\max_{\alpha}\left\{\frac{g(\lambda,\alpha)-g(0,\alpha)}{\lambda}:\ \alpha\in A(0)\right\}.\qquad(3.33)$$

Пусть $\lambda_k\to0$ и $\alpha_k\in A(0)$ такие, что

$$\frac{g(\lambda_k,\alpha_k)-g(0,\alpha_k)}{\lambda_k}=\max_{\alpha}\left\{\frac{g(\lambda_k,\alpha)-g(0,\alpha)}{\lambda_k}:\ \alpha\in A(0)\right\}.\qquad(3.34)$$

Без потери общности можно считать, что $\alpha_k \to \alpha_0 \in A(0)$. Далее, по лемме 1.8

$$\frac{g(0, \alpha_k) - g(-\delta, \alpha_k)}{\delta} \leqslant \frac{g(\lambda_k, \alpha_k) - g(0, \alpha_k)}{\lambda_k} \leqslant \frac{g(\delta, \alpha_k) - g(0, \alpha_k)}{\delta},$$

так что отношение $\lambda_k^{-1}[g(\lambda_k, \alpha_k) - g(0, \alpha_k)]$ в силу непрерывности $g(\lambda, \alpha)$ и компактности A ограничено. Можно считать, что оно стремится к некоторой величине μ.

Пусть $\lambda > 0$ фиксировано. Тогда $\lambda_k < \lambda$ при больших k и по лемме 1.8

$$\frac{g(\lambda, \alpha_k) - g(0, \alpha_k)}{\lambda} \geqslant \frac{g(\lambda_k, \alpha_k) - g(0, \alpha_k)}{\lambda_k}.$$

Переходя к пределу при $k \to \infty$, имеем

$$\frac{g(\lambda, \alpha_0) - g(0, \alpha_0)}{\lambda} \geqslant \mu, \tag{3.35}$$

откуда при $\lambda \to 0$ получаем, что $g'(0, \alpha_0) \geqslant \mu$. Используя теперь неравенства (3.33) и (3.34), получим

$$g'(0) \leqslant \mu \leqslant g'(0, \alpha_0), \qquad \alpha_0 \in A(0).$$

Совмещая это неравенство с ранее доказанным, получим

$$g'(0, \alpha_0) \geqslant g'(0) \geqslant \sup_\alpha \{g'(0, \alpha)\colon \alpha \in A(0)\}, \qquad \alpha_0 \in A(0),$$

т. е.

$$g'(0) = \max_\alpha \{g'(0, \alpha)\colon \alpha \in A(0)\}.$$

Последнее соотношение эквивалентно утверждению теоремы.

Теорема 3.14. *Пусть A — компакт, $f(x, \alpha)$ выпукла по x при $\alpha \in A$ и непрерывна по x и α для x из некоторой окрестности x_0 и $\alpha \in A$. Тогда*

$$\partial f(x_0) = \overline{\mathrm{co}}\left(\bigcup_{\alpha \in A(x_0)} \partial f(x_0, \alpha)\right).$$

Доказательство. По теореме 3.5

$$f'(x_0, p, \alpha) = \max_{x^*} \{\langle p, x^* \rangle\colon x^* \in \partial f(x_0, \alpha)\}.$$

Используя предыдущую теорему, получаем

$$f'(x_0, p) = \max_{\alpha \in A(x_0)} \max_{x^* \in \partial f(x_0, \alpha)} \langle p, x^* \rangle =$$

$$= \max_{x^*} \left\{ \langle p, x^* \rangle : \ x^* \in \bigcup_{\alpha \in A(x_0)} \partial f(x_0, \alpha) \right\} =$$

$$= \sup_{x^*} \left\{ \langle p, x^* \rangle : \ x^* \in \overline{\mathrm{co}} \bigcup_{\alpha \in A(x_0)} \partial f(x_0, \alpha) \right\}. \qquad (3.36)$$

При выводе этих соотношений было учтено, что точные верхние грани линейной функции на множестве и на его выпуклой оболочке совпадают и что верхняя грань линейной функции на множестве совпадает с верхней гранью на его замыкании.

Снова используя теорему 3.5, получаем

$$f'(x_0, p) = \sup_{x^*} \{ \langle p, x^* \rangle : \ x^* \in \partial f(x_0) \}. \qquad (3.37)$$

Из соотношений (3.36) и (3.37) следует, что опорные функции двух выпуклых замкнутых множеств совпадают, а поэтому совпадают и сами множества. Теорема доказана.

4. Субдифференциал расстояния до множества. Пусть M — выпуклое множество, C — выпуклое ограниченное множество, $0 \in \mathrm{int}\, C$. Тогда для всех $x \in X$ определена функция

$$d_C(x \mid M) = \inf_{\rho} \{ \rho \geqslant 0 : \ x \in M + \rho C \}. \qquad (3.38)$$

Если $C = B$, где B — единичный шар с центром в нуле, то, как легко видеть, $d_B(x|M)$ есть просто расстояние от точки x до множества M. Если $M = \{0\}$, т. е. состоит из единственной точки — начала координат, то

$$d_C(x \mid \{0\}) = \inf_{\rho} \{ \rho \geqslant 0 : \ x \in \rho C \} =$$

$$= \inf_{\rho} \left\{ \rho \geqslant 0 : \ \frac{x}{\rho} \in C \right\} = r_C(x),$$

где согласно определению I.2.1 $r_C(x)$ — функция Минковского множества C.

Теорема 3.15. *Справедлива формула*

$$d_C(x \mid M) = \sup_{x^*} \{ \langle x, x^* \rangle - W_M(x^*) : \ W_C(x^*) \leqslant 1 \}.$$

Доказательство. Так как $0 \in \operatorname{int} C$, то множества $M+\rho C$, $\rho \geqslant 0$, вложены друг в друга и нетрудно убедиться, что если $x \in \overline{M+\rho C}$, то $d_C(x|M) \leqslant \rho$. По теореме I.2.7 $x \in \overline{M+\rho C}$ тогда и только тогда, когда

$$\langle x, \ x^* \rangle \leqslant W_{M+\rho C}(x^*) = W_M(x^*) + \rho W_C(x^*) \tag{3.39}$$

при всех x^*. Если $x \in \overline{M}$, то $d_C(x|M) = 0$ и

$$\langle x, \ x^* \rangle \leqslant W_M(x^*).$$

Поэтому

$$\sup_{x^*} \{\langle x, x^* \rangle - W_M(x^*) \colon \ W_C(x^*) \leqslant 1\} = 0,$$

причем точная верхняя грань в левой части последнего соотношения достигается при $x^* = 0$. Таким образом, если $d_C(x|M) = 0$, то утверждение теоремы справедливо.

Пусть теперь $x \in \overline{M+\rho C}$ и $\rho \geqslant d_C(x|M) > 0$. При $x^* \neq 0$ имеем $W_C(x^*) > 0$, так как $0 \in \operatorname{int} C$. Поскольку относительно x^* неравенство (3.39) положительно однородно, то после перенормировки его можно переписать в виде

$$\langle x, \ x^* \rangle \leqslant W_M(x^*) + \rho, \qquad W_C(x^*) = 1.$$

Поэтому

$$\rho \geqslant \sup_{x^*} \{\langle x, x^* \rangle - W_M(x^*) \colon \ W_C(x^*) = 1\}. \tag{3.40}$$

Пусть теперь ρ удовлетворяет неравенству (3.40), так что $x \in \overline{M+\rho C}$. Покажем, что $\rho \geqslant d_C(x|M)$. Так как $0 \in \operatorname{int} C$, то для любого $\varepsilon > 0$ выполнено соотношение $(x - \varepsilon C) \cap \cap (M + \rho C) \neq \varnothing$, т. е. $x \in (M + (\rho + \varepsilon) C)$, и, значит, $d_C(x|M) \leqslant \rho + \varepsilon$. Ввиду произвольности ε отсюда следует, что $d_C(x|M) \leqslant \rho$. Если же ρ не удовлетворяет неравенству (3.40), то $(M+\rho C) \cap (x - \varepsilon_0 C) = \varnothing$ при некотором $\varepsilon_0 > 0$ и $x \notin (M + (\rho - \varepsilon_0) C)$, т. е. $\rho + \varepsilon_0 \leqslant d_C(x|M)$. Отсюда вытекает, что $d_C(x|M)$ совпадает с правой частью неравенства (3.40).

Из теоремы 3.15 сразу следует, что $d_C(x|M)$ — выпуклая функция x, ибо она есть точная верхняя грань линейных функций. Используя определение (3.38) и то, что $0 \in \operatorname{int} C$, легко показать, что $d_C(x|M)$ ограничена в любой ограниченной области и поэтому непрерывна (и даже липшицева) по теоремам 1.2 и 1.3.

Вычислим субдифференциал $\partial d_C(x_0|M)$.

Теорема 3.16. *Справедлива формула*

$$\partial d_C(x_0|M)=\begin{cases}\partial\delta(x_0|\overline{M})\cap\{x^*:\ W_C(x^*)\leqslant 1\}, \\ \qquad \textit{если}\ d_C(x_0|M)=0, \\ \partial\delta(x_0|\overline{M+d_C(x_0\ M)C})\cap\{x^*:\ W_C(x^*)=1\}, \\ \qquad \textit{если}\ d(x_0|M)>0.\end{cases}$$

Доказательство. По определению $x^*\in\partial d_C(x_0|M)$ тогда и только тогда, когда

$$d_C(x|M)-d_C(x_0|M)\geqslant\langle x-x_0,\ x^*\rangle \tag{3.41}$$

при всех x. Любая точка x представима в виде $y+\rho z$ при некоторых $y\in M$, $z\in C$, $\rho\geqslant 0$, причем $d_C(x|M)$ есть нижняя грань таких ρ. Поэтому неравенство (3.41) эквивалентно неравенству

$$\rho-d_C(x_0|M)\geqslant\langle y+\rho z-x_0,\ x^*\rangle,\ \ \rho\geqslant 0,\ \ y\in M,\ \ z\in C,$$

или

$$\rho-d_C(x_0|M)\geqslant W_M(x^*)+\rho W_C(x^*)-\langle x_0,\ x^*\rangle. \tag{3.42}$$

Неравенство (3.42) справедливо для любых $\rho\geqslant 0$ только в том случае, если

$$W_C(x^*)\leqslant 1. \tag{3.43}$$

Тогда его можно переписать в эквивалентной форме:

$$-d_C(x_0|M)\geqslant W_M(x^*)-\langle x_0,\ x^*\rangle. \tag{3.44}$$

Если $d_C(x_0|M)=0$, то $x_0\in\overline{M}$, а неравенство (3.44) можно записать в виде

$$0\geqslant\langle x-x_0,\ x^*\rangle,\qquad x\in\overline{M},$$

т. е. $-x^*\in[\operatorname{con}(\overline{M}-x_0)]^*=-\partial\delta(x_0|\overline{M})$ (см. формулу (3.25)). Совмещая это с неравенством (3.44), получаем первое утверждение теоремы.

Пусть теперь $d_C(x_0|M)>0$. Тогда $x_0\in\overline{M+\rho_0 C}$, $\rho_0=d_C(x_0|M)$ и неравенство (3.39) справедливо при $x=x_0$, $\rho=\rho_0$:

$$-d_C(x_0|M)W_C(x^*)\leqslant W_M(x^*)-\langle x_0,\ x^*\rangle.$$

Вычитая его из неравенства (3.44), получаем

$$d_C(x_0|M)(W_C(x^*)-1)\geqslant 0.$$

С учетом условия (3.43) это дает $W_C(x^*)=1$. Поэтому неравенство (3.44) можно переписать в виде

$$0 \geqslant W_M(x^*) + d_C(x_0|M)W_C(x^*) - \langle x_0, x^*\rangle$$

или

$$0 \geqslant \langle x - x_0, x^*\rangle, \quad x \in \overline{M + d_C(x_0|M)C},$$

откуда

$$-x^* \in [\mathrm{con}\,(\overline{M + d_C(x_0|M)C} - x_0)]^* = \\ = -\partial\delta(x_0|\overline{M + d_C(x_0|M)C}),$$

что завершает доказательство.

5. Конус допустимых направлений и субдифференциалы. На предыдущих страницах уже неоднократно встречался конус

$$\mathrm{con}\,(M - x_0) = \{p\colon\ p = \lambda(x - x_0), \quad \lambda > 0, \quad x \in M\},$$

определенный для выпуклого множества M и точки $x_0 \in \in M$. В экстремальных задачах конусы такого вида будут возникать постоянно, а вычисление сопряженных к ним будет одной из основных задач. Покажем, как вычисляется $[\mathrm{con}\,(M - x_0)]^*$ в случае, когда множество M задается при помощи неравенства для выпуклой функции.

Теорема 3.17. *Пусть $f(x)$ — выпуклая функция,*

$$M = \{x\colon\ f(x) \leqslant 0\}$$

и существует точка $x_1 \in M$, в которой $f(x_1) < 0$. Если $f(x_0) = 0$ и $f'(x_0, p)$ есть замкнутая функция, то

$$[\mathrm{con}\,(M - x_0)]^* = -\mathrm{con}\,\partial f(x_0).$$

Доказательство. Так как

$$0 > f(x_1) = f(x_1) - f(x_0) \geqslant \langle x_1 - x_0, x^*\rangle, \qquad x^* \in \partial f(x_0),$$

то нуль не принадлежит выпуклому замкнутому множеству $\partial f(x_0)$. Покажем, что

$$\mathrm{con}\,\partial f(x_0) = \{x^*\colon\ x^* = \lambda x_0^*, \qquad \lambda \geqslant 0, \quad x_0^* \in \partial f(x_0)\}$$

есть замкнутое множество. В самом деле, пусть последовательности λ_k и x_k^* таковы, что $\lambda_k x_k^* \to x^*$, $\lambda_k > 0$, $x_k^* \in \partial f(x_0)$, $x^* \neq 0$. Тогда последовательность λ_k ограничена, ибо если предположить противное (т. е. что $\lambda_k \to +\infty$), то

$\lambda_k \| x_k^* \| \to \| x^* \|$, и, значит, $\| x_k^* \| \to 0$, чего не может быть, так как $0 \notin \partial f(x_0)$. В силу ограниченности λ_k можно считать, что $\lambda_k \to \lambda_0 > 0$. Но тогда последовательность x_k^* сходится к $\lambda_0^{-1} x^*$, и так как $\partial f(x_0)$ — замкнутое множество, то $\lambda_0^{-1} x^* \in \partial f(x_0)$. Значит, $x^* \in \operatorname{con} \partial f(x_0)$, что доказывает замкнутость $\operatorname{con} \partial f(x_0)$.

По определению $p \in [-\operatorname{con} \partial f(x_0)]^*$ тогда и только тогда, когда

$$\langle p, -\lambda x_0^* \rangle \geqslant 0, \qquad \lambda > 0, \qquad x_0^* \in \partial f(x_0),$$

т. е.

$$\sup_{x_0^*} \{\langle p, x_0^* \rangle :\ x_0^* \in \partial f(x_0)\} \leqslant 0. \tag{3.45}$$

Так как по предположению $f'(x_0, p)$ — замкнутая функция, то, используя теорему 3.1 и неравенство (3.45), можно заключить, что

$$[-\operatorname{con} \partial f(x_0)]^* = \{p:\ f'(x_0, p) \leqslant 0\}.$$

Положим $p_1 = x_1 - x_0$. Тогда для $0 < \lambda < 1$

$$f'(x_0, p_1) \leqslant \frac{f(x_0 + \lambda p_1) - f(x_0)}{\lambda} \leqslant f(x_1) - f(x_0) < 0$$

и в силу выпуклости $f'(x_0, p)$ по p

$$f'(x_0, \lambda p_1 + (1 - \lambda) p) \leqslant \lambda f'(x_0, p_1) + (1 - \lambda) f'(x_0, p) < 0,$$

если $f'(x_0, p) \leqslant 0$.

Пусть теперь $p = \lambda(x - x_0)$, $x \in M$, $\lambda > 0$, т. е. $p \in \operatorname{con}(M - x_0)$. Тогда $f(x) \leqslant 0$ и

$$0 \geqslant f(x) - f(x_0) \geqslant \langle x - x_0,\ x^* \rangle, \quad x^* \in \partial f(x_0).$$

Поэтому

$$\sup_{x^*} \{\langle p, x^* \rangle :\ x^* \in \partial f(x_0)\} =$$

$$= \frac{1}{\lambda} \sup_{x^*} \{\langle x - x_0, x^* \rangle :\ x^* \in \partial f(x_0)\} \leqslant 0.$$

Отсюда следует, что

$$\operatorname{con}(M - x_0) \subseteq \{p:\ f'(x_0, p) \leqslant 0\}.$$

С другой стороны, если $f'(x_0, p) < 0$, то для малых $\lambda > 0$

$$f(x_0 + \lambda p) = f(x_0) + \lambda f'(x_0, p) + o(\lambda) = \lambda f'(x_0, p) + o(\lambda) < 0,$$

и, значит, $x_0 + \lambda p \in M$. Это означает, что $p \in \operatorname{con}(M - x_0)$. Теперь из неравенства (3.46) следует

$$\lambda p_1 + (1 - \lambda) p \in \operatorname{con}(M - x_0),$$

если $p \in \{p\colon f'(x_0, p) \leqslant 0\}$. Устремляя λ к нулю, получаем, что любая точка конуса $\{p\colon f'(x_0, p) \leqslant 0\}$ есть предельная точка для конуса $\operatorname{con}(M - x_0)$. Поскольку конус $\{p\colon f'(x_0, p) \leqslant 0\}$ замкнут (так как $f'(x_0, p)$ — замкнутая функция p и справедлива теорема 1.5) и

$$\overline{\operatorname{con}(M - x_0)} \supseteq \{p\colon f'(x_0, p) \leqslant 0\},$$

то

$$\{p\colon f'(x_0, p) \leqslant 0\} = \overline{\operatorname{con}(M - x_0)}.$$

Но ранее было доказано, что

$$\{p\colon f'(x_0, p) \leqslant 0\} = [-\operatorname{con} \partial f(x_0)]^*.$$

Тем самым справедливо равенство

$$[-\operatorname{con} \partial f(x_0)]^* = \overline{\operatorname{con}(M - x_0)}.$$

Используя теперь замкнутость конуса $\operatorname{con} \partial f(x_0)$ и леммы I.3.3, I.3.5, получаем

$$-\operatorname{con} \partial f(x_0) = [-\operatorname{con} \partial f(x_0)]^{**} = [\overline{\operatorname{con}(M - x_0)}]^* = [\operatorname{con}(M - x_0)]^*,$$

что и требовалось доказать.

Теорема 3.18. *Если $f(x)$ — выпуклая функция, непрерывная в точке $x_0 \in M$, $M = \{x\colon f(x) \leqslant 0\}$ и существует точка x_1 такая, что $f(x_1) < 0$, то*

$$[\operatorname{con}(M - x_0)]^* = \begin{cases} \{0\}, & \text{если } f(x_0) < 0, \\ -\operatorname{con} \partial f(x_0), & \text{если } f(x_0) = 0. \end{cases}$$

Доказательство. Если $f(x_0) < 0$, то в силу непрерывности f в точке x_0 получаем, что $x_0 + \lambda p \in M$ для любого p при достаточно малом $\lambda > 0$. Поэтому $\operatorname{con}(M - x_0) = X$. Но тогда конус, сопряженный к $\operatorname{con}(M - x_0)$, может состоять только из одной точки —

начала координат,— что доказывает первое утверждение. Второе утверждение является следствием предыдущей теоремы и теоремы 3.5.

Теорема 3.19. *Если $f(x)$ — собственная выпуклая функция и существуют такие точки x_1, x_0, что $f(x_1) < 0$ и $x_0 \in \operatorname{ri} \operatorname{dom} f$, $x_0 \in M = \{x : f(x) \leqslant 0\}$, то*

$$[\operatorname{con}(M - x_0)]^* = \begin{cases} (\operatorname{Lin} \operatorname{dom} f)^{\perp}, & \text{если } f(x_0) < 0, \\ -\operatorname{con} \partial f(x_0), & \text{если } f(x_0) = 0. \end{cases}$$

(*Здесь* $(\operatorname{Lin} \operatorname{dom} f)^{\perp}$ — *подпространство, ортогональное к* $\operatorname{Lin} \operatorname{dom} f$.)

Доказательство. Второе утверждение теоремы следует из теорем 3.6 и 3.17. Первое утверждение получается, если заметить, что в силу теоремы 1.4 для $x_0 \in \operatorname{ri} \operatorname{dom} f$ выполняется соотношение

$$\operatorname{con}(M - x_0) = \operatorname{Lin} \operatorname{dom} f \tag{3.47}$$

и поэтому $[\operatorname{con}(M - x_0)]^* = [\operatorname{Lin} \operatorname{dom} f]^*$. Но если $K = \mathscr{L}$, где $\mathscr{L}$ — подпространство, то $x^* \in K^*$ тогда и только тогда, когда

$$\langle x, x^* \rangle \geqslant 0, \qquad x \in \mathscr{L}.$$

Если же $x \in \mathscr{L}$, то $-x \in \mathscr{L}$, и, значит,

$$\langle -x, x^* \rangle \geqslant 0, \qquad x \in \mathscr{L}.$$

Поэтому $\langle x, x^* \rangle = 0$, $x \in \mathscr{L}$, т. е. $x^* \in \mathscr{L}^{\perp}$. Итак, если $K = \mathscr{L}$, то и $K^* = \mathscr{L}^{\perp}$. Применяя это утверждение к равенству (3.47), получим требуемый результат.

Глава III

ВЫПУКЛЫЕ МНОГОЗНАЧНЫЕ ОТОБРАЖЕНИЯ

Многозначные отображения — объект, сравнительно недавно введенный в широкое пользование. Тем не менее понятие многозначного отображения в последние годы интенсивно исследуется. Введение этого понятия позволяет взглянуть с единой точки зрения на многие теоремы теории экстремальных задач и значительно упростить их доказательства.

§ 1. Основные определения и свойства

В этом параграфе вводятся основные определения и свойства многозначных отображений.

Пусть X и Y — конечномерные евклидовы пространства. Их прямое произведение $X \times Y$ обозначим через Z, так что, если $x \in X$, $y \in Y$, то пара (x, y) обозначает некоторую точку z пространства Z. Скалярные произведения в указанных пространствах будем обозначать через $\langle \cdot, \cdot \rangle_X$, $\langle \cdot, \cdot \rangle_Y$ и т. п., так что $\langle \cdot, \cdot \rangle_Z = \langle \cdot, \cdot \rangle_X + \langle \cdot, \cdot \rangle_Y$. Отсутствие индекса у скалярного произведения, как правило, не будет приводить к недоразумениям, так как всегда будет ясно, о каком пространстве идет речь, и всюду в дальнейшем эти индексы опускаются. Элементы пространства X, Y, Z будут обозначаться буквами x, y, z с индексами внизу. Элементы пространств X^*, Y^*, Z^* — дубликатов пространств X, Y, Z — будут обозначаться буквами x, y, z со звездочкой сверху, т. е. x^*, y^*, z^*. В скалярном произведении будут всегда участвовать элементы исходного пространства и его дубликата:

$$\langle x, x^* \rangle, \quad \langle y, y^* \rangle, \quad \langle z, z^* \rangle.$$

Пусть M — произвольное множество в $Z = X \times Y$. Тогда M определяет *многозначное отображение* a по формуле

$$a(x) = \{y\colon (x, y) \in M\}.$$

Множество M называется *графиком многозначного отображения* и обозначается $\operatorname{gf} a$. Легко видеть, что верна формула

$$\operatorname{gf} a = \{(x,\ y):\ y \in a(x)\}.$$

Положим также

$$\operatorname{dom} a = \{x:\ a(x) \neq \varnothing\},\quad \|a(x)\| = \sup_y \{\|y\|:\ y \in a(x)\}.$$

По определению $\|\varnothing\| = 0$.

Определение 1.1. Многозначное отображение a *выпукло*, если $\operatorname{gf} a$ — выпуклое множество.

Определение 1.2. Многозначное отображение a *выпуклозначно*, если $a(x)$ выпукло в Y.

Определение 1.3. Многозначное отображение a *замкнуто*, если $\operatorname{gf} a$ — замкнутое в Z множество.

Определение 1.4. Многозначное отображение a *ограничено*, если существует такая константа c, что

$$\|a(x)\| \leqslant c(1 + \|x\|).$$

Определение 1.5. Многозначное отображение a *полунепрерывно сверху в* x_0, если для любой окрестности нуля U в Y существует такая окрестность нуля V в X, что

$$a(x) \subseteq a(x_0) + U,\qquad \forall x \in x_0 + V.$$

Определение 1.6. Многозначное отображение a *полунепрерывно снизу в точке* x_0, если для любой окрестности нуля U в Y найдется такая окрестность нуля V в X, что

$$a(x_0) \subseteq a(x) + U,\qquad \forall x \in x_0 + V.$$

Определение 1.7. Многозначное отображение a *непрерывно*, если оно полунепрерывно сверху и снизу.

Определение 1.8. Многозначное отображение a *удовлетворяет условию Липшица* в области Ω, если

$$\rho(a(x_1),\ a(x_2)) \leqslant L\|x_1 - x_2\|,$$

где ρ — *хаусдорфово расстояние*, т. е.

$$\rho(A,\ B) = \max\left\{\sup_{y_1 \in A} \inf_{y_2 \in B} \|y_1 - y_2\|,\ \sup_{y_2 \in B} \inf_{y_1 \in A} \|y_1 - y_2\|\right\},$$

а L — постоянная.

Отметим некоторые свойства выпуклых отображений.

Лемма 1.1. *Пусть a — выпуклое, замкнутое многозначное отображение и в некоторой точке $x_0 \in \operatorname{dom} a$ множество $a(x_0)$ ограничено. Тогда a ограничено.*

Доказательство. Допустим, что это не так. Тогда найдутся последовательности $z_k = (x_k, y_k)$, $y_k \in a(x_k)$ такие, что

$$\frac{\|y_k\|}{1+\|x_k\|} \to +\infty.$$

Пусть $\|x_k\| \leqslant r$. Положим

$$\lambda_k = \frac{1+\|x_k\|}{\|y_k\|}.$$

По предположению $\lambda_k \to 0$. Возьмем $y_0 \in a(x_0)$ и рассмотрим точки

$$\begin{aligned} \bar{x}_k &= \lambda_k x_k + (1-\lambda_k)x_0, \\ \bar{y}_k &= \lambda_k y_k + (1-\lambda_k)y_0. \end{aligned} \tag{1.1}$$

Так как $0 \leq \lambda_k \leq 1$ при больших k, а отображение a выпукло, то $\bar{y}_k \in a(\bar{x}_k)$ при больших k. Заметим теперь, что $\bar{x}_k \to x_0$. Далее,

$$\bar{y}_k = (1+\|x_k\|)\frac{y_k}{\|y_k\|} + (1-\lambda_k)\,y_0. \tag{1.2}$$

Не ограничивая общности, можно считать, что $\frac{y_k}{\|y_k\|} \to w$, $\|w\| = 1$. Аналогично, в силу ограниченности $\|x_k\|$ получаем, что $\|x_k\| \to \alpha$. Тогда

$$\bar{y}_k \to (1+\alpha)w + y_0.$$

В силу замкнутости a $(1+\alpha)w + y_0 \in a(x_0)$. А так как y_0 — произвольная точка $a(x_0)$, то

$$(1+\alpha)w + a(x_0) \subseteq a(x_0),$$

что, как нетрудно видеть, противоречит ограниченности $a(x_0)$.

Пусть теперь $\|x_k\| \to +\infty$. Положим в этом случае

$$\lambda_k = \frac{1+\|x_k\|}{\|x_k\|\|y_k\|}.$$

Тогда, повторяя все предыдущие рассуждения с учетом замены формулы (1.2) на

$$\bar{y}_k = \frac{1+\|x_k\|}{\|x_k\|}\frac{y_k}{\|y_k\|} + (1-\lambda_k)\,y_0$$

и формулы $\frac{1+\|x_k\|}{\|x_k\|} \to 1$, снова получим противоречие. Это завершает доказательство.

Лемма 1.2. *Если* $\operatorname{gf} a = K$, *где* K — *выпуклый замкнутый конус, то* a *ограничено тогда и только тогда, когда* $a(0) = \{0\}$.

Доказательство. Если $a(0) = \{0\}$, то a ограничено по предыдущей лемме. Обратно, если a ограничено, то точка $y_0 \neq 0$, $y_0 \in a(0)$, существовать не может, так как K — конус, откуда $\lambda(0, y_0) = (0, \lambda y_0) \in K$ при всех $\lambda > 0$, т. е. $\lambda y_0 \in a(0)$, $\lambda > 0$, и, значит, $a(0)$ не ограничено.

Замечание. Если $\operatorname{gf} a$ — выпуклый конус, то нетрудно показать, что условие ограниченности a может быть записано в виде $\|a(x)\| \leqslant c\|x\|$.

Определение 1.9. Если $\operatorname{gf} a = K$ — выпуклый конус, то *сопряженное отображение* a^* определяется по правилу

$$a^*(y^*) = \{x^*\colon\ (-x^*,\ y^*) \in K^*\}.$$

Лемма 1.3. *Сопряженное отображение* $a^*(y^*)$ *ограничено тогда и только тогда, когда* $a(x)$ *определено при всех* x, *т. е.* $\operatorname{dom} a = X$.

Доказательство. Заметим, что K^* всегда замкнутый конус и поэтому $a^*(y^*)$ — замкнутое выпуклое отображение. Согласно предыдущей лемме теперь достаточно показать, что $a^*(0) = \{0\}$ тогда и только тогда, когда $\operatorname{dom} a = X$. Пусть $x^* \in a^*(0)$, т. е.

$$-\langle x, x^*\rangle + \langle y, 0\rangle \geqslant 0, \qquad (x, y) \in K$$

(здесь использовано определение двойственного конуса K^*). Тогда

$$\langle x, x^*\rangle \leqslant 0, \qquad x \in \operatorname{dom} a. \tag{1.3}$$

Если $\operatorname{dom} a = X$, то последнее неравенство возможно лишь, если $x^* = 0$, т. е. $a^*(0) = \{0\}$. Обратно, если $x^* \neq 0$, то $\operatorname{dom} a$ лежит в полупространстве, определяемом неравенством (1.3). Таким образом, $a^*(0)$ содержит ненулевой элемент, т. е. не ограничено, но и $\operatorname{dom} a \neq X$.

Теорема 1.1. *Пусть a — выпуклое замкнутое отображение, ограниченное в точке $x_0 \in \operatorname{int} \operatorname{dom} a$. Тогда в некоторой окрестности x_0 оно удовлетворяет условию Липшица.*

Доказательство. Пусть B — шар радиуса 1 с центром в нуле и $x_0 + rB \subseteq \operatorname{dom} a$. Такое $r > 0$ существует, ибо $x_0 \in \operatorname{int} \operatorname{dom} a$. Положим

$$x^+ = x_0 + r\frac{x - x_0}{\|x - x_0\|} = \left(1 - \frac{r}{\|x - x_0\|}\right)x_0 + \frac{r}{\|x - x_0\|}x,$$

$$x^- = x_0 - r\frac{x - x_0}{\|x - x_0\|} = \left(1 + \frac{r}{\|x - x_0\|}\right)x_0 - \frac{r}{\|x - x_0\|}x.$$

Тогда

$$x = \frac{\|x - x_0\|}{r}x^+ + \left(1 - \frac{\|x - x_0\|}{r}\right)x_0,$$

$$x_0 = \frac{\|x - x_0\|}{\|x - x_0\| + r}x^- + \frac{r}{\|x - x_0\| + r}x.$$

Так как $x \in x_0 + rB$, то $\|x - x_0\| \leqslant r$, и поэтому из предыдущих формул следует, что x и x_0 являются выпуклыми комбинациями точек x^+, x_0 и x^-, x соответственно.

В силу выпуклости отображения a

$$a(x) \supseteq \frac{\|x - x_0\|}{r}a(x^+) + \left(1 - \frac{\|x - x_0\|}{r}\right)a(x_0), \tag{1.4}$$

$$a(x_0) \supseteq \frac{\|x - x_0\|}{\|x - x_0\| + r}a(x^-) + \frac{r}{\|x - x_0\| + r}a(x). \tag{1.5}$$

Пусть $y^+ \in a(x^+)$, $y^- \in a(x^-)$, $y_0 \in a(x_0)$, $y \in a(x)$ — любые точки из соответствующих множеств. Тогда из включения (1.4) получаем

$$a(x) \supseteq \frac{\|x - x_0\|}{r}(y^+ - y_0) + y_0$$

или

$$y_0 \in a(x) - \frac{\|x - x_0\|}{r}(y^+ - y_0). \tag{1.6}$$

Но в силу предположений теоремы и леммы 1.1 множество

$a(x)$ ограничено. Поэтому

$$\|y^+\| \leqslant c(1+\|x^+\|), \qquad \|y_0\| \leqslant c(1+\|x_0\|),$$
$$\|y^+ - y_0\| \leqslant c(1+\|x^+\|) + c(1+\|x_0\|).$$

Но $\|x^+\| \leqslant \|x_0\| + r$. Поэтому окончательно получаем

$$\|y^+ - y_0\| \leqslant c(2+2\|x_0\|+r).$$

Теперь из соотношения (1.6) следует

$$y_0 \in a(x) + \frac{c(2+2\|x_0\|+r)}{r}\|x-x_0\|B,$$

или, поскольку y_0 — произвольная точка $a(x_0)$,

$$a(x_0) \subseteq a(x) + \frac{c(2+2\|x_0\|+r)}{r}\|x-x_0\|B. \qquad (1.7)$$

Аналогичные преобразования формулы (1.5) приводят к формуле

$$a(x) \subseteq a(x_0) + \frac{2c(1+\|x_0\|+r)}{r}\|x-x_0\|B. \qquad (1.8)$$

Но из этих включений сразу следует

$$\rho(a(x), a(x_0)) \leqslant \frac{2c(1+\|x_0\|+r)}{r}\|x-x_0\|, \qquad (1.9)$$

что и требовалось доказать.

Теорема 1.2. *Пусть $\Omega \subseteq \operatorname{int}\operatorname{dom} a$ и Ω компактно. Тогда выпуклое замкнутое многозначное отображение a, ограниченное в некоторой точке $\bar{x}$, удовлетворяет условию Липшица в Ω.*

Доказательство. В силу компактности Ω существует такое $r > 0$, что $\Omega + rB \subseteq \operatorname{int}\operatorname{dom} a$. Поэтому в качестве x_0 можно брать любую точку из Ω и выбранное r годится для использования его в предыдущем доказательстве.

Теорема 1.3. *Если выпуклое замкнутое ограниченное многозначное отображение a определено всюду в X, то*

$$\rho(a(x), a(x_0)) \leqslant 2c\|x-x_0\|.$$

Доказательство получается из неравенства (1.9), если устремить r к бесконечности. Это возможно сделать, так как $\operatorname{dom} a = X$.

§ 2. Локально сопряженные отображения

Локальные свойства дифференцируемых функций достаточно хорошо могут быть описаны при помощи понятия производной и связанного с ним понятия градиента функции. Для выпуклых функций понятие градиента заменяется субдифференциалом. В случае многозначного отображения аналогичную роль играет локально сопряженное отображение, которое вводится в этом параграфе.

1. Основные свойства. Введем некоторые обозначения. Положим

$$W_a(x, y^*) = \inf_y \{\langle y, y^*\rangle \colon\ y \in a(x)\}, \tag{2.1}$$

$$\Omega_a(x^*, y^*) = \inf_{(x,y)} \{-\langle x, x^*\rangle + \langle y, y^*\rangle \colon\ (x, y) \in \operatorname{gf} a\}. \tag{2.2}$$

Если $a(x) = \varnothing$, то положим $W_a(x, y^*) = +\infty$.

Очевидно, что $W_a(x, y^*)$ и $\Omega_a(x^*, y^*)$ есть взятые со знаком минус опорные функции множеств $a(x)$ и $\operatorname{gf} a$ соответственно. Тем не менее эти обозначения оказываются очень удобными во многих дальнейших выкладках.

Заметим, что

$$\Omega_a(x^*, y^*) = \inf_x \{-\langle x, x^*\rangle + W_a(x, y^*)\}, \tag{2.3}$$

так что всегда имеет место неравенство

$$\Omega_a(x^*, y^*) \leqslant -\langle x, x^*\rangle + W_a(x, y^*). \tag{2.4}$$

Кроме того, из формулы (2.3) следует, что если рассматривать $W_a(x, y^*)$ при фиксированном y^* как функцию x, то

$$\Omega_a(x^*, y^*) = -(W_a(\cdot, y^*))^*(x^*), \tag{2.5}$$

т. е. Ω_a есть взятая со знаком минус сопряженная к $W_a(\cdot, y^*)$ функция (см. § 2 гл. II).

Л е м м а 2.1. *Пусть a — выпуклое многозначное отображение. Тогда*

$$\begin{gathered} a(\lambda_1 x_1 + \lambda_2 x_2) \supseteq \lambda_1 a(x_1) + \lambda_2 a(x_2), \\ \lambda_1 \geqslant 0, \quad \lambda_2 \geqslant 0, \quad \lambda_1 + \lambda_2 = 1, \end{gathered} \tag{2.6}$$

и $W_a(x, y^*)$ — выпуклая по x функция.

Доказательство. Если

$$(x_1,\ y_1) \in \mathrm{gf}\, a, \qquad (x_2,\ y_2) \in \mathrm{gf}\, a,$$

то

$$(\lambda_1 x_1 + \lambda_2 x_2,\ \lambda_1 y_1 + \lambda_2 y_2) \in \mathrm{gf}\, a,$$

так как $\mathrm{gf}\, a$ — выпуклое множество. Таким образом,

$$\lambda_1 y_1 + \lambda_2 y_2 \in a(\lambda_1 x_1 + \lambda_2 x_2).$$

Так как y_1 и y_2 — произвольные точки множеств $a(x_1)$ и $a(x_2)$ соответственно, то из последнего включения следует (2.6). Так как точная нижняя грань по более широкому множеству меньше, чем по более узкому, то в силу включения (2.6)

$$W_a(\lambda_1 x_1 + \lambda_2 x_2,\ y^*) = \inf_y \{\langle y,\ y^*\rangle :\ y \in a(\lambda_1 x_1 + \lambda_2 x_2)\} \leqslant$$
$$\leqslant \inf_{y_1, y_2} \{\langle \lambda_1 y_1 + \lambda_2 y_2,\ y^*\rangle :\ y_1 \in a(x_1),\ y_2 \in a(x_2)\} =$$
$$= \lambda_1 W_a(x_1,\ y^*) + \lambda_2 W_a(x_2,\ y^*),$$

что доказывает второе утверждение леммы.

Возьмем $z \in \mathrm{gf}\, a$ и введем обозначение

$$K_a(z) \doteq \mathrm{con}\,(\mathrm{gf}\, a - z), \tag{2.7}$$

т. е.

$$K_a(z) = \{\bar z:\ \bar z = \lambda(z_1 - z),\quad \lambda > 0,\quad z_1 \in \mathrm{gf}\, a\}.$$

Нетрудно видеть, что справедлива и другая формула:

$$K_a(z) = \{\bar z:\ z + \lambda \bar z \in \mathrm{gf}\, a \ \text{ при достаточно малых } \lambda > 0\}.$$

Обозначим

$$a(x;\ y^*) = \{y \in a(x):\ \langle y,\ y^*\rangle = W_a(x,\ y^*)\}, \tag{2.8}$$

$$a_z(\bar x) = \{\bar y:\ (\bar x,\ \bar y) \in K_a(z)\}. \tag{2.9}$$

Таким образом, $a(x;\ y^*)$ в соответствии с формулой (2.1) есть множество тех точек y, на которых $\langle y,\ y^*\rangle$ достигает своей нижней грани по $y \in a(x)$. Многозначное отображение $a_z(\bar x)$ полностью определяется конусом $K_a(z)$ и $\mathrm{gf}\, a_z = K_a(z)$.

Определение 2.1. Отображение

$$a^*(y^*;\ z) = \{x^*:\ (-x^*,\ y^*) \in K_a^*(z)\}$$

называется *локально сопряженным в точке z* к выпуклому отображению a.

Сопоставляя данное определение с определением 1.9 и формулой (2.9), получаем

$$a^*(y^*; z) = a_z^*(y^*).$$

Теорема 2.1. *Пусть a — выпуклое отображение. Тогда*

$$a^*(y^*; z) = \begin{cases} \varnothing, & y \notin a(x; y^*), \\ \partial_x W_a(x, y^*), & y \in a(x; y^*). \end{cases}$$

Замечание. Если a — выпуклое многозначное отображение, то функция $W_a(x, y^*)$ выпукла по x. В формулировке теоремы через $\partial_x W_a(x, y^*)$ обозначен субдифференциал $W_a(x, y^*)$ как функции x, т. е. множество таких x^*, что

$$W_a(x_1, y^*) - W_a(x, y^*) \geqslant \langle x_1 - x, x^* \rangle$$

при всех $x_1 \in X$.

Доказательство. Пусть $x^* \in a^*(y^*; z)$, $z = (x, y)$. Тогда по определению сопряженного конуса $K_a^*(z)$ это означает, что

$$-\langle \bar{x}, x^* \rangle + \langle \bar{y}, y^* \rangle \geqslant 0, \quad (\bar{x}, \bar{y}) \in K_a(z).$$

Последнее в силу формулы (2.7) эквивалентно тому, что

$$-\langle x_1 - x, x^* \rangle + \langle y_1 - y, y^* \rangle \geqslant 0, \quad (x_1, y_1) \in \operatorname{gf} a. \tag{2.10}$$

Если $x_1 = x$, $y_1 \in a(x)$, то

$$\langle y_1, y^* \rangle \geqslant \langle y, y^* \rangle, \quad y_1 \in a(x).$$

Поэтому $y \in a(x; y^*)$ и

$$\langle y, y^* \rangle = W_a(x, y^*). \tag{2.11}$$

Тогда из неравенства (2.10) следует, что

$$\langle y_1, y^* \rangle - W_a\langle x, y^* \rangle \geqslant \langle x_1 - x, x^* \rangle.$$

Взяв точную нижнюю грань в левой части по $y_1 \in a(x_1)$, получаем

$$W_a(x_1, y^*) - W_a(x, y^*) \geqslant \langle x_1 - x, x^* \rangle, \tag{2.12}$$

т. е. $x^* \in \partial_x W_a(x, y^*)$.

Если же $x^* \in \partial_x W_a(x, y^*)$, $y \in a(x; y^*)$, то, идя от неравенства (2.12) в обратном направлении, нетрудно убедиться, что $x^* \in a^*(y^*; z)$; это завершает доказательство теоремы.

Замечание. Одновременно показано, что все множества $a^*(y^*; z)$, где $z = (x, y)$, а $y \in a(x; y^*)$ — любое, совпадают между собой и равны $\partial_x W_a(x, y^*)$.

Теорема 2.2. *Отображение* $a^*(y^*; z)$ *ограничено тогда и только тогда, когда множество* $a_z(\bar{x}) = \{\bar{y}: (\bar{x}, \bar{y}) \in \in K_a(z)\}$ *не пусто при любом* $\bar{x}$, *т. е. когда* $\operatorname{dom} a_z = X$.

Доказательство следует из леммы 1.3 и того факта, что $a^*(y^*; z)$ есть отображение, сопряженное к $a_z(\bar{x})$ в смысле определения 1.9.

Из теоремы 2.2 следует, что отображение $a^*(y^*; z)$ будет неограниченным на границе множества $\operatorname{dom} a$.

Теорема 2.3. *Пусть a — выпуклое замкнутое ограниченное отображение и M — компактное множество в Z такое, что его проекция на X целиком лежит в* $\operatorname{int}\operatorname{dom} a$. *Тогда отображения $a^*(y^*; z)$ равномерно ограничены по $z \in M$, т. е. существует такая константа c, для которой*

$$\|a^*(y^*; z)\| \leqslant c\|y^*\|$$

при всех $z \in M$.

Доказательство. В силу компактности M и того, что проекция M целиком лежит в $\operatorname{int}\operatorname{dom} a$, найдется такое число $r > 0$, что для любой точки $z = (x, y) \in M$ точка x принадлежит $\operatorname{int}\operatorname{dom} a$ вместе с шаром радиуса r.

Допустим теперь, что теорема неверна. Тогда найдется такая последовательность точек $(-x_k^*, y_k^*) \in K_a(z_k)$, $z_k \in M$, что

$$\frac{\|x_k^*\|}{\|y_k^*\|} \to \infty.$$

Можно считать, что $z_k \to z_\infty \in \operatorname{int}\operatorname{dom} a$ и что $y_k^*/\|y_k^*\| \to \to y_\infty^*$. По определению (2.7) конуса $K_a(z)$ условие $(-x_k^*, y_k^*) \in K_a^*(z_k)$ эквивалентно условию

$$-\langle x - x_k, x_k^*\rangle + \langle y - y_k, y_k^*\rangle \geqslant 0, \qquad (x, y) \in \operatorname{gf} a,$$

или

$$-\left\langle x - x_k, \frac{x_k^*}{\|y_k^*\|}\right\rangle + \left\langle y - y_k, \frac{y_k^*}{\|y_k^*\|}\right\rangle \geqslant 0.$$

Положим теперь $\widetilde{x}_k = x_k + r x_k^* / \|x_k^*\|$ и возьмем $\widetilde{y}_k \in a(\widetilde{x}_k)$ произвольно. Все $\widetilde{y}_k$ ограничены в силу ограниченности отображения a. Но тогда, подставляя $(\widetilde{x}_k, \widetilde{y}_k)$ в предыдущее неравенство, получим

$$-r\frac{\|x_k^*\|}{\|y_k^*\|} + \left\langle \widetilde{y}_k - y_k, \frac{y_k^*}{\|y_k^*\|} \right\rangle \geqslant 0,$$

что невозможно при больших k, ибо $\|x_k^*\|/\|y_k^*\| \to \infty$. Теорема доказана.

Теорема 2.4. *Пусть a — выпуклое замкнутое ограниченное отображение. Тогда отображение $a^*(y^*; z)$ полунепрерывно сверху для всех $y^* \in Y^*$ и $z = (x, y)$ таких, что $x \in \operatorname{int}\operatorname{dom} a$.*

Доказательство. Допустим противное: для некоторых y_0^* и $z_0 = (x_0, y_0)$ таких, что $x_0 \in \operatorname{int}\operatorname{dom} a$, отображение $a^*(y^*; z)$ не полунепрерывно. Тогда для некоторого $\varepsilon > 0$ существуют последовательности $y_k^* \to y_0^*$, $z_k \to z_0$, $x_k^* \in a^*(y_k^*; z_k)$ такие, что

$$x_k^* \notin a^*(y_0^*; z_0) + \varepsilon B, \tag{2.13}$$

где B — единичный шар с центром в нуле. Поскольку в силу предыдущей теоремы все отображения $a^*(y^*; z_k)$ равномерно ограничены, то можно считать, что все x_k^* ограничены. Поэтому без потери общности допускаем, что $x_k^* \to x_0^*$. По определению отображения $a^*(y^*; z)$ имеем

$$-\langle x - x_k, x_k^* \rangle + \langle y - y_k, y_k^* \rangle \geqslant 0, \qquad (x, y) \in \operatorname{gf} a.$$

Переходя к пределу, получим

$$-\langle x - x_0, x_0^* \rangle + \langle y - y_0, y_0^* \rangle \geqslant 0,$$

т. е. $x_0^* \in a^*(y_0^*; z_0)$, что противоречит формуле (2.13).

Лемма 2.2. *$x^* \in a^*(y^*; z)$ тогда и только тогда, когда*

$$\Omega_a(x^*, y^*) + \langle x, x^* \rangle = W_a(x, y^*).$$

Доказательство. Согласно теореме 2.1 включение $x^* \in a^*(y^*; z)$ эквивалентно включению

$$x^* \in \partial_x W_a(x, y^*), \quad y \in a(x; y^*), \quad z = (x, y),$$

т. е.

$$W_a(x_1, y^*) - W_a(x, y^*) \geqslant \langle x_1 - x, x^* \rangle, \quad \forall x_1 \in X.$$

Переписывая последнее неравенство в виде

$$-\langle x_1, x^* \rangle + W_a(x_1, y^*) \geqslant -\langle x, x^* \rangle + W_a(x, y^*)$$

и беря точную нижнюю грань по x_1, приходим к соотношению

$$\Omega_a(x^*, y^*) \geqslant -\langle x, x^* \rangle + W_a(x, y^*).$$

Сопоставляя его с неравенством (2.4), получаем утверждение леммы.

2. Локально сопряженные отображения и вычисление субдифференциалов. Различные примеры выпуклых отображений будут рассмотрены в следующем параграфе, где также вычислены сопряженные отображения для некоторых конкретных выпуклых многозначных отображений. Здесь же будет проиллюстрировано использование понятия локально сопряженного отображения для вычисления субдифференциала одной достаточно сложной функции.

Пусть a — выпуклое замкнутое ограниченное отображение, а $\varphi(z)$ — выпуклая собственная функция, для которой $\operatorname{dom}\varphi = Z$. По теореме II.1.4 она непрерывна по z и в каждой точке имеет субдифференциал $\partial_z\varphi(z)$ (теорема II.3.5). Так как множество $a(x)$ ограничено и замкнуто, а функция $\varphi(z)$ непрерывна, то функция

$$f(x) = \min_y \{\varphi(x, y): \ y \in a(x)\} \qquad (2.14)$$

конечна для всех $x \in \operatorname{dom} a$ и непрерывна на $\operatorname{ri}\operatorname{dom} a$.

Обозначим $\tilde{Y} = Y \times R^1$, $\tilde{Z} = X \times \tilde{Y}$, так что

$$\tilde{y} = (y, y^0), \qquad \tilde{z} = (x, y, y^0).$$

Пусть $\tilde{a}$ — многозначное отображение, график которого в $\tilde{Z}$ задается следующей формулой:

$$\operatorname{gf}\tilde{a} = \{(x, y, y^0): \ (x, y) \in \operatorname{gf} a, \ \ y^0 \geqslant \varphi(x, y)\}. \quad (2.15)$$

Множество $\operatorname{gf}\tilde{a}$ выпукло и замкнуто, так как $\operatorname{gf} a$ — выпуклое замкнутое множество, а $\varphi(x, y)$ — выпуклая непрерывная функция. Таким образом,

$$\tilde{a}(x) = \{(y, y^0): \ y \in a(x), \ \ y^0 \geqslant \varphi(x, y)\} \qquad (2.16)$$

есть выпуклое замкнутое многозначное отображение. Ясно, что $\operatorname{dom}\tilde{a} = \operatorname{dom} a$ и $W_a = +\infty$ при $x \notin \operatorname{dom} a$.

Вычислим

$$W_{\tilde{a}}(x, y^*, y^{0*}) =$$
$$= \inf_{(y^0, y)} \{\langle y, y^* \rangle + y^0 y^{0*}\colon\ y \in a(x),\ y^0 \geqslant \varphi(x, y)\}$$

для $x \in \operatorname{dom} a$. Имеем

$$W_{\tilde{a}}(x, y^*, y^{0*}) = \begin{cases} -\infty, & \text{если } y^{0*} < 0; \\ W_a(x, y^*), & \text{если } y^{0*} = 0; \\ \min\limits_{y} \{\langle y, y^* \rangle + y^{0*}\varphi(x, y)\colon\ y \in a(x)\}, & \\ & \text{если } y^{0*} > 0. \end{cases}$$

Отсюда видно, что $W_{\tilde{a}}(x, 0, 1) = f(x)$, и, значит, функция $f(x)$ по лемме 2.1 — выпуклая. Из этого факта и из теоремы 2.1 следует, что вычисление субдифференциала $\partial f(x)$ сводится к вычислению локально сопряженного отображения к $\tilde{a}$.

Вычислим конус $K_{\tilde{a}}(x_0, y_0, y_0^0)$, взяв $y_0^0 = \varphi(x_0, y_0)$. По определению $K_a(\tilde{z})$ (см. формулу (2.7)) $\bar{\tilde{z}} = (\bar{x}, \bar{y}, \bar{y}^0) \in K_{\tilde{a}}(\tilde{z}^0)$, $\tilde{z}_0 = (x_0, y_0, y_0^0)$ тогда и только тогда, когда $\tilde{z}_0 + \lambda\bar{\tilde{z}} \in \operatorname{gf}\tilde{a}$ при достаточно малых $\lambda > 0$, т. е. согласно (2.15)

$$(x_0 + \lambda\bar{x},\ y_0 + \lambda\bar{y}) \in \operatorname{gf} a, \tag{2.17}$$

$$y_0^0 + \lambda\bar{y}^0 \geqslant \varphi(x_0 + \lambda x,\ y_0 + \lambda\bar{y}). \tag{2.18}$$

Конус, определяемый соотношением (2.17), есть, очевидно, $K_a(z_0) \times \mathbf{R}^1$, $z_0 = (x_0, y_0)$, если учесть, что $\bar{y}^0$ во включении (2.17) не фигурирует и может быть выбрано произвольно.

Как нетрудно убедиться,

$$(K_a(z_0) \times \mathbf{R}^1)^* = K_a^*(z_0) \times \{0\}. \tag{2.19}$$

Если учесть, что $y_0^0 = \varphi(x_0, y_0)$, то конус, определяемый формулой (2.18), есть $\operatorname{con}(\widetilde{M} - \tilde{z}_0)$ для множества $\widetilde{M} \subseteq \widetilde{Z}$, определяемого следующим образом:

$$\widetilde{M} = \{(x,\ y,\ y^0):\ \varphi(x,\ y) - y^0 \leqslant 0\}.$$

Выбрав y^0 достаточно большим, всегда можно добиться того, чтобы было $\varphi(x,\ y) - y^0 < 0$. Если учесть теперь, что $\varphi(x,\ y)$ — непрерывная функция, то согласно теоремам II.3.5 и II.3.17, примененным к функции $\varphi(x,\ y) - y^0$ в точке $(x_0,\ y_0,\ \varphi(x_0,\ y_0))$, получим

$$[\operatorname{con}(\widetilde{M} - \widetilde{z}_0)]^* =$$
$$= \{(-\lambda x_0^*,\ -\lambda y_0^*,\ \lambda) : (x_0^*,\ y_0^*) \in \partial_z\varphi(z_0),\quad \lambda \geqslant 0\}, \tag{2.20}$$

где

$$z_0 = (x_0,\ y_0),\quad \partial_{\widetilde{z}}(\varphi(z_0) - y_0) = (\partial_z\varphi(z_0),\ -1).$$

Пусть теперь $(\bar{x}_0,\ \bar{y}_0) \in K_a(z_0)$. Выберем $\bar{y}_0^0$ так, чтобы выполнялось неравенство

$$\bar{y}_0^0 > \varphi'(z_0,\ \bar{z}_0),\qquad \bar{z}_0 = (\bar{x}_0,\ \bar{y}_0), \tag{2.21}$$

где $\varphi'(z_0,\ \bar{z}_0)$ — производная от $\varphi(z)$ в точке z_0 по направлению $\bar{z}_0$. Тогда для достаточно малых $\lambda > 0$

$$\bar{y}_0^0 > \frac{\varphi(x_0 + \lambda\bar{x}_0,\ y_0 + \lambda\bar{y}_0) - \varphi(x_0,\ y_0)}{\lambda}. \tag{2.22}$$

С другой стороны, так как $y^0 = \varphi(x_0,\ y_0)$, то неравенство (2.18) можно переписать в виде

$$\bar{y}^0 \geqslant \frac{\varphi(x_0 + \lambda\bar{x},\ y_0 + \lambda\bar{y}) - \varphi(x_0,\ y_0)}{\lambda}.$$

Из непрерывности $\varphi(z)$ в силу неравенства (2.22) теперь следует, что соотношение (2.18) будет выполняться для всех точек $(\bar{x},\ \bar{y},\ \bar{y}_0)$ из окрестности точки $(\bar{x}_0,\ \bar{y}_0, \bar{y}_0^0)$. Это означает, что точка $(\bar{x}_0,\ \bar{y}_0,\ \bar{y}_0^0)$ принадлежит внутренности конуса, определяемого неравенством (2.18), и конусу, определяемому включением (2.17). Поэтому конус $K^*_{\widetilde{a}}(\widetilde{z}_0)$ как сопряженный к пересечению указанных конусов по теореме I.3.2 равен сумме сопряженных. Используя равенства (2.19) и (2.20), получаем, что

$(x^*, y^*, y^{0*}) \in K^*_{\tilde{a}}(\tilde{z}_0)$ тогда и только тогда, когда

$$\begin{gathered} x^* = x_1^* - \lambda x_0^*, \quad y^* = y_1^* - \lambda y_0^*, \quad y^{*0} = \lambda, \\ (x_1^*, y_1^*) \in K_a^*(z_0), \quad (x_0^*, y_0^*) \in \partial_z \varphi(z_0), \qquad \lambda \geqslant 0. \end{gathered} \tag{2.23}$$

В частности, $x^* \in \tilde{a}^*(y^*, y^{0*}; \tilde{z}_0)$, т. е. $(-x^*, y^*, y^{0*}) \in \in K^*_{\tilde{a}}(\tilde{z}_0)$ тогда и только тогда, когда

$$\begin{gathered} -x^* = -x_1^* - \lambda x_0^*, \quad y^* = y_1^* - \lambda y_0^*, \quad y_0^* = \lambda, \\ x_1^* \in a^*(y_1^*; z_0), \quad (x_0^*, y_0^*) \in \partial_z \varphi(z_0), \quad \lambda \geqslant 0. \end{gathered} \tag{2.24}$$

Пусть теперь $x_0 \in \operatorname{dom} a$, $y^* = 0$, $y^{0*} = 1$. Выберем y_0 так, чтобы в этой точке достигался минимум функции $\varphi(x_0, y)$ по $y \in a(x_0)$. Тогда согласно полученному выше выражению для $W_{\tilde{a}}(x, y^*, y^{0*})$ имеем

$$(y_0, \varphi(x_0, y_0)) \in \tilde{a}(x; 0, 1),$$

так как $\langle y, y^* \rangle + y^{0*} y^0$ при $y^* = 0$, $y^{0*} = 1$ достигает своего минимума по $y \in a(x_0)$, $y^0 \geqslant \varphi(x_0, y)$ именно в точке $(y_0, \varphi(x_0, y_0))$.

Так как $f(x) = W_{\tilde{a}}(x, 0, 1)$, то $\partial f(x_0) = \partial_x W_{\tilde{a}}(x_0, 0, 1)$. Из результатов теоремы 2.1 теперь вытекает, что

$$\partial f(x_0) = \tilde{a}^*(0, 1; \tilde{z}_0), \quad \tilde{z}_0 = (x_0, y_0, \varphi(x_0, y_0)).$$

Воспользуемся формулами (2.24). Они показывают, что $x^* \in \tilde{a}^*(0, 1, \tilde{z}_0)$ тогда и только тогда, когда

$$\begin{gathered} x^* = x_1^* + x_0^*, \qquad y_1^* = y_0^*, \\ x_1^* \in a^*(y_1^*; z_0), \qquad (x_0^*, y_0^*) \in \partial_z \varphi(z_0). \end{gathered} \tag{2.25}$$

Теорема 2.5. *Пусть a — выпуклое замкнутое ограниченное отображение, $\varphi(z)$ — выпуклая собственная функция $z = (x, y)$, $\operatorname{dom} \varphi = Z$ и*

$$f(x) = \min_y \{\varphi(x, y) \colon y \in a(x)\}. \tag{2.26}$$

Тогда $f(x)$ — выпуклая функция и для всех $x_0 \in \operatorname{dom} a$ справедливо равенство

$$\partial f(x_0) = \{x^* + a^*(y^*; z_0) \colon (x^*, y^*) \in \partial_z \varphi(z_0)\}, \tag{2.27}$$

где $z_0 = (x_0, y_0)$, *а* $y_0 \in a(x_0)$ — *любая точка, в которой достигается минимум в* (2.26). *В частности, если* $\varphi(z)$ — *дифференцируемая функция, то*

$$\partial f(x_0) = \varphi'_x(z_0) + a^*(\varphi'_y(z_0), z_0), \tag{2.28}$$

где $\varphi'_x(z_0)$ *и* $\varphi'_y(z_0)$ — *векторы частных производных по компонентам* x^i *и* y^j *векторов* x *и* y *соответственно.*

Доказательство. Формула (2.27) представляет собой просто иначе записанные соотношения (2.25). Если же функция $\varphi(z)$ дифференцируема, то $\partial_z\varphi(z_0)$ состоит из одного вектора — градиента функции $\varphi(z)$. Последний распадается, очевидно, на компоненты, соответствующие частным производным по x^i и по y^i. Теперь в формуле (2.27) $x^* = \varphi'_x(z_0)$, $y^* = \varphi'_y(z_0)$.

Замечание. Ограниченность отображения a использовалась только при доказательстве существования точки $y_0 \in a(x_0)$, в которой функция $\varphi(x_0, y)$ достигает минимума. Поэтому, если вместо ограниченности a предположить, что минимум в правой части формулы (2.26) достигается, то теорема останется в силе. Заметим также, что вместо условия $\operatorname{dom} \varphi = Z$ достаточно потребовать, чтобы функция $\varphi(z)$ была непрерывной для x из некоторой окрестности Ω точки x_0 и $y \in a(x)$, $x \in \Omega$.

Рассмотрим теперь постоянное отображение a:

$$a(x) = \{M:\ x \in X\},$$

где M — выпуклое замкнутое множество в Y. Ясно, что

$$\operatorname{gf} a = X \times M,$$

поэтому

$$K_a(z_0) = X \times \operatorname{con}(M - y_0)$$

и легко вычислить $K_a^*(z_0)$:

$$K_a^*(z_0) = \{0\} \times [\operatorname{con}(M - y_0)]^*.$$

Отсюда следует, что

$$a^*(y^*; z_0) = \begin{cases} 0, & \text{если } y^* \in [\operatorname{con}(M - y_0)]^*, \\ \varnothing, & \text{если } y^* \notin [\operatorname{con}(M - y_0)]^*. \end{cases} \tag{2.29}$$

Используя теперь теорему 2.5 и значение к ней, получаем следующий результат.

Теорема 2.6. *Пусть $\varphi(x, y)$ — выпуклая функция, непрерывная по x в окрестности Ω точки x_0, и $y \in M$. Пусть, далее,*

$$f(x) = \min_y \{\varphi(x, y):\ y \in M\} \tag{2.30}$$

и при $x = x_0$ минимум достигается в точке $y_0 \in M$. Тогда

$$\partial f(x_0) = \{x^*:\ (x^*, y^*) \in \partial_z \varphi(z_0),\ \ y^* \in [\operatorname{con}(M - y_0)]^*\}.$$

3. Композиция многозначных отображений. Пусть X, X_1, Y — некоторые конечномерные линейные пространства, и пусть заданы многозначные отображения: a_1 — из X в X_1, a_2 — из X_1 в Y. Тогда

$$\operatorname{gf} a_1 \subseteq X \times X_1, \qquad \operatorname{gf} a_2 \subseteq X_1 \times Y.$$

Определим многозначное отображение $a_2 \circ a_1$ из X в Y по правилу

$$(a_2 \circ a_1)(x) = \{y:\ y \in a_2(x_1) \ \text{ при некотором } x_1 \in a_1(x)\}$$

или

$$(a_2 \circ a_1)(x) = \bigcup_{x_1 \in a_1(x)} a_2(x_1).$$

Так как

$$\operatorname{gf}(a_2 \circ a_1) = \{(x, y):\ (x, x_1) \in \operatorname{gf} a_1,\ (x_1, y) \in \operatorname{gf} a_2\},$$

то легко убедиться, что $a_2 \circ a_1$ есть выпуклое отображение, если a_1 и a_2 выпуклы.

Теорема 2.7. *Пусть существует такая точка $(\bar{x}, \bar{x}_1, \bar{y})$, что либо*

$$(\bar{x}, \bar{x}_1) \in \operatorname{ri} \operatorname{gf} a_1,\ (\bar{x}_1, y) \in \operatorname{ri} \operatorname{gf} a_2,$$

либо

$$(\bar{x}, \bar{x}_1) \in \operatorname{int} \operatorname{gf} a_1,\ (\bar{x}_1, \bar{y}) \in \operatorname{gf} a_2.$$

Тогда для $z_0 = (x_0, y_0)$ выполняется соотношение

$$(a_2 \circ a_1)^* (y^*; z_0) = \{x_1^*:\ x_1^* \in a_1^*(x_2^*; (x_0, x_{10})),\ x_2^* \in a_2^*(y^*; (x_{10}, y_0))\},$$

где x_{10} — любая точка такая, что $x_{10} \in a_1(x_0)$, $y_0 \in a_2(x_{10})$, или, короче,

$$(a_2 \circ a_1)^* (\cdot\,; z_0) = a_1^*(\cdot\,; (x_0, x_{10})) \circ a_2^*(\cdot\,; (x_{10}, y_0)).$$

Доказательство. По определению $x^* \in (a_2 \circ a_1)^*(y^*; z_0)$ тогда и только тогда, когда $(-x^*, y^*) \in K_a^*(z_0)$, или, что эквивалентно, когда

$$-\langle x - x_0, x^* \rangle + \langle y - y_0, y^* \rangle \geqslant 0 \qquad (2.31)$$

для всех $(x, y) \in \mathrm{gf}(a_2 \circ a_1)$. Учитывая выражение для $\mathrm{gf}(a_2 \circ a_1)$ и то, что $(x_0, y_0) \in \mathrm{gf}(a_2 \circ a_1)$, получаем, что существует такая точка x_{10}, что

$$(x_0, x_{10}) \in \mathrm{gf}\, a_1, \quad (x_{10}, y_0) \in \mathrm{gf}\, a_2.$$

Перепишем теперь неравенство (2.31) в эквивалентном виде:

$$\begin{gathered} -\langle x - x_0, x^* \rangle + \langle x_1 - x_{10}, 0 \rangle + \langle y - y_0, y^* \rangle \geqslant 0, \\ (x, x_1) \in \mathrm{gf}\, a_1, \quad (x_1, y) \in \mathrm{gf}\, a_2. \end{gathered} \qquad (2.32)$$

Рассмотрим пространство $X \times X_1 \times Y$, элементами которого являются тройки (x, x_1, y), $x \in X$, $x_1 \in X_1$, $y \in Y$, и два множества в этом пространстве

$$M_1 = \{(x, x_1, y): (x, x_1) \in \mathrm{gf}\, a_1\},$$
$$M_2 = \{(x, x_1, y): (x_1, y) \in \mathrm{gf}\, a_2\}.$$

Неравенство (2.32) означает, что $(-x^*, 0, y^*)$ принадлежит конусу, сопряженному к

$$\begin{aligned} &\mathrm{con}(M_1 \cap M_2 - (x_0, x_{10}, y_0)) = \\ &\quad = \mathrm{con}[(M_1 - (x_0, x_{10}, y_0)) \cap (M_2 - (x_0, x_{10}, y_0))] = \\ &\qquad = \mathrm{con}(M_1 - (x_0, x_{10}, y_0)) \cap \mathrm{con}(M_2 - (x_0, x_{10}, y_0)). \end{aligned}$$

(Здесь были использованы теоремы I.3.8 и I.3.7.) Легко подсчитать, что

$$\mathrm{con}(M_1 - (x_0, x_{10}, y_0)) = K_{a_1}((x_0, x_{10})) \times Y,$$
$$\mathrm{con}(M_2 - (x_0, x_{10}, y_0)) = X \times K_{a_2}((x_{10}, y_0)),$$

и поэтому

$$\begin{gathered} [\mathrm{con}(M_1 - (x_0, x_{10}, y_0))]^* = K_{a_1}^*((x_0, x_{10})) \times \{0\}, \\ [\mathrm{con}(M_2 - (x_0, x_{10}, y_0))] = \{0\} \times K_{a_2}^*((x_{10}, y_0)). \end{gathered} \qquad (2.33)$$

Используя теоремы I.3.2 или II.3.10 применительно к конусам $\mathrm{con}(M_1 - (x_0, x_{10}, y_0))$ и $\mathrm{con}(M_2 - (x_0, x_{10}, y_0))$, получим

$$[\mathrm{con}(M_1 \cap M_2 - (x_0, x_{10}, y_0))]^* =$$
$$= [\mathrm{con}(M_2 - (x_0, x_{10}, y_0))]^* + [\mathrm{con}(M_2 - (x_0, x_{10}, y_0))]^*.$$

Учитывая соотношения (2.33), можно утверждать, что вектор $(-x^*, 0, y^*)$ представим в виде

$$-x^* = -x_1^*, \quad 0 = y_1^* - x_2^*, \quad y^* = y_2^*, \tag{2.34}$$
$$\left(-x_1^*, y_1^*\right) \in K_{a_1}^*((x_0, x_{10})), \quad \left(-x_2^*, y_2^*\right) \in K_{a_2}^*((x_{10}, y_0)).$$

Но соотношения (2.34), очевидно, эквивалентны соотношениям

$$x^* = x_1^*, \quad x_1^* \in a_1^*\left(x_2^*; (x_0, x_{10})\right), \quad x_2^* \in a_2^*\left(y^*; (x_{10}, y_0)\right),$$

что доказывает теорему.

§ 3. Примеры выпуклых многозначных отображений

Покажем на конкретных примерах, как вычисляются величины, связанные с выпуклым многозначным отображением. В дальнейшем эти вычисления будут использованы при решении некоторых экстремальных задач.

