

ОПТИМИЗАЦИЯ
И ИССЛЕДОВАНИЕ
О П Е Р А Ц И Й



В. Ф. ДЕМЬЯНОВ
Л. В. ВАСИЛЬЕВ

Недифференцируемая оптимизация





ОПТИМИЗАЦИЯ И ИССЛЕДОВАНИЕ ОПЕРАЦИЙ

Редактор серии
Н. Н. МОИСЕЕВ

МОСКВА «НАУКА»
ГЛАВНАЯ РЕДАКЦИЯ
ФИЗИКО-МАТЕМАТИЧЕСКОЙ ЛИТЕРАТУРЫ
1981

В. Ф. ДЕМЬЯНОВ, Л. В. ВАСИЛЬЕВ

НЕДИФФЕРЕНЦИРУЕМАЯ ОПТИМИЗАЦИЯ

МОСКВА «НАУКА»
ГЛАВНАЯ РЕДАКЦИЯ
ФИЗИКО-МАТЕМАТИЧЕСКОЙ ЛИТЕРАТУРЫ
1981

22.18
Д 32
УДК 519.6

Демьянов В. Ф., Васильев Л. В. **Недифференцируемая оптимизация.** — М.: Наука, Главная редакция физико-математической литературы, 1981 (Оптимизация и исследование операций). — 384 с.

Многие задачи, возникающие в технике, экономике, в самой математике, приводят к необходимости нахождения экстремальных значений негладких функций.

Книга посвящена систематическому изложению начал теории оптимизации недифференцируемых функций. Приводятся необходимые сведения из теории точно множественных отображений, изучаются свойства некоторых классов недифференцируемых функций, устанавливаются условия оптимальности. Подробно рассматриваются численные методы минимизации выпуклых недифференцируемых функций, а также функций максимума как при наличии, так и при отсутствии ограничений.

Книга предназначена для широкого круга научных работников и инженеров, специализирующихся в области теории управления, оптимизации и проектирования, а также студентов и аспирантов соответствующих специальностей.

Библ. 172. Илл. 55.

Владимир Федорович Демьянов
Леонид Васильевич Васильев

НЕДИФФЕРЕНЦИРУЕМАЯ ОПТИМИЗАЦИЯ

Редактор *И. Е. Морозова*
Техн. редактор *Н. В. Вершинина.* Корректор *Л. Н. Воровина*

ИБ № 11644

Сдано в набор 28.10.80. Подписано к печати 04.12.81. Т-27790. Формат 84×108¹/₃₂. Бумага тип. № 1. Литературная гарнитура. Высокая печать. Условн. печ. л. 20,16. Уч.-изд. л. 19,47. Тираж 6800 экз. Заказ № 968. Цена 2 р. 10 к.

Издательство «Наука», Главная редакция физико-математической литературы. 117071, Москва, В-71, Ленинский проспект, 15

Ленинградская типография № 2 головное предприятие ордена Трудового Красного Знамени Ленинградского объединения «Техническая книга» им. Евгения Соколовой Союзполиграфпрома при Государственном комитете СССР по делам издательств, полиграфии и книжной торговли. 198052, г. Ленинград Л-52, Измайловский проспект, 29.

Д 20204—147
053(02)—81 КБ — 2-31 — 81. 1702070000

© Издательство «Наука».
Главная редакция
физико-математической
литературы. 1981

ОГЛАВЛЕНИЕ

Предисловие	7
Глава I. Элементы выпуклого анализа и смежные вопросы	13
§ 1. Выпуклые множества. Выпуклые оболочки. Теорема отделимости	13
§ 2. Точечно-множественные отображения	22
§ 3. Выпуклый конус. Конус возможных направлений. Со- пряженный конус	31
§ 4. Выпуклые функции. Непрерывность и дифференцируе- мость по направлениям	40
§ 5. Субградиент и субдифференциал выпуклой функции	54
§ 6. Расстояние от множества до конуса. Условия миниму- ма	70
§ 7. ε -субдифференциал	77
§ 8. ε -производные по направлению. Непрерывность ε -суб- дифференциального отображения	87
§ 9. Некоторые свойства и неравенства для выпуклых функ- ций	99
§ 10. Условный ε -субдифференциал	109
§ 11. Условные производные по направлениям. Непрерыв- ность условного ε -субдифференциального отображе- ния	119
§ 12. Задание выпуклого множества с помощью неравенств	130
§ 13. Нормальный конус. Коническое отображение	137
§ 14. Дифференцируемость по направлениям функции суп- ремума	143
§ 15. Дифференцируемость выпуклой функции	150
§ 16*. Сопряженные функции	163
§ 17. Вычисление ε -субградиентов некоторых классов вы- пуклых функций	177
Глава II. Квазидифференцируемые функции	183
§ 1. Определение и примеры квазидифференцируемых функций	183
§ 2. Свойства квазидифференцируемых функций. Основные формулы квазидифференциального исчисления	189
§ 3. Примеры вычисления квазидифференциалов	198
§ 4*. Квазидифференцируемость выпукло-вогнутой функции	207
§ 5. Необходимые условия оптимальности квазидифферен- цируемой функции на E_n	214
§ 6. Квазидифференцируемые множества	220
§ 7. Необходимые условия оптимальности квазидифферен- цируемой функции на квазидифференцируемом множе- стве	229
§ 8*. Функция расстояния до множества	241
§ 9. Неявные функции	251

Глава III. Минимизация на всем пространстве	254
§ 1. Необходимые и достаточные условия минимума выпуклой функции на E_n	254
§ 2. Минимизация гладкой функции	256
§ 3. Метод наискорейшего спуска	258
§ 4. Субградиентный метод для минимизации выпуклой функции	265
§ 5. Многошаговый субградиентный метод	277
§ 6. Релаксационный субградиентный метод	286
§ 7. Релаксационный ε -субградиентный метод	301
§ 8. Метод Келли	310
§ 9. Минимизация функции супремума	315
§ 10. Минимизация выпуклой функции максимума. Метод экстремального базиса	318
Глава IV. Минимизация при наличии ограничений	326
§ 1. Необходимые и достаточные условия минимума выпуклой функции на выпуклом множестве	326
§ 2. ε -стационарные точки	334
§ 3. Метод условного градиента	337
§ 4. Методы наискорейшего спуска для минимизации выпуклых функций	342
§ 5. (ε, μ) -субградиентный метод при наличии ограничений	348
§ 6. Субградиентный метод с постоянным шагом	352
§ 7. Модифицированный (ε, μ) -субградиентный метод при наличии ограничений	355
§ 8. Метод негладких штрафных функций	359
§ 9. Метод Келли для минимизации на выпуклом множестве	366
§ 10. Релаксационный субградиентный метод при наличии ограничений	369
Библиографический комментарий	373
Литература	376

ПРЕДИСЛОВИЕ

Термин «недифференцируемая оптимизация» (НДО), сравнительно недавно вошедший в научный оборот, включает в себя широкий круг вопросов, связанных с нахождением экстремальных значений недифференцируемых функций. Задачи минимизации или максимизации негладких функций естественным образом возникают как при решении практических задач, так и в самой математике. В качестве примера достаточно сослаться на задачи чебышевской аппроксимации. Везде далее, не теряя общности, мы будем обсуждать задачи минимизации. Среди негладких задач минимизации наиболее изученными являются минимаксные и выпуклые задачи (монографии [31], [36], [57], [110], [120]). Интерес к НДО особенно возрос в последние годы. Различным аспектам негладкой оптимизации посвящены работы [30], [81], [127], а также многочисленные сборники и статьи (см., например, [14], [20], [87]—[89], [98], [130], [135], [140]—[142], [152], [153], [160]).

Для решения любой задачи минимизации необходимо:

1. Исследовать свойства минимизируемой функции, в частности ее дифференцируемость и дифференцируемость по направлениям.
2. Установить необходимые (а при возможности и достаточные) условия глобального или локального минимума.
3. Найти направления спуска (наискорейшего или просто «подходящего» в некотором смысле).
4. Разработать методы последовательных приближений.

В книге рассматриваются задачи минимизации негладких функций конечного числа переменных.

Следует отметить принципиальную важность необходимых условий экстремума (их исследованию.

посвящены, например, работы [24], [45], [57], [73], [74], [103], [159], [163], [167], [168]).

Известно значение понятия градиента в случае гладкой функции. У негладких функций уже не существует градиента. Для функции максимума и выпуклой функции аналогом градиента служит субдифференциал: с каждой точкой x_0 связан некоторый выпуклый компакт $\partial f(x_0)$, называемый субдифференциалом функции $f(x)$ в точке x_0 . Любой элемент $v \in \partial f(x_0)$ называется субградиентом функции $f(x)$ в точке x_0 .

С помощью субдифференциала можно:

1. Найти производную функции $f(x)$ в точке x_0 по направлениям:

$$\frac{\partial f(x_0)}{\partial g} \equiv \lim_{\alpha \rightarrow +0} \alpha^{-1} [f(x_0 + \alpha g) - f(x_0)] = \max_{v \in \partial f(x_0)} (v, g).$$

2. Проверить необходимое условие минимума: для того чтобы точка x^* была точкой минимума функции $f(x)$ на E_n , необходимо, чтобы

$$0 \in \partial f(x^*).$$

3. Найти направление наискорейшего спуска: если $0 \notin \partial f(x_0)$, то направление

$$g(x_0) = -v(x_0) \|v(x_0)\|^{-1},$$

где $\|v(x_0)\| = \min_{v \in \partial f(x_0)} \|v\|$, $v(x_0) \in \partial f(x_0)$, есть направление наискорейшего спуска функции $f(x)$ в точке x_0 .

Такая важная роль субдифференциала привела к попыткам расширить понятие субдифференциала на класс липшицевых функций (работы Ф. Кларка [133], [134], Дж. Варги [9], [168], Б. Н. Пшеничного [104], Н. З. Шора [126], [127], А. Голдстейна [139], [140] и др.).

Используя субдифференциалы и субградиенты, удалось построить ряд методов последовательных приближений для минимизации выпуклых функций, функций максимума и некоторых других классов функций ([30], [36], [51], [91], [127], [149]—[151], [155], [156], [170], [171]).

Задача минимизации гладкой функции $f(x)$ на множестве

$$\Omega = \{x \in E_n \mid h_i(x) \leq 0 \quad \forall i \in 1:N\},$$

где $h_i(x)$ — гладкие на E_n функции, тоже является задачей НДО, ибо множество Ω можно задать так:

$$\Omega = \{x \in E_n \mid h(x) \leq 0\},$$

где $h(x) = \max_{i \in 1:N} h_i(x)$ — негладкая уже функция.

Цель настоящей книги — дать систематическое изложение начал теории оптимизации недифференцируемых функций конечного числа переменных. В гл. I приводятся необходимые сведения из теории выпуклых функций, выпуклых множеств и точечно-множественных отображений. Большое внимание уделено ε -субдифференциалам и свойствам ε -субдифференциального отображения. Выпуклые функции важны не только потому, что это богатый класс негладких функций, но и потому, что аппарат теории выпуклых функций позволяет исследовать и другие классы негладких функций.

Мы уже говорили о некоторых обобщениях субдифференциала. Это понятие для выпуклых функций и функций максимума органично связано с производной по направлениям. Большинство из упомянутых выше авторов в своих обобщениях субдифференциала использовали не производную по направлениям, а какую-нибудь другую конструкцию. Для целей оптимизации, однако, представляется более естественным и полезным опираться на производную по направлениям.

В гл. II описан новый класс недифференцируемых функций — квазидифференцируемые функции. Для таких функций существенную роль играет понятие квазидифференциала, тесно связанное с производной по направлениям. Оказывается, каждой точке соответствует пара выпуклых множеств (квазидифференциал). Квазидифференциал является обобщением понятия производной (для гладкой функции) и субдифференциала (для выпуклой функции). С помощью квазидифференциала удастся достаточно просто описать необходимые условия экстремума, найти направления наискорейшего спуска и подъема. Установлены основные формулы квазидифференциального исчисления, являющегося обобщением классического дифференциального исчисления. Класс квазидифференцируемых функций представляет собой линейное пространство, замкнутое относительно всех

«алгебраических» операций, а также операций взятия поточечных максимума и минимума (напомним, что класс выпуклых функций, как и класс функций максимума, не является линейным пространством, а представляет собой выпуклый конус). Введено понятие квазидифференцируемого множества. Получено необходимое условие экстремума квазидифференцируемой функции на квазидифференцируемом множестве в терминах квазидифференциалов, что значительно расширяет класс задач, которые можно аналитически исследовать. Для широкого класса квазидифференцируемых функций можно алгоритмизировать процесс проверки необходимых условий экстремума и нахождение направлений наискорейшего спуска и подъема. Тем самым открыт путь к численному решению подобных задач, хотя методы нахождения экстремума квазидифференцируемых функций требуют еще своей разработки.

Численным методам решения задач НДО посвящены главы III и IV, в которых излагаются основные методы и идеи, используемые при минимизации выпуклых функций и функций максимума. Методы последовательных приближений делятся на релаксационные и нерелаксационные. Метод называется релаксационным, если значение функции на очередном шаге меньше, чем на предыдущем. Среди излагаемых методов есть как релаксационные, так и нерелаксационные. Мы воздерживаемся от обсуждения преимуществ одного метода перед другим, поскольку у дракона оптимизации много голов и против каждой из них нужен свой меч. Так, метод субградиентного спуска весьма прост в реализации, но очень медленно сходится. Многое зависит от поставленной цели и имеющихся ресурсов. Иногда можно ограничиться достаточно грубым приближением к оптимальному решению, но получить его нужно быстро; в другом случае необходима высокая точность, а время счета практически не имеет значения.

Большинство описанных методов являются методами первого порядка, поскольку в них используются аппроксимации первого порядка (производная, субградиент, субдифференциал). Можно надеяться, что дальнейшее развитие теории НДО приведет к появлению методов и более высокого порядка.

Часть материала вынесена в задания для самостоятельного изучения. Мы совершенно не затрагиваем вероятностных подходов (отсылая читателя к монографиям [51], [83], [107], [117]), а также вопросов, связанных с теорией игр ([21], [58], [63], [64]) и многокритериальной оптимизацией ([95]), в которых возникают задачи НДО.

В книге отражены результаты исследований по негладкой оптимизации, проводимых в последние годы на факультете прикладной математики — процессов управления ЛГУ и в НИИ Вычислительной математики и процессов управления ЛГУ.

Часть результатов докладывалась на VII и VIII Всесоюзных школах по оптимизации в Жукине (1977 г.) и Шушенском (1979 г.).

Авторы искренне признательны Н. Н. Моисееву, который способствовал возникновению замысла настоящей монографии и поддерживал работу над ней на различных ее этапах.

А. Б. Певный и А. М. Рубинов внимательно и критически прочли рукопись. Ряд замечаний, сделанных Е. Ф. Войтоном, М. К. Гавуриным, Ю. М. Ермольевым, С. С. Кутателадзе, В. Н. Малоземовым, Б. Н. Пшеничным, В. М. Тихомировым, способствовал улучшению книги. Большую помощь в оформлении рукописи оказала Э. В. Демьянова. Наконец, авторы благодарны всем, кто своими замечаниями, вопросами, советами помогал работе.

В книге пп. 1, 2 § 13 гл. I и § 5 гл. II написаны Л. Н. Поляковой, §§ 15, 16 гл. I — А. М. Рубиновым, § 14 гл. I и пп. 1, 2 § 4 гл. IV — В. К. Шомесовой, § 9 гл. I, §§ 5, 7 гл. IV — Л. В. Васильевым, В. Ф. Демьяновым.

Л. В. Васильевым написаны п. 6 § 4, § 6, пп. 5, 6 § 8, § 17 гл. I, §§ 1, 3, 6, 7 гл. III, §§ 1—3, п. 3 § 4, §§ 6, 8—10 гл. IV.

В. Ф. Демьянов написал §§ 1—3, пп. 1—5 § 4, §§ 5, 7, пп. 1—4 § 8, §§ 10—12, п. 3 § 13 гл. I, §§ 1—4, 6—9 гл. II, §§ 2, 4, 5, 8—10 гл. III.

В книге используются стандартные обозначения и сокращения.

Выражение $\inf_{x \in A} \{f(x) \mid x \in A\}$ есть то же самое, что и $\inf_{x \in A} f(x)$.

Запись

$$\Omega = \{v \in E_n \mid \exists \alpha_0 > 0: x_0 + \alpha v \in A \quad \forall \alpha \in [0, \alpha_0]\}$$

читается так: Ω есть множество таких точек $v \in E_n$, для которых существует $\alpha_0 > 0$ такое, что $x_0 + \alpha v \in A$ для всех $\alpha \in [0, \alpha_0]$.

Символ ■ означает конец доказательства.

Множество целых чисел от p до q обозначается $p : q$.

Нижний и верхний пределы соответственно обозначаются $\underline{\lim}$ и $\overline{\lim}$.

Количество элементов множества A обозначается $|A|$.

Символ $\equiv$ означает «равно по определению».

Звездочкой отмечены параграфы и их части, чтение которых не обязательно для понимания последующего материала.

В. Ф. Демьянов

Л. В. Васильев

ЭЛЕМЕНТЫ ВЫПУКЛОГО АНАЛИЗА И СМЕЖНЫЕ ВОПРОСЫ

§ 1. Выпуклые множества. Выпуклые оболочки. Теорема отделимости

1. Далее рассматривается n -мерное евклидово пространство E_n векторов $x = (x^{(1)}, \dots, x^{(n)})$. Пространство E_n является линейным. Обозначим $\mathbb{O}_n = \mathbb{O} = (0, \dots, 0) \in E_n$, $x_k = (x_k^{(1)}, \dots, x_k^{(n)})$, $(x_1, x_2) = \sum_{i=1}^n x_1^{(i)} x_2^{(i)}$, $\|x\| = \sqrt{(x, x)}$, $x^2 = (x, x)$.

Имеет место *неравенство Коши — Буняковского*:

$$|(x_1, x_2)| \leq \|x_1\| \cdot \|x_2\|.$$

Векторы $x_1, \dots, x_r$ называются *линейно независимыми*, если равенство $\sum_{k=1}^r \alpha_k x_k = \mathbb{O}$ возможно тогда и только тогда, когда все коэффициенты α_k , $k \in 1:r$, равны нулю. Если $r \geq n+1$, то векторы $x_1, \dots, x_r$ *линейно зависимы*, т. е. найдутся числа $\beta_1, \dots, \beta_r$ такие, что $\sum_{k=1}^r \beta_k^2 > 0$ (т. е. они не все одновременно равны нулю) и

$$\sum_{k=1}^r \beta_k x_k = \mathbb{O}. \quad (1.1)$$

Если $r \geq n+2$, то одновременно с (1.1) можно добиться выполнения равенства

$$\sum_{k=1}^r \beta_k = 0. \quad (1.2)$$

Докажем это. Введем векторы

$$\bar{x}_k = (1, x_k^{(1)}, \dots, x_k^{(n)}) \in E_{n+1}, \quad k \in 1:r, \quad r \geq n+2.$$

В E_{n+1} любые $n+2$ вектора линейно зависимы, т. е. существуют числа β_k , $\sum_{k=1}^r \beta_k^2 > 0$, такие, что

$$\sum_{k=1}^r \beta_k \bar{x}_k = 0_{n+1}. \quad (1.3)$$

Из (1.3) следует

$$\sum_{k=1}^r \beta_k x_k = 0_n, \quad \sum_{k=1}^r \beta_k = 0$$

(приравняли нулю отдельно первую координату и n оставшихся).

Множество, не содержащее ни одного элемента, называется *пустым* и обозначается $\emptyset$.

Положим

$$S_\delta(x_0) = \{x \in E_n \mid \|x - x_0\| \leq \delta\}, \quad \delta > 0.$$

Множество $S_\delta(x_0)$ называется δ -окрестностью точки x_0 . Точка x_0 называется *внутренней* точкой множества G , если существует $\delta > 0$ такое, что $S_\delta(x_0) \subset G$. Множество внутренних точек множества G обозначим $\text{int } G$ (это множество может оказаться и пустым).

Множество $G \subset E_n$ будем называть *открытым*, если для любого x_0 существует $\delta > 0$ такое, что $S_\delta(x_0) \subset G$. Очевидно, что в этом случае $G = \text{int } G$.

Замыканием множества $G \subset E_n$ называется множество точек x , допускающих представление $x = \lim_{k \rightarrow \infty} x_k$, где $x_k \in G \ \forall k \in 1: \infty$. Замыкание G будем обозначать $\bar{G}$.

Множество $G \subset E_n$ будем называть *замкнутым*, если из соотношений $x_k \xrightarrow{k \rightarrow \infty} x_0$, $x_k \in G \ \forall k \in 1: \infty$, следует $x_0 \in G$. Очевидно, что для замкнутого множества $G = \bar{G}$.

Точка x_0 называется *граничной* точкой множества $G \subset E_n$, если ее δ -окрестность $S_\delta(x_0)$ при любом $\delta > 0$ содержит хотя бы одну точку, не принадлежащую G , и хотя бы одну точку, принадлежащую G (при этом x_0 может и не принадлежать G). Множество граничных точек множества G обозначим G_{fr} .

Множество G называется *ограниченным*, если существует такое число $K < +\infty$, что $\|x\| \leq K \ \forall x \in G$.

Множество G будем называть *неограниченным*, если для любого $K > 0$ найдется такое $x \in G$, что $\|x\| > K$.

Очевидно, что объединение, пересечение, сумма и разность двух ограниченных множеств есть снова ограниченное множество.

Пересечение двух множеств, одно из которых ограничено, есть ограниченное множество.

Если A и B — замкнутые множества, то их объединение и пересечение — снова замкнутые множества. Но для суммы, разности и алгебраической разности это уже не так.

Пример 1. Пусть

$$A = \left\{ x = (x^{(1)}, x^{(2)}) \in E_2 \mid x^{(2)} \geq \frac{1}{x^{(1)}}, \quad x^{(1)} > 0 \right\},$$

$$B = \{x = (x^{(1)}, x^{(2)}) \in E_2 \mid x^{(1)} = 0, \quad x^{(2)} \leq 0\}.$$

Очевидно, что A и B — замкнутые множества и не являются ограниченными.

Множество

$$C = A + B = \{x = (x^{(1)}, x^{(2)}) \mid x^{(1)} > 0, \quad x^{(2)} \in (-\infty, \infty)\}$$

не является замкнутым, ибо

$$\bar{C} = \{x = (x^{(1)}, x^{(2)}) \mid x^{(1)} \geq 0, \quad x^{(2)} \in (-\infty, \infty)\} \neq C.$$

Однако если множества A и B замкнуты и хотя бы одно из них ограничено, то сумма (и алгебраическая разность) тоже замкнута.

Компактом или *компактным множеством* называется множество, обладающее тем свойством, что из всякой последовательности его элементов можно выбрать сходящуюся подпоследовательность, предел которой принадлежит тому же множеству. Хорошо известно, что в E_n множество компактно тогда и только тогда, когда оно замкнуто и ограничено.

Определение 1. Множество $\Omega \subset E_n$ называется *выпуклым*, если вместе с любыми двумя точками x_1, x_2 оно содержит отрезок, соединяющий эти точки, т. е. $[x_1, x_2] \subset \Omega$, где

$$[x_1, x_2] = \{x \in E_n \mid x = \alpha x_1 + (1 - \alpha) x_2, \quad \alpha \in [0, 1]\}.$$

Выпуклое множество Ω называется *строго выпуклым*, если для любых $x_1, x_2 \in \Omega$, $x_1 \neq x_2$, и любого $\alpha \in (0, 1)$ точка $x_\alpha = \alpha x_1 + (1 - \alpha) x_2 \in \text{int } \Omega$.

Существует другое определение выпуклого множества.

Определение 1*. Множество $\Omega \subset E_n$ называется *выпуклым*, если вместе с любыми двумя точками x_1, x_2 оно содержит точку $2^{-1}(x_1 + x_2)$, т. е. если середина отрезка, соединяющего точки x_1 и x_2 , принадлежит Ω .

Ясно, что если множество Ω выпукло в смысле определения 1, то оно выпукло и в смысле определения 1*. Обратное неверно.

Пример 2. Пусть Ω есть множество рациональных чисел из $[0, 1]$. Очевидно, что это множество выпукло в смысле определения 1*, но не является выпуклым в смысле определения 1.

Задание 1.1. Доказать, что если множество Ω замкнуто и выпукло в смысле определения 1*, то оно выпукло и в смысле определения 1.

Замечание 1. Везде далее, говоря о выпуклом множестве, имеем в виду выпуклость в смысле определения 1.

Лемма 1.1. Пусть $\Omega \subset E_n$ — замкнутое множество и для любых $x_1, x_2 \in \Omega$ найдется $\alpha \in (0, 1)$ (α зависит от x_1 и x_2) такое, что $x_\alpha = \alpha x_1 + (1 - \alpha)x_2 \in \Omega$. Тогда множество Ω выпуклое.

Доказательство. Допустим противное. Тогда существуют $x_1, x_2 \in \Omega$, $x_1 \neq x_2$, и такое $\alpha_0 \in (0, 1)$, что

$$x_{\alpha_0} = \alpha_0 x_1 + (1 - \alpha_0)x_2 \notin \Omega.$$

Положим

$$G_1 = \Omega \cap [x_1, x_{\alpha_0}], \quad G_2 = \Omega \cap [x_{\alpha_0}, x_2].$$

Эти множества замкнуты, причем $x_{\alpha_0} \notin G_1$, $x_{\alpha_0} \notin G_2$. Найдем

$$\min_{x \in G_1} \|x - x_{\alpha_0}\| = \|\bar{x} - x_{\alpha_0}\|,$$

$$\min_{x \in G_2} \|x - x_{\alpha_0}\| = \|\bar{\bar{x}} - x_{\alpha_0}\|,$$

где $\bar{x} \in G_1$, $\bar{\bar{x}} \in G_2$, в частности, $\bar{x} \in \Omega$, $\bar{\bar{x}} \in \Omega$. При этом на $(\bar{x}, x_{\alpha_0})$ и $(x_{\alpha_0}, \bar{\bar{x}})$ нет точек Ω , что противоречит условию леммы (здесь $(x, y) = [x, y] \setminus \{x\}$, $[x, y) = [x, y] \setminus \{y\}$). ■

2. Пусть $G \subset E_n$. Выпуклой оболочкой множества G назовем множество

$$\text{co } G = \left\{ x = \sum_{k=1}^r \alpha_k x_k \mid x_k \in G, \alpha_k \geq 0, \sum_{k=1}^r \alpha_k = 1, \right. \\ \left. r - \text{любое натуральное число} \right\}.$$

Теорема 1.1 (теорема Каратеодори). *Любой вектор $x \in \text{co } G$ может быть представлен в виде выпуклой комбинации не более чем $n+1$ векторов множества G .*
Доказательство. Пусть $x \in \text{co } G$, т. е.

$$x = \sum_{k=1}^r \alpha_k x_k, \quad \alpha_k \geq 0, \quad \sum_{k=1}^r \alpha_k = 1.$$

Без ограничения общности считаем, что все $\alpha_k > 0$. Допустим, что $r \geq n+2$. По доказанному выше (см. (1.2))

существуют такие числа β_k , $\sum_{k=1}^r \beta_k^2 > 0$, что

$$\sum_{k=1}^r \beta_k x_k = 0, \quad \sum_{k=1}^r \beta_k = 0. \quad (1.4)$$

Положим $\varepsilon = \min_{\{k \mid \beta_k > 0\}} \beta_k^{-1} \alpha_k > 0$ и $\bar{\alpha}_k = \alpha_k - \varepsilon \beta_k$, $k \in 1:r$.

Здесь $\varepsilon > 0$, ибо в силу (1.4) среди коэффициентов β_k существуют и положительные, и отрицательные (хотя бы по одному). Теперь ясно, что

$$x = \sum_{k=1}^r \bar{\alpha}_k x_k, \quad \sum_{k=1}^r \bar{\alpha}_k = 1, \quad \bar{\alpha}_k \geq 0 \quad \forall k.$$

Очевидно также, что хотя бы одно из чисел $\bar{\alpha}_k$ обращается в нуль, т. е. x может быть представлен как выпуклая комбинация уже не более чем $r-1$ векторов из G . Продолжая далее аналогично, придем к представлению, в котором $r \leq n+1$. ■

Лемма 1.2. *Если $G \subset E_n$ — ограниченное замкнутое множество, то $\text{co } G$ — ограниченное замкнутое выпуклое множество.*

Доказательство. Пусть $x \in \text{co } G$, т. е.

$$x = \sum_{k=1}^r \alpha_k x_k = (x^{(1)}, \dots, x^{(n)}), \quad \alpha_k \geq 0, \quad \sum_{k=1}^r \alpha_k = 1, \quad x_k \in G. \quad (1.5)$$

Если G ограничено, то для любого вектора из G все его компоненты ограничены по модулю одним и тем же числом M . Тогда для вектора x будет

$$|x^{(i)}| = \left| \sum_{k=1}^r \alpha_k x_k^{(i)} \right| \leq \sum_{k=1}^r \alpha_k |x_k^{(i)}| \leq M \sum_{k=1}^r \alpha_k = M,$$

т. е. компоненты вектора x (см. (1.5)) из $\text{co } G$ ограничены тем же числом M , т. е. $\text{co } G$ — ограниченное множество.

Докажем выпуклость $\text{co } G$. Пусть $x \in \text{co } G$, $y \in \text{co } G$. Без ограничения общности (добавив, если нужно, слагаемые с нулевыми коэффициентами) можно считать, что

$$x = \sum_{k=1}^r \alpha_k x_k, \quad y = \sum_{k=1}^r \alpha'_k x_k,$$

$$x_k \in G, \quad \alpha_k \geq 0, \quad \alpha'_k \geq 0, \quad \sum_{k=1}^r \alpha_k = \sum_{k=1}^r \alpha'_k = 1.$$

Возьмем

$$x_\alpha = \alpha x + (1 - \alpha) y, \quad \alpha \in [0, 1], \quad (1.6)$$

$$x_\alpha = \sum_{k=1}^r (\alpha \alpha_k + (1 - \alpha) \alpha'_k) x_k = \sum_{k=1}^r \bar{\alpha}_k x_k,$$

где $\bar{\alpha}_k = \alpha \alpha_k + (1 - \alpha) \alpha'_k$. Ясно, что

$$\bar{\alpha}_k \geq 0, \quad \sum_{k=1}^r \bar{\alpha}_k = \alpha \sum_{k=1}^r \alpha_k + (1 - \alpha) \sum_{k=1}^r \alpha'_k = 1,$$

т. е. $x_\alpha \in \text{co } G$ при любом $\alpha \in [0, 1]$. Значит, множество $\text{co } G$ — выпуклое.