Пример 3.1. Пусть M — выпуклое множество в Y и $\mathrm{gf}\, a = X \times M$, т. е. $a(x) = M$ при всех $x \in X$. Такое отображение будем называть *постоянным*. Тогда

$$W_a(x, y^*) = \inf_y \{\langle y, y^*\rangle : y \in M\} = -W_M(-y^*),$$

где W_M — опорная функция множества M,

$$\Omega_a(x^*, y^*) = \inf_{(x,y)} \{-\langle x, x^*\rangle + \langle y, y^*\rangle : x \in X, \ y \in M\} =$$
$$= \begin{cases} -\infty, & \text{если} \quad x^* \neq 0, \\ -W_M(-y^*), & \text{если} \quad x^* = 0. \end{cases}$$

Как показано в предыдущем параграфе для постоянного отображения,

$$a^*(y^*; z_0) = \begin{cases} 0, & \text{если} \quad y^* \in [\mathrm{con}(M - y_0)]^*, \\ \varnothing, & \text{если} \quad y^* \notin [\mathrm{con}(M - y_0)]^*. \end{cases}$$

Пример 3.2. Пусть $a(x) = Ax + U$, где U — выпуклое множество в Y, а A — оператор, действующий из пространства X в пространство Y. Тогда

$$W_a(x, y^*) = \langle Ax, y^*\rangle - W_U(-y^*),$$

$$\Omega_a(x^*, y^*) = \begin{cases} -\infty, & \text{если } A^*y^* \neq x^*, \\ -W_U(-y^*), & \text{если } x^* = A^*y^*. \end{cases}$$

Если $y_0 = Ax_0 + u_0$, $u_0 \in U$, то

$$a^*(y^*; z_0) = \begin{cases} A^*y^*, & \text{если } y^* \in [\operatorname{con}(U - u_0)]^*, \\ \varnothing, & \text{если } y^* \notin [\operatorname{con}(U - u_0)]^*. \end{cases}$$

Элементарные выкладки, связанные с этими вычислениями, представляем читателю в качестве упражнения.

Пример 3.3. Пусть $Y = \mathbf{R}^1$, $f(x)$ — собственная выпуклая функция и

$$\operatorname{gf} a = \operatorname{epi} f, \qquad a(x) = \{y\colon\ y \geqslant f(x)\}.$$

Тогда $y^* \in \mathbf{R}^1$ и

$$W_a(x, y^*) = \inf_y \{yy^*\colon\ y \geqslant f(x)\} =$$

$$= \begin{cases} +\infty, & \text{если } x \notin \operatorname{dom} f, \\ -\infty, & \text{если } y^* < 0, \quad x \in \operatorname{dom} f, \\ y^*f(x), & \text{если } y^* \geqslant 0, \quad x \in \operatorname{dom} f, \end{cases}$$

$$\Omega_a(x^*, y^*) = \inf_{(x,y)} \{-\langle x, x^*\rangle + yy^*\colon\ y \geqslant f(x)\} =$$

$$= \begin{cases} -\infty, & \text{если } y^* < 0, \\ -W_{\operatorname{dom} f}(x^*), & \text{если } y^* = 0. \end{cases}$$

Если $y^* > 0$, то

$$\Omega_a(x^*, y^*) = \inf_x \{-\langle x, x^*\rangle + y^*f(x)\} =$$

$$= -y^* \sup_x \left\{\left\langle x, \frac{x^*}{y^*}\right\rangle - f(x)\right\} = -y^*f^*\left(\frac{x^*}{y^*}\right).$$

Пусть теперь $x_0 \in \operatorname{dom} f$, $y_0 = f(x_0)$, $y^* \geqslant 0$. Так как при условии $y \geqslant f(x_0)$ произведение yy^* достигает своего минимума по y, если $y = y_0 = f(x_0)$, то $y_0 \in a(x; y^*)$. По теореме 2.1 и из полученного выражения для $W_a(x, y^*)$ получаем

$$a^*(y^*; z_0) = \begin{cases} y^*\partial f(x_0), & y^* > 0, \\ \partial\delta(x_0 \mid \operatorname{dom} f), & y^* = 0, \end{cases}$$

$$z_0 = (x_0, f(x_0)), \quad x_0 \in \operatorname{dom} f.$$

Пример 3.4. Пусть $f_i(y)$, $i=1, \ldots, m$,— выпуклые функции в пространстве $\mathbf{R}^m$. Положим

$$\operatorname{gf} a = \{(x, y): f_i(y) \leqslant x^i, i = 1, \ldots, m\}$$

или

$$\operatorname{gf} a = \{(x, y): f(y) \leqslant x\},$$

где $f(y) = (f_1(y), \ldots, f_m(y))$, $x = (x^1, \ldots, x^m)$ и неравенство $f(y) \leqslant x$ для векторов означает выполнение неравенства для каждой компоненты: $f_i(y) \leqslant x^i$, $i = 1, \ldots, m$. Тогда

$$a(x) = \{y: f(y) \leqslant x\}.$$

Отсюда следует, что

$$W_a(x, y^*) = \inf_y \{\langle y, y^*\rangle: f(y) \leqslant x\},$$

$$\Omega_a(x^*, y^*) = \inf_{(x,y)} \{-\langle x, x^*\rangle + \langle y, y^*\rangle: f(y) \leqslant x\} = -\infty,$$

если хотя бы одна из компонент вектора x^* больше нуля. Если $x^* \leqslant 0$, то

$$\Omega_a(x^*, y^*) = \inf_y \{-\langle f(y), x^*\rangle + \langle y, y^*\rangle\}.$$

Вычисление сопряженного отображения будет проведено ниже. Отметим здесь справедливость следующего интересного результата.

Теорема 3.1. *Пусть $f_i(y)$ — выпуклые замкнутые функции. Тогда, если множество $a(x_0) = \{y: f(y) \leqslant x_0\}$ ограничено, то существует такая константа c, что*

$$\sup_y \{\|y\|: f(y) \leqslant x\} \leqslant c(1 + \|x\|).$$

Доказательство теоремы получается прямым применением леммы 1.1 к рассматриваемому отображению.

Пример 3.5. Пусть в X задан выпуклый замкнутый конус X^+ — конус положительных элементов. В частности, можно в качестве X^+ выбрать положительный ортант, т. е.

$$X^+ = \{x: x^i \geqslant 0, i = 1, \ldots, m\},$$

если $X = \mathbf{R}^m$.

Будем говорить, что *отображение* F, которое ставит в соответствие точке $y \in Y$ точку $F(y) \in X$, *выпукло* *относительно конуса* X^+, если

$$F(\lambda_1 y_1 + \lambda_2 y_2) \overset{X^+}{\leqslant} \lambda_1 F(y_1) + \lambda_2 F(y_2),$$
$$\lambda_1,\ \lambda_2 \geqslant 0, \qquad \lambda_1 + \lambda_2 = 1,$$

где неравенство $\overset{X^+}{\leqslant}$ понимается в смысле определений, данных в п. 3 § 3 главы I. Так как конус X^+ в данный момент фиксирован, то вместо $\overset{X^+}{\leqslant}$ будем писать просто $\leqslant$.

Положим

$$\operatorname{gf} a = \{(x,\ y):\ F(y) \leqslant x\}.$$

Если $x_1 \geqslant F(y_1)$, $x_2 \geqslant F(y_2)$, то для $\lambda_1,\ \lambda_2 \leqslant 0$, $\lambda_1 + \lambda_2 = 1$,

$$\lambda_1 x_1 + \lambda_2 x_2 \geqslant \lambda_1 F(y_1) + \lambda_2 F(y_2) \geqslant F(\lambda_1 y_1 + \lambda_2 y_2),$$

откуда следует, что $\operatorname{gf} a$ — выпуклое множество и поэтому отображение

$$a(x) = \{y:\ F(y) \leqslant x\}$$

выпукло. Теперь

$$W_a(x,\ y^*) = \inf_y \{\langle y,\ y^*\rangle:\ F(y) \leqslant x\}.$$

Если y фиксировано, то $x \geqslant F(y)$ тогда и только тогда, когда $x - F(y) \in X^+$, т. е. $x \in F(y) + X^+$ или $x = F(y) + x_1$, $x_1 \in X^+$. Поэтому

$$\Omega_a(x^*,\ y^*) = \inf_{(x,y)} \{\langle -x,\ x^*\rangle + \langle y,\ y^*\rangle:\ x \geqslant F(y)\} =$$
$$= \inf_y \inf_{x_1 \in X^+} \{-\langle F(y),\ x^*\rangle - \langle x_1,\ x^*\rangle + \langle y_1,\ y^*\rangle\} =$$
$$= \begin{cases} -\infty, & \text{если } -x^* \notin (X^+)^*, \\ \inf\limits_y \{-\langle F(y),\ x^*\rangle + \langle y,\ y^*\rangle:\ -x^* \in (X^+)^*\}. \end{cases}$$

Переобозначив $-x^*$ через x^*, получаем

$$\Omega_a(-x^*,\ y^*) = \begin{cases} \inf\limits_y \{\langle y,\ y^*\rangle + \langle F(y),\ x^*\rangle:\ x^* \in (X^+)^*\}, \\ -\infty, \quad \text{если } x^* \notin (X^+)^*. \end{cases}$$

Пример 3.6. Пусть $\varphi(z)$, $z \in X \times Y$,— выпуклая функция. Положим

$$\operatorname{gf} a = \{z:\ \varphi(z) \leqslant 0\}, \quad a(x) = \{y:\ \varphi(x,\ y) \leqslant 0\}.$$

Пусть $\varphi(z_0)=0$, функция $\varphi(z)$ непрерывна в точке z_0 и существует точка z_1, в которой $\varphi(z_1)<0$. По теореме II.3.18

$$K_a^*(z_0)=[\operatorname{con}(\operatorname{gf}a-z_0)]^*=-\operatorname{con}\partial_z\varphi(z_0),$$

поэтому

$$a^*(y^*;z_0)=$$
$$=\{\lambda x_0^*:\ y^*=-\lambda y_0^*,\quad (x_0^*,y_0^*)\in\partial_z\varphi(z),\quad \lambda\geqslant 0\}. \quad (3.1)$$

Итак, доказана

Теорема 3.2. *Пусть $\varphi(z)$ — выпуклая функция, непрерывная в точке z_0, $\varphi(z_0)=0$ и существует такая точка z_1, что $\varphi(z_1)<0$. Тогда справедлива формула* (3.1).

Пусть теперь

$$\varphi(z)=\max_{1\leqslant i\leqslant m}\varphi_i(z) \quad (3.2)$$

и функции $\varphi_i(z)$ непрерывны в точке z_0. Согласно теореме II.3.14

$$\partial_z\varphi(z)=\overline{\operatorname{co}}\Big(\bigcup_{i\in I(z)}\partial_z\varphi_i(z)\Big),\quad I(z)=\{i:\ \varphi_i(z)=\varphi(z)\}. \quad (3.3)$$

По теоремам II.3.2 и II.3.5 множества $\partial_z\varphi_i(z)$ ограничены и замкнуты, значит, этими свойствами будет обладать и их объединение. Из теорем I.1.6 и I.1.7 теперь вытекает, что в формуле (3.3) достаточно брать выпуклую оболочку без операции замыкания. Значит,

$$\partial_z\varphi(z)=\operatorname{co}\Big(\bigcup_{i\in I(z)}\partial_z\varphi_i(z)\Big). \quad (3.4)$$

Но по лемме I.1.4

$$z^*\in\operatorname{co}\bigcup_{i\in I(z)}\partial_z\varphi_i(z) \quad (3.5)$$

тогда и только тогда, когда

$$z^*=\lambda_1 z_1^*+\ldots+\lambda_k z_k^*,$$

где каждое z_j^* принадлежит одному из множеств $\partial_z\varphi_i(z)$, а λ_j неотрицательны и в сумме равны единице. Обозначим

$$J_i=\{j:\ z_j^*\in\partial_z\varphi_i(z)\},$$
$$\gamma_i=\sum_{j\in J_i}\lambda_j,\qquad i\in I(z);$$

тогда

$$z^* = \sum_{i=1}^{k} \gamma_i \left(\sum_{j \in J_i} \frac{\lambda_j}{\gamma_i} z_j^* \right). \tag{3.6}$$

Так как $\partial_z \varphi_i(z)$ — выпуклое множество, а

$$\sum_{j \in J_i} \frac{\lambda_j}{\gamma_i} = 1,$$

то

$$\sum_{j \in J_i} \frac{\lambda_j}{\gamma_i} z_j^* \in \partial_z \varphi_i(z). \tag{3.7}$$

Из соотношений (3.7) и (3.6) следует, что включение (3.5) выполняется тогда и только тогда, когда

$$\begin{gathered} z^* = \sum_{i \in I(z)} \gamma_i z_i^*, \quad z_i^* \in \partial_z \varphi_i(z), \\ \gamma_i \geqslant 0, \quad \sum_{i \in I(z)} \gamma_i = 1. \end{gathered} \tag{3.8}$$

Совмещая полученный результат с равенством (3.4), окончательно имеем

$$\partial_z \varphi(z) = \left\{ \sum_{i \in I(z)} \gamma_i z_i^* \colon z_i^* \in \partial_z \varphi_i(z), \quad \gamma_i \geqslant 0, \right. \\ \left. i \in I(z), \quad \sum_{i \in I(z)} \gamma_i = 1 \right\}. \tag{3.9}$$

Теорема 3.3. *Пусть* $\varphi_i(z)$, $i = 1, \ldots, m$, — *непрерывные в точке* z_0 *выпуклые функции,* $\varphi(z_0) = 0$ *и существует такая точка* z_1, *что* $\varphi_i(z_1) < 0$, $i = 1, \ldots, m$. *Пусть также*

$$a(x) = \{y \colon \varphi_i(x, y) \leqslant 0, \, i = 1, \ldots, m\}.$$

Тогда

$$a^*(y^*; z_0) = \left\{ \sum_{i \in I_0} \lambda_i x_i^* \colon y^* = -\sum_{i \in I_0} \lambda_i y_i^*, \right. \\ \left. (x_i^*, y_i^*) \in \partial_z \varphi_i(z), \quad \lambda_i \geqslant 0, \quad i \in I_0 \right\},$$

где $I_0 = \{i \colon \varphi_i(x_0, y_0) = 0\}$.

Доказательство. Если определить $\varphi(z)$ формулой (3.2), то рассматриваемое в теореме отображение $a(x)$ совпадает с рассмотренным выше. Так как $\varphi(z_0) = 0$, то $I(z_0) = I_0$. Подставив выражение для $\partial_z\varphi(z_0)$, даваемое формулой (3.9), в формулу (3.1) и обозначив $\lambda_i = \lambda\gamma_i$, получим требуемый результат.

Теорема 3.4. *Пусть $f_i(y)$, $i = 1, \ldots, m$,— выпуклые непрерывные в точке y_0 функции и*

$$a(x) = \{y\colon f_i(y) \leqslant x^i,\ i = 1, \ldots, m\}.$$

Тогда

$$a^*(y^*; z_0) = \Big\{-\lambda \in \mathbf{R}^m\colon \lambda_i \geqslant 0,\quad \lambda_i\left(f_i(y_0) - x_0^i\right) = 0,$$
$$i = 1, \ldots, m,\quad y^* + \sum_{i=1}^m \lambda_i y_i^* = 0;\quad y_i^* \in \partial_y f_i(y_0)\Big\}. \qquad (3.10)$$

(*Здесь λ означает вектор с компонентами λ_i, $i = 1, \ldots, m$.*)

Доказательство. Пусть в предыдущей теореме

$$\varphi_i(z) = f_i(y) - x^i.$$

Тогда, если выбрать $z_1 = (x_1, y_1)$, где y_1 произвольно, а $x_1^i > f_i(y_1)$, то $\varphi_i(z_1) < 0$, $i = 1, \ldots, m$. Кроме того,

$$\partial_z\varphi_i(z) = (-e_i,\ \partial_y f_i(y)), \qquad (3.11)$$

где $e_i = (\underbrace{0, 0, \ldots, 1}_{i}, 0, \ldots, 0)$. Поэтому $(x_i^*, y_i^*) \in \partial_z\varphi_i(z_0)$ только тогда, когда $x_i^* = -e_i$, $y_i^* \in \partial_y f_i(y)$. Если учесть, что λ_i в формуле (3.10) может быть отлично от нуля, только если $\varphi_i(z_0) = \varphi(z_0) = 0$, то окажется, что фактически сумма в (3.10) распространяется только на индексы i из I_0.

Заметим также, что если $f_i(y_0) < x_0^i$, $i=1, \ldots, m$, то $K_a(z_0) = Z$, $K_a^*(z_0) = \{0\}$ и $x^* \in a^*(y^*; z_0)$ только тогда, когда $x^* = 0$, $y^* = 0$. Но этот же результат дает и формула (3.10).

Так как очевидно, что

$$\sum_{i=1}^m \lambda_i e_i = \lambda,$$

то формула (3.10) получается из формулы для $a^*(y^*; z_0)$ в теореме 3.3 подстановкой в нее выражения (3.11). Теорема 3.4 дает выражение для локально сопряженного отображения, рассмотренного в примере 3.4.

Допустим теперь, что имеется выпуклое множество $M \subseteq X$ и

$$\operatorname{gf} a_1 = \operatorname{gf} a \cap (M \times Y) = \{(x, y)\colon \varphi(x, y) \leqslant 0,\ x \in M\}.$$

Тогда для $z_0 = (x_0, y_0) \in \operatorname{gf} a$

$$K_{a_1}(z_0) = K_a(z_0) \cap \operatorname{con}((M - x_0) \times (Y - y_0)).$$

Если существует точка (x_1, y_1) такая, что $x_1 \in M$, $\varphi(x_1, y_1) < 0$, то

$$\operatorname{int} \operatorname{gf} a \cap (M \times Y) \neq \varnothing,$$

и поэтому

$$\operatorname{int} K_a(z_0) \cap \operatorname{con}((M - x_0) \times (Y - y_0)) \neq \varnothing.$$

По теореме I.3.2

$$K^*_{a_1}(z_0) = K^*_a(z_0) + [\operatorname{con}((M - x_0) \times (Y - y_0))]^*,$$

но легко убедиться, что

$$[\operatorname{con}((M - x_0) \times (Y - y_0))]^* = [\operatorname{con}(M - x_0)]^* \times \{0\}.$$

Поэтому согласно предыдущему, если $\varphi(x_0, y_0) = 0$, то

$$K^*_{a_1}(z_0) = -\operatorname{con} \partial_z \varphi(z_0) + [\operatorname{con}(M - x_0)]^* \times \{0\}$$

и $(-x^*, y^*) \in K^*_{a_1}(z_0)$, т. е. $x^* \in a_1^*(y^*; z_0)$ тогда и только тогда, когда

$$x^* = \lambda x_0^* - x_1^*, \qquad y^* = -\lambda y_0^*,$$

$$(x_0^*, y_0^*) \in \partial_z \varphi(z_0), \qquad x_1^* \in [\operatorname{con}(M - x_0)]^*, \qquad \lambda \geqslant 0.$$

Отсюда получаем, что справедлива следующая

Теорема 3.5. *Пусть* $\varphi(z)$ — *выпуклая функция,* $\varphi(z_0) = 0$, $\varphi(z)$ *непрерывна в точке* z_0 *и существует такая точка* $z_1 = (x_1, y_1)$, *что* $\varphi(z_1) < 0$, $x_1 \in M$, *где* M — *выпуклое множество в* X. *Тогда*

$$a_1^*(y^*; z_0) = \{\lambda x_0^* - x_1^*\colon y^* = -\lambda y_0^*,\ (x_0^*, y_0^*) \in \partial_z \varphi(z_0),\ x_1^* \in [\operatorname{con}(M - x_0)]^*,\ \lambda \geqslant 0\}.$$

Пример 3.7. Пусть $X = \mathbf{R}^n$, $Y = \mathbf{R}^m$, A и B — матрицы размеров $r \times n$ и $r \times m$ соответственно, а d — r-мерный вектор. Будем обозначать через A_i, B_i i-е строки матриц A и B, а через d_i — i-ю компоненту вектора d. Положим

$$\operatorname{gf} a \doteq \{(x, y): Ax - By \leqslant d\}, \tag{3.12}$$

где неравенство для векторов понимается как система неравенств для компонент. Многозначное отображение, график которого задан формулой (3.12), называется *многогранным.*

Пусть $z_0 = (x_0, y_0) \in \operatorname{gf} a$. Обозначим

$$I_0 = \{i: A_i x_0 - B_i y_0 = d_i,\ i = 1, \ldots, r\}.$$

Чтобы вычислить конус $K_a(z_0)$, посмотрим, для каких $\bar{z} = (\bar{x}, \bar{y}) \in Z$ существует достаточно малое $\lambda > 0$ такое, что $z_0 + \lambda\bar{z} \in \operatorname{gf} a$. Если $i \in I_0$, то неравенство

$$A_i(x_0 + \lambda\bar{x}) - B_i(y_0 + \lambda\bar{y}) = d_i + \lambda(A_i\bar{x} - B_i\bar{y}) \leqslant d_i$$

возможно лишь, если

$$A_i\bar{x} - B_i\bar{y} \leqslant 0, \qquad i \in I_0. \tag{3.13}$$

Если $i \notin I_0$, то неравенство

$$A_i(x_0 + \lambda\bar{x}) - B_i(y_0 + \lambda\bar{y}) = (A_i x_0 - B_i y_0) + \lambda(\bar{A}_i x - \bar{B}_i y) < d_i$$

справедливо при малых λ независимо от выбора $\bar{z}$. Отсюда следует, что конус $K_a(z_0)$ полностью определяется неравенствами (3.13).

Применим теперь теорему I.4.9 к системе (3.13), переписанной в виде

$$\langle\bar{x}, -A_i\rangle + \langle\bar{y}, B_i\rangle \geqslant 0, \qquad i \in I_0.$$

Это применение показывает, что $(x^*, y^*) \in K_a^*(z_0)$ тогда и только тогда, когда

$$x^* = -\sum_{i \in I_0} A_i^* \lambda_i, \quad y^* = \sum_{i \in I_0} B_i^* \lambda_i, \qquad \lambda_i \geqslant 0,$$
$$i = 1, \ldots, m, \tag{3.14}$$

где A_i^*, B_i^* — транспонированные векторы A_i и B_i, т. е. вектор-столбцы. Если положить $\lambda_i = 0$, $i \notin I_0$, и обозначить через λ столбец с компонентами λ_i, то формулы

(3.14) можно переписать в эквивалентном виде:

$$K_a^*(z_0) = \{(x^*, y^*):\ x^* = -A^*\lambda,\quad y^* = B^*\lambda,\quad \lambda \geqslant 0,\\ \langle Ax_0 - By_0 - d, \lambda\rangle = 0\}.$$

Отсюда для $a^*(y^*;\ z_0)$ получается формула

$$a^*(y^*;\ z_0) = \\ = \{A^*\lambda:\ y^* = B^*\lambda,\ \lambda \geqslant 0,\ \langle Ax_0 - By_0 - d,\ \lambda\rangle = 0\}. \quad (3.15)$$

Заметим, что из (3.14) и (3.15) следует, что $a^*(y^*; z_0)$ зависит не от точки z_0, а только от множества

$$I(z_0) = \{i:\ A_i x_0 - B_i y_0 = d_i\}.$$

Так как таких множеств лишь конечное число, то и различных отображений $a^*(y^*;\ z_0)$ лишь конечное число.

Исследуем функцию $W_a(x,\ y^*)$. Так как при фиксированном x $a(x)$ есть также многогранное множество, то нижняя грань $\langle y,\ y^*\rangle$ по $y \in a(x)$ либо равна $-\infty$, либо достигается. Этот факт будет доказан в следующей главе, но он легко следует из представления многогранного множества по формулам (I.4.24), (I.4.25). Очевидно поэтому, что

$$(x,\ x^0) \in \operatorname{epi} W_a(\cdot,\ y^*),\ x^0 \in \mathbf{R}^1$$

тогда и только тогда, когда при некотором y

$$\langle y,\ y^*\rangle \leqslant x^0,\quad y \in a(x),$$

т. е. вектор $(x,\ x^0,\ y)$ удовлетворяет системе линейных неравенств

$$Ax - By \leqslant d,\qquad \langle y,\ y^*\rangle \leqslant x^0. \quad (3.16)$$

Согласно теореме I.4.15 существуют такие векторы $(x_j,\ x_j^0,\ y_j)$, $j = 1, \ldots, k$, и подмножество индексов $J_0 \subseteq \{1,\ 2,\ \ldots,\ k\}$, что любое решение системы (3.16) и только такие решения можно представить в виде

$$x = \sum_{j=1}^{k} \gamma_j x_j, \quad (3.17)$$

$$x^0 = \sum_{j=1}^{k} \gamma_j x_j^0, \quad (3.18)$$

$$y = \sum_{j=1}^{k} \gamma_j y_j, \tag{3.19}$$

$$\gamma_j \geqslant 0, \qquad \sum_{j \in J_0} \gamma_j = 1. \tag{3.20}$$

Ясно, что $\operatorname{epi} W_a(\cdot, y^*)$ есть множество точек, которые можно записать в виде (3.17), (3.18) при условиях (3.20). Поэтому на основании теоремы I.4.15 $\operatorname{epi} W_a(\cdot, y^*)$ есть многогранное множество, а так как такое множество может быть описано системой линейных неравенств, то оно замкнуто.

Таким образом, доказана

Теорема 3.6. *Для многогранного отображения $W_a(x, y^*)$ есть замкнутая функция x, а сопряженное отображение дается формулой* (3.15) *и является кусочно постоянным по аргументу z_0.*

Пример 3.8. Пусть K — выпуклый замкнутый конус в $Z = X \times Y$ и $\operatorname{gf} a = K$. Тогда

$$\Omega_a(x^*, y^*) = \inf_{(x,y)} \{-\langle x, x^*\rangle + \langle y, y^*\rangle \colon (x, y) \in K\} =$$
$$= \begin{cases} 0, & \text{если } (-x^*, y^*) \in K^*, \\ \infty, & \text{если } (-x^*, y^*) \notin K^*. \end{cases}$$

Далее, для $\lambda > 0$

$$W_a(\lambda x, y^*) = \inf_y \{\langle y, y^*\rangle \colon (\lambda x, y) \in K\} =$$
$$= \lambda \inf_y \left\{\left\langle \frac{y}{\lambda}, y^* \right\rangle \colon \left(x, \frac{y}{\lambda}\right) \in \frac{1}{\lambda} K\right\}.$$

Но K — конус и поэтому $\lambda^{-1}K = K$. Если обозначить $y_1 = \lambda^{-1}y$, то окончательно получаем

$$W_a(\lambda x, y^*) = \lambda \inf_{y_1} \{\langle y_1, y^*\rangle \colon (x, y_1) \in K\} =$$
$$= \lambda W_a(x, y^*),$$

т. е. $W_a(x, y^*)$ — положительно однородная выпуклая функция относительно x.

Пусть теперь $(x_0, y_0) \in K$. Так как $K_a(z_0) = \operatorname{con}(K - z_0)$, то $(-x^*, y^*) \in K_a^*(z_0)$ согласно теореме I.3.8 тогда и

только тогда, когда
$$-\langle x-x_0,\ x^*\rangle+\langle y-y_0,\ y^*\rangle\geqslant 0$$
при всех $(x,\ y)\in K$. Переписав последнее неравенство в виде
$$-\langle x,\ x^*\rangle+\langle y,\ y^*\rangle\geqslant-\langle x_0,\ x^*\rangle+\langle y_0,\ y^*\rangle,$$
по лемме I.3.8 получаем, что $(-x^*,\ y^*)\in K^*$ и нижняя грань левой части последнего неравенства по $(x,\ y)\in K$ равна нулю. Поэтому
$$0\geqslant-\langle x_0,\ x^*\rangle+\langle y_0,\ y^*\rangle.$$
С другой стороны, так как $(x_0,\ y_0)\in K$, а $(-x^*,\ y^*)\in K^*$, то
$$0\leqslant-\langle x_0,\ x^*\rangle+\langle y_0,\ y^*\rangle,$$
т. е. окончательно получаем $\langle x_0,\ x^*\rangle=\langle y_0,\ y^*\rangle$.

Таким образом,
$$a^*(y^*;\ z_0)=\{x^*:\ (-x^*,\ y^*)\in K^*,\ \langle x_0,\ x^*\rangle=\langle y_0,\ y^*\rangle\}.$$

§ 4. Теорема двойственности для выпуклых многозначных отображений

В этом параграфе будет доказан результат, который можно рассматривать как теорему двойственности для многозначных отображений. Как следствие этого результата могут быть получены многие другие теоремы выпуклого анализа и теории экстремальных задач.

1. Теорема двойственности.

Теорема 4.1. *Пусть a — выпуклое многозначное отображение и при фиксированном y^* функция $W_a(x,\ y^*)$ замкнута и для некоторого x_1 $W_a(x_1,\ y^*)$ конечна. Тогда*
$$\sup_{x^*}\{\Omega_a(x^*,\ y^*)+\langle x,\ x^*\rangle\}=W_a(x,\ y^*).$$

Доказательство. Пусть $W_a(x_0,\ y^*)\neq+\infty$. Обозначим $\gamma_0=W_a(x_0,\ y^*)$. Тогда для любого $\varepsilon>0$ существует выпуклая окрестность U точки x_0 такая, что множество
$$\Gamma_\varepsilon=\{(x,\ y):\ x\in U,\quad \langle y,\ y^*\rangle\leqslant\gamma_0-\varepsilon\}$$
не пересекается с $\operatorname{graf} a$, т. е. $\Gamma_\varepsilon\cap\operatorname{graf} a=\varnothing$. Это есть прямое следствие замкнутости $W_a(x,\ y^*)$.

Далее, ясно, что Γ_ε выпукло. Поэтому по теореме об отделимости выпуклых множеств существует такая пара векторов x_0^*, y_0^*, что они одновременно не равны нулю и

$$-\langle x, x_0^*\rangle + \langle y, y_0^*\rangle \geqslant -\langle \xi, x_0^*\rangle + \langle \eta, y_0^*\rangle,$$
$$(x, y) \in \operatorname{graf} a, \qquad (\xi, \eta) \in \Gamma_\varepsilon. \tag{4.1}$$

Положив $x = \xi = x_0$, получим

$$\langle y, y_0^*\rangle \geqslant \langle \eta, y_0^*\rangle \tag{4.2}$$

для $y \in a(x_0)$, $\langle \eta, y^*\rangle \leqslant \gamma_0 - \varepsilon$.

Покажем, что $y_0^* = \lambda y^*$ и $\lambda > 0$. Действительно, из (4.2) следует, что неравенство $\langle \bar\eta, y^*\rangle \geqslant 0$ влечет неравенство $\langle \bar\eta, y_0^*\rangle \geqslant 0$. На основании теоремы I.4.9 отсюда следует, что $y_0^* = \lambda y^*$.

Докажем, что $\lambda > 0$. Если $\lambda = 0$, то $y_0^* = \lambda y^* = 0$ и (4.1) примет при $x = x_0$ вид

$$-\langle x_0, x_0^*\rangle \geqslant -\langle \xi, x_0^*\rangle, \qquad \xi \in U. \tag{4.3}$$

Отсюда

$$\langle \xi - x_0, x_0^*\rangle \geqslant 0, \qquad \xi \in U,$$

а поэтому $x_0^* \equiv 0$. Но это невозможно, так как x_0^* и y_0^* одновременно нулю не равны.

Итак, $\lambda > 0$, и поэтому можно считать, что $\lambda = 1$, $y_0^* = y^*$. Неравенство (4.1) приобретает вид

$$-\langle x, x_0^*\rangle + \langle y, y^*\rangle \geqslant -\langle \xi, x_0^*\rangle + \langle \eta, y^*\rangle,$$

или

$$\Omega_a(x_0^*, y^*) \geqslant -\langle x_0, x^*\rangle + \gamma_0 - \varepsilon. \tag{4.4}$$

Таким образом, для данного ε нашелся такой вектор x_0^*, что при $\xi = x_0$

$$\Omega_a(x_0^*, y^*) + \langle x_0, x_0^*\rangle \geqslant W_a(x_0, y_0^*) - \varepsilon,$$

или в силу произвольности $\varepsilon > 0$

$$\sup_{x^*} \{\Omega_a(x^*, y^*) + \langle x_0, x^*\rangle\} \geqslant W_a(x_0, y^*).$$

Сравнивая это с (2.4), получаем то, что требовалось.

Теперь рассмотрим случай, когда $W_a(x_0, y^*) = +\infty$, т. е. $a(x_0) = \varnothing$. Пусть γ_0 — любое число. Тогда существует окрестность U точки x_0 такая, что множества Γ_ε и $\operatorname{graf} a$ не пересекаются и снова получается (4.1). Аналогично, $y_0^* = \lambda y^*$. Здесь есть две возможности.

а) Для всех γ_0 $\lambda > 0$, т. е. можно положить $y_0^* = y^*$. Тогда

$$-\langle x, x_0^*\rangle + \langle y, y^*\rangle \geqslant -\langle x_0, x_0^*\rangle + \gamma_0 - \varepsilon,$$

$$\Omega_a(x_0^*, y^*) + \langle x_0, x_0^*\rangle \geqslant \gamma_0 - \varepsilon,$$

$$\sup_{x^*} [\Omega_a(x^*, y^*) + \langle x_0, x^*\rangle] \geqslant \gamma_0 - \varepsilon,$$

и так как γ_0 произвольно, то

$$\sup_{x^*} [\Omega_a(x^*, y^*) + \langle x_0, x^*\rangle] = +\infty.$$

б) Для некоторого γ_0 $\lambda = 0$, т. е. $y_0^* = 0$, и мы имеем

$$-\langle x, x_0^*\rangle \geqslant -\langle \xi, x_0^*\rangle, \quad (x, y) \in \operatorname{graf} a, \qquad \xi \in U,$$

или, так как $x_0^* \neq 0$,

$$-\langle x, x_0^*\rangle \geqslant -\langle x_0, x_0^*\rangle - \inf_{\xi \in U} \langle \xi - x_0, x_0^*\rangle \geqslant$$

$$\geqslant -\langle x_0, x_0^*\rangle + \varepsilon, \qquad \varepsilon > 0,$$

$$-\langle x - x_0, x_0^*\rangle \geqslant \varepsilon, \qquad x \in \operatorname{dom} a.$$

Если теперь взять точку x_1, в которой $W_a(x, y^*)$ конечна, то, воспользовавшись для нее проведенными ранее рассуждениями, получим, что найдется такой вектор x_1^*, что выполняется неравенство (4.4) с заменой в нем x_0, x_0^* на x_1, x_1^*. Поэтому $\Omega_a(x_1, y^*)$ конечно. Для $\rho > 0$, учитывая последнее неравенство, получаем

$$\Omega_a(x_1^* + \rho x_0^*, y^*) + \langle x_0, x_1^* + \rho x_0^*\rangle =$$

$$= \inf_x \{-\langle x, x_1^* + \rho x_0^*\rangle + W_a(x, y^*)\} + \langle x_0, x_1^*\rangle + \rho\langle x_0, x_0^*\rangle =$$

$$= \inf_x \{-\langle x_0, x_1^*\rangle + W_a(x, y^*) - \rho\langle x - x_0, x_0^*\rangle\} +$$

$$+ \langle x_0, x_1^*\rangle \geqslant \Omega_a(x_1^*, y^*) + \rho\varepsilon + \langle x_0, x_1^*\rangle.$$

Поэтому

$$\sup_{x^*}\{\Omega_a(x^*, y^*) + \langle x_0, x^*\rangle\} \geqslant$$
$$\geqslant \Omega_a(x_1^*, y^*) + \rho\varepsilon + \langle x_0, x_1^*\rangle$$

и, устремляя ρ к $+\infty$, получаем

$$\sup_{x^*}\{\Omega_a(x^*, y^*) + \langle x_0, x^*\rangle\} = +\infty,$$

что и требовалось доказать.

Теорема 4.2. *Если $W_a(x, y^*)$ конечна в точке x_0 и полунепрерывна снизу в этой точке, то*

$$\sup_{x^*}\{\Omega_a(x^*, y^*) + \langle x_0, x^*\rangle\} = W_a(x_0, y^*). \tag{4.5}$$

Действительно, для того чтобы множество Γ_ε при малом ε не пересекалось с $\operatorname{gf} a$, достаточно полунепрерывности $W_a(x, y^*)$ в точке x_0. Дальнейшее доказательство повторяет предыдущее.

Теорема 4.3. *Пусть $x_0 \in \operatorname{int}\operatorname{dom} a$ и $W_a(x_0, y^*) > -\infty$. Тогда существует такой вектор x_0^*, что*

$$\Omega_a(x^*, y^*) + \langle x_0, x^*\rangle \leqslant \Omega_a(x_0^*, y^*) + \langle x_0, x_0^*\rangle = W_a(x_0, y^*).$$

Доказательство. Множество

$$\Gamma_0 = \{(x, y)\colon\ x = x_0,\quad \langle y, y^*\rangle < W_a(x_0, y^*)\}$$

и $\operatorname{gf} a$ не пересекаются по определению $W_a(x_0, y^*)$. Применив теорему отделимости для выпуклых множеств $\operatorname{gf} a$ и Γ_0, получим неравенство (4.1), в котором вместо Γ_ε будет фигурировать Γ_0:

$$-\langle x, x_0^*\rangle + \langle y, y_0^*\rangle \geqslant -\langle \xi, x_0^*\rangle + \langle \eta, y_0^*\rangle,$$
$$(x, y) \in \operatorname{gf} a, \qquad (\xi, \eta) \in \Gamma_0. \tag{4.6}$$

Положив $x = \xi = x_0$, получим

$$\langle y, y_0^*\rangle \geqslant \langle \eta, y_0^*\rangle$$

для $y \in a(x_0)$, $\langle \eta, y^*\rangle < W_a(x_0, y^*)$. Точно так же, как при доказательстве теоремы 4.1, отсюда следует, что $y_0^* = \lambda y^*$, $\lambda \geqslant 0$. Покажем, что $\lambda > 0$. Действительно, если

$\lambda=0$, т. е. $y_0^*=0$, то (4.6) можно переписать в виде

$$-\langle x, x_0^*\rangle \geqslant -\langle x_0, x_0^*\rangle, \qquad x\in \operatorname{dom} a,$$

т. е.

$$-\langle x-x_0, x_0^*\rangle \geqslant 0$$

для $x\in\operatorname{dom} a$, и, в частности, для $x\in x_0+\varepsilon B$ при малом $\varepsilon>0$, так как $x_0\in\operatorname{int}\operatorname{dom} a$. (Напомним, что B — единичный шар в X.) Однако такое неравенство возможно лишь, если $x_0^*=0$, т. е. $x_0^*=0$, $y_0^*=0$, в противоречии с тем, что x_0^* и y_0^* не равны нулю одновременно.

Итак, $\lambda>0$ и можно считать, что $\lambda=1$, $y_0^*=y^*$. Теперь (4.6) дает

$$\Omega_a\left(x_0^*, y^*\right) \geqslant -\langle x_0, x_0^*\rangle + W_a(x_0, y^*). \tag{4.7}$$

Однако в силу неравенства (2.4) справедливо противоположное неравенство, поэтому

$$\Omega_a\left(x_0^*, y^*\right)+\langle x_0, x_0^*\rangle = W_a(x_0, y^*),$$

что совместно с неравенством (2.4) доказывает теорему.

Рассмотрим некоторые следствия теоремы 4.1.

Теорема 4.4. *Если $f(x)$ — выпуклая собственная замкнутая функция, то $f(x)=f^{**}(x)$.*

Замечание. Эта теорема повторяет теорему II.2.1. В силу формулы (2.5) теорему 4.1 можно было бы получить как следствие теоремы II.2.1. Однако ввиду важности теоремы 4.1 здесь было проведено ее независимое доказательство.

Доказательство. Согласно примеру 3.3 для отображения

$$a(x)=\{y\in\mathbf{R}^1\colon\ y\geqslant f(x)\}$$

имеем

$$W_a(x, 1)=f(x), \qquad \Omega_a(x^*, 1)=-f^*(x^*).$$

Поэтому согласно теореме 4.1, все условия которой выполнены в силу предположений о $f(x)$, получаем

$$\sup_{x^*}\{\langle x, x^*\rangle - f^*(x^*)\}=f(x),$$

что и требовалось доказать.

Теорема 4.5. *Пусть* $\operatorname{gf} a = K$, *где* K — *выпуклый конус. Тогда, если* $W_a(x, y^*)$ — *собственная замкнутая функция, то*

$$\sup_{x^* \in a^*(y^*)} \langle x, x^* \rangle = \inf_{y \in a(x)} \langle y, y^* \rangle, \tag{4.8}$$

где

$$a^*(y^*) = \{x^* : (-x^*, y^*) \in K^*\}.$$

Доказательство. Согласно примеру 3.8

$$\Omega_a(x^*, y^*) = \begin{cases} 0, & x^* \in a^*(y^*), \\ -\infty, & x^* \overline{\in} a^*(y^*). \end{cases}$$

Поэтому применение теоремы 4.1 дает

$$\sup_{x^*} \{\Omega_a(x^*, y^*) + \langle x, x^* \rangle\} = \\ = \sup_{x^*} \{\langle x, x^* \rangle : x^* \in a^*(y^*)\} = W_a(x, y^*),$$

что и требовалось доказать.

Теорема 4.6. *Если в предыдущей теореме конус* K — *многогранный, то справедливо равенство* (4.8).

Доказательство следует из того, что согласно теореме 3.6 $W_a(x, y^*)$ есть замкнутая функция, если $\operatorname{gf} a$ — многогранное множество.

2. Теоремы о минимаксе. Сделаем несколько предварительных замечаний. Функция $g(x)$ называется *вогнутой*, если $-g(x)$ есть выпуклая функция. Соответственно вогнутая функция *замкнута*, если $-f(x)$ есть замкнутая функция.

Если $f(x)$ — замкнутая функция, то она достигает минимума на компактном множестве M. Действительно, пусть $\mu = \inf_x \{f(x) : x \in M\}$. Рассмотрим множества

$$M_\alpha = \{x : f(x) \leqslant \alpha, \quad x \in M\} = \{x : f(x) \leqslant \alpha\} \cap M, \quad \alpha > \mu.$$

По теореме II.1.5 множества $\{x : f(x) \leqslant \alpha\}$ замкнуты, а поэтому M_α компактны как пересечение компактного и замкнутого множества. Очевидно, что $M_{\alpha_1} \subseteq M_{\alpha_2}$, $\alpha_1 < \alpha_2$, т. е. множества M_α вложены друг в друга, а поэтому их пересечение

$$M_\mu = \bigcap_{\alpha > \mu} M_\alpha$$

не пусто. Но тогда очевидно, что точки множества M_μ доставляют минимум $f(x)$ на M.

Пусть теперь M — выпуклое замкнутое подмножество в X, N — выпуклое подмножество в Y, а $f(x, y)$ — выпуклая по x при фиксированном y и вогнутая по y при фиксированном x функция.

Теорема 4.7. *Если $f(x, y)$ есть собственная замкнутая функция x при фиксированном y, $\operatorname{dom} f(\cdot, y) \cap M \neq \varnothing$ для $y \in N$ и функция*

$$g(x^*) = \inf_{y \in N} \sup_{x \in M} (\langle x, x^* \rangle - f(x, y)) \qquad (4.9)$$

конечна и полунепрерывна снизу при $x^* = 0$, то

$$\inf_{x \in M} \sup_{y \in N} f(x, y) = \sup_{y \in N} \inf_{x \in M} f(x, y). \qquad (4.10)$$

Доказательство. Пусть

$$f_0(x, y) = \begin{cases} f(x, y), & x \in M, \\ +\infty, & x \notin M. \end{cases}$$

Так как при фиксированном y

$$\operatorname{epi} f_0(\cdot, y) = \operatorname{epi} f(\cdot, y) \cap \{(x, x^0) \colon x \in M\},$$

то $\operatorname{epi} f_0(\cdot, y)$ есть пересечение (в силу замкнутости $f(\cdot, y)$ и множества M) выпуклых замкнутых множеств, и поэтому $f_0(x, y)$ выпукла и замкнута по x при фиксированном y. Свойство вогнутости $f_0(x, y)$ по y сохраняется.

В силу условия $\operatorname{dom} f(\cdot, y) \cap M \neq 0$, $y \in N$, $f_0(x, y)$ есть собственная функция при $y \in N$. Очевидно, что (4.10) эквивалентно соотношению

$$\inf_{x \in X} \sup_{y \in N} f_0(x, y) = \sup_{y \in N} \inf_{x \in X} f_0(x, y). \qquad (4.11)$$

Пусть

$$f_0^*(x^*, y) = \sup_x \{\langle x, x^* \rangle - f_0(x, y)\}. \qquad (4.12)$$

Функция $f_0^*(x^*, y)$ выпукла по совокупности аргументов, так как она есть верхняя грань выпуклых функций $\langle x, x^* \rangle - f_0(x, y)$. Последняя функция, очевидно, выпукла по x^*; по y она выпукла, так как $f_0(x, y)$ вогнута по этому аргументу.

Рассмотрим многозначное выпуклое отображение

$$a(x^*) = \{(y, y^0):\ y^0 \geqslant f_0^*(x^*, y),\ y \in N\}.$$

Для этого отображения

$$W_a(x^*, y^*, y^{0*}) = \\ = \inf_{(y, y^0)} \{\langle y, y^*\rangle + y^{0*}y^0:\ y^0 \geqslant f_0^*(x^*, y),\quad y \in N\}.$$

В частности, при $y^* = 0$, $y^{0*} = 1$

$$W_a(x^*, 0, 1) = \inf_y \{f_0^*(x^*, y):\ y \in N\} = g(x^*).$$

Аналогично,

$$\Omega_a(x, 0, 1) = \\ = \inf_{(x^*, y, y^0)} \{-\langle x, x^*\rangle + y^0:\ y^0 \geqslant f_0^*(x^*, y),\quad y \in N\} = \\ = \inf_{(x^*, y)} \{-\langle x, x^*\rangle + f_0^*(x^*, y):\quad y \in N\} = \\ = \inf_{y \in N} \inf_{x^*} \{-\langle x, x^*\rangle + f_0^*(x^*, y)\} = \\ = \inf_{y \in N} (-f_0(x, y)).$$

Здесь было использовано то, что $f_0^*(x^*, y)$ есть (согласно (4.12)) функция, сопряженная к $f_0(x, y)$ относительно x; а так как $f_0(x, y)$ замкнута по x, то в силу теоремы 4.4

$$\inf_{x^*} \{-\langle x, x^*\rangle + f_0^*(x^*, y)\} = \\ = -\sup_{x^*} \{\langle x, x^*\rangle - f_0^*(x^*, y)\} = -f_0(x, y).$$

Так как $g(x^*)$ полунепрерывна в $x^* = 0$, то согласно теореме 4.2

$$\sup_x \left\{\inf_{y \in N} (-f_0(x, y))\right\} = g(0). \tag{4.13}$$

Но

$$g(0) = \inf_{y \in N} \sup_{x \in X} (-f_0(x, y)). \tag{4.14}$$

Сопоставление (4.13) и (4.14) доказывает (4.11), а, значит, и (4.10).

Теорема 4.8. *Пусть N — выпуклое компактное множество в Y, функция $f(x, y)$ есть замкнутая выпуклая собственная функция x при каждом $y \in N$ и замкнутая вогнутая функция y при каждом x, а $g(0) \neq \pm\infty$. Тогда имеет место соотношение*

$$\inf_x \sup_{y \in N} f(x, y) = \sup_{y \in N} \inf_x f(x, y).$$

Доказательство. В рассматриваемом случае $M = X$ и необходимо проверить только полунепрерывность в нуле функции

$$g(x^*) = \inf_{y \in N} f^*(x^*, y),$$

$$f^*(x^*, y) = \sup_x \{\langle x, x^* \rangle - f(x, y)\}.$$

Так как $f(x, y)$ замкнута по y, то $f^*(x^*, y)$ замкнута по x^* и y. Поэтому согласно сказанному в начале этого раздела нижняя грань $f^*(x^*, y)$ по $y \in N$ достигается.

Пусть теперь $x_i^* \to 0$, $g(x_i^*) \to \mu$. Обозначим через y_i то значение $y \in N$, для которого

$$f^*(x_i^*, y_i) = g(x_i^*).$$

Так как $y_i \in N$ и N — компакт, то без ограничения общности можно считать, что $y_i \to y_0 \in N$. Используя это, получаем

$$\mu = \lim_{i \to \infty} g(x_i^*) = \lim_{i \to \infty} f^*(x_i^*, y_i) \geqslant f^*(0, y_0) \geqslant g(0), \qquad (4.15)$$

так как $f^*(x^*, y)$ замкнута по совокупности переменных и

$$f^*(0, y_0) \geqslant \inf\{f^*(0, y) \colon y \in N\} = g(0).$$

Тем самым полунепрерывность $g(x^*)$ в нуле доказана, и это завершает доказательство теоремы.

Глава IV

ВЫПУКЛОЕ ПРОГРАММИРОВАНИЕ

Выпуклое программирование — это область исследований, объектом которых служат экстремальные задачи, связанные с минимизацией выпуклых функций на выпуклых множествах. Это, во-первых, задачи описания точек, в которых выпуклая функция достигает своего минимума на данном множестве, т. е. формулировка необходимых условий экстремума. Во-вторых, описание задач, двойственных к данной задаче, т. е. таких экстремальных задач, нахождение максимума в которых так или иначе тесно связано с решением исходной задачи. Этой проблематике, а также описанию решения некоторых более конкретных вопросов и посвящена данная глава.

§ 1. Линейное программирование

К задачам линейного программирования относятся задачи, в которых ищется минимум линейной функции на множестве, заданном системой линейных неравенств. Таким образом, *задачей линейного программирования* называется задача нахождения минимума целевой функции $\langle x, x_0^* \rangle$ при ограничениях

$$\langle x, x_i^* \rangle - \alpha_i \leqslant 0, \qquad i = 1, \ldots, m, \tag{1.1}$$

или, короче,

$$\inf_x \{\langle x, x_0^* \rangle: \langle x, x_i^* \rangle - \alpha_i \leqslant 0, \quad i = 1, \ldots, m\}. \tag{1.2}$$

В дальнейшем будет предполагаться, что система неравенств (1.1) совместна, т. е. существует по крайней мере одна точка x, ей удовлетворяющая. Множество точек x, удовлетворяющих системе (1.1), будем обозначать через D:

$$D = \{x: \langle x, x_i^* \rangle - \alpha_i \leqslant 0, \quad i = 1, \ldots, m\},$$

и называть *допустимой* областью.

1. Необходимые условия минимума.

Теорема 1.1. *В задаче линейного программирования нижняя грань целевой функции либо достигается и конечна, либо равна $-\infty$.*

Доказательство. Согласно теореме I.4.15 и формулам (I.4.24), (I.4.25) любые решения системы (1.1) и только они представимы в виде

$$x = \sum_{j \in I^+} \gamma_j y_j + \sum_{j \in I^0} \lambda_j x_j, \tag{1.3}$$

$$\sum_{j \in I^+} \gamma_j = 1, \quad \gamma_j \geqslant 0, \quad \lambda_j \geqslant 0, \tag{1.4}$$

где $I^0 \cup I^+ = \{1, \ldots, m\}$. Поэтому для любого $x \in D$ выполняется равенство

$$\langle x, x_0^* \rangle = \sum_{j \in I^+} \gamma_j \langle y_j, x_0^* \rangle + \sum_{j \in I^0} \lambda_j \langle x_j, x_0^* \rangle \tag{1.5}$$

и задача линейного программирования сводится к минимизации выражения в правой части соотношения (1.5) по γ_j и λ_j, удовлетворяющим ограничениям (1.4). Если $\langle x_j, x_0^* \rangle < 0$ для некоторого $j \in I^0$, то, устремляя соответствующее λ_j к $+\infty$, можно добиться неограниченного убывания $\langle x, x_0^* \rangle$. Если же

$$\langle x_j, x_0^* \rangle \geqslant 0, \qquad j \in I^0, \tag{1.6}$$

то при нахождении точной нижней грани $\langle x, x_0^* \rangle$ нужно положить все $\lambda_j \in I^0$ равными нулю. Поэтому

$$\min_x \{\langle x, x_0^* \rangle : x \in D\} =$$

$$= \min_{\gamma_j} \left\{ \sum_{j \in I^+} \gamma_j \langle y_j, x_0^* \rangle : \sum_{j \in I^+} \gamma_j = 1, \quad \gamma_j \geqslant 0 \right\} =$$

$$= \min_{j \in I^+} \langle y_j, x_0^* \rangle \tag{1.7}$$

и минимум достигается на точке y_{j_0}, для которой

$$\langle y_{j_0}, x_0^* \rangle = \min_{j \in I^+} \langle y_j, x_0^* \rangle.$$

Согласно теореме I.4.15 и формулам (1.3), (1.4)

$$D = M_D + K_D,$$

где M_D — многогранник, а K_D — многогранный конус. А по теоремам I.4.6 и I.4.3 точки y_j в соотношении (1.3) можно считать крайними точками многогранника M_D.

Теорема 1.2. *Множество решений задачи линейного программирования есть множество точек x, представимых в виде*

$$x = \sum_{j \in I_0^+} \gamma_j y_j + \sum_{j \in I_0^0} \lambda_j x_j,$$
$$\sum_{j \in I_0^+} \gamma_j = 1, \quad \gamma_j \geqslant 0, \quad \lambda_j \geqslant 0, \tag{1.8}$$

где

$$I_0^+ = \left\{ j \in I^+ \colon \langle y_j, x_0^* \rangle = \min_{k \in I^+} \langle y_k, x_0^* \rangle \right\},$$
$$I_0^0 = \{ j \in I^0 \colon \langle x_j, x_0^* \rangle = 0 \}.$$

Доказательство. Легко проверить, что для любой точки x, представимой в виде (1.8), выполняется соотношение

$$\langle x, x_0^* \rangle = \min_{j \in I^+} \langle y_j, x_0^* \rangle,$$

т. е. действительно в силу равенства (1.7) в этой точке достигается минимум. Так же просто из соотношения (1.5) вытекает, что

$$\langle x, x_0^* \rangle > \min_{j \in I^+} \langle y_j, x_0^* \rangle,$$

если точка x не представима в виде (1.8).

Теорема 1.3. *Если множество D ограничено, то D есть многогранник, минимум целевой функции достигается на некотором множестве его крайних точек и любая точка минимума есть выпуклая комбинация этих крайних точек.*

Теорема непосредственно вытекает из предыдущей с учетом того, что $D = M_D$, $K_D = \{0\}$ и $x_j = 0$, $j \in I^0$.

Следующая теорема полностью характеризует те точки, в которых достигается минимум в задаче линейного программирования.

Теорема 1.4. *Пусть x_0 — точка минимума задачи линейного программирования. Тогда существуют такие* *числа* λ_i, $i = 1, \ldots, m$, *что*

$$\begin{gathered} x_0^* + \sum_{i=1} \lambda_i x_i^* = 0, \\ \lambda_i\left(\langle x_0, x_i^*\rangle - \alpha_i\right) = 0, \qquad \lambda_i \geqslant 0, \quad i = 1, \ldots, m. \end{gathered} \tag{1.9}$$

Обратно, если выполнены условия (1.9), *то* $x_0 \in D$ — *точка минимума задачи линейного программирования.*

Доказательство. Пусть x_0 — точка минимума. Рассмотрим конус

$$K_D(x_0) = \{\bar{x} : x_0 + \lambda\bar{x} \in D \text{ при малых } \lambda > 0\}.$$

Обозначим

$$I_0 = \{i: \langle x_0, x_i^*\rangle - \alpha_i = 0, \qquad i = 1, \ldots, m\}.$$

Если $\bar{x} \in K_D(x_0)$, то при достаточно малых $\lambda > 0$ должны выполняться неравенства

$$\langle x_0 + \lambda\bar{x}, x_i^*\rangle - \alpha_i = \langle x_0, x_i^*\rangle - \alpha_i + \lambda\langle\bar{x}, x_i^*\rangle \leqslant 0.$$

Но легко видеть, что эти неравенства выполняются тогда и только тогда, когда

$$\langle\bar{x}, x_i^*\rangle \leqslant 0, \qquad i \in I_0,$$

так что

$$K_D(x_0) = \{\bar{x}: \langle\bar{x}, -x_i^*\rangle \geqslant 0, \qquad i \in I_0\}. \tag{1.10}$$

Так как x_0 — точка минимума и $x_0 + \lambda\bar{x} \in D$ при достаточно малых $\lambda > 0$, если $\bar{x} \in K_D(x_0)$, то

$$\langle x_0 + \lambda\bar{x}, x_0^*\rangle \geqslant \langle x_0, x_0^*\rangle,$$

т. е.

$$\langle\bar{x}, x_0^*\rangle \geqslant 0, \qquad \bar{x} \in K_D(x_0).$$

Иными словами, $x_0^* \in K_D^*(x_0)$. Но тогда согласно теореме I.4.9 и формуле (1.10) x_0^* можно представить в виде

$$x_0^* = -\sum_{i \in I_0} \lambda_i x_i^*, \qquad \lambda_i \geqslant 0, \quad i \in I_0.$$

Положив $\lambda_i = 0$, $i \notin I_0$, с учетом определения I_0 получим формулы (1.9).

Пусть теперь x_0 принадлежит D и выполнены условия (1.9). Тогда для любой точки $x \in D$ получаем

$$\langle x, x_0^* \rangle \geqslant \langle x, x_0^* \rangle + \sum_{i=1}^{m} \lambda_i \left(\langle x, x_i^* \rangle - \alpha_i \right) =$$

$$= \left\langle x, x_0^* + \sum_{i=1}^{m} \lambda_i x_i^* \right\rangle - \sum_{i=1}^{m} \lambda_i \alpha_i = -\sum_{i=1}^{m} \lambda_i \alpha_i. \qquad (1.11)$$

В то же время в силу условий (1.9)

$$\langle x_0, x_0^* \rangle = \langle x_0, x_0^* \rangle + \sum_{i=1}^{m} \lambda_i \left(\langle x_0, x_i^* \rangle - \alpha_i \right) =$$

$$= -\sum_{i=1}^{m} \lambda_i \alpha_i, \qquad (1.12)$$

т. е. $\langle x, x_0^* \rangle \geqslant \langle x_0, x_0^* \rangle$ для всех $x \in D$, и, следовательно, x_0 — точка минимума.

2. Двойственная задача. Пусть вектор $y^* \in \mathbf{R}^m$ имеет компоненты λ_i, $i = 1, \ldots, m$. Обозначим

$$\varphi(y^*) = -\sum_{i=1}^{m} \lambda_i \alpha_i,$$

$$D_* = \left\{ y^* \colon x_0^* + \sum_{i=1}^{m} \lambda_i x_i^* = 0, \quad \lambda_i \geqslant 0, \quad i = 1, \ldots, m \right\}. \qquad (1.13)$$

Задача максимизации функции $\varphi(y^*)$ на множестве D_* называется *двойственной задачей линейного программирования.*

Теорема 1.5. *Если $x \in D$, $y^* \in D_*$, то $\langle x, x_0 \rangle \geqslant \geqslant \varphi(y^*)$.*

Доказательство следует из формулы (1.11).

Теорема 1.6. *Если нижняя грань целевой функции в задаче линейного программирования равна $-\infty$, то ограничения двойственной задачи несовместны, т. е. $D_* = \varnothing$.*

Действительно, из предыдущей теоремы следует, что если $D_* \neq \varnothing$, то $\langle x, x_0^* \rangle$ ограничено снизу на D.

Теорема 1.7. *Если исходная задача линейного программирования имеет решение x_0, то двойственная задача имеет решение y_0^* и $\langle x_0, x_0^* \rangle = \varphi(y_0^*)$.*

Доказательство. В самом деле, если в качестве вектора y_0^* взять вектор с компонентами λ_i, которые фигурируют в теореме 1.4, то из равенства (1.12) и теоремы 1.5 следует, что

$$\varphi(y_0^*) = \langle x_0, x_0^* \rangle \geqslant \varphi(y^*), \quad y^* \in D_*,$$

т. е. y_0^* — решение двойственной задачи и выполняется требуемое теоремой равенство.

Теорема 1.8. *Если $x \in D$, $y^* \in D_*$, то x является решением исходной, а y^* — двойственной задачи тогда и только тогда, когда*

$$\lambda_i\big(\langle x, x_i^* \rangle - \alpha_i\big) = 0, \qquad i = 1, \ldots, m. \tag{1.14}$$

Доказательство непосредственно следует из соотношения (1.11) и теоремы 1.7.

3. Разрешимость линейных неравенств. Рассмотрим систему линейных неравенств

$$\langle x, x_i^* \rangle \leqslant 0, \qquad i = 1, \ldots, m. \tag{1.15}$$

Обозначим через K_i полупространство, определяемое i-м неравенством (1.15). Очевидно, что $\operatorname{int} K_i$ состоит из точек, для которых $\langle x, x_i^* \rangle$ строго меньше нуля, а $\operatorname{Lin} K_i = \mathbf{R}^n$.

Теорема 1.9. *Система неравенств*

$$\langle x, x_i^* \rangle < 0, \qquad i = 1, \ldots, m, \tag{1.16}$$

разрешима тогда и только тогда, когда не существует таких $\lambda_i \geqslant 0$, $i = 1, \ldots, m$, не всех равных нулю, что

$$\sum_{i=1}^{m} \lambda_i x_i^* = 0. \tag{1.17}$$

Доказательство. Рассмотрим вопрос об неотделимости введенных конусов K_i. Воспользуемся теоремой I.3.5. Так как $\operatorname{Lin} K_i = \mathbf{R}^n$, то условие б) этой теоремы выполняется автоматически. Поэтому конусы неотделимы тогда и только тогда, когда пересечение $\operatorname{int} K_i$, $i = 1, \ldots, m$, не пусто. Но это, очевидно, эквивалентно совместности системы (1.16). Если теперь учесть, что

$$K_i^* = \{-\lambda_i x_i^*\colon \lambda_i \geqslant 0\}$$

и определение I.3.3 отделимости, то получаем, что система (1.16) совместна тогда и только тогда, когда равенство (1.17) невозможно.

Теорема 1.10. *Система неравенств*

$$\begin{aligned} &\langle x, x_i^* \rangle \leqslant 0, \qquad i = 1, \ldots, m, \\ &\langle x, x^* \rangle < 0 \end{aligned} \tag{1.18}$$

совместна тогда и только тогда, когда не существует таких $\lambda_i \geqslant 0$, $i = 1, \ldots, m$, *что*

$$x^* + \sum_{i=1}^{m} \lambda_i x_i^* = 0. \tag{1.19}$$

Доказательство. Очевидно, что система (1.18) не совместна тогда и только тогда, когда из системы (1.15) следует, что произведение $\langle x, x^* \rangle$ неотрицательно, т. е. когда $x^* \in K^*$, где K — конус, определяемый системой (1.15). Но по теореме I.4.9

$$x^* = -\sum_{i=1}^{m} \lambda_i x_i^*, \qquad \lambda_i \geqslant 0, \quad i = 1, \ldots, m,$$

т. е. справедливо соотношение (1.19). Итак, система (1.18) несовместна тогда и только тогда, когда (1.19) выполняется, что доказывает теорему.

Теорема 1.11. *Система* (1.15) *имеет единственное решение* $x = 0$ *тогда и только тогда, когда для любого* $x^* \in \mathbf{R}^n$ *существуют такие* $\lambda_i \geqslant 0$, *что выполняется равенство* (1.19).

Это — элементарное следствие теоремы 1.10.

Рассмотрим теперь неоднородные системы неравенств

$$\langle x, x_i^* \rangle - \alpha_i \leqslant 0, \qquad i = 1, \ldots, m. \tag{1.20}$$

Легко видеть, что вопросы о разрешимости системы (1.20) сводятся к исследованию аналогичных вопросов для системы

$$\begin{aligned} &\langle x, x_i^* \rangle - \alpha_i x^0 \leqslant 0, \qquad i = 1, \ldots, m, \\ &-x^0 < 0, \end{aligned} \tag{1.21}$$

в которой неизвестным является вектор $(x, x^0) \in \mathbf{R}^{n+1}$.

Теорема 1.12. *Система*

$$\langle x, x_i^* \rangle - \alpha_i < 0, \qquad i = 1, \ldots, m,$$

разрешима тогда и только тогда, когда не существует таких чисел $\lambda_i \geqslant 0$, $i = 1, \ldots, m$, *что* λ_i *не все равны нулю и*

$$\sum_{i=1}^{m} \lambda_i x_i^* = 0, \quad \sum_{i=1}^{m} \lambda_i \alpha_i \leqslant 0, \tag{1.22}$$

Доказательство получается применением теоремы 1.9 к системе (1.21) со знаками строгих неравенств.

Теорема 1.13. *Система* (1.20) *совместна тогда и только тогда, когда не существует таких* $\lambda_i \geqslant 0$, *что*

$$\sum_{i=1}^{m} \lambda_i x_i^* = 0, \quad \sum_{i=1}^{m} \lambda_i \alpha_i = -1.$$

Доказательство получается прямым применением теоремы 1.10 к системе (1.21).

Рассмотрим теперь вопрос о том, при каких условиях множество решений системы (1.20) ограничено. Пусть K — конус, определяемый неравенствами (1.15). Если K содержит элементы x, отличные от нуля, то множество решений системы (1.20) неограничено. В самом деле, если $x_0 \in K$, $x_0 \neq 0$, и если x удовлетворяет системе (1.20), то $x + \rho x_0$ удовлетворяет системе (1.20) при всех $\rho \geqslant 0$. Обратно, если x_j, $j = 1, \ldots,$ — решения системы (1.20), $\|x_j\| \to \infty$, то, выбирая из последовательности $y_j = \|x_j\|^{-1} x_j$ сходящуюся (без ограничения общности можно считать, что $y_j \to y_0$, $\|y_0\| = 1$), получаем

$$\langle y_j, x_i^* \rangle - \frac{\alpha_i}{\|x_j\|} \leqslant 0, \qquad i = 1, \ldots, m.$$

Отсюда после перехода к пределу получаем систему неравенств

$$\langle y_0, x_i^* \rangle \leqslant 0, \qquad i = 1, \ldots, m.$$

Таким образом, $y_0 \in K$ и $y_0 \neq 0$. Итак, множество решений системы (1.20) ограничено тогда и только тогда, когда $K = \{0\}$.

На основании утверждения теоремы 1.11 убеждаемся в справедливости следующей теоремы.

Теорема 1.14. *Множество решений* (1.20) *ограничено тогда и только тогда, когда для любого* $x^* \in \mathbf{R}^n$ *существуют такие* $\lambda_i \geqslant 0$, *что выполняется равенство* (1.19).

Теорема 1.15. *Пусть система* (1.20) *несовместна. Тогда существует такая ее подсистема*

$$\langle x, x_i^* \rangle - \alpha_i \leqslant 0, \qquad i \in I, \quad I \subseteq \{1, 2, \ldots, m\},$$

которая также несовместна, а число элементов в множестве индексов I *не превосходит* $n+1$.

Доказательство. Пусть система (1.20) неразрешима. Тогда найдутся такие $\lambda_i \geqslant 0$, что

$$\sum_{i=1}^{m} \lambda_i x_i^* = 0, \qquad \sum_{i=1}^{m} \lambda_i \alpha_i = -1. \tag{1.23}$$

Обозначим

$$J = \{i:\ \lambda_i > 0,\ i = 1, \ldots, m\}. \tag{1.24}$$

Если мощность множества J больше $n+1$, то векторы $(x_i^*, \alpha_i) \in \mathbf{R}^{n+1}$ линейно зависимы. Поэтому существуют такие не все равные нулю γ_i, что

$$\sum_{i \in J} \gamma_i x_i^* = 0, \qquad \sum_{i \in J} \gamma_i \alpha_i = 0. \tag{1.25}$$

Положим

$$\varepsilon_0 = \min_{i \in J} \left\{ \frac{\lambda_i}{\gamma_i} :\ \gamma_i > 0 \right\}.$$

Тогда все числа $\overline{\lambda}_i = \lambda_i - \varepsilon_0 \gamma_i$ неотрицательны и из соотношений (1.23)—(1.25) следует

$$\sum_{i \in J} \overline{\lambda}_i x_i^* = 0, \qquad \sum_{i \in J} \overline{\lambda}_i = -1. \tag{1.26}$$

Заметим, что в соответствии с выбором ε_0 одна из величин $\overline{\lambda}_i$, $i \in J$, равна нулю, и поэтому в суммах (1.26) не все слагаемые отличны от нуля. Исключая равные нулю слагаемые и повторяя указанную процедуру несколько раз, придем после конечного числа шагов к ситуации, когда в суммах (1.26) будет не более $n+1$ отличных от нуля слагаемых. Взяв тогда в качестве I множество индексов, отвечающих ненулевым компонентам λ_i, получим на основании теоремы 1.13, что соответствующая подсистема неравенств несовместна, что и доказывает теорему.

Теорема 1.15 показывает, что система (1.20) совместна тогда и только тогда, когда разрешима любая ее подсистема из $n+1$ неравенств.

§ 2. Необходимые условия экстремума в выпуклом программировании

Пусть x_0 — точка минимума функции $f_0(x)$ на множестве M. Очевидно, что точка x_0 должна выделяться своими свойствами среди всех точек $x \in M$. Условия, которые в определенной степени характеризуют точку минимума, называются *необходимыми условиями экстремума*. Если минимизируемая функция и множество M выпуклы, то эти условия оказываются, как правило, и достаточными. Нижеследующие теоремы дают необходимые условия, форма которых варьируется в зависимости от степени детализации описания множества M.