Покажем замкнутость множества $\text{co } G$. Пусть $x_s \rightarrow x^*$, $x_s \in \text{co } G$. Тогда

$$x_s = \sum_{k=1}^{r_s} \alpha_{sk} x_{sk}, \quad \alpha_{sk} \geq 0, \quad \sum_{k=1}^{r_s} \alpha_{sk} = 1, \\ x_{sk} \in G, \quad k \in 1 : r_s, \quad s \in 1 : \infty. \quad (1.7)$$

По теореме Каратеодори можно выбрать такие представления x_s , что $r_s \leq n+1$. Будем считать, что $r_s = n+1$ (пополнив для этого, если нужно, представления (1.7) произвольными векторами из G с нулевыми коэффициентами). Так как G — ограниченное множество и $\alpha_{sk} \in [0, 1]$, то можно выбрать последовательность $\{s_i\}$ такую, что

$$x_{s_i k} \xrightarrow{s_i \rightarrow \infty} x_{0k}, \quad \alpha_{s_i k} \xrightarrow{s_i \rightarrow \infty} \alpha_{0k}.$$

Из (1.7) ясно, что

$$x^* = \sum_{k=1}^{n+1} \alpha_{0k} x_{0k}, \quad x_{0k} \in G, \quad \alpha_{0k} \geq 0, \quad \sum_{k=1}^{n+1} \alpha_{0k} = 1,$$

т. е. $x^* \in \text{co } G$. Лемма доказана полностью.

Зада н и е 1.2. Доказать, что если G — выпуклое множество, то $\text{co } G = \bar{G}$.

З а м е ч а н и е 2. Очевидно следующее свойство: если $A_i \subset E_n$ — выпуклые множества, $i \in I \equiv 1 : N$, то

$$\text{co } \bigcup_{i \in I} A_i = \left\{ x = \sum_{i \in I} \alpha_i x_i \mid x_i \in A_i, \alpha_i \geq 0, \sum_{i \in I} \alpha_i = 1 \right\}.$$

Теорема 1.2 (теорема отделимости). Пусть $\Omega \subset E_n$ — замкнутое выпуклое множество и $x_0 \notin \Omega$. Тогда найдутся вектор $g_0 \in E_n$, $\|g_0\| = 1$, и число $a > 0$ такие, что для любого $x \in \Omega$ будет

$$(x - x_0, g_0) \leq -a. \quad (1.8)$$

Доказательство. Найдем

$$\min_{z \in \Omega} \|x_0 - z\|^2 = \|x_0 - z_0\|^2 \equiv \rho_0^2,$$

где $\rho_0 = \|x_0 - z_0\|$. Нетрудно убедиться в том, что наименьшее значение здесь достигается. Так как $x_0 \notin \Omega$, то $\rho_0 > 0$. Положим

$$g_0 = -\rho_0^{-1}(z_0 - x_0). \quad (1.9)$$

Покажем, что для всех $x \in \Omega$ будет

$$(x - x_0, g_0) \leq -\rho_0. \quad (1.10)$$

Допустим противное. Тогда найдется $\bar{x} \in \Omega$ такое, что

$$(\bar{x} - x_0, g_0) = -b_0 > -\rho_0. \quad (1.11)$$

Пусть

$$x_\alpha = \alpha \bar{x} + (1 - \alpha) z_0 = z_0 + \alpha (\bar{x} - z_0), \quad \alpha \in [0, 1].$$

Ясно, что $x_\alpha \in [z_0, \bar{x}]$. Тогда

$$\begin{aligned} (x_0 - x_\alpha)^2 &= (x_0 - z_0)^2 - 2\alpha (\bar{x} - z_0, x_0 - z_0) + \alpha^2 (\bar{x} - z_0)^2 = \\ &= \rho_0^2 - 2\alpha [(\bar{x} - x_0, x_0 - z_0) - (z_0 - x_0, x_0 - z_0)] + \\ &\quad + \alpha^2 (\bar{x} - z_0)^2. \end{aligned}$$

Учитывая (1.9), имеем

$$(x_0 - x_\alpha)^2 = \rho_0^2 - 2\alpha [-b_0 \rho_0 + \rho_0^2] + \alpha^2 (\bar{x} - z_0)^2. \quad (1.12)$$

Так как из (1.11) $\rho_0^2 - b_0 \rho_0 > 0$, то из (1.12) следует, что при достаточно малых $\alpha \in (0, 1)$ будет $(x_0 - x_\alpha)^2 < \rho_0^2$, что противоречит определению точки z_0 .

Итак, (1.10) эквивалентно (1.8). Теорема доказана.

Замечание 3. Из доказательства теоремы вытекает, что если $\Omega \subset E_n$ — замкнутое выпуклое множество и $x_0 \in E_n$, то существует единственная точка $z_0 \in \Omega$ такая, что $(z - x_0, z_0 - x_0) \geq \|z_0 - x_0\|^2 \quad \forall z \in \Omega$.

Следствие 1. Если $\Omega \subset E_n$ — замкнутое выпуклое множество и x_0 — граничная точка множества Ω , то существует вектор $g_0 \in E_n$, $\|g_0\| = 1$, такой, что

$$(x - x_0, g_0) \leq 0 \quad \forall x \in \Omega. \quad (1.13)$$

Доказательство. Пусть x_0 — граничная точка Ω . Тогда найдется последовательность $\{x_k\}$ такая, что $x_k \notin \Omega$, $x_k \rightarrow x_0$. Так как $x_k \notin \Omega$, то по теореме 1.2 для всех $x \in \Omega$ будет

$$(x - x_k, g_k) \leq -a_k, \quad (1.14)$$

где

$$a_k = \min_{z \in \Omega} \|x_k - z\| = \|x_k - z_k\| > 0,$$

$$g_k = -\|z_k - x_k\|^{-1} (z_k - x_k), \quad \|g_k\| = 1.$$

Ясно, что $a_k > 0$ и $a_k = \min_{z \in \Omega} \|x_k - z\| \leq \|x_0 - x_k\|$, т. е.

$a_k \rightarrow 0$. Можем без ограничения общности считать, что $g_k \rightarrow g_0$, $\|g_0\| = 1$. Переходя к пределу в (1.14), получим

$$(x - x_0, g_0) \leq 0 \quad \forall x \in \Omega,$$

что и требуется. ■

З а м е ч а н и е 4. Условие (1.8) означает, что плоскость

$$(x - x_0, g_0) = 0, \quad (1.15)$$

проходящая через точку x_0 , обладает тем свойством, что все множество Ω лежит по одну сторону от этой плоскости, причем на расстоянии, не меньшем чем a .

Условие (1.13) означает, что все множество Ω лежит по одну сторону от плоскости (1.15).

С л е д с т в и е 2. Пусть $\Omega \subset E_n$ — выпуклое множество и $x_0 \notin \Omega$. Тогда существует вектор $g_0 \in E_n$, $\|g_0\| = 1$, такой, что

$$(x - x_0, g_0) \leq 0 \quad \forall x \in \bar{\Omega}.$$

З а д а н и е 1.3. Доказать следующие теорему и следствие.

Теорема 1.3. Пусть Ω_1 и Ω_2 — замкнутые выпуклые множества и хотя бы одно из них ограничено. Если множества Ω_1 и Ω_2 не имеют общих точек, то существуют такие вектор g_0 , точка $x_0 \in \Omega_1$ и число $a > 0$, что

$$(x - x_0, g_0) \leq 0 \quad \forall x \in \Omega_1,$$

$$(x - x_0, g_0) \geq a \quad \forall x \in \Omega_2.$$

У к а з а н и е. Надо найти

$$\min_{x \in \Omega_1, y \in \Omega_2} \|x - y\|^2 = \|x_0 - y_0\|^2.$$

С л е д с т в и е. Пусть Ω_1 и Ω_2 — замкнутые выпуклые множества, причем хотя бы одно из них ограничено. Если эти множества не имеют общих внутренних точек, т. е.

$$\text{int } \Omega_1 \cap \text{int } \Omega_2 = \emptyset, \quad (1.16)$$

то существуют такие вектор g_0 и точка $x_0 \in \Omega_1$, что

$$(x - x_0, g_0) \leq 0 \quad \forall x \in \Omega_1,$$

$$(x - x_0, g_0) \geq 0 \quad \forall x \in \Omega_2.$$

З а м е ч а н и е 5. Утверждение в следствии можно усилить, заменив (1.16) одним из условий:

$$\text{int } \Omega_1 \cap \Omega_2 = \emptyset, \quad \Omega_1 \cap \text{int } \Omega_2 = \emptyset. \quad (1.17)$$

§ 2. Точечно-множественные отображения

1. Пусть Ω — некоторое множество. Совокупность всех непустых подмножеств множества Ω обозначим $\Pi(\Omega)$. Пусть Ω_1 и Ω_2 — некоторые множества конечномерных пространств. Отображение $A(x)$, которое переводит каждую точку $x \in \Omega_1$ в некоторое подмножество множества Ω_2 , называется *точечно-множественным отображением* (т. м. о.) Ω_1 в Ω_2 или отображением Ω_1 в $\Pi(\Omega_2)$. Будем обозначать это так: $A: \Omega_1 \rightarrow \Pi(\Omega_2)$.

Пусть $A \subset E_n$, $B \subset E_n$. Положим

$$\rho(A, B) = \sup_{w \in A} \inf_{v \in B} \|v - w\| + \sup_{v \in B} \inf_{w \in A} \|v - w\|. \quad (2.1)$$

Число $\rho(A, B)$ называется *расстоянием* между множествами A и B в метрике Хаусдорфа (или *расстоянием Хаусдорфа*).

Величина $\sup_{w \in A} \inf_{v \in B} \|v - w\|$ называется *уклонением* множества A от множества B (в смысле Хаусдорфа), а величина $\sup_{v \in B} \inf_{w \in A} \|v - w\|$ называется *уклонением* множества B от множества A (в смысле Хаусдорфа).

За д а н и е 2.1. Показать, что если A и B — замкнутые множества пространства E_n , то равенство $\rho(A, B) = 0$ эквивалентно соотношению $A = B$.

З а м е ч а н и е 1. Требование замкнутости в задании 2.1 существенно, что видно из следующего примера. Пусть

$$A = \{x = (x^{(1)}, x^{(2)}) \in E_2 \mid x^{(i)} \text{ — рациональные числа, } |x^{(i)}| \leq 1, i \in 1:2\},$$

$$B = \{x = (x^{(1)}, x^{(2)}) \in E_2 \mid |x^{(i)}| \leq 1, i \in 1:2\}.$$

Множество A не является замкнутым. Ясно, что $\rho(A, B) = 0$, но $A \neq B$.

Отображение $A(x)$ называется *непрерывным по Хаусдорфу* (*H-непрерывным*) в точке x_0 , если

$$\rho(A(x), A(x_0)) \xrightarrow{x \rightarrow x_0} 0. \quad (2.2)$$

Т. м. о. $A: \Omega_1 \rightarrow \Pi(\Omega_2)$, где $\Omega_1 \subset E_n$, $\Omega_2 \subset E_m$, называется *полунепрерывным сверху* (п. н. св.) в точке $x_0 \in \Omega_1$, если из того, что $x_k \rightarrow x_0$, $v_k \in A(x_k)$, $v_k \rightarrow v_0$, следует $v_0 \in A(x_0)$.

Отметим, что существуют разные определения полунепрерывности сверху [86], которые совпадают, если Ω_2 ограничено.

Графиком отображения $A(x)$ называется множество

$$Z = \{[x, y] \in \Omega_1 \times \Omega_2 \mid x \in \Omega_1, y \in A(x)\}. \quad (2.3)$$

Задание 2.2. Доказать, что отображение $A(x)$ полунепрерывно сверху на Ω_1 тогда и только тогда, когда график Z (см. (2.3)) замкнут в $\Omega_1 \times \Omega_2$.

Задание 2.3. Пусть т.м.о. $A: \Omega_1 \rightarrow \Pi(\Omega_2)$ полунепрерывно сверху на Ω_1 . Доказать, что тогда для любого $x \in \Omega_1$ множество $A(x)$ замкнуто.

Обратное неверно, что видно из следующего примера.

Пусть $\Omega_1 = \Omega_2 = E_1$,

$$A(x) = \begin{cases} [0, 1], & x \neq 0, \\ [0, 2^{-1}], & x = 0. \end{cases}$$

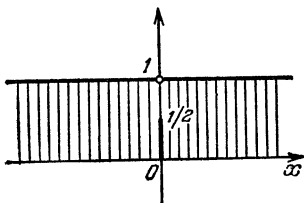


Рис. 1.

При любом x множество $A(x)$ замкнуто, но отображение $A(x)$ не является п. н. св. в точке $x_0 = 0$ (рис. 1).

Отображение $A: \Omega_1 \rightarrow \Pi(\Omega_2)$ называется *полунепрерывным снизу* (п. н. сн.) в точке $x_0 \in \Omega_1$, если для всякого $v_0 \in A(x_0)$ и любой последовательности $\{x_k\}$, $x_k \rightarrow x_0$, $x_k \in \Omega_1$, найдутся такие $v_k \in A(x_k)$, что $v_k \rightarrow v_0$.

Отображение $A(x)$ называется *непрерывным по Какутани* (*K-непрерывным*) в точке x_0 , если оно полунепрерывно в этой точке и снизу, и сверху.

Следующий пример иллюстрирует полунепрерывность сверху и снизу.

Пример 1. Пусть $\Omega_1 = \Omega_2 = E_1$. Рассмотрим отображения

$$A_1(x) = \begin{cases} [x, 1+x], & x \neq 0, \\ [0, 2^{-1}], & x = 0, \end{cases}$$

$$A_2(x) = \begin{cases} [x, 1+x] & x \neq 0, \\ [0, 2], & x = 0. \end{cases}$$

В точке $x_0 = 0$ отображение $A_1(x)$ полунепрерывно снизу, но не полунепрерывно сверху, а отображение $A_2(x)$ полунепрерывно сверху, но не полунепрерывно снизу (рис. 2, 3).

Отображение $A(x)$ называется *ограниченным* на Ω_1 , если для любого ограниченного подмножества $G \subset \Omega_1$ найдется ограниченное подмножество $B \subset \Omega_2$ такое, что $A(x) \subset B \quad \forall x \in G$.

Отображение $A(x)$ называется *замкнутым* в точке x_0 , если множество $A(x_0)$ замкнуто.

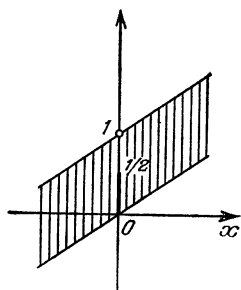


Рис. 2.

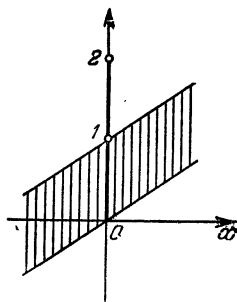


Рис. 3.

Теорема 2.1 [78]. Если отображение $A(x)$ ограничено в некоторой окрестности точки x_0 и замкнуто в точке x_0 , то из непрерывности $A(x)$ по Хаусдорфу в точке x_0 следует непрерывность $A(x)$ по Какутани в точке x_0 , и наоборот, т. е. понятия H -непрерывности и K -непрерывности в этом случае эквивалентны.

Доказательство. Пусть т. м. о. $A(x)$ непрерывно по Хаусдорфу в точке x_0 , т. е. имеет место (2.2). Покажем, что тогда $A(x)$ — п. н. св. Пусть $x_k \rightarrow x_0$, $v_k \in A(x_k)$, $v_k \rightarrow v_0$.

Допустим, что $v_0 \notin A(x_0)$. В силу замкнутости $A(x_0)$ существует $a > 0$ такое, что при достаточно больших k

$$\inf_{v \in A(x_0)} \|v - v_k\| = \min_{v \in A(x_0)} \|v - v_k\| \geq a > 0.$$

Тогда (см. (2.1)) тем более

$$\rho(A(x_k), A(x_0)) \geq \inf_{v \in A(x_0)} \|v - v_k\| \geq a > 0,$$

что противоречит (2.2). Итак, т. м. о. $A(x)$ является п. н. св.

Установим теперь, что т. м. о. $A(x)$ полунепрерывно снизу.

Пусть $v_0 \in A(x_0)$ и $x_k \rightarrow x_0$. Положим

$$b_k = \inf_{v \in A(x_k)} \|v - v_0\|.$$

Поскольку $b_k \leq \rho(A(x_0), A(x_k))$, а $A(x)$ непрерывно по Хаусдорфу, то $\rho(A(x_0), A(x_k)) \rightarrow 0$. Отсюда $b_k \rightarrow 0$.

Тогда для любой последовательности $\{\gamma_k\}$ такой, что $\gamma_k \rightarrow +0$, найдутся $v_k \in A(x_k)$ такие, что $\|v_k - v_0\| \leq b_k + \gamma_k$, т. е. $v_k \rightarrow v_0$, а это и означает (в силу произвольности $v_0 \in A(x_0)$ и последовательности $\{x_k\}$), что $A(x)$ полунепрерывно снизу. Отсюда и из установленной выше полунепрерывности сверху вытекает K -непрерывность.

Пусть теперь т. м. о. $A(x)$ является K -непрерывным в точке x_0 . Допустим, что оно не является H -непрерывным в этой точке. Тогда найдутся число $a > 0$ и последовательность $\{x_k\}$ такие, что

$$x_k \rightarrow x_0, \quad \rho(A(x_k), A(x_0)) \geq 2a > 0. \quad (2.4)$$

В этом случае (см. (2.1)) для каждого фиксированного k либо

$$\sup_{w \in A(x_k)} \inf_{v \in A(x_0)} \|v - w\| \geq a, \quad (2.5)$$

либо

$$\sup_{v \in A(x_0)} \inf_{w \in A(x_k)} \|v - w\| \geq a. \quad (2.6)$$

Без ограничения общности можем считать, что для всей последовательности $\{x_k\}$ имеет место либо (2.5), либо (2.6). Рассмотрим каждый из этих случаев подробно.

В случае (2.5) найдутся $w_k \in A(x_k)$, такие, что

$$\inf_{v \in A(x_0)} \|v - w_k\| \geq 2^{-1}a. \quad (2.7)$$

По предположению т. м. о. $A(x)$ ограничено в окрестности точки x_0 , поэтому последовательность $\{w_k\}$ тоже ограничена. Выделим сходящуюся подпоследовательность $\{w_{k_s}\}$, $w_{k_s} \rightarrow w_0$. Из (2.7) ясно, что $w_0 \notin A(x_0)$, а это противоречит полунепрерывности $A(x)$ сверху в точке x_0 . Итак, случай (2.5) невозможен.

В случае (2.6) для каждого k найдутся $v_k \in A(x_0)$ такие, что

$$\inf_{w \in A(x_k)} \|v_k - w\| \geq 2^{-1}a. \quad (2.8)$$

Так как множество $A(x_0)$ замкнуто и ограничено, то существует подпоследовательность $\{k_s\}$ такая, что $v_{k_s} \rightarrow v_0 \in A(x_0)$. Из (2.8) при достаточно больших k_s

$$\inf_{w \in A(x_{k_s})} \|v_0 - w\| \geq 4^{-1}a,$$

что противоречит полунепрерывности снизу отображения $A(x)$ в точке x_0 . ■

Следствие. Если $A(x)$ H -непрерывно и замкнуто в точке x_0 , то оно и K -непрерывно.

Если же $A(x)$ K -непрерывно в точке x_0 и ограничено в некоторой окрестности точки x_0 , то оно и H -непрерывно.

Замечание 2. Требования в условиях теоремы являются минимальными. Действительно, замкнутость $A(x_0)$ следует из полунепрерывности сверху т. м. о. (см. задание 2.3). Приведем пример, показывающий, что ограниченность $A(x)$ в окрестности точки x_0 (и, естественно, в самой точке x_0) существенна. Пусть $\Omega_1 = \Omega_2 = E_1$,

$$A(x) = \begin{cases} [-1+x, 1+x] \cup \{x^{-1}\}, & x \neq 0, \\ [-1, 1], & x = 0. \end{cases}$$

Отображение $A(x)$ не является ограниченным в окрестности точки $x_0 = 0$. Ясно, что в точке $x_0 = 0$ отображение $A(x)$ полунепрерывно снизу, ибо для любого $v_0 \in [-1, 1]$ и любой последовательности $\{x_k\}$ такой, что $x_k \rightarrow 0$, найдется последовательность $\{v_k\}$ такая, что $v_k = (v_0 + x_k) \in A(x_k)$, $v_k \rightarrow v_0$.

С другой стороны, если $x_k \rightarrow 0$, $v_k \in [-1+x_k, 1+x_k] \subset A(x_k)$, $v_k \rightarrow v_0$, то $v_0 \in [-1, 1]$, т. е. $v_0 \in A(x_0)$; значит, отображение $A(x)$ полунепрерывно сверху.

Итак, $A(x)$ является п. н. св. и п. н. сн., т. е. K -непрерывно в точке $x_0 = 0$.

Но при $|x^{-1}| - |x| > 1$

$\rho(A(x), A(x_0)) \geq |x^{-1}| - 1 - |x| + |x| = |x^{-1}| - 1 \rightarrow +\infty$, т. е. т. м. о. $A(x)$ не является H -непрерывным в точке $x_0 = 0$.

2. Пусть $A \subset E_n$, $B \subset E_n$.

Лемма 2.1. Если

$$\sup_{v \in A} (v, g) = \sup_{v \in B} (v, g) \quad (2.9)$$

для всех $g \in E_n$, то

$$\overline{\text{co } A} = \overline{\text{co } B}. \quad (2.10)$$

Доказательство. Допустим противное. Пусть $\overline{\text{co } A} \neq \overline{\text{co } B}$. Тогда найдется либо точка $v_0 \in \overline{\text{co } A}$ такая, что $v_0 \notin \overline{\text{co } B}$, либо точка $w_0 \in \overline{\text{co } B}$ такая, что $w_0 \notin \overline{\text{co } A}$. Пусть, например, нашлась точка $v_0 \in \overline{\text{co } A}$, $v_0 \notin \overline{\text{co } B}$. Тогда по теореме отделимости существуют число $a > 0$ и вектор g_0 , $\|g_0\| = 1$, такие, что

$$(w - v_0, g_0) \leq -a \quad \forall w \in \overline{\text{co } B}. \quad (2.11)$$

Так как $v_0 \in \overline{\text{co } A}$, то (см. теорему 1.1)

$$v_0 = \lim_{k \rightarrow \infty} v_k, \quad v_k = \sum_{i=1}^{n+1} \alpha_{ki} v_{ki},$$

$$v_{ki} \in A, \quad \alpha_{ki} \geq 0, \quad \sum_{i=1}^{n+1} \alpha_{ki} = 1.$$

Тогда

$$(v_k, g_0) = \sum \alpha_{ki} (v_{ki}, g_0) \leq \sum \alpha_{ki} \sup_{v \in A} (v, g_0) = \sup_{v \in A} (v, g_0).$$

Переходя к пределу, имеем

$$(v_0, g_0) \leq \sup_{v \in A} (v, g_0). \quad (2.12)$$

Поскольку $B \subset \overline{\text{co } B}$, то из (2.11) и (2.12) следует

$$\sup_{w \in B} (w, g_0) \leq (v_0, g_0) - a \leq \sup_{v \in A} (v, g_0) - a,$$

что противоречит (2.9). Лемма доказана.

Следствие. Если A и B — замкнутые выпуклые множества, то (2.9) эквивалентно соотношению $A = B$.

Легко показать, что

$$\sup_{v \in A} (v, g) = \sup_{v \in \text{co } A} (v, g) = \sup_{v \in \overline{\text{co } A}} (v, g). \quad (2.13)$$

Замечание 3. Рассуждая, как при доказательстве леммы 2.1, можно показать, что если

$$\sup_{v \in A} (v, g) \geq \sup_{v \in B} (v, g) \quad \forall g \in E_n,$$

то $\overline{\text{co } A} \supset \overline{\text{co } B}$,

3. Пусть множество $A(x)$ замкнуто, выпукло и ограничено при каждом

$$x \in \text{int } S_\delta(x_0),$$

где

$$\delta > 0, \quad S_\delta(x_0) = \{x \in E_n \mid \|x - x_0\| \leq \delta\}.$$

Положим

$$h(x, g) = \max_{v \in A(x)} (v, g).$$

Лемма 2.2. Если функция $h(x, g)$ непрерывна на $S_\delta(x_0) \times E_n$, то отображение $A(x)$ H -непрерывно в точке x_0 .

Доказательство. Установим вначале, что отображение $A(x)$ является полунепрерывным сверху в точке x_0 . Пусть $x_k \rightarrow x_0$, $v_k \in A(x_k)$, $v_k \rightarrow v_0$. Надо показать, что

$$v_0 \in A(x_0). \quad (2.14)$$

Допустим противное. Пусть $v_0 \notin A(x_0)$. По теореме отделимости найдутся число $a > 0$ и вектор $g_0 \in E_n$, $\|g_0\| = 1$, такие, что

$$(g_0, v - v_0) \leq -a < 0 \quad \forall v \in A(x_0).$$

Отсюда

$$h(x_0, g_0) = \max_{v \in A(x_0)} (v, g_0) \leq (v_0, g_0) - a. \quad (2.15)$$

Но $h(x_k, g_0) = \max_{v \in A(x_k)} (v, g_0) \geq (v_k, g_0)$. Переходя здесь к пределу при $k \rightarrow \infty$ и учитывая непрерывность $h(x, g_0)$ по x и неравенство (2.15), имеем

$$h(x_0, g_0) \geq (v_0, g_0) \geq \max_{v \in A(x_0)} (v, g_0) + a = h(x_0, g_0) + a,$$

что невозможно, ибо $a > 0$. Полученное противоречие доказывает полунепрерывность сверху т. м. о. $A(x)$ в точке x_0 .

Установим теперь полунепрерывность снизу $A(x)$. Пусть $v_0 \in A(x_0)$. Возьмем произвольную последовательность $\{x_k\}$, $x_k \rightarrow x_0$. Надо показать, что найдутся такие $v_k \in A(x_k)$, что $v_k \rightarrow v_0$. Допустим противное. Без ограничения общности можно считать, что

$$\rho(A(x_k), v_0) = \min_{v \in A(x_k)} \|v - v_0\| = \|v_k - v_0\| \geq a > 0.$$

В силу необходимого условия минимума функции $\varphi(v) = \|v - v_0\|^2$ на множестве $A(x_k)$ (см. замечание 3 § 1)

$$(\|v_k - v_0\|^{-1}(v_k - v_0), v - v_0) \geq a \quad \forall v \in A(x_k). \quad (2.16)$$

Положим

$$g_k = -\|v_k - v_0\|^{-1}(v_k - v_0).$$

Без ограничения общности можем считать, что $g_k \rightarrow g_0$. Тогда из (2.16)

$$\begin{aligned} h(x_k, g_k) &= \max_{v \in A(x_k)} (v, g_k) \leq (v_0, g_k) - a \leq \\ &\leq \max_{v \in A(x_0)} (v, g_k) - a = h(x_0, g_k) - a. \end{aligned}$$

Переходя здесь к пределу при $k \rightarrow \infty$ и учитывая непрерывность h по x и g , получим

$$h(x_0, g_0) \leq h(x_0, g_0) - a,$$

что невозможно, так как $a > 0$. Противоречие и доказывает полунепрерывность снизу т.м.о. $A(x)$. Ранее уже установлена полунепрерывность сверху. Значит, $A(x)$ является K -непрерывным. Так как $A(x)$ ограничено в окрестности точки x_0 , то по теореме 2.1 $A(x)$ H -непрерывно. Лемма доказана.

Замечание 4. Ограниченности $A(x)$ можно и не требовать, она будет следовать из непрерывности $h(x, g)$ по g . Правда, тогда надо определить $h(x, g)$ так:

$$h(x, g) = \sup_{v \in A(x)} (v, g). \quad (2.17)$$

Лемма 2.3 Пусть на Ω_1 задано полунепрерывное сверху т.м.о. $A: \Omega_1 \rightarrow \Pi(\Omega_2)$, где Ω_2 — компакт из E_m , а на $\Omega_1 \times \Omega_2$ задано полунепрерывное сверху т.м.о. $B: \Omega_1 \times \Omega_2 \rightarrow \Pi(E_p)$. Тогда точечно-множественное отображение $C(x) = \bigcup_{z \in A(x)} B(x, z)$ полунепрерывно сверху на Ω_1 .

Доказательство. Пусть $x^* \in \Omega_1$, $x_k \rightarrow x^*$, $v_k \rightarrow v^*$, $v_k \in C(x_k)$. Так как $v_k \in C(x_k)$, то при каждом k найдется такая точка $z_k \in A(x_k)$, что $v_k \in B(x_k, z_k)$. Выделим из последовательности $\{z_k\}$ сходящуюся подпоследовательность $\{z_{k_s}\}$, $z_{k_s} \rightarrow z^*$. В силу того, что $A(x)$ полунепрерывно сверху, будет $z^* \in A(x^*)$.

Имеем

$$(x_{k_s}, z_{k_s}) \rightarrow (x^*, z^*), \quad v_{k_s} \rightarrow v^*, \quad v_{k_s} \in B(x_{k_s}, z_{k_s}).$$

Поскольку $B(x, z)$ полунепрерывно сверху, то $v^* \in B(x^*, z^*)$, т. е. $v^* \in C(x^*)$. Лемма доказана.