1. Ограничения, заданные выпуклыми множествами. Пусть $f(x)$ — собственная выпуклая функция, а x_0 — точка ее минимума во всем пространстве:

$$f(x) \geqslant f(x_0), \qquad x \in X.$$

Договоримся, что всюду в дальнейшем, если x_0 — точка минимума, то $f(x_0) > -\infty$. Переписав неравенство для $f(x)$ в виде

$$f(x) - f(x_0) \geqslant \langle x - x_0,\ 0\rangle, \qquad 0 \in X^*, \tag{2.1}$$

получаем, что $0 \in \partial f(x_0)$. Таким образом, справедлива

Теорема 2.1. *Для того чтобы точка x_0 была точкой минимума выпуклой функции $f(x)$ во всем пространстве, необходимо и достаточно, чтобы $0 \in \partial f(x_0)$.*

Эта элементарная теорема в сочетании с доказанными ранее позволяет получить ряд новых результатов.

Теорема 2.2. *Пусть M — выпуклое множество и существует точка $x_1 \in M$, в которой выпуклая функция $f(x)$ непрерывна. Для того чтобы точка x_0 была точкой минимума $f(x)$ на множестве M, необходимо и достаточно, чтобы*

$$\partial f(x_0) \cap K_M^*(x_0) \neq \varnothing, \tag{2.2}$$

где

$K_M(x_0) = \mathrm{con}(M - x_0) = \{\bar{x}: x_0 + \lambda\bar{x} \in M$ *при малых* $\lambda > 0\}$.

Замечание. Так как конус $K_M(x_0)$ будет часто встречаться в этом параграфе, рекомендуем читателю вспомнить п. 4 § 3 главы I и теоремы I.3.6—I.3.9.

Доказательство. Так как по определению

$$\delta(x \mid M) = \begin{cases} 0, & x \in M, \\ +\infty, & x \notin M, \end{cases}$$

то легко видеть, что нахождение минимума $f(x)$ на множестве M эквивалентно нахождению минимума функции $f(x) + \delta(x|M)$ во всем пространстве. По теореме II.3.8 субдифференциал функции $f(x) + \delta(x|M)$ равен сумме субдифференциалов функций $f(x)$ и $\delta(x|M)$. Но согласно формуле (II.3.25) справедливо равенство

$$\partial\delta(x_0 \mid M) = -K_M^*(x_0). \tag{2.3}$$

Таким образом, субдифференциал функции $f(x) + \delta(x|M)$ в точке x_0 равен $\partial f(x_0) - K_M^*(x_0)$. Согласно теореме 2.1 для того, чтобы точка x_0 была точкой минимума функции $f(x) + \delta(x|M)$, необходимо и достаточно, чтобы

$$0 \in \partial f(x_0) - K_M^*(x_0),$$

что эквивалентно соотношению (2.2). Теорема доказана.

Теорема 2.3. *Пусть выполнены условия предыдущей теоремы и*

$$M = M_1 \cap M_2 \cap \ldots \cap M_k,$$

где M_i, $i = 1, \ldots, k$,— *выпуклые множества, причем*

$$\mathrm{int}\, M_1 \cap \mathrm{int}\, M_2 \cap \ldots \cap \mathrm{int}\, M_{k-1} \cap M_k \neq \varnothing.$$

Тогда для того, чтобы точка x_0 *была точкой минимума функции* $f(x)$ *на множестве* M, *необходимо и достаточно, чтобы нашлись такие точки* $x_i^* \in K_{M_i}^*(x_0)$, $i = 1, \ldots, n$ *и точка* $x^* \in \partial f(x_0)$, *что*

$$x^* = x_1^* + \ldots + x_k^*.$$

Доказательство. Согласно теоремам I.3.7 и I.3.6 из условий теоремы следует, что

$$K_M(x_0) = \bigcap_{i=1}^{k} K_{M_i}(x_0), \tag{2.4}$$

и если $x_2 \in \operatorname{int} M_i$, $i = 1, \ldots, k-1$, $x_2 \in M_k$, то

$$\bar{x}_2 = x_2 - x_0 \in \operatorname{int} K_{M_i}(x_0), \qquad i = 1, \ldots, k-1,$$

$$\bar{x}_2 \in K_{M_k}(x_0).$$

Поэтому

$$\operatorname{int} K_{M_1}(x_0) \cap \ldots \cap \operatorname{int} K_{M_{k-1}}(x_0) \cap K_{M_k}(x_0) \neq \varnothing. \tag{2.5}$$

Из соотношений (2.4), (2.5) и теоремы I.3.2 вытекает, что

$$K_M^*(x_0) = K_{M_1}^*(x_0) + \ldots + K_{M_k}^*(x_0). \tag{2.6}$$

Тогда точка $x^* \in \partial f(x_0) \cap K_M^*(x_0)$, которая существует согласно теореме 2.2, допускает разложение

$$x^* = x_1^* + \ldots + x_k^*, \quad x_i^* \in K_{M_i}^*(x_0), \qquad i = 1, \ldots, k, \tag{2.7}$$

что доказывает теорему.

Теорема 2.4. *Пусть $f(x)$ — собственная выпуклая функция,*

$$M = \bigcap_{i=1}^{k} M_i,$$

где M_i — выпуклые множества, и существует точка $x_1 \in M$, в которой функция $f(x)$ непрерывна. Тогда для того, чтобы точка x_0 была точкой минимума $f(x)$ на множестве M, необходимо, чтобы нашлись такие векторы $x^* \in \partial f(x_0)$, $x_i^* \in K_{M_i}^*(x_0)$ и число λ, равное нулю или единице, что

$$\lambda x^* = x_1^* + \ldots + x_k^*. \tag{2.8}$$

Причем, если $\lambda = 0$, то среди точек $x_1^*, \ldots, x_k^*$ есть хотя бы одна ненулевая. Если же $\lambda = 1$, то эти условия являются достаточными.

Доказательство. Согласно теореме 2.2 существует вектор $x^* \in \partial f(x_0)$ такой, что $x^* \in K_M^*(x_0)$. Так как справедлива формула (2.4), то по теореме I.3.3 возможны два случая.

1) Либо верна формула (2.6), а значит и формула (2.7). Последняя совпадает с равенством (2.8) при $\lambda = 1$. Так как теорема 2.2 давала необходимые и достаточные условия, то в этом случае условие (2.8) является одновременно и достаточным.

2) Либо существуют такие $x_i^* \in K_{M_i}^*(x_0)$, не все равные нулю, что

$$x_1^* + \ldots + x_k^* = 0.$$

В этом случае формула (2.8) справедлива при $\lambda = 0$.

2. Ограничения, заданные выпуклыми неравенствами. В задачах выпуклого программирования область, в которой ищется минимум, часто задается системой выпуклых неравенств. Более точно, пусть $f_i(x)$, $i = 0, \ldots, m$,— выпуклые функции, D — выпуклое множество, $\operatorname{dom} f_i \supset D$, $i = 0, \ldots, m$.

Требуется найти минимум выпуклой функции $f_0(x)$ в области M, где

$$M = \{x\colon\ f_i(x) \leqslant 0, \quad i = 1, \ \ldots, \ m, \quad x \in D\}. \tag{2.9}$$

Рассмотрим многозначное отображение, ставящее в соответствие каждой точке $y \in \mathbf{R}^m$ множество

$$a(y) = \{(x, \ x^0)\colon\ f_i(x) \leqslant y^i, \ i = 1, \ \ldots, \ m, \ f_0(x) \leqslant x^0, \ x \in D\}. \tag{2.10}$$

Так как функции $f_i(x)$ выпуклы и множество D выпукло, то нетрудно убедиться в том, что a — выпуклое отображение из $Y = \mathbf{R}^m$ во множество подмножеств пространства $X \times \mathbf{R}^1$.

Вычислим функцию $W_a(y, x^*, x^{0*})$:

$$W_a(y, x^*, x^{0*}) = \inf_{(x, x^0)} \{\langle x, x^* \rangle + x^0 x^{0*}\colon\ (x, x^0) \in a(y)\}.$$

Если положить $x^* = 0$, $x^{0*} = 1$ и обозначить $W_a(y, 0, 1)$

через $V(y)$, то будем иметь

$$V(y) = \inf_{(x,x^0)} \{x^0\colon f_0(x) \leqslant x^0, f_i(x) \leqslant y^i,\ i=1, \ldots, m,\ x \in D\} =$$
$$= \inf_x \{f_0(x)\colon f_i(x) \leqslant y^i,\ i = 1, \ldots, m,\ x \in D\}. \quad (2.11)$$

В частности, если $y=0$, то $V(0)$ совпадает со значением точной нижней грани функции $f_0(x)$ на множестве M.

Вычислим теперь $\Omega_a(y^*, x^*, x^{0*})$ в случае, когда $x^* = 0$, $x^{0*} = 1$:

$$\Omega_a(y^*, 0, 1) = \inf_{(y,x,x^0)} \{-\langle y, y^*\rangle + x^0\colon (x, x^0) \in a(y)\} =$$
$$= \begin{cases} \inf\limits_x \left\{f_0(x) - \sum\limits_{i=1}^m y^{i*} f_i(x)\colon x \in D\right\}, \\ \qquad\qquad\qquad \text{если } -y^* \geqslant 0, \\ -\infty, \quad \text{если } -y^{i*} < 0 \text{ для некоторого } i = 1, \ldots, m. \end{cases}$$

Введем следующее обозначение:

$$L(x, y^*) = f_0(x) + \sum_{i=1}^m y^{i*} f_i(x). \quad (2.12)$$

Функция $L(x, y^*)$ называется *функцией Лагранжа* рассматриваемой *задачи выпуклого программирования*. Обозначим

$$\varphi(y^*) = \inf_x \{L(x, y^*)\colon x \in D\}. \quad (2.13)$$

Тогда можно записать

$$\Omega_a(y^*, 0, 1) =$$
$$= \begin{cases} \varphi(-y^*), & \text{если } -y^* \geqslant 0, \\ -\infty, & \text{если } -y^{i*} < 0 \text{ для некоторого } i. \end{cases} \quad (2.14)$$

Пусть теперь для некоторого $x_1 \in D$ выполняются неравенства $f_i(x_1) < 0$, $i = 1, \ldots, m$, и $V(0)$ — конечное число. Это означает, что в исходной задаче нахождения нижней грани функции $f_0(x)$ на множестве M эта точная нижняя грань есть конечное число. Согласно лемме III.2.1 функция $V(y) = W_a(y, 0, 1)$ выпукла.

Так как для y из некоторой окрестности точки $y=0$ выполняются неравенства $f_i(x_1) \leqslant y^i$, $i = 1, \ldots, m$, $x_1 \in D$, то $V(y)$ не принимает значения $+\infty$ в этой окрестности и, значит, $y=0$ есть внутренняя точка области определения $V(y)$: $0 \in \operatorname{int} \operatorname{dom} V$. Но в начале § 1 главы II было показано, что если выпуклая функция f принимает значение $-\infty$, то она равна $-\infty$ всюду в $\operatorname{ri} \operatorname{dom} f$. Так как $V(0)$ конечно и $0 \in \operatorname{int} \operatorname{dom} V$, то отсюда следует, что функция $V(y)$ не принимает значений $-\infty$, а, значит, $V(y)$ конечна в некоторой окрестности $y=0$. Поэтому согласно теореме II.1.4 функция $V(y)$ непрерывна в окрестности нуля. Это позволяет применить теорему III.4.3. При этом следует учесть, что для рассматриваемого здесь отображения (2.10) по сравнению со случаем, рассмотренным в теореме III.4.3, пространства X и Y поменялись местами. Применение теоремы III.4.3 показывает, что существует такой вектор y_0^*, что

$$V(0)=W_a(0,\,0,\,1)=\Omega_a\left(y_0^*,\,0,\,1\right)+\langle 0,\,y_0^*\rangle \geqslant$$
$$\geqslant \Omega_a(y^*,\,0,\,1)+\langle 0,\,y^*\rangle$$

где учтено, что в формуле (4.5), фигурирующей в теореме III.4.3, следует заменить x^* на y^*, y^* на $\left(x^*,\,x_0^*\right)=(0,\,1)$, а x_0 на $y=0$. Итак, существует такой вектор y_0^*, что

$$V(0)=\Omega_a\left(y_0^*,\,0,\,1\right)\geqslant \Omega_a(y^*,\,0,\,1).$$

Учитывая формулы (2.12)—(2.14), получаем

$$V(0)=\varphi\left(-y_0^*\right)=$$
$$=\inf_x\left\{L\left(x,\,-y_0^*\right):\ x\in D\right\}\geqslant \varphi(-y^*),\quad -y^*\geqslant 0,\qquad (2.15)$$

причем, так как $\varphi\left(-y_0^*\right)=V(0)$ — конечное число, то $-y_0^*\geqslant 0$.

Теорема 2.5. *Если существует такая точка $x_1\in D$, что $f_i(x_1)<0$ для всех $i=1,\,\ldots,\,m$ и*

$$V(0)=\inf_x\{f_0(x):\ x\in M\}$$

есть конечное число, то существует такой вектор $y_0^*\geqslant 0$, что

$$V(0)=\varphi\left(y_0^*\right)=\inf_x\left\{L\left(x,\,y_0^*\right):\ x\in D\right\}.\qquad (2.16)$$

В частности, если x_0 — точка минимума функции $f_0(x)$ на множестве M, то

$$f_0(x_0)\leqslant L\left(x,\,y_0^*\right),\qquad x\in D.\qquad (2.17)$$

Доказательство получается из формулы (2.15), если заменить в ней $-y^*$ на y^* и $-y_0^*$ на y_0^*. Последнее утверждение теоремы следует из того, что по определению x_0 и $V(0)$ (см. формулу (2.11)) $f_0(x_0)=V(0)$.

Теорема 2.6. *Пусть существует такая точка $x_1 \in D$, что $f_i(x_1)<0$, $i=1, \ldots, m$. Тогда для того, чтобы точка x_0 была точкой минимума функции $f_0(x)$ на множестве M, задаваемом формулой* (2.9), *необходимо и достаточно, чтобы нашелся такой вектор $y_0^* \in \mathbf{R}^m$, чтобы*

$$\begin{aligned} &L(x_0, y_0^*) \leqslant L(x, y_0^*), \qquad x \in D, \\ &y_0^* \geqslant 0, \quad y_0^{i*} f_i(x_0)=0, \qquad i=1, \ldots, m. \end{aligned} \tag{2.18}$$

Доказательство. Согласно предыдущей теореме существует такой вектор $y_0^* \in \mathbf{R}^m$, $y_0^* \geqslant 0$, что выполняется неравенство (2.17). Но так как $x_0 \in M$, то $x_0 \in D$ и $f_i(x_0) \leqslant 0$. Поэтому

$$L(x_0, y_0^*) = f_0(x_0) + \sum_{i=1}^m y_0^{i*} f_i(x_0) \leqslant f(x_0) \leqslant L(x_0, y_0^*).$$

Таким образом,

$$f(x_0) = L(x_0, y_0^*) \tag{2.19}$$

и

$$\sum_{i=1}^m y_0^{i*} f_i(x_0) = 0.$$

Так как $y_0^{i*} \geqslant 0$, $f_i(x_0) \leqslant 0$, то из последнего равенства вытекает

$$y_0^{i*} f_i(x_0) = 0, \qquad i=1, \ldots, m. \tag{2.20}$$

Из равенств (2.20) и соотношений (2.19), (2.17) получается последнее из условий (2.18), чем доказана необходимость условий теоремы.

Покажем их достаточность. Пусть $x \in M$ и условия (2.18) выполнены. Тогда

$$f_0(x) \geqslant f_0(x) + \sum_{i=1}^m y_0^{i*} f_i(x) = \\ = L(x, y_0^*) \geqslant L(x_0, y_0^*) = f_0(x_0)$$

для всех $x \in M$, т. е. x_0 — точка минимума.

Замечание. Достаточность условий теоремы 2.6 показана без предположений о выпуклости функций $f_i(x)$, $i=0, \ldots, m$.

Допустим теперь, что $f_i(x)$, $i=0, \ldots, m$,— непрерывные функции. Первое из условий (2.18) означает, что x_0 — точка минимума функции $L(x, y_0^*)$ по x на выпуклом множестве D. Так как $y_0^* \geqslant 0$ и $f_i(x)$— выпуклые непрерывные функции, то $L(x, y_0^*)$— выпуклая непрерывная функция. Согласно теореме 2.2 функция $L(x, y_0^*)$ достигает своего минимума на D в точке x_0 тогда и только тогда, когда

$$\partial_x L(x_0, y_0^*) \cap K_D^*(x_0) \neq \varnothing.$$

Но по теореме II.3.8

$$\partial_x L(x_0, y_0^*) = \partial f_0(x_0) + \sum_{i=1}^{m} y_0^{i*} \partial f_i(x_0),$$

поэтому существует такой вектор $x^* \in K_D^*(x_0)$, что

$$x^* \in \partial f_0(x_0) + \sum_{i=1}^{m} y_0^{i*} \partial f_i(x_0). \qquad (2.21)$$

Включение (2.21) есть необходимое и достаточное условие того, что точка x_0 доставляет минимум функции $L(x, y_0^*)$ на D. Таким образом, справедлива

Теорема 2.7. *Пусть выполнены условия теоремы 2.6 и $f_i(x)$, $i=0, 1, \ldots, m$,— непрерывные выпуклые функции. Точка x_0 есть точка минимума функции $f_0(x)$ на множестве M тогда и только тогда, когда существуют такие точки $y_0^* \in \mathbf{R}^m$ и $x^* \in K_D^*(x_0)$, что выполнено включение* (2.21) *и*

$$y_0^{i*} f_i(x_0) = 0, \quad y_0^{i*} \geqslant 0, \qquad i = 1, \ldots, m.$$

Определение 2.1. Вектор $y_0^* \geqslant 0$ называется *вектором Куна — Таккера,* если выполняется соотношение

$$\inf_x \{f_0(x)\colon\ x \in M\} = \inf_x \{L(x, y_0^*)\colon\ x \in D\}.$$

Теорема 2.5 дает условия, гарантирующие существование такого вектора. Свяжем теперь условие существования этого вектора со свойствами функции $V(y)$ в окрестности нуля.

Теорема 2.8. *Вектор y_0^* является вектором Куна — Таккера для задачи выпуклого программирования тогда и только тогда, когда $-y_0^* \in \partial V(0)$.*

Доказательство. По определению $-y_0^* \in \partial V(0)$ тогда и только тогда, когда

$$V(y) \geqslant V(0) + \langle y, -y_0^* \rangle,$$

т. е. когда

$$\inf_y \{V(y) + \langle y, y_0^* \rangle\} = V(0).$$

Подставим в эту формулу выражение (2.11) для $V(y)$. Получим

$$\inf_y \inf_x \{f_0(x) + \langle y, y_0^* \rangle\colon\ f_i(x) \leqslant y^i, \quad i = 1, \ldots, m,$$
$$x \in D\} = V(0). \quad (2.22)$$

Нетрудно заметить, что если $y^{i*} < 0$ для некоторого $i = 1, \ldots, m$, то в левой части соотношения (2.22) получим значение, равное $-\infty$. Так как $V(0)$ конечно, то этот случай исключается. Поэтому $y_0^* \geqslant 0$, и левая часть формулы (2.22) приобретает вид

$$\inf_x \left\{f_0(x) + \sum_{i=1}^m f_i(x)\, y_0^{i*}\colon\ x \in D\right\} = \inf_x \{L(x, y_0^*)\colon\ x \in D\}.$$

Итак, $-y_0^* \in \partial_y V(0)$ тогда и только тогда, когда

$$\inf_x \{L(x, y_0^*)\colon\ x \in D\} = V(0),$$

т. е. когда y_0^* есть вектор Куна — Таккера.

Все предыдущее изложение велось в предположении существования такой точки $x_1 \in D$, что $f_i(x_1) < 0$, $i = 1, \ldots, m$. Вместо этого теперь предположим, что функции $f_i(x)$ непрерывны. Рассмотрим множество

$$M_i = \{x\colon\ f_i(x) \leqslant 0\}, \qquad i = 0, 1, \ldots, m.$$

Пусть $x_0 \in M_i$. По теореме II.3.18 возможны следующие случаи:

1) точек x, для которых $f_i(x) < 0$, не существует; тогда согласно теореме 2.1 $0 \in \partial f_i(x_0)$, так как $x_0 \in M$, и, значит, $0 = f_i(x_0) \leqslant f_i(x)$, т. е. x_0 — точка минимума функции $f_i(x)$.

2) существует точка x, в которой $f_i(x) < 0$; тогда

$$K^*_{M_i}(x_0) = \begin{cases} 0, & \text{если} \quad f_i(x_0) < 0, \\ -\operatorname{con} \partial f_i(x_0), & \text{если} \quad f_i(x_0) = 0. \end{cases} \tag{2.23}$$

Пусть x_0 — точка минимума функции $f_0(x)$ на множестве M, где M задано формулой (2.9). Так как

$$M = \left(\bigcap_{i=1}^{m} M_i \right) \cap D,$$

то по теореме 2.4 получаем, что существуют такие точки

$$x_i^* \in K^*_{M_i}(x_0), \; x^* \in K^*_D(x_0), \; x_0^* \in \partial f_0(x_0)$$

и число $\lambda_0 \geqslant 0$, что $\lambda_0 x_i^*$, $i = 1, \ldots, m$, и x^* не равны нулю одновременно и

$$\lambda_0 x_0^* = \sum_{i=1}^{m} x_i^* + x^*. \tag{2.24}$$

Если теперь при некотором x $f_i(x) < 0$ для всех i, то верна формула (2.23). Положим

$$\begin{gathered} x_i^* = -\lambda_i x_{i0}^*, \quad \lambda_i \geqslant 0, \quad x_{i0}^* \in \partial f_i(x_0), \\ x_{00}^* = x_0^*, \end{gathered} \tag{2.25}$$

причем $\lambda_i > 0$, только если $f_i(x_0) = 0$, и $\lambda_i = 0$, если $f_i(x_0) < 0$, т. е.

$$\lambda_i f_i(x_0) = 0, \qquad i = 1, \ldots, m. \tag{2.26}$$

Заменяя $x_0^*, \ldots, x_m^*$ в соотношении (2.24) их выражениями (2.25), получаем

$$\begin{gathered} \lambda_0 x_{00}^* + \lambda_1 x_{10}^* + \ldots + \lambda_m x_{m0}^* = x^*, \\ \lambda^i \geqslant 0, \quad \lambda_i f_i(x_0) = 0, \quad x_{i0}^* \in \partial f_i(x_0), \quad i = 0, \ldots, m, \end{gathered} \tag{2.27}$$

$$x^* \in K^*_D(x_0).$$

При этом среди чисел λ_i есть отличные от нуля, так как не все λ, x_i^*, x^* равны нулю.

Если же для некоторого i_0 точек x, для которых $f_{i_0}(x) < 0$, не существует, то, как сказано выше, $0 \in \partial f_{i_0}(x_0)$. Полагая $x_{i_0 0}^* = 0$, $\lambda_{i_0} = 1$ и $\lambda_i = 0$, $i \neq i_0$, $x^* = 0$, получим, что соотношения (2.27) также выполнены. Тем самым доказана

Теорема 2.9. *Пусть функции $f_i(x)$, $i = 0, \ldots, m$, выпуклы и непрерывны. Тогда для того, чтобы точка x_0 доставляла минимум $f_0(x)$ на множестве*

$$M = \{x\colon f_i(x) \leqslant 0, \quad i = 1, \ldots, m, \quad x \in D\},$$

необходимо, чтобы нашлись не все равные нулю числа λ_i, $i = 1, \ldots, m$, такие, чтобы при некоторых $x_i^* \in \partial f_i(x_0)$ выполнялись соотношения

$$\sum_{i=0}^{m} \lambda_i x_i^* \in K_D^*(x_0),$$

где $\lambda_i \geqslant 0$, $\lambda_i f_i(x_0) = 0$, $i = 0, \ldots, m$.

Теоремы 2.6, 2.7 и 2.9 допускают дальнейшую детализацию, если конкретизировать способ задания множества D.

Теорема 2.10. *Пусть $f_i(x)$, $i = 0, 1, \ldots, m$,— выпуклые собственные функции, причем*

$$f_i(x) = \langle x, x_i^* \rangle - \alpha_i$$

для $i = k+1, \ldots, m$. Пусть D — выпуклое множество, $\operatorname{ri} \operatorname{dom} f_i \supseteq D$, $i = 1, \ldots, m$, и существует такая точка $x_1 \in \operatorname{ri} D$, что $f_i(x_1) < 0$, $i = 1, \ldots, k$. Тогда для того, чтобы точка x_0 была точкой минимума функции $f_0(x)$ на множестве

$$M = \{x\colon f_i(x) \leqslant 0, i = 1, \ldots, m, x \in D\},$$

необходимо и достаточно, чтобы нашелся такой вектор $y_0^* \in R^m$, что выполнены соотношения (2.18).

Доказательство. Пусть

$$\tilde{y}^* \in \mathbf{R}^k, \quad \tilde{y}^* = (y_1^*, \ldots, y_k^*),$$

$$\tilde{L}(x, \tilde{y}^*) = f_0(x) + \sum_{i=1}^{k} y_i^* f_i(x),$$

$$D_0 = \{x\colon f_i(x) \leqslant 0, \quad i = k+1, \ldots, m, \quad x \in D\}.$$

Тогда из теоремы 2.6 следует, что существует такой вектор $\widetilde{y}_0^*$, что

$$\begin{gathered}\widetilde{L}(x_0, \widetilde{y}_0^*) \leqslant \widetilde{L}(x, \widetilde{y}_0^*), \quad x \in D_0, \\ \widetilde{y}_0^* \geqslant 0, \quad y_0^{i*} f_i(x_0) = 0, \quad i = 1, \ldots, k.\end{gathered} \tag{2.28}$$

Так как

$$D \subseteq \bigcap_{i=0}^{m} \operatorname{ri} \operatorname{dom} f_i,$$

то функция $\widetilde{L}(x, \widetilde{y}_0^*)$ непрерывна на D относительно сдвинутого подпространства, содержащего $\bigcap\limits_{i=0}^{m} \operatorname{ri} \operatorname{dom} f_i$. Если вести рассуждения для этого подпространства, то на основании теоремы 2.2 получим, что существует такой вектор $x^* \in \partial_x \widetilde{L}(x_0, \widetilde{y}_0^*)$, что $x^* \in K_{D_0}^*(x_0)$. Но очевидно, что

$$K_{D_0}(x_0) = K_D(x_0) \cap K_{D_1}(x_0), \tag{2.29}$$

где $K_{D_1}(x_0)$ — конус, соответствующий множеству

$$D_1 = \{x\colon f_i(x) = \langle x, x_i^* \rangle - \alpha_i \leqslant 0, \quad i = k+1, \ldots, m\}.$$

Легко подсчитать (так же, как это делалось в доказательстве теоремы 1.4), что

$$K_{D_1}^*(x_0) = \\ = \left\{x^*\colon x^* = -\sum_{i=k+1}^{m} \lambda_i x_i^*, \quad \lambda_i \geqslant 0, \quad \lambda_i f_i(x_0) = 0\right\}. \tag{2.30}$$

Так как точка x_1, фигурирующая в условии теоремы, принадлежит и D_1, то $\operatorname{ri} K_D(x_0) \cap K_{D_1}(x_0) \neq \varnothing$. Поэтому на основании теоремы I.4.16 заключаем, что

$$K_{D_0}^*(x_0) = K_D^*(x_0) + K_{D_1}^*(x_0).$$

Таким образом, вектор $x^* \in \partial_x \widetilde{L}(x_0, \widetilde{y}_0^*)$ можно представить в виде

$$x^* = x_1^* - \sum_{i=k+1}^{m} \lambda_i x_i^*, \quad \lambda_i \geqslant 0, \quad \lambda_i f_i(x_0) = 0, \quad i = 1, \ldots, m, \\ x_1^* \in K_D(x_0). \tag{2.31}$$

Положим теперь $y_0^{i*} = \lambda_i$, $i = k+1, \ldots, m$. Тогда

$$L(x, y_0^*) = \widetilde{L}(x, \widetilde{y}_0^*) + \sum_{i=k+1}^{m} \lambda_i f_i(x),$$

и вектор

$$x_1^* = x^* + \sum_{i=k+1}^{m} \lambda_i x_i^*,$$

очевидно, принадлежит $\partial_x L(x_0, y_0^*)$.

Итак, $\partial_x L(x_0, y_0^*) \cap K_D^*(x_0) \neq \varnothing$, и из теоремы 2.2 можно заключить, что x_0 — точка минимума функции $L(x, y_0^*)$ на множестве D. Вместе с соотношениями

$$\lambda_i \geqslant 0, \quad \lambda_i f_i(x_0) = 0, \qquad i = k+1, \ldots, m,$$

вытекающими из формул (2.31) и (2.28), это доказывает теорему.

§ 3. Двойственные задачи выпуклого программирования

Как и в предыдущем параграфе, рассмотрим задачу выпуклого программирования: минимизировать функцию $f_0(x)$ на множестве

$$M = \{x\colon f_i(x) \leqslant 0, \quad i = 1, \ldots, m, \quad x \in D\},$$
$$\operatorname{dom} f_i \supseteq D, \qquad i = 1, \ldots, m.$$

Относительно функций $f_i(x)$ будем предполагать, что они собственные, выпуклые и замкнутые; D — выпуклое множество. Эти предположения считаются выполненными всюду в этом параграфе и специально не оговариваются в формулировках теорем.

Пусть опять

$$L(x, y^*) = f_0(x) + \sum_{i=1}^{m} y^{i*} f_i(x),$$

$$\varphi(y^*) = \inf_x \{L(x, y^*)\colon x \in D\}, \qquad y^* \geqslant 0,$$

$$V(y) = \inf_x \{f_0(x)\colon f_i(x) \leqslant y^i, \quad i = 1, \ldots, m, \quad x \in D\}.$$

Если ввести многозначное отображение

$$a(y) = \{(x, x^0) \in \mathbf{R}^{n+1} : f_i(x) \leqslant y^i,\\ i = 1, \ldots, m, \quad f_0(x) \leqslant x^0, \quad x \in D\},$$

то, как показано в предыдущем параграфе (формулы (2.11), (2.14)),

$$W_a(y, 0, 1) = V(y), \tag{3.1}$$

$$\Omega_a(-y^*, 0, 1) = \begin{cases} \varphi(y^*), & \text{если } y^* \geqslant 0, \\ -\infty, & \text{если } y^{i*} < 0 \\ & \text{для некоторого } i. \end{cases} \tag{3.2}$$

Назовем задачу максимизации $\varphi(y^*)$ по всем $y^* \geqslant 0$ *двойственной задачей*.

Теорема 3.1. *Пусть $V(0) \neq \pm\infty$ и функция $V(y)$ полунепрерывна снизу в точке $y = 0$. Тогда*

$$\inf_x \{f_0(x) \colon x \in M\} = \sup_{y^*} \{\varphi(y^*) \colon y^* \geqslant 0\}, \tag{3.3}$$

т. е. точная нижняя грань в исходной задаче совпадает с точной верхней гранью целевой функции в двойственной задаче.

Доказательство. Согласно теореме III.4.2, примененной к введенному многозначному отображению $a(y)$, с учетом замены обозначений (пространства X и Y нужно поменять местами, вместо x_0 в формуле (4.4) следует взять $y = 0$) получаем

$$\sup_{y^*} \{\Omega_a(y^*, 0, 1) + \langle 0, y^* \rangle\} = W_a(0, 0, 1),$$

так как функция $W_a(y, 0, 1) = V(y)$ по предположению полунепрерывна снизу в нуле, и, значит, условия теоремы III.4.2 выполнены. Используя соотношения (3.1), (3.2), получаем из последнего равенства

$$\sup_{y^*} \{\varphi(y^*) : y^* \geqslant 0\} = V(0),$$

что и требовалось доказать.

Заметим теперь, что если $x \in M, \quad y^* \geqslant 0$, то

$$\varphi(y^*) \leqslant L(x, y^*) = f_0(x) + \sum_{i=1}^{m} y^{i*} f_i(x) \leqslant f_0(x),$$

так что всегда $\varphi(y^*) \leqslant f_0(x)$ для $x \in M$, $y^* \geqslant 0$.

Дадим соотношению (3.3) другую интерпретацию. Легко видеть, что

$$\psi(x) = \sup_{y^*}\{L(x, y^*):\ y^* \geqslant 0\} =$$

$$= \begin{cases} f_0(x), \text{ если } f_i(x) \leqslant 0, & i = 1, \ldots, m, \\ +\infty, \text{ если } f_i(x) > 0 \text{ для некоторого } i. \end{cases}$$

Поэтому соотношение (3.3) эквивалентно следующему:

$$\inf_{x\in D}\sup_{y^*\geqslant 0} L(x, y^*) = \sup_{y^*\geqslant 0}\inf_{x\in D} L(x, y^*), \tag{3.4}$$

так что равенство (3.3) выполнено тогда и только тогда, когда выполнено равенство (3.4).

Теорема 3.2. *Пусть D — замкнутое множество и множество D_* — точек минимума функции $f_0(x)$ на M — ограничено. Тогда выполнены соотношения* (3.3) *и* (3.4).

Доказательство. Определим многозначное отображение

$$\bar{a}(y, y^0) = \{x:\ f_i(x) \leqslant y^i,\quad i = 0, \ldots, m,\quad x \in D\}.$$

В силу замкнутости f_i и D отображение $\bar{a}$ — замкнутое и выпуклое и $D_* = \bar{a}(0, V(0))$, так что $\bar{a}(0, V(0))$ ограничено. Но тогда по лемме III.1.1 $\bar{a}(y, y_0)$ — ограниченное отображение и множество $\bar{a}(y, y_0)$ ограничено и замкнуто, т. е. компактно. Пусть теперь $y_j \to 0$, $\lim\limits_{j\to\infty} V(y_j) = \mu$. Можно предполагать, что $\mu < +\infty$, так как случай $\mu = +\infty$ тривиален. Тогда

$$V(y_j) = \inf_x\{f_0(x):\ x \in \bar{a}(y_j, \mu + \varepsilon)\} \tag{3.5}$$

при больших j. Но $\bar{a}(y_j, \mu + \varepsilon)$ — компактное множество, а $f_0(x)$ — замкнутая собственная функция, поэтому нижняя грань в (3.5) достигается в некоторой точке $x_j \in \bar{a}(y_j, \mu + \varepsilon)$. Так как $\bar{a}$ — ограниченное отображение, то последовательность x_j ограничена и из нее можно выбрать сходящуюся.

Будем считать, что $x_j \to x_0$. Так как $\bar{a}$ — замкнутое отображение, то $x_0 \in \bar{a}(0, \mu + \varepsilon)$. Тогда из определения $\bar{a}$ следует, что $x_0 \in M$ и $f_0(x_0) \leqslant \mu + \varepsilon$. Но в этом случае

$V(0) \leqslant \mu + \varepsilon$, и так как $\varepsilon > 0$ произвольно, то

$$\mu = \lim_{j\to\infty} V(y_j) \geqslant V(0),$$

т. е. $V(y)$ полунепрерывна в точке нуль. Для завершения доказательства остается применить результаты теоремы 3.1.

Теорема 3.3. *Пусть задача выпуклого программирования имеет вектор Куна — Таккера* y_0^*. *Тогда выполнены соотношения* (3.3), (3.4) *и*

$$\varphi(y_0^*) = \sup_{y^*} \{\varphi(y^*):\ y^* \geqslant 0\}.$$

Доказательство. Если y_0^* — вектор Куна — Таккера, то $-y^* \in \partial V(0)$ по теореме 2.8. Поэтому

$$V(y) \geqslant V(0) - \langle y,\, y_0^* \rangle$$

для всех y. В частности, если $y \to 0$, то $\underline{\lim}\, V(y) \geqslant V(0)$, т. е. функция $V(y)$ полунепрерывна снизу в нуле. Отсюда по теореме 3.1 следует, что соотношения (3.3) и (3.4) выполнены. Далее, по определению 2.1

$$\varphi(y_0^*) = V(0) = \inf_x \{f_0(x):\ x \in M\}.$$

В то же время выше было показано, что $\varphi(y^*) \leqslant f_0(x)$ для любого $y^* \geqslant 0$, $x \in M$. Поэтому

$$\varphi(y_0^*) = \inf_x \{f_0(x):\ x \in M\} \geqslant \varphi(y^*), \qquad y^* \geqslant 0,$$

что завершает доказательство.

Комбинируя эту теорему с теоремами 2.5 и 2.10 предыдущего параграфа, гарантировавшими существование вектора Куна — Таккера, можно получить различные более просто проверяемые условия, при которых справедливы соотношения (3.3) и (3.4).

Заметим также, что если функции $f_i(x)$ имеют вид $\langle x, x_i^* \rangle - \alpha_i$, $i = 0, \ldots, m$, $D = X$, то получающаяся задача есть задача линейного программирования. Для нее определенное выше многозначное отображение $a(y)$ многогранно, а поэтому согласно теореме III.3.6 функция $V(y) = W_a(y, 0, 1)$ замкнута. Отсюда следует, что справедлива

Теорема 3.4. *Если все функции $f_i(x)$ имеют вид $\langle x, x_i^* \rangle - \alpha_i$, $i = 0, \ldots, m$, т. е. рассматривается задача линейного программирования, то имеют место соотношения* (3.3) *и* (3.4).

Нетрудно убедиться, что получающаяся при этом двойственная задача совпадает с рассмотренной в § 1.

§ 4. Некоторые задачи теории приближений

Пусть C — выпуклое компактное множество в X и $0 \in \operatorname{int} C$. По определению

$$W_C(x^*) = \max_x \{\langle x, x^* \rangle :\ x \in C\}.$$

Так как $0 \in \operatorname{int} C$, то $C \supseteq \varepsilon B$, где B — единичный шар, а $\varepsilon > 0$. Поэтому

$$W_C(x^*) \geqslant \max_x \{\langle x, x^* \rangle :\ x \in \varepsilon B\} = \varepsilon \| x^* \|. \tag{4.1}$$

Последнее соотношение следует из того, что по известной формуле $\langle x, x^* \rangle \leqslant \|x\| \|x^*\|$, положив $x_0 = \varepsilon \|x^*\|^{-1} x^*$, получим $x_0 \in \varepsilon B$, $\langle x_0, x^* \rangle = \varepsilon \|x^*\|$.

Возьмем теперь функцию Минковского множества C:

$$r_C(x) = \inf_{\rho \geqslant 0} \{\rho :\ x \in \rho C\}.$$

По доказанному в п. 2 § 2 главы I функция $r_C(x)$ положительно однородна, выпукла и конечна для всех x, так что $r_C(x)$ непрерывна по x. Так как C — компакт, то $r_C(x) > 0$ для $x \neq 0$. Поэтому $r_C(x)$ может служить в качестве обобщенного расстояния от точки x до точки 0.

Введем для компактного выпуклого множества A и замкнутого выпуклого множества M величину

$$d_C(A \mid M) = \inf_\rho \{\rho \geqslant 0 :\ A \subseteq M + \rho C\}. \tag{4.2}$$

В частности, если A состоит из одной точки x, то получается функция $d_C(x|M)$, рассмотренная п. 4 § 3 главы II. Вычисление же функции $d_C(x|M)$ сводится к оценке расстояния от точки x до множества M, когда метрика задается при помощи функции $r_C(x)$. Чтобы это стало более ясно, заметим, что

$$d_C(x\,|\,M)=\inf_{\rho}\{\rho\geqslant 0\colon\ x\in M+\rho C\}=$$

$$=\inf_{x_1\in M}\inf_{\rho}\{\rho\geqslant 0\colon\ x\in x_1+\rho C\}=$$

$$=\inf_{x_1\in M}\inf_{\rho}\{\rho\colon\ (x-x_1)\in\rho C\}=\inf_{x_1\in M}r_C(x-x_1),$$

т. е.

$$d_C(x\,|\,M)=\inf_{x_1}\{r_C(x-x_1)\colon\ x_1\in M\}. \tag{4.3}$$

В общем случае

$$d_C(A\,|\,M)=\inf_{\rho}\{\rho\geqslant 0\colon\ A\subseteq M+\rho C\}=$$

$$=\sup_{x_2\in A}\inf_{\rho}\{\rho\geqslant 0\colon\ x_2\in M+\rho C\}=\sup_{x_2\in A}d_C(x_2\,|\,M). \tag{4.4}$$

Таким образом, вычисление $d_C(A|M)$ сводится к нахождению в A точки, которая наихудшим образом аппроксимируется множеством M в метрике, определяемой множеством C. Если $C=B$, где B — единичный шар, то $r_B(x)=\|x\|$, так что в этом случае приближение получается в естественной метрике пространства X.

Пусть теперь множество M состоит из одной точки x. Тогда

$$d_C(A\,|\,x)=\inf_{\rho}\{\rho\geqslant 0\colon\ A\subseteq x+\rho C\}=$$

$$=\inf_{\rho}\{\rho\geqslant 0\colon\ A-x\subseteq\rho C\}.$$

Отсюда видно, что вычисление $d_C(A|x)$ сводится к нахождению наименьшего коэффициента растяжения, при котором множество ρC содержит $A-x$. Если поставить задачу нахождения минимума $d_C(A|x)$ по x, то геометрически задача сведется к нахождению такого сдвига множества A, при котором его можно погрузить в множество ρC с наименьшим коэффициентом растяжения. В частности, если $C=B$ — шар, то задача состоит в нахождении шара наименьшего радиуса, описанного вокруг A.

Рассмотрим, что может дать теория, развитая в предыдущих параграфах, для поставленных задач.

Теорема 4.1. *Справедлива формула*

$$d_C(A\,|\,M) = \sup_{x^*}\{W_A(x^*) - W_M(x^*):\ W_C(x^*) \leqslant 1\}.$$

Доказательство. Согласно теореме II.3.15

$$d_C(x\,|\,M) = \sup_{x^*}\{\langle x, x^*\rangle - W_M(x^*):\ W_C(x^*) \leqslant 1\}. \tag{4.5}$$

Подставляя это выражение в формулу (4.4), получим

$$d_C(A\,|\,M) = \sup_{x_2\in A}\ \sup_{x^*}\{\langle x_2, x^*\rangle - W_M(x^*):\ W_C(x^*) \leqslant 1\} =$$

$$= \sup_{x^*}\left\{\sup_{x_2\in A}\langle x_2, x^*\rangle - W_M(x^*):\ W_C(x^*) \leqslant 1\right\},$$

что и требовалось доказать.

Обратимся теперь к задаче об отыскании условий, характеризующих точку x_1, в которой достигается минимум в правой части формулы (4.3), т. е. характеризующих точку из M, ближайшую к x в метрике $r_C(x)$.

Так как $r_C(x) = d_C(x\,|\,\{0\})$, то согласно теореме II.3.16 справедлива формула

$$\partial r_C(x_0) = \begin{cases} \partial\delta(x_0\,|\,\{0\}) \cap \{x^*:\ W_C(x^*) \leqslant 1\}, \\ \qquad\qquad \text{если} \quad r_C(x_0) = 0, \\ \partial\delta(x_0\,|\,r_C(x_0)\,C) \cap \{x^*:\ W_C(x^*) = 1\}, \\ \qquad\qquad \text{если} \quad r_C(x_0) > 0. \end{cases} \tag{4.6}$$

Но $r_C(x_0) = 0$, только если $x_0 = 0$. Так как по формуле (II.3.25)

$$\partial\delta(x_0|M) = -(\mathrm{con}\,(M - x_0))^*, \tag{4.7}$$

а $(\mathrm{con}\,\{0\})^* = X^*$, то

$$\partial\delta(0|\{0\}) = -X^* = X^*,$$

и в силу формулы (4.6) выполняется соотношение

$$\partial r_C(0) = \{x^*:\ W_C(x^*) \leqslant 1\}. \tag{4.8}$$

Пусть теперь $x_0 \neq 0$, так что $r_C(x_0) > 0$. Тогда

$$\mathrm{con}\,(r_C(x_0)C - x_0) = \{\lambda(r_C(x_0)x - x_0):\ \ \lambda > 0, \quad x \in C\},$$

и поэтому сопряженный ему конус состоит из тех и только тех точек x^*, для которых

$$\langle r_C(x_0)x - x_0,\ x^*\rangle \geqslant 0, \qquad x \in C,$$

или, по другому,

$$-\langle x_0,\ x^*\rangle \geqslant r_C(x_0)W_C(-x^*). \tag{4.9}$$

Учитывая формулы (4.7) и (4.6), получаем

$$\partial r_C(x_0)=\{x^*\colon \langle x_0,\ x^*\rangle \geqslant r_C(x_0),\ \ W_C(x^*)=1\}. \tag{4.10}$$

Но $x_0 \in r_C(x_0)C$ по определению $r_C(x_0)$ и в силу замкнутости C. Поэтому

$$\langle x_0,\ x^*\rangle \leqslant r_C(x_0)W_C(x^*),$$

а, значит, в формуле (4.10) имеет место равенство

$$\langle x_0,\ x^*\rangle = r_C(x_0).$$

Зафиксируем полученный результат.

Теорема 4.2. *Если C — компактное выпуклое множество и $0 \in \operatorname{int} C$, то*

$$\partial r_C(x_0)=$$
$$=\begin{cases}\{x^*\colon W_C(x^*)\leqslant 1\}, & \text{если}\quad x_0=0,\\ \{x^*\colon \langle x_0,\ x^*\rangle = r_C(x_0),\ \ W_C(x^*)=1\}, & \text{если}\quad x_0\neq 0.\end{cases} \tag{4.11}$$

Теперь можно сформулировать условия, характеризующие точку x_1, в которой функция $r_C(x-x_1)$ достигает своего минимума $d_C(x|M)$ на множестве M.

Теорема 4.3. *Для того чтобы точка $x_1 \in M$ была наилучшим приближением к $x \notin M$ в метрике $r_C(x-x_1)$, необходимо и достаточно, чтобы нашелся такой вектор $x^* \in K_M^*(x_1)$, что*

$$\langle x_1-x,\ x^*\rangle = r_C(x-x_1)=d_C(x|M),\quad W_C(-x^*)=1.$$

Доказательство. Точка наилучшего приближения x_1 по определению является точкой минимума функции $g(x_2)=r_C(x-x_2)$ при $x_2\in M$. По теореме 2.2 для этого необходимо и достаточно, чтобы нашелся такой вектор $x^*\in K_M^*(x_1)$, что $x^*\in\partial g(x_1)$. Но $x^*\in\partial g(x_1)$ тогда и только тогда, когда

$$g(x_2)-g(x_1)=r_C(x-x_2)-r_C(x-x_1)\geqslant \langle x_2-x_1,\ x^*\rangle$$

для всех x_2. Обозначая $z=x-x_2$, $z_0=x-x_1$, получаем

$$r_C(z)-r_C(z_0)\geqslant \langle z-z_0,\ -x^*\rangle,$$

т. е. $-x^* \in \partial r_C(x - x_1)$. Согласно теореме 4.2 для выполнения этого неравенства необходимо и достаточно, чтобы

$$\langle x - x_1,\ -x^* \rangle = r_C(x - x_1) = d_C(x \mid M),$$
$$W_C(-x^*) = 1,$$

что доказывает теорему.

Множество $x + \rho C$ будем для удобства называть *обобщенным шаром радиуса* ρ *с центром в точке* x. Рассмотрим задачу нахождения обобщенного шара минимального радиуса, содержащего данный компакт A. Эта задача сводится, как говорилось выше, к минимизации функции $d_C(A \mid x)$ по x. По определению

$$\begin{aligned} d_C(A \mid x) &= \\ &= \inf_{\rho} \{\rho \geqslant 0\colon\ A \subseteq x + \rho C\} = \inf_{\rho} \{\rho \geqslant 0\colon\ A - x \in \rho C\} = \\ &= \sup_{x_1 \in A} \inf_{\rho} \{\rho \geqslant 0\colon\ x_1 - x \in \rho C\} = \sup_{x_1 \in A} r_C(x_1 - x). \end{aligned}$$

Таким образом, поставленная задача сводится к нахождению минимума функции

$$f(x) = \sup_{x_1 \in A} r_C(x_1 - x). \tag{4.12}$$

Заметим, что $f(x)$ есть радиус минимального шара с фиксированным центром x, содержащего множество A. Так как A — компакт, а r_C — непрерывная функция, то верхняя грань в правой части формулы (4.12) достигается.

Обозначим

$$A(x) = \{x_1\colon\ r_C(x_1 - x) = f(x),\ x_1 \in A\}.$$

Согласно теореме II.3.14, все условия которой в рассматриваемом случае выполняются,

$$\partial f(x) = \overline{\mathrm{co}} \left(\bigcup_{x_1 \in A(x)} \partial_x r_C(x_1 - x) \right). \tag{4.13}$$

Из выкладок, проделанных в предыдущей теореме, и формулы (4.11) следует, что

$$\begin{aligned} \partial_x r_C(x_1 - x) = \{x^*\colon\ \langle x_1 - x,\ x^* \rangle &= \\ &= r_C(x - x_1),\ \ W_C(-x^*) = 1\}. \end{aligned} \tag{4.14}$$

Так как множество x^*, для которых $W_C(-x^*) = 1$, в силу неравенства (4.1) ограничено, а $A(x)$ компактно, то нетрудно показать, используя теоремы I.1.6 и I.1.1, что в формуле (4.13) вместо замкнутой выпуклой оболочки можно взять просто выпуклую оболочку, т. е.

$$\partial f(x) = \mathrm{co}\Big(\bigcup_{x_1 \in A(x)} \partial_x r_C(x_1 - x)\Big). \tag{4.15}$$

Теорема 2.1 утверждает, что точка x_0 есть точка минимума функции $f(x)$ тогда и только тогда, когда $0 \in \partial f(x)$. Из формул (4.15), (4.14) и теоремы I.1.1 теперь следует, что если x_0 есть точка минимума $f(x)$, то существуют такие числа $\lambda_i \geqslant 0$, $i = 0, \ldots, n$, и такие векторы x_i^*, $x_{1i} \in A(x_0)$, $i = 0, \ldots, n$, что

$$\sum_{i=0}^{n} \lambda_i x_i^* = 0, \quad \sum_{i=0}^{n} \lambda_i = 1, \tag{4.16}$$

$$\langle x_{1i} - x_0, x_i^* \rangle = r_C(x_0 - x_{1i}) = f(x_0),$$

$$W_C(-x_i^*) = 1, \quad i = 0, \ldots, n.$$

Сформулируем полученный результат.

Теорема 4.4. *Для того чтобы обобщенный шар $x_0 + f(x_0)C$ был шаром минимального радиуса, описанным вокруг компактного множества A, необходимо и достаточно, чтобы выполнялись условия* (4.16).

Заметим, что в предыдущих рассуждениях никак не использовалась выпуклость множества A, так что этого условия нет и в формулировке теоремы.

Теорема 4.5. *Пусть A — компактное множество. Тогда существует такое его подмножество A_0, состоящее не более чем из $n+1$ точек, что радиус минимального шара, описанного вокруг A_0, совпадает с радиусом минимального шара, описанного вокруг A.*

Доказательство. Если рассмотреть множество A_0, состоящее из точек x_{1i}, $i = 0, \ldots, n$, фигурирующих в условиях (4.16), то, поскольку эти условия являются и достаточными, шар $x_0 + f(x_0)C$ будет шаром минимального радиуса, описанным вокруг A_0.

§ 5. Задачи наилучшего равномерного приближения

Классической задачей наилучшего равномерного приближения является задача аппроксимации непрерывной функции на отрезке многочленами.

Пусть $g(t)$, $t \in [0, 1]$,— непрерывная функция,

$$P_n(x, t) = x^1 + x^2 t + \ldots + x^n t^{n-1},$$

где x — вектор с компонентами $x^1, \ldots, x^n$, являющимися коэффициентами многочлена $P_n(x, t)$ степени $n-1$ относительно переменной t. Наилучшее равномерное приближение сводится к отысканию такого многочлена, для которого величина

$$f(x) = \max_{t \in [0,1]} |g(t) - P_n(x, t)| \tag{5.1}$$

минимальна.

Другая классическая задача, которая с иных позиций рассматривалась в предыдущем параграфе, есть задача нахождения центра шара, описанного вокруг компакта $A \subseteq \mathbf{R}^n$, для которого радиус был бы минимален. Нетрудно видеть, что такая задача сводится к минимизации функций

$$f(x) = \max_y \{ \|x - y\| : \ y \in A \}.$$

Как ни различна на первый взгляд природа сформулированных задач, тем не менее они могут быть решены одним и тем же приемом. Этот прием основан на использовании теоремы II.3.14.

Пусть $f(x, \alpha)$, $\alpha \in A$, где A — компакт, семейство выпуклых по $x \in \mathbf{R}^n$, непрерывных по x и α функций. Будем предполагать также, что при каждом $\alpha \in A$ существует градиент функции $f(x, \alpha)$ по x, т. е. существует вектор $f'_x(x, \alpha)$. Предположим, что он непрерывен по α при каждом x.

Положим

$$\begin{gathered} f(x) = \max_\alpha \{ f(x, \alpha) : \ \alpha \in A \}, \\ A(x) = \{ \alpha \in A : \ f(x, \alpha) = f(x) \}. \end{gathered} \tag{5.2}$$

Ясно, что $f(x)$ — выпуклая функция, а $A(x)$ — компактное множество. Согласно теореме II.3.14

$$\partial f(x) = \overline{\mathrm{co}} \Big(\bigcup_{\alpha \in A(x)} f'_x(x, \alpha) \Big).$$

Легко видеть, что в силу сделанных предположений все условия теоремы II.3.14 выполнены, а так как функции $f(x, \alpha)$ дифференцируемы по x, то $\partial f(x, \alpha) = \{f'_x(x, \alpha)\}$. Множество

$$\bigcup_{\alpha \in A(x)} f'_x(x, \alpha)$$

есть образ множества $A(x)$ при отображении $f'_x(x, \alpha)$: $A \to \mathbf{R}^n$, т. е.

$$\bigcup_{\alpha \in A(x)} f'_x(x, \alpha) = f'_x(x, A(x)).$$

Но $A(x)$ — компакт, а отображение $f'_x(x, \alpha)$ непрерывно по α в силу сделанных предположений. Поэтому $f'_x(x, A(x))$ — компактное множество. Итак,

$$\partial f(x) = \overline{\text{co}}\, f'_x(x, A(x)).$$

Но в силу теоремы I.1.7 (если M — компакт, то $\text{co}\, M = \overline{\text{co}}\, M$) имеем

$$\partial f(x) = \text{co}\, f'_x(x, A(x)).$$

Применяя теорему I.1.1, окончательно получим

$$\partial f(x) = \left\{ \sum_{i=0}^{n} \lambda_i f'_x(x, \alpha_i) \colon \ \alpha_i \in A(x), \quad \lambda_i \geqslant 0, \right.$$
$$\left. i = 0, \ldots, n, \quad \sum_{i=0}^{n} \lambda_i = 1 \right\}. \tag{5.3}$$

Из формулы (5.3) и теоремы 2.2 непосредственно вытекает следующая

Теорема 5.1. *Для того чтобы функция $f(x)$, определенная соотношением (5.2), достигала своего минимума на выпуклом множестве M в точке x_0, необходимо и достаточно, чтобы нашлись такие точки $\alpha_i \in A(x_0)$, $i = 0, 1, \ldots, n$, и такие числа λ_i, что*

$$\begin{gathered} \sum_{i=0}^{n} \lambda_i f'_x(x_0, \alpha_i) \in K^*_M(x_0), \\ \lambda_i \geqslant 0, \quad i = 0, 1, \ldots, n, \quad \lambda_0 + \lambda_1 + \ldots + \lambda_n = 1. \end{gathered} \tag{5.4}$$

Замечание. Некоторые точки α_i могут совпадать между собой так, что на самом деле число слагаемых в формулах (5.4) может быть меньше $n + 1$.

Рассмотрим теперь множество точек $\alpha_i \in A(x_0)$, $i=0, \ldots, n$, фигурирующих в теореме, и обозначим его через A_0:

$$A_0 = \{\alpha_i\colon\ i=0, 1, \ldots, n\}.$$

Положим

$$f_0(x) = \max_{\alpha} \{f(x, \alpha)\colon\ \alpha \in A_0\}. \tag{5.5}$$

Если точка x_0 является точкой минимума функции $f(x)$, то выполнены условия (5.4), которые являются необходимыми и достаточными условиями минимума. Легко видеть, что применение теоремы 5.1 к функции $f_0(x)$ показывает, что x_0 есть также точка минимума и для функции $f_0(x)$.

Таким образом, справедлива

Теорема 5.2. *При сделанных выше предположениях существует множество $A_0 \subseteq A$, состоящее не более чем из $n+1$ точек α_i и такое, что точка минимума функции $f(x)$ является одновременно и точкой минимума функции $f_0(x)$, определенной соотношением* (5.5).

Используем полученные результаты для решения поставленных выше задач. Пусть

$$f(x, y) = \|x - y\|$$

и $y \in A$, где A — компакт. Если $x \neq y$, то

$$f'_x(x, y) = \frac{x-y}{\|x-y\|} = \frac{x-y}{f(x, y)}. \tag{5.6}$$

Пусть

$$f(x) = \max_{y} \{\|x-y\|\colon\ y \in A\}, \tag{5.7}$$

$$A(x) = \{y \in A\colon\ \|x-y\| = f(x)\}. \tag{5.8}$$

Применение теоремы 5.1 в случае $M = \mathbf{R}^n$ показывает, что точка x_0 тогда и только тогда является точкой минимума функции (5.7), когда существуют такие точки $y_i \in A$, $i=0, \ldots, n$, и такие числа $\lambda_i \geqslant 0$, что

$$\sum_{i=0}^{n} \lambda_i \frac{x_0 - y_i}{\|x_0 - y_i\|} = 0, \qquad \sum_{i=0}^{n} \lambda_i = 1,$$

$$\|x_0 - y_i\| = f(x_0).$$

Отсюда легко получаем

$$x_0 = \sum_{i=0}^{n} \lambda_i y_i,$$
$$\lambda_i \geqslant 0, \quad \|x_0 - y_i\| = f(x_0), \quad i = 0, \dots, n, \quad \sum_{i=0}^{n} \lambda_i = 1. \tag{5.9}$$

Напомним, что рассматриваемая задача означала поиск центра шара наименьшего радиуса, описанного вокруг компакта A. Геометрическая интерпретация условий (5.9) позволяет сформулировать следующую теорему.

Теорема 5.3. *Для того чтобы точка x_0 была центром шара наименьшего радиуса, описанного вокруг компакта A, необходимо и достаточно, чтобы нашлись такие точки $y_i \in A$, $i = 0, 1, \dots, n$, лежащие на поверхности шара, что x_0 принадлежит симплексу, натянутому на эти точки.*

Тот факт, что точки y_i лежат на поверхности шара, выражается вторым из соотношений (5.9). Геометрическая интерпретация теоремы 5.2 для данного случая достаточно очевидна.

Теорема 5.4. *В данном компакте $A \subseteq \mathbf{R}^n$ существует такое подмножество A_0, состоящее не более чем из $n+1$ точек, что шар минимального радиуса, описанный вокруг A, одновременно является шаром минимального радиуса для A_0.*

Пусть теперь на некотором компакте $\Omega \subseteq \mathbf{R}^m$ заданы непрерывные функции $\varphi_i(y)$, $y \in \Omega$, $i = 1, \dots, n$. Назовем *обобщенным многочленом* выражение

$$P_n(x, y) = \sum_{i=1}^{n} x^i \varphi_i(y). \tag{5.10}$$

Если теперь $g(y)$ — произвольная непрерывная функция, то задача ее приближения при помощи обобщенного многочлена состоит в минимизации функции

$$f(x) = \max_{y \in \Omega} |g(y) - P_n(x, y)| \tag{5.11}$$

по $x \in \mathbf{R}^n$. Пусть

$$A = \{(y, \xi) \in \mathbf{R}^{m+1} \colon y \in \Omega, \ |\xi| \leqslant 1\}.$$

Положим

$$f(x,\ y,\ \xi)=\xi(g(y)-P_n(x,\ y)).$$

Тогда

$$f'_x(x,\ y,\ \xi)=-\xi\varphi(y), \tag{5.12}$$

где $\varphi(y)$ — n-мерный вектор с компонентами $\varphi_i(y)$. Нетрудно убедиться в том, что

$$f(x)=\max_{y,\xi}\{f(x,\ y,\ \xi):\ (y,\ \xi)\in A\}. \tag{5.13}$$

Далее, если $(y_1,\ \xi_1)\in A(x)$, т. е.

$$\xi_1(g(y_1)-P_n(x,\ y_1))=$$
$$=\max_{(y,\xi)}\{\xi(g(y)-P_n(x,\ y)):\ y\in\Omega,\quad |\xi|\leqslant 1\},$$

то

$$\xi_1=\operatorname{sign}(g(y_1)-P_n(x,\ y_1)). \tag{5.14}$$

Применение теоремы 5.1 к функции $f(x)$, определяемой соотношением (5.11) или соотношением (5.13), дает следующий результат. Точка x_0 доставляет наименьшее значение функции $f(x)$ тогда и только тогда, когда существуют такие точки $y_i\in\Omega$, $i=0,\ 1,\ \ldots,\ n$, что

$$\sum_{i=0}^{n}\lambda_i\xi_i\varphi(y_i)=0, \tag{5.15}$$

$$\xi_i=\operatorname{sign}(g(y_i)-P_n(x_0,\ y_i)), \tag{5.16}$$

$$|g(y_i)-P_n(x_0,\ y_i)|=f(x_0),\qquad i=0,\ 1,\ \ldots,\ n, \tag{5.17}$$

$$\sum_{i=0}^{n}\lambda_i=1,\qquad \lambda_i\geqslant 0. \tag{5.18}$$

Теорема 5.5. *Для того чтобы обобщенный многочлен $P_n(x_0,\ y)$ был многочленом наилучшего равномерного приближения непрерывной функции $g(y)$ на компакте Ω, необходимо и достаточно, чтобы нашлись такие точки $y_i\in\Omega$, $i=0,\ 1,\ \ldots,\ n$, в которых отклонение многочлена от функции $g(y)$ максимально по модулю и выполнены условия*

$$\sum_{i=0}^{n}\lambda_i\xi_i\varphi(y_i)=0,$$

$$\xi_i=\operatorname{sign}(g(y_i)-P_n(x_0,\ y_i)),$$

при некоторых неотрицательных и не обращающихся в нуль одновременно λ_i.

Доказательство теоремы сводится к рассмотрению формул (5.15)—(5.18). То, что числа λ_i неотрицательны и не все равны нулю, следует из условий (5.18). Соотношение (5.17) выражает тот факт, что в точках y_i достигается максимальное уклонение, ибо $f(x_0)$ определяется по формуле (5.11).

Теорема 5.2 применительно к рассмотренной задаче получает следующую интерпретацию.

Теорема 5.6. *Существует подмножество* Ω_0 *компакта* Ω, *состоящее не более чем из* $n+1$ *точек и такое, что многочлен наилучшего приближения функции* $g(y)$ *на* Ω *является одновременно многочленом наилучшего приближения на* Ω_0.

§ 6. Модели экономической динамики

Развитие линейного и выпуклого программирования с самого начала тесно связано с решением экономических задач. При этом имеется в виду не только решение методами линейного и выпуклого программирования конкретных задач минимизации расходов при ограниченных ресурсах, но и получение некоторых результатов, связанных с уточнением основных понятий, встречающихся в экономической науке. В частности, оказывается, что такое понятие, как цена, тесно связано с решением двойственной задачи выпуклого программирования и с вектором Куна — Таккера.

1. Вектор Куна — Таккера и цены. Чтобы проиллюстрировать связь между вектором Куна — Таккера и ценами, дадим следующую интерпретацию задаче выпуклого программирования, рассмотренной в § 2. Пусть создание некоторого объекта, который характеризуется технологическими параметрами $x^1, x^2, \ldots, x^n$, т. е. вектором $x \in \mathbf{R}^n$, имеет стоимость $f_0(x)$. При этом имеется m ресурсов в количествах $y_0^1, \ldots, y_0^m$, и при выборе вектора технологических параметров x количество расходуемого i-го ресурса равно $f_i(x)$. Таким образом, возникают ограничения

$$f_i(x) \leqslant y_0^i, \qquad i = 1, \ldots, m. \tag{6.1}$$

Вектор x по техническим соображениям может меняться в некотором множестве D. Если теперь поставить задачу выбора технологических параметров x так, чтобы минимизировать стоимость создания объекта, то получим задачу

$$\min_x \{f_0(x): f_i(x) \leqslant y_0^i, \quad i = 1, \ldots, m, \quad x \in D\}, \qquad (6.2)$$

т. е. типичную задачу выпуклого программирования в случае, если выполнены соответствующие требования о выпуклости функции $f_i(x)$ и множества D.

Пусть, как и в § 2,

$$V(y) = \inf_x \{f_0(x): f_i(x) \leqslant y^i, \quad i = 1, \ldots, m, \quad x \in D\} \qquad (6.3)$$

и существует вектор Куна — Таккера y_0^* для задачи (6.2). Согласно теореме 2.8

$$-y_0^* \in \partial V(y_0). \qquad (6.4)$$

Более того, согласно этой же теореме любой вектор y_0^*, принадлежащий $-\partial V(y_0)$, есть вектор Куна — Таккера. В частности, любой такой вектор имеет неотрицательные компоненты согласно определению 2.1.

Представим себе теперь, что имеется свободный рынок, на котором можно купить дополнительные ресурсы по ценам y^{i*}, $i = 1, \ldots, m$. Пусть y^* есть этот вектор цен. Спрашивается, при каких ценах y^* имеет смысл покупать эти ресурсы в количество $y - y_0 \geqslant 0$? Очевидно, что тогда стоимость создания объекта будет равна $V(y) + \langle y - y_0, y^* \rangle$ и покупка дополнительных ресурсов будет иметь смысл лишь, если

$$V(y) + \langle y - y_0, y^* \rangle < V(y_0).$$

Обозначим через P множество таких векторов y^*, для которых существует такой вектор $y_0^* \in (-\partial V(y_0))$, что $y^* \geqslant y_0^*$. Так как $-y_0^* \in \partial V(y_0)$, то

$$V(y) - V(y_0) \geqslant -\langle y - y_0, y_0^* \rangle,$$

$$V(y) + \langle y - y_0, y^* \rangle - V(y_0) \geqslant \langle y - y_0, y^* - y_0^* \rangle \geqslant 0$$

для $y \geqslant y_0$, т. е. по таким ценам никакая дополнительная покупка ресурсов не целесообразна, так как не приводит к удешевлению общей стоимости разработки.

Пусть теперь $\tilde{y}^* \notin P$. Так как $y_0^* \geqslant 0$, если $y_0^* \in -\partial V(y_0)$, то нетрудно показать, что множество P замкнуто в силу замкнутости $\partial V(y_0)$. Кроме того, P выпукло, так как выпукло $\partial V(y_0)$. По теореме I.2.1 существуют такие вектор $\bar{y}$ и число $\varepsilon > 0$, что

$$\langle \bar{y}, \tilde{y}^* \rangle \leqslant \langle \bar{y}, z^* \rangle - \varepsilon, \quad z^* \in P. \tag{6.5}$$

По определению P

$$P = \{z^*: z^* \geqslant y_0^* \quad \text{для некоторого} \quad y_0^* \in -\partial V(y_0)\}.$$

Поэтому векторы z^* с достаточно большими компонентами всегда принадлежат P. Тогда из неравенства (6.5) следует, что $\bar{y} \geqslant 0$, ибо, в противном случае, устремив z^* к $+\infty$, получили бы $-\infty$ в правой части (6.5). Кроме того, из неравенства (6.5) вытекает, что

$$\langle \bar{y}, \tilde{y}^* \rangle + \langle \bar{y}, -y_0^* \rangle \leqslant -\varepsilon, \quad -y_0^* \in \partial V(y_0)$$

или

$$\langle \bar{y}, \tilde{y}^* \rangle + \sup_{z^*} \{\langle \bar{y}, z^* \rangle: z^* \in \partial V(y_0)\} \leqslant -\varepsilon. \tag{6.6}$$

Если функция $V(y)$ непрерывна в точке y_0, то согласно теореме II.3.5 левая часть неравенства (6.6) есть производная по направлению $\bar{y}$ от функции $V(y) + \langle y - y_0, \tilde{y}^* \rangle$ в точке y_0. Так как эта производная меньше отрицательного числа, то при достаточно малых $\lambda > 0$

$$V(y_0 + \lambda \bar{y}) + \langle (y_0 + \lambda \bar{y}) - y_0, \tilde{y}^* \rangle < V(y_0). \tag{6.7}$$

Таким образом, если $\tilde{y}^* \notin P$, то существует такой вектор $y = y_0 + \lambda \bar{y} \geqslant y_0$, что дополнительная покупка ресурсов в количестве $y - y_0 \geqslant 0$ по ценам $\tilde{y}^*$ приводит к общему сокращению стоимости создания объекта.

Сформулируем полученный результат, используя введенную экономическую терминологию.

Теорема 6.1. *Пусть стоимость создания объекта при векторе ресурсов y определяется формулой* (6.2). *Пусть y_0 — вектор наличных ресурсов и функция $V(y)$ непрерывна в точке y_0. Тогда приобретение дополнительных ресурсов $y - y_0 \geqslant 0$ по ценам y^* целесообразно в том* *и только том случае, когда не существует такого вектора Куна — Таккера задачи* (6.2), *что* $y^* \geqslant y_0^*$.

Из приведенной теоремы следует, что вектор Куна — Таккера целесообразно рассматривать как некоторые цены, обусловленные ограниченностью определенных ресурсов.

2. Абстрактная модель экономической динамики. Попытаемся теперь на абстрактном уровне описать функционирование некоторой экономической системы. Будем считать, что это функционирование происходит дискретно во времени, т. е. время t принимает значения $0, 1, 2, \ldots$ $\ldots, T$. При этом, если в момент t имеется вектор ресурсов $x \in \mathbf{R}^n$, то благодаря деятельности производства этот вектор в момент времени $t+1$ может быть переработан в один из векторов $y \in a(x)$, где $a(x)$ — выпуклое многозначное отображение, $a(x) \subseteq \mathbf{R}^n$. Таким образом, возможные количества ресурсов x_t, $t = 0, 1, \ldots, T$, по периодам времени связаны соотношениями

$$x_{t+1} \in a(x_t), \qquad t = 0, 1, \ldots, T-1. \tag{6.8}$$

Допустим, что вектор начальных ресурсов x_0 задан. Назовем *траекторией системы* последовательность точек $x_0, x_1, \ldots, x_T$, связанных соотношением (6.8). Очевидно, что траектория задана неоднозначно. Поставим задачу отыскания траектории $\{\overset{*}{x}_t\}_{t=0}^{T}$, минимизирующей функцию

$$f = \sum_{i=1}^{T} g(x_t, t) \tag{6.9}$$

и такой, что $x_T \in M$, где M — выпуклое множество, а $g(x, t)$, $t = 1, \ldots, T$ — выпуклые функции x при каждом t. Если функцию $g(x, t)$ интерпретировать как расходы, то поставленная задача является задачей выбора траектории, приводящей на заданное множество конечных ресурсов M с минимальными суммарными расходами.

В качестве основы для решения поставленной задачи будет служить теорема 2.4. Чтобы вложить поставленную задачу в схему, лежащую в основе теоремы 2.4, введем векторы $w = (x_1, x_2, \ldots, x_T) \in \mathbf{R}^{nT}$. Положим

$$f(w) = \sum_{t=1}^{T} g(x_t, t).$$

Если определить в $\mathbf{R}^{nT}$ множества

$$\widetilde{M}_t = \{w\colon (x_t, x_{t+1}) \in \operatorname{gf} a\}, \qquad t = 0, 1, \ldots, T-1, \tag{6.10}$$

$$\widetilde{M}_T = \{w\colon x_T \in M\}, \tag{6.11}$$

то поставленная задача сведется к минимизации выпуклой функции $f(w)$ на выпуклом множестве

$$\widetilde{M} = \bigcap_{t=0}^{T} \widetilde{M}_t.$$

Чтобы применить теорему 2.4, вычислим конусы $K_{\widetilde{M}_t}(w)$, которые будем обозначать просто через $\widetilde{K}_t(w)$. Если

$$w + \lambda\bar{w} \in \widetilde{M}_t, \quad t = 1, \ldots, T-1,$$

при достаточно малых $\lambda > 0$, т. е.

$$(x_t + \lambda\bar{x}_t,\ x_{t+1} + \lambda\bar{x}_{t+1}) \in \operatorname{gf} a,$$

то $\bar{w} \in \widetilde{K}_t(w)$. Таким образом,

$$\widetilde{K}_t(w) = \{\bar{w}\colon (\bar{x}_t,\ \bar{x}_{t+1}) \in K_a((x_t,\ x_{t+1}))\}.$$

Поэтому, как нетрудно подсчитать,

$$\widetilde{K}_t^*(w) = \{w^*\colon (x_t^*, x_{t+1}^*) \in K_a^*((x_t, x_{t+1})),$$
$$x_k^* = 0, \quad k \neq t, t+1\}, \tag{6.12}$$

где $w^* \in \mathbf{R}^{nT}$, $w^* = (x_1^*, \ldots, x_T^*)$.

Действительно, по определению $w^* \in \widetilde{K}_t^*(w)$ тогда и только тогда, когда

$$\langle\bar{w}, w^*\rangle = \sum_{k=1}^{T} \langle\bar{x}_k, x_k^*\rangle \geqslant 0 \tag{6.13}$$

для всех $\bar{w} \in \widetilde{K}_t(w)$. Но компоненты $\bar{x}_k$ вектора $\bar{w} \in \widetilde{K}_t(w)$ при $k \neq t$, $t+1$ произвольны. Поэтому соотношение (6.13) возможно лишь, если $x_k^* = 0$, $k \neq t$, $t+1$. В этом случае оно принимает вид

$$\langle\bar{x}_t, x_t^*\rangle + \langle\bar{x}_{t+1}, x_{t+1}^*\rangle \geqslant 0,$$
$$(\bar{x}_t,\ \bar{x}_{t+1}) \in K_a((x_t,\ x_{t+1})),$$

т. е.

$$\left(x_t^*,\, x_{t+1}^*\right) \in K_a^*\left((x_t,\, x_{t+1})\right).$$

Тем самым формула (6.12) доказана.

Если $t = 0$, то, так как x_0 фиксировано, $w + \lambda \bar{w} \in \widetilde{M}_0$ тогда и только тогда, когда $x_1 + \lambda \bar{x}_1 \in a(x_0)$, т. е. $\bar{x}_1 \in \in K_{a(x_0)}(x_1)$. Итак,

$$\widetilde{K}_0(w) = \{\bar{w}\colon\ \bar{x}_1 \in K_{a(x_0)}(x_1)\},$$

откуда

$$\widetilde{K}_0^*(w) = \{w^*\colon\ x_1^* \in K_{a(x_0)}^*(x_1),\quad x_t^* = 0,\quad t = 2, \ldots, T\}. \tag{6.14}$$

Наконец, из соотношения

$$\widetilde{K}_T(w) = \{\bar{w}\colon\ \bar{x}_T \in K_M(x_T)\}$$

вытекает

$$\widetilde{K}_T(w) = \{w^*\colon\ x_T^* \in K_M^*(x_T),\quad x_t^* = 0,\quad t < T\}. \tag{6.15}$$

Пусть теперь существует траектория $\{x_t\}_{t=0}^T$ такая, что функции $g(x,\, t)$ непрерывны в ее точках x_t. Тогда на основании теоремы 2.4, если $\{\tilde{x}_t\}_{t=0}^T$ — оптимальная траектория, то существуют такие векторы

$$w^*(t) \in \widetilde{K}_t^*(\widetilde{w}),\quad t = 0, 1, \ldots, T,\quad \widetilde{w} = (\tilde{x}_1, \ldots, \tilde{x}_T) \tag{6.16}$$

и число λ, принимающее значение нуль или единица, что

$$\lambda w^* = \sum_{t=0}^{T} w^*(t), \tag{6.17}$$

где $w^* \in \partial_w f(\widetilde{w})$. При этом λ и $w^*(t)$ не все равны нулю одновременно. Из вида функции $f(w)$ следует, что если $w^* \in \partial_w f(w)$, то вектор w^* имеет вид

$$w^* = \left(x_{10}^*,\, x_{20}^*,\, \ldots,\, x_{T0}^*\right),\qquad x_{t0}^* \in \partial_x g\left(\tilde{x}_t,\, t\right). \tag{6.18}$$

Далее, согласно формулам (6.12), (6.14) и (6.15) имеем

$$\left.\begin{aligned} &w^*(t) = \left(0, 0, \ldots, 0, x_t^*(t), x_{t+1}^*(t), 0, \ldots, 0\right),\\ &\left(x_t^*(t), x_{t+1}^*(t)\right) \in K_a^*\left((\tilde{x}_t, \tilde{x}_{t+1})\right),\quad t = 1, \ldots, T-1, \end{aligned}\right\} \tag{6.19}$$

$$w^*(0) = \left(x_1^*(0), \ldots, 0\right), \quad x_1^*(0) \in K^*_{a(x_0)}\left(\tilde{x}_1\right), \tag{6.20}$$

$$w^*(T) = \left(0, 0, \ldots, 0, x_T^*(T)\right), \quad x_T^*(T) \in K_M^*\left(\tilde{x}_T\right). \tag{6.21}$$

Подставив выражения (6.19)—(6.21) для $w^*(t)$ в формулу (6.17), получим в покомпонентном виде

$$\lambda x_{t0}^* = x_t^*(t-1) + x_t^*(t), \qquad t = 1, 2, \ldots, T. \tag{6.22}$$

Введем следующие обозначения:

$$x_{t+1}^* \equiv x_{t+1}^*(t), \quad t = 0, 1, \ldots, T-1, \quad x^* \equiv x_T^*(T).$$

Из соотношений (6.19) следует, что (см. определение III.2.1)

$$-x_t^*(t) \in a^*\left(x_{t+1}^*(t); \left(\tilde{x}_t, \tilde{x}_{t+1}\right)\right), \qquad t = 1, \ldots, T-1,$$

или, учитывая равенства (6.22) и новые обозначения,

$$x_t^* - \lambda x_{t0}^* \in a^*\left(x_{t+1}^*; \left(\tilde{x}_t, \tilde{x}_{t+1}\right)\right), \qquad t = 1, \ldots, T-1. \tag{6.23}$$

Из формулы (6.20) следует, что

$$x_1^* \in K^*_{a(x_0)}\left(\tilde{x}_1\right).$$

Согласно теореме I.3.8 это можно записать следующим образом:

$$\langle y - \tilde{x}_1, x_1^* \rangle \geqslant 0, \qquad y \in a(x_0),$$

т. е.

$$\langle \tilde{x}_1, x_1^* \rangle = \inf_y \left\{\langle y, x_1^* \rangle : \; y \in a(x_0)\right\},$$

или, в обозначениях § 2 главы III,

$$\tilde{x}_1 \in a\left(x_0; x_1^*\right). \tag{6.24}$$

Пусть теперь $x_{00}^* = 0$, x_0^* —произвольный элемент множества $a^*\left(x_1^*; \left(x_0, \tilde{x}_1\right)\right)$. В силу включения (6.24) и теоремы III.2.1 последнее множество не пусто. В силу этой же теоремы верно и обратное, т. е. если $a^*\left(x_1^*; \left(x_0, \tilde{x}_1\right)\right)$ не пусто, то выполнено (6.24). Таким образом, формула (6.20) эквивалентна включению (6.24), которое, в свою очередь, эквивлентно существованию вектора x_0^*

такого, что

$$x_0^* - \lambda x_{00}^* \in a^*\left(x_1^*; \left(x_0, \tilde{x}_1\right)\right), \tag{6.25}$$

т. е. распространение соотношения (6.23) на случай $t = 0$ охватывает условие (6.20).

Наконец, при $t = T$ соотношение (6.22) с учетом обозначения $x^* \equiv x_T^*(T)$ переходит в равенство

$$\lambda x_{T0}^* = x_T^* + x^*, \qquad x^* \in K_M^*\left(\tilde{x}_T\right). \tag{6.26}$$

Итак, применение теоремы 2.4 к рассматриваемой задаче позволяет получить следущий результат.

Теорема 6.2. *Пусть a — выпуклое отображение, $g(x, t)$ — выпуклые функции по x при $t = 1, 2, \ldots, T$, x_0 — фиксированный вектор. Пусть существует последовательность точек $\{x_t\}_{t=0}^T$, являющаяся траекторией, попадающей на выпуклое множество M, т. е.*

$$x_{t+1} \in a(x_t), \qquad t = 0,\ 1,\ \ldots,\ T-1, \quad x_T \in M,$$

и функция $g(x, t)$ непрерывна в точке x_t этой траектории, $t = 0, 1, \ldots, T$. Тогда для того, чтобы траектория $\{\tilde{x}_t\}_{t=0}^T$, начинающаяся в x_0 и заканчивающаяся в x_T, минимизировала функцию

$$f = \sum_{t=1}^{T} g(x_t, t)$$

по всем траекториям, необходимо, чтобы нашлись такие векторы x^*, x_t^*, $t = 0, 1, \ldots, T$, и число $\lambda = 0, 1$, что

$$\left.\begin{aligned} &x_t^* \in a^*\left(x_{t+1}^*; \left(\tilde{x}_t, \tilde{x}_{t+1}\right)\right) + \lambda \partial_x g\left(\tilde{x}_t, t\right), \\ &t = 0, 1, \ldots, T-1, \quad \partial_x g(x_0, 0) \equiv 0, \end{aligned}\right. \tag{6.27}$$

$$x_T^* + x^* \in \lambda \partial_x g\left(\tilde{x}_T, T\right), \tag{6.28}$$

$$x^* \in K_M^*\left(x_T^*\right). \tag{6.29}$$

При этом λ, x_t^* и x^* не все равны нулю одновременно. Если $\lambda = 1$, то эти условия достаточны.

Ясно, что соотношения (6.28) и (6.29) эквивалентны формулам (6.26), а включение (6.27) есть просто условие (6.23), переписанное по-другому с учетом (6.18).

Перепишем включение (6.27) в несколько ином виде. Для этого напомним, что согласно § 2 главы III

$$W_a(x, y^*) = \inf_y \{\langle y, y^* \rangle : y \in a(x)\},$$

$$a(x; y^*) = \{y \in a(x): \langle y, y^* \rangle = W_a(x, y^*)\}$$

и по теореме III.2.1 выполняется равенство

$$a^*(y^*; z) = \partial_x W_a(x, y^*),$$

если $y \in a(x; y^*)$, $z = (x, y)$. Согласно формулам (6.27) множество $a^*\left(x^*_{t+1}; (\tilde{x}_t, \tilde{x}_{t+1})\right)$ не пусто, а поэтому можно написать

$$a^*\left(x^*_{t+1}; (\tilde{x}_t, \tilde{x}_{t+1})\right) = \partial_x W_a\left(\tilde{x}_t, x^*_{t+1}\right), \tag{6.30}$$

$$\tilde{x}_{t+1} \in a\left(\tilde{x}_t; x^*_{t+1}\right). \tag{6.31}$$

Далее, так как

$$-W_a(x, y^*) = \sup_y \{-\langle y, y^* \rangle : y \in a(x)\},$$

то $-W_a(x, y^*)$ есть выпуклая функция y^* при условии замкнутости множества $a(x)$ и по теореме II.3.11

$$\partial_{y^*}(-W_a(x, y^*)) = -a(x; y^*).$$

Так как функция $-W_a(x, y^*)$ выпукла, то $W_a(x, y^*)$ — вогнутая функция, а для вогнутых функций естественно определить субдифференциал формулой

$$\partial_{y^*} W_a(x, y^*) = -\partial_{y^*}(-W_a(x, y^*)).$$

Отсюда и из предыдущей формулы получаем, что

$$\partial_{y^*} W_a(x, y^*) = a(x; y^*).$$

Тогда включения (6.31) можно переписать в виде

$$\tilde{x}_{t+1} \in \partial_{y^*} W_a\left(\tilde{x}_t, x^*_{t+1}\right). \tag{6.32}$$

С учетом формул (6.30) и (6.32) результат предыдущей теоремы можно записать в более симметричной форме.

Теорема 6.3. *Пусть выполнены условия предыдущей теоремы и $a(x)$ — замкнутое множество при каждом x. Тогда для того, чтобы траектория $\{\tilde{x}_t\}_{t+0}^T$ была оптимальной, необходимо, чтобы нашлись такие не все* *равные нулю векторы* x^*, x_t^*, $t=0, 1, \ldots, T$, *и число* $\lambda=0, 1$, *что*

$$\tilde{x}_{t+1} \in \partial_{y^*} W_a(\tilde{x}_t, x_{t+1}^*),$$

$$x_t^* \in \partial_x W_a(\tilde{x}_t, x_{t+1}^*) + \lambda \partial_x g(\tilde{x}_t, t),$$

$$t=0, 1, \ldots, T-1, \quad \partial_x g(\tilde{x}_0, 0)=0,$$

$$x_T^* + x^* \in \lambda \partial_x g(\tilde{x}_T, T),$$

$$x^* \in K_M^*(\tilde{x}_T).$$

Если $\lambda=1$, *то эти условия достаточны.*

Заметим, что благодаря тому, что функции $g(x, t)$ могут принимать значение $+\infty$, теоремы 6.2 и 6.3 охватывают также случай наличия так называемых фазовых ограничений. А именно, пусть при каждом t задано выпуклое множество D_t. Пусть $g_0(x, t)$ — непрерывная выпуклая функция x и

$$g(x, t) = g_0(x_t, t) + \delta(x|D_t).$$

Тогда очевидно, что траектория, минимизирующая функцию

$$f = \sum_{t=1}^{T} g(x_t, t),$$

должна минимизировать и функцию

$$f_0 = \sum_{t=1}^{T} g_0(x_t, t)$$

при дополнительном условии $x_t \in D_t$. Чтобы условия теоремы 6.2 были выполнены, достаточно потребовать, чтобы существовала траектория $\{x_t\}_{t=0}^{T}$, исходящая из начальной точки x_0 и попадающая на M и такая, что $x_t \in \operatorname{int} D_t$. Ясно, что при этом функции $g(x, t)$ будут непрерывны в точках x_t этой траектории.

Применение теоремы II.3.8 с учетом формулы (II.3.25) показывает, что

$$\partial_x g(x, t) = \partial_x g_0(x, t) - K_{D_t}^*(x). \tag{6.33}$$

Теперь теорема 6.2 трансформируется в следующую теорему.

Теорема 6.4. *Пусть $g_0(x, t)$, $t=1, \ldots, T$,— выпуклые непрерывные функции, D_t — выпуклые множества и существует такая траектория $\{x_t\}_{t=0}^T$, исходящая из x_0 и оканчивающаяся на M, что*

$$x_t \in \operatorname{int} D_t, \qquad t=1, \ldots, T.$$

Тогда для того, чтобы траектория $\{\tilde{x}_t\}_{t=0}^T$ минимизировала функцию

$$f_0 = \sum_{t=1}^{T} g_0(x_t, t)$$

среди всех траекторий, начинающихся в x_0 и оканчивающихся на M, удовлетворяющих условиям $x_t \in D_t$, $t=1, \ldots, T$, необходимо, чтобы нашлись такие не все равные нулю число $\lambda = 0, 1$ и векторы x^*, x_t^*, $t=0, \ldots$ $\ldots, T$, что

$$x_t^* \in a^*\left(x_{t+1}^*; (\tilde{x}_t, \tilde{x}_{t+1})\right) + \lambda \partial_x g_0(\tilde{x}_t, t) - K_{D_t}^*(\tilde{x}_t),$$

$$t=0, 1, \ldots, T-1, \quad \partial_x g_0(x_0, t) = 0, \quad K_{D_0}^*(x_0) = \{0\},$$

$$x_T^* + x^* \in \lambda \partial_x g_0(\tilde{x}_T, T) - K_{D_T}^*(\tilde{x}_T),$$

$$x^* \in K_M^*(\tilde{x}_T).$$

Заметим также, что фазовые ограничения могут быть учтены другим способом. А именно, если все точки траектории должны принадлежать некоторому множеству D, то можно переопределить многозначное отображение, введя новое отображение a^0:

$$a^0(x) = \begin{cases} a(x), & x \in D, \\ \varnothing, & x \notin D. \end{cases}$$

Это приведет несколько к другим формулировкам результатов.

3. Примеры. Чтобы проиллюстрировать применение предыдущих теорем, рассмотрим некоторые модели, задаваемые конкретными многозначными отображениями. При этом во всех моделях будем считать, что множество M совпадает со всем пространством, а функции $g(x, t)$ непрерывно дифференцируемы. В этом случае

$$\partial_x g(x, t) = g'_x(x, t),$$

где $g'_x(x, t)$ — градиент функции $g(x, t)$ по x, а $K^*_M(\widetilde{x}_T) = \{0\}$, так как конус $K_M(\widetilde{x}_T)$ совпадает со всем пространством.

Пример 6.1. Пусть $a(x) = Ax + U$, где A — $n \times n$-матрица, а U — выпуклое множество в $\mathbf{R}^n$.

В рассматриваемом случае оптимизационная задача приобретает следующий вид: требуется выбрать управление $u_t \in U$, $t = 0, 1, \ldots, T-1$, таким образом, чтобы траектория системы минимизировала функцию

$$f = \sum_{t=1}^{T} g(x_t, t).$$

Пусть $\{\widetilde{x}_t\}_{t=0}^{T}$ — оптимальная траектория, а $\widetilde{u}_t \in \in U$ — соответствующее управление, т. е.

$$\widetilde{x}_{t+1} = A\widetilde{x}_t + \widetilde{u}_t, \qquad t = 0, 1, \ldots, T-1.$$

Согласно примеру 3.2 из главы III имеем

$$a^*(y^*; z) = \begin{cases} A^*y^*, & \text{если} \quad y^* \in [\operatorname{con}(U - u_0)]^*, \\ \varnothing, & \text{если} \quad y^* \notin [\operatorname{con}(U - u_0)]^*, \end{cases} \tag{6.34}$$
$$z = (x, Ax + u_0).$$

Будем считать, что $a^* \neq \varnothing$.

Условие $y^* \in [\operatorname{con}(U - u_0)]^*$ эквивалентно тому, что

$$\langle u - u_0, y^* \rangle \geqslant 0, \qquad u \in U,$$

т. е.

$$\langle u_0, y^* \rangle = \inf_u \{\langle u, y^* \rangle : u \in U\}. \tag{6.35}$$

С учетом формул (6.34), (6.35) и того, что $K^*_M(\widetilde{x}_T) = \{0\}$ соотношения (6.27)—(6.29) могут быть переписаны в виде

$$\begin{gathered} x^*_t = A^*x^*_{t+1} + \lambda g'_x(\widetilde{x}_t, t), \\ \langle \widetilde{u}_t, x^*_{t+1} \rangle = \min_u \{\langle u, x^*_{t+1} \rangle : u \in U\}, \\ t = 0, 1, \ldots, T-1, \\ x^*_T = \lambda g'_x(\widetilde{x}_T, T), \end{gathered} \tag{6.36}$$

Из первого соотношения (6.36) следует, что $\lambda \neq 0$. В самом деле, если $\lambda = 0$, то тогда $x^*_t = 0$ при всех

t, чего не должно быть согласно утверждению теоремы 6.2. Поэтому $\lambda = 1$.

Итак, необходимыми и достаточными условиями оптимальности траектории в рассматриваемом случае являются условия (6.36) при $\lambda = 1$.

Пример 6.2. Пусть

$$X = \mathbf{R}^n, \quad Y = \mathbf{R}^n, \quad a(x) = \{y\colon \varphi(x, y) \leqslant 0\},$$

где $\varphi(z)$, $z = (x, y)$, — выпуклая непрерывная функция и существует такая точка z_1, что $\varphi(z_1) < 0$.

Оптимизационная задача приобретает следующий вид: найти последовательность $x_t \in \mathbf{R}^n$, $t = 0, 1, \ldots, T$, удовлетворяющую соотношениям

$$\varphi(x_t, x_{t+1}) \leqslant 0, \qquad t = 0, 1, \ldots, T-1,$$

и минимизирующую функцию

$$f = \sum_{t=0}^{T} g(x_t, t).$$

Согласно примеру 3.6 главы III

$$a^*(y^*; z) =$$
$$= \{\lambda x_0^*\colon y^* = -\lambda y_0^*, \ \ (x_0^*, y_0^*) \in \partial_z \varphi(z), \ \ \lambda \geqslant 0\}, \tag{6.37}$$

если $\varphi(z) = 0$. Если же $\varphi(z) < 0$, то, так как функция $\varphi(z)$ непрерывна и $\operatorname{gf} a = \{z\colon \varphi(z) < 0\}$, выполняются равенства $K_a(z) = \mathbf{R}^n \times \mathbf{R}^n$ и $K_a^*(z) = \{0\}$. Поэтому

$$a^*(y^*; z) = \begin{cases} 0, & \text{если } y^* = 0, \\ \varnothing, & \text{если } y^* \neq 0. \end{cases} \tag{6.38}$$

Формулы (6.37) и (6.38) можно объединить в одну формулу следующего вида:

$$a^*(y^*; z) =$$
$$= \{\lambda x_0^*\colon y^* = -\lambda y_0^*, \ \ (x_0^*, y_0^*) \in \partial_z \varphi(z), \ \ \lambda \geqslant 0,$$
$$\lambda \varphi(z) = 0\}.$$

Использование этой формулы и того факта, что $K_M^*(\tilde{x}_T) = \{0\}$, позволяет записать соотношения (6.27)—

(6.29) в виде

$$\begin{gathered} x_t^* = \lambda_t x_{0t}^* + \lambda g_x'(\tilde{x}_t, t), \\ x_{t+1}^* = -\lambda_t y_{0t}^*, \quad (x_{0t}^*, y_{0t}^*) \in \partial_z \varphi(\tilde{x}_t, \tilde{x}_{t+1}), \\ \lambda_t \varphi(\tilde{x}_t, \tilde{x}_{t+1}) = 0, \qquad \lambda_t \geqslant 0, \\ t = 0, 1, \ldots, T-1, \quad x_T^* = \lambda g_x'(\tilde{x}_T, T). \end{gathered} \tag{6.39}$$

Так как

$$\begin{gathered} \lambda_t x_{0t}^* = x_t^* - \lambda g_x'(\tilde{x}_t, t), \qquad \lambda_t y_{0t}^* = -x_{t+1}^*, \\ (\lambda_t x_{0t}^*, \lambda_t y_{0t}^*) \in \lambda_t \partial_z \varphi(\tilde{x}_t, \tilde{x}_{t+1}), \end{gathered}$$

то соотношения (6.39) можно записать следующим образом:

$$\begin{gathered} (x_t^* - \lambda g_x'(\tilde{x}_t, t), -x_{t+1}^*) \in \lambda_t \partial_z \varphi(\tilde{x}_t, \tilde{x}_{t+1}), \\ \lambda_t \varphi(\tilde{x}_t, \tilde{x}_{t+1}) = 0, \qquad \lambda_t \geqslant 0, \\ t = 0, 1, \ldots, T-1, \quad x_T^* = \lambda g_x'(\tilde{x}_T, T). \end{gathered} \tag{6.40}$$

Итак, для того чтобы траектория $\{\tilde{x}_t\}_{t=0}^T$ была оптимальной в рассматриваемой задаче, необходимо, чтобы нашлись такие числа $\lambda = 0, 1$, λ_t и векторы x_t^*, чтобы выполнялись соотношения (6.40).

Пример 6.3. Пусть $K \subseteq \mathbf{R}^n \times \mathbf{R}^n$ — конус, $\operatorname{gf} a = K$,

$$a(x) = \{y\colon (x, y) \in K\}.$$

Пусть, кроме того, $M = \mathbf{R}^n$, $g(x, t) \equiv 0$, $t = 1, 2, \ldots, T-1$. Таким образом,

$$\begin{gathered} \partial_x g(x, t) = 0, \qquad t = 1, 2, \ldots, T-1, \\ K_M^*(x) = \{0\}. \end{gathered} \tag{6.41}$$

В моделях экономической динамики чаще всего рассматривается ситуация, соответствующая этому примеру, а сама модель носит название *модели Неймана — Гейла*.

Согласно примеру 3.8 главы III в исследуемом случае

$$a^*(y^*; z) = \{x^*\colon (-x^*, y^*) \in K^*, \quad \langle x, x^*\rangle = \langle y, y^*\rangle\},$$
$$z = (x, y).$$

С учетом формул (6.41) соотношения (6.27), (6.28) приобретают следующий вид:

$$\begin{aligned} &\left(-x_t^*, x_{t+1}^*\right) \in K^*, \quad \langle \tilde{x}_t, x_t^* \rangle = \langle \tilde{x}_{t+1}, x_{t+1}^* \rangle, \\ &t = 0, 1, \ldots, T-1, \\ &x_T^* = \lambda g_x'\left(\tilde{x}_T, T\right), \quad \lambda = 0, 1. \end{aligned} \tag{6.42}$$

Рассмотрим теперь частный случай — так называемую *модель Неймана.* Пусть имеется m технологических способов производства, причем использование j-го технологического способа производства с единичной интенсивностью ведет к производству вектора a_j товаров, $a_j \in \mathbf{R}^n$, так что число различных производимых товаров равно n и при единичной интенсивности использования j-го технологического способа производится i-го товара в количестве a_j^i, $i = 1, \ldots, n$. Естественно считать, что $a_j^i \geqslant 0$. При этом при единичной интенсивности использования j-го технологического способа расходуется вектор $b_j \in \mathbf{R}^n$ товаров. Если теперь в данный момент времени имеется вектор x товаров, то интенсивности λ^j каждого способа производства должны, очевидно, удовлетворять неравенству

$$x \geqslant \sum_{j=1}^{m} b_j \lambda^j, \qquad \lambda^j \geqslant 0.$$

При этом будет выпущен вектор y товаров, удовлетворяющий уравнению

$$y = \sum_{j=1}^{m} a_j \lambda^j.$$

Взяв теперь в качестве столбцов матрицы $\mathscr{A}$ векторы a_j и аналогично образовав матрицу $\mathscr{B}$, можно сказать, что определено некоторое многозначное отображение $a(x)$, графиком которого является многогранный конус

$$K = \{(x, y): \ x \geqslant \mathscr{B}\lambda, \ y = \mathscr{A}\lambda, \quad \lambda \geqslant 0, \quad \lambda \in \mathbf{R}^m\},$$

где λ — вектор с компонентами λ^j. Из нашей интерпретации следуют неравенства $\mathscr{B} \geqslant 0$, $\mathscr{A} \geqslant 0$, т. е. матрицы $\mathscr{A}$ и $\mathscr{B}$ имеют неотрицательные коэффициенты.

Так как для $(x, y) \in K$

$$x = \mathscr{B}\lambda + v, \quad y = \mathscr{A}\lambda, \quad \lambda \geqslant 0, \quad v \geqslant 0, \quad v \in \mathbf{R}^n,$$

то нетрудно вычислить K^*. Действительно, $(x^*, y^*) \in K^*$ тогда и только тогда, когда

$$\langle x, x^* \rangle + \langle y, y^* \rangle \geqslant 0, \qquad (x, y) \in K,$$

т. е.

$$\langle \mathscr{B}\lambda + v, x^* \rangle + \langle \mathscr{A}\lambda, y^* \rangle \geqslant 0$$

для всех $v \geqslant 0$ и $\lambda \geqslant 0$. Переписав последнее неравенство в виде

$$\langle v, x^* \rangle + \langle \lambda, \mathscr{B}^* x^* + \mathscr{A}^* y^* \rangle \geqslant 0, \ v \geqslant 0, \ \lambda \geqslant 0,$$

где $\mathscr{B}^*$ и $\mathscr{A}^*$ — транспонированные матрицы $\mathscr{B}$ и $\mathscr{A}$, получаем очевидным образом неравенства

$$x^* \geqslant 0, \qquad \mathscr{B}^* x^* + \mathscr{A}^* y^* \geqslant 0.$$

Поэтому

$$K^* = \{(x^*, y^*) \colon \mathscr{B}^* x^* + \mathscr{A}^* y^* \geqslant 0, \ x^* \geqslant 0\}. \tag{6.43}$$

Пусть в начальный момент имеется вектор товаров x_0. Тогда к моменту T может быть произведено x_T товаров, где x_T получается из цепочки включений

$$x_{t+1} \in a(x_t), \qquad t = 0, 1, \ldots, T-1. \tag{6.44}$$

Если в момент T заданы цены товаров p^{i*}, $i = 1, \ldots, n$, то общая стоимость товаров в момент T будет определяться следующей формулой:

$$\sum_{i=1}^{n} p^{i*} x_T^i = \langle x_T, p^* \rangle. \tag{6.45}$$

Теперь можно рассмотреть задачу о выборе такого способа функционирования экономики, т. е. такой траектории, удовлетворяющей соотношениям (6.44), чтобы максимизировать $\langle x_T, p \rangle$. Последнее эквивалентно минимизации функции

$$g(x_T, T) = -\langle x_T, p^* \rangle.$$

Мы приходим к задаче, рассмотренной выше, со специальным конусом K и функцией $g(x_T, T)$, и для характеристики оптимальной траектории можно использовать соотношения (6.42). С учетом (6.43) эти соотношения могут быть записаны в следующем виде:

$$\begin{gathered} -x_t^* \geqslant 0, \quad -\mathscr{B}^* x_t^* + \mathscr{A}^* x_{t+1}^* \geqslant 0, \\ \langle \tilde{x}_t, x_t^* \rangle = \langle \tilde{x}_{t+1}, x_{t+1}^* \rangle, \\ t = 0, 1, \ldots, T-1, \\ x_T^* = -\lambda p^*, \quad \lambda = 0, 1. \end{gathered} \tag{6.46}$$

Заметим теперь, что если множества M_i в теореме 2.4 многогранные, то при ее доказательстве можно использовать теорему I.4.14, из которой следует, что в формулировке теоремы 2.4 можно положить $\lambda = 1$. Так как теорема 6.2 целиком основывалась на теореме 2.4, то в случае многогранного отображения a в теореме 2.4 λ следует брать равным 1. Поэтому $\lambda = 1$ и в формулах (6.46).

Положим теперь $p_t^* = -x_t^*$, $t = 0, 1, \ldots, T$, и запишем, соотношения (6.46) в следующем виде:

$$\begin{gathered} p_t^* \geqslant 0, \quad \mathscr{B}^* p_t^* - \mathscr{A}^* p_{t+1}^* \geqslant 0, \\ \langle \tilde{x}_t, p_t^* \rangle = \langle \tilde{x}_{t+1}, p_{t+1}^* \rangle, \\ t = 0, 1, \ldots, T-1, \\ p_T^* = p^*. \end{gathered} \tag{6.47}$$

Проделанные формальные преобразования приводят к следующему выводу: траектория $\{\tilde{x}_t\}_{t=0}^T$ модели Неймана, удовлетворяющая соотношениям (6.44), тогда и только тогда максимизирует функцию конечного дохода (6.45), когда существуют такие векторы p_t^*, что выполнены соотношения (6.47).

Рассмотрим возможную экономическую интерпретацию полученного результата. Назовем векторы p_t^* *ценами в момент t*. Это имеет смысл, поскольку $p_t^* \geqslant 0$.

Пусть y — произвольный вектор из $a(\tilde{x}_t)$. Это означает, что

$$y = \mathscr{A}\lambda, \quad \tilde{x}_t \geqslant \mathscr{B}\lambda, \quad \lambda \geqslant 0.$$

Основываясь на этих соотношениях и формулах (6.47), легко убедиться в справедливости следующей цепочки неравенств:

$$\langle y, \overset{*}{p}_{t+1}\rangle = \langle \mathscr{A}\lambda, \overset{*}{p}_{t+1}\rangle = \langle \lambda, \mathscr{A}^* \overset{*}{p}_{t+1}\rangle \leqslant \leqslant \langle \lambda, \mathscr{B}^* \overset{*}{p}_t\rangle = \langle \mathscr{B}\lambda, \overset{*}{p}_t\rangle \leqslant \langle \tilde{x}_t, \overset{*}{p}_t\rangle.$$

Но так как

$$\langle \tilde{x}_{t+1}, \overset{*}{p}_{t+1}\rangle = \langle \tilde{x}_t, \overset{*}{p}_t\rangle,$$

то

$$\langle \tilde{x}_{t+1}, \overset{*}{p}_{t+1}\rangle = \max_y \{\langle y, \overset{*}{p}_{t+1}\rangle\colon\ y \in a(\tilde{x}_t)\}.$$

Из полученного соотношения следует важный вывод: для оптимальной траектории существуют такие цены, что при выборе в момент t интенсивностей технологических способов производства оптимальной траектории соответствует тот, который обеспечивает максимальный доход в ценах момента $t+1$.

Глава V

НЕОБХОДИМЫЕ УСЛОВИЯ ЭКСТРЕМУМА

В этой главе строятся необходимые условия экстремума для задач, в которых исходные данные задаются множествами и функциями, не обязательно являющимися выпуклыми, но в то же время и не совершенно произвольными. Локально эти множества и функции должны обладать некоторыми свойствами. Эти свойства выражаются в понятиях конусов касательных направлений и шатров для множеств. Локальные свойства функций описываются верхней выпуклой аппроксимацией, определение которой дается во втором параграфе.

§ 1. Конусы касательных направлений и шатры

1. Касательные направления. Пусть M — произвольное множество в пространстве X.

Определение 1.1. Вектор $\bar{x} \in X$ называется *касательным* к множеству M в точке $x \in M$, если существует такая функция $\varphi(\lambda) \in X$, что

$$x + \lambda\bar{x} + \varphi(\lambda) \in M$$

при достаточно малых $\lambda \geqslant 0$ и $\lambda^{-1}\varphi(\lambda) \to 0$, когда $\lambda \downarrow 0$.

Касательные направления можно собирать в некоторые множества — обычно конусы, так как легко видеть, что если $\bar{x}$ — касательный вектор, то и вектор $\alpha\bar{x}$, $\alpha \geqslant 0$, также является касательным.

Определение 1.2. Выпуклый конус $K_M(x)$ называется *конусом касательных направлений* в точке x к множеству M, если из включения $\bar{x} \in K_M(x)$ следует, что $\bar{x}$ — касательный вектор к множеству M в точке $x \in M$.

Важно отметить, что конус $K_M(x)$ определен неоднозначно. В этом есть свои положительные стороны, которые будут более ясны из дальнейшего изложения. Во всяком случае необходимые условия минимума не будут зависеть по своей форме от выбора конуса $K_M(x)$. Однако чем этот конус шире, тем более существенны будут эти условия.

Если M — выпуклое множество, то легко видеть, что

$$K_M(x) = \operatorname{con}(M - x) = \{\bar{x}\colon \bar{x} = \lambda(x_1 - x),\\ x_1 \in M,\ \lambda > 0\} \tag{1.1}$$

есть конус касательных направлений к M. Для этого достаточо положить $\varphi(\lambda) \equiv 0$ в определение 1.1. Условимся, что если множество M — выпукло, то всегда в качестве $K_M(x)$ будет браться конус, задаваемый формулой (1.1).

Приведем примеры конусов касательных направлений.

Пример 1.1. Пусть множество M задано системой уравнений

$$f_i(x) = 0, \qquad i \in I, \tag{1.2}$$

где I — конечное множество индексов, а функции $f_i(x)$ непрерывно дифференцируемы.

Лемма 1.1. *Если $x_0 \in M$, т. е. x_0 удовлетворяет системе* (1.2), *и градиенты $f_i'(x_0)$, $i \in I$, линейно независимы, то*

$$K_M(x_0) = \{\bar{x}\colon \langle \bar{x}, f_i'(x_0)\rangle = 0, \quad i \in I\} \tag{1.3}$$

есть конус касательных направлений.

Доказательство. Пусть $f'(x_0)$ есть матрица, строками которой служат векторы $f_i'(x_0)$, $i \in I$. Пусть мощность множества I равна m, так что $f'(x_0)$ имеет размерность $m \times n$ ($X = \mathbf{R}^n$). Тогда условие $\bar{x} \in K_M(x_0)$ можно записать в виде

$$f'(x_0)\bar{x} = 0.$$

Пусть $(f'(x_0))^*$ — матрица, транспонированная к $f'(x_0)$, и

$$\mathscr{A} = f'(x_0)(f'(x_0))^*.$$

Квадратная матрица $\mathscr{A}$ размерности $m \times m$ невырождена. В самом деле, если предположить, что она вырождена, то найдется ненулевой вектор $y \in \mathbf{R}^m$ такой, что

$\mathcal{A}y=0$. Тогда

$$\langle y, \mathcal{A}y\rangle = \|(f'(x_0))^* y\|^2 = \Big\| \sum_{j\in I} f_j'(x_0)\, y^j \Big\|^2 = 0,$$

т. е.

$$\sum_{j\in I} f_j'(x_0)\, y^j = 0$$

при условии, что не все y^j равны нулю. Последнее равенство означает, что векторы $f_i'(x_0)$, $i\in I$, линейно зависимы в противоречии с предположением. Итак, $\mathcal{A}$ — невырожденная матрица.

Рассмотрим систему уравнений

$$g_i(\lambda,\ y) \equiv f_i(x_0 + \lambda\bar{x} + (f'(x_0))^* y) = 0, \qquad i\in I, \quad (1.4)$$

где y — неизвестное, а λ — переменный параметр. Нетрудно посчитать, что

$$\frac{\partial g_i(0,\ 0)}{\partial\lambda} = \langle\bar{x},\ f_i'(x_0)\rangle = 0, \qquad i\in I,$$

а матрица с элементами $\dfrac{\partial g_i(0,\ 0)}{\partial y^j}$, $i,\ j=1,\ \ldots,\ m$, совпадает с $\mathcal{A}$ и поэтому невырождена. По доказываемой ниже теореме система (1.4) имеет при достаточно малых $\lambda>0$ решение $y(\lambda)$, причем $\lambda^{-1}y(\lambda)\to 0$, если $\lambda\to 0$. Положим

$$\varphi(\lambda) = (f'(x_0))^* y(\lambda).$$

Тогда $\lambda^{-1}\varphi(\lambda)\to 0$ при $\lambda\to 0$ и согласно системе (1.4)

$$x_0 + \lambda\bar{x} + \varphi(\lambda) \in M$$

при достаточно малых $\lambda>0$, т. е. x — действительно касательный вектор.

Вычислим конус, сопряженный к конусу, задаваемому соотношением (1.3). Поскольку система уравнений

$$\langle\bar{x},\ f_i'(x_0)\rangle = 0, \qquad i\in I,$$

эквивалентна системе неравенств

$$\langle\bar{x},\ f_i'(x_0)\rangle \geqslant 0, \qquad i\in I,$$
$$\langle\bar{x},\ -f_i'(x_0)\rangle \geqslant 0, \qquad i\in I,$$

то согласно теореме I.4.9 конус $K_M^*(x_0)$ состоит из элементов вида

$$x^* = \sum_{i\in I} \lambda_i^+ f_i'(x_0) - \sum_{i\in I} \lambda_i^- f_i'(x_0),$$
$$\lambda_i^+ \geqslant 0, \qquad \lambda_i^- \geqslant 0.$$

Обозначая $\lambda_i = \lambda_i^+ - \lambda_i^-$ получаем

$$K_M^*(x_0) = \left\{x^*\colon\ x^* = \sum_{i\in I} \lambda_i f_i'(x_0), \quad \lambda_i \in \mathbf{R}^1, \quad i \in I\right\}. \tag{1.5}$$

Пример 1.2. Пусть множество M задано системой равенств и неравенств:

$$M = \{x\colon\ f_i(x) \leqslant 0, \quad i \in I^-, \quad f_i(x) = 0,\ i \in I\},$$

где I^- и I — конечные множества индексов, а функции $f_i(x)$ непрерывно дифференцируемы. Пусть $x_0 \in M$ и

$$I^-(x_0) = \{i \in I^-\colon\ f_i(x_0) = 0\}.$$

Если векторы $f_i(x_0)$, $i \in I$, линейно независимы, то направление $\bar{x}$, удовлетворяющее соотношениям

$$\begin{aligned} &\langle \bar{x}, f_i'(x_0)\rangle < 0, \qquad && i \in I^-(x_0), \\ &\langle \bar{x}, f_i'(x_0)\rangle \doteq 0, \qquad && i \in I, \end{aligned} \tag{1.6}$$

является касательным. Действительно, согласно предыдущему примеру существует такая функция $\varphi(\lambda)$ $(\lambda^{-1}\varphi(\lambda) \to 0$, если $\lambda \downarrow 0)$, что

$$f_i(x_0 + \lambda\bar{x} + \varphi(\lambda)) = 0, \qquad i \in I.$$

С другой стороны, для $i \in I^-$ согласно известной формуле анализа имеем

$$f_i(x_0 + \lambda\bar{x} + \varphi(\lambda)) = f_i(x_0) + \lambda\langle \bar{x} + \lambda^{-1}\varphi(\lambda), \quad f_i'(\xi_i)\rangle, \tag{1.7}$$

где ξ_i — точка отрезка, соединяющего x_0 и $x_0 + \lambda\bar{x} + \varphi(\lambda)$. Из формулы (1.7) следует, что если $i \in I^- \setminus I^-(x_0)$, то $f_i(x_0) < 0$, и

$$f_i(x_0 + \lambda\bar{x} + \varphi(\lambda)) < 0$$

при достаточно малых $\lambda > 0$. Если же $i \in I^-(x_0)$, то формула (1.7) показывает, что

$$f_i(x_0 + \lambda\bar{x} + \varphi(\lambda)) =$$
$$= \lambda\langle\bar{x}, f_i'(x_0)\rangle + \lambda\langle\bar{x}, f_i'(\xi_i) - f_i'(x_0)\rangle + \langle\varphi(\lambda), f_i'(\xi_i)\rangle.$$

Так как $\xi_i \to x_0$ при $\lambda \to 0$, то последние два слагаемых стремятся к нулю, причем быстрее, чем λ. Поэтому в силу неравенства (1.6)

$$f_i(x_0 + \lambda\bar{x} + \varphi(\lambda)) < 0, \quad i \in I^-(x_0),$$

при малых $\lambda > 0$.

Итак, при достаточно малых $\lambda > 0$ выполняются соотношения

$$f_i(x_0 + \lambda\bar{x} + \varphi(\lambda)) < 0, \quad i \in I^-,$$
$$f_i(x_0 + \lambda\bar{x} + \varphi(\lambda)) = 0, \quad i \in I,$$

т. е. $x_0 + \lambda\bar{x} + \varphi(\lambda) \in M$. Полученное включение означает, что вектор $\bar{x}$ — касательное направление к множеству M. Таким образом,

$$K_M(x_0) =$$
$$= \{\bar{x} : \langle\bar{x}, f_i'(x_0)\rangle < 0,\ i \in I^-(x_0),\ \langle\bar{x}, f_i'(x_0)\rangle = 0,\ i \in I\}. \tag{1.8}$$

Применяя теорему I.4.9, нетрудно показать, что

$$K_M^*(x_0) = \Big\{x^* : x^* =$$
$$= -\sum_{i \in I^-(x_0)} \lambda_i f_i'(x_0) - \sum_{i \in I} \lambda_i f_i'(x_0),\ \lambda_i \geqslant 0,\ i \in I^-(x_0)\Big\}. \tag{1.9}$$

2. Теорема о неявных функциях. Как было видно из приведенного выше примера, вычисление касательных направлений невозможно без применения теорем о разрешимости тех или иных систем нелинейных уравнений. Такого рода теоремы необходимы и в дальнейшем — при построении шатров и получении необходимых условий экстремума.

Т е о р е м а 1.1. *Пусть $x \in \mathbf{R}^n$, $y \in \mathbf{R}^k$ и $g(y, x)$ — непрерывная вектор-функция с компонентами $g_i(y, x)$, $i =$*

$=1, \ldots, n$, *удовлетворяющая следующему условию: существует такая невырожденная* $n \times n$-*матрица* g'_x, *что*

$$\|g(y, x) - g'_x x\| \leqslant \bar{r}\left(\sqrt{\|x\|^2 + \|y\|^2}\right),$$

где $\bar{r}(\lambda) = o(\lambda)$. *Тогда при достаточно малых* y *система уравнений* $g(y, x) = 0$ *разрешима относительно* x, *причем существует такое решение* $x(y)$, *что*

$$\lim_{y \to 0} \frac{\|x(y)\|}{\|y\|} = 0.$$

Замечание. Из условий теоремы легко следует, что:

а) $g(0, 0) = 0$;

б) существуют первые частные производные функции $g(y, x)$ в точке $y = 0$, $x = 0$, причем

$$g'_{iy}(0, 0) = 0, \qquad i = 1, \ldots, n;$$

в) матрица g'_x — матрица частных производных $\dfrac{\partial g_i(0, 0)}{\partial x^j}$, $i, j = 1, \ldots, n$,— невырождена.

Если предположить, что функции $g_i(y, x)$ непрерывно дифференцируемы в окрестности точки $y = 0$, $x = 0$, то легко убедиться, что из условий а), б) и в) следует выполнение условий теоремы 1.1. Более того, в этом случае она совпадает с обычной теоремой о неявных функциях, из которой следует, что $x(y)$ есть непрерывно дифференцируемая функция y в окрестности начала координат.

Доказательство. Из условий теоремы следует, что

$$\begin{gathered} g(y, x) = g'_x x + r(y, x), \\ \|r(y, x)\| \leqslant \bar{r}\left(\sqrt{\|x\|^2 + \|y\|^2}\right). \end{gathered} \tag{1.10}$$

Рассмотрим отображение

$$f(y, x) = x - (g'_x)^{-1} g(y, x).$$

Используя соотношение (1.10), получаем

$$\begin{gathered} f(y,\ x) = -\left(g_x'\right)^{-1} r(y,\ x), \\ \|f(y,\ x)\| \leqslant \left\|\left(g_x'\right)^{-1}\right\| \bar{r}\left(\sqrt{\|x\|^2 + \|y\|^2}\right). \end{gathered} \tag{1.11}$$

Заметим теперь, что без ограничения общности можно считать $\bar{r}(\lambda)$ неубывающей функцией λ. В самом деле, в случае, если это не так, заменим функцию $\bar{r}(\lambda)$ на функцию

$$\omega(\lambda) = \sup_{0 \leqslant t \leqslant \lambda} \bar{r}(t)$$

и покажем, что $\omega(\lambda)$ — неубывающая функция и

$$\frac{\omega(\lambda)}{\lambda} \to 0.$$

Действительно, так как $\lambda^{-1}\bar{r}(\lambda) \to 0$, то для всякого $\varepsilon > 0$ найдется такое $\delta > 0$, что

$$\frac{\bar{r}(t)}{t} \leqslant \varepsilon,$$

как только $t \leqslant \delta$. Если теперь $\lambda < \delta$, то

$$\frac{\omega(\lambda)}{\lambda} = \frac{1}{\lambda} \sup_{0 \leqslant t \leqslant \lambda} \bar{r}(t) \leqslant \sup_{0 \leqslant t \leqslant \lambda} \frac{\bar{r}(t)}{t} \leqslant \varepsilon,$$

т. е. $\lambda^{-1}\omega(\lambda) \to 0$.

Пусть теперь

$$\tau(y) = \inf_{\tau} \left\{\tau : c\bar{r}\left(\sqrt{\|y\|^2 + \tau^2}\right) \leqslant \tau\right\},$$

где $c = \left\|\left(g_x'\right)^{-1}\right\|$. При достаточно малых y выполняется неравенство

$$c\bar{r}\sqrt{\|y\|^2 + \|y\|^2} = c\bar{r}\left(\sqrt{2}\,\|y\|\right) \leqslant \|y\|,$$

поэтому $\tau(y) \leqslant \|y\|$. Кроме того, по определению точной нижней грани для каждого y существует такое $\tau^*(y)$, что

$$\begin{gathered} \tau(y) \leqslant \tau^*(y) \leqslant \tau(y) + \|y\|^2, \\ cr\left(\sqrt{\|y\|^2 + (\tau^*(y))^2}\right) \leqslant \tau^*(y). \end{gathered} \tag{1.12}$$

Покажем, что

$$\frac{\tau(y)}{\|y\|} \to 0,$$

когда $y \to 0$. По определению $\tau(y)$ имеем

$$\tau(y) - \|y\|^2 < c\bar{r}\left(\sqrt{\|y\|^2 + (\tau(y) - \|y\|^2)^2}\right) \leqslant$$
$$\leqslant c\bar{r}(\|y\| + |\tau(y) - \|y\|^2|), \qquad (1.13)$$

где использовано то, что $\bar{r}(\lambda)$ — неубывающая функция и

$$\sqrt{\alpha^2 + \beta^2} \leqslant \alpha + \beta, \quad \alpha > 0, \quad \beta > 0.$$

Далее, так как $\tau(y) \leqslant \|y\|$ для малых y, то правую часть неравенства (1.13) можно оценить следующим образом:

$$c\bar{r}(\|y\| + |\tau(y) - \|y\|^2|) \leqslant c\bar{r}(2\|y\|),$$

откуда

$$\tau(y) - \|y\|^2 \leqslant c\bar{r}(2\|y\|),$$

или

$$\frac{\tau(y)}{\|y\|} \leqslant 2c\,\frac{\bar{r}(2\|y\|)}{2\|y\|} + \|y\|,$$

т. е.

$$\frac{\tau(y)}{\|y\|} \to 0$$

при $y \to 0$. Но тогда из неравенства (1.12) следует, что

$$\frac{\tau^*(y)}{\|y\|} \to 0.$$

Если теперь

$$\|x\| \leqslant \tau^*(y),$$

то согласно неравенствам (1.11) и (1.12) справедливы неравенства

$$\|f(y, x)\| \leqslant c\bar{r}\left(\sqrt{\|y\|^2 + \|x\|^2}\right) \leqslant c\bar{r}\left(\sqrt{\|y\|^2 + (\tau^*(y))^2}\right) \leqslant \tau^*(y)$$

Отсюда следует, что непрерывное отображение $f(y, x)$ отображает шар $\|x\| \leqslant \tau^*(y)$ в себя. На основании теоремы Брауэра можно утверждать, что отображение $f(y, x)$ имеет неподвижную точку в этом шаре, т. е. существует такая точка $x(y)$, что

$$x(y) = f(y, x(y)), \quad \|x(y)\| \leqslant \tau^*(y).$$

Но из определения $f(y, x)$ следует, что

$$g(y, x(y)) = 0.$$

Кроме того,

$$\frac{\|x(y)\|}{\|y\|} \leqslant \frac{\tau^*(y)}{\|y\|},$$

т. е. $\|y\|^{-1}x(y) \to 0$ при $y \to 0$. Теорема доказана.

Покажем, что проверка условий теоремы может быть упрощена.

Лемма 1.2. *Пусть функция* $f(x)$, $x \in \mathbf{R}^n$, *имеет в точке* x_0 *градиент* $f'(x_0)$ *и*

$$\lim_{\lambda \downarrow 0} \frac{f\left(x_0 + \lambda\bar{x} + \varphi(\lambda)\right) - f(x_0)}{\lambda} = \langle \bar{x}, f'(x_0)\rangle$$

для любого вектора $\bar{x} \in \mathbf{R}^n$ *и любой функции* $\varphi(\lambda) \in \mathbf{R}^n$, *определенной для малых* $\lambda > 0$ *и удовлетворяющей условию* $\lambda^{-1}\varphi(\lambda) \to 0$ *при* $\lambda \downarrow 0$. *Тогда существует такая функция* $\bar{r}(\alpha)$, $\alpha^{-1}\bar{r}(\alpha) \to 0$ *при* $\alpha \downarrow 0$, *что*

$$|f(x) - f(x_0) - \langle x - x_0, f'(x_0)\rangle| \leqslant \bar{r}(\|x - x_0\|).$$

Доказательство. Без ограничения общности можно считать, что $x_0 = 0$, $f(x_0) = 0$, так что

$$\lim_{\lambda \downarrow 0} \frac{f(\lambda\bar{x} + \varphi(\lambda))}{\lambda} = \langle \bar{x}, f'(0)\rangle. \tag{1.13'}$$

Допустим, что лемма неверна. Тогда найдется такая последовательность точек $x_k \to 0$, что

$$\frac{|f(x_k) - \langle x_k, f'(0)\rangle|}{\|x_k\|} \geqslant \varepsilon > 0. \tag{1.14}$$

Векторы

$$\bar{x}_k = \frac{x_k}{\|x_k\|}$$

имеют единичную норму, и поэтому из них можно выбрать сходящуюся подпоследовательность. Без ограничения общности можно предполагать, что $\bar{x}_k \to \bar{x}$. Положим

$$\lambda_k = \|x_k\|, \quad \varphi(\lambda_k) = \lambda_k(\bar{x}_k - \bar{x}). \tag{1.15}$$

Можно считать, что $\lambda_{k+1} < \lambda_k$. На промежутках $[\lambda_{k+1}, \lambda_k]$ зададим функцию $\varphi(\lambda)$ при помощи линейной интерполяции. Тогда из соотношений (1.15) и того, что $\bar{x}_k \to \bar{x}$, следует, что $\lambda^{-1}\varphi(\lambda) \to 0$. Поэтому справедлива формула (1.13′). С другой стороны,

$$x_k = \lambda_k \bar{x} + \varphi(\lambda_k)$$

и в силу оценки (1.14) выполняется неравенство

$$\left| \lim_{k\to\infty} \frac{f\left(\lambda_k \bar{x} + \varphi(\lambda_k)\right)}{\lambda_k} - \langle \bar{x}, f'(0)\rangle \right| \geqslant \varepsilon > 0,$$

что противоречит соотношению (1.13′). Полученное противоречие доказывает лемму.

Из результатов леммы 1.2 вытекает, что если функции $g_i(y, x)$ удовлетворяют условиям а), б) и в) замечания к теореме 1.1 и, кроме того, выполняется соотношение

$$\lim_{\lambda \downarrow 0} \frac{g_i\left(\lambda \bar{y} + \varphi_1(\lambda), \lambda \bar{x} + \varphi_2(\lambda)\right)}{\lambda} = \langle \bar{x}, g'_{ix}(0, 0)\rangle$$

для любых $\varphi_1(\lambda)$ и $\varphi_2(\lambda)$, стремящихся к нулю быстрее, чем λ, то справедлива теорема 1.1. Чтобы в этом убедиться, достаточно вместо функции $f(x)$ в условиях теоремы взять $g_i(y, x)$, в качестве аргумента x взять пару (y, x), вместо x_0 — точку (0, 0) и учесть условия а) и б) замечания.

3. Шатры. Конус касательных направлений содержит направления, для каждого из которых в определении 1.1 существует своя функция $\varphi(\lambda)$. Этого оказывается недостаточно для того, чтобы судить о свойствах множества M. Вводимое ниже понятие шатра позволяет судить об отображении в M близких между собой касательных направлений.

О п р е д е л е н и е 1.3. Конус касательных направлений $K_M(x_0)$ в точке $x_0 \in M$ называется *локальным шатром*, если для каждого направления $\bar{x}_0 \in \operatorname{ri} K_M(x_0)$ существует выпуклый конус Q и определенное в окрестности начала координат непрерывное отображение $\psi(\bar{x})$ такие, что:

а) $\bar{x}_0 \in \operatorname{ri} Q$, $\operatorname{Lin} Q = \operatorname{Lin} K_M(x_0)$ и $Q \subseteq K_M(x_0)$;

б) $\psi(\bar{x}) = \bar{x} + r(\bar{x})$, $\|\bar{x}\|^{-1} r(\bar{x}) \to 0$, если $\bar{x} \to 0$;

в) $x + \psi(\bar{x}) \in M$ для $\bar{x} \in Q \cap (\varepsilon B)$ при некотором $\varepsilon > 0$. (Напомним, что B — единичный шар в пространстве $X = \mathbf{R}^n$.)

Пример 1.3. Пусть множество M задано, как в примере 1.1, и в точке x_0 выполнены условия леммы 1.1. Покажем, что конус $K_M(x_0)$, определяемый формулой (1.3), является локальным шатром. Для этого составим систему уравнений

$$g_i(\bar{x}, y) \equiv f_i(x_0 + \bar{x} + (f'(x_0))^* y) - \langle \bar{x}, f_i'(x_0) \rangle = 0, \quad i \in I, \tag{1.16}$$

в которой неизвестным является y, а $\bar{x}$ — параметр. Нетрудно видеть, что

$$g_i(0,\ 0) = 0, \qquad i \in I,$$

$$g'_{i\bar{x}}(0,\ 0) = 0, \qquad i \in I,$$

а матрица производных $\dfrac{\partial g_i(0,\ 0)}{\partial y^j}$, $i, j \in I$, совпадает с матрицей

$$\mathcal{A} = f'(x_0)(f'(x_0))^*$$

и поэтому невырождена. Так как функции $f_i(x)$ непрерывно дифференцируемы, то выполнены условия теоремы 1.1. Согласно замечанию к теореме при достаточно малых $\bar{x}$ существует гладкое решение $y(\bar{x})$ такое, что

$$\frac{y(\bar{x})}{\|\bar{x}\|} \to 0. \tag{1.17}$$

Положим $Q = K_M(x_0)$ и

$$\psi(\bar{x}) = \bar{x} + (f'(x_0))^* y(\bar{x}). \tag{1.18}$$

Возьмем $\bar{x} \in Q \cap (\varepsilon B)$, где $\varepsilon > 0$, так, чтобы $y(\bar{x})$ было определено в шаре εB. В силу формулы (1.3) соотношение (1.16) переходит в равенство

$$f_i(x_0 + \psi(\bar{x})) = 0, \qquad i \in I,$$

т. е. $x_0+\psi(\bar{x})\in M$. Отсюда и из формул (1.17), (1.18) следует, что $K_M(x_0)$ — локальный шатер.

Пример 1.4. Пусть множество M задано, как в примере 1.2. Покажем, что конус $K_M(x_0)$, определяемый формулой (1.8), является локальным шатром. Нетрудно видеть, что $\operatorname{ri} K_M(x_0)=K_M(x_0)$.

Пусть $\bar{x}_0\in K_M(x_0)$. Положим

$$\delta=\frac{1}{\|\bar{x}_0\|}\max_i\{\langle\bar{x}_0, f_i'(x_0)\rangle:\ i\in I^-(x_0)\}<0$$

и

$$Q=\{\bar{x}:\ \langle\bar{x}, f_i'(x_0)\rangle\leqslant\delta\|\bar{x}\|,\qquad i\in I^-(x_0),$$
$$\langle\bar{x}, f_i'(x_0)\rangle=0,\qquad i\in I\}. \quad (1.19)$$

Тогда $\bar{x}_0\in\operatorname{ri} Q$ и $Q\subseteq K_M(x_0)$. Выберем функцию $\psi(\bar{x})$, как в примере 1.3. На основании теоремы о среднем нетрудно убедиться в том, что

$$f_i\left(x_0+\psi(\bar{x})\right)=f_i(x_0)+\langle\bar{x}, f_i'(x_0)\rangle+o(\|x\|),\qquad i\in I^-. \quad (1.20)$$

Если $i\in I^-\setminus I^-(x_0)$, то $f_i(x_0)<0$. Тогда из формулы (1.20) следует, что

$$f_i(x_0+\psi(\bar{x}))<0,\qquad i\in I^-\setminus I^-(x_0), \quad (1.21)$$

при малых $\bar{x}$. Если $i\in I^-(x_0)$, то для $\bar{x}\in Q$

$$f_i(x_0+\psi(\bar{x}))\leqslant\delta\|\bar{x}\|+o(\|\bar{x}\|)<0,\qquad i\in I^-(x_0), \quad (1.22)$$

при малых $\bar{x}$. Наконец,

$$f_i(x_0+\psi(\bar{x}))=0,\qquad i\in I, \quad (1.23)$$

по построению функции $\psi(\bar{x})$. Из формул (1.21) — (1.23) следует, что при малых $\bar{x}\in Q$ сумма $x_0+\psi(\bar{x})\in M$, т. е. $K_M(x_0)$ — локальный шатер.

Итак, если M задано, как в примере 1.2, и в точке x_0 векторы $f_i'(x_0)$, $i\in I$, линейно независимы, то конус $K_M(x_0)$, задаваемый формулой (1.8), есть локальный шатер.

Пример 1.5. Пусть M — выпуклое множество. Тогда

$$K_M(x_0)=\operatorname{con}(M-x_0),\qquad x_0\in M,$$

есть локальный шатер. Не ограничивая общности, можно считать, что $x_0 = 0$, $0 \in M$ и M имеет внутренние точки. В случае, если $\operatorname{int} M = \varnothing$, нужно рассматривать M относительно пространства $\operatorname{Lin} M$.

Итак,

$$K_M(0) = \{\bar{x}:\ \bar{x} = \lambda x,\ \lambda > 0,\ x \in M\}.$$

Пусть $\bar{x}_0 \in \operatorname{int} K_M(0)$, $x_0 \neq 0$. Тогда

$$\bar{x}_0 + \varepsilon B \subseteq K_M(0), \qquad 0 \notin (\bar{x}_0 + \varepsilon B)$$

для достаточно малых $\varepsilon > 0$ (напомним, что B — единичный шар с центром в нуле). Выберем точки $\bar{x}_i$, $i = 1, \ldots$ $\ldots, n+1$, так, чтобы $\bar{x}_i \in (\bar{x}_0 + \varepsilon B)$ и чтобы точка $\bar{x}_0$ была центром симплекса S, натянутого на точки $\bar{x}_i$. Это возможно, как неоднократно было показано в главе I. По доказанному в главе I (см., например, доказательство теоремы I.1.2 и леммы I.1.6) существует такое $\varepsilon_1 > 0$, что

$$\bar{x}_0 + \varepsilon_1 B \subseteq S \subseteq K_M(0).$$

Так как $\bar{x}_i \in K_M(0)$, то $\bar{x}_i = \lambda_i x_i$, $x_i \in M$. Ясно, что $\lambda_i > 0$, так как $\bar{x}_i \neq 0$, ибо $0 \notin (\bar{x}_0 + \varepsilon B)$, а $\bar{x}_i \in (\bar{x}_0 + \varepsilon B)$. Положим

$$\lambda_0 = \min\left\{\frac{1}{\lambda_i}:\ i = 1, \ldots, n+1\right\}.$$

Так как $0 \in M$, $\bar{x}_i = \lambda_i x_i$, $x_i \in M$, то

$$\lambda_0 \bar{x}_i = (1 - \lambda_0\lambda_i)\cdot 0 + (\lambda_0\lambda_i) x_i \in M,$$

ибо $\lambda_0\lambda_i \leqslant 1$ по определению λ_0. Поэтому $\lambda_0 S \subseteq M$. Значит,

$$\lambda_0(\bar{x}_0 + \varepsilon_1 B) \subseteq \lambda_0 S \subseteq M.$$

Положим теперь

$$Q = \operatorname{con}(\bar{x}_0 + \varepsilon_1 B) = \{\bar{x}:\ \bar{x} = \gamma(\bar{x}_0 + \varepsilon_1 u), \quad u \in B, \quad \gamma > 0\}.$$

Очевидно, что $\bar{x}_0 \in \operatorname{int} Q$, $Q \subseteq K_M(0)$. Так как $0 \notin (\bar{x}_0 + \varepsilon B)$ и $\varepsilon_1 \leqslant \varepsilon$, то

$$\delta = \min_u \{\|\bar{x}_0 + \varepsilon_1 u\|:\ u \in B\} > 0.$$

Поэтому для $\bar{x} \in Q$ выполняется соотношение

$$\|\bar{x}\| = \gamma \|\bar{x}_0 + \varepsilon_1 u\| \geqslant \gamma\delta,$$

или

$$\gamma \leqslant \frac{1}{\delta} \|\bar{x}\|.$$

Выберем теперь $\varepsilon_2 > 0$ настолько малым, чтобы $\varepsilon_2/\delta < \lambda_0$. Тогда, если $\bar{x} \in Q$, $\|\bar{x}\| < \varepsilon_2$, то

$$\bar{x} = \gamma(\bar{x}_0 + \varepsilon_1 u), \qquad \gamma \leqslant \frac{\|\bar{x}\|}{\delta} \leqslant \frac{\varepsilon_2}{\delta} < \lambda_0.$$

Поэтому

$$\bar{x} = \gamma(\bar{x}_0 + \varepsilon_1 u) = \\ = \left(\frac{\gamma}{\lambda_0}\right)(\lambda_0(\bar{x}_0 + \varepsilon_1 u)) \in \left(\frac{\gamma}{\lambda_0}\right)(\lambda_0(\bar{x}_0 + \varepsilon_1 B)) \subseteq \left(\frac{\gamma}{\lambda_0}\right) M,$$

откуда в силу того, что $\gamma < \lambda_0$, $0 \in M$, получаем

$$\bar{x} \in \left(\frac{\gamma}{\lambda_0}\right) M \subseteq M,$$

т. е. $\bar{x} \in M$. Тем самым доказано, что если $\bar{x} \in Q$, $\|\bar{x}\| < \varepsilon_2$, то $\psi(\bar{x}) = \bar{x} \in M$. Это означает, что $K_M(0)$ — гладкий локальный шатер к M в точке нуль.

Если $\bar{x}_0 \in \operatorname{int} K_M(0)$, $\bar{x}_0 = 0$, то, повторяя предыдущие рассуждения, убеждаемся, что $\varepsilon_1 B \subseteq M$ для достаточно малого $\varepsilon_1 > 0$, а $K_M(0) = X$. Поэтому, полагая в данном случае $Q = X$, для $\bar{x}$, $\|\bar{x}\| < \varepsilon_1$, получаем, что $\psi(\bar{x}) = \bar{x} \in M$, что завершает доказательство.

Важнейшее свойство шатров состоит в том, что при достаточно общих предположениях пересечение локальных шатров есть снова локальный шатер.

Теорема 1.2. *Пусть M_i, $i \in I$, — множества в $\mathbf{R}^n$, I — конечное множество индексов, $x_0 \in M = \bigcap_{i \in I} M_i$ и $K_{M_i}(x_0)$, $i \in I$, — локальные шатры множеств M_i в точке x_0. Если шатры $K_{M_i}(x_0)$, $i \in I$, неотделимы, то для любого вектора $\bar{x}_0 \in \mathbf{ri}\, K$, где*

$$K = \bigcap_{i \in I} K_{M_i}(x_0),$$

существует такой конус Q и такая функция $\psi(\bar{x})$, **что:**

а) $\bar{x}_0 \in \operatorname{ri} Q$, $\operatorname{Lin} Q = \operatorname{Lin} K$ *и* $Q \subseteq K$;

б) $\psi(\bar{x}) = \bar{x} + r(\bar{x})$, $\|\bar{x}\|^{-1} r(\bar{x}) \to 0$, *если* $\bar{x} \to 0$;

в) $x_0 + \psi(\bar{x}) \in M$ *для* $\bar{x} \in Q \cap (\varepsilon B)$ *при некотором* $\varepsilon > 0$.