Лемма 2.4. Пусть точечно-множественное отображение

$$G: \Omega_1 \rightarrow \Pi(\Omega_2), \quad \Omega_1 \subset E_n, \quad \Omega_2 \subset E_m,$$

полунепрерывно сверху и ограничено на Ω_1 . Тогда $\text{co } G(x)$ полунепрерывно сверху на Ω_1 .

Доказательство. Пусть

$$x^* \in \Omega_1, \quad x_s \in \Omega_1, \quad x_s \rightarrow x^*, \quad v_s \rightarrow v^*, \quad v_s \in \text{co } G(x_s).$$

Требуется установить, что $v^* \in \text{co } G(x^*)$.

По теореме Каратеодори (теорема 1.1) при каждом s справедливо представление

$$v_s = \sum_{k=1}^{m+1} \alpha_{sk} v_{sk},$$

где $\alpha_{sk} \geq 0$, $\sum_{k=1}^{m+1} \alpha_{sk} = 1$, $v_{sk} \in G(x_s) \quad \forall k \in 1:(m+1)$. Так как $G(x)$ ограничено, то можно выделить такую последовательность $\{s_i\}$, что

$$v_{s_i k} \xrightarrow{i \rightarrow \infty} v_k^*, \quad \alpha_{s_i k} \rightarrow \alpha_k^*, \quad k \in 1:(m+1).$$

Тогда

$$v^* = \sum_{k=1}^{m+1} \alpha_k^* v_k^*, \quad \alpha_k^* \geq 0, \quad \sum_{k=1}^{m+1} \alpha_k^* = 1,$$

и поскольку

$$x_{s_i} \xrightarrow{i \rightarrow \infty} x^*, \quad v_{s_i k} \rightarrow v_k^*, \quad v_{s_i k} \in G(x_{s_i}),$$

а $G(x)$ полунепрерывно сверху, то $v_k^* \in G(x^*)$ при каждом $k \in 1:(m+1)$, что завершает доказательство. ■

Задание 2.4. Доказать, что если $G_1(x)$ и $G_2(x)$ полунепрерывны сверху и ограничены на Ω_1 , то отображение $\text{co}\{G_1(x) \cup G_2(x)\}$ полунепрерывно сверху на Ω_1 .

§ 3. Выпуклый конус. Конус возможных направлений. Сопряженный конус

1. Определение 1. Множество $\Gamma \subset E_n$ называется *конусом*, если вместе с точкой x оно содержит и точку $v = \lambda x$ для любого $\lambda > 0$.

Пусть $\Omega \subset E_n$ — замкнутое выпуклое множество и $x_0 \in \Omega$. Рассмотрим множество

$$\gamma(x_0) = \{v = \lambda(x - x_0) \mid \lambda > 0, x \in \Omega\}.$$

Ясно, что $\gamma(x_0)$ — конус, причем $0 \in \gamma(x_0)$.

Замыкание конуса $\gamma(x_0)$ называется *конусом возможных направлений* множества Ω в точке x_0 и обозначается $\Gamma(x_0)$.

Замечание 1. Если $x_0 \in \text{int } \Omega$, то очевидно, что

$$\gamma(x_0) = \Gamma(x_0) = E_n.$$

Лемма 3.1. Если Ω — замкнутое выпуклое множество и $x_0 \in \Omega$, то $\Gamma(x_0)$ — замкнутый выпуклый конус.

Доказательство. То, что $\Gamma(x_0)$ — замкнутый конус, видно из определения замкнутого множества. Для доказательства выпуклости $\Gamma(x_0)$ достаточно убедиться в выпуклости $\gamma(x_0)$. Пусть $v_1, v_2 \in \gamma(x_0)$, т. е. $v_1 = \lambda_1(x_1 - x_0)$, $v_2 = \lambda_2(x_2 - x_0)$, где $\lambda_1 > 0$, $\lambda_2 > 0$, $x_1, x_2 \in \Omega$. Надо показать, что

$$v_\alpha = \alpha v_1 + (1 - \alpha) v_2 \in \gamma(x_0) \quad \text{при} \quad \alpha \in [0, 1].$$

Положим

$$\bar{\lambda} = \alpha \lambda_1 + (1 - \alpha) \lambda_2 > 0, \quad \alpha_0 = \bar{\lambda}^{-1} \alpha \lambda_1, \quad 0 \leq \alpha_0 \leq 1.$$

В силу выпуклости Ω точка $\bar{x} = \alpha_0 x_1 + (1 - \alpha_0) x_2 \in \Omega$. Тогда

$$\begin{aligned} v_\alpha &= \alpha \lambda_1 (x_1 - x_0) + (1 - \alpha) \lambda_2 (x_2 - x_0) = \\ &= \bar{\lambda} [\alpha_0 (x_1 - x_0) + (1 - \alpha_0) (x_2 - x_0)] = \bar{\lambda} (\bar{x} - x_0) \in \gamma(x_0), \end{aligned}$$

что и требовалось доказать. ■

Определение 2. Пусть $\Gamma \subset E_n$ — конус. *Сопряженным ему конусом* Γ^+ называется множество

$$\Gamma^+ = \{w \in E_n \mid (w, v) \geq 0 \quad \forall v \in \Gamma\}.$$

Примеры.

1. Если $\Gamma = E_n$, то $\Gamma^+ = \{0\}$.
2. Если $\Gamma = \{x = (x^{(1)}, \dots, x^{(n)}) \mid x^{(i)} \geq 0 \forall i \in 1:n\}$, то $\Gamma^+ = \Gamma$.
3. Если $\Gamma = \{x \in E_n \mid (A, x) \leq 0\}$, где $A \in E_n$, то $\Gamma^+ = \{x = -\lambda A \mid \lambda \geq 0\}$.
4. Если $\Gamma = \{0\}$, то $\Gamma^+ = E_n$.
5. Если $\Gamma(x_0)$ — конус возможных направлений множества Ω в точке x_0 и $x_0 \in \text{int } \Omega$, то, как уже отмечалось, $\Gamma(x_0) = E_n$, и тогда $\Gamma^+(x_0) = \{0\}$.

Задание 3.1. Пусть $\Gamma = \{x \in E_n \mid (A, x) = 0\}$. Показать, что $\Gamma^+ = \{v = \lambda A \mid \lambda \in (-\infty, \infty)\}$.

Лемма 3.2. Если $\Gamma \subset E_n$ — конус, то $\Gamma^+ \subset E_n$ является замкнутым выпуклым конусом.

Доказательство следует из определения сопряженного конуса.

Задание 3.2. Показать, что если $\Gamma(x_0)$ — конус возможных направлений множества Ω в точке x_0 , то справедливо соотношение

$$\Gamma^+(x_0) = \{w \in E_n \mid (x - x_0, w) \geq 0 \forall x \in \Omega\}. \quad (3.1)$$

Теорема 3.1. Пусть $\Gamma \subset E_n$ — замкнутый выпуклый конус, $G \subset E_n$ — выпуклый компакт. Для того чтобы Γ и G не имели общих точек, т. е.

$$\Gamma \cap G = \emptyset, \quad (3.2)$$

необходимо и достаточно, чтобы нашелся вектор $w_0 \in \Gamma^+$ такой, что

$$\max_{x \in G} (w_0, x) < 0. \quad (3.3)$$

Доказательство. Достаточность. Так как $w_0 \in \Gamma^+$, то по определению сопряженного конуса

$$(w_0, v) \geq 0 \quad \forall v \in \Gamma. \quad (3.4)$$

По предположению (см. (3.2))

$$(w_0, x) < 0 \quad \forall x \in G. \quad (3.5)$$

Неравенства (3.4) и (3.5) означают, что Γ и G не имеют общих точек.

Необходимость. Пусть Γ и G не имеют общих точек. Найдем (рис. 4)

$$\rho = \min_{x \in G, v \in \Gamma} \|x - v\| = \|x_0 - v_0\|. \quad (3.6)$$

В силу условия (3.2) $\|x_0 - v_0\| > 0$. Возьмем $g_0 = \|v_0 - x_0\|^{-1}(v_0 - x_0)$. Покажем вначале, что

$$(v_0, g_0) = 0. \quad (3.7)$$

Допустим противное, т. е. что $(v_0, g_0) = a \neq 0$. Тогда $v_0 \neq 0$. Рассмотрим точки $v_\lambda = \lambda v_0$. При $\lambda > 0$ будет $v_\lambda \in \Gamma$ (ибо Γ — конус).

Имеем

$$\begin{aligned} (v_\lambda - x_0)^2 &= (v_0 - x_0 + \\ &+ (\lambda - 1)v_0)^2 = \\ &= \|x_0 - v_0\|^2 + \\ &+ 2(\lambda - 1)a\|v_0 - x_0\| + \\ &+ (\lambda - 1)^2 v_0^2. \end{aligned}$$

При λ , достаточно близких к единице и таких, что

$$\text{sign}(\lambda - 1) = \text{sign } a,$$

будет $\|v_\lambda - x_0\|^2 \leq \|x_0 - v_0\|^2$, что противоречит (3.6).

Так же как и при доказательстве теоремы 1.2 (неравенство (1.10)), показывается, что

$$(x, g_0) \leq -\|x_0 - v_0\| \quad \forall x \in G. \quad (3.8)$$

Установим, что

$$(v, g_0) \geq 0 \quad \forall v \in \Gamma. \quad (3.9)$$

Действительно, с учетом (3.7) имеем

$$\begin{aligned} (v, g_0) &= (v - x_0, g_0) + \\ &+ (x_0 - v_0, g_0) + (v_0, g_0) \geq \rho - \rho = 0. \end{aligned} \quad (3.10)$$

Положим $w_0 = g_0$. Тогда из (3.8) вытекает (3.3), а из (3.9) заключаем, что $w_0 = g_0 \in \Gamma^+$. Теорема доказана.

Теорема 3.2 Пусть $\Gamma \subset E_n$ — замкнутый выпуклый конус, $G \subset E_n$ — выпуклый компакт. Для того чтобы

$$\text{int } \Gamma \cap G = \emptyset, \quad \Gamma \cap \text{int } G = \emptyset, \quad (3.11)$$

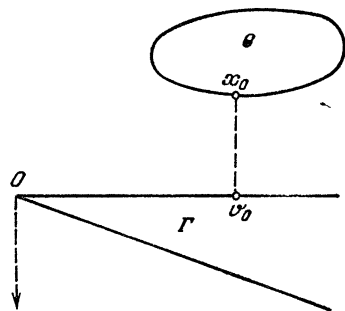


Рис. 4.

необходимо и достаточно, чтобы нашелся вектор $w_0 \in \Gamma^+$ такой, что

$$\max_{x \in G} (w_0, x) \leq 0. \quad (3.12)$$

Задание 3.3. Доказать теорему 3.2.

2. Лемма 3.3. Если $\Gamma \subset E_n$ — замкнутый выпуклый конус, то

$$\Gamma^{++} = \Gamma. \quad (3.13)$$

Доказательство. Пусть $v_0 \in \Gamma$. Тогда по определению сопряженного конуса для любого $w \in \Gamma^+$ будет $(w, v_0) \geq 0$, т. е. $v_0 \in \Gamma^{++}$; значит, $\Gamma \subset \Gamma^{++}$.

Докажем обратное включение. Допустим противное. Пусть $v_0 \in \Gamma^{++}$, но $v_0 \notin \Gamma$. Положим $G = \{v_0\}$. По теореме 3.1 существует $w_0 \in \Gamma^+$ такой, что $(w_0, v_0) < 0$, т. е. $v_0 \notin \Gamma^{++}$, что противоречит предположению. ■

Следствие 1. Если $\Gamma \subset E_n$ — выпуклый конус и $\bar{\Gamma}$ — его замыкание, то

$$\Gamma^{++} = \bar{\Gamma}. \quad (3.14)$$

Доказательство. $\bar{\Gamma}$ — замкнутый конус, и по (3.13) $\bar{\Gamma}^{++} = \bar{\Gamma}$. Но $\bar{\Gamma}^+ = \Gamma^+$. ■

Следствие 2. Если $\Gamma_i \subset E_n$, $i \in 1:s$, — выпуклые конусы и

$$\sum_{i=1}^s \Gamma_i = \left\{ v = \sum_{i=1}^s v_i \mid v_i \in \Gamma_i, i \in 1:s \right\},$$

то

$$\left(\sum_{i=1}^s \Gamma_i \right)^+ = \bigcap_{i=1}^s \Gamma_i^+. \quad (3.15)$$

Это следует из определения сопряженного конуса и того факта, что $0 \in \bar{\Gamma}_i$, $i \in 1:s$. ■

Следствие 3. Если $\Gamma_i \subset E_n$, $i \in 1:s$, — замкнутые выпуклые конусы, то

$$\left(\bigcap_{i=1}^s \Gamma_i \right)^+ = \overline{\sum_{i=1}^s \Gamma_i^+}. \quad (3.16)$$

Доказательство.

$$\left(\bigcap_{i=1}^s \Gamma_i \right)^+ = \left(\bigcap_{i=1}^s (\Gamma_i^+)^+ \right)^+ = \left(\sum_{i=1}^s \Gamma_i^+ \right)^{++} = \overline{\sum_{i=1}^s \Gamma_i^+}. \quad \blacksquare$$

Задание 3.4. Доказать, что если $\mathcal{K}_1, \dots, \mathcal{K}_m$ — замкнутые конусы в E_n , то либо $\mathcal{K}_1 + \dots + \mathcal{K}_m$ — замкнутый конус, либо существует нетривиальное представление нуля. Последнее означает, что найдутся $x_i \in \mathcal{K}_i$ такие, что $x_1 + \dots + x_m = 0$ и не все x_i равняются нулю.

Замечание 2. В (3.16) знак замыкания существен. Приведем пример, из которого следует, что сумма двух замкнутых выпуклых конусов может не быть замкнутым множеством.

Пример 1. Рассмотрим два конуса в E_3 :

$$\mathcal{K}_1 = \{x = (x^{(1)}, x^{(2)}, x^{(3)}) \mid x^{(1)} = x^{(3)} = 0, x^{(2)} \leq 0\},$$

$$\mathcal{K}_2 = \{x = (x^{(1)}, x^{(2)}, x^{(3)}) \mid (x^{(1)})^2 + (x^{(3)})^2 \leq 2x^{(2)}x^{(3)}, x^{(2)} \geq 0\}$$

(образующей конуса $\mathcal{K}_2$ является окружность $(x^{(1)})^2 + (x^{(3)})^2 = 2x^{(3)}$ в плоскости $x^{(2)} = 1$). Очевидно, что $\mathcal{K}_1$ и $\mathcal{K}_2$ — замкнутые выпуклые конусы. Однако их сумма

$$\mathcal{K}_1 + \mathcal{K}_2 = \{x \in E_3 \mid x^{(3)} > 0\} \cup \{x \mid x^{(1)} = x^{(3)} = 0\}$$

не является замкнутым конусом.

Задание 3.5. Доказать, что если множества $\Omega_1, \dots, \Omega_s$ многогранные (т.е. каждое из них задано конечным числом неравенств вида $(x, A_i) \leq b_i \forall i \in 1:m$), то их сумма (множество $\Omega_1 + \dots + \Omega_s$) является замкнутым множеством.

3. Пусть G — произвольное множество из E_n . Через $\mathcal{K}(G)$ обозначим выпуклый конус, натянутый на G :

$$\mathcal{K}(G) = \left\{ v = \sum_{k=1}^r \alpha_k v_k \mid v_k \in G, \alpha_k \geq 0, \right. \\ \left. k \in 1:r, r \text{ — любое натуральное число} \right\}.$$

Конус $\mathcal{K}(G)$ называется *выпуклой конической оболочкой* множества G .

Задание 3.6. Показать, что если Γ — выпуклый конус, то $\Gamma + \Gamma = \Gamma$. Почему важна выпуклость?

Лемма 3.4. Любая точка $v \in \mathcal{K}(G)$ может быть представлена в виде

$$v = \sum_{k=1}^r \alpha_k v_k, \quad v_k \in G, \quad \alpha_k > 0, \quad k \in 1:r,$$

где $1 \leq r \leq n$ и векторы $v_k, k \in 1:r$, линейно независимы.

Доказательство. Пусть $v \in \mathcal{K}(G)$; тогда

$$v = \sum_{k=1}^r \alpha_k v_k, \quad v_k \in G, \quad \alpha_k \geq 0.$$

Можем считать, что $\alpha_k > 0 \quad \forall k \in 1:r$. Если v_k линейно зависимы, то существуют $\beta_k, k \in 1:r$, такие, что

$$\sum_{k=1}^r \beta_k^2 > 0, \quad \sum_{k=1}^r \beta_k v_k = 0. \quad (3.17)$$

Можем также считать, что среди $\{\beta_k\}$ есть положительные (из (3.17) следует, что среди $\{\beta_k\}$ есть хотя бы один ненулевой коэффициент, будем считать его положительным, изменив для этого, если необходимо, знаки всех β_k на противоположные; при этом (3.17) останется справедливым). Положим

$$\varepsilon = \min_{\{k \mid \beta_k > 0\}} \beta_k^{-1} \alpha_k, \quad \bar{\alpha}_k = \alpha_k - \varepsilon \beta_k, \quad k \in 1:r.$$

Ясно, что $v = \sum_{k=1}^r \bar{\alpha}_k v_k$ и среди $\{\bar{\alpha}_k\}$ есть хотя бы один нулевой коэффициент (а остальные неотрицательны). Таким образом, если в представлении вектора $v \in \mathcal{K}(G)$ имеются линейно зависимые v_k , то можно получить новое представление, в котором участвует меньше векторов. Продолжая так далее, в конце концов придем к представлению, в котором все v_k линейно независимы. При этом ясно, что их число не более n . Лемма доказана.

Лемма 3.5. Пусть $G \subset E_n$ — ограниченное замкнутое множество и $0 \notin \text{co } G$. Тогда $\mathcal{K}(G)$ — замкнутый выпуклый конус.

Доказательство. Ясно, что $\mathcal{K}(G)$ — выпуклый конус. Докажем его замкнутость. Прежде всего заметим, что, как легко проверить,

$$\mathcal{K}(G) = \{v = \lambda v' \mid \lambda \geq 0, v' \in \text{co } G\}. \quad (3.18)$$

Пусть теперь $v_k \rightarrow v^*, v_k \in \mathcal{K}(G)$. Надо доказать, что $v^* \in \mathcal{K}(G)$. В силу (3.18) имеем

$$v_k = \lambda_k v'_k, \quad \lambda_k \geq 0, \quad v'_k \in \text{co } G. \quad (3.19)$$

Так как G — ограниченное множество, то и $\text{co } G$ — ограниченное множество; кроме того, по условию $0 \notin \text{co } G$. Отсюда

$$0 < m \leq \|v'\| \leq M < \infty \quad \forall v' \in \text{co } G.$$

Из (3.19) заключаем, что $\{\lambda_k\}$ — ограниченная последовательность. Последовательность $\{v'_k\}$ тоже ограниченная, ибо $v'_k \in \text{co } G$. Поэтому можно выбрать подпоследовательность $\{k_s\}$ такую, что

$$\lambda_{k_s} \rightarrow \lambda^*, \quad v'_{k_s} \rightarrow v'.$$

Ясно, что $v^* = \lambda^* v'$, $\lambda^* \geq 0$, $v' \in \text{co } G$. Значит, $v^* \in \mathcal{K}(G)$, что и требовалось доказать. ■

З а м е ч а н и е 3. Условие $0 \notin \text{co } G$ в лемме 3.5 существенно, что видно из следующего примера.

Пусть

$$G = \{x = (x^{(1)}, x^{(2)}) \in E_2 \mid (x^{(1)} - 1)^2 + (x^{(2)})^2 \leq 1\}.$$

Здесь $0 \in \text{co } G = G$, $\mathcal{K}(G) = \{0\} \cup \{x = (x^{(1)}, x^{(2)}) \mid x^{(1)} > 0\}$, т. е. $\mathcal{K}(G)$ не является замкнутым множеством.

Но если G состоит из конечного числа точек, то условие $0 \notin \text{co } G$ в лемме 3.5 может быть опущено.

Л е м м а 3.6. Если множество G состоит из конечного числа точек, то $\mathcal{K}(G)$ является замкнутым выпуклым конусом.

Доказательство. Пусть $v_k \rightarrow v^*$, $v_k \in \mathcal{K}(G)$. Требуется доказать, что $v^* \in \mathcal{K}(G)$. ~~При $v^* = 0$ это очевидно следует из определения $\mathcal{K}(G)$. Пусть $v^* \neq 0$.~~

По лемме 3.4 $v_k = \sum_{i=1}^{r_k} \lambda_{ki} v_{ki}$, $\lambda_{ki} > 0$, $v_{ki} \in G$. Поскольку G состоит из конечного числа точек, можно выбрать подпоследовательность наборов $\{v_{k_s i}\}$, содержащую одни и те же векторы, т. е.

$$v_{k_s} = \sum_{i=1}^r \lambda_{k_s i} v_i, \quad \lambda_{k_s i} > 0, \quad v_i \in G, \quad r \leq n.$$

где v_i — линейно независимы

Докажем, что последовательность $\{\lambda_{k_s} = (\lambda_{k_s 1}, \dots, \lambda_{k_s r}) \in E_r\}$ ограничена. Допустим противное. Пусть это не так. Без ограничения общности можем считать,

что ~~вся последовательность $\{\lambda_{k_s}\}$ неограничена, т. е.~~

$\gamma_{k_s} = \sum_{i=1}^r \lambda_{k_s i} \rightarrow \infty$ (напомним, что все $\lambda_{k_s i} > 0$). Тогда

$$v_{k_s} = \gamma_{k_s}^{-1} \sum_{i=1}^r \lambda'_{k_s i} v_i = \gamma_{k_s}^{-1} v'_{k_s},$$

где

$$v'_{k_s} = \sum_{i=1}^r \lambda'_{k_s i} v_i, \quad \lambda'_{k_s i} = \gamma_{k_s}^{-1} \lambda_{k_s i} > 0.$$

Так как

$$\sum_{i=1}^r \lambda'_{k_s i} = 1,$$

то

Можно считать $\lambda'_{k_s i} \rightarrow \lambda_i$, $v'_{k_s} \in \text{co } G$, $\lambda_i \geq 0$, $\sum \lambda_i = 1$

Множество $\text{co } G$ ограничено, поэтому $v'_{k_s} \rightarrow 0$, т. е. $\sum \lambda_i v_i^* = 0$, что противоречит допущению. *линейной независимости* И так, $\{\lambda_{k_s}\}$ — ограниченная последовательность. Выбирая сходящуюся

подпоследовательность, получим, что $v^* = \sum_{i=1}^r \lambda_i^* v_i$, где $v_i \in G$, $\lambda_i^* \geq 0$, т. е. $v^* \in \mathcal{X}(G)$. Лемма доказана.

4. Выше для $x \in \Omega$ были определены точечно-множественные отображения $\gamma(x)$, $\Gamma(x) \equiv \bar{\gamma}(x)$ и $\Gamma^+(x)$.

Лемма 3.7. *Отображение $\gamma(x)$ полунепрерывно снизу на Ω .*

Доказательство. Пусть $x_0 \in \Omega$. Надо показать, что если $v_0 \in \gamma(x_0)$ и $x_i \rightarrow x_0$, $x_i \in \Omega$, то найдется последовательность точек $\{v_i\}$, $v_i \in \gamma(x_i)$, такая, что $v_i \rightarrow v_0$. Если $v_0 = 0$, то утверждение очевидно. Предположим, что $v_0 \neq 0$. Так как $v_0 \in \gamma(x_0)$, то $v_0 = \lambda_0(\bar{x} - x_0)$, где $\lambda_0 > 0$, $\bar{x} \in \Omega$, $\bar{x} \neq x_0$. Тогда $v_i = \lambda_0(\bar{x} - x_i) \in \gamma(x_i)$. Очевидно, что $v_i \rightarrow v_0$. ■

Лемма 3.8. *Отображение $\Gamma^+(x)$ полунепрерывно сверху на Ω .*

Доказательство. Пусть $x_0 \in \Omega$ и $x_k \rightarrow x_0$, $y_k \rightarrow y_0$, $x_k \in \Omega$, $y_k \in \Gamma^+(x_k)$. Надо показать, что $y_0 \in \Gamma^+(x_0)$.

Так как $y_k \in \Gamma^+(x_k)$, то (см. задание 3.2)

$$(x - x_k, y_k) \geq 0 \quad \forall x \in \Omega.$$

Переходя здесь к пределу при фиксированном $x \in \Omega$, имеем $(x - x_0, y_0) \geq 0$. Это неравенство справедливо для любого $x \in \Omega$, а это и означает, что $y_0 \in \Gamma^+(x_0)$. Лемма доказана.

5. Пусть $A \subset E_n$ — компакт, и пусть

$$\Gamma = \{v \in E_n \mid (z, v) \leq 0 \quad \forall z \in A\}.$$

Легко проверить, что Γ — выпуклый конус.

Лемма 3.9. Справедливо соотношение

$$\Gamma^+ = -\overline{\mathcal{K}}(\text{co } A), \quad (3.20)$$

где $\mathcal{K}(G)$ — выпуклая коническая оболочка множества G .

Доказательство. Правую часть (3.20) обозначим B . Пусть $w \in B$, т. е.

$$w = \lim_{k \rightarrow \infty} w_k, \quad w_k = -\lambda_k z_k,$$

$$\lambda_k \geq 0, \quad z_k = \sum_{i=1}^{r_k} \alpha_{ki} z_{ki},$$

$$\alpha_{ki} \geq 0, \quad \sum_{i=1}^{r_k} \alpha_{ki} = 1, \quad z_{ki} \in A \quad \forall i \in 1:r_k.$$

В силу леммы 3.4 без ограничения общности можно считать, что

$$r_k = r \leq n \quad \forall k.$$

Возьмем любое $v \in \Gamma$. По определению множества Γ будет

$$(v, z_{ki}) \leq 0 \quad \forall i \in 1:r.$$

Поэтому

$$(v, w_k) = -\lambda_k \sum_{i=1}^r \alpha_{ki} (z_{ki}, v) \geq 0 \quad \forall v \in \Gamma,$$

а это и значит, что $w_k \in \Gamma^+$. Так как Γ^+ — замкнутое множество, то $w \in \Gamma^+$, т. е.

$$B \subset \Gamma^+. \quad (3.21)$$

Докажем теперь, что

$$\Gamma^+ \subset B. \quad (3.22)$$

Допустим противное. Тогда найдется такое $w_0 \in \Gamma^+$, что $w_0 \notin B$. По теореме отделимости существует $g \in E_n$, $\|g\| = 1$, такое, что

$$(w_0, g) < 0, \quad (3.23)$$

$$(v, g) \geq 0 \quad \forall v \in B. \quad (3.24)$$

Так как для любого $z \in A$ будет $-z \in B$, то из (3.24)

$$(z, g) \leq 0 \quad \forall z \in A.$$

Значит, $g \in \Gamma$. Но $w_0 \in \Gamma^+$. Следовательно, $(g, w_0) \geq 0$, что противоречит (3.23). Итак, (3.22) установлено.

Соотношение (3.20) вытекает теперь из (3.21) и (3.22). ■

§ 4. Выпуклые функции. Непрерывность и дифференцируемость по направлениям

1. Пусть на выпуклом множестве $S \subset E_n$ задана функция $f(x)$. Везде в дальнейшем, если не оговорено противное, функция $f(x)$ предполагается конечной в области своего определения, т. е. принимающей конечные значения в любой точке области определения.

Графиком функции $f(x)$ называется множество

$$\{[\beta, x] \in E_{n+1} \mid x \in S, \beta \in E_1, \beta = f(x)\}.$$

Надграфиком функции $f(x)$, определенной на множестве S , называется множество

$$\text{epi } f = \{[\beta, x] \in E_{n+1} \mid x \in S, \beta \in E_1, \beta \geq f(x)\}.$$

Функцию $f(x)$ будем называть *выпуклой* на S , если

$$f(\alpha x_1 + (1 - \alpha)x_2) \leq \alpha f(x_1) + (1 - \alpha)f(x_2) \quad \forall x_1, x_2 \in S, \quad \forall \alpha \in [0, 1]. \quad (4.1)$$

Надграфик выпуклой функции — выпуклое множество. Функция $f(x)$ называется *строго выпуклой* на S , если

$$f(\alpha x_1 + (1 - \alpha)x_2) < \alpha f(x_1) + (1 - \alpha)f(x_2) \quad (4.2)$$

$$\forall x_1, x_2 \in S, \quad x_1 \neq x_2, \quad \forall \alpha \in (0, 1).$$

Функция $f(x)$ называется *сильно выпуклой* на S , если существует $m > 0$ такое, что

$$f(\alpha x_1 + (1 - \alpha)x_2) \leq \alpha f(x_1) + (1 - \alpha)f(x_2) - \alpha(1 - \alpha)m \|x_1 - x_2\|^2 \quad \forall x_1, x_2 \in S, \quad \forall \alpha \in [0, 1]. \quad (4.3)$$

Функция $f(x)$ называется *вогнутой* на S , если

$$f(\alpha x_1 + (1 - \alpha)x_2) \geq \alpha f(x_1) + (1 - \alpha)f(x_2) \quad \forall x_1, x_2 \in S, \quad \forall \alpha \in [0, 1].$$

Аналогично определяются *строго вогнутые* и *сильно вогнутые* функции. Заметим, что *линейная* функция $f(x) = (A, x) + b$, где $A \in E_n$, $b \in (-\infty, \infty)$, является одновременно и выпуклой, и вогнутой. Из определения также ясно, что если $f_i(x)$, $i \in 1:N$, — выпуклые на S функции, то и функция

$$f(x) = \sum_{i=1}^N \alpha_i f_i(x), \quad \alpha_i \geq 0, \quad i \in 1:N,$$

тоже выпуклая. В частности, выпуклой является и функция $F(x) = Af(x)$, если $A > 0$, а $f(x)$ — выпуклая функция.

Задача 4.1. Доказать, что если функция $f(x)$ выпуклая, $x_i \in S$, $\alpha_i \geq 0$, $i \in 1:p$, $\sum_{i=1}^p \alpha_i = 1$, то

$$f\left(\sum_{i=1}^p \alpha_i x_i\right) \leq \sum_{i=1}^p \alpha_i f(x_i). \quad (4.4)$$

Неравенство (4.4) называется *неравенством Иенсена*.