Замечание. Функция $\psi(\bar{x})$ может и не быть непрерывной, и поэтому K, вообще говоря, не есть локальный шатер к M в точке x_0.

Доказательство. Ограничимся случаем двух множеств M_1 и M_2. Для упрощения будем считать, что $x_0 = 0$, и обозначим $K_i = K_{M_i}(x_0)$, $i = 1, 2$.

Рассмотрим конус $K = K_1 \cap K_2$. Так как конусы K_1 и K_2 неотделимы, то $\operatorname{ri} K_1 \cap \operatorname{ri} K_2 \neq \varnothing$ и по теореме I.1.4

$$\operatorname{Lin} K = \operatorname{Lin} K_1 \cap \operatorname{Lin} K_2, \quad \operatorname{ri} K = \operatorname{ri} K_1 \cap \operatorname{ri} K_2. \tag{1.24}$$

Кроме того, по теореме I.3.4

$$\operatorname{Lin} K_1 + \operatorname{Lin} K_2 = \mathbf{R}^n. \tag{1.25}$$

Пусть теперь $\bar{x}_0 \in \operatorname{ri} K$. Согласно формулам (1.24) $\bar{x}_0 \in \operatorname{ri} K_i$, $i = 1, 2$. Так как K_i — локальный шатер к M_i в точке нуль, то существуют такие конусы Q_i, что

$$\bar{x}_0 \in \operatorname{ri} Q_i, \qquad \operatorname{Lin} Q_i = \operatorname{Lin} K_i, \qquad Q_i \subseteq K_i, \qquad i = 1, 2.$$

Кроме того, существуют функции

$$\psi_i(\bar{x}) = \bar{x} + r_i(\bar{x}), \qquad \|\bar{x}\|^{-1} r_i(\bar{x}) \to 0$$

при $\bar{x} \to 0$ такие, что

$$\psi_i(\bar{x}) \in M_i, \qquad \bar{x} \in Q_i \cap (\varepsilon B). \tag{1.26}$$

Пусть теперь векторы $e_1, \ldots, e_k$ образуют базис в $\operatorname{Lin} Q_1 = \operatorname{Lin} K_1$, а векторы $e_{k+1}, \ldots, e_n$ принадлежат $\operatorname{Lin} Q_2 = \operatorname{Lin} K_2$ и дополняют базис $e_1, \ldots, e_k$ до базиса во всем $\mathbf{R}^n$. Это возможно, так как справедлива формула (1.25).

Составим систему уравнений, записав ее в векторной форме:

$$g(\bar{x}, y) = \psi_1(y^1 e_1 + \ldots + y^k e_k + \bar{x}) - \\ - \psi_2(\bar{x} - y^{k+1} e_{k+1} - \ldots - y^n e_n) = 0. \tag{1.27}$$

Неизвестным здесь является y с компонентами y^i, $i = 1, \ldots, n$, а вектор $\bar{x}$ — параметром. Если обозначить через $\mathscr{E}$ матрицу со столбцами e_i, $i = 1, \ldots, n$, то имеет место формула

$$g(\bar{x}, y) = \mathscr{E}y + r_1\left(\bar{x} + \sum_{i=1}^{k} y^i e_i\right) - r_2\left(\bar{x} - \sum_{i=k+1}^{n} y^i e_i\right), \tag{1.28}$$

так как $\psi_i(\bar{x}) = \bar{x} + r_i(\bar{x})$. В силу свойств функции r_1

$$\frac{r_1\left(\bar{x} + \sum_{i=1}^{k} y^i e_i\right)}{\sqrt{\|\bar{x}\|^2 + \|y\|^2}} \to 0, \tag{1.29}$$

когда x и y стремятся к нулю, и аналогично для r_2. Поэтому

$$\|g(\bar{x}, y) - \mathscr{E}y\| \leqslant \bar{r}\left(\sqrt{\|\bar{x}\|^2 + \|y\|^2}\right).$$

Так как $e_1, \ldots, e_n$ — базис в $\mathbf{R}^n$, то матрица $\mathscr{E}$ невырождена и можно применить теорему 1.1, из которой следует, что существует такая $y(\bar{x})$, удовлетворяющая системе (1.27) и такая, что

$$\frac{y(\bar{x})}{\|x\|} \to 0. \tag{1.30}$$

Положим теперь

$$\psi(\bar{x}) = \psi_1\left(\bar{x} + \sum_{i=1}^{k} y^i(\bar{x}) e_i\right) = \psi_2\left(\bar{x} - \sum_{i=k+1}^{n} y^i(\bar{x}) e_i\right). \tag{1.31}$$

Тогда

$$\psi(\bar{x}) = \psi_1\left(\bar{x} + \sum_{i=1}^{k} y^i(\bar{x}) e_i\right) = \bar{x} + \sum_{i=1}^{k} y^i(\bar{x}) e_i + + r_1\left(\bar{x} + \sum_{i=1}^{k} y^i(\bar{x}) e_i\right) = \bar{x} + r(\bar{x}), \tag{1.32}$$

где

$$r(\bar{x}) = \sum_{i=1}^{k} y^i(\bar{x}) e_i + r_1\left(\bar{x} + \sum_{i=1}^{k} y^i(\bar{x}) e_i\right).$$

В силу (1.30) и того, что $\|\bar{x}\|^{-1}r_1(\bar{x}) \to 0$ при $x \to 0$, $r(\bar{x})$ обладает тем же свойством.

Предположим, что $\bar{x}_0 \neq 0$. Так как $\bar{x}_0 \in \operatorname{ri} Q_i$, то существует такое $\varepsilon_0 > 0$, что $\bar{x}_0 + \operatorname{Lin} Q_i \cap (\varepsilon_0 B) \subseteq Q_i$. Введем в рассмотрение конус

$$K_\delta = \left\{\bar{x}\colon \langle \bar{x}, \bar{x}_0\rangle > \|\bar{x}\|\|\bar{x}_0\|\left(1 - \frac{\delta^2}{2\|\bar{x}_0\|^2}\right)\right\}. \quad (1.33)$$

Нетрудно видеть, что если $\bar{x} \in K_\delta$, то

$$\left\|\|\bar{x}_0\|\frac{\bar{x}}{\|\bar{x}\|} - \bar{x}_0\right\| < \delta. \quad (1.34)$$

Поэтому, если $\bar{x} \in \operatorname{Lin} Q_i \cap K_{\varepsilon_0}$, то в силу неравенства (1.34) получаем, что

$$\frac{\|\bar{x}_0\|}{\|\bar{x}\|}\bar{x} = \bar{x}_0 + \left(\frac{\|\bar{x}_0\|}{\|\bar{x}\|}\bar{x} - \bar{x}_0\right) \in \bar{x}_0 + \operatorname{Lin} Q_i \cap (\varepsilon_0 B),$$

и поэтому $\bar{x} \in Q_i$.

Пусть теперь

$$Q = Q_1 \cap Q_2 \cap K_{\frac{1}{2}\varepsilon_0}.$$

Введем следующие обозначения:

$$z_1(\bar{x}) = \bar{x} + \sum_{i=1}^{k} y^i(\bar{x})\, e_i, \quad z_2(\bar{x}) = \bar{x} - \sum_{i=k+1}^{n} y^i(\bar{x})\, e_i.$$

В силу равенства (1.31)

$$\psi(\bar{x}) = \psi_1(z_1(x)) = \psi_2(z_2(\bar{x})). \quad (1.35)$$

Выберем $\varepsilon_1 > 0$ настолько малым, чтобы из условий $\|\bar{x}\| < \varepsilon_1$, $\bar{x} \in Q$, следовало, что

$$\|z_i(\bar{x})\| < \varepsilon, \quad z_i(\bar{x}) \in K_{\varepsilon_0}. \quad (1.36)$$

Из соотношений (1.30) и (1.33) следует возможность такого выбора.

Заметим теперь, что так как векторы e_i, $i = 1, \ldots, k$, принадлежат $\operatorname{Lin} Q_1$, то $z_1(\bar{x}) \in \operatorname{Lin} Q_1$ для $\bar{x} \in Q$.

Аналогично, если $\bar{x} \in Q$, то $z_2(\bar{x}) \in \operatorname{Lin} Q_2$. Поэтому для

$$\bar{x} \in Q \cap (\varepsilon_1 B)$$

из формул (1.36) и (1.26) следует, что

$$z_i(\bar{x}) \in \operatorname{Lin} Q_i \cap K_{\varepsilon_0}, \qquad z_i(\bar{x}) \in Q_i \cap (\varepsilon B),$$
$$\psi_i(z_i(\bar{x})) \in M_i, \qquad i = 1, \; 2.$$

Но тогда равенство (1.35) показывает, что $\psi(\bar{x}) \in M_1 \cap M_2$ для $\bar{x} \in Q \cap (\varepsilon_1 B)$, т. е. для $\bar{x}_0 \in \operatorname{ri} K$, $\bar{x}_0 \neq 0$, построены множество Q и функция ψ, удовлетворяющие условиям теоремы.

Если $\bar{x}_0 = 0$, то $\operatorname{Lin} Q_i \cap (\varepsilon_0 B) \subseteq Q_i$ при достаточно малом $\varepsilon_0 > 0$, а отсюда сразу следует, что $Q_i = \operatorname{Lin} Q_i$, т. е. Q_i — подпространство. Причем, так как $\operatorname{Lin} Q_i = \operatorname{Lin} K_i$, то $Q_i = \operatorname{Lin} K_i$.

Положим

$$Q = Q_1 \cap Q_2 = \operatorname{Lin} Q_1 \cap \operatorname{Lin} Q_2$$

и выберем $\varepsilon_1 > 0$ настолько малым, чтобы из $\|\bar{x}\| < \varepsilon_1$ следовало неравенство $\|z_i(\bar{x})\| < \varepsilon$. Тогда для $\bar{x} \in Q \cap (\varepsilon_1 B)$ выполняется включение

$$z_i(\bar{x}) \in Q_i \cap (\varepsilon B), \qquad i = 1, \; 2,$$

и формулы (1.26) показывают, что

$$\psi_i(z_i(\bar{x})) \in M_i, \qquad i = 1, 2.$$

Теперь из равенства (1.35) получаем соотношение

$$\psi(\bar{x}) = \psi_1(z_1(\bar{x})) = \psi_2(z_2(\bar{x})) \in M_1 \cap M_2.$$

Теорема доказана.

Теорема 1.3. *Пусть выполнены условия теоремы 1.2 и хотя бы один из конусов $K_{M_i}(x_0)$ не есть подпространство. Тогда существует точка $x_1 \in M$, отличная от x_0.*

Доказательство. Пусть $x_0 = 0$ и $K_i \equiv K_{M_i}(0)$. Так как конусы K_i неотделимы, то

$$\operatorname{ri} K = \bigcap_{i \in I} \operatorname{ri} K_i \neq \varnothing, \qquad K = \bigcap_{i \in I} K_i.$$

Допустим, что конус K_j, $j \in I$, не есть подпространство. Тогда $0 \notin \operatorname{ri} K_j$, ибо, как мы видели в конце доказательства предыдущей теоремы, из условия, что $0 \in \operatorname{ri} K_j$, вытекает равенство $K_j = \operatorname{Lin} K_j$. Так как $\operatorname{ri} K \neq \varnothing$, то можно указать направление $\bar{x}_0 \in \operatorname{ri} K$ такое, что $\bar{x}_0 \neq 0$.

Согласно предыдущей теореме существует такой конус Q, что

$$\bar{x}_0 \in \operatorname{ri} Q, \quad Q \subseteq K, \quad \operatorname{Lin} Q = \operatorname{Lin} K,$$

и такая функция $\psi(\bar{x}) = \bar{x} + r(\bar{x})$, что $\|\bar{x}\|^{-1} r(\bar{x}) \to 0$ при $\bar{x} \to 0$ и

$$\psi(\bar{x}) \in M, \quad \bar{x} \in Q \cap (\varepsilon B), \quad \varepsilon > 0.$$

Так как $\|\bar{x}\|^{-1} r(\bar{x}) \to 0$ при $\bar{x} \to 0$, то $\psi(\bar{x}) \neq 0$ при достаточно малых $\bar{x}$. Поэтому при достаточно малых $\bar{x}$, принадлежащих Q, $\psi(\bar{x}) \in M$, $\psi(\bar{x}) \neq 0$, что и требовалось доказать.

Теорема 1.4. *Пусть выполнены все условия теоремы* 1.2, *и пусть, кроме того,* $K_{M_i}(x_0)$ *— гладкие локальные шатры, т. е. соответствующие им функции* $\psi_i(\bar{x})$, *фигурирующие в определении* 1.3 *локального шатра, непрерывно дифференцируемы в окрестности начала координат. Тогда конус*

$$K = \bigcap_{i \in I} K_{M_i}(x_0)$$

есть гладкий локальный шатер к M *в точке* x_0.

Доказательство практически полностью совпадает с доказательством теоремы 1.2. Однако теперь в силу замечания к теореме 1.1 можно утверждать, что построенная в доказательстве теоремы 1.2 функция $\psi(\bar{x})$ — гладкая, и поэтому K — локальный шатер.

§ 2. Функции, допускающие верхнюю выпуклую аппроксимацию

Для того чтобы можно было выписать необходимые условия экстремума, необходимо, чтобы функции, фигурирующие в задаче, обладали некоторыми специальными свойствами. В частности, в окрестности точки минимума они должны допускать аппроксимацию некоторыми более простыми и сравнительно легко вычислимыми функциями. Так, например, гладкие функции допускают линейную аппроксимацию. Как было показано в § 2 главы I, выпуклые функции достаточно хорошо аппроксимируются выпуклыми положительно однородными функциями — производными по направлению. Однако негладкие и невыпуклые функции уже невозможно приблизить в окрестности некоторой точки выпуклыми положительно однородными функциями. Для таких функций вводится понятие верхней выпуклой аппроксимации. Следует сразу заметить, что верхняя выпуклая аппроксимация определяется неоднозначно, и для получения содержательных условий экстремума, как правило, необходимо знать достаточно широкое семейство верхних выпуклых аппроксимаций.

1. Определение верхней выпуклой аппроксимации и субдифференциала. Пусть $f(x)$, $x \in X$,— произвольная функция, принимающая конечные значения и значения $\pm\infty$. Положим

$$\operatorname{dom} f = \{x\colon\ |f(x)| < +\infty\}.$$

Определение 2.1. Положим для $x \in \operatorname{dom} f$, $\bar{x} \in X$, $\bar{x} \neq 0$,

$$F(\bar{x}, x) = \sup_{r(\cdot)} \limsup_{\lambda \downarrow 0} \frac{f(x + \lambda\bar{x} + r(\lambda)) - f(x)}{\lambda},$$

где внешняя верхняя грань берется по всем функциям $r(\lambda) \in X$ таким, что $\lambda^{-1} r(\lambda) \to 0$ при $\lambda \downarrow 0$. Очевидно, что величина $F(\bar{x}, x)$ — конечная или бесконечная — существует всегда. Легко также проверить, что функция $F(\bar{x}, x)$ положительно однородна по $\bar{x}$, т. е. для $\lambda > 0$

$$F(\lambda\bar{x}, x) = \lambda F(\bar{x}, x).$$

Замечание. Если функция $f(x)$ в окрестности точки x удовлетворяет условию Липшица, то

$$|f(x + \lambda\bar{x} + r(\lambda)) - f(x + \lambda\bar{x})| \leqslant L\|r(\lambda)\|.$$

Поэтому

$$\frac{f(x + \lambda\bar{x} + r(\lambda)) - f(x + \lambda\bar{x})}{\lambda} \to 0,$$

$$F(\bar{x}, x) = \sup_{r(\cdot)} \limsup_{\lambda \downarrow 0} \left[\frac{f(x+\lambda\bar{x}) - f(x)}{\lambda} + \right.$$
$$\left. + \frac{f(x+\lambda\bar{x}+r(\lambda)) - f(x+\lambda\bar{x})}{\lambda} \right] = \limsup_{\lambda \downarrow 0} \frac{f(x+\lambda\bar{x}) - f(x)}{\lambda},$$

т. е. для функций, удовлетворяющих условию Липшица, справедлива формула

$$F(\bar{x}, x) = \limsup_{\lambda \downarrow 0} \frac{f(x+\lambda\bar{x}) - f(x)}{\lambda}.$$

Определение 2.2. Функция $h(\bar{x}, x)$ называется *верхней выпуклой аппроксимацией* функции $f(x)$ в точке x, если:

1) $h(\bar{x}, x) \geqslant F(\bar{x}, x)$ для всех $\bar{x} \neq 0$;

2) $h(\bar{x}, x)$ — выпуклая замкнутая положительно однородная функция $\bar{x}$.

Для краткости в дальнейшем вместо слов «верхняя выпуклая аппроксимация» будем писать просто в. в. а. Ясно, что в. в. а. для функции $f(x)$ в точке x определена неоднозначно и может существовать много различных в. в. а.

Очевидна следующая лемма.

Лемма 2.1. *Если* $h_1(\bar{x}, x)$ *и* $h_2(\bar{x}, x)$ *есть в. в. а. для* f *в точке* x, *то* $\lambda_1 h_1 + \lambda_2 h_2$, $\lambda_1 + \lambda_2 = 1$, λ_1, $\lambda_2 \geqslant 0$, *и* $\max(h_1, h_2)$ *также являются в. в. а.*

Определение 2.3. Если $h(\bar{x}, x)$ есть в. в. а. для f в точке x, то множество

$$\partial h(0, x) = \{x^* \in X^*\colon\ h(\bar{x}, x) \geqslant \langle \bar{x}, x^* \rangle, \quad \bar{x} \in X\} \tag{2.1}$$

называется *субдифференциалом функции* f в точке x и обозначается $\partial f(x)$.

Согласно определению 2.2 и теореме II.3.12 субдифференциал $\partial h(0, x)$ выпуклой замкнутой положительно однородной функции $h(\bar{x}, x)$ относительно $\bar{x}$ существует всегда. При этом

$$\partial h(0, x) = \operatorname{dom} h^*(\cdot, x), \tag{2.2}$$

где

$$h^*(\bar{x}^*, x) = \sup_{\bar{x}} \{\langle \bar{x}, \bar{x}^* \rangle - h(\bar{x}, x)\}.$$

С другой стороны, из соотношения (2.2) и теоремы II.2.5 следует, что

$$h(\bar{x}, x) = \sup_{x^*} \{\langle \bar{x}, x^* \rangle \colon x^* \in \partial h(0, x)\},$$

или, с учетом определения 2.3,

$$h(\bar{x}, x) = \sup_{x^*} \{\langle \bar{x}, x^* \rangle \colon x^* \in \partial f(x)\}. \qquad (2.3)$$

При этом ясно, что субдифференциал $\partial f(x)$ в формуле (2.3) вычислен по функции $h(\bar{x}, x)$. Из формулы (2.3) также видно, что $h(\bar{x}, x)$ есть опорная функция множества $\partial f(x)$.

Формулы (2.1) и (2.3) показывают, что $h(\bar{x}, x)$ и $\partial f(x)$ однозначно определяют друг друга. Если $\partial f(x)$ — субдифференциал, то $\partial f(x)$ — выпуклое замкнутое множество, а функция $h(\bar{x}, x)$, определенная при помощи формулы (2.3), должна быть верхней выпуклой аппроксимацией для f в точке x. Выпуклость и замкнутость субдифференциала непосредственно следуют из определения 2.3 и теоремы II.3.2.

Из сказанного следует, что возможно другое определение субдифференциала, эквивалентное определению 2.3: выпуклое замкнутое множество $\partial f(x)$ называется *субдифференциалом функции* f в точке x, если функция $h(\bar{x}, x)$, определенная формулой (2.3), есть в. в. а. для f в точке x. Как и в. в. а., субдифференциал определен неоднозначно.

Использование одного и того же символа $\partial f(x)$, введенного для общего случая определением 2.3, а для выпуклой функции — определением II.3.2, не приводит, как правило, к недоразумениям, так как из контекста всегда достаточно ясно, о чем идет речь. Более того, если функция f выпукла и непрерывна в точке x, то следствие 2 формулируемой ниже теоремы 2.2 показывает, что субдифференциал такой функции в смысле определения II.3. 2 есть одновременно и субдифференциал в используемом здесь смысле.

Лемма 2.2. *Если h_1 и h_2 — в. в. а. для f в точке x и $h_1 \geqslant h_2$, то*

$$\partial_1 f(x) \supseteq \partial_2 f(x),$$

где $\partial_1 f(x)$ и $\partial_2 f(x)$ — субдифференциалы, определяемые h_1 и h_2 соответственно.

Действительно, согласно определению 2.3 и формуле (2.1) $x^* \in \partial_2 f(x)$ тогда и только тогда, когда

$$\langle \bar{x}, x^* \rangle \leqslant h_2(\bar{x}, x).$$

В силу условия леммы отсюда вытекает, что

$$\langle \bar{x}, x^* \rangle \leqslant h_2(\bar{x}, x) \leqslant h_1(\bar{x}, x),$$

т. е. $x^* \in \partial_1 f(x)$.

Проиллюстрируем данные простым примером. Пусть $x \in \mathbf{R}^1$,

$$f_c(0) = \begin{cases} 0, & x < 0, \\ c, & x \geqslant 0. \end{cases}$$

Нетрудно подсчитать, что для $c > 0$

$$F(\bar{x}, 0) = \begin{cases} -\infty, & \bar{x} < 0, \\ 0, & \bar{x} > 0. \end{cases}$$

Поэтому любая функция

$$h(\bar{x}, \ 0) = a\bar{x}, \qquad a \geqslant 0,$$

является в. в. а. для f_c в точке $x = 0$. Соответственно $\partial f_c(0) = \{a\}$, $a \geqslant 0$. Итак, если $c > 0$, то f_c в точке нуль имеет целое множество субдифференциалов, каждый из которых состоит из единственного числа $a \geqslant 0$.

Если $c < 0$, то

$$F(\bar{x}, 0) = \begin{cases} +\infty, & \bar{x} < 0, \\ 0, & \bar{x} > 0. \end{cases}$$

Положим

$$h_0(\bar{x}, 0) = \begin{cases} +\infty, & \bar{x} < 0, \\ 0, & \bar{x} \geqslant 0. \end{cases}$$

Функция $h_0(\bar{x}, 0)$ выпукла, замкнута и положительно однородна, так что h_0 есть в. в. а. для f_c, $c < 0$, в точке нуль. Простое вычисление показывает, что

$$\partial f_c(0) = \partial h_0(0, 0) = (-\infty, 0].$$

Итак, множество $(-\infty, 0]$ есть субдифференциал f_c в точке нуль. Так как

$$h_0(\bar{x}, 0) = F(\bar{x}, 0), \qquad \bar{x} \neq 0,$$

то для любой другой в. в. а. h будет выполнено неравенство $h \geqslant h_0$. Из леммы 2.2 следует, что

$$\partial f_c(0) \supseteq (-\infty, 0]$$

для любого субдифференциала $\partial f_c(0)$.

Приведенный пример поучителен тем, что он показывает возможность существования субдифференциалов даже у разрывных функций.

Из определений следует, что $\partial(cf) = c\partial f$ для $c \geqslant 0$.

Следующая теорема показывает наличие субдифференциала суммы двух функций.

Теорема 2.1. *Пусть $f = f_1 + f_2$, h_1 и h_2 — в. в. а. для f_1 и f_2 в точке x. Тогда*

$$h(\bar{x}, x) = h_1(\bar{x}, x) + h_2(\bar{x}, x)$$

есть в. в. а. для f в точке x. Если при этом

$$\operatorname{int}\operatorname{dom} h_1(\cdot, x) \cap \operatorname{dom} h_2(\cdot, x) \neq \varnothing,$$

то

$$\partial f(x) = \partial f_1(x) + \partial f_2(x).$$

Доказательство. По определению

$$F(\bar{x}, x) = \sup_{r(\cdot)} \limsup_{\lambda \downarrow 0} \left[\frac{f_1(x + \lambda\bar{x} + r(\lambda))}{\lambda} + \right.$$
$$\left. + \frac{f_2(x + \lambda x + r(\lambda)) - f_1(x) - f_2(x)}{\lambda} \right] \leqslant$$
$$\leqslant \sup_{r_1(\cdot)} \limsup_{\lambda \downarrow 0} \frac{f_1(x + \lambda\bar{x} + r_1(\lambda)) - f_1(x)}{\lambda} +$$
$$+ \sup_{r_2(\cdot)} \limsup_{\lambda \downarrow 0} \frac{f_2(x + \lambda\bar{x} + r_2(\lambda)) - f_2(x)}{\lambda} =$$
$$= F_1(\bar{x}, x) + F_2(\bar{x}, x).$$

Так как $h_1 \geqslant F_1$, $h_2 \geqslant F_2$, то $h \geqslant F$. Кроме того, функция h выпукла, положительно однородна и замкнута как сумма выпуклых, положительно однородных и замкнутых функций. Из сказанного следует, что h есть в. в. а. для f в точке x и первая часть теоремы доказана.

На основании теорем II.3.8 и II.1.4 имеем

$$\partial h(0, x) = \partial h_1(0, x) + \partial h_2(0, x),$$

т. е.

$$\partial f(x) = \partial f_1(x) + \partial f_2(x),$$

что и требовалось доказать.

2. Классы функций, допускающих верхнюю выпуклую аппроксимацию. Рассмотрим вопрос о том, как для некоторых классов функций вычислить верхнюю аппроксимацию и субдифференциал.

Теорема 2.2. *Пусть для данной функции f в точке x функция $F(\bar{x}, x)$ выпукла по $\bar{x}$ и замкнута, если положить $F(0, x) = 0$. Тогда $\partial F(0, x)$ есть субдифференциал функции f в точке x и для любого другого субдифференциала $\partial f(x)$ выполняется включение*

$$\partial f(x) \supseteq \partial F(0, x).$$

Доказательство следует из того, что в предположениях теоремы $F(\bar{x}, x)$ есть в.в.а. и для любой другой в.в.а. h выполняется неравенство $h \geqslant F$.

Следствие 1. *Если функция $f(x)$ непрерывно дифференцируема в точке x, то*

$$h(\bar{x}, x) = \langle \bar{x}, f'(x) \rangle$$

есть в.в.а., а $\partial f(x) = \{f'(x)\}$.

Действительно, для непрерывно дифференцируемой функции

$$F(\bar{x}, x) = \langle \bar{x}, f'(x) \rangle.$$

Следствие 2. *Если f — выпуклая функция, непрерывная в точке x, то*

$$h(\bar{x}, x) = f'(x, \bar{x}) = \lim_{\lambda \downarrow 0} \frac{f(x + \lambda\bar{x}) - f(x)}{\lambda}$$

есть в.в.а., а обычный субдифференциал выпуклой функции, задаваемой определением II.3.2, есть субдифференциал в смысле определения 2.3.

Доказательство. Так как по теореме II.1.3 непрерывная выпуклая функция удовлетворяет условию Липшица, то легко проверить, что

$$F(\bar{x}, x) = f'(x, \bar{x}).$$

Далее, по теоремам II.3.2 и II.3.5 $\partial f(x_0)$ есть ограниченное замкнутое выпуклое множество и

$$f'(x, \bar{x}) = \max_{x^*} \{\langle \bar{x}, x^* \rangle : x^* \in \partial f(x_0)\}.$$

Так как $f'(x, \bar{x})$ есть верхняя грань линейных (а значит, непрерывных и тем более замкнутых) функций, то $f'(x, \bar{x})$ есть замкнутая функция $\bar{x}$. Применение теоремы II.3.11 теперь показывает, что

$$\partial F(0, x) = \partial f'(x, 0) = \partial f(x).$$

Итак, $F(\bar{x}, x)$ есть выпуклая положительно однородная замкнутая функция $\bar{x}$, т. е. она удовлетворяет условиям теоремы 2.2, а ее субдифференциал совпадает с обычным субдифференциалом выпуклой функции f. Это завершает доказательство.

Теорема 2.3. *Пусть f — вогнутая функция, непрерывная в точке x. Тогда любая функция вида*

$$h(\bar{x}, x) = -\langle \bar{x}, x^* \rangle, \qquad x^* \in \partial(-f(x))$$

есть в.в.а. для f в точке x, а $\{-x^*\}$, $x^* \in \partial(-f(x))$, есть субдифференциал f в x. Здесь $\partial(-f(x))$ обозначает обычный субдифференциал выпуклой функции $-f$.

Доказательство. По определению вогнутости функция $f_0(x) = -f(x)$ выпукла. Так как по теореме II.1.3 функция $f_0(x)$ удовлетворяет условию Липшица, то ему удовлетворяет и $f(x)$. Поэтому

$$F(\bar{x}, x) = \lim_{\lambda \downarrow 0} \frac{f(x + \lambda \bar{x}) - f(x)}{\lambda} = f'(x, \bar{x}).$$

Но тогда

$$F(\bar{x}, x) = f'(x, \bar{x}) = -f_0'(x, \bar{x}) =$$
$$= -\max_{x^*} \{\langle \bar{x}, x^* \rangle : x^* \in \partial f_0(x)\} =$$
$$= \min_{x^*} \{-\langle \bar{x}, x^* \rangle : x^* \in \partial f_0(x)\},$$

и, значит,

$$F(\bar{x}, x) \leqslant -\langle \bar{x}, x^* \rangle, \qquad x^* \in \partial f_0(x).$$

Поэтому $h(\bar{x}, x) = -\langle \bar{x}, x^* \rangle$ есть в.в.а. Так как

$$\partial h(0, x) = \{-x^*\},$$

то $\{-x^*\}$ — субдифференциал, что и требовалось доказать.

Теорема 2.4. *Пусть I — произвольное множество индексов и при каждом $i \in I$ функция f_i имеет в точке x в.в.а., равную $h_i(\bar{x}, x)$. Пусть*

$$f(x) = \inf_i \{f_i(x)\colon\ i \in I\},$$

$$I(x) = \{i \in I\colon\ f_i(x) = f(x)\}.$$

Тогда $h_i(\bar{x}, x)$ есть в.в.а. функции f в точке x, если $i \in I(x)$, а $\partial f_i(x)$ — субдифференциал $f_i(x)$, соответствующий $h_i(\bar{x}, x)$, $i \in I(x)$, одновременно является субдифференциалом функции f в x.

Доказательство. Так как

$$\frac{f(x + \lambda\bar{x} + r(\lambda)) - f(x)}{\lambda} \leqslant \frac{f_i(x + \lambda\bar{x} + r(\lambda)) - f_i(x)}{\lambda},$$

$$i \in I(x),$$

то

$$F(\bar{x}, x) \leqslant F_i(\bar{x}, x), \qquad i \in I(x).$$

Поэтому $h_i(\bar{x}, x)$, $i \in I(x)$, есть в.в.а., а $\partial f_i(x)$ — субдифференциал функции f в точке x.

Пусть теперь A — компактное множество и для каждого $\alpha \in A$ определена функция $f(x, \alpha)$. Положим

$$f(x) = \sup_\alpha \{f(x, \alpha)\colon\ \alpha \in A\}. \tag{2.4}$$

Теорема 2.5. *Пусть функции $f(x, \alpha)$ непрерывны по совокупности аргументов, когда x меняется в некоторой окрестности точки x_0, а $\alpha \in A$. Пусть, кроме того, при каждом $\alpha \in A$ существуют производные по направлению $\bar{x}$ в точке x_0, т. е. определены $f'(x_0, \bar{x}, \alpha)$, причем* *разностное отношение*

$$\frac{f\left(x_0+\lambda x,\, \alpha\right)-f\left(x_0,\, \alpha\right)}{\lambda} \tag{2.5}$$

стремится к $f'(x_0, \bar{x}, \alpha)$ *равномерно по* $\alpha \in A$ *при* $\lambda \downarrow 0$. *Тогда функция* $f(x)$, *определенная формулой* (2.4), *дифференцируема по направлению* $\bar{x}$ *и*

$$f'\left(x_0, \bar{x}\right)=\max_{\alpha}\left\{f'\left(x_0, \bar{x}, \alpha\right):\ \alpha \in A\left(x_0\right)\right\},$$

где $A(x)=\{\alpha \in A:\ f(x, \alpha)=f(x)\}$.

Доказательство. Обозначим

$$\gamma(\lambda, \alpha)=\frac{f\left(x_0+\lambda \bar{x},\, \alpha\right)-f\left(x_0,\, \alpha\right)}{\lambda}-f'\left(x_0, \bar{x}, \alpha\right). \tag{2.6}$$

Так как при каждом $\lambda \geq 0$ отношение (2.5) непрерывно по α и сходится к $f'(x_0, \bar{x}, \alpha)$ равномерно на компакте A, то $f'(x_0, \bar{x}, \alpha)$ — непрерывная функция α.

Пусть теперь $x(\lambda)=x_0+\lambda \bar{x}$. Легко проверить, что имеет место неравенство

$$\frac{f(x(\lambda), \alpha)-f\left(x_0, \alpha\right)}{\lambda} \geqslant \frac{f(x(\lambda))-f\left(x_0\right)}{\lambda} \geqslant \\ \geqslant \frac{f\left(x(\lambda), \alpha_0\right)-f\left(x_0, \alpha_0\right)}{\lambda} \tag{2.7}$$

для любых $\alpha \in A(x(\lambda))$, $\alpha_0 \in A(x_0)$. Из правого неравенства, устремляя λ к нулю, получаем

$$\liminf_{\lambda \downarrow 0} \frac{f(x(\lambda))-f\left(x_0\right)}{\lambda} \geqslant f'\left(x_0, \bar{x}, \alpha_0\right),$$

или, так как α_0 — произвольный элемент $A(x_0)$,

$$\liminf_{\lambda \downarrow 0} \frac{f(x(\lambda))-f\left(x_0\right)}{\lambda} \geqslant \max_{\alpha}\left\{f'\left(x_0, \bar{x}, \alpha\right):\ \alpha \in A\left(x_0\right)\right\}. \tag{2.8}$$

Выберем теперь в качестве C произвольное открытое подмножество A, содержащее $A(x_0)$, $C \supseteq A(x_0)$, и покажем, что

$$\inf_{C \supseteq A(x_0)} \sup_{\alpha \in C} f'\left(x_0, \bar{x}, \alpha\right)=\max_{\alpha \in A(x_0)} f'\left(x_0, \bar{x}, \alpha\right). \tag{2.9}$$

Так как $f'(x_0, \bar{x}, \alpha)$ — непрерывная по α функция, то для любого $C \supseteq A(x_0)$ выполняется неравенство

$$\sup_{\alpha \in C} f'(x_0, \bar{x}, \alpha) \geqslant \max_{\alpha \in A(x_0)} f'(x_0, \bar{x}, \alpha)$$

и поэтому левая часть соотношения (2.9) всегда не меньше правой.

С другой стороны, $A(x_0)$ есть замкнутое подмножество компактного множества A и поэтому $A(x_0)$ само компактно. Выберем $\varepsilon > 0$ и поставим в соответствие каждому $\alpha \in A(x_0)$ окрестность U_α так, чтобы

$$|f'(x_0, \bar{x}, \bar{\alpha}) - f'(x_0, \bar{x}, \alpha)| < \varepsilon, \quad \bar{\alpha} \in U_\alpha. \tag{2.10}$$

Так как $A(x_0)$ — компактное множество, а объединение открытых множеств U_α, очевидно, покрывает $A(x_0)$, то существует конечный набор множеств U_α, например, $U_{\alpha_1}, \ldots, U_{\alpha_m}$, покрывающих $A(x_0)$. Обозначим

$$C_\varepsilon = \bigcup_{i=1}^{m} U_{\alpha_i}.$$

Для любой точки $\alpha \in C_\varepsilon$ найдется такой номер i, что $\alpha \in U_{\alpha_i}$. В силу неравенства (2.10) можно записать

$$f'(x_0, \bar{x}, \alpha) \leqslant f'(x_0, \bar{x}, \alpha_i) + \varepsilon \leqslant \\ \leqslant \max_{i=1,\ldots,m} f'(x_0, \bar{x}, \alpha_i) + \varepsilon \leqslant \max_{\alpha \in A(x_0)} f'(x_0, \bar{x}, \alpha) + \varepsilon.$$

Поэтому

$$\inf_{C \supseteq A(x_0)} \sup_{\alpha \in C} f'(x_0, \bar{x}, \alpha) \leqslant \sup_{\alpha \in C_\varepsilon} f'(x_0, \bar{x}, \alpha) \leqslant \\ \leqslant \max_{\alpha \in A(x_0)} f'(x_0, \bar{x}, \alpha) + \varepsilon.$$

Так как $\varepsilon > 0$ произвольно, то получается, что левая часть соотношения (2.9) не больше правой. Сопоставляя это с полученным ранее противоположным неравенством, получаем равенство (2.9).

Для завершения доказательства теоремы воспользуемся теперь леммой II.3.6, в силу которой при достаточно малом λ справедливо включение $A(x(\lambda)) \subseteq C$. Поэтому из левого неравенства (2.7) следует, что

$$\frac{f(x(\lambda))-f(x_0)}{\lambda}\leqslant\sup_{\alpha\in C}\frac{f(x(\lambda),\alpha)-f(x_0,\alpha)}{\lambda}=$$

$$=\sup_{\alpha\in C}\left[f'\left(x_0,\bar{x},\alpha\right)+\gamma(\lambda,\alpha)\right].$$

По предположению теоремы $\gamma(\lambda,\alpha)\leqslant r(\lambda)$, $r(\lambda)\to 0$, так как отношение (2.5) стремится к $f'(x_0,\bar{x},\alpha)$ равномерно по $\alpha\in A$. Поэтому

$$\frac{f(x(\lambda))-f(x_0)}{\lambda}\leqslant\sup_{\alpha\in C}f'\left(x_0,\bar{x},\alpha\right)+r(\lambda).$$

Переходя к пределу сначала по $\lambda\downarrow 0$, а потом беря нижнюю грань по всем $C\supseteq A(x_0)$, получим, используя равенство (2.9),

$$\limsup_{\lambda\downarrow 0}\frac{f(x(\lambda))-f(x_0)}{\lambda}\leqslant\max_{\alpha\in A(x_0)}f'\left(x_0,\bar{x},\alpha\right).\qquad(2.11)$$

Сопоставление формул (2.8) и (2.11) завершает доказательство теоремы.

Теорема 2.6. *Пусть A — компакт, функции $f(x,\alpha)$, $\alpha\in A$, дифференцируемы по x в окрестности точки x_0 и градиент $f'_x(x,\alpha)$ непрерывен по совокупности аргументов x и α. Тогда для функции $f(x)$, определяемой формулой (2.4), выполняется соотношение*

$$f'\left(x_0,\bar{x}\right)=\max_{\alpha}\left\{\langle\bar{x},f'_x(x_0,\alpha)\rangle:\ \alpha\in A(x_0)\right\}.\qquad(2.12)$$

Более того, $f'(x_0,\bar{x})$ есть в.в.а. для $f(x)$ в точке x_0, и соответствующий субдифференциал может быть задан формулой

$$\partial f(x_0)=\operatorname{co}\Big(\bigcup_{\alpha\in A(x_0)}f'_x(x_0,\alpha)\Big)=\operatorname{co}f'_x(x_0,A(x_0)).\qquad(2.13)$$

Доказательство. Так как $f'_x(x,\alpha)$ непрерывно зависит от x в окрестности x_0 и $\alpha\in A$, причем A — компакт, то

$$\|f'_x(x,\alpha)-f'_x(x_0,\alpha)\|\leqslant\varepsilon(\|x-x_0\|),$$

где функция $\varepsilon(\lambda)$, монотонно убывая, стремится к нулю, когда $\lambda\downarrow 0$.

Используя известную в анализе теорему о среднем, получаем, что

$$\frac{f(x_0+\lambda\bar{x},\alpha)-f(x_0,\alpha)}{\lambda}=\langle\bar{x},f'(x_0+\theta_\alpha\lambda\bar{x},\alpha)\rangle,\quad 0\leqslant\theta_\alpha\leqslant 1.$$

Так как в рассматриваемом случае

$$f'(x_0,\bar{x},\alpha)=\langle\bar{x},f'(x_0,\alpha)\rangle, \tag{2.14}$$

то из соотношения (2.6) получаем, что

$$|\gamma(\lambda,\alpha)|=|\langle\bar{x},f'(x_0+\theta_\alpha\lambda\bar{x},\alpha)-f'(x_0,\alpha)\rangle|\leqslant$$
$$\leqslant\|\bar{x}\|\|f'(x_0+\theta_\alpha\lambda\bar{x},\alpha)-f'(x_0,\alpha)\|\leqslant$$
$$\leqslant\|\bar{x}\|\varepsilon(\theta_\alpha\lambda\|\bar{x}\|)\leqslant\|\bar{x}\|\varepsilon(\lambda\|\bar{x}\|).$$

Отсюда следует, что величина $\gamma(\lambda,\alpha)$ равномерно стремится к нулю при $\lambda\downarrow 0$ и все условия теоремы 2.5 выполнены. Используя результат этой теоремы и формулу (2.14), приходим к соотношению (2.12).

Заметим теперь, что так как $f'_x(x,\alpha)$ непрерывно зависит от своих аргументов и A — компакт, то

$$\|f'_x(x,\alpha)\|\leqslant L$$

для всех x из малой окрестности x_0 и $\alpha\in A$. Воспользовавшись снова теоремой о среднем значении, получаем, что

$$|f(x,\alpha)-f(y,\alpha)|=$$
$$=|\langle x-y,f'(y+\theta(x-y),\alpha)\rangle|\leqslant L\|x-y\|,$$

т. е. функция $f(x,\alpha)$ удовлетворяет в окрестности x_0 условию Липшица. Но

$$f(x,\alpha)-f(y,\alpha)\leqslant f(x)-f(y)\leqslant f(x,\alpha_0)-f(y,\alpha_0),$$
$$\alpha\in A(y),\quad \alpha_0\in A(x).$$

Так как

$$f(x,\alpha)-f(y,\alpha)\geqslant -L\|x-y\|,$$
$$f(x,\alpha_0)-f(y,\alpha_0)\leqslant L\|x-y\|,$$

то

$$|f(x)-f(y)|\leqslant L\|x-y\|,$$

т. е. $f(x)$ также удовлетворяет условию Липшица.

В силу замечания, сделанного после определения 2.1, и того факта, что $f'(x_0, \bar{x})$ существует и определяется формулой (2.12), получаем

$$F(\bar{x}, x_0) = f'(x_0, \bar{x}) = \max_{\alpha} \{\langle \bar{x}, f'_x(x_0, \alpha)\rangle :\ \alpha \in A(x_0)\}.$$

Последнюю формулу можно записать и иначе:

$$F(\bar{x}, x_0) = \max_{x^*} \{\langle \bar{x}, x^*\rangle :\ x^* \in f'_x(x_0, A(x_0))\}.$$

Сделаем несколько замечаний:

1) множество $f'_x(x_0, A(x_0))$ компактно как непрерывный образ компактного множества $A(x_0)$;

2) в силу теоремы I.1.7 множество $\operatorname{co} f'_x(x_0, A(x_0))$ выпукло и компактно;

3) максимум линейной функции на множестве и на его выпуклой оболочке один и тот же.

Поэтому

$$F(\bar{x}, x_0) = \max_{x^*} \{\langle \bar{x}, x^*\rangle :\ x^* \in \operatorname{co} f'_x(x_0, A(x_0))\}.$$

Итак, $F(\bar{x}, x_0)$ есть выпуклая, положительно однородная и замкнутая функция $\bar{x}$. На самом деле она даже непрерывна, так как из компактности $\operatorname{co} f'_x(x_0, A(x_0))$ следует, что $F(\bar{x}, x_0)$ определена для всех $\bar{x}$ и, значит, по теореме II.1.4 непрерывна. Ясно, что функцию $F(\bar{x}, x_0)$ можно рассматривать как верхнюю выпуклую аппроксимацию функции f в точке x_0.

По теореме II.3.11 $x^* \in \partial F(0, x_0)$ тогда и только тогда, когда

$$x^* \in \operatorname{co} f'_x(x_0, A(x_0))$$

и $\langle 0, x^*\rangle = F(0, x_0)$. Но последнее выполняется, очевидно, для всех x^*, ибо $F(0, x_0) = 0$. Поэтому

$$\partial f(x_0) = \partial F(0, x_0) = \operatorname{co} f'_x(x_0, A(x_0)),$$

что и требовалось доказать.

3. Функция расстояния до множества. Пусть M — произвольное множество. Рассмотрим следующую функцию:

$$d(x \mid M) = \inf_{y} \{\|x - y\| :\ y \in M\}, \tag{2.15}$$

где $\|x-y\| = \langle x-y,\ x-y\rangle^{1/2}$ — обычное евклидово расстояние между точками. Тогда $d(x|M)$ есть расстояние от точки x до множества M. Поскольку расстояния до множества и до его замыкания одинаковы, то без ограничения общности будем в дальнейшем считать, что M замкнуто. Это позволяет переписать формулу (2.15) в следующем виде:

$$d(x|M) = \min_y \{\|x-y\|:\ y \in M\}.$$

Множество

$$M(x) = \{y \in M:\ \|x-y\| = d(x|M)\}$$

замкнуто и ограничено, а поэтому компактно.

Лемма 2.3. *Функция $d(x|M)$ удовлетворяет условию Липшица с константой $L=1$.*

Доказательство. Если $y_1 \in M(x_1)$, $y_2 \in M(x_2)$, то

$$\|x_1-y_1\| - \|x_2-y_1\| \leqslant d(x_1|M) - d(x_2|M) \leqslant \\ \leqslant \|x_1-y_2\| - \|x_2-y_2\|.$$

Из неравенства треугольника вытекает, что

$$\|x_1-y_2\| - \|x_2-y_2\| \leqslant \|x_1-x_2\|,$$
$$\|x_2-y_1\| - \|x_1-y_1\| \leqslant \|x_1-x_2\|,$$

поэтому

$$|d(x_1|M) - d(x_2|M)| \leqslant \|x_1-x_2\|,$$

что и требовалось доказать.

Теорема 2.7. *Если $d(x|M) > 0$, то*

$$d'(x, \bar{x}|\ M) = \lim_{\lambda \downarrow 0} \frac{d(x+\lambda\bar{x}|M) - d(x|M)}{\lambda} = \\ = \min_y \left\{\frac{\langle \bar{x},\ x-y\rangle}{d(x|M)}:\ y \in M(x)\right\}.$$

Доказательство. Если $x \neq y$, то $\|x-y\|$ есть дифференцируемая функция x, градиент которой равен

$$\|x-y\|'_x = \frac{x-y}{\|x-y\|}. \tag{2.16}$$

Без ограничения общности можно считать, что M компактно, т. е. ограничено, так как, если отбросить точки

M, лежащие достаточно далеко, то расстояние до вновь полученного множества при малых изменениях x будет совпадать с $d(x|M)$. Записав $d(x|M)$ в виде

$$d(x\,|\,M) = -\max_{y}\{-\|x-y\|:\ y \in M\}$$

и воспользовавшись теоремой 2.6, формулой (2.12) и выражением (2.16) для градиента, немедленно получим утверждение теоремы.

Теорема 2.8. *Если $d(x|M) > 0$, то функция*

$$h(\bar{x}, x) = \frac{\langle \bar{x}, x-y\rangle}{d(x\,|\,M)}, \qquad y \in M(x), \tag{2.17}$$

есть в.в.а. для d в точке x, а $[d(x|M)]^{-1}(x-y)$ — соответствующий субдифференциал.

Доказательство. Так как функция $d(x|M)$ удовлетворяет условию Липшица, то

$$F(\bar{x}, x) = d'\ (x, \bar{x}\,|\,M) = \min_{y}\left\{\frac{\langle \bar{x}, x-y\rangle}{d(x\,|\,M)}:\ y \in M(x)\right\}. \tag{2.18}$$

Из этого выражения утверждение теоремы следует непосредственно.

Теорема 2.9. *Пусть $d(x|M)=0$, и пусть $K_M(x)$ — выпуклый конус касательных направлений для M в x. Тогда функция*

$$h(\bar{x}, x) = \inf_{\bar{y}}\{\|\bar{x}-\bar{y}\|:\ \bar{y} \in K_M(x)\} = d(\bar{x}\,|\,K_M(x))$$

есть верхняя выпуклая аппроксимация для d в точке x, а $B \cap (-K_M^*(x))$ — соответствующий субдифференциал.

Доказательство. Пусть $\bar{y} \in K_M(x)$. По определению существует такая функция $r(\lambda)$, $\lambda^{-1}r(\lambda) \to 0$ при $\lambda \to 0$, что $x + \lambda\bar{y} + r(\lambda) \in M$ при малых λ.

Так как $d(x|M)$ удовлетворяет условию Липшица, то

$$F(\bar{x}, x) = \limsup_{\lambda \downarrow 0} \frac{d(x+\lambda\bar{x}\,|\,M) - d(x\,|\,M)}{\lambda} \leqslant$$
$$\leqslant \limsup_{\lambda \downarrow 0} \frac{\|(x+\lambda\bar{x}) - (x+\lambda\bar{y}+r(\lambda))\|}{\lambda} \leqslant$$
$$\leqslant \|\bar{x}-\bar{y}\| + \lim_{\lambda \downarrow 0}\frac{\|r(\lambda)\|}{\lambda} = \|\bar{x}-\bar{y}\|.$$

Так как $\bar{y}$ — произвольный элемент $K_M(x)$, то

$$F(\bar{x}, x) \leqslant \inf_{\bar{y}} \{\|\bar{x} - \bar{y}\| : \bar{y} \in K_M(x)\} = d(\bar{x} \mid \overline{K_M(x)}). \tag{2.19}$$

Согласно п. 4 § 3 главы II функция $d(\bar{x} | \overline{K_M(x)})$ выпукла. Теорема 3.5 показывает, что эта функция есть точная верхняя грань линейных функций и, значит, замкнута. Наконец, положительная однородность $d(\bar{x} | \overline{K_M(x)})$ следует из того, что $K_M(x)$ — конус. В этих условиях формула (2.19) показывает, что функция $d(\bar{x} | \overline{K_M(x)})$ есть в. в. а. для d в точке x.

Вычислим субдифференциал функции $d(\bar{x} | \overline{K_M(x)})$ при $\bar{x} = 0$. Заметим, что эта функция совпадает с введенной для выпуклого случая в п. 4 § 3 главы II функцией $d_C(\cdot | \overline{K_M(x)})$, когда $C = B$, т. е. C есть единичный шар. Поэтому можно воспользоваться теоремой II.3.15, которая показывает, что

$$\partial d(0 | \overline{K_M(x)}) = \partial \delta(0 | \overline{K_M(x)}) \cap \{x^* : W_B(x^*) \leqslant 1\}.$$

Согласно формулам (II.3.25) и (II.3.26)

$$W_B(x^*) = \max_x \{\langle x, x^* \rangle : \|x\| \leqslant 1\} = \|x^*\|,$$

$$\partial \delta(0 | \overline{K_M(x)}) = -K_M^*(x),$$

поэтому

$$\partial d(0 | \overline{K_M(x)}) = (-K_M^*(x)) \cap \{x^* : \|x^*\| \leqslant 1\},$$

что и требовалось доказать.

4. Главные верхние аппроксимации и главные субдифференциалы. Как показывают приведенные выше примеры, для данной функции f в точке x может существовать много верхних выпуклых аппроксимаций. Однако, естественно, что если h_1 и h_2 — в. в. а. функции f в точке x и $h_1 \geqslant h_2$, то h_1 хуже приближает функцию f в окрестности точки x. Это мотивирует следующее

Определение 2.4. В. в. а. h функции f в точке x называется *главной*, если не существует другой в. в. а. h_1 такой, что

$$h(\bar{x}, x) \geqslant h_1(\bar{x}, x)$$

для всех x. Соответствующий в. в. а. h субдифференциал называется *главным*.

Теорема 2.10. *Если $F(\bar{x}, x)$ после доопределения $F(0, x)=0$ является выпуклой замкнутой функцией, то существует единственный главный субдифференциал. В частности, если f — выпуклая непрерывная функция, то ее обычный субдифференциал является единственным главным субдифференциалом.*

Теорема очевидным образом следует из определения главного субдифференциала.

Условимся, что в дальнейшем всегда, когда выполнены условия теоремы 2.10 под рассматриваемым в контексте субдифференциалом понимается главный субдифференциал.

Теорема 2.11. *Если $h(\bar{x}, x)=\langle\bar{x}, x^*\rangle$ есть в. в. а. функции f в точке x, то $h(\bar{x}, x)$ — главная в. в. а., а x^* — главный субдифференциал.*

Доказательство. Пусть существует в. в. а. h_1 такая, что

$$\langle\bar{x}, x^*\rangle \geqslant h_1(\bar{x}, x)$$

для всех $\bar{x}$. Возьмем $x_1^* \in \partial f(x)$, где $\partial f(x)$ соответствует $h_1(\bar{x}, x)$. Тогда в силу формулы (2.3) получаем, что

$$\langle\bar{x}, x^*\rangle \geqslant h_1(\bar{x}, x) \geqslant \langle\bar{x}, x_1^*\rangle.$$

Нетрудно убедиться, что одна линейная функция может быть всюду больше другой лишь в том случае, если они совпадают. Поэтому

$$\langle\bar{x}, x^*\rangle = h_1(\bar{x}, x) = \langle\bar{x}, x_1^*\rangle,$$

т. е. $h=h_1$.

В заключение приведем несколько абстрактную теорему о существовании главного дифференциала. Она не используется в дальнейшем изложении. При ее доказательстве предполагается, что читатель знаком с частично упорядоченными множествами в объеме § 2 главы I книги Н. Данфорда и Дж. Т. Шварца [1].

Теорема 2.12. *Пусть $F(\bar{x}, x)$ как функция $\bar{x}$ замкнута и не принимает значений, равных $-\infty$. Тогда для каждого x такого, что $F(\bar{x}, x)<+\infty$, существует главная* *верхняя выпуклая аппроксимация* h *такая, что* $h(\bar{x}, x) = F(\bar{x}, x)$, *и*

$$F(\bar{x}, x) = \inf_h h(\bar{x}, x),$$

где нижняя грань берется по всем главным выпуклым аппроксимациям.

Доказательство. Последнее утверждение теоремы очевидным образом следует из первого, поэтому кратко проведем доказательство только первого утверждения.

Пусть $F(\bar{x}_0, x)$ — конечное число. Построим функцию

$$h_{\bar{x}_0}(\bar{x}, x) = \begin{cases} \lambda F(\bar{x}_0, x), & \text{если } \bar{x} = \lambda\bar{x}_0, \quad \lambda \geqslant 0, \\ +\infty, & \text{если } \bar{x} \neq \lambda\bar{x}_0, \quad \lambda \geqslant 0. \end{cases}$$

Легко видеть, что $h_{\bar{x}_0}$ — в. в. а. и $h_{\bar{x}_0}$ совпадает с F, когда $\bar{x} = \bar{x}_0$.

Рассмотрим множество всех в. в. а. h таких, что $h \leqslant h_{\bar{x}_0}$. Очевидно, что это есть частично упорядоченное множество, где порядок задается неравенством $h_1 \leqslant h_2$, если соотношение

$$h_1(\bar{x}, \ x) \leqslant h_2(\bar{x}, \ x)$$

выполнено для всех $\bar{x}$. Так как $F \leqslant h \leqslant h_{x_0}$, то $h(\bar{x}_0, x) = F(\bar{x}_0, x)$. Заметим, что функция h выпукла, замкнута и положительно однородна тогда и только тогда, когда $\operatorname{epi} h$ есть выпуклый замкнутый конус. Пусть теперь h_α — линейно упорядоченное подмножество рассматриваемого множества, т. е. для любых α_1 и α_2 либо $h_{\alpha_1} \leqslant h_{\alpha_2}$, либо $h_{\alpha_2} \leqslant h_{\alpha_1}$.

Рассмотрим множество

$$H = \bigcup_\alpha \operatorname{epi} h_\alpha(\cdot, x).$$

Нетрудно убедиться, что в силу линейной упорядоченности семейства h_α множество H есть выпуклый конус. Так как $h \geqslant F$, то $H \subseteq \operatorname{epi} F(\cdot, x)$, а так как F — замкнутая функция, то $\operatorname{epi} F(\cdot, x)$ есть замкнутое множество и $\bar{H} \subseteq \operatorname{epi} F(\cdot, x)$. Зададим теперь функцию h при помощи ее надграфика:

$$\operatorname{epi} h = \overline{H}.$$

$\overline{H}$ — выпуклый замкнутый конус и поэтому h выпуклая замкнутая положительно однородная функция. Так как

$$\operatorname{epi} F(\cdot,\ x) \supseteq \operatorname{epi} h \supseteq \operatorname{epi} h_\alpha,$$

то $F \leqslant h \leqslant h_\alpha$. Таким образом, h есть в. в. а. и h является минорантой линейно упорядоченного множества. Согласно лемме Цорна в этом случае рассматриваемое множество верхних выпуклых аппроксимаций имеет минимальный элемент h_0, т. е. в данном случае это означает, что не существует другого h такого, что $h_0 \geqslant h$. Это как раз и показывает, что h_0 — главная в. в. а. Теорема доказана.

§ 3. Отображения, локально сопряженные к многозначным отображениям

В главе III было дано определение многозначных отображений и были изучены свойства выпуклых отображений. Одним из основных понятий, введенных там, было понятие локально сопряженного отображения. В этом параграфе будут рассмотрены многозначные отображения, график которых уже не обязательно является выпуклым. Читателю рекомендуется перед дальнейшим чтением еще раз просмотреть основные определения главы III.

Пусть X и Y — конечномерные пространства и $Z = X \times Y$ — их прямое произведение. Элементами пространства Z являются векторы z, представляющие собой пары $(x,\ y)$. Многозначное отображение a задано, если задано множество $\operatorname{gf} a \subseteq Z$, называемое *графиком* a; при этом

$$a(x) = \{y\colon\ (x,\ y) \in \operatorname{gf} a\}.$$

В дальнейшем будут изучаться, в основном, выпуклозначные отображения, т. е. такие, для которых множество $a(x)$ выпукло при любом $x \in \operatorname{dom} a$. Для таких отображений имеет смысл ввести следующие обозначения:

$$W_a(x,\ y^*) = \inf_y \{\langle y,\ y^*\rangle\colon\ y \in a(x)\}, \qquad (3.1)$$

$$a(x;\ y^*) = \{y \in a(x)\colon\ \langle y,\ y^*\rangle = W_a(x,\ y^*)\}. \qquad (3.2)$$

Лемма 3.1. *Пусть a — выпуклозначное замкнутое ограниченное непрерывное отображение. Тогда $W_a(x, y^*)$ есть непрерывная функция, а множество $a(x; y^*)$ полунепрерывно сверху зависит от своих аргументов.*

Доказательство. Пусть $x_0 \in \operatorname{dom} a$. Так как отображение $a(x)$ непрерывно в точке x_0, то для любого открытого шара $U \subseteq Y$ с центром в точке 0 и радиуса $\varepsilon > 0$ найдется такая окрестность V точки x_0, что

$$a(x) \subseteq a(x_0) + U, \quad a(x_0) \subseteq a(x) + U \tag{3.3}$$

для всех $x \in V$. Зафиксируем y_0^*. Тогда из включений (3.3) следует, что

$$W_a(x, y^*) \geqslant \inf_{y,u} \{\langle y + u, y^*\rangle :\ y \in a(x_0),\ \ \|u\| < \varepsilon\} =$$

$$= \inf_{y,u} \{\langle y, y_0^*\rangle + \langle u, y^*\rangle + \langle y, y^* - y_0^*\rangle :$$

$$y \in a(x_0),\ \|u\| < \varepsilon\} \geqslant W_a(x_0, y_0^*) + \inf_u \{\langle u, y^*\rangle :\ \|u\| < \varepsilon\} +$$

$$+ \inf_y \{\langle y, y^* - y_0^*\rangle :\ y \in a(x_0)\} \geqslant$$

$$\geqslant W_a(x_0, y_0^*) - \varepsilon \|y^*\| - \|a(x_0)\| \|y^* - y_0^*\|,$$

и, аналогично,

$$W_a(x_0, y_0^*) \geqslant W_a(x, y^*) - \varepsilon \|y_0^*\| - \|a(x)\| \|y^* - y_0^*\|.$$

Поэтому

$$|W_a(x, y^*) - W_a(x_0, y_0^*)| \leqslant$$

$$\leqslant \varepsilon \max(\|y_0^*\|, \|y^*\|) + (\|a(x_0)\| + \varepsilon) \|y^* - y_0^*\|. \tag{3.4}$$

При выводе формулы (3.4) было использовано, что в силу первого соотношения (3.3)

$$\|a(x)\| = \sup_y \{\|y\| :\ y \in a(x)\} \leqslant$$

$$\leqslant \sup_{y,u} \{\|y + u\| :\ y \in a(x_0),\ u \in U\} \leqslant \|a(x_0)\| + \varepsilon.$$

Из неравенства (3.4) непосредственно следует непрерывность функции $W_a(x, y^*)$.

Докажем второе утверждение. Для этого необходимо показать, что для любой окрестности U нуля в Y, найдется такое $\varepsilon > 0$, что

$$a(x; y^*) \subseteq a(x_0; y_0^*) + U, \tag{3.5}$$

как только $\|x - x_0\| < \varepsilon$, $\|y^* - y_0^*\| < \varepsilon$.

Доказательство проведем от противного, т. е. предположим, что включение (3.5) не имеет места. Заметим, что, так как множество $a(x)$ ограничено и замкнуто, то множество $a(x; y^*)$ также ограничено и замкнуто. Пусть теперь $x_i \to x_0$, $y_i^* \to y_0$, но при каждом i существует такая точка $y_i \in a(x_i; y_i^*)$, что

$$y_i \notin (x_0; y_0^*) + U. \tag{3.6}$$

Так как отображение a непрерывно, то $a(x)$ лежит в некоторой малой окрестности множества $a(x_0)$, а в силу ограниченности последнего все $a(x)$ ограничены при x, достаточно близких к x_0. Но $a(x; y^*) \subseteq a(x)$. Поэтому все $a(x_i; y_i^*)$ ограничены в совокупности и последовательность y_i ограничена. Без ограничения общности можно считать, что $y_i \to y_0$. По определению множества $a(x; y^*)$ имеем

$$\langle y_i, y_i^* \rangle = W_a(x_i, y_i^*).$$

Переходя к пределу по непрерывности получаем, что

$$\langle y_0, y_0^* \rangle = W_a(x_0, y_0^*). \tag{3.7}$$

Из включения $y_i \in a(x_i)$ в силу замкнутости отображения a вытекает, что $y_0 \in a(x_0)$. Совместно с равенством (3.7) это означает, что $y_0 \in a(x_0; y_0^*)$. Но последнее противоречит условию (3.6), так как если последовательность не принадлежит некоторой окрестности множества, то и предельная точка не может ей принадлежать. Полученное противоречие доказывает лемму.

Лемма 3.2. *Пусть выполнены условия предыдущей леммы и пусть кроме того, a удовлетворяет условию Липшица. Тогда $W_a(x, y^*)$ также удовлетворяет условию Липшица.*

Доказательство. Тот факт, что a удовлетворяет условию Липшица означает, что радиус ε шара U в формулах (3.3) может быть оценен сверху величиной

$L\|x - x_0\|$. Тогда из неравенства (3.4) следует, что

$$\left|W_a(x, y^*) - W_a(x_0, y_0^*)\right| \leqslant L\|x - x_0\| \max(\|y_0^*\|, \|y^*\|) + \\ + (\|a(x_0)\| + L\|x - x_0\|)\|y - y_0^*\|, \quad (3.8)$$

откуда вытекает утверждение леммы.

Пусть теперь $\operatorname{gf} a$ имеет в каждой своей точке z выпуклый конус касательных направлений $K_{\operatorname{gf} a}(z)$. Для сокращения будем его обозначать через $K_a(z)$. Напомним, что $K_a(z)$ есть конус касательных направлений к $\operatorname{gf} a$ в точке z, если он выпуклый и для каждого $\bar{z} \in K_a(z)$ существует такая функция r: $[0, 1] \to Z$, что $z + \lambda \bar{z} + r(\lambda) \in \in \operatorname{gf} a$ при малых $\lambda \geqslant 0$.

Определение 3.1. Отображение

$$a^*(y^*; z) = \{x^* : (-x^*, y^*) \in K_a^*(z)\} \quad (3.9)$$

называется *локально сопряженным* к a в точке z.

Так как конус $K_a(z)$ определен не однозначно, то и локально сопряженное отображение определено не однозначно. Однако, если отображение a — выпуклое, т. е. $\operatorname{gf} a$ есть выпуклое множество, то в качестве $K_a(z)$ всегда будет браться конус

$$K_a(z) = \operatorname{con}(\operatorname{gf} a - z); \quad (3.10)$$

поэтому локально сопряженное отображение в этом случае будет совпадать с введенным в § 2 главы III.

Если отображение $a_z(\bar{x})$ определить по формуле

$$a_z(\bar{x}) = \{\bar{y} : (\bar{x}, \bar{y}) \in K_a(z)\}, \quad (3.11)$$

то

$$a^*(y^*; z) = a_z^*(y^*); \quad (3.12)$$

отображение a_z^* задается определением III.1.9, если в нем в качестве конуса K взять конус $K_a(z)$.

В дальнейшем, как правило, будут рассматриваться выпуклозначные отображения. Для таких отображений направления

$$\gamma(0, y_1 - y), \quad y_1 \in a(x), \quad z = (x, y), \quad \gamma > 0,$$

в силу выпуклости $a(x)$ всегда будут касательными, так как при достаточно малом $\lambda > 0$ выполняется соотношение

$$(x,\ \ y + \lambda\gamma(y_1 - y)) \in \operatorname{gf} a.$$

Поэтому для выпуклозначных отображений всюду в дальнейшем будет предполагаться, что

$$K_a(z) \supseteq (0,\ \operatorname{con}(a(x) - y)), \quad z = (x,\ y) \in \operatorname{gf} a. \qquad (3.13)$$

Лемма 3.3. *Если выполнено соотношение* (3.13), *то* $a^*(y^*;\ z) \neq \varnothing$ *в том случае, если* $y \in a(x;\ y^*)$.

Доказательство. Если $x^* \in a^*(y^*;\ z)$, то по определению

$$-\langle \bar{x},\ x^*\rangle + \langle \bar{y},\ y^*\rangle \geqslant 0, \qquad (\bar{x},\ \bar{y}) \in K_a(z).$$

В частности, если выполнено включение (3.13), то

$$\langle y_1 - y,\ y^*\rangle \geqslant 0, \qquad y_1 \in a(x),$$

т. е.

$$W_a(x,\ y^*) \geqslant \langle y,\ y^*\rangle.$$

Но $y \in a(x)$, и поэтому

$$W_a(x,\ y^*) \leqslant \langle y,\ y^*\rangle.$$

Отсюда следует, что $W_a(x,\ y^*) = \langle y,\ y^*\rangle$, т. е. $y \in a(x;\ y^*)$.

Определение 3.2. Отображение

$$A^*(y^*;\ x) = \overline{\operatorname{co}}\Big(\bigcup_{y \in a(x;\, y^*)} a^*(y^*;\ (x,\ y))\Big) \qquad (3.14)$$

называется *сопряженным* к a в точке $x \in \operatorname{dom} a$.

Лемма 3.4. *Если a — выпуклое многозначное отображение, то*

$$A^*(y^*;\ x) = a^*(y^*;\ (x,\ y)) = \partial_x W_a(x,\ y^*), \qquad (3.15)$$

где y — любая точка $a(x;\ y^*)$.

Доказательство непосредственно следует из теоремы III.2.1.

Установим теперь связь между локально сопряженными отображениями и субдифференциалами функций, допускающих верхнюю выпуклую аппроксимацию.

Пусть $f_i(x)$, $i = 1,\ \ldots,\ m$, — функции, непрерывные в каждой точке $x \in \operatorname{dom} f_i$ и допускающие в каждой точке верхнюю выпуклую аппроксимацию $h_i(\bar{x},\ x)$. Составим вектор $f(x) \in \mathbf{R}^m$ с компонентами $f_i(x)$, положим

$$a(x) = \{y:\ y \geqslant f(x)\},$$
$$I(z) = \{i:\ y^i = f_i(x),\quad i = 1, \ldots, m\} \tag{3.16}$$

и вычислим $a^*(y^*;\ z)$. Пусть

$$K_a(z) = \{\bar{z} = (\bar{x}, \bar{y}):\ \bar{y}^i > h_i(\bar{x}, x),\quad i \in I(z)\}. \tag{3.17}$$

Из выпуклости и положительной однородности функции $h_i(\bar{x}, x)$ следует, что $K_a(z)$ — выпуклый конус. Если $(\bar{x}, \bar{y}) \in K_a(z)$, то

$$\bar{x} \in \bigcap_{i \in I(z)} \operatorname{dom} h_i(\cdot, x),$$

ибо иначе для конечного числа $\bar{y}^i$ невозможно неравенство $\bar{y}^i > h_i(\bar{x}, x)$.

Если $i \notin I(z)$, то в силу непрерывности $f_i(x)$ при достаточно малых $\lambda \geqslant 0$ выполняется неравенство

$$y^i + \lambda\bar{y}^i \geqslant f_i(x + \lambda\bar{x}),\qquad i \notin I(z),$$

так как $y^i > f_i(x)$. Если же $i \in I(z)$, то $y^i = f_i(x)$ и

$$\bar{y}^i > h_i(\bar{x}, x) \geqslant F_i(\bar{x}, x) \geqslant \limsup_{\lambda \downarrow 0} \frac{f_i(x + \lambda\bar{x}) - f_i(x)}{\lambda},$$

т. е. при достаточно малых λ справедливы неравенства

$$\bar{y}^i \geqslant \frac{f_i(x + \lambda\bar{x}) - f_i(x)}{\lambda},$$
$$y^i + \lambda\bar{y}^i \geqslant f_i(x + \lambda\bar{x}),\qquad i \in I(z).$$

Из сказанного следует, что конус $K_a(z)$ действительно есть конус касательных направлений к $\operatorname{gf} a$ в точке z.

Вычислим двойственный к нему конус $K_a^*(z)$. По определению $(-x^*,\ y^*) \in K_a^*(z)$ тогда и только тогда, когда

$$-\langle\bar{x}, x^*\rangle + \langle\bar{y}, y^*\rangle \geqslant 0,\qquad (\bar{x}, \bar{y}) \in K_a(z),$$

или

$$-\langle\bar{x}, x^*\rangle + \sum_{i \in I(z)} y^{i*}\bar{y}^i + \sum_{i \notin I(z)} y^{i*}\bar{y}^i \geqslant 0,$$
$$(\bar{x}, \bar{y}) \in K_a(z). \tag{3.18}$$

В силу формулы (3.17) y^i произвольно, в случае, если $i \notin I(z)$. Поэтому неравенство (3.18) возможно лишь при

$$y^{i*} = 0, \qquad i \notin I(z). \tag{3.19}$$

Так как для $i \in I(z)$, $\overline{y}^i > h_i(\overline{x}, x)$, то неравенство (3.18) возможно лишь, если

$$y^{i*} \geqslant 0, \qquad i \in I(z). \tag{3.20}$$

С учетом условий (3.19) и (3.20) неравенство (3.18) можно переписать в следующем эквивалентном виде:

$$-\langle \overline{x}, x^* \rangle + \sum_{i \in I(z)} y^{i*} h_i(\overline{x}, x) \geqslant 0,$$

$$\overline{x} \in \bigcap_{i \in I(z)} \operatorname{dom} h_i(\cdot, x). \tag{3.21}$$

Из соотношения (3.21) следует, что выпуклая функция

$$\varphi(\overline{x}) = -\langle \overline{x}, x^* \rangle + \sum_{i \in I(z)} y^{i*} h_i(\overline{x}, x) \tag{3.22}$$

достигает минимума по $\overline{x}$ на множестве

$$M = \bigcap_{i \in I(z)} \operatorname{dom} h_i(\cdot, x)$$

в точке $\overline{x} = 0$. Предположим, что

$$\operatorname{int} M = \bigcap_{i \in I(z)} \operatorname{int} \operatorname{dom} h_i(\cdot, x) \neq \varnothing. \tag{3.23}$$

Тогда, так как $\operatorname{dom} \varphi = M$, по теореме II.1.4 функция φ непрерывна в одной из точек M. Функции h_i положительно однородны и поэтому $\operatorname{dom} h_i(\cdot, x)$ есть выпуклый конус. Отсюда следует, что M есть также выпуклый конус и $K_M(0) = M$.

Применение теоремы IV.2.2 показывает, что в точке $\overline{x} = 0$ должно быть выполнено соотношение

$$\partial\varphi(0) \cap M^* \neq \varnothing. \tag{3.24}$$

Но по теореме II.3.8 субдифференциал $\partial\varphi$ состоит из векторов, имеющих вид

$$-x^* + \sum_{i \in I(z)} y^{i*} x_i^*, \qquad x_i^* \in \partial f_i(x), \qquad i \in I(z).$$

(Напомним, что $\partial f_i(x) \equiv \partial h_i(0, x)$ по определению.) С другой стороны, по теореме I.3.2

$$M^* = \sum_{i \in I(z)} (\operatorname{dom} h_i(\cdot, x))^*.$$

Из неравенства (3.24) следует, что существуют такие x_i^*, x_{0i}^*, $i \in I(z)$, что

$$-x^* + \sum_{i \in I(z)} y^{i*} x_i^* = \sum_{i \in I(z)} x_{i0}^*,$$

$$x_i^* \in \partial f_i(x), \quad x_{i0}^* \in (\operatorname{dom} h_i(\cdot, x))^*,$$

или

$$\begin{gathered} x^* = \sum_{i \in I(z)} \left(y^{i*} x_i^* - x_{0i}^* \right), \\ x_i^* \in \partial f_i(x), \quad x_{0i}^* \in (\operatorname{dom} h_i(\cdot, x))^*, \quad i \in I(z). \end{gathered} \tag{3.25}$$

Резюмируем полученный результат.

Теорема 3.1. *Пусть $f_i(x)$, $i = 1, \ldots, m$,— непрерывные функции, допускающие верхние выпуклые аппроксимации $h_i(\bar{x}, x)$. Пусть далее*

$$a(x) = \{y \in \mathbf{R}^m \colon y^i \geqslant f_i(x), \quad i = 1, \ldots, m\}.$$

Тогда если

$$I(z) = \{i \colon y^i = f_i(x), \ i = 1, \ldots, m\},$$

$$\bigcap_{i \in I(z)} \operatorname{int} \operatorname{dom} h_i(\cdot, x) \neq \varnothing,$$

то $a^*(y^*; z) = \varnothing$, когда $y^{i*} \neq 0$ для некоторого $i \notin I(z)$ или $y^{i*} < 0$ для некоторого $i \in I(z)$.

Если же $y^{i*} = 0$, $i \notin I(z)$, и $y^{i*} \geqslant 0$, $i \in I(z)$, то множество $a^*(y^*; z)$ состоит из элементов x^*, представимых в виде (3.25).

Следствие. *Если выполнены условия предыдущей теоремы и, кроме того,* $\operatorname{dom} h_i(\cdot, x) = X$, $i \in I(z)$, *то формула* (3.25) *может быть заменена формулой*

$$x^* = \sum_{i \in I(z)} y^{i*} x_i^*, \quad x_i^* \in \partial f_i(x), \quad i \in I(z).$$

В самом деле, в рассматриваемом случае

$$(\operatorname{dom} h_i(\cdot, x))^* = (X)^* = \{0\},$$

и поэтому вектор x_{0i}^* в формуле (3.25) может быть только нулевым.

Покажем теперь, что $K_a(z)$ есть не только конус касательных направлений к gf a в точке z, но и локальный шатер.

Лемма 3.5. *Если $h(\bar{x}, x)$ есть верхняя выпуклая аппроксимация f в точке x, то существует такая функция $r(\bar{x})$, $\|\bar{x}\|^{-1}r(\bar{x}) \to 0$ при $\bar{x} \to 0$, что*

$$f(x+\bar{x}) \leqslant f(x) + h(\bar{x}, x) + r(\bar{x}). \tag{3.26}$$

Доказательство. Допустим противное. Тогда найдется такая последовательность точек $\bar{x}_i \to 0$, что

$$\frac{f\left(x+\bar{x}_i\right) - f(x) - h\left(\bar{x}_i, x\right)}{\|\bar{x}_i\|} \geqslant \varepsilon > 0. \tag{3.27}$$

Положим $\lambda_i = \|\bar{x}_i\|$ и

$$\bar{x}_{i0} = \frac{\bar{x}_i}{\|\bar{x}_i\|}.$$

Без ограничения общности можно считать, что $\bar{x}_{i0} \to \bar{x}_0$, $\|\bar{x}_0\| = 1$, и что последовательность λ_i монотонно убывает. Положим

$$\varphi(\lambda_i) = \lambda_i(\bar{x}_{i0} - \bar{x}_0)$$

и доопределим $\varphi(\lambda)$ для $\lambda \in [\lambda_{i+1}, \lambda_i]$ линейной интерполяцией. Тогда $\lambda^{-1}\varphi(\lambda) \to 0$, ибо $\bar{x}_{i0} - \bar{x}_0$ стремится к нулю. Заметим теперь, что

$$\limsup_{i\to\infty} \frac{f\left(x+\bar{x}_i\right) - f(x)}{\|\bar{x}_i\|} =$$

$$= \limsup_{i\to\infty} \frac{f\left(x+\lambda_i\bar{x}_0 + \varphi(\lambda_i)\right) - f(x)}{\lambda_i} \leqslant F\left(\bar{x}_0, x\right).$$

С другой стороны, из замкнутости $h(\bar{x}, x)$ следует, что

$$\liminf_{i\to\infty} h\left(\bar{x}_{i0}, x\right) \geqslant h\left(\bar{x}_0, x\right).$$

Поэтому имеет место следующая цепочка неравенств:

$$\varepsilon \leqslant \limsup_{i\to\infty}\left[\frac{f\left(x+\bar{x}_i\right)-f\left(x_i\right)}{\left\|\bar{x}_i\right\|}-h\left(\frac{\bar{x}_i}{\left\|\bar{x}_i\right\|}, x\right)\right] \leqslant$$

$$\leqslant \limsup_{i\to\infty}\frac{f\left(x+\bar{x}_i\right)-f\left(x_i\right)}{\left\|\bar{x}_i\right\|}-\liminf_{i\to\infty} h\left(\bar{x}_{i0}, x\right) \leqslant$$

$$\leqslant F\left(\bar{x}_0, x\right)-h\left(\bar{x}_0, x\right)\leqslant 0.$$

Получаем противоречие. Лемма доказана.

Теорема 3.2. *Пусть $f_i(x)$, $i=1, \ldots, m$,— непрерывные функции. Тогда конус $K_a(z)$, определяемый формулой* (3.17), *есть гладкий локальный шатер к* gf a *в точке z.*

Доказательство. Пусть $\bar{z}_0=(\bar{x}_0, \bar{y}_0)\in \operatorname{ri} K_a(z)$. Рассмотрим функции

$$\varphi_i(\bar{z})=h_i(\bar{x}, x)-y^i, \quad i\in I(z).$$

Эти функции выпуклы и принимают на относительно открытом в $\operatorname{Lin} K_a(z)$ множестве конечные значения, ибо по определению $\varphi_i(\bar{z})<0$, если $\bar{z}\in K_a(z)$. По теореме II.1.4 $\varphi_i(\bar{z})$ непрерывны в $\operatorname{ri} K_a(z)$. Выберем теперь $\varepsilon>0$ настолько малым, чтобы

$$\Omega\equiv \bar{z}_0+(\varepsilon B_z)\cap \operatorname{Lin} K_a(z)\subseteq K_a(z), \tag{3.28}$$

где B_z обозначает единичный шар в Z, и для $\bar{z}\in\Omega$ выполнялись условия

$$\varphi_i\left(\bar{z}\right)\leqslant-\frac{\delta}{2}, \qquad \delta=\min_{i\in I(z)}\left[\bar{y}_0^i-h_i\left(\bar{x}_0, x\right)\right]. \tag{3.29}$$

Положим

$$Q=\operatorname{con}\Omega.$$

Если теперь $\bar{z}\in Q$, то для некоторого $\gamma>0$

$$\bar{z}=\gamma(\bar{z}_0+u), \quad \|u\|<\varepsilon, \quad u\in \operatorname{Lin} K_a(z),$$

т. е.

$$\left\|\frac{\bar{z}}{\gamma}-\bar{z}_0\right\|<\varepsilon.$$

Отсюда следует, что

$$\gamma > \frac{\|\bar{z}\|}{\|\bar{z}_0\| + \varepsilon}. \tag{3.30}$$

В силу леммы 3.5 и с учетом неравенства $\|\bar{x}\| \leqslant \|\bar{z}\|$ получаем

$$f_i(x + \bar{x}) - (y^i + \bar{y}^i) \leqslant f_i(x) + h_i(\bar{x}, x) + r(\bar{x}) - f_i(x) - \\ - \bar{y}^i = r(\bar{x}) + \varphi_i(\bar{z}) = r(\bar{x}) + \gamma\varphi_i(\bar{z}_0 + u) \leqslant \\ \leqslant r(\bar{x}) - \gamma\frac{\delta}{2} \leqslant \|\bar{x}\|\left(\frac{r(\bar{x})}{\|\bar{x}\|} - \frac{\|\bar{z}\|}{\|\bar{x}\|}\frac{\delta}{2(\|\bar{z}_0\| + \varepsilon)}\right) \leqslant \\ \leqslant \|\bar{x}\|\left(\frac{r(\bar{x})}{\|\bar{x}\|} - \frac{\delta}{2(\|\bar{z}_0\| + \varepsilon)}\right).$$

Пусть $\varepsilon_1 > 0$ выбрано настолько малым, что

$$\frac{r(\bar{x})}{\|\bar{x}\|} < \frac{\delta}{2(\|\bar{z}_0\| + \varepsilon)},$$

$$f_i(x + \bar{x}) - (y^i + \bar{y}^i) < 0, \quad i \notin I(z),$$

если $\|\bar{x}\| \leqslant \|\bar{z}\| < \varepsilon_1$ (в силу леммы 3.5 это возможно). Тогда для $\bar{z} \in (\varepsilon_1 B_z) \cap Q$ из предыдущего неравенства следует

$$f_i(x + \bar{x}) - (y^i + \bar{y}^i) < 0, \quad i \in I(z),$$

т. е. $z + \bar{z} = (x + \bar{x},\ y + \bar{y}) \in \operatorname{gf} a$.

Таким образом, построен конус Q, $\bar{z}_0 \in \operatorname{ri} Q$, $\operatorname{Lin} Q = \operatorname{Lin} K_a(z)$, $Q \subseteq K_a(z)$, и функция $\psi(\bar{z}) = \bar{z}$ такие, что все условия определения 1.3 выполнены. Это доказывает, что $K_a(z)$ — локальный шатер, причем гладкий, так как ψ — тождественное, а значит, гладкое отображение.

Приведем еще один пример вычисления сопряженного отображения.

Теорема 3.3. *Пусть $f(z)$ обладает в точке z_0, $f(z_0) = 0$, выпуклой верхней аппроксимацией $h(\bar{z}, z_0)$ и пусть $h(\bar{z}, z_0)$ непрерывна по $\bar{z}$ и существует такой вектор $\bar{z}_1$, что $h(\bar{z}_1, z_0) < 0$. Тогда:*

1) *конус*

$$K_a(z_0) = \{\bar{z}\colon h(\bar{z}, z_0) < 0\}$$

есть гладкий локальный шатер, являющийся конусом касательных направлений к gf a *в точке* z_0, *где* a *задается соотношением*

$$\operatorname{gf} a = \{z\colon f(z) \leqslant 0\},\ \ a(x) = \{y\colon f(x, y) \equiv f(z) \leqslant 0\};$$

2) *соответствующее* $K_a(z_0)$ *локально сопряженное отображение может быть задано формулой*

$$a^*(y^*;\ z_0) = \{x^*\colon\ (-x^*, y^*) \in (-\operatorname{con} \partial f(z_0))\},$$

где $\partial f(z_0)$ — *субдифференциал, соответствующий* $h(\bar{z}, z_0)$.

Доказательство. Очевидно, что конус $K_a(z_0)$ не пуст $(\bar{z}_1 \in K_a(z_0))$ и открыт ($h(\bar{z}, z_0)$ непрерывна по $\bar{z}$). Поэтому $K_a(z) = \operatorname{int} K_a(z)$.

Пусть $\bar{z}_0 \in K_a(z)$. Выберем $\varepsilon > 0$ настолько малым, что для всех $\bar{z}$ из окрестности $\Omega = \bar{z}_0 + \varepsilon B_z$, точки $\bar{z}_0$ выполняется неравенство

$$h(\bar{z}, z_0) < \frac{1}{2} h(\bar{z}_0, z_0) < 0.$$

Положим $Q = \operatorname{con}\Omega$. По лемме 3.5 для $\bar{z} \in Q$, т. е. для $\bar{z}$, имеющих вид $\bar{z} = \gamma(\bar{z}_0 + u)$, $\|u\| < \varepsilon$, с учетом неравенства (3.30) получаем

$$f(z_0 + \bar{z}) \leqslant f(z_0) + h(\bar{z}, z_0) + r(\bar{z}) =$$
$$= \gamma h(\bar{z}_0 + u, z_0) + r(\bar{z}) \leqslant \gamma \frac{1}{2} h(\bar{z}_0, z_0) + r(\bar{z}) \leqslant$$
$$\leqslant \|\bar{z}\| \left(\frac{h(\bar{z}_0, z_0)}{2(\|\bar{z}_0\| + \varepsilon)} + \frac{r(\bar{z})}{\|\bar{z}\|} \right).$$

Если теперь ε_1 выбрано так, что

$$\frac{h(\bar{z}_0, z_0)}{2(\|\bar{z}_0\| + \varepsilon)} + \frac{r(\bar{z})}{\|\bar{z}\|} < 0$$

для $\|\bar{z}\| < \varepsilon_1$, то согласно предыдущему неравенству

$$f(z_0 + \bar{z}) < 0, \qquad z \in (\varepsilon_1 B_z) \cap Q,$$

т. е. $z_0 + \bar{z} \in \operatorname{gf} a$ для $\bar{z} \in (\varepsilon_1 B_z) \cap Q$. Отсюда следует, что $K_a(z_0)$ — локальный шатер к $\operatorname{gf} a$ в точке z_0, а значит и конус касательных направлений.

Перейдем к доказательству второй части теоремы. Так как $h(\bar{z}, z_0)$ — непрерывная выпуклая функция, то замыкание конуса $K_a(z)$ легко вычислить по формуле

$$\overline{K_a(z_0)} = \{\bar{z}: h(\bar{z}, z_0) \leqslant 0\}. \tag{3.31}$$

Действительно, если $h(\bar{z}_0, z_0) = 0$, то для $0 < \lambda < 1$ выполняется неравенство

$$h(\lambda\bar{z}_1 + (1-\lambda)\bar{z}_0, z_0) \leqslant$$
$$\leqslant \lambda h(\bar{z}_1, z_0) + (1-\lambda) h(\bar{z}_0, z_0) = \lambda h(\bar{z}_0, z_0) < 0.$$

Откуда следует, что $\lambda\bar{z}_1 + (1-\lambda)\bar{z}_0 \in K_a(z)$. Устремляя λ к нулю получаем, что $\bar{z}_0$ есть предельная точка $K_a(z)$.

Вспомним теперь, что

$$h(\bar{z}, z_0) = \max_{z_0^*} \{\langle \bar{z}, z_0^* \rangle: z_0^* \in \partial f(z_0)\}$$

и $\partial f(z_0) = \partial h(0, z_0)$ — обычный субдифференциал выпуклой функции $h(\bar{z}, z_0)$. Тогда неравенство $h(\bar{z}, z_0) \leqslant 0$ эквивалентно неравенству

$$\langle \bar{z}, -z_0^* \rangle \geqslant 0, \qquad z_0^* \in \partial f(z_0).$$

Последнее же эквивалентно неравенству

$$\langle \bar{z}, z^* \rangle \geqslant 0, \quad z^* \in K, \tag{3.32}$$

где

$$K = -\operatorname{con} \partial f(z_0) =$$
$$= \{-z^*: z^* = \gamma z_0^*, \quad \gamma \geqslant 0, \quad z_0^* \in \partial f(z_0)\}. \tag{3.33}$$

Но соотношение (3.32) означает, что $\bar{z} \in K^*$. Таким образом, в силу формул (3.31), (3.32) $\bar{z} \in \overline{K_a(z_0)}$ тогда и только тогда, когда $\bar{z} \in K^*$, т. е.

$$\overline{K_a(z_0)} = K^*.$$

Покажем, что K — замкнутый конус. Пусть $\gamma_i z_i^* \to z^*$, $z^* \neq 0$, $\gamma_i \geqslant 0$, $z_i^* \in \partial f(z_0)$. Так как $\partial f(z_0)$ есть субдифференциал выпуклой непрерывной функции $h(\overline{z}, z_0)$, то согласно теоремам II.3.2 и II.3.5 $\partial f(z_0)$ есть выпуклое замкнутое ограниченное множество. Далее,

$$h(\overline{z}_1, z_0) = \max_{z_0^*}\{\langle \overline{z}_1, z_0^*\rangle: \ z_0^* \in \partial f(z_0)\} < 0,$$

и поэтому $0 \notin \partial f(z_0)$. Значит $\|z_0^*\| \geqslant \delta > 0$ для всех $z_0^* \in \partial f(z_0)$. Так как $\partial f(z_0)$ ограничено, то $\|\overline{z}_0^*\| \leqslant \Delta$ для всех $z_0^* \in \partial f(z_0)$. Заметим теперь, что для больших i выполняется неравенство

$$\frac{1}{2}\|z^*\| \leqslant \gamma_i \|z_i^*\| < 2\|z^*\|.$$

Поэтому

$$\frac{1}{2}\frac{\|z^*\|}{\|z_i^*\|} \leqslant \gamma_i \leqslant 2\frac{\|z^*\|}{\|z_i^*\|},$$

$$\frac{1}{2}\frac{\|z^*\|}{\Delta} \leqslant \gamma_i \leqslant 2\frac{\|z^*\|}{\delta}.$$

Отсюда следует, что последовательность γ_i ограничена. Без потери общности можно считать, что $\gamma_i \to \gamma_0$, причем

$$0 < \frac{1}{2}\frac{\|z^*\|}{\Delta} < \gamma_0.$$

Поэтому

$$z_i^* = \frac{(\gamma_i z_i^*)}{\gamma_i} \to \frac{z^*}{\gamma_0}.$$

Так как $\partial f(z_0)$ — замкнутое множество, то из $z_i^* \in \partial f(z_0)$ следует, что

$$\frac{z^*}{\gamma_0} = z_0^* \in \partial f(z_0).$$

Итак, $z^* = \gamma_0 z_0^*$, $\gamma_0 > 0$, $z_0^* \in \partial f(z_0)$, что доказывает замкнутость множества K.

Вспомним теперь, что согласно § 3 главы I для замкнутого конуса K справедливо тождество $K^{**} = K$. Отсюда получаем

$$(K_a(z_0))^* = (\overline{K_a(z_0)})^* = (K^*)^* = K,$$

где использовано ранее доказанное соотношение $\overline{K_a(z_0)} = K^*$.

Итак,

$$K_a^*(z_0) = -\operatorname{con} \partial f(z_0).$$

Отсюда согласно определению 3.1 получаем

$$a^*(y^*;\ z_0) = \{x^*:\ (-x^*,\ y^*) \in (-\operatorname{con} \partial f(z_0))\},$$

что завершает доказательство.

Теорема 3.4. *Пусть $f_i(z)$, $i = 1, \ldots, m$,— заданные функции и многозначное отображение a определяется формулами*

$$\operatorname{gf} a = \{z:\ f_i(z) \leqslant 0,\ i = 1, \ldots, m\},$$
$$a(x) = \{y:\ f_i(x,\ y) \equiv f_i(z) \leqslant 0,\ \ i = 1, \ldots, m\}.$$

Положим для $z \in \operatorname{gf} a$

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

Тогда, если $I(z) = \varnothing$, то $K_a(z) = Z$, $K_a^*(z) = \{0\}$. Если же $I(z) \neq \varnothing$, $h_i(\bar{z},\ z)$, $i \in I(z)$, а верхние выпуклые аппроксимации для f_i в точке z непрерывны по $\bar{z}$ и существует такое $\bar{z}_1$, что

$$h_i(\bar{z}_1,\ z) < 0,\ i \in I(z),$$

то конус

$$K_a(z) = \{\bar{z}:\ h_i(\bar{z},\ z) < 0,\ i \in I(z)\}$$

является гладким локальным шатром для $\operatorname{gf} a$ в точке z и

$$K_a^*(z) = -\sum_{i \in I(z)} \operatorname{con} \partial f_i(z),$$

$$a^*(y^*;\ z) = \left\{x^*:\ (-x^*,\ y^*) \in \left(-\sum_{i \in I(z)} \operatorname{con} \partial f_i(z)\right)\right\},$$

где $\partial f_i(z)$ — субдифференциалы, соответствующие $h_i(\bar{z},\ z)$.

Доказательство в основном повторяет доказательство теоремы 3.3 и поэтому мы его опустим. Заметим только, что

$$K_a(z) = \bigcap_{i \in I(z)} K_i,$$

где $K_i = \{\bar{z}: h_i(\bar{z}, z) < 0\}$ (согласно доказанному выше) — открытые конусы. В силу предположения теоремы существует точка z_1, принадлежащая всем этим конусам. Поэтому в силу теоремы I.3.2 выполняется равенство

$$K_a^*(z) = \sum_{i \in I(z)} K_i^*,$$

что с учетом выражения для K_i^*:

$$K_i^* = \{z^*: z^* \in (-\operatorname{con} \partial f_i(z))\},$$

следующего из доказательства предыдущей теоремы, завершает доказательство.

Пусть теперь фиксированы многозначное отображение a и функция $\varphi(z)$. Положим

$$f(x) = \inf_y \{\varphi(x, y): y \in a(x)\}, \tag{3.34}$$

$$a_\varphi(x) = \{y \in a(x): \varphi(x, y) = f(x)\}. \tag{3.35}$$

Теорема 3.5. *Пусть $\varphi(z)$ непрерывно дифференцируема по z и $\varphi'_x(z)$, $\varphi'_y(z)$ векторы ее частных производных по x и y, $z = (x, y)$ соответственно. Допустим, что функция*

$$l(\bar{x}) = \inf_{\bar{y}} \{\langle \bar{y}, y^* \rangle: (\bar{x}, \bar{y}) \in K_a(z)\}$$

замкнута, а функция $f(x)$, определяемая соотношением (3.34), удовлетворяет условию Липшица. Тогда для любого $y \in a_\varphi(x)$, $z = (x, y)$, функция

$$h(\bar{x}, x) = \langle \bar{x}, \varphi'_x(z) \rangle + \inf_{\bar{y}} \{\langle \bar{y}, \varphi'_y(z) \rangle: (\bar{x}, \bar{y}) \in K_a(z)\}$$

есть верхняя выпуклая аппроксимация для функции f, а

$$\partial f(x) = \varphi'_x(z) + a^*(\varphi'_y(z); z)$$

— соответствующий субдифференциал.

Доказательство. Пусть $\bar{z} = (\bar{x}, \bar{y}) \in K_a(z)$, $z = (x, y)$, $y \in a_\varphi(z)$. По определению конуса касательных направлений существует такая функция $r(\lambda) \in z$,

$\lambda^{-1}r(\lambda) \to 0$ при $\lambda \downarrow 0$, что

$$z + \lambda\bar{z} + r(\lambda) \in \operatorname{gf} a$$

при достаточно малых $\lambda \geqslant 0$, или

$$y + \lambda\bar{y} + r_y(\lambda) \in a(x + \lambda\bar{x} + r_x(\lambda)), \tag{3.36}$$

где r_x, r_y — компоненты функции r, соответствующие пространствам X и Y, в прямое произведение которых разлагается Z.

Согласно замечанию, сделанному после определения 2.1, если функция $f(x)$ удовлетворяет условию Липшица, то

$$F(\bar{x}, x) = \limsup_{\lambda \downarrow 0} \frac{f(x + \lambda\bar{x}) - f(x)}{\lambda}.$$

Точно также можно было бы показать, что

$$F(\bar{x}, x) = \limsup_{\lambda \downarrow 0} \frac{f(x + \lambda\bar{x} + r(\lambda)) - f(x)}{\lambda}$$

независимо от выбора функции $r(\lambda)$, $\lambda^{-1}r(\lambda) \to 0$ при $\lambda \to 0$. По определению формулой (3.34) функции f и из того, что $y \in a_\varphi(x)$, вытекает неравенство

$$\frac{f(x + \lambda\bar{x} + r_x(\lambda)) - f(x)}{\lambda} \leqslant \frac{\varphi(z + \lambda\bar{z} + r(\lambda)) - \varphi(z)}{\lambda}.$$

Поэтому

$$F(\bar{x}, x) = \limsup_{\lambda \downarrow 0} \frac{f(x + \lambda\bar{x} + r_x(\lambda)) - f(x)}{\lambda} \leqslant$$
$$\leqslant \limsup_{\lambda \downarrow 0} \frac{\varphi(z + \lambda\bar{z} + r(\lambda)) - \varphi(z)}{\lambda} =$$
$$= \langle \bar{x}, \varphi_x'(z) \rangle + \langle \bar{y}, \varphi_y'(z) \rangle.$$

Если x зафиксировано, то $\bar{y}$ может быть выбрано любым из множества

$$a_z(\bar{x}) = \{\bar{y}\colon (\bar{x}, \bar{y}) \in K_a(z)\}.$$

Поэтому

$$F(\bar{x}, x) \leqslant \langle \bar{x}, \varphi_x'(z) \rangle + \inf_{\bar{y}} \{\langle \bar{y}, \varphi_y'(z) \rangle\colon \bar{y} \in a_z(\bar{x})\}. \tag{3.37}$$

Если $a_z(\bar{x}) = \varnothing$, то по соглашению

$$\inf_{\bar{y}} \{\langle \bar{y}, \varphi_y'(z)\rangle: \bar{y} \in a_z(\bar{x})\} = +\infty,$$

т. е. соотношение (3.37) выполняется тривиальным образом.

Заметим теперь, что если применить теорему III.4.5 к отображению a_z, задаваемому конусом $\operatorname{gf} a_z = K_a(z)$, то придем к следующему соотношению:

$$\inf_{\bar{y}} \{\langle \bar{y}, \varphi_y'(z)\rangle: \bar{y} \in a_z(\bar{x})\} = \\ = \sup_{x^*} \{\langle \bar{x}, x^*\rangle: x^* \in a_z^*(\varphi_y'(z))\},$$

или, учитывая формулы (3.11), (3.12),

$$\inf_{\bar{y}} \{\langle \bar{y}, \varphi_y'(z)\rangle: \bar{y} \in a_z(\bar{x})\} = \\ = \sup_{x^*} \{\langle \bar{x}, x^*\rangle: x^* \in a^*(\varphi_y'(z); z)\}.$$

Итак, если обозначить

$$h(\bar{x}, x) = \\ = \langle \bar{x}, \varphi_x'(z)\rangle + \sup_{x^*} \{\langle \bar{x}, x^*\rangle: x^* \in a^*(\varphi_y'(z), z)\},$$

то выполняется неравенство $F(\bar{x}, x) \leqslant h(\bar{x}, x)$. Из выражения для функции h видно, что это — положительно однородная выпуклая, замкнутая функция $\bar{x}$.

Итак, $h(\bar{x}, x)$ есть верхняя выпуклая аппроксимация f в точке x. Использование теоремы II.3.11 показывает, что

$$\partial f(x) \equiv \partial f(0, x) = \varphi_x'(z) + a^*(\varphi_y'(z), z).$$

Доказательство теоремы завершено.

§ 4. Общие необходимые условия минимума

Построение необходимых условий экстремума тесно связано с классами множеств и функций, которые участвуют в задаче. В предыдущих параграфах такие классы были введены и исследованы, так что теперь формулировка необходимых условий экстремума может быть дана сравнительно просто.

1. Ограничения, задаваемые произвольными множествами.

Теорема 4.1. *Пусть x_0 — точка минимума функции $f(x)$ на множестве M и пусть $h(\bar{x}, x_0)$ — верхняя выпуклая аппроксимация f в точке x_0. Тогда, если выполнено условие*

$$\operatorname{int} \operatorname{dom} h(\cdot, x_0) \cap K_M(x_0) \neq \varnothing,$$

то $\partial f(x_0) \cap K_M^*(x_0) \neq \varnothing$.

Доказательство. Так как x_0 — точка минимума, то $h(\bar{x}, x_0) \geqslant 0$ для всех $\bar{x} \in K_M(x_0)$. В самом деле, если $\bar{x} \in K_M(x_0)$ и $h(\bar{x}, x_0) < 0$, то существует такая функция $r(\lambda)$, $\lambda^{-1} r(\lambda) \to 0$ при $\lambda \downarrow 0$, что

$$x_0 + \lambda\bar{x} + r(\lambda) \in M.$$

Поэтому

$$\limsup_{\lambda \downarrow 0} \frac{f\left(x_0 + \lambda\bar{x} + r(\lambda)\right) - f(x_0)}{\lambda} \leqslant F(\bar{x}, x_0) \leqslant h(\bar{x}, x_0) < 0,$$

т. е. при достаточно малых $\lambda > 0$ $f(x_0 + \lambda\bar{x} + r(\lambda)) < f(x_0)$ в противоречии с тем, что x_0 — точка минимума.

Таким образом, выпуклая функция $h(\bar{x}, x_0)$ достигает своего минимума на выпуклом множестве $K_M(x_0)$ в точке $\bar{x} = 0$. Согласно теореме IV.2.2 это возможно лишь если

$$\partial f(x_0) \cap K_M^*(x_0) \neq \varnothing,$$

что и требовалось доказать.

Следствие. *Пусть функция f допускает верхнюю выпуклую аппроксимацию $h(\bar{x}, x_0)$ в точке x_0. Тогда для того, чтобы точка x_0 была точкой минимума функции f необходимо выполнение условия $0 \in \partial f(x_0)$.*

Важно отметить, что это условие должно выполняться для любого субдифференциала. Для иллюстрации рассмотрим функцию

$$f_c(x) = \begin{cases} 0, & x < 0, \\ c, & x \geqslant 0. \end{cases}$$

Тогда, если $c > 0$, то согласно вычислениям, проведенным в § 2, любое число $a \geqslant 0$ является субдифференциалом. Таким образом, условие $0 \in \partial f_c(0)$ не выполнено. С другой стороны, если $c < 0$, то любой субдифференциал $\partial f_c(0)$ содержит множество $(-\infty, 0]$, и, значит, $0 \in \partial f_c(0)$. Этот пример показывает, что полученные необходимые условия минимума могут быть эффектными даже для разрывных функций.

Теорема 4.2. *Пусть x_0 — точка минимума функции $f(x_0)$ на множестве*

$$M = \bigcap_{i=1}^{m} M_i.$$

Пусть, далее, f допускает в x_0 верхнюю выпуклую аппроксимацию $h(\bar{x}, x_0)$, конусы $K_{M_i}(x_0)$ являются локальными шатрами и

$$\operatorname{int} \operatorname{dom} h(\cdot, x_0) \cap \left(\bigcap_{i=1}^{m} K_{M_i}(x_0) \right) \neq \varnothing.$$

Тогда существует такое число $\lambda \geqslant 0$ и такие векторы $x_i^* \in K_{M_i}^*(x_0)$, не все одновременно равные нулю, что

$$\lambda x_0^* = \sum_{i=1}^{m} x_i^*, \qquad x_0^* \in \partial f(x_0).$$

Доказательство. Если конусы $K_i \equiv K_{M_i}(x_0)$ отделимы, то существуют такие не все равные нулю векторы x_i^*, $i = 1, \ldots, m$, что

$$\sum_{i=1}^{m} x_i^* = 0, \qquad x_i^* \in K_i^*,$$

и требуемый теоремой результат получается, если положить $\lambda = 0$.

Если же конусы K_i, $i = 1, \ldots, m$, неотделимы, то согласно теореме 1.2 для любого вектора

$$\bar{x}_0 \in \operatorname{ri} K, \qquad K = \bigcap_{i=1}^{m} K_i,$$

существуют такие конус Q и функция $\psi(\bar{x})$, что

$$\bar{x}_0 \in \operatorname{ri} Q, \qquad \operatorname{Lin} Q = \operatorname{Lin} K, \qquad Q \subseteq K,$$

$$\psi(\bar{x}) = \bar{x} + r(\bar{x}), \qquad \|\bar{x}\|^{-1} r(\bar{x}) \to 0,$$

если $\bar{x} \to 0$, и, кроме того, $x_0 + \psi(\bar{x}) \in M$ для $\bar{x} \in Q \cap (\varepsilon B)$, $\varepsilon > 0$. Поэтому при достаточно малых $\lambda > 0$ выполняется соотношение

$$x_0 + \psi(\lambda \bar{x}_0) = x_0 + \lambda \bar{x}_0 + r(\lambda \bar{x}_0) \in M,$$

$\lambda^{-1} r(\lambda \bar{x}_0) \to 0$ при $\lambda \downarrow 0$ и, значит, $\bar{x}_0$ — касательное направление.

Таким образом, конус $\mathrm{ri}\, K$ есть конус касательных направлений к M в точке x_0. По предположению теоремы

$$\mathrm{int}\,\mathrm{dom}\, h(\cdot,\, x_0) \cap K \neq \varnothing,$$

т. е. существует вектор $\bar{x}_1 \in \mathrm{int}\,\mathrm{dom}\, h(\cdot,\, x_0)$ и $\bar{x}_1 \in K$. Так как любой вектор из выпуклого множества может быть приближен векторами из его относительной внутренности и $\bar{x}_1$ есть внутренняя точка $\mathrm{dom}\, h(\cdot,\, x_0)$, то найдется вектор $\bar{x}_2 \in \mathrm{ri}\, K$, одновременно принадлежащий $\mathrm{int}\,\mathrm{dom}\, h(\cdot,\, x_0)$. Таким образом, $\mathrm{ri}\, K$ есть конус касательных направлений к M в точке x_0, для которого выполнены предположения теоремы 4.1. Поэтому

$$\partial f(x_0) \cap (\mathrm{ri}\, K)^* \neq \varnothing.$$

Но согласно лемме I.1.6 $\bar{K} = \overline{(\mathrm{ri}\, K)}$ и $(\mathrm{ri}\, K)^* = K^*$. Следовательно,

$$\partial f(x_0) \cap K^* \neq \varnothing.$$

Так как конусы K_i, $i = 1, \ldots, m$, неотделимы, то в соответствии с теоремой I.3.3 можем записать, что

$$K^* = \sum_{i=1}^{m} K_i^*.$$

Это означает, что найдется такой вектор $x_0^* \in \partial f(x_0)$ и такие векторы $x_i^* \in K_i^*$, что

$$x_0^* = \sum_{i=1}^{m} x_i^*.$$

Доказательство завершено.

Следствие 1. *Пусть x_0 точка минимума функции $f_0(x)$ при дополнительном условии*

$$f_i(x) = 0, \qquad i = 1, \ldots, m,$$

где $f_i(x)$ — непрерывно дифференцируемые функции для $i = 0, 1, \ldots, m$. Тогда существуют такие числа y^{i*}, $i = 0, 1, \ldots, m$, что

$$\sum_{i=0}^{m} y^{i*} f_i'(x_0) = 0, \qquad y^{0*} \geqslant 0,$$

и не все числа y^{i*} равны нулю одновременно.

Доказательство. Рассмотрим множество

$$M = \{x\colon\ f_i(x) = 0, \quad i = 1, \ldots, m\}.$$

Возможны два случая: градиенты $f_i'(x_0)$, $i = 1, \ldots, m$, линейно зависимы и градиенты линейно независимы. В первом случае положим $y^{0*} = 0$, а в качестве y^{i*}, $i = 1, \ldots, m$, возьмем числа из линейной комбинации

$$\sum_{i=1}^{m} y^{i*} f_i'(x_0) = 0.$$

Во втором случае согласно примеру 1.1 имеем

$$K_M(x_0) = \{\bar{x} : \langle \bar{x},\ f_i'(x_0)\rangle = 0,\ i = 1, \ldots, m\},$$

$$K_M^*(x_0) = \left\{x^*\colon\ x^* = \sum_{i=1}^{m} \lambda_i f_i'(x_0), \quad \lambda_i \in \mathbf{R}^1, \quad i = 1, \ldots, m\right\}.$$

Так как $f_0(x)$ — гладкая функция, то $\partial f_0(x_0) = \{f_0'(x_0)\}$, и применение теоремы 4.1 дает равенство

$$f_0'(x_0) = \sum_{i=1}^{m} \lambda_i f_i'(x_0),$$

откуда получается требуемый результат, если положить

$$y^{0*} = 1, \quad y^{i*} = -\lambda_i, \quad i = 1, \ldots, m.$$

Следствие 2. *Пусть x_0 — точка минимума функции $f_0(x)$ при ограничениях*

$$f_i(x) \leqslant 0, \quad i \in I^-, \quad f_i(x) = 0, \quad i \in I,$$

где функции $f_i(x)$, $i \in \{0\} \cup I^- \cup I$, непрерывно дифференцируемы. Тогда существуют такие числа y^{i*}, $i \in \{0\} \cup \cup I^- \cup I$, что

$$\sum_{i \in \{0\} \cup I^- \cup I} y^{i*} f_i'(x_0) = 0,$$

причем, не все числа y^{i*} равны нулю, и

$$y^{i*} \geqslant 0, \quad i \in \{0\} \cup I^-, \quad y^{i*} f_i(x_0) = 0, \quad i \in I^-.$$

Доказательство. Рассмотрим множества

$$M_i = \begin{cases} \{x:\ f_i(x) \leqslant 0\}, & i \in I^-, \quad f_i(x_0) = 0, \\ X, & i \in I^-, \quad f_i(x_0) < 0, \\ \{x:\ f_i(x) = 0\}, & i \in I. \end{cases}$$

В соответствии с примерами 1.3 и 1.4 конусы

$$K_i = \begin{cases} \{\bar{x}:\ \langle \bar{x}, f_i'(x_0)\rangle < 0\}, & i \in I^-, \quad f_i(x_0) = 0, \\ X, & i \in I^-, \quad f_i(x_0) < 0, \\ \{\bar{x}:\ \langle \bar{x}, f_i'(x_0)\rangle = 0\}, & i \in I \end{cases}$$

являются локальными шатрами к множествам M_i в точке x_0, если $f_i'(x_0) \neq 0$ для $i \in I^-$, $f_i(x_0) = 0$, или $i \in I$. Воспользовавшись тем, что

$$K_i^* = \begin{cases} \{-\lambda_i f_i'(x_0):\ \lambda_i \geqslant 0\}, & i \in I^-, \quad f_i(x_0) = 0, \\ \{0\}, & i \in I^-, \quad f_i(x_0) < 0, \\ \{-\lambda_i f_i'(x_0):\ \lambda_i \in \mathbf{R}^1\}, & i \in I, \end{cases}$$

результат теоремы 4.2 можно переписать в виде

$$\lambda x_0^* = -\sum_{i \in \{0\} \cup I^- \cup I} \lambda_i f_i'(x_0),$$

$$\lambda_i \geqslant 0, \quad i \in I^-, \quad \lambda_i f_i(x_0) = 0, \quad i \in I^-.$$

Положив $y^{0*} = \lambda$, $y^{i*} = \lambda_i$, получаем требуемый результат.

Если $f_{i_0}'(x_0) = 0$ для некоторого $i_0 \in I^-$, $f_i(x_0) = 0$ или $i \in I$, то достаточно положить $y^{i_0*} = 1$ и $y^{i*} = 0$ для всех остальных i.

Следствие 3. *Пусть x_0 — точка минимума функции $f_0(x)$ при ограничениях*

$$f_i(x) \leqslant 0, \qquad i \in I^-,$$
$$f_i(x) = 0, \qquad i \in I, \qquad x \in M,$$

где функции $f_i(x)$ непрерывно дифференцируемы, а множество M выпукло. Тогда существуют такие не все рав-

ные нулю числа y^{i*}, $i \in \{0\} \cup I^- \cup I$, *что*

$$\sum_{i \in \{0\} \cup I^- \cup I} y^{i*} f_i'(x_0) \in (\operatorname{con}(M - x_0))^*,$$

$$y^{i*} \geqslant 0, \quad i \in \{0\} \cup I^-, \quad y^{i*} f_i(x_0) = 0, \quad i \in I^-.$$

Доказательство получается из теоремы 4.2 совершенно аналогично предыдущему, если помимо множеств M_i рассмотреть и множество M, для которого, согласно примеру 1.5, $\operatorname{con}(M - x_0)$ есть локальный шатер. Действительно, в этом случае найдутся такие числа $\lambda_i \geqslant 0$, $i \in I^-$, что

$$\lambda f_0'(x_0) = -\sum_{i \in I^- \cup I} \lambda_i f_i'(x_0) + x^*, \quad x^* \in (\operatorname{con}(M - x_0))^*,$$

и $\lambda_i = 0$, если $f_i(x_0) < 0$. Из этой формулы и следует нужный результат, если положить $y^{0*} = \lambda$, $y^{i*} = \lambda_i$.

2. Ограничения, задаваемые многозначными отображениями. Исследуем теперь следующую задачу. Пусть a — многозначное отображение, W — некоторое множество в пространстве Y, M — некоторое множество в пространстве X. Требуется найти минимум функции $f_0(x)$ при условиях

$$a(x) \cap W \neq \varnothing, \qquad x \in M. \tag{4.1}$$

Ниже эта задача будет анализироваться при различных предположениях относительно a, W и M.

Теорема 4.3. *Пусть x_0 — точка минимума функции $f_0(x)$ при ограничениях (4.1). Предположим, что $f_0(x)$ в точке x_0 допускает верхнюю выпуклую аппроксимацию $h(\bar{x}, x_0)$, непрерывную по $\bar{x}$; $y_0 \in a(x_0) \cap W$ и $K_a(x_0, y_0)$, $K_W(y_0)$, $K_M(x_0)$ — соответственно локальные шатры к множествам* $\operatorname{gf} a$, W, M. *Тогда существует такое число $\lambda \geqslant 0$ и такие векторы $x_0^* \in \partial f(x_0)$, $y^* \in K_W^*(y_0)$, $x^* \in K_M^*(x_0)$, что*

$$x^* - \lambda x_0^* \in a^*(-y^*; z_0),$$

где $z_0 = (x_0, y_0)$, а λ, y^* и x^* не равны нулю одновременно.

Доказательство. Поставленную задачу в пространстве $Z = X \times Y$ можно интерпретировать как задачу минимизации функции $\varphi(z)$, $z = (x, y)$, определенной соотношением $\varphi(z) = f(x)$, на пересечении множеств $\operatorname{gf} a$,

$X \times W$, $M \times Y$, лежащих в пространстве Z. Легко видеть, что

$$\limsup_{\lambda \downarrow 0} \frac{\varphi\left(z_0 + \lambda \bar{z} + r(\lambda)\right) - \varphi(z_0)}{\lambda} \leqslant F(\bar{x}, x_0),$$

так что $h(\bar{x}, x_0)$ есть верхняя выпуклая аппроксимация для $\varphi(z)$ в точке $z_0 = (x_0, y_0)$. Так как $h(\bar{x}, x_0)$ от $\bar{y}$ не зависит, то субдифференциал, соответствующий этой верхней выпуклой аппроксимации функции $\varphi(z)$ в точке z_0, имеет следующий вид:

$$\partial\varphi(z_0) = \partial f(x_0) \times \{0\}, \tag{4.2}$$

т. е. состоит из векторов $z^* = (x^*, y^*)$, для которых $x^* \in \partial f(x_0)$, $y^* = 0$.

В силу предположений теоремы множества $K_a(z_0)$, $X \times K_W(y_0)$, $K_M(x_0) \times Y$ являются локальными шатрами к множествам gf a, $X \times Y$, $M \times Y$ в точке z_0 соответственно. Легко вычислить сопряженные конусы:

$$(X \times K_W(y_0))^* = \{0\} \times K_W^*(y_0), \tag{4.3}$$

$$(K_M(x_0) \times Y)^* = K_M^*(x_0) \times \{0\}. \tag{4.4}$$

Воспользуемся теперь теоремой 4.2. Из этой теоремы и формул (4.2)—(4.4) следует, что найдутся такие не равные нулю одновременно число $\lambda \geqslant 0$ и векторы y^*, x^*, (x_1^*, y_1^*), что

$$\lambda(x_0^*, 0) = (x_1^*, y_1^*) + (0, y^*) + (x^*, 0), \tag{4.5}$$

$$x_0^* \in \partial f(x_0), \quad (x_1^*, y_1^*) \in K_a^*(z_0), \tag{4.6}$$

$$y^* \in K_W^*(y_0), \quad x^* \in K_M^*(x_0). \tag{4.7}$$

Перепишем равенство (4.5) в следующем виде:

$$\lambda x_0^* = x_1^* + x^*, \quad y_1^* + y^* = 0. \tag{4.8}$$

Тогда $y_1^* = -y^*$ и

$$\left(-(-x_1^*), -y^*\right) \in K_a^*(z_0).$$

Согласно определению 3.1 последнее означает, что

$$-x_1^* \in a^*(-y^*; z_0).$$

Если переписать первое из равенств (4.8) в виде

$$-x_1^* = x^* - \lambda x_0^*,$$

то получим соотношение

$$x^* - \lambda x_0^* \in a^*(-y^*;\ z_0).$$

Согласно теореме 4.2 не все векторы, входящие в равенство (4.5), равны нулю одновременно. Поэтому величины λ, y^* и x^* одновременно нулю равны быть не могут. Это завершает доказательство.

Следствие 1. *Пусть $a(x) = \{f(x)\}$, где $f\colon X \to Y$ — однозначное гладкое отображение, W и M — выпуклые множества, $f_0(x)$ — гладкая функция. Тогда, если x_0 — точка минимума функции $f_0(x)$ при ограничениях*

$$f(x) \in W, \qquad x \in M, \tag{4.9}$$

то существуют такие одновременно не равные нулю числа $\lambda \geqslant 0$ и векторы x^* и y^*, что

$$\lambda f_0'(x_0) - f'^*(x_0)\,y^* = x^*,$$

$$x^* \in (\operatorname{con}(M - x_0))^*, \qquad y^* \in (\operatorname{con}(W - f(x_0)))^*.$$

Доказательство. Если $f\colon\ X \to Y$ — однозначное гладкое отображение, то его график есть множество $\{x,\ f(x)\}$, $x \in X$. По определению производная отображения f есть линейный оператор $f'(x)$ такой, что

$$\lim_{\|\bar{x}\| \to 0} \frac{f(x+\bar{x}) - f(x) - f'(x)\,\bar{x}}{\|\bar{x}\|} = 0,$$

или, иными словами,

$$f(x+\bar{x}) = f(x) + f'(x)\bar{x} + r(\bar{x}), \tag{4.10}$$

$\|\bar{x}\|^{-1} r(\bar{x}) \to 0$, когда $\bar{x} \to 0$. Так как здесь рассматриваются конечномерные пространства $X = \mathbf{R}^n$, $Y = \mathbf{R}^m$, то $f(x)$ есть просто вектор в $\mathbf{R}^m$ с компонентами $f^i(x)$, $i = 1, \ldots, m$, а $f'(x)$ есть $m \times n$-матрица с компонентами $\partial f^i/\partial x_j$, $i = 1, \ldots, m$, $j = 1, \ldots, n$.

Если теперь $z_0 = (x_0,\ y_0) \in \operatorname{gf} f$, т. е. $y_0 = f(x_0)$, то

$$K_a(x_0,\ y_0) = \{(\bar{x},\ \bar{y})\colon\ \bar{x} \in X,\ \bar{y} = f'(x_0)\bar{x}\} \tag{4.11}$$

есть локальный шатер к $\operatorname{gf} f$. В самом деле, если положить

$$\psi(\bar{x}, \bar{y}) = (\bar{x}, f(x_0 + \bar{x}) - f(x_0)),$$

то согласно формулам (4.10) и (4.11) имеем

$$(x_0, y_0) + \psi(\bar{x}, \bar{y}) = (x_0 + \bar{x}, f(x_0 + \bar{x})) \in \operatorname{gf} f,$$
$$\psi(\bar{x}, \bar{y}) = (\bar{x}, f'(x_0)\bar{x} + r(\bar{x})) = (\bar{x}, \bar{y}) + (0, r(\bar{x})),$$

так что определение локального шатра удовлетворяется тривиальным образом.

Вычислим $K_a^*(x_0, y_0)$. Из соотношения (4.11) следует, что

$$K_a^*(x_0, y_0) =$$
$$= \{(x^*, y^*) : \langle \bar{x}, x^* \rangle + \langle \bar{y}, y^* \rangle \geqslant 0, \quad \bar{x} \in X, \bar{y} = f'(x_0)\bar{x}\} =$$
$$= \{(x^*, y^*) : \langle \bar{x}, x^* \rangle + \langle f'(x_0)\bar{x}, y^* \rangle \geqslant 0, \quad \bar{x} \in X\} =$$
$$= \{(x^*, y^*) : \langle \bar{x}, x^* + (f'(x_0))^* y^* \rangle \geqslant 0, \quad \bar{x} \in X\}.$$

Итак, $(x^*, y^*) \in K_a^*(x_0, y_0)$ только в том случае, если

$$\langle \bar{x}, x^* + f'^*(x_0) y^* \rangle \geqslant 0, \quad \bar{x} \in X, \tag{4.12}$$

где f'^* — оператор (в данном случае — матрица, сопряженная к матрице f'). Легко видеть, что неравенство (4.12) возможно лишь если

$$x^* = -f'^*(x_0) y^*.$$

Таким образом,

$$K_a^*(x_0, y_0) = \{(-f'^*(x_0) y^*, y^*) : y^* \in Y^*\}.$$

В частности, согласно определению 3.1 получаем

$$a^*(y^*; z_0) = \{x^* : (-x^*, y^*) \in K_a^*(z_0)\} = \{f'^*(x_0) y^*\}. \tag{4.13}$$

Итак, в рассматриваемом случае локально сопряженное отображение однозначно. Воспользуемся тем, что согласно примеру 1.5 $\operatorname{con}(M - x_0)$ и $\operatorname{con}(W - f(x_0))$ есть локальные шатры к M и W в точках x_0 и $y_0 = f(x_0)$ соответственно. Применение теоремы 4.3 с учетом формулы (4.13) дает возможность утверждать, что существует число

$\lambda \geqslant 0$ и векторы x^* и y^* такие, что

$$x^* - \lambda f_0'(x_0) = -f'^*(x_0)y^*,$$

$$x^* \in (\operatorname{con}(M - x_0))^*, \quad y^* \in (\operatorname{con}(W - f(x_0)))^*,$$

что и требовалось доказать.

Рассмотрим еще одно применение теоремы 4.3.

Следствие 2. *Пусть x_0 — точка минимума функции $f_0(x)$ при ограничениях*

$$f_i(x) \leqslant 0, \qquad i = 1, \ldots, m, \quad x \in M. \tag{4.14}$$

Предположим, что функции f_i, $i = 0, 1, \ldots, m$, допускают в точке x_0 верхние выпуклые аппроксимации $h_i(\bar{x}, x_0)$, непрерывные по $\bar{x}$, $K_M(x_0)$ — локальный шатер к M в x_0. Тогда существуют такие числа $y^{i*} \geqslant 0$, $i = 0, 1, \ldots, m$, одновременно не равные нулю, и векторы $x_i^* \in \partial f_i(x_0)$, $i = 0, 1, \ldots, m$, что

$$\sum_{i=0}^{m} y^{i*} x_i^* \in K_M^*(x_0), \quad y^{i*} f_i(x_0) = 0, \qquad i = 1, \ldots, m. \tag{4.15}$$

Доказательство. Рассмотрим многозначное отображение

$$a(x) = \{y \in \mathbf{R}^m:\ y^i \geqslant f_i(x), \quad i = 1, \ldots, m\}$$

и положим $W = \{0\}$. Тогда система ограничений (4.14) эквивалентна системе (4.1). Воспользуемся теоремой 4.3. Так как $0 \in a(x_0)$, то в качестве точки y_0 можно взять точку 0. Ясно, что в рассматриваемом случае

$$K_W(0) = \{0\}, \qquad K_W^*(0) = \mathbf{R}^m. \tag{4.16}$$

Обратимся теперь к вычислению $a^*(y^*; z_0)$. Рассмотрим случай, когда $f_i(x_0) < 0$ для всех $i = 1, \ldots, m$. Согласно лемме 3.5

$$f_i(x_0 + \bar{x}) \leqslant f_i(x_0) + h_i(\bar{x}, x_0) + r_i(\bar{x}).$$

Так как $f_i(x_0) < 0$ и $h_i(\bar{x}, x)$ непрерывны по $\bar{x}$, а $r_i(\bar{x})$ стремится к нулю быстрее, чем $\|\bar{x}\|$, то для достаточно малых $\bar{x}$ получаем, что

$$f_i(x_0 + \bar{x}) < 0, \qquad i = 1, \ldots, m.$$

Поэтому

$$K_a(z_0) = \{(\bar{x}, \bar{y}): \ \bar{x} \in X, \quad y \in Y\}$$

есть конус касательных направлений и даже локальный шатер к $\underline{\mathrm{gf}}\, a$ в точке $z_0 = (x_0, 0)$. Действительно, для малых $\bar{x}$ и $\bar{y}$ выполняется неравенство

$$f_i(x_0 + \bar{x}) < \bar{y}^i, \qquad i = 1, \ldots, m,$$

т. е. $(x_0 + \bar{x}, \bar{y}) \in \mathrm{gf}\, a$. Так как

$$K_a^*(z_0) = (X \times Y)^* = \{0\} \times \{0\},$$

то согласно определению 3.1 $a^*(y^*; z_0)$ не пусто лишь в случае, если $y^* = 0$, и при этом содержит единственный вектор $x^* = 0$.

Применение теоремы 4.3 теперь показывает, что найдутся такие числа $\lambda \geqslant 0$, $\quad x_0^* \in \partial f_0(x_0)$ и $x^* \in K_M^*(x_0)$, что

$$x^* - \lambda x_0^* = 0.$$

Положив $y^{0*} = \lambda$ и $y^{i*} = 0$, $i = 1, \ldots, m$, получаем формулы (4.15).

Рассмотрим теперь случай, когда $f_i(x_0) = 0$ для некоторого i. Множество

$$I(z_0) = \{i: \ f_i(x_0) = 0, \quad i = 1, \ldots, m\}$$

не пусто. Кроме того, выполнены предположения теоремы 3.1, при доказательстве которой было установлено, что

$$K_a(z_0) = \{\bar{z} = (\bar{x}, \bar{y}): \bar{y}^i > h_i(\bar{x}, x_0), \quad i \in I(z_0)\}$$

есть локальный шатер к $\mathrm{gf}\, a$ в точке z_0. Там же было показано, что $a^*(y^*; z_0)$ не пусто лишь в случае, если $y^{i*} = 0$, $i \notin I(z_0)$, $\quad y^{i*} \geqslant 0$, $\quad i \in I(z_0)$, и при этом состоит из векторов вида

$$x^* = \sum_{i=1}^{m} y^{i*} x_i^*, \qquad x_i^* \in \partial f_i(x_0). \tag{4.17}$$

В соответствии с теоремой 4.3, условия которой выполнены согласно только что сказанному, найдутся такие векторы $y^* \in K_W^*(0)$, $x^* \in K_M^*(x_0)$ и число $\lambda \geqslant 0$, не равные все нулю одновременно, что

$$x^* - \lambda x_0^* \in a^*(-y^*;\, z_0). \tag{4.18}$$

Так как $K_W^*(0) = \mathbf{R}^m$ (см. формулы (4.16)), то условие $y^* \in K_W^*(0)$ не накладывает никаких ограничений на выбор вектора y^*, и вектор $-y^*$ в соотношении (4.18) можно заменить на y^*:

$$x^* - \lambda x_0^* \in a^*(y^*;\, z_0).$$

Из этой формулы следует, что $a^*(y^*;\ z_0) \neq \varnothing$. Откуда согласно (4.17) получаем, что $y^{i*} = 0$, если $f_i(x_0) < 0$, $y^{i*} \geqslant 0$, если $f_i(x_0) = 0$ и

$$x^* - \lambda x_0^* = \sum_{i=1}^{m} y^{i*} x_i^*, \quad x_i^* \in \partial f_i(x_0), \qquad i = 1, \ldots, m.$$

Полагая $y^{0*} = \lambda$, получаем, что

$$\sum_{i=0}^{m} y^{i*} x_i^* = x^* \in K_M^*(x_0),$$

$$y^{i*} f_i(x_0) = 0, \quad i = 1, \ldots, m.$$

3. Ограничения, задаваемые равенствами и неравенствами. Рассмотрим теперь задачу минимизации функции $f_0(x)$ при ограничениях

$$f_i(x) \leqslant 0, \quad i \in I^-; \quad f_i(x) = 0, \quad i \in I; \quad x \in M. \tag{4.19}$$

Сформулируем предположения, при которых будет решаться поставленная задача. Пусть x_0 — точка минимума.

Предположение 1. Функции $f_i(x)$, $i \in \{0\} \cup I^-$, допускают в точке x_0 верхнюю выпуклую аппроксимацию $h_i(\bar{x}, x_0)$.

Предположение 2. Функции $f_i(x)$, $i \in I$, непрерывно дифференцируемы в некоторой окрестности точки x_0, т. е. обладают непрерывными градиентами $f_i'(x)$.

Предположение 3. В точке x_0 существует выпуклый конус $K_M(x_0)$ касательных направлений к множеству M, обладающий следующим свойством.

Пусть $L = \operatorname{Lin} K_M(x_0)$ — наименьшее линейное подпространство, содержащее $K_M(x_0)$, а $\bar{x}_0 \in K_M(x_0)$, $\bar{x}_0 \neq 0$, и $e_j \in L$, $j=1, \ldots, m$,— фиксированные векторы. Обозначим через L_0 линейное подпространство, содержащее $\bar{x}_0$ и e_j, $j=1, \ldots, m$, т. е. множество всех векторов $\bar{x}$ вида

$$\bar{x} = \gamma \bar{x}_0 + \sum_{j=1}^{m} \delta_j e_j,$$

где γ и δ_j — действительные числа. Тогда существуют такие $\varepsilon_1 > 0$ и $\varepsilon_2 > 0$, что:

а) множество

$$\Omega = \left\{ \bar{x} = \bar{x}_0 + \sum_{j=1}^{m} \delta_j e_j : |\delta_j| \leqslant \varepsilon_1, \quad j = 1, \ldots, m \right\}$$

содержится в $K_M(x_0)$;

б) существует функция $\psi(\bar{x})$, которая определена для всех достаточно малых $\bar{x} \in L_0$, непрерывно дифференцируема в области определения и

$$\psi(\bar{x}) = \bar{x} + r(\bar{x}), \quad x_0 + \psi(\bar{x}) \in M$$

для $\bar{x} \in Q$ и $\|\bar{x}\| < \varepsilon_2$, где $r(\bar{x})$ такова, что $\|\bar{x}\|^{-1} r(\bar{x}) \to 0$ при $\bar{x} \to 0$, $\bar{x} \in L_0$, а множество Q определяется соотношением

$$Q = \operatorname{con} \Omega = \left\{ \bar{x} : \bar{x} = \gamma \left(\bar{x}_0 + \sum_{i=1}^{m} \delta_j e_j \right), \ \gamma \geqslant 0, \ |\delta_j| \leqslant \varepsilon_1 \right\}.$$

Ясно, что из трех предположений наиболее трудно проверяемым является предположение 3, однако оно необходимо для целого ряда задач.

Заметим, что если $K_M(x_0)$ есть шатер множества M, лежащего в конечномерном пространстве $\mathbf{R}^n$, то предположение 3 выполняется, как это нетрудно усмотреть из определения 1.3 шатра, если вместо $K_M(x_0)$ рассмотреть $\operatorname{ri} K_M(x_0)$; последнее обстоятельство не влияет на окончательный результат.

Предположение 4.

$$\left(\bigcap_{i \in \{0\} \cup I^-} \operatorname{dom} h_i(\cdot, x_0) \right) \cap K_M(x_0) \neq \varnothing.$$

Прежде чем сформулировать основную теорему, докажем предварительно несколько лемм.

Лемма 4.1. *Пусть* $l_i(x)$, $i=1, \ldots, m$,— *линейные функции, определенные на подпространстве* $L \subseteq X$. *Тогда либо существуют такие не все равные нулю числа* α_i, *что*

$$\sum_{i=1}^{m} \alpha_i l_i(x) = 0, \qquad x \in L, \tag{4.20}$$

либо существуют такие векторы $e_j \in L$, $j=1, \ldots, m$, *что*

$$l_i(e_j) = \delta_{ij} = \begin{cases} 1, & i=j, \\ 0, & i \neq j. \end{cases} \tag{4.21}$$

Доказательство. Рассмотрим вектор-функцию $l(x) \in \mathbf{R}^m$ с компонентами $l_i(x)$. Пусть $A = l(L)$ — образ подпространства L при отображении l.

Если $A = \mathbf{R}^m$, то любой вектор $y \in \mathbf{R}^m$ можно представить в виде

$$y = l(x), \qquad x \in L.$$

В частности, если $q_j \in \mathbf{R}^m$ — единичные орты $\mathbf{R}^m$, т. е. для компонент q_j^i, $i=1, \ldots, m$, выполняется соотношение $q_j^i = \delta_{ij}$, то существуют такие векторы $e_j \in L$, что

$$q_j = l(e_j),$$

или, в покомпонентной записи, $l_i(e_j) = \delta_{ij}$, так что соотношения (4.21) выполняются. При этом соотношение (4.20) возможно только при нулевых α_i, так как

$$\sum_{i=1}^{m} \alpha_i l_i(e_j) = \alpha_j = 0, \qquad j = 1, \ldots, m.$$

Если же A не совпадает со всем $\mathbf{R}^m$, то A есть некоторое собственное подпространство пространства $\mathbf{R}^m$. Отсюда следует, что A лежит в некоторой гиперплоскости

$$\langle y, y^* \rangle = 0, \quad y \in A, \quad y^* \neq 0.$$

Учитывая, что $y = l(x)$, $x \in L$, получаем

$$\sum_{i=1}^{m} y^{i*} l_i(x) = 0, \qquad x \in L.$$

Полагая $\alpha_i = y^{i*}$, приходим к соотношению (4.20).

Лемма 4.2. *Пусть выполнены предположения* 2 *и* 3. *Тогда либо*

$$\sum_{i\in I} \alpha_i \langle \bar{x}, f_i'(x_0)\rangle = 0, \qquad \sum_{i\in I} |\alpha_i| = 1 \tag{4.22}$$

для всех $\bar{x} \in L$, $L = \operatorname{Lin} K_M(x_0)$, *либо для всякого* $\bar{x}_0 \in K_M(x_0)$, $\bar{x}_0 \neq 0$, *удовлетворяющего условиям*

$$\langle \bar{x}_0, f_i'(x_0)\rangle = 0, \qquad i = 1, \ldots, m, \tag{4.23}$$

существует функция $r_0(\gamma) \in X$, *определенная для всех достаточно малых* $\gamma \geqslant 0$ *и такая, что*

$$x_0 + \gamma\bar{x}_0 + r_0(\gamma) \in M,$$

$$f_i(x_0 + \gamma\bar{x}_0 + r_0(\gamma)) = 0, \qquad i = 1, \ldots, m,$$

и $\gamma^{-1}r_0(\gamma) \to 0$ *при* $\gamma \downarrow 0$.

Доказательство. Согласно предыдущей лемме либо выполнено соотношение (4.22), либо найдутся такие векторы $e_j \in L$, $j \in I$, что

$$\langle e_j, f_i'(x_0)\rangle = \delta_{ij}, \qquad i, j \in I. \tag{4.24}$$

Ясно, что необходимо рассмотреть только вторую возможность.

Возьмем функцию $\psi(\bar{x})$, определенную на подпространстве L_0 (см. предположение 3), и составим систему уравнений

$$g_i(\gamma, \delta) \equiv f_i\left(x_0 + \psi\left(\gamma\bar{x}_0 + \sum_{j\in I} \delta_j e_j\right)\right) = 0, \qquad i \in I. \tag{4.25}$$

Здесь δ — вектор с компонентами δ_j. Согласно предположениям 2 и 3 функции $g_i(\gamma, \delta)$ непрерывно дифференцируемы при всех достаточно малых γ и δ. Вычислим первые производные функций g_i при $\gamma = 0$, $\delta = 0$. Для этого заметим, что в силу свойств функции $\psi(\bar{x})$ справедлива формула

$$\psi_\gamma'(0) \equiv \frac{\partial}{\partial\gamma}\psi\left(\gamma\bar{x}_0 + \sum_{j\in I}\delta_j e_j\right)\Bigg|_{\gamma=0,\delta=0} =$$

$$= \lim_{\gamma\to 0}\frac{\psi(\gamma\bar{x}_0) - \psi(0)}{\gamma} = \lim_{\gamma\to 0}\left(\bar{x}_0 + \frac{r(\gamma\bar{x}_0)}{\gamma}\right) = \bar{x}_0. \tag{4.26}$$

Аналогично получаем

$$\psi'_{\delta_j}(0) = \frac{\partial}{\partial \delta_j} \psi\left(\gamma \bar{x}_0 + \sum_{j \in I} \delta_j e_j\right)\Bigg|_{\gamma=0, \delta=0} = e_j. \quad (4.27)$$

Поэтому по правилу дифференцирования сложной функции имеем

$$\frac{\partial}{\partial \gamma} g_i(\gamma, \delta)\Bigg|_{\gamma=0, \delta=0} = \langle \psi'_\gamma(0), f'_i(x_0) \rangle =$$
$$= \langle \bar{x}_0, f'_i(x_0) \rangle = 0, \qquad i \in I, \quad (4.28)$$

в силу условий (4.23). Аналогично получаем

$$\frac{\partial}{\partial \delta_j} g_i(\gamma, \delta)\Bigg|_{\gamma=0, \delta=0} = \langle \psi'_{\delta_j}(0), f'_i(x_0) \rangle =$$
$$= \langle e_j, f'_i(x_0) \rangle = \delta_{ij}. \quad (4.29)$$

Если воспользоваться теперь замечанием к теореме 1.1, то можно заключить, что для достаточно малых γ определена непрерывная функция $\delta(\gamma)$, причем такая, что выполняется соотношение

$$\lim_{\gamma \to 0} \frac{\delta(\gamma)}{\gamma} = 0$$

и удовлетворяются уравнения (4.25).