Задача 4.2. Доказать, что если $f(x)$ — непрерывная на S функция и

$$f(2^{-1}(x_1 + x_2)) \leq 2^{-1}[f(x_1) + f(x_2)] \quad \forall x_1, x_2 \in S, \quad (4.5)$$

то функция $f(x)$ — выпуклая на S .

Утверждение перестает быть верным, если $f(x)$ не является непрерывной.

Замечание 1. Очевидно, если $f_i(x)$, $i \in 1:N$, — выпуклые функции и хотя бы одна из них *сильно* (строго) выпуклая, то и функция $f(x) = \sum_{i=1}^N f_i(x)$ является *сильно* (строго) выпуклой.

Задание 4.3. Доказать, что если $f(x)$ — выпуклая на S функция и $f(x) \geq 0 \forall x \in S$, то функция $f_1(x) = f^2(x)$ тоже выпуклая. Отметим, что условие $f(x) \geq 0$ в этом задании существенно, что видно из примера.

Пример 1. Пусть $x \in E_1$, $f(x) = x^2 - 1$, $f_1(x) = f^2(x)$. Ясно, что $f_1(x)$ не является выпуклой функцией.

Лемма 4.1. Если $f(x)$ — выпуклая на S функция, то для всех $\alpha \in [0, 1]$ и $x_1, x_2 \in S$ таких, что $x_\alpha = \alpha x_1 + (1 - \alpha)x_2 \in S$, будет

$$f(\alpha x_1 + (1 - \alpha)x_2) \geq \alpha f(x_1) + (1 - \alpha)f(x_2). \quad (4.6)$$

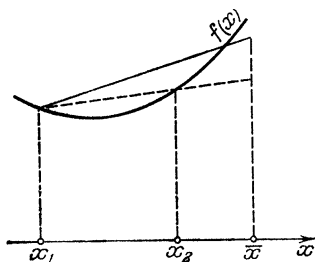


Рис. 5.

Доказательство. Допустим противное. Тогда найдутся $x_1, x_2 \in S$ и точка $\bar{x} = \alpha x_1 + (1 - \alpha)x_2 \in S$, где $\alpha \notin [0, 1]$, такие, что

$$f(\bar{x}) < \alpha f(x_1) + (1 - \alpha)f(x_2). \quad (4.7)$$

Пусть для определенности $\alpha < 0$. Тогда, как нетрудно проверить (рис. 5),

$$x_2 = \beta x_1 + (1 - \beta)\bar{x}, \quad \text{где } \beta = -\alpha(1 - \alpha)^{-1} \in (0, 1).$$

В силу выпуклости $f(x)$ из (4.7) имеем

$$\begin{aligned} f(x_2) &\leq \beta f(x_1) + (1 - \beta)f(\bar{x}) < \\ &< \beta f(x_1) + (1 - \beta)(\alpha f(x_1) + (1 - \alpha)f(x_2)) = \\ &= (\beta + (1 - \beta)\alpha)f(x_1) + (1 - \beta)(1 - \alpha)f(x_2) = f(x_2), \end{aligned}$$

что невозможно. Аналогично приходим к противоречию, если $\alpha > 1$ (в этом случае на рис. 5 надо поменять местами точки x_1 и x_2). ■

Лемма 4.2. Если $f(x)$ — сильно выпуклая на S функция (см. (4.3)), то

$$\begin{aligned} f(\alpha x_1 + (1 - \alpha)x_2) &\geq \\ &\geq \alpha f(x_1) + (1 - \alpha)f(x_2) - \alpha(1 - \alpha)m \|x_1 - x_2\|^2 \end{aligned} \quad (4.8)$$

для всех $x_1, x_2 \in S$ и $\alpha \notin [0, 1]$ таких, что $[\alpha x_1 + (1 - \alpha)x_2] \in S$.

Доказательство аналогично доказательству леммы 4.1. Допустим противное; тогда найдутся $x_1, x_2 \in S$, $x_1 \neq x_2$, $\alpha \notin [0, 1]$ такие, что

$$f(\bar{x}) < \alpha f(x_1) + (1 - \alpha) f(x_2) - \alpha(1 - \alpha) m \|x_1 - x_2\|^2, \quad (4.9)$$

где $\bar{x} = \alpha x_1 + (1 - \alpha) x_2 \in S$.

Пусть для определенности $\alpha < 0$. Тогда снова

$$x_2 = \beta x_1 + (1 - \beta) \bar{x}, \quad \beta = -\alpha(1 - \alpha)^{-1} \in (0, 1).$$

Из (4.3) и (4.9) имеем

$$\begin{aligned} f(x_2) &\leq \beta f(x_1) + (1 - \beta) f(\bar{x}) - \beta(1 - \beta) m \|x_1 - \bar{x}\|^2 < \\ &< \beta f(x_1) + (1 - \beta) [\alpha f(x_1) + (1 - \alpha) f(x_2) - \\ &- \alpha(1 - \alpha) m \|x_1 - x_2\|^2] - \beta(1 - \beta) m \|x_1 - \bar{x}\|^2 = \\ &= f(x_2) + m(1 - \beta) [-\alpha(1 - \alpha) \|x_1 - x_2\|^2 - \\ &- \beta \|x_1 - \bar{x}\|^2]. \end{aligned} \quad (4.10)$$

Но $x_1 - \bar{x} = (1 - \alpha)(x_1 - x_2)$, поэтому выражение в последних квадратных скобках в (4.10) равно нулю. Тогда из (4.10) следует $f(x_2) < f(x_2)$, что невозможно. Аналогично приходим к противоречию при $\alpha > 1$. ■

Задание 4.4. Показать, что если $f(x)$ — строго выпуклая на S функция, то

$$f(\alpha x_1 + (1 - \alpha) x_2) > \alpha f(x_1) + (1 - \alpha) f(x_2) \quad (4.11)$$

при $x_1 \neq x_2$, $\alpha \notin [0, 1]$ и таких, что $\alpha x_1 + (1 - \alpha) x_2 \in S$.

2. Пусть функция $f(x)$ — выпуклая на выпуклом множестве $S \subset E_n$.

Теорема 4.1. Функция $f(x)$ непрерывна в любой внутренней точке множества S .

Доказательство. Пусть $x_0 \in \text{int } S$. Через e_i , $i \in 1:n$, обозначим координатные орты: $e_i = (0, \dots, 0, 1, 0, \dots, 0)$. Поскольку $x_0 \in \text{int } S$, то найдутся $\beta_i > 0$

такие, что $x_0 \pm g_i \in S$, где $g_i = \beta_i e_i$. Положим $g_{n+i} = -g_i$. Пусть $x \rightarrow x_0$. Точку x можно представить в виде $x = x_0 + \sum_{i=1}^{2n} \alpha_i g_i$, где $\alpha_i \geq 0$, причем $\min\{\alpha_i, \alpha_{i+n}\} = 0$

для всех $i \in 1:n$. Тогда $\sum_{i=1}^{2n} \alpha_i \rightarrow 0$ при $x \rightarrow x_0$ (ясно, что

α_i зависят от x). Пусть уже $\sum_{i=1}^{2n} \alpha_i < 1$. Так как

$$x = x_0 + \sum_{i=1}^{2n} \alpha_i g_i = \left(1 - \sum_{i=1}^{2n} \alpha_i\right) x_0 + \sum_{i=1}^{2n} \alpha_i (x_0 + g_i),$$

то в силу выпуклости $f(x)$ имеем

$$f(x) \leq \left(1 - \sum_{i=1}^{2n} \alpha_i\right) f(x_0) + \sum_{i=1}^{2n} \alpha_i f(x_0 + g_i),$$

или

$$f(x) - f(x_0) \leq \sum_{i=1}^{2n} \alpha_i (f(x_0 + g_i) - f(x_0)). \quad (4.12)$$

Точку x_0 можно представить в виде $x_0 = 0,5(x + x')$, где

$$x' = x_0 + \sum_{i=1}^{2n} \alpha'_i g_i, \quad \alpha'_i = \begin{cases} \alpha_{i+n}, & i \leq n, \\ \alpha_{i-n}, & i > n. \end{cases}$$

Ясно, что $\alpha'_i \geq 0$. Поэтому

$$f(x_0) \leq 0,5[f(x) + f(x')].$$

Отсюда

$$\begin{aligned} f(x_0) - f(x) &\leq f(x') - f(x_0) \leq \\ &\leq \sum_{i=1}^{2n} \alpha'_i [f(x_0 + g_i) - f(x_0)]. \end{aligned} \quad (4.13)$$

Непрерывность функции $f(x)$ в точке x_0 следует теперь из (4.12) и (4.13), ибо

$$\alpha_i \geq 0, \quad \alpha'_i \geq 0, \quad \sum_{i=1}^{2n} \alpha_i = \sum_{i=1}^{2n} \alpha'_i \xrightarrow{x \rightarrow x_0} 0. \blacksquare$$

3. Пусть $g \in E_n$, S — выпуклое множество, $x_0 \in S$, существует $\alpha_0 > 0$ такое, что $x_0 + \alpha_0 g \in S$, и пусть на S задана функция $f(x)$.

Функция $f(x)$ называется *дифференцируемой* в точке x_0 по направлению g , если существует конечный предел

$$\frac{\partial f(x_0)}{\partial g} \equiv \lim_{\alpha \rightarrow +0} \alpha^{-1} [f(x_0 + \alpha g) - f(x_0)].$$

Число $\frac{\partial f(x_0)}{\partial g}$ называют в этом случае *производной* функции $f(x)$ в точке x_0 по направлению g . Очевидно, что

$$f(x_0 + \alpha g) = f(x_0) + \alpha \frac{\partial f(x_0)}{\partial g} + o(x_0, g, \alpha),$$

где $\frac{o(x_0, g, \alpha)}{\alpha} \xrightarrow{\alpha \rightarrow +0} 0$.

Если функция $f(x)$ дифференцируема в точке x_0 по любому направлению $g \in E_n$, то будем говорить, что она *дифференцируема в точке x_0 по направлениям*. Из определения ясно, что если $g_1 = \beta g$, где $\beta > 0$, то

$$\frac{\partial f(x_0)}{\partial g_1} = \beta \frac{\partial f(x_0)}{\partial g}.$$

Заметим, что из дифференцируемости по направлениям не следует непрерывность функции в точке.

Пример 2. Пусть $x = (x^{(1)}, x^{(2)}) \in E_2$. Через G обозначим (рис. 6) множество, ограниченное осью $\mathcal{L} = \{(x^{(1)}, x^{(2)}) | x^{(1)} \geq 0, x^{(2)} = 0\}$ и кривой

$$\mathcal{K} = \{(x^{(1)}, x^{(2)}) | x^{(1)} = \cos \varphi \cos \frac{\varphi}{4};$$

$$x^{(2)} = \sin \varphi \cos \frac{\varphi}{4}; \quad \varphi \in [0, 2\pi]\}.$$

Положим

$$f(x) = \begin{cases} 1, & x \in G, \\ 0, & x \notin G. \end{cases}$$

Возьмем $x_0 = (0, 0) = \theta$. Ясно, что для любого $g = (g^{(1)}, g^{(2)})$ существует $\frac{\partial f(x_0)}{\partial g} = 0$, однако функция $f(x)$ разрывна в точке x_0 .

Функция $f(x)$ называется *дифференцируемой* в точке $x_0 \in \text{int } S$, если существует такой вектор $v(x_0)$, что

$$f(x_0 + \alpha g) = f(x_0) + \alpha (g, v(x_0)) + o(x_0, g, \alpha)$$

$$\forall g \in E_n, \quad \|g\| = 1,$$

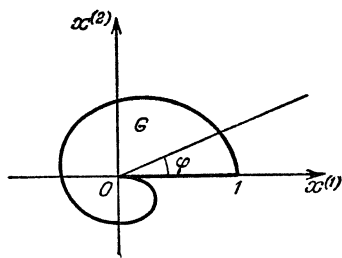


Рис. 6.

где $\frac{o(x_0, g, \alpha)}{\alpha} \xrightarrow{\alpha \rightarrow 0} 0$ равномерно по g . Если функция $f(x)$ дифференцируема в точке x_0 , то она дифференцируема в этой точке по любому направлению g , причем $\frac{\partial f(x_0)}{\partial g} = (g, v(x_0))$. Кроме того, в этом случае функция $f(x)$ имеет частные производные первого порядка в точке x_0 , причем

$$v(x_0) = \frac{\partial f(x_0)}{\partial x} \equiv \left(\frac{\partial f(x_0)}{\partial x^{(1)}}, \dots, \frac{\partial f(x_0)}{\partial x^{(n)}} \right).$$

Вектор $v(x_0) = \frac{\partial f(x_0)}{\partial x}$ называется *градиентом* функции $f(x)$ в точке x_0 . Для функции $f(x)$ в примере 2 имело место представление

$$f(x_0 + \alpha g) = f(x_0) + o(x_0, g, \alpha),$$

но функция $f(x)$ не является дифференцируемой в точке $x_0 = (0, 0)$, ибо $\frac{o(x_0, g, \alpha)}{\alpha} \rightarrow 0$ неравномерно по g , $\|g\| = 1$.

Теорема 4.2. Пусть $f(x)$ — выпуклая на S функция, точка $x_0 \in \text{int } S$. Тогда функция $f(x)$ является дифференцируемой в точке x_0 по любому направлению $g \in E_n$.

Доказательство. Требуется доказать существование конечного предела

$$\frac{\partial f(x_0)}{\partial g} \equiv \lim_{\alpha \rightarrow +0} h(\alpha),$$

где $h(\alpha) = \alpha^{-1} [f(x_0 + \alpha g) - f(x_0)]$. Поскольку $x_0 \in \text{int } S$, то существует $\delta > 0$ такое, что интервал $(x_0 - \delta g, x_0 + \delta g) \subset S$. Пусть $\alpha_0 \in (0, \delta)$. Для $\alpha \in (0, \alpha_0]$ будет

$$f(x_0 + \alpha g) = f(\beta x_0 + (1 - \beta)(x_0 + \alpha_0 g)),$$

где $\beta = \alpha_0^{-1}(\alpha_0 - \alpha) \in (0, 1]$. В силу выпуклости $f(x)$

$$f(x_0 + \alpha g) \leq \beta f(x_0) + (1 - \beta) f(x_0 + \alpha_0 g).$$

Отсюда

$$\begin{aligned} h(\alpha) &= \alpha^{-1} [f(x_0 + \alpha g) - f(x_0)] \leq \\ &\leq \alpha^{-1} (1 - \beta) (f(x_0 + \alpha_0 g) - f(x_0)) = \\ &= \alpha_0^{-1} [f(x_0 + \alpha_0 g) - f(x_0)] \equiv h(\alpha_0). \end{aligned}$$

Таким образом, $h(\alpha)$ не возрастает при $\alpha \rightarrow +0$. С другой стороны,

$$\begin{aligned} f(x_0) &= f((\alpha + \alpha_0)^{-1} \alpha_0 (x_0 + \alpha g) + (\alpha + \alpha_0)^{-1} \alpha (x_0 - \alpha_0 g)) \leq \\ &\leq (\alpha + \alpha_0)^{-1} \alpha_0 f(x_0 + \alpha g) + (\alpha + \alpha_0)^{-1} \alpha f(x_0 - \alpha_0 g). \end{aligned}$$

Поэтому

$$f(x_0 + \alpha g) \geq \alpha_0^{-1} (\alpha + \alpha_0) f(x_0) - \alpha_0^{-1} \alpha f(x_0 - \alpha_0 g).$$

Отсюда

$$h(\alpha) \geq \alpha_0^{-1} [f(x_0) - f(x_0 - \alpha_0 g)],$$

т.е. функция $h(\alpha)$ ограничена снизу при $\alpha \rightarrow +0$. Существование конечного предела $\frac{\partial f(x_0)}{\partial g} \equiv \lim_{\alpha \rightarrow +0} h(\alpha)$ следует теперь из монотонности и ограниченности функции $h(\alpha)$. Теорема доказана.

Следствие. Из доказательства теоремы ясно, что

$$\frac{\partial f(x_0)}{\partial g} = \inf_{\alpha > 0} \alpha^{-1} [f(x_0 + \alpha g) - f(x_0)].$$

Лемма 4.3. Если $f(x)$ выпукла на открытом выпуклом множестве S и дифференцируема там, то

$$f(x_0 + g) \geq f(x_0) + (f'(x_0), g)$$

для всех $x_0 \in S$ и $g \in E_n$ таких, что $x_0 + g \in S$. Здесь

$$f'(x_0) = \frac{\partial f(x_0)}{\partial x} \equiv \left(\frac{\partial f(x_0)}{\partial x^{(1)}}, \dots, \frac{\partial f(x_0)}{\partial x^{(n)}} \right).$$

Доказательство. Положим $x_1 = x_0 + g$. В силу выпуклости $f(x)$

$$f(x_0 + \alpha g) = f(\alpha x_1 + (1 - \alpha) x_0) \leq$$

$$\leq \alpha f(x_1) + (1 - \alpha) f(x_0) \quad \forall \alpha \in [0, 1].$$

Отсюда

$$\alpha^{-1} [f(x_0 + \alpha g) - f(x_0)] \leq f(x_1) - f(x_0).$$

Переходя к пределу при $\alpha \rightarrow +0$, имеем

$$(f'(x_0), g) \leq f(x_1) - f(x_0) = f(x_0 + g) - f(x_0),$$

что и требовалось доказать.

4. Пусть

$$f(x) = \sup_{y \in G} \varphi(x, y), \quad (4.14)$$

где $x \in \Omega \subset E_n$, Ω — выпуклое множество, G — произвольное множество любого пространства, функция $\varphi(x, y)$ — выпуклая по x на Ω при каждом фиксированном $y \in G$. Будем также предполагать, что функция $f(x)$, заданная соотношением (4.14), конечна на Ω .

Легко убедиться, что тогда $f(x)$ тоже выпукла по x на Ω . Действительно, положим $x_\alpha = \alpha x_1 + (1 - \alpha)x_2$, где $x_1, x_2 \in \Omega$, $\alpha \in [0, 1]$.

Так как в силу выпуклости $\varphi(x, y)$ по x

$$\varphi(x_\alpha, y) \leq \alpha \varphi(x_1, y) + (1 - \alpha) \varphi(x_2, y),$$

то отсюда

$$\begin{aligned} f(x_\alpha) &= \sup_{y \in G} \varphi(x_\alpha, y) \leq \sup_{y \in G} [\alpha \varphi(x_1, y) + (1 - \alpha) \varphi(x_2, y)] \leq \\ &\leq \alpha \sup_{y \in G} \varphi(x_1, y) + (1 - \alpha) \sup_{y \in G} \varphi(x_2, y) = \\ &= \alpha f(x_1) + (1 - \alpha) f(x_2), \end{aligned}$$

а это и означает выпуклость $f(x)$ на Ω . ■

5. Пусть теперь

$$f(x) = \inf_{y \in G} \varphi(x, y), \quad (4.15)$$

где $x \in \Omega \subset E_n$, Ω — выпуклое множество, G — выпуклое множество произвольного пространства, функция $\varphi(x, y)$ — выпуклая по совокупности переменных $[x, y]$ на $\Omega \times G$. Предположим также, что функция $f(x)$, заданная (4.15), конечна на Ω .

Лемма 4.4. *Функция $f(x)$, заданная соотношениями (4.15), выпукла на Ω .*

Доказательство. Возьмем любые $x_1, x_2 \in \Omega$, и пусть последовательности $\{y_{1k}\}$ и $\{y_{2k}\}$ таковы, что

$$\begin{aligned} y_{ik} &\in \Omega \quad \forall i \in 1:2, \quad \forall k \in 1:\infty, \\ \varphi(x_i, y_{ik}) &\xrightarrow{k \rightarrow \infty} f(x_i) = \inf_{y \in G} \varphi(x_i, y). \end{aligned} \quad (4.16)$$

Положим

$$x_\alpha = \alpha x_1 + (1 - \alpha)x_2, \quad y_{\alpha k} = \alpha y_{1k} + (1 - \alpha)y_{2k}, \quad \alpha \in [0, 1].$$

Поскольку функция $\varphi(x, y)$ выпукла по совокупности переменных, то

$$\begin{aligned}\varphi(x_\alpha, y_{\alpha k}) &= \varphi(\alpha x_1 + (1 - \alpha)x_2, \alpha y_{1k} + (1 - \alpha)y_{2k}) \leq \\ &\leq \alpha \varphi(x_1, y_{1k}) + (1 - \alpha) \varphi(x_2, y_{2k}).\end{aligned}$$

Поэтому

$$\begin{aligned}f(x_\alpha) &= \inf_{y \in G} \varphi(x_\alpha, y) \leq \varphi(x_\alpha, y_{\alpha k}) \leq \\ &\leq \alpha \varphi(x_1, y_{1k}) + (1 - \alpha) \varphi(x_2, y_{2k}).\end{aligned}$$

Отсюда и из (4.16) получаем, переходя к пределу,

$$f(x_\alpha) \leq \alpha f(x_1) + (1 - \alpha) f(x_2) \quad \forall \alpha \in [0, 1],$$

а это и означает, что функция $f(x)$ выпукла по x на Ω . ■

Замечание 2. Для выпуклости функции (4.14) не требовалась выпуклость множества G , а для выпуклости функции (4.15) это условие существенно. В условии леммы предполагалось, что функция $\varphi(x, y)$ выпукла по совокупности переменных x и y . Приводимый ниже пример показывает, что выпуклости по каждой из переменных x и y недостаточно для выпуклости по совокупности переменных.

Пример 3. Пусть $x \in E_1$, $y \in E_1$, $\varphi(x, y) = xy$. Функция $\varphi(x, y)$ является выпуклой по x при каждом фиксированном y и выпуклой по y при каждом фиксированном x . Положим $x_\alpha = \alpha x_1 + (1 - \alpha)x_2$, $y_\alpha = \alpha y_1 + (1 - \alpha)y_2$, где $\alpha \in [0, 1]$. Для выпуклости $\varphi(x, y)$ по совокупности переменных необходимо, чтобы для всех x_1, x_2, y_1, y_2 и $\alpha \in [0, 1]$ выполнялось неравенство

$$\varphi(x_\alpha, y_\alpha) \leq \alpha \varphi(x_1, y_1) + (1 - \alpha) \varphi(x_2, y_2). \quad (4.17)$$

Имеем

$$\begin{aligned}F(\alpha) &\equiv \varphi(x_\alpha, y_\alpha) - \alpha \varphi(x_1, y_1) - (1 - \alpha) \varphi(x_2, y_2) = \\ &= (\alpha x_1 + (1 - \alpha)x_2)(\alpha y_1 + (1 - \alpha)y_2) - \alpha x_1 y_1 - \\ &\quad - (1 - \alpha)x_2 y_2 = \alpha(1 - \alpha)(y_1 - y_2)(x_2 - x_1).\end{aligned}$$

Очевидно, что если $\alpha \in (0, 1)$, $x_1 \neq x_2$, $y_1 \neq y_2$ и $\text{sign}(y_1 - y_2) = \text{sign}(x_2 - x_1)$, то $F(\alpha) > 0$, а тогда неравенство (4.17) не выполняется, т. е. функция $\varphi(x, y)$ не является выпуклой по совокупности переменных.

Замечание 3. Вообще говоря, операция $\inf$ не сохраняет выпуклости, т. е. если функция $\varphi(x, y)$ выпукла по x при любом фиксированном $y \in G$, то функция $f(x)$, заданная (4.15), не обязательно выпукла.

Пример 4. Пусть $x \in E_1$, $y \in 1:2$, $\varphi(x, 1) = x$, $\varphi(x, 2) = -x$. Тогда

$$f(x) = \inf_{y \in 1:2} \varphi(x, y) = \begin{cases} -x, & x \geq 0, \\ x, & x < 0. \end{cases}$$

Очевидно, что функция $f(x)$ не является выпуклой. Таким образом, в то время как операция $\sup$ сохраняет свойство выпуклости (см. п. 4), операция $\inf$, вообще говоря, его не сохраняет.

Задание 4.5. Пусть функция $f(x)$ задана на E_n и непрерывно дифференцируема там. Доказать эквивалентность следующих утверждений:

1. $f(x)$ — выпуклая на E_n функция.
2. $f(x_2) - f(x_1) \geq (f'(x_1), x_2 - x_1) \quad \forall x_1, x_2 \in E_n$.
3. $(f'(x_0 + \alpha g), g)$ — неубывающая при увеличении α функция при любых фиксированных x_0 и g .
4. Пусть $f(x)$ — дважды непрерывно дифференцируемая на E_n функция. Матрица вторых производных $f''(x)$ неотрицательно определена, т. е.

$$(f''(x)g, g) \geq 0 \quad \forall x, g \in E_n.$$

Здесь

$$f''(x) = \begin{pmatrix} \frac{\partial^2 f}{\partial x^{(1)} \partial x^{(1)}} & \frac{\partial^2 f}{\partial x^{(1)} \partial x^{(2)}} & \cdots & \frac{\partial^2 f}{\partial x^{(1)} \partial x^{(n)}} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^2 f}{\partial x^{(n)} \partial x^{(1)}} & \frac{\partial^2 f}{\partial x^{(n)} \partial x^{(2)}} & \cdots & \frac{\partial^2 f}{\partial x^{(n)} \partial x^{(n)}} \end{pmatrix}.$$

Проверить

Задание 4.6. Пусть $f(x)$ — непрерывно дифференцируемая на E_n функция. Доказать эквивалентность следующих утверждений:

1. $f(x)$ — строго выпуклая функция.
2. $f(x_2) - f(x_1) > (f'(x_1), x_2 - x_1) \quad \forall x_1, x_2 \in E_n, x_1 \neq x_2$.
3. $(f'(x + \alpha g), g)$ — возрастающая при увеличении α функция α .
4. Пусть $f(x)$ — дважды непрерывно дифференцируемая на E_n функция. Матрица вторых производных $f''(x)$ положительно определенная, т. е. $(f''(x)g, g) > 0 \quad \forall x, g \in E_n, g \neq 0$.

Задание 4.7. Доказать, что если $f(x)$ — сильно выпуклая дважды непрерывно дифференцируемая на E_n функция, то существует $m > 0$ такое, что

$$(f''(x)g, g) \geq m \|g\|^2 \quad \forall x, g \in E_n.$$

Задание 4.8. Доказать, что для квадратичной функции

$$f(x) = 2^{-1}(Ax, x) + (b, x),$$

где A — $(n \times n)$ -матрица, $b \in E_n$, понятия строгой выпуклости и сильной выпуклости совпадают (т. е. всякая сильно выпуклая функция является строго выпуклой, и наоборот).

За д а н и е 4.9. Показать, что если $f(x)$ — непрерывная на E_n функция и

$$f(2^{-1}(x_1 + x_2)) \leq 2^{-1}[f(x_1) + f(x_2)] - \gamma \|x_1 - x_2\|^2 \quad \forall x_1, x_2 \in E_n,$$

где $\gamma > 0$ одно и то же для всех $x_1, x_2 \in E_n$, то функция $f(x)$ сильно выпукла на E_n , и наоборот.

6. Квазивыпуклые функции. Пусть на выпуклом множестве $S \subset E_n$ определена функция $f(x)$. Будем считать, что она конечна.

Функция $f(x)$ называется *квазивыпуклой* на S , если

$$f(\alpha x + (1 - \alpha)y) \leq \max\{f(x), f(y)\} \quad (4.18)$$

$$\forall \alpha \in [0, 1], \quad \forall x, y \in S.$$

Если при $\alpha \in (0, 1)$, $x \neq y$ неравенство (4.18) строгое, то функция $f(x)$ называется *строго квазивыпуклой*.

Очевидно, что любая выпуклая функция является квазивыпуклой.

Если функции $f_i(x)$, $i \in 1:m$, квазивыпуклы на S , то функция $f(x) = \max_{i \in 1:m} f_i(x)$ также квазивыпукла на S .

Из (4.18) вытекает, что если $x_i \in S$, $\alpha_i \geq 0$, $\sum_{i=1}^m \alpha_i = 1$, то $f\left(\sum_{i=1}^m \alpha_i x_i\right) \leq \max_{i \in 1:m} f(x_i)$.

Лемма 4.5. Функция $f(x)$ квазивыпукла на S тогда и только тогда, когда при любом $c \in E_1$ множество $\mathcal{D}_c = \{x \in S | f(x) \leq c\}$ выпукло.

Доказательство. Необходимость. Зафиксируем $c \in E_1$. Если $x \in \mathcal{D}_c$ и $y \in \mathcal{D}_c$, то из (4.18)

$$f(\alpha x + (1 - \alpha)y) \leq \max\{f(x), f(y)\} \leq c \\ \forall \alpha \in [0, 1],$$

а это означает, что множество $\mathcal{D}_c$ выпукло.

Достаточность. Пусть при любом $c \in E_1$ множество $\mathcal{D}_c$ выпукло. Возьмем любые точки $x \in S$ и $y \in S$. Положим $c_0 = \max\{f(x), f(y)\}$. Тогда $x \in \mathcal{D}_{c_0}$, $y \in \mathcal{D}_{c_0}$.

и, следовательно,

$$\alpha x + (1 - \alpha)y \in \mathcal{D}_{\alpha_0} \quad \forall \alpha \in [0, 1],$$

т. е. выполнено (4.18). ■

Точку $\bar{x} \in S$ назовем точкой *действительного локального минимума* функции $f(x)$ на S , если из того, что $\bar{x} \in S$ и $f(\bar{x}) < f(\bar{x})$, вытекает, что на отрезке $[\bar{x}, \bar{x}]$ найдется такая точка y , что $f(y) > f(\bar{x})$, и такое $\delta > 0$, что $f(\bar{x}) \leq f(x) \quad \forall x \in S_\delta(\bar{x})$.

Положим $\mathcal{M}^* = \{x \in S \mid f(x) = \inf_{y \in S} f(y)\}$.

Следствие. Если функция $f(x)$ квазивыпукла на S , то множество $\mathcal{M}^*$ выпукло и любая точка действительного локального минимума функции $f(x)$ на S есть точка глобального минимума $f(x)$ на S .