В силу предположения 3 имеем

$$\psi\left(\gamma \bar{x}_0 + \sum_{j \in I} \delta_j(\gamma) e_j\right) = \gamma \bar{x}_0 + \sum_{j \in I} \delta_j(\gamma) e_j +$$
$$+ r\left(\gamma \bar{x}_0 + \sum_{j \in I} \delta_j(\gamma) e_j\right) = \gamma \bar{x}_0 + r_0(\gamma),$$

где

$$r_0(\gamma) = \sum_{j \in I} \delta_i(\gamma) e_j + r\left(\gamma \bar{x}_0 + \sum_{j \in I} \delta_j(\gamma) e_j\right).$$

Из свойств функций $\delta(\gamma)$ и $r(\bar{x})$ (они стремятся к нулю быстрее, чем γ и $\bar{x}$) следует, что

$$\frac{r_0(\gamma)}{\gamma} \to 0 \quad (4.30)$$

и при достаточно малых $\gamma \geqslant 0$

$$\gamma \bar{x}_0 + \sum_{j \in I} \delta_j(\gamma) e_j = \gamma \left(x_0 + \sum_{j \in I} \frac{\delta_j(\gamma)}{\gamma} e_j \right) \in \gamma \Omega \subseteq Q, \quad (4.31)$$

так как

$$\left| \frac{\delta_j(\gamma)}{\gamma} \right| \leqslant \varepsilon_1, \qquad j \in I.$$

Согласно предположению 3 из включения (4.31) вытекает, что

$$x_0 + \psi\left(\gamma \bar{x}_0 + \sum_{j \in I} \delta_j(\gamma) e_j \right) \in M,$$

при малом γ, т. е.

$$x_0 + \gamma \bar{x}_0 + r_0(\gamma) \in M. \quad (4.32)$$

Учитывая, что система (4.25) может быть переписана в виде

$$f_i(x_0 + \gamma \bar{x}_0 + r_0(\gamma)) = 0, \quad i \in I, \quad (4.33)$$

из соотношений (4.32) и (4.30) получаем все утверждения леммы.

Теперь можно приступить к формулировке и доказательству основной теоремы.

Теорема 4.4. *Пусть для точки x_0, являющейся точкой минимума функции $f_0(x)$ при ограничениях (4.19), выполнены предположения 1—4. Тогда существуют такие числа y^{i*}, $i \in \{0\} \cup I^- \cup I$, что*

$$\sum_{i \in \{0\} \cup I^-} y^{i*} h_i(\bar{x}, x_0) + \sum_{i \in I} y^{i*} \langle \bar{x}, f_i'(x_0) \rangle \geqslant 0 \quad (4.34)$$

для всех

$$\bar{x} \in K_M(x_0) \cap \left(\bigcap_{i \in \{0\} \cup I^-} \operatorname{dom} h_i(\cdot, x_0) \right).$$

При этом $y^{i*} \geqslant 0$ для $i \in \{0\} \cup I^-$ и $y^{i*} = 0$ для $i \in I_0^-$, где

$$I_0^- = \{ i \in I^- : f_i(x_0) < 0 \}.$$

Доказательство. Если соотношение (4.22) выполняется, то утверждение теоремы выполняется тривиальным образом.

Достаточно для $i \in I$ положить $y^{i*} = \alpha_i$, взяв α_i из соотношения (4.22), а все остальные y^{i*} выбрать равными нулю.

Предположим теперь, что соотношение (4.22) не выполняется. Введем следующие обозначения:

$$J^- = \{0\} \cup (I^- \setminus I_0^-), \quad J = J^- \cup I,$$
$$K = K_M(x_0) \cap \Big(\bigcap_{i \in \{0\} \cup I^-} \operatorname{dom} h_i(\cdot, x_0) \Big).$$

Пусть b — число индексов в J. Определим множество P следующим образом: $\bar{y} \in P$, $P \subseteq \mathbf{R}^b$, тогда и только тогда, когда существует такой вектор $\bar{x} \in K$, что

$$\bar{y}^i > h_i(\bar{x}, x_0), \quad i \in J^-, \qquad \bar{y}^i = \langle \bar{x}, f_i'(x_0) \rangle, \quad i \in I.$$

Из выпуклости функций h_i легко следует, что P — выпуклое множество. В силу предположения 4 оно не пусто.

Рассмотрим множество

$$N = \{\bar{y} \in \mathbf{R}^b \colon \bar{y}^i \leqslant 0, \quad i \in J^-, \quad \bar{y}^i = 0, \quad i \in I\}$$

и покажем, что P и N не пересекаются. Допустим противное, тогда найдутся вектор $\bar{y}_0 \in P \cap N$ и вектор $\bar{x}_0 \in K$, такие, что

$$\begin{aligned} &h_i(\bar{x}_0, x_0) < \bar{y}^i \leqslant 0, \qquad i \in J^-; \\ &\langle \bar{x}_0\, f_i'(x_0) \rangle = 0, \qquad i \in I. \end{aligned} \tag{4.35}$$

Так как $\bar{x}_0 \in K$, а $K \subseteq K_M(x_0)$, то $\bar{x}_0 \in K_M(x_0)$. Из первого соотношения (4.35) вытекает, что $\bar{x}_0 \neq 0$, ибо $h_i(0, x_0) = 0$.

Из второго соотношения (4.35) и леммы (4.2) следует существование такой функции $r_0(\gamma)$, что $\gamma^{-1} r_0(\gamma) \to 0$ при $\gamma \downarrow 0$, для которой выполнены соотношения (4.32), (4.33).

В силу предположения 1 для $i \in J^-$ получаем, что

$$\limsup_{\gamma \downarrow 0} \frac{f_i(x_0 + \gamma \bar{x}_0 + r_0(\gamma)) - f_i(x_0)}{\gamma} \leqslant h_i(\bar{x}_0, x_0) < 0,$$

т. е. для всех достаточно малых $\gamma > 0$ выполняется

$$f_i(x_0 + \gamma \bar{x}_0 + r_0(\gamma)) \leqslant f_i(x_0) + \frac{1}{2} \gamma h_i(\bar{x}_0, x_0), \quad i \in J^-.$$

Отсюда, так как $h_i(\bar{x}_0, x_0) < 0$ для $i \in J^-$, при достаточно малых $\gamma > 0$ получаем следующие соотношения:

$$f_0(x_0 + \gamma\bar{x}_0 + r_0(\gamma)) < f_0(x_0), \tag{4.36}$$

$$f_i\left(x_0 + \gamma\bar{x}_0 + r_0(\gamma)\right) < 0, \qquad i \in \left(I^- \setminus I_0^-\right). \tag{4.37}$$

Величины $h_i(\bar{x}_0, x_0), i \in I_0^-$, конечны, поскольку $\bar{x}_0 \in K$, а $K \subseteq \operatorname{dom} h_i(\cdot, \ x_0)$. Поэтому

$$\limsup_{\gamma \downarrow 0} \frac{f_i\left(x_0 + \gamma\bar{x}_0 + r_0(\gamma)\right) - f_i(x_0)}{\gamma} \leqslant h_i\left(\bar{x}_0, x_0\right).$$

Откуда вытекает, что неравенство

$$f_i(x_0 + \gamma\bar{x}_0 + r_0(\gamma)) \leqslant f_i(x_0) + \gamma(h_i(\bar{x}_0, \ x_0) + \varepsilon) < 0, \quad i \in I_0, \tag{4.38}$$

выполняется для любого $\varepsilon > 0$ и при достаточно малых $\gamma > 0$, так как $f_i(x_0) < 0$ для $i \in I_0^-$.

Сопоставляя соотношения (4.32), (4.33), (4.36)—(4.38), видим, что найдутся точки $x = x_0 + \gamma\bar{x}_0 + r_0(\gamma)$, отличные от x_0, в которых значения $f_0(x)$ меньше, а все ограничения (4.19) выполняются. Это противоречит тому, что x_0 — точка минимума функции $f_0(x)$. Таким образом, введенные выше множества P и N не пересекаются.

На основании теоремы отделимости (теорема I.2.3) существует такой вектор $y^* \in \mathbf{R}^b$, что

$$\langle \bar{y}, y^* \rangle \geqslant \langle \bar{y}_1, y^* \rangle \tag{4.39}$$

для всех $\bar{y} \in P$, $\bar{y}_1 \in N$, или, в компонентной записи,

$$\sum_{i \in J^-} \bar{y}^i y^{i*} + \sum_{i \in I} \bar{y}^i y^{i*} \geqslant \sum_{i \in J^-} \bar{y}_1^i y^{i*} + \sum_{i \in I} \bar{y}_1^i y^{i*}, \qquad \bar{y} \in P, \quad \bar{y}_1 \in N. \tag{4.40}$$

Из определения множества P следует, что величины $\bar{y}^i$, $i \in J^-$, могут неограниченно расти при фиксированном $\bar{x}$. Поэтому $y^{i*} \geqslant 0$, $i \in J^-$, так как противоположное неравенство приводило бы к противоречию с (4.40) при $\bar{y}^i \to +\infty$. Итак,

$$y^{i*} \geqslant 0, \qquad i \in \{0\} \cup \left(I^- \setminus I_0^-\right). \tag{4.41}$$

Пусть $\bar{x} \in K$. Положим в неравенстве (4.40) $\bar{y}_1 = 0$. Тогда устремив $\bar{y}^i$ к $h_i(\bar{x}, x_0)$ для $i \in J^-$ и положив для $i \in I$ $\bar{y}^i = \langle \bar{x}, f_i'(x_0) \rangle$, получим

$$\sum_{i \in J^-} y^{i*} h_i(\bar{x}, x_0) + \sum_{i \in I} y^{i*} \langle \bar{x}, f_i'(x_0) \rangle \geqslant 0$$

для $\bar{x} \in K$. Выбрав $y^{i*} = 0$ для $i \in I_0^-$, переходим к соотношениям

$$\sum_{i \in \{0\} \cup I^-} y^{i*} h_i(\bar{x}, x_0) + \sum_{i \in I} y^{i*} \langle \bar{x}, f_i'(x_0) \rangle \geqslant 0,$$

$$\bar{x} \in K.$$

$$y^{i*} \geqslant 0, \qquad i \in \{0\} \cup I^-;$$

$$y^{i*} = 0, \qquad i \in I_0^-,$$

что и требовалось доказать.

Теорема 4.5. *Пусть выполнены предположения 1—3 и существует точка* $\bar{x}_1 \in K_M(x_0)$, *в которой функции* $h_i(\bar{x}, x_0)$, $i \in \{0\} \cup I^-$, *непрерывны. Тогда существуют не все равные нулю числа* y^{i*}, $i \in \{0\} \cup I^- \cup I$, *и векторы* $x_i^* \in \partial f_i(x_0)$, $i \in \{0\} \cup I^-$, $x_{i0}^* \in (\mathrm{dom}\, h_i(\cdot, x_0))^*$, $x^* \in \in K_M^*(x_0)$, *такие, что*

$$\sum_{i \in \{0\} \cup I^-} y^{i*} x_i^* + \sum_{i \in I} y^{i*} f_i'(x_0) = x^* + \sum_{i \in \{0\} \cup I^-} x_{i0}^*,$$

$$y^{i*} \geqslant 0, \quad i \in \{0\} \cup I^-; \quad y^{i*} = 0, \quad i \in I_0^-. \tag{4.42}$$

Доказательство. В силу предположений выполнены все условия теоремы 4.4. Поэтому найдутся такие числа y^{i*}, что $y^{i*} \geqslant 0$, $i \in \{0\} \cup I^-$; $y^{i*} = 0$, $i \in I_0^-$ и будет выполнено неравенство (4.34). Обозначим функцию в левой части неравенства (4.34) через $h(\bar{x})$. Ясно, что

$$\mathrm{dom}\, h = \bigcap_{i \in \{0\} \cup I^-} \mathrm{dom}\, h_i(\cdot, x_0),$$

и в силу условий доказываемой теоремы выполнено включение

$$\bar{x}_1 \in \mathrm{dom}\, h, \tag{4.43}$$

функция $h(\bar{x})$ непрерывна в точке $\bar{x}_1$.

Неравенство (4.34) показывает, что $\bar{x}=0$ есть точка минимума функции $h(\bar{x})$ на конусе

$$K = K_M(x_0) \cap \left(\bigcap_{i\in\{0\}\cup I^-} \operatorname{dom} h_i(\cdot, x_0)\right).$$

Согласно теореме IV.2.2

$$\partial h(0) \cap K^* \neq \varnothing, \tag{4.44}$$

а в силу теоремы II.3.8

$$\partial h(0) = \sum_{i\in\{0\}\cup I^-} y^{i*}\partial f_i(x) + \sum_{i\in I} y^{i*} f_i'(x_0), \tag{4.45}$$

так как согласно сделанным предположениям функции h_i непрерывны в точке $\bar{x}_1$ и

$$\bar{x}_1 \in K_M(x_0) \cap \left(\bigcap_{i\in\{0\}\cup I^-} \operatorname{int}\operatorname{dom} h_i(\cdot, x_0)\right),$$

то по теореме I.3.3 получаем, что

$$K^* = K_M^*(x_0) + \sum_{i\in\{0\}\cup I^-} (\operatorname{dom} h_i(\cdot, x_0))^*. \tag{4.46}$$

Из соотношений (4.44)—(4.46) следует равенство (4.42).

Следствие 1. *Если функция $h_i(\bar{x}, x_0)$ непрерывна по $\bar{x}$ во всем пространстве X, то в предположениях 1—3 верна теорема 4.5, а равенство (4.42) может быть записано в виде*

$$\sum_{i\in\{0\}\cup I^-} y^{i*}x_i^* + \sum_{i\in I} y^{i*} f_i'(x_0) = x^*.$$

В самом деле, если $h_i(\bar{x}, x_0)$ непрерывна во всем пространстве, то

$$\operatorname{dom} h_i(\cdot, x_0) = X, \quad (\operatorname{dom} h_i(\cdot, x_0))^* = \{0\},$$

так что из включения $x_{i0}^* \in (\operatorname{dom} h_i(\cdot, x_0))^*$ следует, что $x_{i0}^* = 0$.

Глава VI

НЕОБХОДИМЫЕ УСЛОВИЯ В ЗАДАЧАХ ОПТИМАЛЬНОГО УПРАВЛЕНИЯ

Задачи оптимального управления и вариационного исчисления — традиционная область применения теории необходимых условий экстремума. При этом, как правило, приходится рассматривать необходимые условия для задач, определенных в достаточно общих функциональных пространствах, так как в задачах оптимального управления минимизируемая функция зависит от аргумента, который сам является функцией времени.

Большинство результатов предыдущей главы о необходимых условиях экстремума после соответствующей переформулировки могут быть перенесены на общие функциональные пространства. После этого они непосредственно применимы к задачам вариационного исчисления и оптимального управления. Это обычный путь. В этой книге до сих пор рассматривались конечномерные пространства, и использование в последней главе сложных функциональных пространств означало бы отклонение от общего плана. Поэтому при исследовании задач оптимального управления в этой главе используются методы, основанные на переходе от исходной задачи к последовательности аппроксимирующих ее конечномерных задач. При этом оказывается удобным трактовать задачу оптимального управления как задачу оптимизации на множестве траекторий некоторого дифференциального включения. Понятие дифференциального включения, которое будет определено ниже, позволяет охватить многие из рассмотренных в литературе задач оптимального управления единым методом решения. В то же время такая общая постановка задачи, давая возможность получать новые результаты, не всегда позволяет получить уже известные во всей полноте, так как общность постановки неизбежно ведет к тому, что игнорируются некоторые детали и особенности конкретной задачи.

Как уже сказано выше, аргументом в задачах оптимального управления являются функции. Поэтому совсем избежать понятий, связанных со свойствами некоторых простейших функциональных пространств, невозможно. Мы предполагаем, что читатель знает, что такое абсолютно непрерывная функция и измеримая функция, что производная от абсолютно непрерывной функции существует почти всюду и абсолютно непрерывная функция совпадает с интегралом от своей производной и тому подобное. Все эти и другие используемые ниже свойства можно найти в учебниках.

§ 1. Дифференциальные включения

Рассмотрим конечномерные пространства X и Y. Всюду в дальнейшем предполагается, что $X = Y = \mathbf{R}^n$. Пусть задано многозначное отображение $a(x)$, $\operatorname{gf} a \subseteq Z$, $Z = X \times Y$.

В дальнейшем изложении фиксируем стандартный промежуток изменения одномерного аргумента t. В качестве такого промежутка выберем отрезок $[0,1]$. Пусть для $t \in [0,1]$ определена абсолютно непрерывная функция $x(t) \in \mathbf{R}^n$. Такая функция имеет производную почти всюду и может случиться так, что при этом выполняется соотношение

$$\frac{d}{dt} x(t) \in a(x(t)) \tag{1.1}$$

для почти всех t. Соотношение (1.1) называется *дифференциальным включением*, а при указанных условиях функция $x(t)$ называется *решением* этого *дифференциального включения*.

Ясно, что дифференциальное включение есть обобщение обычных дифференциальных уравнений. Если $a(x)$ — однозначное отображение и единственным элементом, принадлежащим множеству $a(x)$, является $f(x)$, то включение (1.1) переходит в обычное дифференциальное уравнение

$$\frac{d}{dt} x(t) = f(x).$$

Если $f(x, u) \in \mathbf{R}^n$ — вектор-функция, определенная для $x \in \mathbf{R}^n$ и $u \in U$, где U — некоторое подмножество из $\mathbf{R}^s$, то можно определить многозначное отображение $a(x)$ следующим образом:

$$a(x) = \{y\colon\ y = f(x, u),\ u \in U\} = f(x, U). \tag{1.2}$$

При таком выборе $a(x)$ включение (1.1) переходит в соотношение

$$\frac{d}{dt} x(t) \in f(x(t), U),$$

или

$$\frac{d}{dt} x(t) = f(x(t), u(t)), \qquad u(t) \in U. \tag{1.3}$$

Читатель, знакомый с теорией оптимального управления, сразу увидит в уравнении (1.3) обычное уравнение динамики управляемого объекта и заметит, что любое решение уравнения (1.3) при некотором *управлении* $u(t) \in U$ одновременно является решением и дифференциального включения (1.1). Можно показать, что при достаточно естественных предположениях решение включения (1.1), когда $a(x)$ определено соотношением (1.2), одновременно является решением уравнения (1.3) при некотором управлении $u(t) \in U$.

Из приведенного примера ясно, что дифференциальное включение (1.1) может иметь не единственное решение. Возникает вопрос: при каких условиях существует хотя бы одно решение дифференциального включения и каковы семейства решений? Ответ на этот вопрос дает следующая теорема. Заметим, что она является далеко не самой общей теоремой такого рода, но она вполне достаточна для наших нужд.

Теорема 1.1. *Если $a(x)$ — выпуклозначное замкнутое ограниченное отображение, полунепрерывное сверху по x, и* $\mathrm{dom}\, a$ *содержит шар радиуса*

$$r = (1 + \|x_0\|) e^c - 1,$$

где c — константа такая, что

$$\|a(x)\| \leqslant c(1 + \|x\|),$$

то существует решение $x(t)$ дифференциального включения (1.1), удовлетворяющее неравенству

$$\|x(t)\| \leqslant (1 + \|x_0\|) e^{ct} - 1, \qquad x(0) = x_0.$$

При этом решение $x(t)$ определено на всем отрезке $[0, 1]$ *и удовлетворяет на этом отрезке условию Липшица.*

Доказательство. Разобьем интервал $[0, 1]$ на 2^m частей (m — натуральное число) и положим $\delta = 2^{-m}$. Запишем вместо соотношения (1.1) *разностное включение*

$$x_\delta(t+\delta) \in x_\delta(t) + \delta a(x_\delta(t)),\quad t = 0,\ \delta,\ 2\delta,\ \ldots,\ (2^m-1)\delta. \tag{1.4}$$

Выберем $x_\delta(0) = x_0$ и будем определять $x_\delta(t)$, $t = 0, \delta, \ldots$, шаг за шагом, выбирая каждый раз в качестве $x_\delta(t+\delta)$ произвольный элемент множества $x_\delta(t) + \delta a(x_\delta(t))$.

Из включения (1.4) следует, что

$$\|x_\delta(t+\delta)\| \leqslant \|x_\delta(t)\| + \delta\|a(x_\delta(t))\| \leqslant \|x_\delta(t)\| + \delta c(1 + \|x_\delta(t)\|),$$

т. е.

$$\|x_\delta(t+\delta)\| \leqslant (1+\delta c)\|x_\delta(t)\| + c\delta. \tag{1.5}$$

По индукции, полагая $t = 0, \delta, \ldots$, из неравенства (1.5) нетрудно получить

$$\|x_\delta(t)\| \leqslant (1+\delta c)^{t/\delta}(1 + \|x_0\|) - 1. \tag{1.6}$$

Из элементарных курсов анализа известно, что

$$(1+\delta c)^{t/\delta} \leqslant e^{ct}.$$

Поэтому из неравенства (1.6) получаем

$$\|x_\delta(t)\| \leqslant e^{ct}(1 + \|x_0\|) - 1. \tag{1.7}$$

Из оценки (1.7) и предположений теоремы вытекает, что $x_\delta(t)$ при всех $t = 0, \delta, \ldots, (2^m-1)\delta$ не выходит из области определения отображения a, так что включение (1.4) всегда имеет решение.

Пусть $x_\delta(t)$ удовлетворяет соотношению (1.4). Заметим, что тогда $x_\delta(t)$ определено только для $t = k\delta$, $k = 0, 1, \ldots$ $\ldots, (2^m-1)$. Доопределим $x_\delta(t)$ для всех $t \in [0, 1]$, положив для $t \in [k\delta, (k+1)\delta]$

$$x_\delta(t) = x_\delta(k\delta) + (t - k\delta)\frac{x_\delta((k+1)\delta) - x_\delta(k\delta)}{\delta},$$

т. е. построим линейную интерполяцию. Функция $x_\delta(t)$ удовлетворяет условию Липшица с константой, не зависящей от δ. В самом деле, в силу неравенства (1.7) можно записать

$$\|x_\delta(t)\| \leqslant e^{c}(1 + \|x_0\|) - 1 = r. \tag{1.8}$$

Так как $a(x)$ — ограниченное отображение, то

$$\|a(x_\delta(t))\| \leqslant c(1 + \|x_\delta(t)\|) \leqslant c(1 + r) = L;$$

поэтому

$$\|x_\delta(t + \delta) - x_\delta(t)\| \leqslant \delta L$$

для всех $t = 0,\ \delta,\ \ldots$ Отсюда уже нетрудно получить, что

$$\|x_\delta(t_1) - x_\delta(t_2)\| \leqslant |t_1 - t_2| L$$

для любых $t_1,\ t_2 \in [0,1]$.

Устремим теперь $\delta = 2^{-m}$ к нулю и рассмотрим любую последовательность $x_\delta(\cdot)$, удовлетворяющую включению (1.4). Так как все эти функции равномерно ограничены одной и той же константой L, то к ним применима лемма Арцела, из которой следует, что из последовательности функций $x_\delta(\cdot)$ можно выбрать равномерно сходящуюся подпоследовательность. Без ограничения общности можно предполагать, что сама последовательность $x_\delta(\cdot)$ сходится к некоторой функции $x_0(\cdot)$, причем

$$\gamma(\delta) = \max_{0 \leqslant t \leqslant 1} \|x_\delta(t) - x_0(t)\| \to 0. \tag{1.9}$$

Функция $x_0(t)$ удовлетворяет условию Липшица с константой L. В самом деле,

$$\begin{aligned}\|x_0(t_1) - x_0(t_2)\| \leqslant \|x_0(t_1) - x_\delta(t_1)\| + \\ + \|x_\delta(t_1) - x_\delta(t_2)\| + \|x_\delta(t_2) - x_0(t_2)\| \leqslant \\ \leqslant \gamma(\delta) + |t_1 - t_2| L + \gamma(\delta).\end{aligned}$$

Устремляя δ к нулю, получаем требуемый результат.

Итак, $x_0(t)$ удовлетворяет условию Липщица и, значит почти всюду дифференцируема. Покажем, что $x_0(t)$ удовлетворяет выражению (1.1). Действительно, пусть $t_0 \in (0,1)$ — точка, в которой существует производная $x_0(t)$. Тогда для $\varepsilon > 0$ найдется Δ такое, что

$$\left\| \frac{d}{dt} x_0(t_0) - \frac{x_0(t_2) - x_0(t_1)}{t_2 - t_1} \right\| < \varepsilon, \tag{1.10}$$

как только $|t_1 - t_2| < \Delta$. Так как

$$\begin{aligned}\|x_\delta(t) - x_0(t_0)\| \leqslant \|x_\delta(t) - x_\delta(t_0)\| + \\ + \|x_\delta(t_0) - x_0(t_0)\| \leqslant L|t - t_0| + \gamma(\delta),\end{aligned}$$

то в силу полунепрерывности сверху отображения a справедливо включение

$$a(x_\delta(t)) \subseteq a(x_0(t_0)) + \varepsilon B \tag{1.11}$$

(напомним, что B — единичный шар в $\mathbf{R}^n$), как только разность $(t - t_0)$ и δ достаточно малы. Допустим, что включение (1.11) выполняется, если выполнены следующие условия:

$$|t - t_0| < \Delta_1, \quad \delta < \delta_0. \tag{1.12}$$

Пусть t_1 и t_2 — фиксированные точки, имеющие вид $t_1 = k_1\delta_1$, $t_2 = k_2\delta_1$, $\delta_1 < \delta_0$, причем $t_1 < t_0 < t_2$, $|t_1 - t_2| < \Delta_1$. Так как $\delta = 2^{-m}$, то при $\delta \leqslant \delta_1$ точки t_1 и t_2 будут входить в разбиение отрезка и будут иметь аналогичный вид при всяком δ. Выберем δ настолько малым, чтобы были справедливы неравенства

$$\left\| \frac{x_0(t_1) - x_\delta(t_1)}{t_2 - t_1} \right\| < \varepsilon, \qquad \left\| \frac{x_0(t_2) - x_\delta(t_2)}{t_2 - t_1} \right\| < \varepsilon. \tag{1.13}$$

Теперь

$$\frac{x_\delta(t_2) - x_\delta(t_1)}{t_2 - t_1} = \sum_{t = t_1, t_1+\delta, \ldots, t_2-\delta} \frac{\delta}{t_2 - t_1} \, \frac{x_\delta(t+\delta) - x_\delta(t)}{\delta}.$$

Тогда в силу включений (1.4) и (1.11) выполняется соотношение

$$\frac{x_\delta(t_2) - x_\delta(t_1)}{t_2 - t_1} \in \sum_{t = t_1, t_1+\delta, \ldots, t_2-\delta} \frac{\delta}{t_2 - t_1} (a(x_0(t_0)) + \varepsilon B),$$

или, так как $a(x_0(t_0))$ — выпуклое множество,

$$\frac{x_\delta(t_2) - x_\delta(t_1)}{t_2 - t_1} \in (a(x_0(t_0)) + \varepsilon B). \tag{1.14}$$

Из очевидного соотношения

$$\frac{x_0(t_2) - x_0(t_1)}{t_2 - t_1} = \\ = \frac{x_0(t_2) - x_\delta(t_2)}{t_2 - t_1} + \frac{x_\delta(t_2) - x_\delta(t_1)}{t_2 - t_1} + \frac{x_\delta(t_1) - x_0(t_1)}{t_2 - t_1} \tag{1.15}$$

и формул (1.10), (1.13), (1.14) вытекает, что

$$\frac{d}{dt} x_0(t_0) \in a(x_0(t_0)) + 4\varepsilon B.$$

Так как $\varepsilon > 0$ произвольно, то из последнего включения получаем

$$\frac{d}{dt} x_0(t_0) \in a(x_0(t_0)),$$

откуда вытекает, что $x_0(t)$ есть решение включения (1.1).

Выше было доказано, что $x_0(t)$ удовлетворяет условию Липшица. Наконец, предельный переход в неравенстве (1.7) дает оценку

$$\|x_0(t)\| \leqslant e^{ct}(1 + \|x_0\|) - 1.$$

Теорема 1.1 показывает, что некоторые траектории дифференциального включения (1.1) могут быть аппроксимированы решением разностного включения 1.4. Следующая теорема утверждает, что при дополнительных предположениях любое решение (1.1) допускает такую аппроксимацию.

Определение 1.1. Если $x(t) \in \mathbf{R}^n$, $t \in [0,1]$,— произвольная кривая, то ее *ε-трубкой* называется множество точек x таких, что

$$\|x - x(t)\| \leqslant \varepsilon$$

при некотором $t \in [0,1]$.

Теорема 1.2. *Пусть a — выпуклозначное замкнутое отображение, $a(x)$ — ограничено и $x(t)$, $t \in [0,1]$,— решение дифференциального включения (1.1). Пусть, далее, отображение a удовлетворяет условию Липшица в некоторой ε-трубке траектории $x(t)$. Тогда существует такое решение $x_\delta(t)$ разностного включения (1.4), что*

$$\|x(t) - x_\delta(t)\| \leqslant c_1\delta, \qquad t = 0, \delta, \ldots, 1,$$

и константа c_1 не зависит от δ.

Доказательство. Положим

$$u(t) = \frac{d}{dt} x(t) \in a(x(t)).$$

Тогда

$$x(t+\delta) - x(t) = \int_t^{t+\delta} u(t)\,dt.$$

Поэтому для любого $x^* \in \mathbf{R}^n$ выполняется соотношение

$$\langle x(t+\delta) - x(t), x^*\rangle = \\ = \int_t^{t+\delta} \langle u(\tau), x^*\rangle\, d\tau \geqslant \int_t^{t+\delta} W_a(x(\tau), x^*)\, d\tau. \quad (1.16)$$

Так как отображение a удовлетворяет условию Липшица, то по лемме V.3.2 имеем

$$|W_a(x_1, x^*) - W_a(x_2, x^*)| \leqslant L\|x_1 - x_2\|\|x^*\| \quad (1.17)$$

для x_1 и x_2 из ε-трубки траектории $x(t)$.

По предположению a удовлетворяет условию Липшица в ε-трубке и так как $a(x_0)$ ограничено, то все множества $a(x(t))$ ограничены в совокупности. Значит, для $t_1 < t_2$ выполняется неравенство

$$\|x(t_1) - x(t_2)\| = \\ = \left\| \int_{t_1}^{t_2} u(t)\, dt \right\| \leqslant \int_{t_1}^{t_2} \|u(t)\|\, dt \leqslant L_2 |t_1 - t_2|, \quad (1.18)$$

где $L_1 = \sup\limits_{0 \leqslant t \leqslant 1} \|a(x(t))\|$, т. е. $x(t)$ удовлетворяет условию Липшица с константой L_1.

Из соотношений (1.16)—(1.18) вытекают неравенства

$$\langle x(t+\delta) - x(t), x^*\rangle \geqslant \\ \geqslant \int_t^{t+\delta} [W_a(x(\tau), x^*) - W_a(x(t), x^*)]\, d\tau + \delta W_a(x(t), x^*) \geqslant \\ \geqslant \delta W_a(x(t), x^*) - L\|x^*\| \int_t^{t+\delta} \|x(\tau) - x(t)\|\, d\tau \geqslant \\ \geqslant \delta W_a(x(t), x^*) - LL_1\|x^*\| \int_t^{t+\delta} (\tau - t)\, d\tau = \\ = \delta W_a(x(t), x^*) - \frac{1}{2} LL_1\|x^*\|\delta^2. \quad (1.19)$$

Согласно формуле V.3.1

$$W_a(x, y^*) = \inf_y \{\langle y, y^*\rangle :\ y \in a(x)\},$$

и поэтому легко проверить, что

$$\delta W_a(x(t), x^*) - \frac{LL_1}{2}\|x^*\|\delta^2 = \\ = \inf_y \left\{\langle y, x^*\rangle:\ y \in \delta a(x(t)) + \frac{LL_1\delta^2}{2} B\right\},$$

где, как обычно, B — единичный шар в $\mathbf{R}^n$. Аналогично доказательству теоремы I.2.7 в силу замкнутости $a(x(t))$, можно показать, что выполнение неравенства (1.19) при всех x^* означает, что

$$x(t+\delta) - x(t) \in \delta a(x(t)) + L_2\delta^2 B,$$

или

$$x(t+\delta) \in x(t) + \delta a(x(t)) + L_2\delta^2 B, \tag{1.20}$$

где $L_2 = \frac{1}{2}LL_1$.

Построим теперь траекторию $x_\delta(t)$. Пусть

$$\delta < \frac{2\varepsilon}{L^2 e^L}.$$

Положим $x_\delta(0) = x_0$. Дальнейшие точки строятся рекуррентно. Если точки $x_\delta(t)$, $t = 0, \delta, \ldots, k\delta$, уже построены, то в качестве $x_\delta((k+1)\delta)$ выбирается точка из множества $x_\delta(k\delta) + \delta a(x_\delta(k\delta))$, ближайшая к $x((k+1)\delta)$. При этом, в силу включения (1.20) выполняются соотношения

$$\begin{gathered} x((k+1)\delta) \in x(k\delta) + \delta a(x(k\delta)) + L_2\delta^2 B, \\ x_\delta((k+1)\delta) \in x_\delta(k\delta) + \delta a(x_\delta(k\delta)). \end{gathered} \tag{1.21}$$

Так как a удовлетворяет условию Липшица в ε-трубке решения $x(t)$, а $x(k\delta)$, как будет показано ниже, не выходит за ε-трубку, то

$$a(x(k\delta)) \subseteq a(x_\delta(k\delta)) + L\|x(k\delta) - x_\delta(k\delta)\|B.$$

Из соотношений (1.21) и выбора $x_\delta((k+1)\delta)$ следует, что

$$\Delta_{k+1} \equiv \|x((k+1)\delta) - x_\delta((k+1)\delta)\| \leqslant \\ \leqslant \|x(k\delta) - x_\delta(k\delta)\| + L\delta\|x(k\delta) - x_\delta(k\delta)\| + L_2\delta^2 = \\ = (1 + L\delta)\Delta_k - L_2\delta^2.$$

Итак,

$$\Delta_{k+1} \leqslant (1 + L\delta)\Delta_k + L_2\delta^2, \qquad k = 0, 1, \ldots, 2^m - 1;$$

откуда с учетом того, что $\Delta_0 = 0$, получаем неравенство

$$\Delta_k = \|x(k\delta) - x_\delta(k\delta)\| \leqslant$$

$$\leqslant \frac{1}{2} L_1 (1 + L\delta)^k \delta \leqslant \frac{1}{2} L_1 (1 + L\delta)^{1/\delta} \delta. \qquad (1.22)$$

Так как $(1 + L\delta)^{1/\delta} \leqslant e^L$, то

$$\Delta_k \leqslant \frac{1}{2} L_1 e^L \delta < \varepsilon,$$

и построенная траектория не выходит за ε-трубки решения $x(t)$, так что приведенные рассуждения были правомочны.

Если обозначить $c_1 = \frac{1}{2} L_1 e^L$, то неравенство (1.22) показывает, что

$$\Delta_k \leqslant c_1\delta, \qquad k = 0, 1, \ldots, 2^m,$$

или

$$\|x(t) - x_\delta(t)\| \leqslant c_1\delta, \qquad t = 0, \delta, \ldots, 1,$$

что и требовалось доказать.

§ 2. Задача оптимального управления с дискретным временем

В § 6 главы IV была рассмотрена задача оптимального управления моделью экономической динамики. В этом параграфе будет рассмотрена аналогичная задача, однако уже без предположения о выпуклости входящих в задачу функций и множеств. Поскольку рассуждения, используемые ниже, в существенном аналогичны рассуждениям § 6 главы IV, то читателю рекомендуется еще раз ознакомиться с материалом этого параграфа.

Пусть $X = Y = \mathbf{R}^n$ и a — некоторое многозначное отображение. Рассмотрим систему включений

$$x_{t+1} \in a(x_t), \qquad t = 0, 1, \ldots, T - 1, \qquad (2.1)$$

где T — фиксированное целое число. Всякую последовательность векторов $\{x_t\}_{t=0,\ldots,T}$, удовлетворяющих включению (2.1), назовем *траекторией*. Среди всех траекторий требуется выбрать такую, которая начинается на множестве N и оканчивается на множестве M, где N и M — заданные множества, т. е. $x_0 \in N$, $x_T \in M$, и минимизирует сумму

$$\sum_{t=0}^{T} g(x_t, t). \tag{2.2}$$

Поскольку в этой книге проблемы существования решения не рассматриваются, будем предполагать, что оптимальная траектория $\{\tilde{x}_t\}_{t=0,\ldots,T}$ существует и наша задача состоит в том, чтобы охарактеризовать ее, т. е. написать необходимые условия минимума.

Сформулируем явно предположения, при которых будет решаться задача.

Основное предположение.

1. Отображение a таково, что конусы касательных направлений $K_a(\tilde{x}_t, \tilde{x}_{t+1})$, $t = 0, 1, \ldots, T-1$, являются локальными шатрами.

2. Конусы касательных направлений $K_N(\tilde{x}_0)$ и $K_M(\tilde{x}_T)$ являются локальными шатрами.

3. Функции $g(x, t)$, $t = 0, \ldots, T$, в точках $\tilde{x}_t$ допускают верхнюю выпуклую аппроксимацию $h_t(x, \tilde{x}_t)$, которая непрерывна по $\bar{x}$. Таким образом, определены субдифференциалы

$$\partial g(\tilde{x}_t, t) = \partial h_t(0, \tilde{x}_t).$$

В дальнейшем будем следовать схеме исследования, рассмотренной в § 6 главы IV. При этом в основу рассуждений будет положена теорема V.4.2.

Рассмотрим пространство траекторий. Пусть w есть вектор из пространства $\mathbf{R}^{(T+1)n}$ с n-мерными компонентами x_t, $t = 0, 1, \ldots, T$. Пусть

$$f(w) = \sum_{t=0}^{T} g(x_t, t), \tag{2.3}$$

$$\begin{gathered} \tilde{N} = \{w\colon\ x_0 \in N\}, \qquad \tilde{M} = \{w\colon\ x_T \in M\}, \\ \tilde{M}_t = \{w\colon\ (x_t, x_{t+1}) \in \operatorname{gf} a\}, \quad t = 0, 1, \ldots, T-1. \end{gathered} \tag{2.4}$$

Тогда, как легко проверить, исходя из определений, поставленная задача оптимального управления совпадает с задачей минимизации функции $f(w)$ на множестве

$$\widetilde{N} \cap \left(\bigcap_{t=0}^{T-1} \widetilde{M}_t \right) \cap \widetilde{M}. \tag{2.5}$$

Обозначим $\widetilde{w} = (\widetilde{x}_0, \widetilde{x}_1, \ldots, \widetilde{x}_T)$. Из основного предположения вытекает, что функция $f(w)$ допускает верхнюю выпуклую аппроксимацию в точке $\widetilde{w}$ и при этом

$$\partial f(\widetilde{w}) = (\partial g(\widetilde{x}_0, 0), \ldots, \partial g(\widetilde{x}_T, T)).$$

Легко вычислить конусы касательных направлений в точке $\widetilde{w}$ к $\widetilde{N}$, $\widetilde{M}_t$ и $\widetilde{M}$. Если $\overline{w} = (\overline{x}_0, \overline{x}_1, \ldots, \overline{x}_T)$, то

$$\widetilde{K}_{\widetilde{N}}(\widetilde{w}) = \{\overline{w}\colon \overline{x}_0 \in K_N(\widetilde{x}_0)\},$$

$$\widetilde{K}_t(\widetilde{w}) = \{\overline{w}\colon (\overline{x}_t, \overline{x}_{t+1}) \in K_a((\widetilde{x}_t, \widetilde{x}_{t+1}))\}, \quad t = 0, 1, \ldots, T-1,$$

$$\widetilde{K}_{\widetilde{M}}(\widetilde{w}) = \{\overline{w}\colon \overline{x}_T \in K_M(\widetilde{x}_T)\}.$$

В силу основного предположения все эти конусы являются локальными шатрами. Аналогично выкладкам, проделанным в § 6 главы IV, получаем

$$\widetilde{K}^*_{\widetilde{N}}(\widetilde{w}) = \{w^*\colon x_0^* \in K_N^*(\widetilde{x}_0), \quad x_t^* = 0, \quad t = 1, \ldots, T\}, \tag{2.6}$$

$$\widetilde{K}_t^*(\widetilde{w}) = \{w^*\colon (x_t^*, x_{t+1}^*) \in K_a^*((\widetilde{x}_t, \widetilde{x}_{t+1})), \quad x_k^* = 0, \quad k \neq t, \ t+1\}, \quad t = 0, 1, \ldots, T-1, \tag{2.7}$$

$$\widetilde{K}^*_{\widetilde{M}}(\widetilde{w}) = \{w^*\colon x_T^* \in K_M^*(\widetilde{x}_T), \quad x_t^* = 0, \quad t = 0, 1, \ldots, T-1\}, \tag{2.8}$$

где $w^* = (x_0^*, x_1^*, \ldots, x_T^*)$.

Воспользуемся теперь теоремой V.4.2. Заметим предварительно, что если в этой теореме $\operatorname{dom} h(\cdot, x_0) = X$, то предположение о том, что

$$\operatorname{int} \operatorname{dom} h(\cdot, x_0) \cap \left(\bigcap_{i=1}^{m} K_{M_i}(x_0) \right) \neq \varnothing,$$

можно опустить. Дело в том, что в рассматриваемом случае оно эквивалентно следующему соотношению:

$$K = \bigcap_{i=1}^{m} K_{M_i}(x_0) \neq \varnothing.$$

Так как согласно основному предположению функция $f(w)$ допускает верхнюю выпуклую аппроксимацию, непрерывную во всем пространстве, то в соответствии со сделанным выше замечанием никаких дополнительных предположений о пересечении конусов $\widetilde{K}_t(\widetilde{w})$ не нужно.

Итак, на основании теоремы V.4.2 существуют не все одновременно равные нулю векторы w_t^*, $t = 0, 1, \ldots, T-1$, w_b^* и w_e^* и число $\lambda \geqslant 0$ такие, что

$$\begin{gathered}\lambda w_f^* = w_b^* + \sum_{t=0}^{T-1} w_t^* + w_e^*,\\ w_f^* \in \partial f(\widetilde{w}), \quad w_b^* \in \widetilde{K}^*_{\widetilde{N}}(\widetilde{w}), \quad w_e^* \in \widetilde{K}^*_{\widetilde{M}}(\widetilde{w}), \\ w_t^* \in K_t^*(\widetilde{w}), \qquad t = 0, 1, \ldots, T-1.\end{gathered} \tag{2.9}$$

Согласно формулам (2.6)—(2.8) и выражению для $\partial f(\widetilde{w})$ векторы, входящие в соотношения (2.9), имеют следующую структуру:

$$\begin{gathered}w_f^* = (x_{f0}^*, x_{f1}^*, \ldots, x_{fT}^*), \quad x_{ft}^* \in \partial g(\widetilde{x}_t, t), \quad t = 0, 1, \ldots, T,\\ w_b^* = (x_b^*, 0, \ldots, 0), \quad x_b^* \in K_N^*(\widetilde{x}_0),\\ w_e^* = (0, 0, \ldots, x_e^*), \quad x_e^* \in K_M^*(\widetilde{x}_T),\\ w_t^* = \{0, 0, \ldots, x_t^*(t), \; x_{t+1}^*(t), 0, \ldots, 0\},\\ (x_t^*(t), x_{t+1}^*(t)) \in K_a^*((\widetilde{x}_t, \widetilde{x}_{t+1})), \quad t = 0, 1, \ldots, T-1.\end{gathered} \tag{2.10}$$

Записав (2.9) в покомпонентном виде, получим

$$\lambda x_{f0}^* = x_b^* + x_0^*(0), \tag{2.11}$$

$$\lambda x_{ft}^* = x_t^*(t-1) + x_t^*(t), \qquad t = 1, 2, \ldots, T-1, \tag{2.12}$$

$$\lambda x_{fT}^* = x_T^*(T-1) + x_e^*. \tag{2.13}$$

Введем следующее обозначение: $x_{t+1}^* \equiv x_{t+1}^*(t)$, $t = 0, 1, \ldots$ $\ldots, T-1$. Из формул (2.10) с учетом определения V.3.1

получаем, что

$$-x_t^*(t) \in a^*\left(x_{t+1}^*; \left(\tilde{x}_t, \tilde{x}_{t+1}\right)\right).$$

Теперь равенства (2.11)—(2.13) можно записать в следующем виде:

$$x_t^* - \lambda x_{ft}^* \in a^*\left(x_{t+1}^*; \left(\tilde{x}_t, \tilde{x}_{t+1}\right)\right), \quad t = 1, \ldots, T-1;$$

$$x_b^* - \lambda x_{f0}^* \in a^*\left(x_1^*; \left(\tilde{x}_0, \tilde{x}_1\right)\right), \quad \lambda x_{fT}^* = x_T^* + x_e^*.$$

Для компактности записи удобно обозначить $x_0^* \equiv x_b^*$. Тогда, учтя формулы (2.10), окончательно получим

$$x_t^* \in a^*\left(x_{t+1}^*; \left(\tilde{x}_t, \tilde{x}_{t+1}\right)\right) + \lambda \partial g\left(\tilde{x}_t, t\right), \quad t = 0, 1, \ldots, T-1, \tag{2.14}$$

$$x_0^* \in K_N^*\left(\tilde{x}_0\right), \quad x_T^* + x_e^* \in \lambda \partial g\left(\tilde{x}_T, T\right), \quad x_e^* \in K_M^*\left(\tilde{x}_T\right). \tag{2.15}$$

Заметим, что λ и x_t^*, x_e^* не могут быть все равными нулю, так как в противном случае из равенств (2.11)—(2.13) следовало бы, что все векторы w_b^*, w_t^*, w_e^* равны нулю в противоречии с утверждением теоремы V.4.2.

Таким образом, доказана

Теорема 2.1. *Пусть выполнены условия основного предположения. Тогда для того, чтобы траектория* $\{\tilde{x}_t\}_{t=0,1,\ldots,T}$ *системы* (2.1) *минимизировала функцию* (2.2) *при краевых условиях* $x_0 \in N$, $x_T \in M$, *необходимо, чтобы нашлись такие не все равные нулю векторы* x_t^*, $t = 0$, $1, \ldots, T$, x_e *и число* $\lambda \geqslant 0$, *чтобы выполнялись соотношения* (2.14), (2.15).

Проиллюстрируем применение теоремы 2.1 на одной из задач оптимального управления.

Пусть $X = \mathbf{R}^n$, $Y = \mathbf{R}^n$, $Z = X \times Y$. Предположим, что заданы функции $\varphi_i(z) \equiv \varphi_i(x, y)$, $i \in I^- \cup I^0$, которые непрерывно дифференцируемы. Их градиенты относительно переменных векторов z, x и y будем обозначать соответственно через

$$\varphi_{iz}'(z) \in \mathbf{R}^{2n}, \; \varphi_{ix}'(z) \in \mathbf{R}^n, \quad \varphi_{iy}'(z) \in \mathbf{R}^n,$$

так что

$$\varphi'_{iz}(z) = \left(\varphi'_{ix}(z),\ \varphi'_{iy}(z)\right). \tag{2.16}$$

Функции $\varphi_i(z)$ задают многозначное отображение

$$a(x) = \{y:\ \varphi_i(x,\ y) \leqslant 0,\ \ i \in I^-,\ \ \varphi_i(x,\ y) = 0,\ \ i \in I^0\}. \tag{2.17}$$

Требуется найти траекторию $\{x_t\}_{t=0,\,1,\,\ldots,\,T}$, удовлетворяющую включению

$$x_{t+1} \in a(x_t),$$

или, учитывая соотношение (2.17), удовлетворяющую соотношениям

$$\begin{gathered} \varphi_i(x_t,\ x_{t+1}) \leqslant 0, \qquad i \in I^-, \\ \varphi_i(x_t,\ x_{t+1}) = 0, \qquad i \in I^0, \\ t = 0,\ 1,\ \ldots,\ T-1 \end{gathered} \tag{2.18}$$

и краевым условиям $x_0 \in N$, $x_T \in M$, и минимизирующую функцию

$$\sum_{t=0}^{T} g(x_t,\ t).$$

Для простоты будем предполагать, что множества N и M выпуклы, а $g(x,\ t)$ имеет непрерывный градиент $g'_x(x,\ t)$. Ясно, что поставленная задача есть частный случай уже рассмотренной задачи, в которой отображение a задается соотношением (2.17).

Для применения теоремы 2.1 необходимо проверить выполнение основного предположения. Условие 3 очевидным образом выполнено, так как функция $g(x,\ t)$ непрерывно дифференцируемы. Условие 2 выполнено, так как N и M — выпуклые множества и согласно примеру V.1.5 конусы

$$K_N(\tilde{x}_0) = \operatorname{con}(N - \tilde{x}_0), \qquad K_M(\tilde{x}_T) = \operatorname{con}(M - \tilde{x}_T)$$

являются локальными шатрами. Таким образом, остается выяснить, при каких дополнительных ограничениях на $\varphi_i(z)$ выполнено условие 1 предположения. Так как

$$\operatorname{gf} a = \{z:\ \varphi_i(z) \leqslant 0,\ \ i \in I^-;\ \ \varphi_i(z) = 0,\ \ i \in I^0\},$$

то в соответствии с примерами V.1.2 и V.1.4 конус

$$K_a(z) = \{\bar{z}\colon \langle \bar{z}, \varphi'_{iz}(z)\rangle < 0, \quad i \in I^-(z);$$
$$\langle \bar{z}, \varphi'_{iz}(z)\rangle = 0, \quad i \in I^0\},$$

где $I^-(z) = \{i \in I^-\colon \varphi_i(z) = 0\}$, является локальным шатром к $\operatorname{gf} a$ в точке $z \in \operatorname{gf} a$, если только он не пуст и градиенты $\varphi'_{iz}(z)$, $i \in I^0$, линейно независимы. Будем предполагать выполнение этих условий на оптимальной траектории $\{\tilde{x}_t\}_{t=0,1,\ldots,T}$, т. е. для всех z, имеющих вид $z = (\tilde{x}_t, \tilde{x}_{t+1})$.

Согласно примеру V.1.2 и очевидным образом переписанной формуле (V.1.9) получаем

$$K_a^*(z) = \Big\{z^*\colon z^* = -\sum_{i \in I^- \cup I^0} \lambda_i \varphi'_{iz}(z), \quad \lambda_i \geqslant 0,$$
$$\lambda_i \varphi_i(z) = 0, \quad i \in I^-\Big\}. \quad (2.19)$$

Учитывая формулу (2.16) и то, что $z^* = (x^*, y^*)$, из соотношения (2.19) получаем, что $(x^*, y^*) \in K_a^*(z)$ тогда и только тогда, когда

$$x^* = -\sum_{i \in I^- \cup I^0} \lambda_i \varphi'_{ix}(z), \quad y^* = -\sum_{i \in I^- \cup I^0} \lambda_i \varphi'_{iy}(z),$$
$$\lambda_i \geqslant 0, \quad \lambda_i \varphi_i(z) = 0, \quad i \in I^-.$$

Из этих формул вытекает следующая формула для локально сопряженного отображения:

$$a^*(y^*; z) = \Big\{\sum_{i \in I^- \cup I^0} \lambda_i \varphi'_{ix}(z)\colon y^* +$$
$$+ \sum_{i \in I^- \cup I^0} \lambda_i \varphi'_{iy}(z) = 0, \ \lambda_i \geqslant 0, \ \lambda_i \varphi_i(z) = 0, \ i \in I^-\Big\}. \quad (2.20)$$

Таким образом, для отображения a, описываемого формулами (2.17), локальное сопряженное отображение в предположении линейной независимости векторов $\varphi'_{iz}(z)$, $i \in I^0$, задается формулой (2.20). В этих же предположениях $K_a(z)$ есть локальный шатер и можно применить теорему 2.1. Согласно ей найдутся такие векторы x_t^*, $t = 0, 1, \ldots, T$, x_e^* и число $\lambda_0 \geqslant 0$, что выполняются соотношения (2.14), (2.15). С учетом формулы (2.20) это позволяет написать:

$$x_t^* = \sum_{i \in I^- \cup I^0} \lambda_{it} \varphi'_{ix}(\tilde{x}_t, \tilde{x}_{t+1}) + \lambda_0 g'(\tilde{x}_t, t), \tag{2.21}$$

$$x_{t+1}^* + \sum_{i \in I^- \cup I^0} \lambda_{it} \varphi'_{iy}(\tilde{x}_t, \tilde{x}_{t+1}) = 0. \tag{2.22}$$

$$\begin{aligned} \lambda_{it} \geqslant 0, \quad i \in I^-, \quad \lambda_{it} \varphi_i(\tilde{x}_t, \tilde{x}_{t+1}) = 0, \\ t = 0, 1, \ldots, T-1, \end{aligned} \tag{2.23}$$

$$\begin{aligned} x_0^* \in K_N^*(\tilde{x}_0), \quad x_T^* + x_e^* = \lambda_0 g'(\tilde{x}_T, T), \\ x_e^* \in K_M^*(\tilde{x}_T). \end{aligned} \tag{2.24}$$

Сформулируем полученный результат.

Теорема 2.2. *Пусть выполнены следующие предположения:*

1) *множества N и M выпуклы;*

2) *функции $g(x, t)$, $t = 0, 1, \ldots, T$, непрерывно дифференцируемы по x;*

3) *функции $\varphi_i(z)$, $i \in I^- \cup I^0$, непрерывно дифференцируемы по z и векторы $\varphi'_{iz}(z)$, $i \in I^0$, линейно независимы, а конусы*

$$K_a(z) = \{\bar{z}\colon \langle z, \varphi'_{iz}(z)\rangle < 0,\ i \in I^-(z),$$
$$\langle \bar{z}, \varphi'_{iz}(z)\rangle = 0,\ i \in I^0\}$$

не пусты при всех $z = (\tilde{x}_t, \tilde{x}_{t+1})$, $t = 0, 1, \ldots, T-1$, где $\{\tilde{x}_t\}_{t=0, 1, \ldots, T}$ — траектория, минимизирующая сумму

$$\sum_{t=0}^{T} g(x_t, t)$$

по всем траекториям, удовлетворяющим соотношениям (2.18) *и краевым условиям $x_0 \in N$, $x_T \in M$.*

Тогда существуют такие не все равные нулю векторы x_t^*, $t = 0, 1, \ldots, T$, x_e^* и число $\lambda_0 \geqslant 0$, что выполнены соотношения (2.21)—(2.24).

Заметим, что теорему 2.2 можно было бы доказать при более слабых предположениях, чем этого требуют условия, наложенные выше на функции $\varphi_i(z)$. Однако, при этом не удалось бы показать, что величины x_t^*, x_e^* и λ_0 не могут быть одновременно равными нулю. Таким образом, условия теоремы 2.2 гарантируют некоторую невырожденность получаемых необходимых условий минимума. С другой стороны, предположения теоремы 2.2 обусловлены теоремой 2.1, которая охватывает случай весьма общего многозначного отображения a, вообще говоря, не обладающего какими-либо свойствами гладкости. В таких условиях существенное ослабление предположений, при которых доказывается теорема 2.1, представляется невозможным.

§ 3. Необходимые условия минимума для дифференциальных включений

Возвратимся к рассмотрению дифференциальных включений, определенных в § 1. Как было сказано выше, дифференциальное включение даже при заданной начальной точке определяет не единственную траекторию, ему удовлетворяющую. Поэтому имеет смысл ставить задачу отыскания среди всех траекторий дифференциального включения такой, которая обладала бы дополнительными свойствами. Например, можно поставить задачу отыскания траектории, минимизирующей некоторый функционал, зависящий от этой траектории.

Пусть снова $X = Y = \mathbf{R}^n$ и a — многозначное отображение, $a(x) \subseteq Y$. Если заданы множества $N \subseteq X$, $M \subseteq X$, то можно сформулировать следующую задачу: среди всех траекторий $x(t)$, $t \in [0, 1]$, удовлетворяющих почти всюду дифференциальному включению

$$\frac{d}{dt} x(t) \in a(x(t)), \qquad t \in [0, 1] \tag{3.1}$$

и условиям $x(0) \in N$, $x(1) \in M$, найти такую, которая минимизирует выражение

$$I(x(\cdot)) = \int_0^1 g(x(t), t)\, dt + \varphi_0(x(1)). \tag{3.2}$$

Поставленную задачу будем называть *задачей оптимального управления*. Условия, которые надо наложить на входящие в постановку задачи данные и при которых она будет решаться, будут сформулированы ниже по мере необходимости.

Для сокращения обозначений условимся, что в дальнейшем производная по параметру t будет обозначаться точкой над функцией:

$$\dot{x}(t) \equiv \frac{d}{dt} x(t).$$

Через $\tilde{x}(t)$ обозначим решение поставленной задачи оптимизации, существование которого предполагается.

1. Аппроксимация краевых условий. Замкнутое множество N, задающее ограничение на левый конец траектории, всегда может быть задано при помощи неравенства

$$N = \{x\colon\ d(x|N) \leqslant 0\}, \tag{3.3}$$

где $d(x|N)$ — евклидово расстояние от точки x до N, определенное в п. 3 § 2 главы V. Согласно лемме V.2.3 $d(x|N)$ удовлетворяет условию Липшица с константой, равной единице.

В некоторых задачах множество N задается при помощи конечной системы неравенств и равенств

$$\varphi_b^i(x) \leqslant 0, \quad i \in I^-; \qquad \varphi_b^i(x) = 0, \quad i \in I^0. \tag{3.4}$$

Эти неравенства и равенства могут быть сведены к одному, если положить

$$\varphi_b(x) = \max\left\{\max_{i\in I^-} \varphi_b^i(x),\ \max_{i\in I^0} \left|\varphi_b^i(x)\right|\right\}.$$

Тогда множество N может быть задано одним неравенством:

$$N = \{x\colon\ \varphi_b(x) \leqslant 0\}. \tag{3.5}$$

В обоих случаях можно определить для $\delta \geqslant 0$ семейство множеств

$$N_\delta = \{x\colon\ d(x|N) \leqslant c\delta\}, \tag{3.6}$$

или

$$N_\delta = \{x\colon\ \varphi_b(x) \leqslant c\delta\}, \tag{3.7}$$

где константа c будет выбрана в дальнейшем. Независимо от того, совпадает ли функция $\varphi_b(x)$ с $d(x|N)$ или нет, всегда будем предполагать, что $\varphi_b(x)$ удовлетворяет условию Липшица. Легко проверить, что если функции $\varphi_b^i(x)$, входящие в формулы (3.4), удовлетворяют этому условию, то и функция $\varphi_\delta(x)$, определенная выше, ему удовлетворяет.

Совершенно аналогично определяются множества M_δ, $\delta \geqslant 0$ при помощи неравенства

$$M_\delta = \{x\colon\ \varphi_l(x) \leqslant c\delta\}, \tag{3.8}$$

где $\varphi_l(x)$ — функция, удовлетворяющая условию Липшица и такая, что

$$M = \{x\colon\ \varphi_l(x) \leqslant 0\}.$$

В частности, в качестве $\varphi_l(x)$ может быть взята функция $d(x|M)$.

Предположение A. Множества N и M замкнуты. Множества N_δ, $\delta \geqslant 0$, в каждой точке x обладают локальным шатром $K_{N_\delta}(x)$, причем, если последовательности $x_k \in N_{\delta_k}$, $x_k^* \in K_{\delta_k}^*(x_k)$ сходятся к x_0 и x_0^* соответственно, а $\delta_k \to 0$, то

$$x_0^* \in K_N^*(x_0).$$

Аналогичным свойством обладают множества M_δ.

Покажем, что предположение A достаточно естественно, и если N — выпуклое множество, то оно выполняется.

Лемма 3.1. *Если $\varphi(x)$ — непрерывная выпуклая функция, и*

$$N_\delta = \{x\colon\ \varphi(x) \leqslant c\delta\},$$

то множества N_δ удовлетворяют предположению A, если в качестве $K_{N_\delta}(x)$ естественным образом взят конус

$$K_{N_\delta}(x) = \{\bar{x}\colon\ \bar{x} = \gamma(x_1 - x),\ \gamma > 0,\ x_1 \in N_\delta\}. \tag{3.9}$$

Доказательство. По определению

$$K_{N_\delta}^*(x) = \{x^*\colon\ \langle\bar{x}, x^*\rangle \geqslant 0,\ \bar{x} \in K_{N_\delta}(x)\}.$$

Поэтому, если $x^* \in K_{N_\delta}^*(x)$, то

$$\langle x_1 - x,\ x^*\rangle \geqslant 0, \qquad x_1 \in N_\delta,$$

и, в частности,

$$\langle x_1 - x,\ x^*\rangle \geqslant 0, \qquad x_1 \in N, \tag{3.10}$$

так как $N \subseteq N_\delta$. Если теперь $x \to x_0$, $x^* \to x_0^*$ и $\delta \to 0$, то по непрерывности функции $\varphi(x)$ из неравенства $\varphi(x) \leqslant c\delta$ получаем, что $\varphi(x_0) \leqslant 0$, т. е. $x_0 \in N$, а неравенство (3.10) переходит в неравенство

$$\langle x_1 - x_0, x_0^* \rangle \geqslant 0, \qquad x_1 \in N.$$

Но последнее неравенство эквивалентно включению

$$x_0^* \in K_N^*(x_0),$$

что легко усматривается из формулы (3.9).

Согласно примеру V.1.5 конус, определенный формулой (3.9), есть локальный шатер. Это завершает доказательство леммы.

2. Аппроксимация функционала.

Предположение В. Функции $g(x, t)$ и $\varphi_0(x)$, входящие в соотношение (2.2), непрерывны по x и t и удовлетворяют условию Липшица

$$|g(x_1, t) - g(x_2, t)| < L\|x_1 - x_2\|,$$
$$|\varphi_0(x_1) - \varphi(x_2)| < L\|x_1 - x_2\|$$

в любой ограниченной области пространства X. Константа L может зависеть от этой области, но не зависит от $t \in [0, 1]$. При этом $g(x, t)$ и $\varphi_0(x)$ допускают в каждой точке (x, t) верхнюю выпуклую аппроксимацию, а их субдифференциалы по x $\partial g(x, t)$ и $\partial\varphi_0(x)$ равномерно ограничены в каждой ограниченной области и полунепрерывно сверху зависят от x и t.

Из предположения В вытекает, что соответствующая субдифференциалу $\partial g(x, t)$ верхняя выпуклая аппроксимация $h_g(\bar{x}, (x, t))$, которая в силу формулы V.2.3 имеет вид:

$$h_g(\bar{x}, (x, t)) = \sup_{x^*} \{\langle \bar{x}, x^* \rangle: \ x^* \in \partial g(x, t)\},$$

непрерывно зависит от $\bar{x}$. Действительно, $\partial g(x, t)$ — ограниченное множество, и, значит, в. в. а. h_g определена и принимает конечное значение при всех $\bar{x}$. Так как h_g — выпуклая по $\bar{x}$ функция, определенная во всем пространстве X, то в силу теоремы II.1.4 она непрерывна и

$$\operatorname{dom} h_g(\cdot, (x, t)) = X.$$

Аналогичные рассуждения справедливы для верхней выпуклой аппроксимации, соответствующей функции φ_0.

Лемма 3.2. *Если выполнено предположение* B *и последовательность непрерывных функций* $x_k(t)$ *равномерно сходится к непрерывной функции* $x_0(t)$ *на интервале* $t \in [0, 1]$, *то*

$$\sum_{t=0,\delta,\ldots,1-\delta} \delta g(x_k(t), t) \to \int_0^1 g(x_0(t), t)\, dt$$

при $\delta = 2^{-m} \to 0$ *и* $k \to \infty$.

Доказательство. Так как функция $g(x_0(t), t)$ непрерывна на замкнутом интервале $[0, 1]$, то она равномерно непрерывна на этом интервале. Поэтому для заданного $\varepsilon > 0$ существует такое $\delta(\varepsilon)$, что для $\delta < \delta(\varepsilon)$ выполняется неравенство

$$|g(x_0(t_1), t_1) - g(x_0(t_2), t_2)| < \varepsilon, \tag{3.11}$$

как только $|t_1 - t_2| < \delta(\varepsilon)$. Далее,

$$\int_0^1 g(x_0(t), t)\, dt = \sum_{j=0,1,\ldots,2^m-1} \int_{j\delta}^{(j+1)\delta} [g(x_0(t), t) - \\ - g(x_0(j\delta), j\delta)]\, dt + \sum_{j=0,1,\ldots,2^m-1} \delta g(x_0(j\delta), j\delta). \tag{3.12}$$

Так как $2^m\delta = 1$, то в силу неравенства (3.11) первая сумма меньше ε при $\delta < \delta(\varepsilon)$. Пусть

$$\gamma_k = \max_{t\in[0,1]} \| x_k(t) - x_0(t) \|.$$

Тогда

$$\left| \sum_{j=0,1,\ldots,2^m-1} \delta [g(x_0(j\delta), j\delta) - g(x_k(j\delta), j\delta)] \right| \leqslant \\ \leqslant 2^m \delta L \gamma_k = L\gamma_k.$$

Если k достаточно велико, то $L\gamma_k < \varepsilon$ и поэтому вторая сумма в правой части формулы (3.12) отличается от

$$\sum_{j=0,1,\ldots,2^m-1} \delta g(x_k(j\delta), j\delta)$$

не больше чем на ε.

Итак, для малых δ и больших k

$$\left| \int_0^1 g(x_0(t),t)\,dt - \sum_{t=0,\delta,\ldots,1-\delta} \delta g(x_k(t),\,t) \right| < 2\varepsilon,$$

что завершает доказательство.

3. Аппроксимация дифференциального включения. Необходимые условия минимума для поставленной в начале параграфа задачи оптимизации будут получены путем предельного перехода от дискретной аппроксимации задачи к задаче с непрерывным временем. Поэтому необходимо наложить некоторые условия на отображение a, которые бы гарантировали возможность такого перехода.

Предположение С. Отображение a выпуклозначно и замкнуто, удовлетворяет условию Липшица в ε-трубке траектории $\tilde{x}(t)$, $t \in [0,\ 1]$, и множество $a(\tilde{x}(0))$ ограничено. Кроме того, конусы $K_a((x,\ y))$ являются локальными шатрами и локально сопряженное отображение $a^*(y^*;\ (x,\ y))$ полунепрерывно сверху зависит от своих аргументов и равномерно ограничено для всех x из ε-трубки траектории $\tilde{x}(\cdot)$ и $y \in a(x)$.

Равномерная ограниченность $a^*(y^*;\ (x,\ y))$ означает, что существует такая константа c, для которой

$$\|a^*(y^*;\ (x,\ y))\| \leqslant c\|y^*\| \tag{3.13}$$

в указанной выше области изменения x и y.

Рассмотрим теперь следующую задачу оптимального управления с дискретным временем. Пусть, как и раньше, $\delta = 2^{-m}$, где m — целое число. Требуется минимизировать величину

$$I_\delta(x_\delta(\cdot)) = \sum_{t=0,\delta,\ldots,1-\delta} \delta g(x_\delta(t),\,t) + \\ + \varphi_0(x_\delta(1)) + \delta \sum_{t=0,\delta,\ldots,1} \|\tilde{x}(t) - x_\delta(t)\|^2 \tag{3.14}$$

при следующих условиях:

$$x_\delta(t+\delta) \in x_\delta(t) + \delta a(x_\delta(t)), \quad t = 0,\ \delta,\ \ldots,\ 1-\delta, \tag{3.15}$$

$$x_\delta(0) \in N_\delta, \qquad x_\delta(1) \in M_\delta, \tag{3.16}$$

и $x_\delta(t)$, $t = 0,\ \delta,\ \ldots,\ 1$, принадлежит ε-трубке траектории $\tilde{x}(\cdot)$.

Покажем, что решение задачи существует, если только константа c в неравенствах (3.7) и (3.8) выбрана правильно, а число δ достаточно мало. Действительно, согласно предположению С и теореме 1.2 при малом δ существует траектория $x_\delta(t)$, удовлетворяющая включению (3.15) и лежащая в ε-трубке. При этом

$$\|\widetilde{x}(t) - x_\delta(t)\| \leqslant c_1\delta, \qquad t = 0,\ \delta,\ \ldots,\ 1.$$

Так как функции φ_b и φ_l в ε-трубке удовлетворяют условию Липшица, то

$$\varphi_b(x_\delta(0)) \leqslant \varphi_b(\widetilde{x}(0)) + L\|\widetilde{x}(0) - x_\delta(0)\| \leqslant c_1 L\delta$$

и, аналогично,

$$\varphi_l(x_\delta(1)) \leqslant c_1 L\delta.$$

Поэтому, если в неравенствах (3.7) и (3.8) $c = c_1 L$, то $x_\delta(0) \in N_\delta$, $x_\delta(1) \in M_\delta$ и построенная траектория удовлетворяет включениям (3.15), (3.16).

Таким образом, множество допустимых траекторий дискретной оптимизационной задачи (3.14)—(3.16) не пусто. Так как все траектории этой задачи принадлежат ограниченному множеству — ε-трубке,— а отображение a замкнуто, то нетрудно убедиться, что множество таких траекторий компактно. Отсюда следует, что в дискретной задаче оптимизации ищется минимум непрерывной функции I_δ на компактном множестве и поэтому этот минимум достигается на некоторой траектории.