Доказательство. Выпуклость множества $\mathcal{M}^*$ вытекает из леммы 4.5.

Пусть $\bar{x} \in S$ — точка действительного локального минимума функции $f(x)$ на S , а $x \in \mathcal{M}^*$. Пусть $\bar{x} \notin \mathcal{M}^*$. Поскольку $\mathcal{M}^*$ выпукло, то $f(\bar{x}) < f(\bar{x})$. Поэтому найдется такое $\alpha \in [0, 1]$, что

$$f(\alpha \bar{x} + (1 - \alpha)\bar{x}) > f(\bar{x}). \quad (4.19)$$

В силу квазивыпуклости функции $f(x)$ на множестве S

$$f(\alpha \bar{x} + (1 - \alpha)\bar{x}) \leq \max \{f(\bar{x}), f(\bar{x})\} = f(\bar{x}),$$

а это противоречит (4.19). ■

Замечание 4. Точка $x_0 \in S$ называется *точкой локального минимума* функции $f(x)$ на S , если существует такое $r > 0$, что

$$f(x) \geq f(x_0) \quad \forall x \in \{x \in S \mid \|x - x_0\| \leq r\}.$$

Отметим, что точка локального минимума $f(x)$ на S может оказаться точкой глобального максимума $f(x)$ на S , чего не может случиться с точкой действительного локального минимума $f(x)$ на S , если только $f(x)$ отлична от постоянной на S .

Лемма 4.6. Пусть функция $f(x)$ выпукла на S , а функция $g(x)$ линейна на S . Тогда функция $h(x) = f(x)(g(x))^{-1}$ квазивыпукла на множестве $\Omega = \{x \in S \mid g(x) > 0\}$.

Доказательство. Зафиксируем $x \in \Omega$, $y \in \Omega$ и $\alpha \in [0, 1]$. Очевидно, что $\alpha x + (1 - \alpha)y \in \Omega$, ибо множество Ω выпукло. Имеем

$$\begin{aligned} h(\alpha x + (1 - \alpha)y) &= \\ &= f(\alpha x + (1 - \alpha)y)/g(\alpha x + (1 - \alpha)y) \leq \\ &\leq [\alpha f(x) + (1 - \alpha)f(y)] [\alpha g(x) + (1 - \alpha)g(y)]^{-1} = \\ &= [\alpha g(x)(f(x)/g(x)) + (1 - \alpha)g(y)(f(y)/g(y))] \times \\ &\quad \times [\alpha g(x) + (1 - \alpha)g(y)] \leq \\ &\leq \max \{f(x)/g(x), f(y)/g(y)\} = \max \{h(x), h(y)\}. \quad \blacksquare \end{aligned}$$

Следствие. Множество точек минимума функции $h(x)$ на S выпукло.

Пусть квазивыпуклая на E_n функция $f(x)$ дифференцируема по направлениям. Тогда функция $\frac{\partial f(x)}{\partial g}$ квазивыпукла по $g \in E_n$. Действительно, пусть $x_0 \in E_n$, $g = \beta g_1 + (1 - \beta)g_2$, $g_1, g_2 \in E_n$, $\beta \in [0, 1]$. Имеем

$$\begin{aligned} \frac{\partial f(x_0)}{\partial g} &= \lim_{\alpha \rightarrow +0} \alpha^{-1} [f(x_0 + \alpha g) - f(x_0)] = \\ &= \lim_{\alpha \rightarrow +0} \alpha^{-1} [f(\beta(x_0 + \alpha g_1) + (1 - \beta)(x_0 + \alpha g_2)) - f(x_0)] \leq \\ &\leq \lim_{\alpha \rightarrow 0} \alpha^{-1} [\max \{f(x_0 + \alpha g_1), f(x_0 + \alpha g_2)\} - f(x_0)] = \\ &= \max \left\{ \lim_{\alpha \rightarrow +0} \alpha^{-1} [f(x_0 + \alpha g_1) - f(x_0)], \right. \\ &\quad \left. \lim_{\alpha \rightarrow 0} \alpha^{-1} [f(x_0 + \alpha g_2) - f(x_0)] \right\}. \quad \blacksquare \end{aligned}$$

Положим

$$\tilde{f}(x) = \inf \{ \mu \in E_1 \mid (\mu, x) \in \text{соепр} f \}.$$

Ясно, что функция $\tilde{f}(x)$ выпукла на E_n и $\inf_{x \in E_n} \tilde{f}(x) = \inf_{x \in E_n} f(x)$. Это означает, что в принципе задача нахождения минимума квазивыпуклой функции $f(x)$ сводится к задаче нахождения минимума выпуклой функции $\tilde{f}(x)$.

Важным примером квазивыпуклой задачи (но невыпуклой) является задача наилучшей равномерной дробно-линейной аппроксимации

$$\max_{t \in [a, b]} \left| a_0(t) - \frac{\sum_{i=1}^n x_i a_i(t)}{\sum_{i=1}^n x_i b_i(t)} \right| \rightarrow \min_{x \in E_n},$$

где $a_0(t)$, $a_i(t)$, $b_i(t)$, $i \in 1:n$, — непрерывные на $[a, b]$ функции.

§ 5. Субградиент и субдифференциал выпуклой функции

1. Пусть функция $f(x)$ выпукла и конечна на E_n , $x_0 \in E_n$. Множество

$$\partial f(x_0) = \{v \in E_n \mid f(x) - f(x_0) \geq (v, x - x_0) \forall x \in E_n\} \quad (5.1)$$

называется *субдифференциалом* функции $f(x)$ в точке x_0 . Элемент $v \in \partial f(x_0)$ называется *субградиентом* (или *обобщенным градиентом*) функции $f(x)$ в точке x_0 .

Теорема 5.1. Множество $\partial f(x_0)$ непусто, выпукло, замкнуто и ограничено.

Доказательство. Введем множество

$$\bar{Z} = \{[\beta, x] \in E_{n+1} \mid \beta \in E_1, x \in E_n, \beta \geq f(x)\}.$$

Множество $\bar{Z}$ выпуклое, замкнутое и непустое (это надграфик функции $f(x)$). Точка $[f(x_0), x_0]$ — граничная точка множества $\bar{Z}$. По следствию 2 из теоремы 1.2 существуют такие число c и вектор v , одновременно не равные нулю, т. е.

$$c^2 + v^2 > 0, \quad (5.2)$$

что

$$c\beta + (v, x) \geq cf(x_0) + (v, x_0) \quad \forall [\beta, x] \in \bar{Z}. \quad (5.3)$$

При $x = x_0$ имеем $c(\beta - f(x_0)) \geq 0 \quad \forall \beta \geq f(x_0)$. Отсюда $c \geq 0$. Но при $y = 0$ из (5.3)

$$(v, x - x_0) \geq 0 \quad \forall x \in E_n. \quad (5.4)$$

Из (5.4) следует $v = 0$, что противоречит (5.2). Итак, $c > 0$. Полагая в (5.3) $\beta = f(x)$, получим

$$f(x) - f(x_0) \geq (-c^{-1}v, x - x_0) \quad \forall x \in E_n,$$

т. е. $v_0 = -c^{-1}v \in \partial f(x_0)$. Значит, множество $\partial f(x_0)$ не пусто. Выпуклость и замкнутость $\partial f(x_0)$ очевидны. Осталось показать ограниченность.

Допустим противное. Тогда существует последовательность $\{v_k\}$ такая, что $v_k \in \partial f(x_0)$, $\|v_k\| \rightarrow \infty$. Положим $x_k = x_0 + z_k$, где $z_k = \delta_0 \|v_k\|^{-1} v_k$. Из (5.1)

$$f(x_k) - f(x_0) \geq (v_k, x_k - x_0) = (v_k, z_k) = \delta_0 \|v_k\| \rightarrow +\infty. \quad (5.5)$$

Поскольку $\|x_k - x_0\| = \|z_k\| = \delta_0$, а $f(x)$ — непрерывная функция в силу выпуклости, то $f(x)$ ограничена на множестве $S_{\delta_0}(x_0)$, что противоречит (5.5). Полученное противоречие и доказывает теорему.

Следствие. *Отображение $\partial f(x)$ ограничено на любом ограниченном множестве.*

Доказательство аналогично доказательству теоремы.

Замечание 1. Из определения (5.1) ясно, что отображение $\partial f(x)$ является полунепрерывным сверху в любой точке $x_0 \in E_n$. Однако, как будет показано ниже, оно, вообще говоря, не является полунепрерывным снизу.

Лемма 5.1. *Пусть $x_0 \in E_n$, $f(x)$ — выпуклая на E_n функция. Для любого $\varepsilon > 0$ найдется такое $\delta > 0$, что*

$$\partial f(x) \subset \partial f(x_0) + S_\varepsilon(0) \quad \forall x \in S_\delta(x_0). \quad (5.6)$$

Доказательство. Допустим противное. Тогда найдутся число $\varepsilon_0 > 0$ и последовательности $\{x_k\}$, $\{v_k\}$ такие, что

$$\begin{aligned} x_k &\rightarrow x_0, \quad v_k \in \partial f(x_k), \\ \rho(v_k, \partial f(x_0)) &= \min_{v \in \partial f(x_0)} \|v_k - v\| \geq \varepsilon_0. \end{aligned} \quad (5.7)$$

Поскольку множества $\partial f(x)$ ограничены в совокупности, то последовательность $\{v_k\}$ ограничена. Без ограничения общности можем считать, что $v_k \rightarrow v_0$. Из (5.7) тогда следует, что $v_0 \notin \partial f(x_0)$, а это противоречит полунепрерывности сверху отображения $\partial f(x)$ в точке x_0 . ■

Теорема 5.2. *Справедливо соотношение*

$$\frac{\partial f(x_0)}{\partial g} = \max_{v \in \partial f(x_0)} (v, g).$$

Доказательство. Дифференцируемость $f(x)$ по направлениям установлена в теореме 4.2. Положим

$h(\alpha) = \alpha^{-1} [f(x_0 + \alpha g) - f(x_0)]$ и найдем $\lim_{\alpha \rightarrow +0} h(\alpha)$.
В силу (5.1) для любого $v \in \partial f(x_0)$

$$f(x_0 + \alpha g) - f(x_0) \geq (v, \alpha g).$$

Отсюда

$$\frac{\partial f(x_0)}{\partial g} \equiv \lim_{\alpha \rightarrow +0} h(\alpha) \geq \max_{v \in \partial f(x_0)} (v, g). \quad (5.8)$$

Введем множества

$$Z = \{[\beta, x] \in E_{n+1} \mid x \in E_n, \beta > f(x)\},$$

$$\mathcal{L} = \left\{ [\beta, x] \mid \beta = f(x_0) + \alpha \frac{\partial f(x_0)}{\partial g}, x = x_0 + \alpha g, \alpha > 0 \right\}.$$

При доказательстве теоремы 4.2 было показано, что $h(\alpha)$ не возрастает при $\alpha \rightarrow +0$; поэтому

$$\frac{\partial f(x_0)}{\partial g} \equiv \lim_{\alpha' \rightarrow +0} h(\alpha') \leq h(\alpha) \quad \forall \alpha > 0, \quad (5.9)$$

т. е.

$$f(x_0 + \alpha g) \geq f(x_0) + \alpha \frac{\partial f(x_0)}{\partial g}.$$

Луч $\mathcal{L}$ и множество Z не имеют общих точек. Действительно, в противном случае существовала бы точка $[\beta_1, x_1] \in Z \cap \mathcal{L}$, т. е.

$$\beta_1 > f(x_1), \quad \beta_1 = f(x_0) + \alpha_1 \frac{\partial f(x_0)}{\partial g},$$

$$x_1 = x_0 + \alpha_1 g,$$

где $\alpha_1 > 0$. Отсюда

$$f(x_0) + \alpha_1 \frac{\partial f(x_0)}{\partial g} > f(x_0 + \alpha_1 g),$$

что противоречит (5.9). Образует множество

$$P = Z - \mathcal{L} = \{w = v - v' \mid v \in Z, v' \in \mathcal{L}\}.$$

Очевидно, что P — выпуклое множество и $0 \notin P$.

По следствию 2 из теоремы отделимости существует точка $[c, v] \in E_{n+1}$ такая, что

$$c^2 + v^2 > 0,$$

$$c\beta + (v, x) \geq c \left(f(x_0) + \alpha \frac{\partial f(x_0)}{\partial g} \right) + (v, x_0 + \alpha g) \quad (5.10)$$

для всех α, β, x таких, что $\beta \geq f(x)$, $\alpha \geq 0$. Как и при доказательстве теоремы 5.1, устанавливаем, что $c > 0$. При $\beta = f(x)$, $\alpha = 0$ будет

$$f(x) - f(x_0) \geq -c^{-1}(v, x - x_0) \quad \forall x \in E_n,$$

т. е. $v_0 = -c^{-1}v \in \partial f(x_0)$. Из (5.10) следует

$$f(x) - f(x_0) \geq (v_0, x - x_0) + \alpha \left(\frac{\partial f(x_0)}{\partial g} - (v_0, g) \right)$$

При $x = x_0$ и $\alpha > 0$ получим

$$\frac{\partial f(x_0)}{\partial g} \leq (v_0, g) \leq \max_{v \in \partial f(x_0)} (v, g).$$

Отсюда и из (5.8) следует утверждение теоремы. ■

Следствие. Пусть $\frac{\partial f(x_0)}{\partial g} = -a < 0$ и $v_1 \in \partial f(x_0)$ таково, что

$$(v_1, g) = \max_{v \in \partial f(x_0)} (v, g) = \frac{\partial f(x_0)}{\partial g}.$$

Тогда для $g_1 = -g$ будет

$$\frac{\partial f(x_0)}{\partial g_1} = \max_{v \in \partial f(x_0)} (v, g_1) \geq (v_1, g_1) = -(v_1, g) = -\frac{\partial f(x_0)}{\partial g} = a > 0,$$

т. е. если в каком-то направлении функция $f(x)$ убывает, то в противоположном направлении она непременно должна возрастать.

Замечание 2. В дальнейшем будет использоваться следующее представление выпуклой функции:

$$f(x) = \max_{z \in E_n} [f(z) + (v(z), x - z)], \quad (5.11)$$

где $v(z)$ — произвольный вектор из $\partial f(z)$.

Докажем (5.11). Если $v(z) \in \partial f(z)$, то

$$f(x) \geq f(z) + (v(z), x - z).$$

Это неравенство справедливо для всех $z \in E_n$. Поэтому

$$f(x) \geq \sup_{z \in E_n} [f(z) + (v(z), x - z)]. \quad (5.12)$$

Соотношение

$$\sup_{z \in E_n} [f(z) + (v(z), x - z)] \geq f(x)$$

очевидно. Отсюда и из (5.12) и следует (5.11) (вместо супремума можно написать максимум, ибо супремум достигается при $z = x$). Ясно также, что

$$f(x) = \max_{z \in S} [f(z) + (v(z), x - z)], \quad (5.13)$$

где $S \subset E_n$ — любое множество, содержащее x в качестве своей внутренней точки.

Итак, *выпуклая функция есть максимум линейных функций.*

Замечание 3. Из теоремы 5.1 следует, что если $f(x)$ — выпуклая функция, то для любого $x_0 \in E_n$ существует такое $v_0 = v(x_0) \in E_n$ (возможно, даже и не одно), что

$$f(x) - f(x_0) \geq (v(x_0), x - x_0) \quad \forall x \in E_n.$$

Докажем, что верно и обратное. Именно, *если для любого $x_0 \in E_n$ существует $v(x_0) \in E_n$ такое, что выполнено приведенное выше неравенство, то функция $f(x)$ — выпуклая.*

Возьмем произвольные $x_1, x_2 \in E_n$ и $\alpha \in [0, 1]$. Положим $x_0 = \alpha x_1 + (1 - \alpha)x_2$. По предположению найдется $v_0 = v(x_0)$ такое, что

$$f(x) - f(x_0) \geq (v_0, x - x_0) \quad \forall x \in E_n.$$

При $x = x_1$ отсюда следует

$$f(x_1) \geq f(x_0) + (1 - \alpha)(v_0, x_1 - x_2).$$

При $x = x_2$ аналогично имеем

$$f(x_2) \geq f(x_0) - \alpha(v_0, x_1 - x_2).$$

Умножая первое из последних двух неравенств на α , а второе — на $(1 - \alpha)$ и складывая, получаем

$$f(x_0) = f(\alpha x_1 + (1 - \alpha)x_2) \leq \alpha f(x_1) + (1 - \alpha)f(x_2).$$

Это и значит в силу произвольности $x_1, x_2 \in E_n$ и $\alpha \in [0, 1]$, что $f(x)$ — выпуклая функция, что и требовалось доказать. Установленное только что свойство может быть взято в качестве определения выпуклой функции.

2. Субдифференциал суммы и функции максимума. Пусть $f_1(x)$ и $f_2(x)$ — выпуклые на E_n функции. Функция

$$f(x) = f_1(x) + f_2(x) \quad (5.14)$$

тоже выпуклая.

Лемма 5.2. *Имеет место соотношение*

$$\partial f(x) = \partial f_1(x) + \partial f_2(x). \quad (5.15)$$

Доказательство. Для любого $g \in E_n$

$$\frac{\partial f(x)}{\partial g} = \max_{v \in \partial f(x)} (v, g). \quad (5.16)$$

С другой стороны, из (5.14) ясно, что

$$\begin{aligned} \frac{\partial f(x)}{\partial g} &= \lim_{\alpha \rightarrow +0} \alpha^{-1} [f(x + \alpha g) - f(x)] = \\ &= \lim_{\alpha \rightarrow +0} \alpha^{-1} [f_1(x + \alpha g) - f_1(x)] + \\ &+ \lim_{\alpha \rightarrow +0} \alpha^{-1} [f_2(x + \alpha g) - f_2(x)] = \frac{\partial f_1(x)}{\partial g} + \frac{\partial f_2(x)}{\partial g}. \end{aligned} \quad (5.17)$$

Но

$$\frac{\partial f_1(x)}{\partial g} = \max_{v \in \partial f_1(x)} (v, g), \quad \frac{\partial f_2(x)}{\partial g} = \max_{v \in \partial f_2(x)} (v, g).$$

Отсюда, из (5.16) и (5.17) имеем

$$\begin{aligned} \max_{v \in \partial f(x)} (v, g) &= \max_{v \in \partial f_1(x)} (v, g) + \max_{v \in \partial f_2(x)} (v, g) = \\ &= \max_{v \in [\partial f_1(x) + \partial f_2(x)]} (v, g). \end{aligned}$$

Поскольку это равенство имеет место для всех $g \in E_n$, а множества $\partial f(x)$ и $\partial f_1(x) + \partial f_2(x)$ замкнутые и выпуклые, то по следствию из леммы 2.1

$$\partial f(x) = \partial f_1(x) + \partial f_2(x),$$

что и требовалось доказать. ■

Задание 5.1. Доказать, что если $f(x)$ — выпуклая на E_n функция, $A > 0$, $f_1(x) = Af(x)$, то

$$\partial f_1(x) = A \partial f(x). \quad (5.18)$$

Лемма 5.3. Пусть $f(x)$ — выпуклая непрерывно дифференцируемая на E_n функция. Тогда

$$\partial f(x_0) = \{f'(x_0)\}. \quad (5.19)$$

Доказательство. Зафиксируем $x_0 \in E_n$. В лемме 4.3 показано, что в этом случае

$$f(x) - f(x_0) \geq (f'(x_0), x - x_0) \quad \forall x \in E_n.$$

Значит, $v = f'(x_0) \in \partial f(x_0)$.

Пусть $v \in \partial f(x_0)$, т. е.

$$f(x) - f(x_0) \geq (v, x - x_0) \quad \forall x \in E_n. \quad (5.20)$$

Тогда $h(x) \equiv f(x) - f(x_0) + (v, x_0 - x)$ — непрерывно дифференцируемая на E_n функция, причем из (5.20) $h(x) \geq 0 \quad \forall x \in E_n$. Но поскольку $h(x_0) = 0$, то точка x_0 — точка минимума функции $h(x)$ на E_n . По необходимому условию минимума $h'(x_0) = 0$, т. е. $f'(x_0) - v = 0$, или $v = f'(x_0)$, что и требовалось доказать. ■

Лемма 5.4. Пусть $f_i(x)$, $i \in I \equiv 1 : N$, — выпуклые на E_n функции. Пусть также $f(x) = \max_{i \in I} f_i(x)$. Функция $f(x)$ является выпуклой, причем

$$\partial f(x) = \text{co} \{ \partial f_i(x) \mid i \in R(x) \}, \quad (5.21)$$

где

$$R(x) = \{ i \in I \mid f_i(x) = f(x) \}.$$

Доказательство. Правую часть в (5.21) обозначим $D(x)$. Выпуклость функции $f(x)$ очевидна. По теореме 5.2

$$\frac{\partial f(x)}{\partial g} = \max_{v \in \partial f(x)} (v, g). \quad (5.22)$$

С другой стороны, при фиксированных x и g

$$\begin{aligned} f(x + \alpha g) &= f(x) + \alpha \frac{\partial f(x)}{\partial g} + o(\alpha) = \\ &= \max_{i \in I} \left[f_i(x) + \alpha \frac{\partial f_i(x)}{\partial g} + o_i(\alpha) \right], \end{aligned}$$

где $\alpha^{-1} o(\alpha) \xrightarrow{\alpha \rightarrow +0} 0$, $\alpha^{-1} o_i(\alpha) \xrightarrow{\alpha \rightarrow +0} 0 \quad \forall i \in I$.

При достаточно малых $\alpha > 0$ будет

$$\begin{aligned} f(x + \alpha g) &= \max_{i \in R(x)} f_i(x + \alpha g) = \\ &= \max_{i \in R(x)} \left[f_i(x) + \alpha \frac{\partial f_i(x)}{\partial g} + o_i(\alpha) \right] = \\ &= f(x) + \alpha \max_{i \in R(x)} \frac{\partial f_i(x)}{\partial g} + o(\alpha), \end{aligned}$$

где

$$\alpha^{-1}o(\alpha) \xrightarrow{\alpha \rightarrow +0} 0, \quad o(\alpha) \in [\bar{o}(\alpha), \overline{o}(\alpha)],$$

$$\bar{o}(\alpha) = \min_{i \in R(x)} o_i(\alpha), \quad \overline{o}(\alpha) = \max_{i \in R(x)} o_i(\alpha).$$

Поэтому

$$\frac{\partial f(x)}{\partial g} = \max_{i \in R(x)} \frac{\partial f_i(x)}{\partial g} =$$

$$= \max_{i \in R(x)} \max_{v_i \in \partial f_i(x)} (v_i, g) = \max_{v \in D(x)} (v, g).$$

Отсюда и из (5.22) следует

$$\max_{v \in \partial f(x)} (v, g) = \max_{v \in D(x)} (v, g).$$

Поскольку это соотношение справедливо для всех $g \in E_n$, то по следствию из леммы 2.1 получаем $\partial f(x) = D(x)$, что и требовалось доказать. ■

Пример 1. Пусть $x \in E_1$, $f(x) = |x| = \max\{f_1(x), f_2(x)\}$, где $f_1(x) = -x$, $f_2(x) = x$. Здесь $I = 1:2$. Так как функции $f_1(x)$ и $f_2(x)$ линейные (а потому и выпуклые), то и функция $f(x)$ выпуклая. Поскольку $f'_1(x) = -1$, $f'_2(x) = 1$,

$$R(x) = \begin{cases} \{1\}, & x < 0, \\ \{2\}, & x > 0, \\ \{1, 2\}, & x = 0, \end{cases}$$

то по формуле (5.21) и лемме 5.3

$$\partial f(x) = \text{co}\{f'_i(x) | i \in R(x)\},$$

т. е.

$$\partial f(x) = \begin{cases} -1, & x < 0, \\ +1, & x > 0, \\ [-1, 1], & x = 0. \end{cases}$$

Для этой же функции теперь покажем, что отображение $\partial f(x)$ не является полунепрерывным снизу. Выберем точки $x_0 = 0$, $v_0 = 0,5$ и последовательность $\{x_i\}$ такую, что $x_i \rightarrow -0$. Функции $f_1(x)$ и $f_2(x)$ дифференцируемы. Ясно, что $v_0 \in \partial f(x_0)$, так как

$$\partial f(0) = \text{co}\{f'_1(0), f'_2(0)\} = \text{co}\{-1, 1\} = [-1, 1].$$

Однако для выбранной последовательности $\{x_i\}$ имеем

$$\partial f(x_i) = \{f'_i(x_i)\} = \{-1\} \quad \forall i.$$

Поэтому не существует последовательности $\{v_i\}$, $v_i \in \partial f(x_i)$, которая сходилась бы к $v_0 = 0,5$. Это и означает, что отображение $\partial f(x)$ не является полунепрерывным снизу в точке $x_0 = 0$, а тем самым оно не является непрерывным по Какутани.

Пример 2. Пусть

$$x = (x^{(1)}, \dots, x^{(n)}) \in E_n, \\ f(x) = \max_{i \in 1:n} |x^{(i)}| = \max_i f_i(x),$$

где $f_i(x) = |x^{(i)}|$. Из примера 1 следует, что

$$\partial f_i(x) = \begin{cases} -e_i, & \text{если } x^{(i)} < 0, \\ e_i, & \text{если } x^{(i)} > 0, \\ \text{co}\{e_i, -e_i\}, & \text{если } x^{(i)} = 0, \end{cases}$$

где e_i — единичные орты, $e_i = (0, \dots, 0, 1, 0, \dots, 0)$.

По формуле (5.21)

$$\partial f(x) = \text{co}\{w_i \mid i \in R(x)\}, \\ R(x) = \{i \in 1:n \mid |x^{(i)}| = f(x)\}, \\ w_i = \begin{cases} e_i, & \text{если } x^{(i)} \geq 0, \\ -e_i, & \text{если } x^{(i)} \leq 0. \end{cases}$$

При этом $\partial f(0) = \text{co}\{\pm e_i \mid i \in 1:n\}$. Отметим, что $f(x) = (w, x)$, где $w \in \partial f(x)$.

Задание 5.2. Доказать формулу, аналогичную (5.21), для случая, когда

$$f(x) = \max_{y \in G} \varphi(x, y), \quad (5.23)$$

где $G \subset E_m$ — компакт, функция $\varphi(x, y)$ непрерывна по y и выпукла по x при любом фиксированном $y \in G$.

3. Лемма 5.5. *Выпуклая функция является липшицевой на любом выпуклом ограниченном множестве, т. е. для любого ограниченного $G \subset E_n$ найдется $L < \infty$ такое, что*

$$|f(x) - f(z)| \leq L \|x - z\| \quad \forall x \in G, \quad \forall z \in G.$$

Доказательство. По определению $\partial f(z)$ имеем

$$f(x) - f(z) \geq (v, x - z), \quad \text{где } v \in \partial f(z).$$

В силу следствия из теоремы 5.1 найдется такое $L < \infty$, что

$$\|v\| \leq L \quad \forall v \in \partial f(z), \quad \forall z \in G.$$

Отсюда

$$f(x) - f(z) \geq -L\|x - z\|.$$

Поменяв ролями z и x , имеем

$$f(z) - f(x) \geq -L\|z - x\|.$$

Из последних двух неравенств вытекает

$$|f(x) - f(z)| \leq L\|x - z\| \quad \forall x \in G, \quad \forall z \in G,$$

а это и означает, что функция $f(x)$ является липшицевой на G . Лемма доказана.

Зафиксируем $x_0 \in E_n$ и $g \in E_n$. Пусть $\mathbb{O}(\alpha)$ — произвольная вектор-функция такая, что

$$\mathbb{O}(\alpha) \in E_n, \quad \alpha^{-1}\|\mathbb{O}(\alpha)\| \xrightarrow{\alpha \rightarrow +0} 0.$$

Лемма 5.6. Если функция $f(x)$ липшицева в окрестности точки x_0 и дифференцируема в точке x_0 по направлениям, то

$$\begin{aligned} \lim_{\alpha \rightarrow +0} \alpha^{-1}[f(x_0 + \alpha g + \mathbb{O}(\alpha)) - f(x_0)] &= \\ &= \frac{\partial f(x_0)}{\partial g} \equiv \lim_{\alpha \rightarrow +0} \alpha^{-1}[f(x_0 + \alpha g) - f(x_0)]. \end{aligned}$$

Доказательство. Поскольку $f(x)$ — липшицева функция, то

$$|f(x_0 + \alpha g + \mathbb{O}(\alpha)) - f(x_0 + \alpha g)| \leq L\|\mathbb{O}(\alpha)\| = L|o(\alpha)|,$$

$$\text{т. е. } f(x_0 + \alpha g + \mathbb{O}(\alpha)) - f(x_0 + \alpha g) = o(\alpha),$$

$$\text{где } o(\alpha) = \|\mathbb{O}(\alpha)\|, \quad \alpha^{-1}o(\alpha) \xrightarrow{\alpha \rightarrow +0} 0.$$

Отсюда

$$\begin{aligned} \lim_{\alpha \rightarrow +0} \alpha^{-1}[f(x_0 + \alpha g + \mathbb{O}(\alpha)) - f(x_0)] &= \\ &= \lim_{\alpha \rightarrow +0} [\alpha^{-1}[f(x_0 + \alpha g) - f(x_0)] + \alpha^{-1}o(\alpha)] = \\ &= \lim_{\alpha \rightarrow +0} \alpha^{-1}[f(x_0 + \alpha g) - f(x_0)] + \lim_{\alpha \rightarrow +0} \alpha^{-1}o(\alpha) = \frac{\partial f(x_0)}{\partial g}, \end{aligned}$$

что и требуется. ■

Отметим, что лемма 5.6 справедлива и для любой выпуклой функции, ибо выпуклая функция дифференцируема по направлениям и липшицева в окрестности любой точки x_0 (см. теорему 4.2 и лемму 5.5).

Определение. Функция одной переменной $\psi(t)$, заданная на $[0, T]$, $T < \infty$, называется *абсолютно непрерывной* там, если для любого $\varepsilon > 0$ найдется $\delta > 0$ такое, что

$$\sum_{k=1}^N |\psi(t_k'') - \psi(t_k')| < \varepsilon,$$

как только

$$\sum_{k=1}^N (t_k'' - t_k') < \delta,$$

где $0 \leq t_1' < t_1'' \leq t_2' < t_2'' \leq \dots \leq t_k' < t_k'' \leq \dots \leq t_N' < t_N'' \leq T$, N — любое натуральное число.