Обозначим через $\widetilde{x}_\delta(t)$ решение дискретной задачи оптимизации. Оно определено для $t = 0,\ \delta,\ \ldots,\ 1$. Доопределим $\widetilde{x}_\delta(t)$ для всех t из $[0,\ 1]$ путем линейной аппроксимации так, как это делалось в § 1.

Лемма 3.3. *Если выполнено предположение* С, *то траектории* $\widetilde{x}_\delta(t)$ *равномерно сходятся к* $\widetilde{x}(t)$ *при* $\delta \to 0$.

Доказательство. Допустим противное:

$$\gamma(\delta) = \max_{0 \leqslant t \leqslant 1} \|\widetilde{x}(t) - \widetilde{x}_\delta(t)\| \geqslant \Delta > 0 \tag{3.17}$$

при всех δ. Так как в силу предположения С множества $a(x)$ равномерно ограничены в ε-трубке, то из соотношения (3.15) следует, что $\widetilde{x}_\delta(t)$ удовлетворяет условию Липшица с некоторой константой, не зависящей от δ.

Доказательство этого факта было намечено при доказательстве теоремы 1.1.

Таким образом, все траектории $\widetilde{x}_\delta(t)$ равномерно ограничены (они лежат в ε-трубке) и равностепенно непрерывны. Поэтому из них можно выбрать сходящуюся подпоследовательность. Без ограничения общности можно считать, что $x_\delta(t) \to x_0(t)$, причем сходимость равномерная.

Повторяя рассуждения, уже использованные при доказательстве теоремы 1.1, можно убедиться, что функция $x_0(t)$ удовлетворяет дифференциальному включению (3.1) и $x_0(0) \in N$, $x_0(1) \in M$, так как $x_\delta(0) \in N_\delta$, $x_\delta(1) \in M_\delta$. Отсюда следует, что $x_0(t)$ удовлетворяет ограничениям поставленной в начале параграфа задачи оптимального управления.

Из неравенства (3.17) следует, что

$$\max_{0 \leqslant t \leqslant 1} \|\widetilde{x}(t) - x_0(t)\| \geqslant \Delta;$$

поэтому существует такой промежуток $[a, b] \subseteq [0, 1]$, что

$$\|\widetilde{x}(t) - x_0(t)\| \geqslant \frac{\Delta}{2}, \qquad t \in [a, b].$$

Так как последовательность $\widetilde{x}_\delta(t)$ равномерно сходится к $x_0(t)$, то при малых δ

$$\|\widetilde{x}(t) - \widetilde{x}_\delta(t)\| \geqslant \frac{\Delta}{4}, \qquad t = q\delta \in [a, b]. \tag{3.18}$$

Заметим, что число точек $t = q\delta$, попадающих в промежуток $[a, b]$ при малых δ будет равно $[(b-a)/\delta]$, где квадратные скобки означают, что берется целая часть числа. Поэтому

$$\delta \sum_{t=0,\delta,\ldots,1} \|\widetilde{x}(t) - \widetilde{x}_\delta(t)\|^2 \geqslant$$
$$\geqslant \delta \sum_{t=q\delta\in[a,b]} \|\widetilde{x}(t) - \widetilde{x}_\delta(t)\|^2 \geqslant \delta \left[\frac{b-a}{\delta}\right] \left(\frac{\Delta}{4}\right)^2 \geqslant$$
$$\geqslant \frac{b-a}{2} \left(\frac{\Delta}{4}\right)^2. \tag{3.19}$$

Рассмотрим теперь траектории $x_\delta(t)$, которые согласно теореме 1.2 сходятся к $\widetilde{x}(t)$ и

$$\|\widetilde{x}(t)-x_\delta(t)\|\leqslant c_1\delta.$$

Как было сказано выше, они при малых δ удовлетворяют ограничениям (3.15), (3.16) дискретной задачи оптимизации. Из предположения В и леммы 3.2 вытекает, что

$$\sum_{t=0,\delta,\ldots,1-\delta}\delta g(x_\delta(t),\ t)+\varphi_0(x_\delta(1))\to I\left(\widetilde{x}(\cdot)\right),$$

а так как

$$\delta\sum_{t=0,\delta,\ldots,1}\left\|\widetilde{x}(t)-x_\delta(t)\right\|\leqslant\delta(c_1\delta)\,2^m=c_1\delta,$$

то

$$I_\delta(x_\delta(\cdot))\to I(\widetilde{x}(\cdot)). \tag{3.20}$$

С другой стороны, $\widetilde{x}_\delta(\cdot)$ минимизирует I_δ и поэтому

$$I_\delta(x_\delta(\cdot))\geqslant I_\delta\left(\widetilde{x}_\delta(\cdot)\right)=\sum_{t=0,\delta,\ldots,1-\delta}\delta g\left(\widetilde{x}_\delta(t),\ t\right)+\\+\varphi_0\left(\widetilde{x}_\delta(1)\right)+\delta\sum_{t=0,\delta,\ldots,1}\left\|\widetilde{x}(t)-\widetilde{x}_\delta(t)\right\|.$$

Снова воспользовавшись леммой 3.2 и неравенством (3.19) и устремляя δ к нулю, получаем, что

$$I\left(\widetilde{x}(\cdot)\right)\geqslant I\left(x_0(\cdot)\right)+\frac{b-a}{2}\left(\frac{\Delta}{4}\right)^2,$$

что противоречит оптимальности траектории $\widetilde{x}(\cdot)$. Лемма доказана.

Итак, траектории $\widetilde{x}_\delta(\cdot)$ сходятся к $\widetilde{x}(\cdot)$ и, значит, целиком лежат внутри ε-трубки. Это позволяет применить к дискретной задаче оптимизации теорему 2.1 предыдущего параграфа.

Теорема 3.1. *Пусть выполнены предположения* А, В, С. *Тогда существует такое число* $\lambda_\delta\geqslant 0$ *и векторы* $x_\delta^*(t)$, $t=0,\ \delta,\ \ldots,\ 1$, $x_{l\delta}^*$, *не все равные нулю одновременно, что*

$$-\Delta_\delta x_\delta^*(t)\in a^*\left(x_\delta^*(t+\delta);\quad\left(\widetilde{x}_\delta(t),\ \Delta_\delta\widetilde{x}(t)\right)\right)+\\+2\lambda_\delta\left(\widetilde{x}(t)-\widetilde{x}_\delta(t)\right)+\lambda_\delta\partial g\left(\widetilde{x}_\delta(t),\ t\right), \tag{3.21}$$

$t = 0, \delta, \ldots, 1 - \delta,$

$$\begin{gathered} x_\delta^*(0) \in K_{N_\delta}^*\left(\tilde{x}_\delta(0)\right), \\ x_\delta^*(1) + x_{l_\delta}^* \in \lambda_\delta \partial\varphi_0\left(\tilde{x}_\delta(1)\right), \\ x_{l_\delta}^* \in K_{M_\delta}^*\left(\tilde{x}_\delta(1)\right), \end{gathered} \tag{3.22}$$

где по определению

$$\Delta_\delta x(t) \equiv \frac{x(t+\delta) - x(t)}{\delta}.$$

Доказательство получается прямым применением теоремы 2.1. Единственное, что нужно сделать — это вычислить отображение, локально сопряженное к отображению $x + \delta a(x)$, ибо включение (3.15) эквивалентно соотношению (2.1), если в последнем вместо $a(x)$ взять $x + \delta a(x)$. Обозначим отображение $x + \delta a(x)$ через $I + \delta a$, где I — тождественное отображение ($I(x) = x$).

Тогда, если $K_{I+\delta a}(z)$, $z = (x, y)$,— локальный шатер к $\mathrm{gf}(I + \delta a)$ в точке z, то существуют такие функции $r_1(\bar{x}, \bar{y})$, $r_2(\bar{x}, \bar{y})$, что r_i стремится к нулю быстрее, чем $\|\bar{x}\| + \|\bar{y}\|$, и для всех достаточно малых $\bar{x}$, $\bar{y}$ из некоторого конуса $Q \subseteq \mathrm{ri}\, K_{I+\delta a}(z)$ выполняется включение

$$y + \bar{y} + r_2(\bar{x}, \bar{y}) \in (x + \bar{x} + r_1(\bar{x}, \bar{y}) + \delta a(x + \bar{x} + r_1(\bar{x}, \bar{y}))).$$

Простое преобразование этого включения приводит к соотношению

$$\frac{y - x}{\delta} + \frac{\bar{y} - \bar{x}}{\delta} + \\ + \frac{1}{\delta}\left[r_2(\bar{x}, \bar{y}) - r_1(\bar{x}, \bar{y})\right] \in a\left(x + \bar{x} + r_1(\bar{x}, \bar{y})\right),$$

из которого следует, что $K_a\left(\left(x, \frac{y-x}{\delta}\right)\right)$ — локальный шатер к $\mathrm{gf}\, a$ в точке $\left(x, \frac{y-x}{\delta}\right)$ и

$$\left(\bar{x}, \frac{\bar{y} - \bar{x}}{\delta}\right) \in K_a\left(\left(x, \frac{y-x}{\delta}\right)\right). \tag{3.23}$$

Точно также можно показать, что если выполнено включение (3.23) и K_a — локальный шатер, то

$$(\bar{x}, \bar{y}) \in K_{I+\delta a}((x, y)) \tag{3.24}$$

и $K_{I+\delta a}$ — локальный шатер. Таким образом, соотношения (3.23) и (3.24) эквивалентны.

Пусть теперь $x^* \in (I + \delta a)^*(y^*; z)$, $z = (x, y)$. По определению это значит, что

$$(-x^*, y^*) \in K^*_{I+\delta a}(z),$$

т. е.

$$-\langle \bar{x}, x^* \rangle + \langle \bar{y}, y^* \rangle \geqslant 0, \quad (\bar{x}, \bar{y}) \in K_{I+\delta a}(z).$$

Перепишем последнее соотношение в следующем виде:

$$-\left\langle \bar{x}, \frac{x^* - y^*}{\delta} \right\rangle + \left\langle \frac{\bar{y} - \bar{x}}{\delta}, y^* \right\rangle \geqslant 0, \quad (\bar{x}, \bar{y}) \in K_{I+\delta a}.$$

Если теперь учесть эквивалентность включений (3.23) и (3.24), то получим, что

$$\left(-\frac{x^* - y^*}{\delta}, y^*\right) \in K_a^*\left(\left(x, \frac{y - x}{\delta}\right)\right),$$

т. е.

$$\frac{x^* - y^*}{\delta} \in a^*\left(y^*; \left(x, \frac{y - x}{\delta}\right)\right), \tag{3.25}$$

или

$$x^* \in y^* + \delta a^*\left(y^*; \left(x, \frac{y - x}{\delta}\right)\right). \tag{3.26}$$

Итак, включение $x^* \in (I + \delta a)^*(y^*; z)$ эквивалентно включению (3.26).

Применение теоремы 2.1 к задаче (3.14)—(3.16) после очевидных переобозначений показывает, что должны выполняться соотношения (2.14), (2.15):

$$x_\delta^*(t) \in (I + \delta a)^*\left(x_\delta^*(t + \delta); \left(\tilde{x}_\delta(t), \tilde{x}_\delta(t + \delta)\right)\right) +$$
$$+ \lambda_\delta \partial g\left(\tilde{x}_\delta(t), t\right)\delta + 2\lambda_\delta \delta\left(\tilde{x}(t) - \tilde{x}_\delta(t)\right),$$
$$t = 0, \delta, \ldots, 1 - \delta,$$

$$x_\delta^*(0) \in K^*_{N_\delta}\left(\tilde{x}(0)\right), x_\delta^*(1) + x_l^* \in \lambda_\delta \partial \varphi_0\left(\tilde{x}_\delta(1)\right),$$
$$x_l^* \in K^*_{M_\delta}\left(\tilde{x}_\delta(1)\right).$$

Если теперь учесть, что включение $x^* \in (I + \delta a)^*(y^*;\ z)$ эквивалентно (3.26), то только что приведенные соотношения приобретают вид (3.21), (3.22).

4. Необходимые условия минимума. Теперь можно перейти к формулировке и доказательству основного результата.

Теорема 3.2. *Пусть выполнены предположения* A, B, C *и* $\tilde{x}(\cdot)$ *минимизирует функционал* (3.2), *среди всех траекторий, удовлетворяющих включению* (3.1) *и краевым условиям* $x(0) \in N$, $x(1) \in M$.

Тогда существует число $\lambda_0 \geqslant 0$, *вектор* x_l^* *и функция* $x^*(t)$, $t \in [0, 1]$, *не равные нулю одновременно, такие что*:

1) $x^*(1) + x_l^* \in \lambda_0 \partial\varphi_0(\tilde{x}(1))$;

$$x_l^* \in K_M^*(\tilde{x}(1)), \quad x^*(0) \in K_N^*(\tilde{x}(0));$$

2) *функция* $x^*(t)$ *удовлетворяет условию Липшица и*

$$-\dot{x}^*(t) \in A^*(x^*(t);\ \tilde{x}(t))$$

почти всюду на отрезке [0, 1];

3) $\langle \dot{\tilde{x}}(t), x^*(t)\rangle = W_a(\tilde{x}(t), x^*(t))$ *почти всюду на* [0, 1].

Напомним, что W_a и A^* определены формулами V.3.1 и V.3.14.

Доказательство теоремы опирается на следующую лемму.

Лемма 3.4. *Если выполнено предположение* C, *то множество* $A^*(y^*;\ x)$ *полунепрерывно сверху зависит от* $y^* \in \mathbf{R}^n$ *и* x *из* ε*-трубки траектории* $\tilde{x}(\cdot)$. *Кроме того,*

$$\|A^*(y^*;\ x)\| \leqslant c\|y^*\|$$

для всех x *из* ε*-трубки траектории* $\tilde{x}(\cdot)$.

Доказательство. Допустим противное. Тогда для некоторых $\Delta > 0$ и x_0, y_0^* найдутся такие последовательности $x_k \to x_0$, $y_k^* \to y_0^*$ и точки $x_k^* \in a^*(y_k^*;\ (x_k, y_k))$, $y_k \in a(x_k;\ y_k^*)$, что

$$x_k^* \notin A^*(y_0^*;\ x_0) + \Delta B, \tag{3.27}$$

где B — единичный шар в $\mathbf{R}^n$. Так как $y_k \in a(x_k)$, а множества $a(x)$ равномерно ограничены при всех x из ε-трубки в силу предположения C, то все y_k также равномерно ограничены, и можно считать, что $y_k \to y_0$. По лемме V.3.1 множества $a(x;\ y^*)$ полунепрерывно сверху зависят от своих аргументов и замкнуты, а поэтому из включения $y_k \in a\left(x_k;\ y_k^*\right)$ следует, что

$$y_0 \in a\left(x_0;\ y_0^*\right). \tag{3.28}$$

Далее, по предположению C множества $a^*(y^*;\ (x,\ y))$ равномерно ограничены, так что

$$\| x_k^* \| \leqslant \| a^*\left(y_k^*;\ (x_k,\ y_k)\right) \| \leqslant c \| y_k^* \|.$$

В силу того, что $y_k^* \to y_0^*$, все y_k^* ограничены, а значит ограничены и x_k^*. Значит, без ограничения общности можно считать, что $x_k^* \to x_0^*$. Так как по предположению C множество $a^*(y^*;\ (x,\ y))$ полунепрерывно сверху зависит от своих аргументов, то из включения $x_k^* \in a^*\left(y_k^*;\ (x_k,\ y_k)\right)$ вытекает, что

$$x_0^* \in a^*\left(y_0^*;\ (x_0,\ y_0)\right).$$

Последнее соотношение вместе с включением (3.28) и определением V.3.2 отображения A^* показывает, что $x_0^* \in A^*\left(y_0^*,\ x_0\right)$. Но последнее противоречит соотношению (3.27) и тому, что $x_k^* \to x_0^*$.

Наконец, последнее утверждение леммы сразу следует из соотношения (3.13) и определения V.3.2 отображения A^*.

Пусть теперь $\tilde{x}_\delta(t)$ — последовательность решений дискретной задачи оптимизации. Согласно лемме 3.3 она равномерно сходится к $\tilde{x}(t)$. Кроме того, существует последовательность чисел $\lambda_\delta \geqslant 0$ и векторов $x_\delta^*(t)$, $t = 0,\ \delta,\ \ldots,\ 1$, $x_{l\delta}^*$, одновременно не равных нулю и таких, что выполняются соотношения (3.21), (3.22).

Заметим, что λ_δ, $x_\delta^*(1)$ и $x_{l\delta}^*$ не могут быть равны нулю одновременно. Если это было бы так, то из включения (3.21) и ограниченности отображения $a^*(y^*;\ z)$ следовало бы, что и все $x_\delta^*(t)$, $t = 0,\ \delta,\ \ldots$, равны нулю, что противоречит вышесказанному. Поэтому после соответствующей перенормировки всегда можно считать, что

$$\lambda_\delta + \|x_\delta^*(1)\| + \|x_{l\delta}^*\| = 1. \tag{3.29}$$

Далее, так как все траектории $\widetilde{x}_\delta(t)$ равномерно ограничены, и в силу соотношения (3.15) справедливо включение

$$\Delta_\delta \widetilde{x}_\delta(t) \in a(\widetilde{x}_\delta(t)),$$

то все величины $\Delta_\delta \widetilde{x}_\delta(t)$ также равномерно ограничены. На основании предположения С получаем, что

$$\|a^*(y^*;\ (\widetilde{x}_\delta(t),\ \Delta_\delta \widetilde{x}_\delta(t)))\| \leqslant c\|y^*\|.$$

Далее, согласно предположению В, все множества $\partial g(\widetilde{x}_\delta(t),\ t)$ равномерно ограничены и из включения (3.21) теперь вытекает, что

$$\|x_\delta^*(t)\| \leqslant \|x_\delta^*(t+\delta)\| + \delta c\|\widetilde{x}_\delta^*(t+\delta)\| + \\ + 2\lambda_\delta \delta \gamma(\delta) + \lambda_\delta \delta c_1, \tag{3.30}$$

где $\gamma(\delta)$ определено соотношением (3.17), а число c_1 ограничивает нормы векторов из множеств $\partial g(\widetilde{x}_\delta(t),\ t)$. Если переписать неравенство (3.30) в виде

$$\|x_\delta^*(t)\| \leqslant (1+\delta c)\|x_\delta^*(t+\delta)\| + \delta(2\lambda_\delta\gamma(\delta) + c_1),$$

то нетрудно получить, что

$$\|x_\delta^*(t)\| \leqslant (1+\delta c)^{(1-t)/\delta}\|x_\delta^*(1)\| + \\ + \delta(2\lambda_\delta\gamma(\delta) + T)\frac{(1+\delta c)^{(1-t)/\delta} - 1}{\delta c} \leqslant \\ \leqslant (1+\delta c)^{1/\delta}\left[\|x_\delta^*(1)\| + \frac{2\lambda_\delta\gamma(\delta) + T}{c}\right],$$

откуда следует, что все траектории $x_\delta^*(t)$ равномерно ограничены. Но тогда из включения (3.21) вытекает, что все величины $\Delta_\delta x_\delta^*(t)$ равномерно ограничены. Поэтому ломаная $x_\delta^*(t)$, $t \in [0, 1]$, проведенная через точки $x_\delta^*(t)$, $t = 0, \delta, \ldots, 1$, удовлетворяет условию Липшица с константой, не зависящей от δ.

Итак, функции $x_\delta^*(t)$ равномерно ограничены и удовлетворяют условию Липшица, и поэтому из этой последовательности можно выбрать равномерно сходящуюся. Без ограничения общности можно считать, что $x_\delta^*(t) \to x^*(t)$ равномерно. Точно также в силу условия (3.29) можно считать, что $\lambda_\delta \to \lambda_0, x_{l\delta}^* \to x_l^*$.

Из предположений A и B следует, что

$$x^*(0) \in K_N^*(\widetilde{x}(0)), \quad x_l^* \in K_M^*(\widetilde{x}(1)),$$
$$x^*(1) + x_l^* \in \lambda_0 \partial\varphi_0(\widetilde{x}(1)). \tag{3.31}$$

Проанализируем теперь соотношение (3.21). Используя лемму V.3.3, его можно переписать в виде

$$-\Delta_\delta x_\delta^*(t) \in A^*(x_\delta^*(t+\delta);\ \widetilde{x}_\delta(t)), \tag{3.32}$$

$$\Delta_\delta \widetilde{x}(t) \in a(\widetilde{x}_\delta(t);\ x_\delta^*(t+\delta)), \tag{3.33}$$

$$t = 0, \delta, \ldots, 1-\delta.$$

Согласно только что доказанной лемме и лемме V.3.1, многозначные отображения $A^*(y^*, x)$ и $a(x; y^*)$ полунепрерывны сверху по включению. Поэтому, почти дословно повторяя рассуждения, приведенные при доказательстве теоремы 1.1, из равномерной сходимости $x_\delta^*(t) \to x^*(t)$, $\widetilde{x}_\delta(t) \to \widetilde{x}(t)$ и соотношений (3.32), (3.33) получаем, что почти всюду выполнены соотношения

$$-\dot{x}^*(t) \in A^*(x^*(t); \widetilde{x}(t)), \tag{3.34}$$

$$\dot{\widetilde{x}}(t) \in a(\widetilde{x}(t); x^*(t)). \tag{3.35}$$

Из включений (3.31), (3.34) и (3.35) с учетом формулы (V.3.2), определяющей $a(x; y^*)$, получаем все требуемые теоремой утверждения. Доказательство завершено.

5. Выпуклые дифференциальные включения. Вопрос о применимости теоремы 3.2 сводится к проверке предположений A, B и C. Покажем, что эти предположения выполнены, если входящие в задачу множества и функции выпуклы.

Теорема 3.3. *Пусть $\varphi_0(x)$ — непрерывная выпуклая функция, множества N и M — выпуклы, а отображение a выпукло, замкнуто и ограничено. Для того, чтобы траектория $\widetilde{x}(t)$, $t \in [0, 1]$, целиком лежащая внутри множества* $\operatorname{dom} a$, *минимизировала* $\varphi_0(x(1))$ *среди всех тра-*

екторий дифференциального включения $\dot{x}(t) \in a(x(t))$, *удовлетворяющих краевым условиям* $x(0) \in N$, $x(1) \in M$, *необходимо существование числа* $\lambda_0 \geqslant 0$, *вектора* x_l^* *и абсолютно непрерывной функции* $x^*(t)$, *не равных тождественно нулю и таких, что:*

1) $x^*(1) + x_l^* \in \lambda_0 \partial\varphi_0(\tilde{x}(1))$, $x_l^* \in K_M^*(\tilde{x}(1))$, $x^*(0) \in K_N^*(\tilde{x}(0))$;

2) $-\dot{x}^*(t) \in a^*(x^*(t); (\tilde{x}(t), \dot{\tilde{x}}(t)))$ *почти всюду на* $[0, 1]$;

3) $\dot{\tilde{x}}(t) \in a(\tilde{x}(t); x^*(t))$ *почти всюду на* $[0, 1]$.

Доказательство. В рассматриваемой задаче $g(x, t) \equiv 0$. По предположению теоремы $\tilde{x}(t) \in \operatorname{int} \operatorname{dom} a$ и поэтому эта функция принадлежит $\operatorname{int} \operatorname{dom} a$ вместе с некоторой ε-трубкой.

По лемме 3.1 предположение A выполнено в силу выпуклости множеств N и M и того, что семейства N_δ, M_δ можно определить при помощи выпуклых неравенств

$$N_\delta = \{x\colon\ d(x|N) \leqslant c\delta\},$$

$$M_\delta = \{x\colon\ d(x|M) \leqslant c\delta\}.$$

По теореме III.1.2 условия доказываемой теоремы гарантируют то, что $a(x)$ удовлетворяет условию Липшица. Теоремы III.2.3 и III.2.4 обеспечивают выполнение условий предположения C. Наконец, субдифференциал $\partial\varphi_0(x)$ непрерывной выпуклой функции равномерно ограничен и полунепрерывен сверху по включению. Этот факт легко устанавливается из определения субдифференциала и того, что непрерывная выпуклая функция удовлетворяет условию Липшица. Доказательство оставляем читателю в качестве простого упражнения. Таким образом, выполнено и предположение B.

Применим теорему 3.2. Ее утверждения 1 и 3 сразу же обеспечивают выполнение условий 1 и 3 теоремы 3.3. Кроме того, в силу леммы V.3.4 справедливо равенство

$$A^*(y^*;\ x) = a^*(y^*;\ (x,\ y)),$$

где y — любая точка из $a(x;\ y^*)$. Поэтому

$$A^*(x^*(t); \tilde{x}(t)) = a^*(x^*(t); (\tilde{x}(t), \dot{\tilde{x}}(t))),$$

и утверждение 2 теоремы 3.2 обеспечивает выполнение условия 2 теоремы 3.3.

Следствие. *Если выполнены предположения предыдущей теоремы, то условия* 2 *и* 3 *можно переписать в виде*

$$-\dot{x}^*(t) \in \partial_x W_a(\tilde{x}(t),\ x^*(t)), \tag{3.36}$$

$$\dot{\tilde{x}}(t) \in \partial_{y^*} W_a(\tilde{x}(t),\ x^*(t)), \tag{3.37}$$

где

$$\partial_x W_a(x,\ y^*) = \\ = \{x^*:\ W_a(x_1,\ y^*) - W_a(x,\ y^*) \geqslant \langle x_1 - x,\ x^*\rangle\},$$

$$\partial_{y^*} W_a(x,\ y^*) = \\ = \{u : W_a(x,\ y_1^*) - W_a(x,\ y^*) \leqslant \langle u,\ y_1^* - y^*\rangle\}.$$

Действительно, согласно теореме III.2.1 справедливо равенство

$$a^*(y^*;\ z) = \partial_x W_a(x,\ y^*),$$

если $z = (x,\ y)$, $y \in a(x;\ y^*)$. Кроме того, при доказательстве теоремы IV.6.3 было показано, что

$$\partial_{y^*} W_a(x,\ y^*) = a(x;\ y^*).$$

Поэтому утверждения 2 и 3 теоремы эквивалентны включениям (3.36) и (3.37).

Теорема 3.4. *Если выполнены предположения теоремы* 3.3 *и* $\lambda_0 > 0$, *то условия* 1—3 *теоремы* 3.3 *являются достаточными.*

Доказательство. Без ограничения общности можно считать, что $\lambda_0 = 1$. Пусть $x(t)$ — произвольная траектория, удовлетворяющая дифференциальному включению и краевым условиям. Тогда

$$\langle \dot{x}(t),\ x^*(t)\rangle \geqslant W_a(x(t),\ x^*(t)) \geqslant \\ \geqslant W_a(\tilde{x}(t),\ x^*(t)) + \langle x(t) - \tilde{x}(t),\ -x^*(t)\rangle,$$

так как $-\dot{x}^*(t) \in \partial_x W_a(\tilde{x}(t),\ x^*(t))$ согласно следствию теоремы 3.3. Поэтому

$$\frac{d}{dt}\langle x(t),\ x^*(t)\rangle = \langle \dot{x}(t),\ x^*(t)\rangle + \langle x(t),\ \dot{x}^*(t)\rangle \geqslant$$

$$\geqslant W_a(\tilde{x}(t), x^*(t)) + \langle \tilde{x}(t), \dot{x}^*(t)\rangle =$$
$$= \langle \dot{\tilde{x}}(t), x^*(t)\rangle + \langle \tilde{x}(t), \dot{x}^*(t)\rangle = \frac{d}{dt}\langle \tilde{x}(t), x^*(t)\rangle.$$

где использовано условие 3 теоремы 3.3. Проинтегрировав предыдущее неравенство, получим

$$\langle x(1), x^*(1)\rangle - \langle x(0), x^*(0)\rangle \geqslant$$
$$\geqslant \langle \tilde{x}(1), x^*(1)\rangle - \langle \tilde{x}(0), x^*(0)\rangle.$$

Но $x^*(0) \in K_N^*(\tilde{x}(0))$, так что

$$\langle x(0), \ x^*(0)\rangle \geqslant \langle \tilde{x}(0), \ x^*(0)\rangle.$$

Отсюда и из предыдущего неравенства вытекает, что

$$\langle x(1), \ x^*(1)\rangle \geqslant \langle \tilde{x}(1), \ x^*(1)\rangle. \tag{3.38}$$

В силу условия 1 теоремы 3.3 $x_l^* \in K_M^*$ $(\tilde{x}(1))$, т. е. выполняется неравенство

$$\langle x(1), x_l^*\rangle \geqslant \langle \tilde{x}(1), x_l^*\rangle, \tag{3.39}$$

так как $x(1) \in M$. Кроме того, теорема 3.3 показывает, что $x^*(1) + x_l^* \in \partial\varphi_0(\tilde{x}(1))$. Поэтому с использованием неравенств (3.38) и (3.39) получаем

$$\varphi_0(x(1)) - \varphi_0(\tilde{x}(1)) \geqslant \langle x(1) - \tilde{x}(1), x^*(1) + x_l^*\rangle =$$
$$= \langle x(1) - \tilde{x}(1), x^*(1)\rangle + \langle x(1) - \tilde{x}(1), x_l^*\rangle \geqslant 0.$$

Итак, для любой допустимой траектории, удовлетворяющей дифференциальному включению и краевым условиям, справедливо неравенство

$$\varphi_0(x(1)) \geqslant \varphi_0(\tilde{x}(1)),$$

что и доказывает оптимальность траектории $\tilde{x}(\cdot)$.

6. Примеры. Проиллюстрируем применение теорем 3.2 и 3.3 на некоторых конкретных задачах. Так как основную сложность, как правило, составляет вычисление сопряженного отображения A^* и проверка предположения С, то допустим, что множества N и M выпуклы, функция $g(x, t)$ тождественно равна нулю, а функция $\varphi_0(x)$ непрерывно дифференцируема. Таким образом, предположения А и В во всех нижеследующих примерах выполнены.

При проверке того факта, что отображение a удовлетворяет условию Липшица, будет полезна следующая лемма.

Лемма 3.5. *Пусть* $\varphi(x, y)$ — *непрерывная функция, выпуклая по* y *при фиксированном* x *и удовлетворяющая условию Липшица по* x:

$$|\varphi(x_1, y) - \varphi(x_2, y)| \leqslant L\|x_1 - x_2\|.$$

Пусть, кроме того, для каждого x *существует такое* $y(x)$, *что*

$$\varphi(x, \; y(x)) = \mathscr{E}(x) < 0$$

и $\mathscr{E}(x)$ — *непрерывная функция. Если множества*

$$a(x) = \{y\colon \; \varphi(x, \; y) \leqslant 0\}$$

равномерно ограничены на некотором связном компакте, то отображение $a(x)$ *удовлетворяет на нем условию Липшица.*

Доказательство. Если $y_1 \in a(x_1)$, т. е. $\varphi(x_1, y_1) \leqslant 0$, то при $\lambda \in [0, 1]$ получаем

$$\begin{aligned}\varphi(x_2, \; (1-\lambda)y_1 + \lambda y(x_2)) &\leqslant \\ &\leqslant (1-\lambda)\varphi(x_2, \; y_1) + \lambda\varphi(x_2, \; y(x_2)) = \\ = (1-\lambda)\varphi(x_1, \; y_1) + (1-\lambda)[\varphi(x_2, \; y_1) &- \varphi(x_1, \; y_1)] + \lambda\mathscr{E}(x_2) \leqslant \\ &\leqslant L\|x_2 - x_1\| + \lambda\mathscr{E}(x_2).\end{aligned}$$

Если теперь положить

$$\lambda = -\frac{L}{\mathscr{E}(x_2)}\|x_1 - x_2\|, \tag{3.40}$$

то из предыдущего неравенства вытекает соотношение

$$\varphi\left(x_2, \; y_1 - \frac{L}{\mathscr{E}(x_2)}\|x_1 - x_2\|(y(x_2) - y_1)\right) \leqslant 0.$$

Таким образом, любой точке $y_1 \in a(x_1)$ можно поставить в соответствие точку $y_2 \in a(x_2)$ по следующей формуле:

$$y_2 = y_1 - \frac{L}{\mathscr{E}(x_2)}\|x_1 - x_2\|(y(x_2) - y_1).$$

Так как $\mathscr{E}(x_2) < 0$, а $a(x)$ ограничены на компакте, то

$$\|y_2 - y_1\| \leqslant \frac{L}{\|\mathscr{E}(x_2)\|} \|y(x_2) - y_1\| \|x_1 - x_2\| \leqslant L_1 \|x_1 - x_2\| \tag{3.41}$$

для некоторого L_1. Кроме того, если норма $\|x_1 - x_2\|$ достаточно мала, то λ, определяемое формулой (3.40), лежит в промежутке $[0, 1]$, так что проведенные выкладки законны.

Итак, если только разность $x_1 - x_2$ достаточно мала, то каждой точке $y_1 \in a(x_1)$ можно поставить в соответствие точку $y_2 \in a(x_2)$, так что будет выполнено неравенство (3.41). Если учесть полную равноправность x_1 и x_2, то формула (3.41) показывает, что расстояние между множествами $a(x_1)$ и $a(x_2)$ не превосходит $L_1\|x_1 - x_2\|$, т. е. отображение a удовлетворяет локальному условию Липшица. Из локального условия Липшица с одной и той же константой L_1 следует глобальное условие на всем связном компакте.

Пример 3.1. Пусть $\varphi(x, y)$ — непрерывно дифференцируемая функция, удовлетворяющая условиям леммы 3.5. Так как

$$\operatorname{gf} a = \{z:\ \varphi(z) \leqslant 0\},$$

а $\varphi(z)$, $z = (x, y)$,— гладкая функция, то в соответствии с формулой (2.20) предыдущего параграфа получаем

$$a^*(y^*; z) =$$
$$= \{\lambda\varphi'_x(z):\ y^* = -\lambda\varphi'_y(z),\quad \lambda \geqslant 0,\quad \lambda\varphi(z) = 0\} \tag{3.42}$$

и при этом конусы

$$K_a(z) = \begin{cases} \{\bar{z}:\ \langle \bar{z}, \varphi'_z(z)\rangle < 0\}, & \text{если}\quad \varphi(z) = 0, \\ Z, & \text{если}\quad \varphi(z) < 0 \end{cases}$$

являются локальными шатрами к $\operatorname{gf} a$ в точке z. Рассмотрим зависимость локально сопряженного отображения от аргументов. Так как $a^*(y^*; z)$ не пусто лишь, если $y^* = -\lambda\varphi'_y(z)$, то требование ограниченности a^* эквивалентно тому, что существует такая константа c, что

$$\|\varphi'_x(z)\| \leqslant c\|\varphi'_y(z)\|.$$

Если $\varphi'_y(z) \neq 0$ в некотором компактном множестве, то в силу непрерывности $\varphi'_y(z)$ и $\varphi'_x(z)$ отношение

$$\frac{\|\varphi'_x(z)\|}{\|\varphi'_y(z)\|}$$

ограничено. Таким образом, если $\varphi'_y(z) \neq 0$, то a^* — равномерно ограниченное отображение. Нетрудно проверить, что при этом условии оно полунепрерывно сверху (даже непрерывно) зависит от своих аргументов.

Таким образом, предположение C выполнено. Допустим, что множество $a(x;\ y^*)$ состоит из одной точки. Это будет справедливо, если функция $\varphi(x,\ y)$ строго выпукла по y, т. е.

$$\varphi(x,\ \lambda_1 y_1 + \lambda_2 y_2) < \lambda_1 \varphi(x,\ y_1) + \lambda_2 \varphi(x,\ y_2),$$

$$\lambda_1, \lambda_2 > 0, \quad \lambda_1 + \lambda_2 = 1.$$

При таком предположении $A^*(\dot{x}^*;\ \tilde{x}) = a(x^*;\ (\tilde{x},\ \dot{\tilde{x}}))$, так как в силу утверждения 3 теоремы 3.2 $\dot{\tilde{x}}$ и есть тот единственный элемент из $a(\tilde{x})$, который минимизирует $\langle y,\ x^* \rangle$.

Теперь утверждение 2 теоремы 3.2 можно переписать в виде

$$-\dot{x}^* \in a^*(x^*;\ (\tilde{x},\ \dot{\tilde{x}}))$$

или, учитывая выражение (3.42),

$$\begin{aligned} -\dot{x}^*(t) &= \lambda(t)\, \varphi'_x(\tilde{x}(t),\ \dot{\tilde{x}}(t)), \\ x^*(t) &= -\lambda(t)\, \varphi'_y(\tilde{x}(t),\ \dot{\tilde{x}}(t)), \\ \lambda(t)\, \varphi(\tilde{x}(t),\ &\dot{\tilde{x}}(t)) = 0. \end{aligned} \tag{3.43}$$

Из этих соотношений следует

$$\begin{gathered} \frac{d}{dt}\lambda(t)\, \varphi'_y(\tilde{x}(t),\ \dot{\tilde{x}}(t)) = \lambda(t)\, \varphi'_x(\tilde{x}(t),\ \dot{\tilde{x}}(t)), \\ \lambda(t)\, \varphi(\tilde{x}(t),\ \dot{\tilde{x}}(t)) = 0. \end{gathered} \tag{3.44}$$

Подведем некоторые итоги.

Теорема 3.5. *Пусть $\varphi(x, y)$ — непрерывно дифференцируемая функция, строго выпуклая по y, и множества $\{y: \varphi(x, y) \leqslant 0\}$ равномерно ограничены в ε-трубке траектории $\tilde{x}(t)$, $t \in [0, 1]$. Кроме того, существует такой вектор $y(x)$ для x из ε-трубки, что*

$$\mathscr{E}(x) = \varphi(x, y(x))$$

— непрерывная строго отрицательная функция.

Тогда для того, чтобы траектория $\tilde{x}$ минимизировала $\varphi_0(x(1))$ среди всех траекторий, удовлетворяющих соотношению

$$\varphi(x(t), \dot{x}(t)) \leqslant 0$$

и краевым условиям $x(0) \in N$, $x(1) \in M$, необходимо существование функции $\lambda(t)$, числа λ_0 и вектора x_1^*, не равных нулю одновременно и таких, что

$$\begin{aligned} x_1^* - \lambda(1)\varphi_y'\left(\tilde{x}(1), \dot{\tilde{x}}(1)\right) = \lambda_0\varphi_0'\left(\tilde{x}(1)\right), \\ -\varphi_y'\left(\tilde{x}(0), \dot{\tilde{x}}(0)\right) \in K_N^*\left(\tilde{x}(0)\right), \quad x_1^* \in K_M^*\left(\tilde{x}(1)\right), \end{aligned} \tag{3.45}$$

и выполнены соотношения (3.44).

Заметим, что соотношения (3.45) есть просто утверждение 1 теоремы 3.2, переписанное с использованием соотношения (3.43).

Пример 3.2. Пусть $\varphi(x, y)$ непрерывная функция, выпуклая по совокупности аргументов. Тогда отображение

$$a(x) = \{y: \varphi(x, y) \leqslant 0\} \tag{3.46}$$

выпукло и замкнуто. Если хотя бы для одного x множество $a(x)$ ограничено, то отображение a ограничено согласно лемме III.1.1. Если траектория $\tilde{x}(t)$, $t \in [0, 1]$, целиком лежит в $\operatorname{int}\operatorname{dom} a$, то применима теорема 3.3. Чтобы записать конкретный вид необходимых условий, надо вычислить локально сопряженное отображение к отображению a. Но согласно теореме III.3.2

$$a^*(y^*; z) =$$
$$= \{\lambda x_0^*: y^* = -\lambda y_0^*, \ (x_0^*, y_0^*) \in \partial\varphi(z), \ \lambda \geqslant 0, \ \lambda\varphi(z) = 0\},$$

если только существует такая точка z_1, что $\varphi(z_1) < 0$.

Поэтому условие 2 теоремы 3.3 может быть записано в виде

$$-\dot{x}^*(t) = \lambda(t)\, x_0^*(t); \quad x^*(t) = -\lambda y_0^*(t);$$
$$\left(x_0^*(t),\, y_0^*(t)\right) \in \partial\varphi\left(\tilde{x}(t),\, \dot{\tilde{x}}(t)\right)$$

и при этом

$$\lambda(t)\varphi(\tilde{x}(t),\, \dot{\tilde{x}}(t)) = 0, \qquad \lambda(t) \geqslant 0,$$

или, по-иному,

$$-(\dot{x}^*(t),\, x^*(t)) \in \lambda(t)\partial\varphi(\tilde{x}(t),\, \dot{\tilde{x}}(t)).$$

Таким образом, применение теоремы 3.3 для случая, когда отображение a задается при помощи формулы (3.46), приводит к следующему результату.

Теорема 3.6. *Пусть $\varphi(x, y)$ — непрерывная выпуклая по совокупности аргументов функция, множество $a(x)$ ограничено хотя бы для одного значения x, и существует такая точка $z_1 = (x_1, y_1)$, что $\varphi(x_1, y_1) < 0$.*

Тогда для того, чтобы траектория $\tilde{x}(t)$, $t \in [0, 1]$, целиком лежащая в $\operatorname{int}\operatorname{dom} a$, *минимизировала $\varphi_0(x(1))$ среди всех траекторий, удовлетворяющих неравенству*

$$\varphi(x(t),\, \dot{x}(t)) \leqslant 0$$

и краевым условиям $x(0) \in N$, $x(1) \in M$, необходимо, чтобы нашлась абсолютно непрерывная функция $x^*(t)$, функция $\lambda(t)$, вектор x_l^* и число $\lambda_0 \geqslant 0$ такие, что

$$-(\dot{x}^*(t),\, x^*(t)) \in \lambda(t)\partial\varphi(\tilde{x}(t),\, \dot{\tilde{x}}(t)),$$
$$\lambda(t) \geqslant 0, \quad \lambda(t)\varphi(\tilde{x}(t),\, \dot{\tilde{x}}(t)) = 0,$$
$$x^*(0) \in K_N^*\left(\tilde{x}(0)\right), \quad x_l^* \in K_M^*\left(\tilde{x}(1)\right),$$
$$x^*(1) + x_l^* = \lambda_0\varphi_0'\left(\tilde{x}(1)\right).$$

При этом $x^*(\cdot)$, x_l^* и λ_0 одновременно в нуль не обращаются.

Пример 3.3. Пусть

$$a(x) = \mathscr{A}x + U,$$

где $\mathscr{A}$ — $n\times n$-матрица, а U — выпуклое замкнутое множество в $\mathbf{R}^n$. Многозначное отображение a очевидным образом выпукло, замкнуто и ограничено. Согласно примеру III.3.2, если $z=(x,\ y)$ и $y=\mathscr{A}x+u$, то

$$a^*(y^*;\ z)=\begin{cases}\mathscr{A}^*y^*, & \text{если}\quad y^*\in[\operatorname{con}(U-u)]^*,\\ \varnothing, & \text{если}\quad y^*\notin[\operatorname{con}(U-u)]^*.\end{cases}$$

Поэтому условие 2 теоремы 3.3 может быть переписано в следующем виде:

$$-\dot{x}^*(t)=\mathscr{A}^*x^*(t), \tag{3.47}$$

$$\dot{\tilde{x}}(t)=\mathscr{A}\tilde{x}(t)+\tilde{u}(t),$$

$$x^*(t)\in[\operatorname{con}(U-\tilde{u}(t))]^*, \tag{3.48}$$

где использован тот факт, что выполнение дифференциального включения $\dot{\tilde{x}}\in a(\tilde{x}(t))$ в рассматриваемом случае означает, что

$$\dot{\tilde{x}}=\mathscr{A}\tilde{x}(t)+\tilde{u}(t),\qquad \tilde{u}(t)\in U.$$

Заметим теперь, что согласно определению сопряженного конуса и того, что

$$\operatorname{con}(U-\tilde{u}(t))=\{\lambda(u-\tilde{u}(t)):\ u\in U,\quad \lambda>0\},$$

соотношение (3.48) эквивалентно неравенству

$$\langle u-\tilde{u}(t),\ x^*(t)\rangle\geqslant 0,\quad u\in U. \tag{3.49}$$

Из полученных результатов и теоремы 3.3 вытекает справедливость следующей теоремы.

Теорема 3.6. *Для того, чтобы траектория $\tilde{x}(t)$ и соответствующее ей управление $\tilde{u}(t)$ минимизировали $\varphi_0(x(1))$ среди всех траекторий и управлений, удовлетворяющих уравнению*

$$\dot{x}(t)=\mathscr{A}x(t)+u(t),\quad u(t)\in U,$$

и краевым условиям $x(0)\in N$, $x(1)\in M$, необходимо, чтобы существовала функция $x^*(t)$, вектор x_1^* и число $\lambda_0\geqslant 0$

такие, что

$$-\dot{x}^*(t) = \mathscr{A}^* x^*(t),$$
$$\langle u, x^*(t)\rangle \geqslant \langle \tilde{u}(t), x^*(t)\rangle, \qquad u \in U,$$
$$x^*(0) \in K_N^*\left(\tilde{x}(0)\right), \quad x_l^* \in K_M^*\left(\tilde{x}(1)\right),$$
$$x^*(1) + x_l^* = \lambda_0 \varphi_0'\left(\tilde{x}(1)\right).$$

Пример 3.4. Допустим, что многозначное отображение a задается выпуклым конусом $K \subseteq \mathbf{R}^n \times \mathbf{R}^n$:

$$a(x) = \{y\colon\ (x,\ y) \in K\}. \tag{3.50}$$

Если $a(x) = \{0\}$, то согласно лемме III.1.2 a — ограниченное отображение, так что при исследовании оптимизационных задач для дифференциального включения

$$\dot{x}(t) \in a(x(t)),$$

где a задается формулой (3.50), или, что то же самое, формулой

$$(x(t),\ \dot{x}(t)) \in K, \tag{3.51}$$

применима теорема 3.3. В соответствии с примером III.3.8

$$a^*(y^*;\ z) = \{x^*\colon\ (-x^*,\ y^*) \in K^*, \quad \langle x,\ x^*\rangle = \langle y,\ y^*\rangle\}.$$

Отсюда вытекает, что условие 2 теоремы 3.3 можно в данном случае записать в виде

$$\left.\begin{aligned} &(\dot{x}^*(t),\ x^*(t)) \in K^*,\\ &\langle \dot{\tilde{x}}(t),\ x^*(t)\rangle = \langle \tilde{x}(t),\ \dot{x}^*(t)\rangle. \end{aligned}\right\} \tag{3.52}$$

Сформулируем полученный результат.

Теорема 3.7. *Пусть K — выпуклый замкнутый конус и не существует $y \neq 0$ такого, что $(0,\ y) \in K$. Тогда для того, чтобы траектория $\tilde{x}(t)$, $t \in [0,\ 1]$, целиком лежащая в* $\operatorname{int}\operatorname{dom} a$, *минимизировала $\varphi_0(x(1))$ среди всех траекторий, удовлетворяющих соотношению*

$$(x(t),\ \dot{x}(t)) \in K$$

и краевым условиям $x(0)\in N$, $x(1)\in M$, *необходимо, чтобы нашлись абсолютно непрерывная функция* $x^*(t)$, *вектор* x_l^* *и число* $\lambda_0\geqslant 0$, *не равные одновременно нулю и такие, что*

$$(\dot{x}^*(t),\ x^*(t))\in K^*,$$

$$\langle\dot{\tilde{x}}(t),\ x^*(t)\rangle=\langle\tilde{x}(t),\ \dot{x}^*(t)\rangle$$

и

$$x^*(0)\in K_N^*\left(\tilde{x}(0)\right),\quad x_l^*\in K_M^*\left(\tilde{x}(1)\right),$$

$$x^*(1)+x_l^*=\lambda_0\varphi_0'\left(\tilde{x}(1)\right).$$

БИБЛИОГРАФИЧЕСКИЙ КОММЕНТАРИЙ

Список литературы по линейному программированию, выпуклому анализу, общему математическому программированию, теории оптимального управления и их применения насчитывает сотни наименований. Обзор всей этой литературы сам по себе должен был бы составить целую книгу. Поэтому здесь мы ограничимся только указаниями монографий и статей обзорного характера, в которых можно найти тот или иной результат, приведенный в этой книге. Такие указания дадут возможность читателю в случае необходимости более детально ознакомиться с результатом и рассмотреть его обобщения и применения в различных задачах. Эти монографии содержат также дополнительные ссылки на литературу.

Ссылки на статьи, не носящие обзорного характера, даются только в случае, если рассматриваемый результат нельзя найти в монографической литературе.

Исчерпывающему изложению теории выпуклых множеств для конечномерных пространств посвящена книга Р. Т. Рокафеллара [4]. Книга содержит также большой список литературы, посвященной выпуклому анализу. Различные свойства выпуклых множеств рассматривались также в книгах И. В. Гирсанова [1], А. Д. Иоффе и В. М. Тихомирова [2], С. Карлина [1], П. Ж. Лорана [1], К.— Г. Эльстера, Р. Рейнхардта, М. Шойбле, Г. Доната [1]. Общие теоремы отделимости и их связь с теоремой Хана — Банаха изложены в книге Н. Данфорда и Дж. Т. Шварца [1]. Теоремы отделимости для общей системы конусов в конечномерном пространстве получены В. Г. Болтянским [3].

Теория многогранных множеств и систем линейных неравенств подробно рассмотрена в книгах Д. Гейла [1], Дж. Данцига [1], Г. Куна и А. Таккера (редакторы) [1], Р. Т. Рокафеллара [4].

Выпуклые функции детально изучались в книге Р. Т. Рокафеллара [4]. Двойственность выпуклых функций и свойства субдифференциалов наиболее часто используются в вопросах оптимизации.

Двойственности выпуклых функций и ее применениям посвящена статья А. Д. Иоффе и В. М. Тихомирова [1]. Различные теоремы о свойствах субдифференциалов и способах их вычисления приведены в книгах И. В. Гирсанова [1], Е. Г. Гольштейна [2], А. Д. Иоффе и В. Л. Левина [1], А. Д. Иоффе и В. М. Тихомирова [1, 2], Б. Н. Пшеничного [1], Р. Т. Рокафеллара [4].

Крупный вклад в теорию выпуклых множеств и функций внес французский математик Моро (Moreau G. G.). Подробные ссылки на его работы и труды руководимого им семинара можно найти в монографиях А. Д. Иоффе и В. М. Тихомирова [2], П. Ж. Лорана [1], Р. Т. Рокафеллара [4].

Многозначные отображения стали в последние годы предметом интенсивного изучения. Различные аналитические свойства многозначных отображений и их связь с теорией оптимизации рассмотрены в книге А. Д. Иоффе и В. М. Тихомирова [2] и в статье В. И. Аркина и В. Л. Левина [1]. Там же имеются ссылки на многочисленные работы французских математиков, посвященные этим вопросам, в частности, на работы Валадье (Voladier M.) и Кастена (Costaing Ch.). В связи с теорией экономических моделей многозначные отображения и связанные с ними экстремальные задачи исследовались в работах В. Л. Макарова и А. М. Рубинова [1, 2]. В статье А. М. Рубинова [1] рассмотрена связь многозначных отображений с различными вопросами функционального анализа. Изложение теории многозначных отображений, приведенное в главе III данной книги, основано на работах В. В. Береснева [1], В. В. Береснева и Б. Н. Пшеничного [1]. Некоторые из приведенных в главе III результатов ранее не публиковались.

Теория линейного программирования и его приложений изложена в многочисленных монографиях, в частности, в книгах Д. Гейла [1], Дж. Б. Данцига [1], С. Карлина [1], Г. Куна и А. Таккера (редакторы) [1].

Различные вопросы выпуклого программирования, в частности необходимые условия экстремума, содержатся в монографиях Е. Г. Гольштейна [1, 2], К. Карлина [1], Б. Н. Пшеничного [1], Р. Т. Рокафеллара [4]. Много места проблемам двойственности в выпуклом программировании уделено в книгах Е. Г. Гольдштейна [1, 2], Р. Т. Рокафеллара [4] и в статье А. Д. Иоффе и В. М. Тихомирова [1].

Теория наилучшего приближения функции уже стала классическим объектом применения выпуклого программирования. Эти применения к самым различным вопросам теории аппроксимации можно найти в книгах Е. Г. Гольштейна [2], В. Ф. Демьянова и В. Н. Малоземова [1], П. Ж. Лорана [1], Б. Н. Пшеничного [1], В. М. Тихомирова [1], а также в статье А. Д. Иоффе и В. М. Тихомирова [1].

Использование идей выпуклого программирования применительно к решению проблемы моментов, близко связанной с теорией аппроксимации, рассматривалось в книгах М. Г. Крейна и А. А. Нудельмана [1], Б. Н. Пшеничного [1]. Многочисленные связи между общей теорией выпуклого программирования и методами стохастического программирования отражены в монографии Ю. М. Ермольева [1].

Математическим моделям экономической динамики посвящена обширная литература. Укажем здесь лишь на монографии Х. Никайдо [1] и В. Л. Макарова и А. М. Рубинова [2], где читатель найдет дальнейшие ссылки.

Побудительной причиной современного развития общей теории необходимых условий экстремума было создание математической теории оптимального управления, нашедшее свое первое законченное изложение в монографии ее создателей Л. С. Понтрягина, В. Г. Болтянского, Р. В. Гамкрелидзе, Е. Ф. Мищенко [1].

Современной теории оптимального управления посвящена обширная монографическая литература, среди которой отметим здесь книги Дж. Варги [1], Н. Н. Красовского [1], Н. Н. Моисеева [1]. В первой рассмотрены различные обобщения задачи оптимального управления. В третьей уделено большое внимание численным методам расчета и проблеме оптимального управления.

В книге Н. Н. Красовского [1] проведен глубокий анализ линейных систем управления. Связь между классическим вариационным исчислением и теорией оптимального управления подробно рассматривается в книгах М. Хестена [1] и Л. Янга [1].

Конус касательных направлений является одним из основных объектов в теории необходимых условий экстремума. Различные определения конуса касательных направлений, часто эквивалентные или близкие, можно найти в монографиях И. В. Гирсанова [1], В. Ф. Демьянова и А. М. Рубинова [1], П. Ж. Лорана [1]. Однако этого понятия часто бывает не достаточно при рассмотрении задач вариационного исчисления. В связи с этим Л. Нейштадтом [4], М. Хестенсом [1] были введены конусы касательных направлений, обладающие более тонкими свойствами. Близкое к их определениям понятие используется в § 4 главы V этой книги. Понятие локального шатра было введено В. Г. Болтянским [1—3]. Он исследовал основные свойства шатров и интенсивно использовал его для решения различных экстремальных задач.

Свойства негладких функций применительно к экстремальным задачам изучались в работах А. Я. Дубовицкого и А. А. Милютина [1], А. Д. Иоффе и В. М. Тихомирова [2], Ф. Кларка [1], [2], Е. А. Нурминского [1], Б. Н. Пшеничного [1, 4]. Содержание § 2 главы V основывается на статье Б. Н. Пшеничного [1]. Следует отметить, что понятия, аналогичные понятию функции, допускающей верхнюю выпуклую аппроксимацию, по существу уже содержались в работах Л. Нейштадта [1, 2] и Г. Халкина и Л. Нейштадта [1]. Детально дифференциальные свойства негладких функций применительно к задачам оптимизации рассмотрены в монографии Л. Нейштадта [3]. Однако понятие субдифференциала им не было введено. Для так называемых квазидифференцируемых функций оно было введено Б. Н. Пшеничным [1], для функций, удовлетворяющих условию Липшица,— Ф. Кларком [1—3], а для функций, допускающих верхнюю выпуклую аппроксимацию,— Б. Н. Пшеничным [4].

Различные общие схемы получения необходимых условий экстремума были развиты Р. В. Гамкрелидзе [1, 2], Р. В. Гамкрелидзе и Г. Л. Харатишвили [1], А. Я. Дубовицким и А. А. Милютиным [1], Л. Нейштадтом [3], Г. Халкиным и Л. Нейштадтом [1]. Обратим также внимание на близкие к указанным работы Г. Халкина [1, 2].

В работе И. Екелана [1] приводится интересная теорема, при помощи которой удается охарактеризовать точки, близкие к экстремальным. Используя эту теорему, Ф. Кларк [3] сформулировал необходимые условия экстремума для задач, в которых имеются ограничения типа равенства, задаваемые негладкими функциями и множеством, для которого не предполагается существование конуса касательных направлений.

Дифференциальные включения в последние годы привлекают к себе внимание многих математиков. В настоящее время получены тонкие результаты по теоремам существования решений для дифференциальных включений. Ряд таких теорем содержится в работах А. Ф. Филиппова [1—3]. Подробный обзор полученных результатов можно найти в статье В. И. Благодатских [2].

Задачам оптимального управления, когда имеются ограничения, связывающие в один и тот же момент времени фазовые координаты и управление, посвящена значительная литература. Решению этой задачи в различных постановках посвящены работы А. Я. Дубовицкого и А. А. Милютина [2, 3], А. Н. Дюкалова [1], А. Н. Дюкалова и А. Е. Илютовича [1], К. Маковского и Л. Нейштадта [1], Л. Нейштадта [3].

В работах Р. Т. Рокафеллара [1—3] подробно исследован случай, когда задача выпуклая. А. М. Тер-Крикоров [1] исследовал задачу при линейных ограничениях.

Экстремальные задачи для дифференциальных включений, тесно связанные с задачами оптимального управления со смешанными ограничениями, изучались в работах В. И. Благодатских [1], В. Г. Болтянского [3], Ф. Кларка [1, 4], Б. Н. Пшеничного [3], Р. П. Федоренко [1]. Задачи с дискретным временем подробно рассмотрены В. Г. Болтянским [1, 3], Л. Нейштадтом [3], А. Н. Пропоем [1].

Изложенные выше указания должны дать читателю возможность ориентироваться в современной литературе по оптимизации и ни в коей мере не претендуют на полноту или историческую последовательность. Более полные обзоры литературы с освещением истории вопроса можно найти в книгах А. Д. Иоффе и В. М. Тихомирова [2], Л. Нейштадта [3], Р. Т. Рокафеллара [4].

ЛИТЕРАТУРА

Аркин В. И., Левин В. Л.

1. Выпуклость значений векторных интегралов, теоремы измеримого выбора и вариационные задачи.— УМН, 1972, **27**, 3, с. 21—77.

Береснев В. В.

1. Об отображениях, сопряженных к выпуклым многозначным отображениям.— Кибернетика, 1973, № 5, с. 79—83.
2. О необходимых условиях оптимальности для систем дискретных включений.— Кибернетика, 1977, № 2, с. 58—64.

Береснев В. В., Пшеничный Б. Н.

1. О дифференциальных свойствах функции минимума. ЖВМ и МФ, 1974, **14**, 3, с. 639—651.

Благодатских В. И.

1. Достаточные условия оптимальности для дифференциальных включений.— Изв. АН СССР, сер. матем., 1974, **33**, 3, с. 615—624.
2. Некоторые результаты по теории дифференциальных включений.— Summer School on Ordinary Differential Equations, Brno, Part 1 (1975), 29—67.

Болтянский В. Г.

1. Оптимальное управление дискретными системами.— М.: Наука, 1973.
2. О пересечении локальных шатров.— ДАН СССР, 1974, **219**, 5, с. 1042—1044.
3. Метод шатров в теории экстремальных задач.— УМН, 1975, **30**, 3, с. 1—55.

Бурбаки Н.

1. Топологические векторные пространства.— М.: ИЛ, 1959.

Варга Дж.

1. Оптимальное управление дифференциальными и функциональными управлениями.— М.: Наука, 1977, с. 624.

Гамкрелидзе Р. В.

1. Extremal problems in finite-dimensional spaces.— J. Opt. Theory Appl., 1967 **1**, p. 173—193.
2. Необходимые условия первого порядка и аксиоматика экстремальных задач.— Труды МИАН СССР, 1971, **112**, с. 152—180.

Гамкрелидзе Р. В., Харатишвили Г. Л.

1. Экстремальные задачи в линейных топологических пространствах.— Изв. АН СССР, сер. матем., 1969, **33**, 4, с. 781—839.

Гейл Д.

1. Теория линейных экономических моделей.— М.: ИЛ, 1963.

Гирсанов И. В.
1. Математическая теория экстремальных задач.— М.: Изд-во МГУ, 1970.
Гольштейн Е. Г.
1. Выпуклое программирование. Элементы теории.— М.: Наука, 1970.
2. Теория двойственности в математическом программировании.— М.: Наука, 1971.
Данфорд Н., Шварц Дж. Т.
1. Линейные операторы. Общая теория.— М.: ИЛ, 1964.
Данциг Дж. Б.
1. Линейное программирование, его приложения и обобщения.— М., Прогресс, 1966.
Демьянов В. Ф., Рубинов А. М.
1. Приближенные методы решения экстремальных задач.— Л.: Изд-во ЛГУ, 1968.
Демьянов В. Ф., Малоземов В. Н.
1. Введение в минимакс.— М.: Наука, 1972.
Дубовицкий А. Я., Милютин А. А.
1. Задачи на экстремум при наличии ограничений.— ЖВМ и МФ, 1965, **5**, 3, с. 395—453.
2. Необходимые условия экстремума в задачах оптимального управления со смешанными ограничениями типа неравенств.— ЖВМ и МФ, 1968, **8**, 4, с. 725—770.
3. Необходимые условия экстремума в общей задаче оптимального управления.— М.: Наука, 1971.
Дюкалов А. Н.
1. Признак оптимальности в линейных динамических задачах оптимального управления со смешанными ограничениями.— ЖВМ и МФ, 1976, **16**, 4, с. 856—873.
Дюкалов А. Н., Илютович А. Е.
1. Признак оптимальности в нелинейных задачах оптимального управления со смешанными ограничениями. I, II.— Автоматика и телемеханика, 1977, № 3, с. 96—106; 1977, № 5, с. 11—20.
Дьедонне Ж.
1. Основы современного анализа.— М.: Мир, 1964.
Екелан И. (Ekeland I.)
1. On the Variational Principle.— J. Math. Anal. Appl., 1974, **47**, 2, p. 324—353.
Ермольев Ю. М.
1. Методы стохастического программирования.— М.: Наука, 1976.
Иоффе А. Д., Левин В. Л.
1. Субдифференциалы выпуклых функций, Труды ММО, 1972, **26**, с. 3—73.
Иоффе А. Д., Тихомиров В. М.
1. Двойственность выпуклых функций и экстремальные задачи.— УМН, 1968, **23**, 6, с. 51—116.
2. Теория экстремальных задач.— М.: Наука, 1974.
Карлин С.
1. Математические методы в теории игр, программировании и экономике.— М.: ИЛ, 1964.

Кларк Ф. (Clarke F. H.)
1. Necessary conditions for nonsmooth problems in optimal control and the colculus of variations. Ph. D. thesis.— Seattle, Washington: University of Washington, 1973.
2. Generalized gradients and applications.— Trans. Amer. Math. Soc., 1975, 205, p. 247—262.
3. A new approach to Lagrange multipliers.— Mathematics of operations research, 1976, 1, 2, p. 165—174.
4. The generalized problem of Bolza.— SIAM J. Control and optimization, 1976, 14, 4, p. 682—699.

Красовский Н. Н.
1. Теория управления движением.— М.: Наука, 1968.

Колмогоров А. Н., Фомин С. В.
1. Элементы теории функций и функциональный анализ.— М.: Наука, 1968.

Крейн М. Г., Нудельман А. А.
1. Проблема моментов Маркова и экстремальные задачи.— М.: Наука, 1973.

Лоран П. Ж.
1. Аппроксимация и оптимизация.— М.: Мир, 1975.

Кун Г., Таккер А. (ред.)
1. Линейные неравенства и смежные вопросы.— М.: ИЛ, 1959.

Кутателадзе С. С., Рубинов А. М.
1. Двойственность Минковского и ее приложение.— Новосибирск: Наука, 1976.

Макаров В. Л., Рубинов А. М.
1. Суперлинейные точечно-множественные отображения и модели экономической динамики.— УМН, 1970, 25, 5, с. 126—169.
2. Математическая теория экономической динамики и равновесия.— М.: Наука, 1973.

Маковски К., Нейштадт Л. (Makovski K., Neustadt L.)
1. Optimal control problems with mixed control-phase variable equality and inequality constraints.— SIAM J. Control, 1974, 12, 2, p. 184—228.

Моисеев Н. Н.
1. Численные методы в теории оптимальных систем.— М.: Наука, 1971.

Нейштадт Л. (Neustadt L.)
1. An abstract variational theory with applications to a broad class of optimization problems. I: General theory.— SIAM J. Control, 1966, 4, p. 505—527.
2. An abstract variational theory with applications to a broad class of optimization problems. II: Applications.— SIAM J. Control, 1967, 5, p. 90—137.
3. Optimization: A Theory of Necessary Conditions.— Princeton University Press., 1976.

Никайдо Х.
1. Выпуклые структуры и математическая экономика.— М.: Мир, 1972.

Нурминский Е. А.
1. Квазиградиентный метод решения задачи нелинейного программирования.— Кибернетика, 1973, № 1, с. 122—125.
Понтрягин Л. С., Болтянский В. Г., Гамкрелидзе Р. В., Мищенко Е. Ф.
1. Математическая теория оптимальных процессов.— М.: Физматгиз, 1961.
Пропой А. И.
1. Элементы теории оптимальных дискретных процессов.— М.: Наука, 1973.
Пшеничный Б. Н.
1. Необходимые условия экстремума.— М.: Наука, 1969.
2. Выпуклые многозначные отображения и им сопряженные.— Кибернетика, 1972, № 3, с. 94—102.
3. Необходимые условия экстремума для дифференциальных включений.— Кибернетика, 1976, № 6, с. 60—73.
4. О необходимых условиях экстремума для негладких функций.— Кибернетика, 1977, № 6, с. 92—96.
Рокафеллар Р. Т. (Rockafellar R. T.)
1. Existence and duality theorems for convex problem of Bolza.— Trans. Amer. Math. Soc., 1971, **159**, 1, p. 1—40.
2. State constraints in convex problems of Bolza.— SIAM J. Control, 1972, **10**, 4, p. 691—715.
3. Optimal arcas and the minimum value function in problems of Lagrange.— Trans. Amer. Math. Soc., 1973, **180**, p. 53—84.
4. Выпуклый анализ.— М.: Мир, 1973.
Рубинов А. М.
1. Сублинейные операторы и их приложения.— УМН, 1977, **32**, 4, с. 113—174.
Тер-Крикоров А. М.
1. Оптимальное управление и математическая экономика.— М.: Наука, 1977.
Тихомиров В. М.
1. Некоторые вопросы теории приближений.— М.: Изд-во МГУ, 1976.
Федоренко Р. П.
1. Принцип максимума для дифференциальных включений.— ЖВМ и МФ, 1971, **11**, 4, с. 885—893.
Филиппов А. Ф.
1. Дифференциальные уравнения с разрывной правой частью.— Матем. сборник, 1960, **51**, 1, с. 99—128.
2. Дифференциальные уравнения с многозначной разрывной правой частью.— ДАН СССР, 1963, **151**, 1, с. 65—68.
3. Классические решения дифференциальных уравнений с многозначной правой частью.— Вестник МГУ, сер. матем. и мех., 1967, 3, с. 16—26.
Халкин Г. (Halkin H.)
1. A satisfactory treatment of equality and operator constraints in the Dubovitski — Milutin optimization formalism.— JOTA, 1970, **6**, p. 138—149.

2. A maximum principle of the Pontryagin type for systems described by nonlinear difference equations.— SIAM J. Control, 1966, 4, p. 90—111.

Халкин Г., Нейштадт А. (Halkin H., Neustadt L.)

1. General necessary conditions for optimization problems.— Proc. Nat. Acad. Sci. USA, 1966, 56, p. 1066—1071.

Хестенс М. Р. (Hestenes M. R.)

1. Calculus of Variations and optimal Control Theory.— New York: John Wiley and Sons, 1966.

Эльстер К.-Г., Рейнхардт Р., Шойбле М., Донат Г. (Elster K. H., Reinhardt R., Schauble M., Donath G.)

1. Einfuhrung in die nichtlineare Optimierung.— BSB B. G. Teubner Verlagsgesellschaft, 1977.

Янг Л.

1. Лекции по вариационному исчислению и теории оптимального управления.— М.: Мир, 1974.

ПРЕДМЕТНЫЙ УКАЗАТЕЛЬ

УКАЗАТЕЛЬ ОСНОВНЫХ ОБОЗНАЧЕНИЙ

Борис Николаевич Пшеничный

ВЫПУКЛЫЙ АНАЛИЗ
И ЭКСТРЕМАЛЬНЫЕ ЗАДАЧИ

(Серия: «Нелинейный анализ и его приложения»)

М., 1980 г., 320 стр.

Редакторы *Е. Ю. Ходан, Е. В. Шикин*

Техн. редактор *С. Я. Шкляр*

Корректоры *Н. Д. Дорохова,*
Е. В. Сидоркина

ИБ № 11552

Сдано в набор 09.08.79. Подписано к печати 02.01.80. Т-01101. Бумага $84\times108^1/_{32}$. Тип. № 1. Обыкновенная гарнитура. Высокая печать. Условн. печ. л. 16,8. Уч.-изд. л. 16,23. Тираж 7000 экз. Заказ № 644. Цена книги 2 р. 10 к.

Издательство «Наука»
Главная редакция физико-математической литературы
117071, Москва, В-71, Ленинский проспект, 15

4-я типография издательства «Наука»
630077, Новосибирск, 77, Станиславского, 25