Пусть $f(x)$ — выпуклая на E_n функция, $x(t)$ — непрерывная n -мерная вектор-функция, заданная на $[0, T]$ и удовлетворяющая там условию Липшица, т. е. существует $K_1 > 0$ такое, что

$$\|x(t') - x(t'')\| \leq K_1 |t' - t''| \quad \forall t', t'' \in [0, T].$$

Положим $\psi(t) = f(x(t))$. В силу леммы 5.5, выпуклая функция $f(x)$ является липшицевой по x на любом ограниченном множестве $G \subset E_n$. Функция $\psi(t)$ является липшицевой на $[0, T]$, ибо

$$\begin{aligned} |\psi(t') - \psi(t)| &= |f(x(t')) - f(x(t))| \leq \\ &\leq K \|x(t') - x(t)\| \leq KK_1 |t' - t| \quad \forall t', t \in [0, T], \end{aligned}$$

поскольку $f(x)$ и $x(t)$ является липшицевым по x и t соответственно.

Нетрудно отсюда видеть, что $\psi(t)$ — абсолютно непрерывная на $[0, T]$ функция.

Как известно (см., например, [115]), тогда существует суммируемая на $[0, T]$ функция $\omega(t)$ такая, что

$$\psi(t) = \psi(0) + \int_0^t \omega(\tau) d\tau. \quad (5.24)$$

Интеграл понимается в смысле Лебега. При этом почти для всех $t \in [0, T]$

$$\omega(t) = \lim_{\Delta \rightarrow 0} \Delta^{-1} (\psi(t + \Delta) - \psi(t)) \equiv \psi'(t), \quad (5.25)$$

т. е. функция $\psi(t)$ является почти везде дифференцируемой. Если вдобавок $x(t)$ — дифференцируемая вектор-функция, т. е. $x(t + \Delta) = x(t) + \Delta \dot{x}(t) + \mathcal{O}(\Delta)$, где $\Delta^{-1} \|\mathcal{O}(\Delta)\| \xrightarrow{\Delta \rightarrow 0} 0$, то из (5.24), (5.25) и леммы 5.6 следует

$$f(x(t)) = f(x(0)) + \int_0^t \frac{\partial f(x(\tau))}{\partial \dot{x}(\tau)} d\tau. \quad (5.26)$$

Лемма 5.7. Пусть $S \subset E_n$ — выпуклое открытое множество, и пусть $f_1(x)$ и $f_2(x)$ — выпуклые на S функции. Если

$$\partial f_1(x) = \partial f_2(x) \quad \forall x \in S,$$

то $f_1(x) - f_2(x) = C$, где C — постоянная величина.

Доказательство. Возьмем любое $x_0 \in S$. Из (5.26) для произвольного $x \in S$ будем иметь

$$f_1(x) = f_1(x_0) + \int_0^1 \frac{\partial f_1(x_0 + \tau g)}{\partial g} d\tau,$$

$$f_2(x) = f_2(x_0) + \int_0^1 \frac{\partial f_2(x_0 + \tau g)}{\partial g} d\tau,$$

где $g = x - x_0$. Но

$$\frac{\partial f_1(x_0 + \tau g)}{\partial g} = \max_{v \in \partial f_1(x_0 + \tau g)} (v, g),$$

$$\frac{\partial f_2(x_0 + \tau g)}{\partial g} = \max_{v \in \partial f_2(x_0 + \tau g)} (v, g).$$

Поскольку $\partial f_1(x) = \partial f_2(x) \quad \forall x \in S$, то

$$f_1(x) - f_2(x) = f_1(x_0) - f_2(x_0) \equiv C,$$

что и требовалось доказать. ■

Таким образом, выпуклая функция определяется своим субдифференциальным отображением с точностью до постоянного слагаемого.

Лемма 5.8. Пусть $f(x)$ — выпуклая на E_n функция. Пусть также $x_1 \neq x_2$. Для того чтобы

$$\partial f(x_1) \cap \partial f(x_2) \neq \emptyset, \quad (5.27)$$

необходимо, чтобы для всех $\alpha \in (0, 1)$, и достаточно, чтобы хотя бы для одного $\alpha \in (0, 1)$ было выполнено равенство

$$f(\alpha x_1 + (1 - \alpha)x_2) = \alpha f(x_1) + (1 - \alpha)f(x_2). \quad (5.28)$$

Доказательство. Необходимость. Пусть $v \in \partial f(x_1)$, $v \in \partial f(x_2)$. По определению субдифференциала

$$f(z) - f(x_1) \geq (v, z - x_1) \quad \forall z \in E_n, \quad (5.29)$$

$$f(z) - f(x_2) \geq (v, z - x_2) \quad \forall z \in E_n. \quad (5.30)$$

Подставляя в (5.29) $z = x_2$, а в (5.30) $z = x_1$, имеем

$$f(x_2) - f(x_1) = (v, x_2 - x_1). \quad (5.31)$$

Возьмем любое $\alpha \in (0, 1)$ и положим $x_\alpha = \alpha x_1 + (1 - \alpha)x_2$. Из (5.29) и (5.31)

$$f(x_\alpha) - f(x_1) \geq (1 - \alpha)(v, x_2 - x_1) = (1 - \alpha)(f(x_2) - f(x_1)),$$

т. е.

$$f(x_\alpha) \geq \alpha f(x_1) + (1 - \alpha)f(x_2). \quad (5.32)$$

Но функция $f(x)$ выпуклая, поэтому

$$f(x_\alpha) \leq \alpha f(x_1) + (1 - \alpha)f(x_2) \quad \forall \alpha \in [0, 1].$$

Отсюда и из (5.32) следует (5.28).

Достаточность. Пусть для некоторого $\alpha \in (0, 1)$ имеет место (5.28). Возьмем любое $v \in \partial f(x_\alpha)$. Тогда

$$f(z) \geq f(x_\alpha) + (v, z - x_\alpha) \quad \forall z \in E_n.$$

Отсюда и из (5.28)

$$\begin{aligned} f(z) &\geq \alpha f(x_1) + (1 - \alpha)f(x_2) + (v, z - \alpha x_1 - (1 - \alpha)x_2) = \\ &= \alpha [f(x_1) + (z - x_1, v)] + \\ &+ (1 - \alpha)[f(x_2) + (z - x_2, v)] \quad \forall z \in E_n. \end{aligned} \quad (5.33)$$

При $z = x_1$

$$f(x_1) \geq \alpha f(x_1) + (1 - \alpha)[f(x_2) + (x_1 - x_2, v)],$$

т. е.

$$(1 - \alpha)[f(x_1) - f(x_2)] \geq (1 - \alpha)(x_1 - x_2, v).$$

Так как $1 - \alpha > 0$, то

$$f(x_1) - f(x_2) \geq (x_1 - x_2, v).$$

Аналогично при $z = x_2$ из (5.33) получаем

$$f(x_2) - f(x_1) \geq (x_2 - x_1, v).$$

Отсюда и из предыдущего неравенства

$$f(x_1) - f(x_2) = (x_1 - x_2, v). \quad (5.34)$$

Из (5.33)

$$f(z) - f(x_1) \geq (z - x_1, v) + \\ + (1 - \alpha)[f(x_2) - f(x_1) + (x_1 - x_2, v)] \quad \forall z \in E_n. \quad (5.35)$$

В силу (5.34) выражение в квадратных скобках в (5.35) равно нулю. Поэтому

$$f(z) - f(x_1) \geq (z - x_1, v) \quad \forall z \in E_n,$$

т. е.

$$v \in \partial f(x_1). \quad (5.36)$$

Аналогично из (5.33)

$$f(z) - f(x_2) \geq (z - x_2, v) + \\ + \alpha[f(x_1) - f(x_2) - (x_1 - x_2, v)] \quad \forall z \in E_n,$$

и снова из (5.34) получаем

$$f(z) - f(x_2) \geq (z - x_2, v) \quad \forall z \in E_n,$$

т. е. $v \in \partial f(x_2)$. Отсюда и из (5.36)

$$\partial f(x_\alpha) \subset [\partial f(x_1) \cap \partial f(x_2)].$$

Тем самым (5.27) установлено. Лемма доказана.

4. Зафиксируем $x_0 \in E_n$. Тогда для любого g , $\|g\| = 1$, имеем

$$f(x_0 + \alpha g) = f(x_0) + \alpha \frac{\partial f(x_0)}{\partial g} + o(\alpha, g), \quad (5.37)$$

где

$$\alpha^{-1} o(\alpha, g) \xrightarrow{\alpha \rightarrow +0} 0. \quad (5.38)$$

Покажем, что стремление к нулю в (5.38) равномерно по g , $\|g\| = 1$. Действительно, из (5.26)

$$f(x_0 + \alpha g) = f(x_0) + \int_0^\alpha \frac{\partial f(x_0 + \tau g)}{\partial g} d\tau. \quad (5.39)$$

По теореме 5.2

$$\frac{\partial f(x)}{\partial g} = \max_{v \in \partial f(x)} (v, g).$$

Отсюда и из (5.39)

$$f(x_0 + \alpha g) - f(x_0) \leq \alpha \max_{v \in B(\alpha)} (v, g), \quad (5.40)$$

где $B(\alpha) = \bigcup_{\|x - x_0\| \leq \alpha} \partial f(x)$. Поскольку отображение $\partial f(x)$ полунепрерывно сверху (см. замечание 1), то

$$\begin{aligned} \max_{v \in B(\alpha)} (v, g) &\leq \max_{v \in \partial f(x_0)} (v, g) + q(\alpha) = \\ &= \frac{\partial f(x_0)}{\partial g} + q(\alpha), \end{aligned} \quad (5.41)$$

где $q(\alpha) \xrightarrow{\alpha \rightarrow +0} 0$, причем $q(\alpha)$ не зависит уже от g , $\|g\| = 1$. В силу (5.1)

$$f(x_0 + \alpha g) - f(x_0) \geq \alpha \frac{\partial f(x_0)}{\partial g}.$$

Отсюда, из (5.40) и (5.41) получаем

$$0 \leq f(x_0 + \alpha g) - f(x_0) - \alpha \frac{\partial f(x_0)}{\partial g} \leq \alpha q(\alpha).$$

Так как $q(\alpha)$ не зависит от g , $\|g\| = 1$, то из (5.37) следует равномерность соотношения (5.38) по g , $\|g\| = 1$.

Отметим, что функция $o(\alpha, g) \equiv o(\alpha, g, x_0)$ не является, вообще говоря, даже непрерывной по x_0 .

5. Пусть $f(x)$ — вогнутая конечная на E_n функция, т. е.

$$\begin{aligned} f(\alpha x_1 + (1 - \alpha) x_2) &\geq \alpha f(x_1) + (1 - \alpha) f(x_2) \\ \forall x_1, x_2 \in E_n, \quad \forall \alpha \in [0, 1]. \end{aligned} \quad (5.42)$$

Множество

$$\begin{aligned} \overline{\partial f(x_0)} &= \\ &= \{v \in E_n \mid f(x) - f(x_0) \leq (v, x - x_0) \quad \forall x \in E_n\} \end{aligned} \quad (5.43)$$

называется супердифференциалом функции $f(x)$ в точке x_0 , а элемент $v \in \overline{\partial f(x_0)}$ называется суперградиентом $f(x)$ в точке x_0 . Поскольку функция $f_1(x) = -f(x)$ является выпуклой, то ясно, что

$$\overline{\partial f(x_0)} = -\partial f_1(x_0).$$

Поэтому для супердифференциалов и суперградиентов справедливы утверждения, аналогичные доказанным для выпуклых функций:

Множество $\partial f(x_0)$ непусто, выпукло, замкнуто и ограничено.

Вогнутая функция $f(x)$ дифференцируема по любому направлению, причем

$$\frac{\partial f(x_0)}{\partial g} = \min_{v \in \partial f(x_0)} (v, g). \quad (5.44)$$

Задание 5.3. Доказать формулу, аналогичную (5.15), для суммы вогнутых функций.

Задание 5.4. Найти супердифференциал функции

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

где $I = 1 : N$; $f_i(x)$ — вогнутые на E_n функции.

Задание 5.5. Доказать, что для вогнутой функции $f(x)$ и непрерывно дифференцируемой вектор-функции $x(t)$ справедливо представление

$$f(x(t)) = f(x(0)) + \int_0^t \frac{\partial f(x(\tau))}{\partial x(\tau)} d\tau. \quad (5.45)$$

Интеграл в этом равенстве понимается в смысле Лебега.

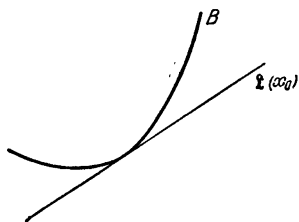


Рис. 7.

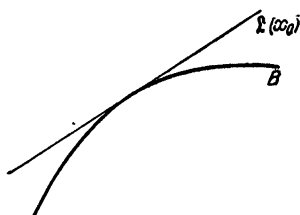


Рис. 8.

Замечание 4. Смысл названия субградиент (и субдифференциал) станет ясным, если обратить внимание на то, что (см. определение (5.1)) плоскость

$$\mathcal{L}(x_0) = \{[\beta, x] \in E_{n+1} \mid x \in E_n, \beta = f(x_0) + (v, x - x_0)\},$$

где $v \in \partial f(x_0)$, проходит через точку $[f(x_0), x_0]$ и лежит «под» графиком функции $f(x)$ (рис. 7):

$$B = \{[\beta, x] \mid x \in E_n, \beta = f(x)\}.$$

Аналогично для суперградиента вогнутой функции плоскость $\mathcal{P}(x_0)$, где $v \in \partial f(x_0)$, лежит «над» графиком функции $f(x)$ (рис. 8) и тоже проходит через точку $[f(x_0), x_0]$.

З а м е ч а н и е 5. Пусть $f(x)$ — выпуклая на E_1 функция. Имеет место следующая геометрическая интерпретация субдифференциала в точке x_0 . Через точку $(f(x_0), x_0) \in E_2$ проведем опорную прямую к надграфику функции $f(x)$. Пусть α — тангенс угла наклона этой прямой к оси Ox . Тогда $\alpha \in \partial f(x_0)$. Обратно, если $\alpha \in \partial f(x_0)$, то существует прямая, проходящая через точку $(f(x_0), x_0)$, являющаяся опорной к надграфику функции $f(x)$ и такая, что тангенс угла ее наклона к оси Ox равен α .

§ 6. Расстояние от множества до конуса.

Условия минимума

1. Пусть на выпуклом множестве $S \subset E_n$ задана выпуклая конечная функция $f(x)$. В § 4 было установлено, что функция $f(x)$ дифференцируема по любому направлению в каждой внутренней точке множества S . Пусть $\Omega \subset \text{int } S$ — выпуклое замкнутое множество.

Л е м м а 6.1. Для того чтобы в точке $x^* \in \Omega$ функция $f(x)$ достигала наименьшего на Ω значения, необходимо и достаточно, чтобы

$$\min_{g \in \Gamma(x^*), \|g\|=1} \frac{\partial f(x^*)}{\partial g} \geq 0, \quad (6.1)$$

где $\Gamma(x^*)$ — конус возможных направлений множества Ω в точке x^* .

Доказательство. Необходимость. Пусть $x^* \in \Omega$ — точка минимума $f(x)$ на Ω . Допустим, что (6.1) не выполняется. Тогда, поскольку $\Gamma(x^*) = \gamma(x^*)$, существует вектор

$$g_0 \in \gamma(x^*) = \{v = \lambda(z - x^*) \mid \lambda > 0, z \in \Omega\}, \\ \|g_0\| = 1,$$

такой, что

$$\frac{\partial f(x^*)}{\partial g_0} = -a < 0, \quad (6.2)$$

Так как

$$f(x^* + \alpha g_0) = f(x^*) + \alpha \frac{\partial f(x^*)}{\partial g_0} + o(\alpha),$$

$$o(\alpha) \cdot \alpha^{-1} \xrightarrow{\alpha \rightarrow +0} 0 \quad (6.3)$$

и при достаточно малых $\alpha > 0$ будет $x^* + \alpha g_0 \in \Omega$, то из (6.2) и (6.3) вытекает существование такого $\alpha_0 > 0$, что

$$f(x^* + \alpha g_0) < f(x^*), \quad x^* + \alpha g_0 \in \Omega \quad \forall \alpha \in (0, \alpha_0].$$

Это противоречит тому, что x^* — точка минимума $f(x)$ на Ω .

Достаточность. Пусть выполнено (6.1). В силу определения $\partial f(x^*)$ имеем для $x \neq x^*$

$$f(x) \geq f(x^*) + \max_{v \in \partial f(x^*)} (v, x - x^*) = f(x^*) + \|x - x^*\| \frac{\partial f(x^*)}{\partial g(x)},$$

где $g(x) = \|x - x^*\|^{-1}(x - x^*)$. Отсюда и из (6.1)

$$f(x) \geq f(x^*) + \|x - x^*\| \min_{g \in \Gamma(x^*), \|g\|=1} \frac{\partial f(x^*)}{\partial g} \geq f(x^*)$$

$$\forall x \in \Omega,$$

что и требовалось доказать. ■

Следствие. Условие (6.1) эквивалентно условию

$$\min_{g \in \Gamma(x^*), \|g\| \leq 1} \frac{\partial f(x^*)}{\partial g} = 0.$$

Замечание 1. Необходимым условием минимума для любой дифференцируемой по направлениям функции является условие

$$\inf_{g \in \gamma(x^*), \|g\|=1} \frac{\partial f(x^*)}{\partial g} \geq 0,$$

где $\gamma(x^*) = \{g = \lambda(x - x^*) \mid \lambda > 0, x \in \Omega\}$.

Задача 6.1. Доказать это утверждение.

Замечание 2. Отметим, что в (6.1) фигурирует минимум, ибо (см. п. 4 § 4) функция $\max_{v \in \partial f(x^*)} (v, g)$ выпукла по g . Очевидно, она конечна, и, следовательно, непрерывна по g на E_n , и достигает минимума на любом компакте из E_n , в частности на множестве $\{g \in \Gamma(x^*) \mid \|g\|=1\}$.

2. Приведем несколько лемм геометрического характера, которые будут часто использоваться в дальнейшем.

Пусть $B \subset E_n$ — выпуклый компакт, $\Gamma \subset E_n$ — выпуклый замкнутый конус, Γ^+ — конус, сопряженный конусу Γ . Положим

$$\begin{aligned} C &= B - \Gamma^+, \quad C_{fr} = C \setminus \text{int } C, \\ \psi &= \min_{\|g\|=1, g \in \Gamma} \max_{v \in B} (v, g), \quad \varphi = \inf_{\|g\|=1} \sup_{v \in C} (v, g), \\ \rho &= \inf_{v \in C_{fr}} \|v\|, \quad d = \inf_{v \in C} \|v\|. \end{aligned}$$

Поскольку множества B и Γ^+ замкнутые, а B ограничено, то инфимум в определении ρ и d достигается, т. е.

$$\rho = \min_{v \in C_{fr}} \|v\|, \quad d = \min_{v \in C} \|v\|.$$

Если $d > 0$, то $\rho = d$, т. е. ρ — расстояние от начала координат до множества C (или, что то же самое, расстояние от множества B до конуса Γ^+), а если $d = 0$, то ρ есть радиус максимального шара с центром в $\mathbb{O}$, содержащегося в C .

Лемма 6.2. *Справедливо равенство*

$$\psi = \varphi. \quad (6.4)$$

Доказательство. Покажем, что при любом $g \in E_n$

$$\sup_{v \in C} (v, g) = \begin{cases} \max_{v \in B} (v, g), & \text{если } g \in \Gamma, \\ +\infty, & \text{если } g \notin \Gamma. \end{cases} \quad (6.5)$$

Действительно,

$$\begin{aligned} \sup_{v \in C} (v, g) &= \sup_{v_1 \in B, v_2 \in [-\Gamma^+]} (v_1 + v_2, g) = \\ &= \max_{v_1 \in B} (v_1, g) + \sup_{v_2 \in [-\Gamma^+]} (v_2, g). \end{aligned}$$

Если $g \in \Gamma = (\Gamma^+)^+$, то $(-v_2, g) \geq 0 \quad \forall v_2 \in [-\Gamma^+]$, а поскольку $\mathbb{O} \in \Gamma^+$, то $\sup_{v \in [-\Gamma^+]} (v, g) = 0$. Если же $g \notin \Gamma$, то найдется такой $v'_2 \in [-\Gamma^+]$, что $(-v'_2, g) < 0$, а так

как Γ^+ — конус, то $\sup_{v \in [-\Gamma^+]} (v, g) = +\infty$. Итак,

$$\sup_{v \in [-\Gamma^+]} (v, g) = \begin{cases} 0, & \text{если } g \in \Gamma, \\ +\infty, & \text{если } g \notin \Gamma. \end{cases} \quad (6.6)$$

Отсюда вытекают (6.5) и (6.4), ибо (6.5) справедливо для любого $g \in E_n$. Лемма доказана.

Из доказательства следует, что инфимум в определении φ достигается, т. е. $\varphi = \min_{\|g\|=1} \sup_{v \in C} (v, g)$.

Лемма 6.3. Условие $\psi \geq 0$ эквивалентно условию $d = 0$.

Доказательство. Пусть $\psi \geq 0$. Тогда в силу (6.4) $\varphi \geq 0$. Допустим, что $d > 0$. Пусть $v_0 \in C$ и $\|v_0\| = d$. Как было установлено при доказательстве теоремы 1.2 (см. неравенство (1.10)),

$$(v, v_0) \geq \|v_0\|^2. \quad (6.7)$$

Положим $g_0 = -\|v_0\|^{-1}v_0$. Из (6.7) имеем

$$\begin{aligned} \varphi &= \min_{\|g\|=1} \sup_{v \in C} (v, g) \leq \\ &\leq \sup_{v \in C} (v, g_0) = -\|v_0\| = -d < 0. \end{aligned} \quad (6.8)$$

Получили противоречие, т. е. из условия $\psi \geq 0$ вытекает $d = 0$.

Предположим теперь, что $d = 0$. Это означает, что $0 \in C$. Тогда

$$\psi = \varphi = \min_{\|g\|=1} \sup_{v \in C} (v, g) \geq \min_{\|g\|=1} (0, g) = 0,$$

т. е. из условия $d = 0$ следует $\psi \geq 0$. Лемма доказана полностью.

Следствие. Условие $\psi < 0$ эквивалентно тому, что $d > 0$.

Лемма 6.4. Справедливо соотношение

$$\psi = \begin{cases} -\rho = -d, & \text{если } d > 0, \\ \rho, & \text{если } d = 0. \end{cases} \quad (6.9)$$

Доказательство. Рассмотрим сначала случай, когда $d > 0$. Учитывая (6.8) и то, что в этом случае

$\rho = d$, имеем $\psi \leq -\rho$. Установим справедливость неравенства противоположного знака. Пусть $v_0 \in C$ и $\|v_0\| = d = \rho$. Положим $g_0 = -\|v_0\|^{-1}v_0$. Покажем, что для любого $g \in E_n$, $\|g\| = 1$, $g \neq g_0$, будет

$$\sup_{v \in C} (v, g) > -\|v_0\| = -d. \quad (6.10)$$

Действительно, пусть $g \in E_n$, $\|g\| = 1$, $g \neq g_0$. Тогда

$$\|v_0\| > (-v_0, g) \geq -\sup_{v \in C} (v, g).$$

Из (6.4) и (6.10) заключаем, что

$$\psi = \min_{\|g\|=1} \sup_{v \in C} (v, g) \geq -\rho.$$

Отметим, что в силу (6.5) $g_0 \in \Gamma$. Итак, если $d > 0$, то $\psi = -\rho = -d$. Пусть теперь $d = 0$. Тогда по лемме 6.3 $\psi \geq 0$. В этом случае, как уже отмечалось, ρ есть радиус максимального шара с центром в начале координат, содержащегося в C , т. е. $S_\rho(0) \subset C$. Для любого $\delta > 0$ такого, что $S_\delta(0) \subset C$, будет

$$\delta = \max_{v \in S_\delta(0)} (v, g) \leq \sup_{v \in C} (v, g) \quad \forall g \in E_n, \quad \|g\| = 1.$$

Отсюда $\delta \leq \varphi$. В частности, $\rho \leq \varphi$. Из (6.4)

$$\rho \leq \varphi = \psi. \quad (6.11)$$

Установим теперь, что

$$S_\psi(0) \subset C. \quad (6.12)$$

Допустим противное. Тогда найдется вектор $v_0 \in E_n$ такой, что $v_0 \in S_\psi(0)$, но $v_0 \notin C$. Очевидно,

$$\|v_0\| \leq \psi. \quad (6.13)$$

По теореме отделимости существуют вектор $g_0 \in E_n$, $\|g_0\| = 1$, и число $a > 0$ такие, что $(v, g_0) \leq (v_0, g_0) - a$ $\forall v \in C$. Поэтому

$$\psi = \varphi = \min_{\|g\|=1} \sup_{v \in C} (v, g) \leq \sup_{v \in C} (v, g_0) \leq (v_0, g_0) - a.$$

Отсюда и из (6.13) $\psi \leq \|v_0\| - a \leq \psi - a < \psi$. Получили противоречие. Итак, включение (6.12), а вместе с ним и неравенство

$$\psi \leq \rho \quad (6.14)$$

установлены. Объединяя (6.11) и (6.14), приходим к выводу, что если $d = 0$, то $\psi = \rho$. Лемма доказана полностью.

Следствие 1. Пусть $\psi < 0$. Тогда, если вектор $g_0 \in \Gamma$, $\|g_0\| = 1$, такой, что

$$\max_{v \in B} (v, g_0) = \psi,$$

то

$$\begin{aligned} (v, g_0) &\leq -\rho \quad \forall v \in B, \\ (\omega, g_0) &\geq 0 \quad \forall \omega \in \Gamma^+. \end{aligned} \quad (6.15)$$

Замечание 3. Вектор $v_0 \in C$, для которого $\|v_0\| = d$, единствен. Это вытекает из того, что сильно выпуклая функция (v, v) достигает минимума на любом выпуклом замкнутом множестве в единственной точке.

Следствие 2. Если C — выпуклое замкнутое множество (не обязательно ограниченное),

$$\psi = \min_{\|g\|=1} \sup_{v \in C} (v, g), \quad d = \min_{v \in C} \|v\|, \quad \rho = \min_{v \in C_{fr}} \|v\|,$$

то условие $\psi \geq 0$ эквивалентно условию $d = 0$,

$$\psi = \begin{cases} -\rho = -d, & \text{если } d > 0, \\ \rho, & \text{если } d = 0. \end{cases}$$

Кроме того, вектор $v_0 \in C$, для которого $\|v_0\| = d$, единствен.

Следствие 3. Пусть множество B состоит из единственной точки b . Тогда, если $b \notin \Gamma^+$, то

$$\min_{v \in \Gamma^+} \|b - v\| = \max_{g \in [-\Gamma], \|g\|=1} (b, g) = \max_{g \in [-\Gamma], \|g\| \leq 1} (b, g).$$

Это очевидным образом следует из (6.9). Ясно также, что

$$\min_{v \in \Gamma^+} \|b - v\| = \max_{g \in [-\Gamma], \|g\| \leq 1} (b, g) \quad \forall b \in E_n.$$

3. Геометрическая интерпретация условий минимума.

Теорема 6.1. Условие (6.1) эквивалентно условию

$$\partial f(x^*) \cap \Gamma^+(x^*) \neq \emptyset. \quad (6.16)$$

Доказательство. Заметим, что (6.16) эквивалентно тому, что

$$0 \in [\partial f(x^*) - \Gamma^+(x^*)]. \quad (6.17)$$

Пусть выполнено (6.1), т. е.

$$\min_{g \in \Gamma(x^*), \|g\|=1} \max_{v \in \partial f(x^*)} (v, g) \geq 0. \quad (6.18)$$

Положим $B = \partial f(x^*)$, $\Gamma = \Gamma(x^*)$. Из (6.18) тогда $\psi \geq 0$. По лемме 6.3 заключаем, что $d = \min_{v \in B - \Gamma^+} \|v\| = 0$, т. е. имеет место (6.17) и тем самым (6.16).

Пусть теперь выполнены (6.16) и (6.17), тогда $d = 0$ и по лемме 6.3 $\psi \geq 0$, что и требовалось доказать. ■

Следствие. Если $\Omega = E_n$, то условие (6.16) эквивалентно условию

$$0 \in \partial f(x^*). \quad (6.19)$$

Если же вдобавок $f(x)$ — непрерывно дифференцируемая функция, то (6.16) превращается в классическое условие

$$f'(x^*) = 0. \quad (6.20)$$

4. Направление наискорейшего спуска. Пусть $x_0 \in \Omega$. Положим

$$d(x_0) = \min_{\substack{v \in \partial f(x_0) \\ w \in \Gamma^+(x_0)}} \|v - w\|. \quad (6.21)$$

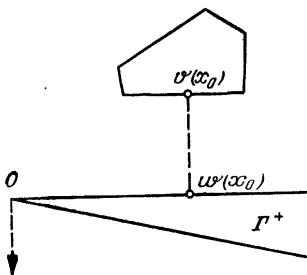


Рис. 9.

Тогда существуют такие точки $v(x_0) \in \partial f(x_0)$ и $w(x_0) \in \Gamma^+(x_0)$, что

$$d(x_0) = \|v(x_0) - w(x_0)\|. \quad (6.22)$$

Вектор $v(x_0) - w(x_0)$, удовлетворяющий (6.22), единствен (см. замечание 3), хотя точки $v(x_0)$ и $w(x_0)$ могут быть и неединственными (рис. 9).

Если $d(x_0) = 0$, то по теореме 6.1 точка x_0 — точка минимума функции $f(x)$ на Ω .

Рассмотрим случай $d(x_0) > 0$. Положим

$$g(x_0) = \|w(x_0) - v(x_0)\|^{-1} (w(x_0) - v(x_0)). \quad (6.23)$$

Определение. Пусть $d(x_0) > 0$. Вектор $g_0 \in \Gamma$, $\|g_0\| = 1$, такой, что

$$\frac{\partial f(x_0)}{\partial g_0} = \inf_{g \in \Gamma(x_0), \|g\|=1} \frac{\partial f(x_0)}{\partial g}, \quad (6.24)$$

называется направлением наискорейшего спуска функции $f(x)$ на множестве Ω в точке x_0 .

Теорема 6.2. Если $d(x_0) > 0$, то вектор $g(x_0)$ является направлением наискорейшего спуска функции $f(x)$ на множестве Ω в точке x_0 , при этом

$$\frac{\partial f(x_0)}{\partial g(x_0)} = -d(x_0). \quad (6.25)$$

Это направление единственно.

Доказательство. Положим $B = \partial f(x)$, $\Gamma = \Gamma(x_0)$. Тогда из (6.22), (6.23) и (6.8) имеем

$$\sup_{v \in C} (v, g(x_0)) = -\|v(x_0) - w(x_0)\| = -d(x_0).$$

Поскольку $d(x_0) < +\infty$, то из (6.5) заключаем, что

$$g(x_0) \in \Gamma, \quad -d(x_0) = \max_{v \in \partial f(x_0)} (v, g(x_0)) = \frac{\partial f(x_0)}{\partial g(x_0)}. \quad (6.26)$$

Из (6.10) и (6.5) для всех $g \in \Gamma(x_0)$, $\|g\| = 1$, получаем

$$\frac{\partial f(x_0)}{\partial g} = \max_{v \in \partial f(x_0)} (v, g) \geq -d(x_0) = \frac{\partial f(x_0)}{\partial g(x_0)},$$

т. е. направление $g(x_0)$ — направление наискорейшего спуска функции $f(x)$ на Ω в точке x_0 . Из (6.10) вытекает также и единственность направления наискорейшего спуска. Равенство (6.25) следует из (6.26). ■

З а м е ч а н и е 4. Направление $g(x_0) \in \Gamma(x_0)$, но может оказаться, что $g(x_0) \notin \gamma(x_0)$, т. е. это направление может выводиться из множества Ω .

З а д а н и е 6.2. Пусть $\mathcal{M}^* = \{x \in E_n \mid f(x) = \inf_{z \in E_n} f(z)\}$. Показать, что если $\mathcal{M}^* \neq \emptyset$ и $\text{int } \mathcal{M}^* \neq \emptyset$, то

$$\partial f(x) = \{0\} \quad \forall x \in \text{int } \mathcal{M}^*.$$

§ 7. ε-субдифференциал

1. Пусть $f(x)$ — выпуклая конечная на E_n функция. Зафиксируем $x_0 \in E_n$, $\varepsilon \geq 0$ и положим

$$\partial_\varepsilon f(x_0) = \{v \in E_n \mid f(x) - f(x_0) \geq (v, x - x_0) - \varepsilon \quad \forall x \in E_n\}. \quad (7.1)$$

Множество $\partial_\varepsilon f(x_0)$ называется ε -субдифференциалом функции $f(x)$ в точке x_0 . Элемент $v \in \partial_\varepsilon f(x_0)$ называется ε -субградиентом функции $f(x)$ в точке x_0 .

Теорема 7.1. Для любых фиксированных $x_0 \in E_n$ и $\varepsilon \geq 0$ множество $\partial_\varepsilon f(x_0)$ непусто, замкнуто, выпукло и ограничено.

Доказательство. Поскольку $\partial f(x_0) \subset \partial_\varepsilon f(x_0)$ $\forall \varepsilon \geq 0$, то множество $\partial_\varepsilon f(x_0)$ непустое. Замкнутость и выпуклость очевидны из определения. Докажем ограниченность. Зафиксируем $\delta > 0$. В силу непрерывности функции $f(x)$ найдется такое $c < \infty$, что

$$\delta^{-1} [f(x) - f(x_0) + \varepsilon] \leq c$$

$$\forall x \in S_\delta(x_0) \equiv \{x \in E_n \mid \|x - x_0\| \leq \delta\}. \quad (7.2)$$

Возьмем любое $v \in \partial_\varepsilon f(x_0)$ и точку $x_1 = x_0 + \delta v \|v\|^{-1}$. Тогда

$$f(x_1) \geq f(x_0) + (v, x_1 - x_0) - \varepsilon = f(x_0) + \delta \|v\| - \varepsilon. \quad (7.3)$$

Отсюда

$$\|v\| \leq \delta^{-1} [f(x_1) - f(x_0) + \varepsilon] \leq c \quad \forall v \in \partial_\varepsilon f(x_0), \quad (7.4)$$

что и требовалось доказать. ■

Следствие. Отображение $\partial_\varepsilon f: [0, \infty) \times E_n \rightarrow \Pi(E_n)$ ограничено на любом ограниченном множестве из $[0, \infty) \times E_n$.

Доказательство аналогично доказательству ограниченности в теореме 7.1.

Пример 1. Пусть $f(x) = \delta(x - a)^2$, где $a \in E_n$, $\delta > 0$ фиксированы. Найдём $\partial_\varepsilon f(x)$. Пусть $v \in \partial_\varepsilon f(x)$. Представим v в виде $v = f'(x) + r = 2\delta(x - a) + r$. По определению

$$f(z) - f(x) \geq (v, z - x) - \varepsilon \quad \forall z \in E_n,$$

т. е.

$$\delta(z - a)^2 - \delta(x - a)^2 \geq (2\delta(x - a) + r, z - x) - \varepsilon \quad \forall z \in E_n. \quad (7.5)$$

Отсюда

$$\delta(z - x)^2 - (r, z - x) + \varepsilon \geq 0 \quad \forall z \in E_n. \quad (7.6)$$

Найдём минимум в левой части этого неравенства. Он достигается при $z^* = x + 0,5\delta^{-1}r$. Из (7.6) при $z = z^*$ имеем $-0,25\delta^{-1}r^2 + \varepsilon \geq 0$. Значит, $\|r\| \leq 2\sqrt{\varepsilon\delta}$. Это условие является необходимым и достаточным для выполнения (7.6). Таким образом,

$$\partial_\varepsilon f(x) = \{v = 2\delta(x - a) + r \mid r \in E_n, \|r\| \leq 2\sqrt{\varepsilon\delta}\}. \quad (7.7)$$

Задание 7.1. Построить множество $\partial_\varepsilon f(x)$ для функции $f(x) = (Ax, x)$, где A — симметричная положительно определенная $n \times n$ -матрица.

Пример 2. Пусть $f(x) = (a, x) + b$, где $a \in E_n$, $b \in E_1$. Возьмем $v \in \partial_\varepsilon f(x)$. Тогда

$$f(z) - f(x) \geq (v, z - x) - \varepsilon,$$

или

$$(a - v, z - x) \geq -\varepsilon \quad \forall z \in E_n. \quad (7.8)$$

Ясно, что $v = a$; поэтому $\partial_\varepsilon f(x) = \partial f(x) = \{a\}$.

2. Зафиксируем $x_0 \in E_n$ и $\varepsilon \geq 0$. Положим

$$B_\varepsilon f(x_0) = \{v \in E_n \mid \exists x \in E_n: v \in \partial f(x), f(x) - f(x_0) \geq (v, x - x_0) - \varepsilon\}. \quad (7.9)$$

Лемма 7.1. Справедливо соотношение

$$\partial_\varepsilon f(x_0) = \overline{B_\varepsilon f(x_0)}. \quad (7.10)$$

Доказательство мы опускаем, ибо соотношение (7.10) является частным случаем доказываемой ниже теоремы 10.1 при $\Omega = E_n$.

Из (7.10) вытекает

Следствие. Пусть x_0 зафиксировано. Тогда для любых $x \in E_n$ и $v \in \partial f(x)$ найдется $\varepsilon > 0$ такое, что $v \in \partial_\varepsilon f(x_0)$.

Замечание 1. Из (7.10) получаем следующую геометрическую интерпретацию ε -субдифференциала: субградиент v функции $f(x)$ в точке $\bar{x}$

принадлежит $\partial_\varepsilon f(x_0)$, если линейная функция $l(y) = (v, y - \bar{x}) + f(\bar{x})$, задающая опорную гиперплоскость к множеству $\text{epi } f$ в точке $\bar{x}$, принимает в точке x_0 значение $\geq f(x_0) - \varepsilon$ (рис. 10).

Приведем пример, показывающий, что замыкание в (7.10) существенно.

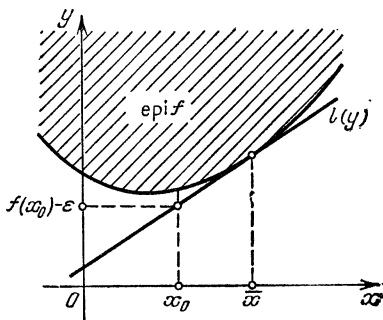


Рис. 10.

Пример 3. Пусть $x \in E_1$. Рассмотрим функцию

$$f(x) = \begin{cases} 3 - 2x, & x \leq 1, \\ x^{-1}, & x \geq 1. \end{cases}$$

Нетрудно проверить, что $f(x)$ — выпуклая функция. Возьмем точку $x_0 = 2$ и $\varepsilon = 0,5$. Имеем

$$f(x) - f(2) = \begin{cases} 3 - 2x - 0,5, & x \leq 1, \\ x^{-1} - 0,5, & x \geq 1. \end{cases} \quad (7.11)$$

Ясно, что

$$f(x) - f(2) \geq -0,5 = (0, x - 2) - 0,5 \quad \forall x \in E_1.$$

Отсюда следует, что

$$0 \in \partial_{0,5} f(2). \quad (7.12)$$

С другой стороны,

$$\partial f(x) = \begin{cases} \{2\}, & x < 1, \\ [-2, -1], & x = 1, \\ \{-x^{-2}\}, & x > 1. \end{cases}$$

Поэтому не существует такого $x \in E_1$, что $0 \in \partial f(x)$. Вспоминая определение множества $B_\varepsilon f(x_0)$ (см. (7.9)), заключаем, что $0 \notin B_{0,5} f(2)$, хотя, конечно, $0 \in \overline{B_{0,5} f(2)}$. Таким образом, знак замыкания в (7.10) убрать нельзя. Однако, как будет сейчас показано, в некоторых случаях это можно сделать.

Лемма 7.2. Если $f(x)$ — сильно выпуклая (см. п. 1 § 4) функция, то

$$\partial_\varepsilon f(x_0) = B_\varepsilon f(x_0). \quad (7.13)$$

Действительно, включение

$$B_\varepsilon f(x_0) \subset \partial_\varepsilon f(x_0) \quad (7.14)$$

очевидно из (7.10).

Докажем справедливость противоположного включения. Пусть $v \in \partial_\varepsilon f(x_0)$, т. е.

$$f(z) - f(x_0) \geq (v, z - x_0) - \varepsilon \quad \forall z \in E_n. \quad (7.15)$$

Положим $h(z) = f(z) - f(x_0) + (v, x_0 - z)$. Поскольку $f(x)$ — сильно выпуклая функция, то и $h(z)$ — сильно выпуклая функция, и потому она достигает своего

минимального значения в некоторой точке $x' \in E_n$:
 $h(x') = \min_{z \in E_n} h(z)$. Тогда $h(x') \leq h(z) \quad \forall z \in E_n$, т. е.

$f(x') - (v, x') \leq f(z) - (v, z)$ или $f(z) - f(x') \geq (v, z - x')$
 $\forall z \in E_n$, а это и означает, что $v \in \partial f(x')$. Из (7.15)
 $(v, x_0 - x') + f(x') - f(x_0) \geq -\varepsilon$, т. е. $v \in B_\varepsilon f(x_0)$,
 значит,

$$\partial_\varepsilon f(x_0) \subset B_\varepsilon f(x_0). \quad (7.16)$$

Отсюда и из (7.14) следует (7.13). ■

3. Пусть $f_1(x)$ и $f_2(x)$ — выпуклые конечные функции. Положим $f(x) = f_1(x) + f_2(x)$. По лемме 5.2

$$\partial f(x) = \partial f_1(x) + \partial f_2(x). \quad (7.17)$$

Для ε -субдифференциалов аналогичное соотношение уже не имеет места.

Теорема 7.2. *Справедливо представление*

$$\partial_\varepsilon f(x_0) = \bigcup_{\substack{\varepsilon_1 + \varepsilon_2 = \varepsilon \\ \varepsilon_1 \geq 0, \varepsilon_2 \geq 0}} [\partial_{\varepsilon_1} f_1(x_0) + \partial_{\varepsilon_2} f_2(x_0)]. \quad (7.18)$$

Доказательство. Выражение в правой части (7.18) обозначим $\mathcal{D}$. Включение

$$\mathcal{D} \subset \partial_\varepsilon f(x_0) \quad (7.19)$$

очевидно. Возьмем $v \in \partial_\varepsilon f(x_0)$. По лемме 7.1 существует последовательность $\{v_k\}$ такая, что

$$v_k \in B_\varepsilon f(x_0), \quad v_k \xrightarrow{k \rightarrow \infty} v. \quad (7.20)$$

По определению $B_\varepsilon f(x_0)$ для каждого v_k найдется x_k такое, что $v_k \in \partial f(x_k)$, $(v_k, x_0 - x_k) + f(x_k) - f(x_0) \geq -\varepsilon$. Из (7.17)

$$v_k = v'_k + v''_k, \quad (7.21)$$

где $v'_k \in \partial f_1(x_k)$, $v''_k \in \partial f_2(x_k)$, и при этом из (7.20)

$$\begin{aligned} & [(v'_k, x_0 - x_k) + f_1(x_k) - f_1(x_0)] + \\ & + [(v''_k, x_0 - x_k) + f_2(x_k) - f_2(x_0)] \geq -\varepsilon. \end{aligned} \quad (7.22)$$

Поскольку выражения в квадратных скобках в левой части (7.22) неположительны, то (7.22) возможно лишь

тогда, когда существуют такие $\varepsilon_{k1} \geq 0$ и $\varepsilon_{k2} \geq 0$, что $\varepsilon_{k1} + \varepsilon_{k2} = \varepsilon$,

$$(v'_k, x_0 - x_k) + f_1(x_k) - f_1(x_0) \geq -\varepsilon_{k1},$$

$$(v''_k, x_0 - x_k) + f_2(x_k) - f_2(x_0) \geq -\varepsilon_{k2},$$

т. е. $v'_k \in B_{\varepsilon_{k1}} f_1(x_0)$, $v''_k \in B_{\varepsilon_{k2}} f_2(x_0)$. Поскольку $f_1(x)$ и $f_2(x)$ — конечные функции,

$$B_{\varepsilon_{k1}} f_1(x_0) \subset \partial_{\varepsilon_{k1}} f_1(x_0) \subset \partial_{\varepsilon} f_1(x_0),$$

$$B_{\varepsilon_{k2}} f_2(x_0) \subset \partial_{\varepsilon_{k2}} f_2(x_0) \subset \partial_{\varepsilon} f_2(x_0),$$

множества $\partial_{\varepsilon} f_1(x_0)$ и $\partial_{\varepsilon} f_2(x_0)$ ограничены, то, не теряя общности, можем считать, что

$$v'_k \rightarrow v', \quad v''_k \rightarrow v'', \quad \varepsilon_{k1} \rightarrow \varepsilon' \geq 0, \quad \varepsilon_{k2} \rightarrow \varepsilon'' \geq 0.$$

При этом $\varepsilon' + \varepsilon'' = \varepsilon$. Так как $v'_k \in \partial_{\varepsilon_{k1}} f_1(x_0)$, то

$$f_1(z) - f_1(x_0) \geq (v'_k, z - x_0) - \varepsilon_{k1} \quad \forall z \in E_n.$$

Переходя здесь к пределу при каждом фиксированном $z \in E_n$, получим $f_1(z) - f_1(x_0) \geq (v', z - x_0) - \varepsilon'$, т. е. $v' \in \partial_{\varepsilon'} f_1(x_0)$. Аналогично $v'' \in \partial_{\varepsilon''} f_2(x_0)$. Поэтому из (7.20) и (7.21)

$$v = v' + v'' \in [\partial_{\varepsilon'} f_1(x_0) + \partial_{\varepsilon''} f_2(x_0)] \subset \mathcal{D}.$$

Значит, $\partial_{\varepsilon} f(x_0) \subset \mathcal{D}$. Отсюда и из (7.19) следует (7.18). ■

Следствие 1. Если $f_i(x)$, $i \in I \equiv 1:N$, — конечные выпуклые функции, то

$$\partial_{\varepsilon} \left(\sum_{i \in I} f_i(x) \right) = \bigcup_{\substack{i \in I \\ \varepsilon_i = \varepsilon, \varepsilon_i \geq 0}} \left[\sum_{i \in I} \partial_{\varepsilon_i} f_i(x) \right].$$

Следствие 2. Если $f_2(x)$ — линейная функция, то $\partial_{\varepsilon} f(x) = \partial_{\varepsilon} f_1(x) + \partial f_2(x)$.

Это вытекает из того, что для линейной функции $f_2(x) = (a, x) + b$ (см. пример 2)

$$\partial_{\varepsilon} f_2(x) = \partial f_2(x) = \{a\}.$$

Задание 7.2. Построить множество $\partial_{\varepsilon} f(x)$ для функции

$$f(x) = (Ax, x) + (a, x),$$

где A — положительно определенная $n \times n$ -матрица, $a \in E_n$.

З а м е ч а н и е 2. Отметим очевидное соотношение

$$\partial_\varepsilon [bf(x)] = b\partial_{\varepsilon b^{-1}} f(x).$$

Здесь $f(x)$ — выпуклая функция, $b > 0$.

З а м е ч а н и е 3. Положим

$$f_\varepsilon(x) = \sup_{z \in E_n} [f(z) + (v_\varepsilon(z), x - z)], \quad (7.23)$$

где $v_\varepsilon(z) \in \partial_\varepsilon f(z)$ (для каждого z выбирается один, все равно какой, вектор $v_\varepsilon(z)$ из $\partial_\varepsilon f(z)$). Из определения $\partial_\varepsilon f(z)$ очевидно, что

$$0 \leq f_\varepsilon(x) - f(x) \leq \varepsilon \quad (7.24)$$

(ср. с (5.11)).

4. Найдем ε -субдифференциал функции максимума. Пусть $f_i(x)$, $i \in I \equiv 1:N$, — выпуклые на E_n функции. Положим

$$f(x) = \max_{i \in I} f_i(x). \quad (7.25)$$

Возьмем $x_0 \in E_n$, $\varepsilon \geq 0$ и построим множество

$$\mathcal{H}_\varepsilon(x_0) = \left\{ v \in E_n \left| \begin{array}{l} v = \sum_{i \in I_1} \alpha_i v_i, \quad \alpha_i > 0, \\ \sum_{i \in I_1} \alpha_i = 1, \quad I_1 \subset I, \\ v_i \in \partial_{\varepsilon_i \alpha_i^{-1}} f_i(x_0), \\ \sum_{i \in I_1} \varepsilon_i + \varepsilon_0 = \varepsilon, \quad \varepsilon_i \geq 0, \quad \varepsilon_0 \geq 0, \\ \sum_{i \in I_1} \alpha_i f_i(x_0) \geq f(x_0) - \varepsilon_0. \end{array} \right. \right. \quad (7.26)$$

Теорема 7.3. Для функции максимума (7.25) имеет место представление

$$\partial_\varepsilon f(x_0) = \overline{\mathcal{H}_\varepsilon(x_0)}. \quad (7.27)$$

Доказательство. Возьмем любое $v \in \mathcal{H}_\varepsilon(x_0)$. Тогда

$$v = \sum_{i \in I_1} \alpha_i v_i, \quad I_1 \subset I, \quad v_i \in \partial_{\varepsilon_i \alpha_i^{-1}} f_i(x_0).$$

Значит,

$$f_i(z) - f_i(x_0) \geq (v_i, z - x_0) - \varepsilon_i \alpha_i^{-1} \quad \forall z \in E_n, \quad \forall i \in I_1. \quad (7.28)$$

Умножая обе части (7.28) на α_i и складывая, получим

$$\sum_{i \in I_1} \alpha_i f_i(z) - \sum_{i \in I_1} \alpha_i f_i(x_0) \geq (v, z - x_0) - \sum_{i \in I_1} \varepsilon_i \quad (7.29)$$

$$\forall z \in E_n.$$

Поскольку $\sum_{i \in I_1} \alpha_i f_i(z) \leq f(z)$, то из (7.26) и (7.29)

$$\begin{aligned} f(z) - f(x_0) &\geq (v, z - x_0) - \sum_{i \in I_1} \varepsilon_i - \varepsilon_0 = \\ &= (v, z - x_0) - \varepsilon \quad \forall z \in E_n. \end{aligned}$$

Отсюда $v \in \partial_e f(x_0)$. Итак, доказано, что $\mathcal{H}_\varepsilon(x_0) \subset \subset \partial_e f(x_0)$. В силу замкнутости $\partial_e f(x_0)$ отсюда вытекает, что

$$\overline{\mathcal{H}_\varepsilon(x_0)} \subset \partial_e f(x_0). \quad (7.30)$$

Пусть теперь $v \in \partial_e f(x_0)$. Тогда по теореме 7.1 $v \in \overline{B_\varepsilon f(x_0)}$, т. е.

$$v = \lim_{k \rightarrow \infty} w_k, \quad w_k \in B_\varepsilon f(x_0). \quad (7.31)$$

По определению $B_\varepsilon f(x_0)$ для каждого k найдется точка $x_k \in E_n$ такая, что

$$w_k \in \partial f(x_k), \quad (w_k, x_0 - x_k) + f(x_k) - f(x_0) \geq -\varepsilon. \quad (7.32)$$

По формуле (5.21)

$$\partial f(x_k) = \text{co} \{ \partial f_i(x_k) \mid i \in R(x_k) \},$$

где

$$R(x_k) = \{ i \in I \mid f_i(x_k) = f(x_k) \}.$$

Отсюда

$$\begin{aligned} w_k &= \sum_{i \in R(x_k)} \alpha_{ki} v_{ki}, \quad v_{ki} \in \partial f_i(x_k), \\ \sum_{i \in R(x_k)} \alpha_{ki} &= 1, \quad \alpha_{ki} \geq 0, \end{aligned}$$

т. е.

$$f_i(z) - f_i(x_k) \geq (v_{ki}, z - x_k) \quad \forall z \in E_n. \quad (7.33)$$

Пусть

$$R_k = \{ i \in R(x_k) \mid \alpha_{ki} > 0 \}.$$

Из (7.33) для каждого $i \in R_k$ имеем

$$\begin{aligned} f_i(z) - f_i(x_0) &\geq (v_{ki}, z - x_0) + \\ &+ [f_i(x_k) - f_i(x_0) + (v_{ki}, x_0 - x_k)] \quad \forall z. \end{aligned} \quad (7.34)$$

Положим

$$N_{ki} \equiv f_i(x_k) - f_i(x_0) + (v_{ki}, x_0 - x_k) = -\alpha_{ki}^{-1} e_{ki}. \quad (7.35)$$

Тогда из (7.34) для каждого $i \in R_k$ получим

$$f_i(z) - f_i(x_0) \geq (v_{ki}, z - x_0) - \alpha_{ki}^{-1} e_{ki} \quad \forall z, \quad (7.36)$$

т. е.

$$v_{ki} \in \partial_{\alpha_{ki}^{-1} e_{ki}} f_i(x_0). \quad (7.37)$$

Из (7.36) при $z = x_0$ получаем

$$N_{ki} = -\alpha_{ki}^{-1} e_{ki} \leq 0. \quad (7.38)$$

Отсюда

$$\begin{aligned} \sum_{i \in R_k} \alpha_{ki} N_{ki} &= - \sum_{i \in R_k} e_{ki} = \\ &= (w_k, x_0 - x_k) + \sum_{i \in R_k} \alpha_{ki} f_i(x_k) - \sum_{i \in R_k} \alpha_{ki} f_i(x_0). \end{aligned} \quad (7.39)$$

Так как

$$\sum_{i \in R_k} \alpha_{ki} f_i(x_k) = f(x_k), \quad \sum_{i \in R_k} \alpha_{ki} f_i(x_0) \leq f(x_0),$$

то из (7.39) и (7.32)

$$\begin{aligned} \sum_{i \in R_k} \alpha_{ki} N_{ki} &= - \sum_{i \in R_k} e_{ki} \geq \\ &\geq (w_k, x_0 - x_k) + f(x_k) - f(x_0) \geq -\varepsilon. \end{aligned}$$

Итак,

$$\sum_{i \in R_k} e_{ki} \leq \varepsilon. \quad (7.40)$$

Из (7.38) также вытекает

$$e_{ki} \geq 0. \quad (7.41)$$

Умножая (7.35) на α_{ki} и суммируя, получаем

$$f(x_k) - \sum_{i \in R_k} \alpha_{ki} f_i(x_0) + (w_k, x_0 - x_k) = - \sum_{i \in R_k} e_{ki},$$

т. е.

$$\sum_{i \in R_k} \alpha_{ki} f_i(x_0) = f(x_k) + (w_k, x_0 - x_k) + \sum_{i \in R_k} e_{ki}.$$

Положим $\varepsilon_{k0} = \varepsilon - \sum_{i \in R_k} \varepsilon_{ki}$. Из (7.40) ясно, что

$$\varepsilon_{k0} \geq 0. \quad (7.42)$$

Из (7.32)

$$\sum_{i \in R_k} \alpha_{ki} f_i(x_0) \geq f(x_0) - \varepsilon + \sum_{i \in R_k} \varepsilon_{ki} = f(x_0) - \varepsilon_{k0}. \quad (7.43)$$

При этом

$$\sum_{i \in R_k} \varepsilon_{ki} + \varepsilon_{k0} = \sum_{i \in R_k} \varepsilon_{ki} + \varepsilon - \sum_{i \in R_k} \varepsilon_{ki} = \varepsilon.$$

Из (7.37), (7.40), (7.43) и определения (7.26) следует

$$\omega_k \in \mathcal{H}_\varepsilon(x_0).$$

Отсюда и из (7.31) вытекает $v \in \overline{\mathcal{H}_\varepsilon(x_0)}$, т. е.

$$\partial_\varepsilon f(x_0) \subset \overline{\mathcal{H}_\varepsilon(x_0)}. \quad (7.44)$$

Утверждение теоремы следует из (7.30) и (7.44). ■

Задание 7.3. Найти вид $\partial_\varepsilon f(x_0)$ для функции $f(x) = \max_{y \in G} \varphi(x, y)$, где $\varphi(x, y)$ выпукла по x при любом $y \in G$, непрерывна по x и y на $E_n \times G$, G — компактное множество.

Пример 4. Проиллюстрируем применение формулы (7.27). Пусть $x \in E_1$, $f(x) = |x|$. Найдем $\partial_\varepsilon f(x)$, где $\varepsilon \geq 0$.

Так как $|x| = \max\{x, -x\}$, то $f(x) = \max_{i \in 1:2} f_i(x)$, где $f_1(x) = x$, $f_2(x) = -x$. В примере 2 было показано, что

$$\begin{aligned} \partial_\varepsilon f_1(x) &= \{+1\} \quad \forall \varepsilon \geq 0, \quad \forall x \in E_1, \\ \partial_\varepsilon f_2(x) &= \{-1\} \quad \forall \varepsilon \geq 0, \quad \forall x \in E_1. \end{aligned} \quad (7.45)$$

В (7.26) в качестве I_1 могут выступать множества:

- 1) $I_1 = \{1\}$; 2) $I_1 = \{2\}$; 3) $I_1 = \{1, 2\}$.

Поэтому из (7.45)

$$\mathcal{H}_\varepsilon(x) = \mathcal{H}_1 \cup \mathcal{H}_2 \cup \mathcal{H}_3, \quad (7.46)$$

где

$$\begin{aligned} \mathcal{H}_1 &= \{v \in E_1 \mid v = 1, \quad x > |x| - \varepsilon_0, \quad \varepsilon_0 \in [0, \varepsilon]\}, \\ \mathcal{H}_2 &= \{v \in E_1 \mid v = -1, \quad -x \geq |x| - \varepsilon_0, \quad \varepsilon_0 \in [0, \varepsilon]\}, \\ \mathcal{H}_3 &= \{v \in E_1 \mid v = \alpha + (1 - \alpha)(-1) = 2\alpha - 1, \quad \alpha \in (0, 1), \\ &\quad \alpha x + (1 - \alpha)x \geq |x| - \varepsilon_0, \quad \varepsilon_0 \in [0, \varepsilon]\}. \end{aligned}$$

Если $x > 0$, то $|x| = x$ и потому

$$\mathcal{K}_1 = \{1\}, \quad \mathcal{K}_2 = \begin{cases} \{-1\}, & \text{если } x \leq 0,5\varepsilon, \\ \emptyset, & \text{если } x > 0,5\varepsilon, \end{cases}$$

$$\mathcal{K}_3 = \{v \in E_1 \mid v = 2\alpha - 1, \alpha \in (0, 1), \alpha \geq 1 - \frac{\varepsilon_0}{2x},$$

$$\varepsilon_0 \in [0, \varepsilon]\}.$$

Отсюда ясно, что

$$\partial_\varepsilon f(x) = \overline{K_\varepsilon(x)} = \begin{cases} [1 - \varepsilon x^{-1}, 1], & \text{если } x > 0,5\varepsilon, \\ [-1, 1], & \text{если } 0 \leq x \leq 0,5\varepsilon. \end{cases} \quad (7.47)$$

Аналогично в случае $x < 0$ имеем $|x| = -x$ и потому

$$\mathcal{K}_2 = \{-1\}, \quad \mathcal{K}_1 = \begin{cases} \{1\}, & \text{если } x \geq -0,5\varepsilon, \\ \emptyset, & \text{если } x < -0,5\varepsilon, \end{cases}$$

$$\mathcal{K}_3 = \left\{ v \in E_1 \mid v = 2\alpha - 1, \alpha \in (0, 1), \alpha \leq -\frac{\varepsilon_0}{2x}, \right.$$

$$\left. \varepsilon_0 \in [0, \varepsilon] \right\}.$$

Отсюда

$$\partial_\varepsilon f(x) = \begin{cases} [-1, 1], & \text{если } 0 \geq x \geq -0,5\varepsilon, \\ [-1, -1 - \varepsilon x^{-1}], & \text{если } x < -0,5\varepsilon. \end{cases} \quad (7.48)$$

Из (7.47) и (7.48) окончательно получаем

$$\partial_\varepsilon f(x) = \begin{cases} [1 - \varepsilon x^{-1}, 1], & \text{если } x > 0,5\varepsilon, \\ [-1, 1], & \text{если } |x| \leq 0,5\varepsilon, \\ [-1, -1 - \varepsilon x^{-1}], & \text{если } x < -0,5\varepsilon. \end{cases}$$

§ 8. ε-производные по направлению. Непрерывность ε-субдифференциального отображения

1. Пусть $f(x)$ — выпуклая конечная на E_n функция. Зафиксируем $\varepsilon \geq 0$. Величину

$$\frac{\partial_\varepsilon f(x_0)}{\partial g} = \max_{v \in \partial_\varepsilon f(x_0)} (v, g) \quad (8.1)$$

будем называть *ε-производной функции $f(x)$ в точке x_0 по направлению g* .

Теорема 8.1. *Справедливо соотношение*

$$\frac{\partial_{\varepsilon} f(x_0)}{\partial g} = \inf_{\alpha > 0} \alpha^{-1} [f(x_0 + \alpha g) - f(x_0) + \varepsilon]. \quad (8.2)$$

Доказательство. Положим

$$h(\alpha) = \alpha^{-1} [f(x_0 + \alpha g) - f(x_0) + \varepsilon], \quad \mathcal{D} = \inf_{\alpha > 0} h(\alpha). \quad (8.3)$$

Поскольку по определению

$$\begin{aligned} f(x_0 + \alpha g) - f(x_0) &\geq (v, \alpha g) - \varepsilon \\ \forall v \in \partial_{\varepsilon} f(x_0), \quad \forall \alpha \geq 0, \end{aligned}$$

то

$$(v, g) \leq h(\alpha) \quad \forall \alpha \geq 0, \quad \forall v \in \partial_{\varepsilon} f(x_0).$$

Следовательно,

$$\frac{\partial_{\varepsilon} f(x_0)}{\partial g} \equiv \max_{v \in \partial_{\varepsilon} f(x_0)} (v, g) \leq \inf_{\alpha > 0} h(\alpha) \equiv \mathcal{D}. \quad (8.4)$$

Из (8.4) следует, что $\mathcal{D} > -\infty$.

Зафиксируем x_0 и g . Введем множества

$$\begin{aligned} Z &= \{[\beta, x] \in E_{n+1} \mid x \in E_n, \beta > f(x)\}, \\ \mathcal{L} &= \{[\beta, x] \in E_{n+1} \mid x = x_0 + \alpha g, \beta = f(x_0) + \alpha \mathcal{D} - \varepsilon, \alpha > 0\}. \end{aligned}$$

Покажем, что

$$\mathcal{L} \cap Z = \emptyset. \quad (8.5)$$

Допустим противное, т. е. что при некоторых α_1 , β_1 и x_1 окажется

$$\begin{aligned} \beta_1 &> f(x_1), \quad \beta_1 = f(x_0) + \alpha_1 \mathcal{D} - \varepsilon, \\ x_1 &= x_0 + \alpha_1 g, \quad \alpha_1 > 0. \end{aligned} \quad (8.6)$$

Из (8.6), в частности, имеем

$$f(x_0) + \alpha_1 \mathcal{D} - \varepsilon > f(x_1).$$

Поэтому

$$\begin{aligned} \mathcal{D} &> \frac{1}{\alpha_1} [f(x_1) - f(x_0) + \varepsilon] = \\ &= \frac{1}{\alpha_1} [f(x_0 + \alpha_1 g) - f(x_0) + \varepsilon] = h(\alpha_1). \end{aligned} \quad (8.7)$$

Но (8.7) противоречит (8.3). Таким образом, (8.5) доказано. Образует множество $P = Z - \mathcal{L}$. Ясно, что оно вы-

пукло, причем $0 \notin P$. По следствию 2 из теоремы отделимости существует такая точка $[c, v] \in E_{n+1}$, что

$$c \in E_1, \quad v \in E_n, \quad c^2 + v^2 > 0, \quad (8.8)$$

$$c\beta + (v, x) \geq c(f(x_0) + \alpha\mathcal{D} - \varepsilon) + (v, x_0 + \alpha g) \\ \forall \beta, x, \alpha: \beta \geq f(x), \quad \alpha \geq 0. \quad (8.9)$$

Покажем, что $c > 0$. При $\alpha = 0, x = x_0$ из (8.9) имеем

$$c(\beta - f(x_0) + \varepsilon) \geq 0 \quad \forall \beta > f(x_0).$$

Отсюда $c \geq 0$. Если бы оказалось

$$c = 0, \quad (8.10)$$

то из (8.8) было бы

$$\|v\| > 0. \quad (8.11)$$

При $\alpha = 0$ из (8.9) и (8.10) получаем

$$(v, x - x_0) \geq 0 \quad \forall x \in E_n,$$

что возможно лишь при $v = 0$, а это противоречит (8.11).

Итак, $c > 0$. Тогда из (8.9) при $\beta = f(x), \alpha = 0$

$$f(x) - f(x_0) \geq (-c^{-1}v, x - x_0) - \varepsilon \quad \forall x \in E_n,$$

т. е.

$$v^* = (-c^{-1}v) \in \partial_\varepsilon f(x_0).$$

Из (8.9) при $\beta = f(x)$

$$f(x) - f(x_0) \geq (v^*, x - x_0) - \varepsilon + \alpha(\mathcal{D} - (v^*, g)).$$

При $x = x_0$

$$\mathcal{D} \leq (v^*, g) + \alpha^{-1}\varepsilon \quad \forall \alpha > 0.$$

При $\alpha \rightarrow +\infty$ получаем

$$\mathcal{D} \leq (v^*, g) \leq \max_{v \in \partial_\varepsilon f(x_0)} (v, g) = \frac{\partial_\varepsilon f(x_0)}{\partial g}.$$

Отсюда и из (8.4) следует (8.2). Теорема доказана.

Замечание 1. При $\varepsilon = 0$ функция $h(\alpha) = \alpha^{-1}[f(x_0 + \alpha g) - f(x_0)]$ является, как следует из доказательства теоремы 4.2, невозрастающей при $\alpha \rightarrow +0$; поэтому

$$\frac{\partial f(x_0)}{\partial g} = \lim_{\alpha \rightarrow +0} \alpha^{-1}[f(x_0 + \alpha g) - f(x_0)] = \\ = \inf_{\alpha > 0} \alpha^{-1}[f(x_0 + \alpha g) - f(x_0)],$$

что совпадает с (8.2) при $\varepsilon = 0$.

2. Пусть $f(x)$ — конечная выпуклая функция. Рассмотрим отображение $\partial_\varepsilon f: [0, \infty) \times E_n \rightarrow \Pi(E_n)$.

Теорема 8.2. Отображение $\partial_\varepsilon f(x)$ H -непрерывно по ε и x на $(0, \infty) \times E_n$.

Доказательство. Зафиксируем любые $\varepsilon_0 > 0$ и $x_0 \in E_n$. Вначале рассмотрим случай, когда $f(x)$ — сильно выпуклая функция, т. е. существует такое $m > 0$, что

$$f(0,5(x+y)) \leq 0,5[f(x) + f(y)] - 0,25m\|x-y\|^2 \quad \forall x, y \in E_n. \quad (8.12)$$

Возьмем $v \in \partial f(x)$, т. е.

$$f(z) \geq f(x) + (v, z-x) \quad \forall z \in E_n. \quad (8.13)$$

Положим $z = 0,5(x+y)$. Из (8.12) и (8.13)

$$\begin{aligned} 0,5(f(x) + f(y)) - 0,25m\|x-y\|^2 &\geq f(0,5(x+y)) = f(z) \geq \\ &\geq f(x) + (v, z-x) = f(x) + 0,5(v, y-x) \quad \forall y \in E_n. \end{aligned}$$

Отсюда

$$f(y) \geq f(x) + (v, y-x) + 0,5m\|x-y\|^2 \quad \forall y \in E_n. \quad (8.14)$$

(В действительности здесь можно вместо $0,5m$ поставить m , как будет показано в § 9 (см. (9.4)).)

В теореме 8.1 установлено, что

$$\begin{aligned} \max_{v \in \partial_\varepsilon f(x)} (v, g) &= \inf_{\alpha > 0} \alpha^{-1} [f(x + \alpha g) - f(x) + \varepsilon] = \\ &= \inf_{\alpha > 0} h(x, g, \alpha, \varepsilon), \end{aligned}$$

где

$$h(x, g, \alpha, \varepsilon) = \frac{1}{\alpha} [f(x + \alpha g) - f(x) + \varepsilon].$$

Заметим, что при $\varepsilon > 0$

$$h(x, g, \alpha, \varepsilon) \xrightarrow{\alpha \rightarrow +0} +\infty.$$

Поскольку $\varepsilon_0 > 0$, а $f(x)$ — непрерывная функция, то существуют $\alpha_0 > 0$ и $\beta > 0$ такие, что

$$\inf_{\alpha > 0} h(x, g, \alpha, \varepsilon) = \inf_{\alpha \geq \alpha_0} h(x, g, \alpha, \varepsilon) \quad \forall x \in S_\beta(x_0),$$

$$\forall \varepsilon: |\varepsilon - \varepsilon_0| \leq \beta, \quad \forall g: \|g\| = 1.$$

С другой стороны, из (8.14) вытекает, что $h(x, g, \alpha, \varepsilon) \xrightarrow{\alpha \rightarrow +\infty} +\infty$; поэтому существует $\alpha_1 > \alpha_0$ такое, что

$$\inf_{\alpha > 0} h(x, g, \alpha, \varepsilon) = \inf_{\alpha_0 \leq \alpha \leq \alpha_1} h(x, g, \alpha, \varepsilon) \quad \forall x \in S_\beta(x_0),$$

$$\forall \varepsilon: |\varepsilon - \varepsilon_0| \leq \beta, \quad \forall g: \|g\| = 1.$$

Поскольку h — непрерывная по всем переменным функция, то

$$\max_{v \in \partial_\varepsilon f(x)} (v, g) = \min_{\alpha_0 \leq \alpha \leq \alpha_1} h(x, g, \alpha, \varepsilon) \equiv H(x, g, \varepsilon).$$

Функция $H(x, g, \varepsilon) \equiv \max_{v \in \partial_\varepsilon f(x)} (v, g)$ тоже непрерывна

по x , ε и g . Отсюда, по лемме 2.2, следует, что отображение $\partial_\varepsilon f(x)$ H -непрерывно по ε и x в точке $[\varepsilon_0, x_0]$, если $f(x)$ — сильно выпуклая функция.

Пусть теперь $f(x)$ — выпуклая (но не обязательно сильно) функция. Образует функцию $\mathcal{F}_\delta(x) = f(x) + \delta(x - x_0)^2 \equiv f(x) + f_\delta(x)$, где $\delta > 0$. Эта функция является сильно выпуклой, поэтому отображение $\partial_\varepsilon \mathcal{F}_\delta(x)$ H -непрерывно по ε и x в точке $[\varepsilon_0, x_0]$. По теореме 7.2

$$\partial_\varepsilon \mathcal{F}_\delta(x) = \bigcup_{\substack{\varepsilon_1 + \varepsilon_2 = \varepsilon \\ \varepsilon_1 \geq 0, \varepsilon_2 \geq 0}} [\partial_{\varepsilon_1} f(x) + \partial_{\varepsilon_2} f_\delta(x)]. \quad (8.15)$$

Поскольку (см. (7.7))

$$\partial_\varepsilon f_\delta(x) = \{v = 2\delta(x - x_0) + r \mid \|r\| \leq 2\sqrt{\varepsilon\delta}\},$$

то отсюда и из (8.15) ясно, что при $\delta \rightarrow 0$

$$\rho(\partial_\varepsilon \mathcal{F}_\delta(x), \partial_\varepsilon f(x)) \rightarrow 0, \quad (8.16)$$

где $\rho(A, B)$ — расстояние между множествами A и B в метрике Хаусдорфа (см. (2.1)). Поскольку отображение $\partial_\varepsilon \mathcal{F}_\delta(x)$ H -непрерывно по ε и x , то из (8.16) заключаем, что $\partial_\varepsilon f(x)$ тоже H -непрерывно по ε и x в точке $[\varepsilon_0, x_0]$. Теорема доказана.

3. Определение. Пусть $\varepsilon \geq 0$. Точку $x_0 \in E_n$ будем называть ε -стационарной, если

$$0 \in \partial_\varepsilon f(x_0). \quad (8.17)$$

Лемма 8.1. Соотношение (8.17) эквивалентно соотношению

$$0 \leq f(x_0) - f^* \leq \varepsilon, \quad (8.18)$$

где

$$f^* = \inf_{x \in E_n} f(x).$$

Доказательство. В силу определения $\partial_\varepsilon f(x_0)$ (8.17) эквивалентно неравенству

$$f(x) \geq f(x_0) + (0, x - x_0) - \varepsilon = f(x_0) - \varepsilon \quad \forall x \in E_n,$$

что и требовалось доказать.

Пусть точка x_0 не является ε -стационарной. Тогда (см. доказательство теоремы 6.2)

$$\inf_{\|g\|=1} \frac{\partial_\varepsilon f(x_0)}{\partial g} = \min_{\|g\|=1} \max_{v \in \partial_\varepsilon f(x_0)} (v, g) = \max_{v \in \partial_\varepsilon f(x_0)} (v, g) < 0.$$

При этом

$$g_0 = -\|v_0\|^{-1} v_0, \quad \text{где} \quad \|v_0\| = \min_{v \in \partial_\varepsilon f(x_0)} \|v\|.$$

Направление $g_0 = g(x_0)$ называется направлением *е-наискорейшего спуска* функции $f(x)$ в точке x_0 . Оно единственно. Это доказывается так же, как доказывалась единственность в теореме 6.2 (см. замечание 4 § 6).

Из (8.2) имеем

$$\inf_{\alpha > 0} \alpha^{-1} [f(x_0 + \alpha g_0) - f(x_0) + \varepsilon] < 0.$$

Отсюда следует существование $\alpha_0 > 0$ такого, что

$$f(x_0 + \alpha_0 g_0) - f(x_0) + \varepsilon < 0,$$

т. е.

$$\inf_{\alpha > 0} f(x_0 + \alpha g_0) < f(x_0) - \varepsilon. \quad (8.19)$$

Значит, в направлении g_0 функция $f(x)$ может быть уменьшена не меньше, чем на ε .

Это свойство справедливо для любого $g \in E_n$ такого, что

$$\frac{\partial_\varepsilon f(x_0)}{\partial g} < 0.$$

4. Пусть $\Omega \subset E_n$ — выпуклое замкнутое множество, $f(x)$ — выпуклая на Ω функция. Было доказано (теорема 6.1), что для того чтобы в точке $x^* \in \Omega$ функция $f(x)$ достигала своего наименьшего на Ω значения, необходимо и достаточно, чтобы

$$\partial f(x^*) \cap \Gamma^+(x^*) \neq \emptyset, \quad (8.20)$$

где $\Gamma^+(x^*)$ — конус, сопряженный конусу возможных направлений.

Лемма 8.2. Если $x_0 \in \Omega$ и

$$\partial_\varepsilon f(x_0) \cap \Gamma^+(x_0) \neq \emptyset, \quad (8.21)$$

то

$$0 \leq f(x_0) - f^* \leq \varepsilon, \quad (8.22)$$

где

$$f^* = \inf_{x \in \Omega} f(x).$$

Доказательство. Из (8.21) следует, что найдется такое $v_0 \in \partial_\varepsilon f(x_0)$, что $v_0 \in \Gamma^+(x_0)$, т. е.

$$(x - x_0, v_0) \geq 0 \quad \forall x \in \Omega. \quad (8.23)$$

Из определения $\partial_\varepsilon f(x_0)$

$$f(x) - f(x_0) \geq (v, x - x_0) - \varepsilon \quad \forall x \in E_n. \quad (8.24)$$

Отсюда и из (8.23)

$$f(x) - f(x_0) \geq -\varepsilon \quad \forall x \in \Omega.$$

Последнее неравенство эквивалентно (8.22). ■

5. Пусть на E_n задана линейная функция $g(x)$, а на множестве $\Omega = \{x \in E_n \mid g(x) > 0\}$ — выпуклая функция $f(x)$. Положим

$$h(x) = f(x) g^{-1}(x) \text{ (здесь } g^{-1}(x) = 1/g(x)),$$

$$h^* = \inf_{x \in \Omega} h(x), \quad \mathcal{M}^* = \{x \in \Omega \mid h(x) = h^*\}.$$

Лемма 8.3. Всякая точка локального минимума функции $h(x)$ на Ω является точкой ее глобального минимума на Ω . Множество $\mathcal{M}^*$ выпукло.

Доказательство. Пусть $x_0 \in \Omega$ — точка локального минимума функции $h(x)$ на Ω . Это означает, что найдется такое $\delta > 0$, что $S_\delta(x_0) \subset \Omega$, ибо Ω открыто, и

$$h(x) \geq h(x_0) \quad \forall x \in S_\delta(x_0). \quad (8.25)$$

Надо показать, что $h(x) \geq h(x_0)$ при всех $x \in \Omega$. Возьмем произвольную точку $\bar{x} \in \Omega \setminus S_\delta(x_0)$. Пусть прямая, проходящая через точки $\bar{x}$ и x_0 , пересекает сферу $S_\delta(x_0) = \{x \mid \|x - x_0\| = \delta\}$ в точках x_1 и x_2 . Тогда из (8.25) следует

$$f(x_1) \geq g(x_1) h(x_0), \quad f(x_2) \geq g(x_2) h(x_0). \quad (8.26)$$

Ясно, что найдется такое $\beta \notin [0, 1]$, что $\bar{x} = \beta x_1 + (1 - \beta)x_2$. В силу выпуклости функций $f(x)$ и $-g(x)$ будет (см. лемму 4.1)

$$f(\bar{x}) \geq \beta f(x_1) + (1 - \beta)f(x_2), \quad g(\bar{x}) \leq \beta g(x_1) + (1 - \beta)g(x_2).$$

Отсюда и из (8.26) получим $f(\bar{x}) \geq \beta g(x_1)h(x_0) + (1 - \beta)g(x_2)h(x_0) = (\beta g(x_1) + (1 - \beta)g(x_2))h(x_0) \geq g(\bar{x})h(x_0)$, т. е. $h(\bar{x}) \geq h(x_0)$. Поэтому, так как $\bar{x}$ — произвольная точка из $\Omega \setminus S_\delta(x_0)$ и выполнено (8.25), то

$$h(x) \geq h(x_0) = h^* \quad \forall x \in \Omega. \quad (8.27)$$

Из (8.27) вытекает, что x_0 — точка глобального минимума функции $h(x)$ на множестве Ω . Так как $h(x)$ квазивыпукла, то множество $\mathcal{M}^*$ выпукло (см. п. 6 § 4). Лемма доказана.

Возьмем $x_0 \in E_n$, $v \in E_n$, $\|v\| = 1$, и $\varepsilon \geq 0$.

Следствие. Любой локальный минимум функции $h(\alpha) = \alpha^{-1}[f(x_0 + \alpha v) - f(x_0) + \varepsilon]$ на $(0, \infty)$ является глобальным минимумом функции $h(\alpha)$ на $(0, \infty)$, причем множество этих минимумов выпукло.

6. Лемма 8.4. Пусть $x_0, g \in E_n$. Для того чтобы

$$f(x_0 + \alpha g) = f(x_0) + \alpha \frac{\partial f(x_0)}{\partial g} \quad \forall \alpha \geq 0, \quad (8.28)$$

необходимо, чтобы для любого $\varepsilon > 0$, и достаточно, чтобы хотя бы для одного $\varepsilon > 0$ было

$$\frac{\partial f(x_0)}{\partial g} = \frac{\partial_\varepsilon f(x_0)}{\partial g}. \quad (8.29)$$

Доказательство. Необходимость. В силу теоремы 8.1 из (8.28) имеем

$$\begin{aligned} \frac{\partial_\varepsilon f(x_0)}{\partial g} &= \inf_{\alpha > 0} \alpha^{-1} [f(x_0 + \alpha g) - f(x_0) + \varepsilon] = \\ &= \frac{\partial f(x_0)}{\partial g} + \inf_{\alpha > 0} \alpha^{-1} \varepsilon = \frac{\partial f(x_0)}{\partial g}. \end{aligned}$$

Достаточность. Покажем, что если выполнено (8.29), то множество $G = \{\alpha \geq 0 \mid f(x_0 + \alpha g) = f(x_0) + \alpha \frac{\partial_\varepsilon f(x_0)}{\partial g} - \varepsilon\}$ пусто. Допустим противное. Пусть при некотором $\lambda > 0$

$$f(x_0 + \lambda g) = f(x_0) + \lambda \frac{\partial_\varepsilon f(x_0)}{\partial g} - \varepsilon. \quad (8.30)$$

Для любого $v \in \partial f(x_0)$

$$f(x_0 + \alpha g) - f(x_0) \geq (v, \alpha g).$$

Отсюда

$$f(x_0 + \alpha g) - f(x_0) \geq \alpha \max_{v \in \partial f(x_0)} (v, g) = \alpha \frac{\partial f(x_0)}{\partial g}. \quad (8.31)$$

Из (8.30) и (8.29) имеем

$$\begin{aligned} \frac{\partial f(x_0)}{\partial g} &= \frac{\partial_e f(x_0)}{\partial g} = \lambda^{-1} [f(x_0 + \lambda g) - f(x_0) + \varepsilon] \geq \\ &\geq \lambda^{-1} \left[f(x_0) + \lambda \frac{\partial f(x_0)}{\partial g} - f(x_0) + \varepsilon \right] = \frac{\partial f(x_0)}{\partial g} + \lambda^{-1} \varepsilon, \end{aligned}$$

что невозможно. Итак, $G = \emptyset$, следовательно,

$$\frac{\partial_e f(x_0)}{\partial g} = \lim_{\alpha \rightarrow +\infty} \alpha^{-1} [f(x_0 + \alpha g) - f(x_0) + \varepsilon]. \quad (8.32)$$

Предположим, что (8.28) не имеет места, т. е. нашлось такое $\gamma > 0$, что

$$f(x_0 + \gamma g) > f(x_0) + \gamma \frac{\partial f(x_0)}{\partial g}. \quad (8.33)$$

Для любого $v \in \partial f(x_0 + \gamma g)$ будет

$$f(x_0) - f(x_0 + \gamma g) \geq -\gamma (v, g).$$

Отсюда

$$\gamma^{-1} [f(x_0 + \gamma g) - f(x_0)] \leq \max_{v \in \partial f(x_0 + \gamma g)} (v, g) = \frac{\partial f(x_0 + \gamma g)}{\partial g}. \quad (8.34)$$

Из (8.33), (8.29) и (8.34) получаем

$$\frac{\partial f(x_0 + \gamma g)}{\partial g} > \frac{\partial f(x_0)}{\partial g} = \frac{\partial_e f(x_0)}{\partial g}. \quad (8.35)$$

Как и при выводе (8.31), при $\alpha > \gamma$ получаем

$$f(x_0 + \alpha g) - f(x_0 + \gamma g) \geq (\alpha - \gamma) \frac{\partial f(x_0 + \gamma g)}{\partial g}. \quad (8.36)$$

Из (8.32) и (8.36)

$$\begin{aligned} \frac{\partial_e f(x_0)}{\partial g} &= \lim_{\alpha \rightarrow +\infty} \alpha^{-1} [f(x_0 + \alpha g) - f(x_0) + \varepsilon] \geq \\ &\geq \lim_{\alpha \rightarrow +\infty} \alpha^{-1} \left[f(x_0 + \gamma g) + (\alpha - \gamma) \frac{\partial f(x_0 + \gamma g)}{\partial g} - f(x_0) + \varepsilon \right] = \\ &= \frac{\partial f(x_0 + \gamma g)}{\partial g}, \end{aligned}$$

что противоречит (8.35). Полученное противоречие и доказывает достаточность. Лемма доказана.

Следствие. Если при некоторых $x_0, g \in E_n$ и $\varepsilon > 0$

$$\frac{\partial f(x_0)}{\partial g} = \frac{\partial_\varepsilon f(x_0)}{\partial g} < 0,$$

то

$$\inf_{x \in E_n} f(x) = -\infty.$$

Замечание 2. Если нашлось хотя бы одно $\varepsilon > 0$, при котором выполнено (8.29), то (8.29) справедливо при всяком $\varepsilon > 0$.

Лемма 8.5. Для любых $x_0, g \in E_n$ и $\varepsilon > 0$ найдется такое $\beta > 0$, зависящее от x_0, g и ε , что

$$f(x_0 + \alpha g) \leq f(x_0) + \alpha \frac{\partial_\varepsilon f(x_0)}{\partial g} \quad \forall \alpha \in [0, \beta]. \quad (8.37)$$

Доказательство. Если $\frac{\partial_\varepsilon f(x_0)}{\partial g} = \frac{\partial f(x_0)}{\partial g}$, то по лемме 8.4 $\beta = +\infty$. Поскольку

$$\frac{\partial_\varepsilon f(x_0)}{\partial g} \geq \frac{\partial f(x_0)}{\partial g},$$

то остается рассмотреть случай, когда

$$\frac{\partial_\varepsilon f(x_0)}{\partial g} = \frac{\partial f(x_0)}{\partial g} + \delta, \quad (8.38)$$

где

$$\delta > 0. \quad (8.39)$$

Допустим, что утверждение леммы не имеет места. Покажем, что тогда

$$f(x_0 + \alpha g) > f(x_0) + \alpha \frac{\partial_\varepsilon f(x_0)}{\partial g} \quad \forall \alpha > 0. \quad (8.40)$$

Действительно, в противном случае нашлось бы такое $\beta_0 > 0$, что

$$f(x_0 + \beta_0 g) \leq f(x_0) + \beta_0 \frac{\partial_\varepsilon f(x_0)}{\partial g}. \quad (8.41)$$

Так как $f(x)$ — выпуклая функция, то для $\gamma \in [0, 1]$

$$\begin{aligned} f(x_0 + \gamma \beta_0 g) &= f(\gamma(x_0 + \beta_0 g) + (1 - \gamma)x_0) \leq \\ &\leq \gamma f(x_0 + \beta_0 g) + (1 - \gamma)f(x_0) \end{aligned}$$

Из (8.41)

$$f(x_0 + \gamma \beta_0 g) \leq f(x_0) + \gamma \beta_0 \frac{\partial_\varepsilon f(x_0)}{\partial g} \quad \forall \gamma \in [0, 1],$$

т. е. нашлось $\beta_0 > 0$ такое, что

$$f(x_0 + \alpha g) \leq f(x_0) + \alpha \frac{\partial_\varepsilon f(x_0)}{\partial g} \quad \forall \alpha \in [0, \beta_0],$$

что противоречит предположению. Итак, если утверждение леммы не имеет места, то должно быть выполнено (8.40). Из (8.40) и (8.38)

$$\alpha^{-1} [f(x_0 + \alpha g) - f(x_0)] > \frac{\partial f(x_0)}{\partial g} + \delta \quad \forall \alpha > 0. \quad (8.42)$$

Так как (см. следствие из теоремы 4.2)

$$\frac{\partial f(x_0)}{\partial g} = \inf_{\alpha > 0} \alpha^{-1} [f(x_0 + \alpha g) - f(x_0)],$$

то из (8.42) $\delta \leq 0$, что противоречит (8.39). Лемма доказана.

Геометрически лемма 8.5 означает, что прямая

$$\mathcal{L} = \left\{ (\alpha, z) \in E_2 \mid z = f(x_0) + \alpha \frac{\partial_\varepsilon f(x_0)}{\partial g} \right\}$$

является либо секущей графика функции $h(\alpha) = f(x_0 + \alpha g)$, либо совпадает с ним при $\alpha \geq 0$.

Лемма 8.6. Пусть $x_0, g \in E_n$, $\varepsilon > 0$. Положим

$$G = \left\{ \alpha \geq 0 \mid f(x_0 + \alpha g) = f(x_0) + \alpha \frac{\partial_\varepsilon f(x_0)}{\partial g} - \varepsilon \right\},$$

$$\bar{\lambda} = \begin{cases} \varepsilon \left[\frac{\partial_\varepsilon f(x_0)}{\partial g} - \frac{\partial f(x_0)}{\partial g} \right]^{-1}, & \text{если } G \text{ ограничено,} \\ +\infty, & \text{если } G \text{ пусто или} \\ & \text{неограничено.} \end{cases}$$

Тогда

$$f(x_0 + \alpha g) \leq f(x_0) + \alpha \frac{\partial_\varepsilon f(x_0)}{\partial g} \quad \forall \alpha \in [0, \bar{\lambda}]. \quad (8.43)$$

Доказательство. Если $G \neq \emptyset$ и $\gamma \in G$, то, в силу выпуклости функции $f(x)$, при $\alpha \in [0, \gamma]$ выполняется (8.43). Поэтому, если G неограничено, то (8.43) справедливо при любых $\lambda \geq 0$. В случае же,

когда множество G ограничено, для любых $\gamma \in G$ получим

$$f(x_0) + \gamma \frac{\partial_e f(x_0)}{\partial g} - \varepsilon = f(x_0 + \gamma g) \geq f(x_0) + \gamma \frac{\partial f(x_0)}{\partial g}.$$

Отсюда вытекает

$$\gamma \geq \varepsilon \left[\frac{\partial_e f(x_0)}{\partial g} - \frac{\partial f(x_0)}{\partial g} \right]^{-1} = \bar{\lambda}.$$

Обратимся теперь к случаю, когда множество $G = \emptyset$. Прежде всего заметим, что тогда

$$\begin{aligned} \frac{\partial_e f(x_0)}{\partial g} &= \inf_{\alpha \geq 0} \alpha^{-1} [f(x_0 + \alpha g) - f(x_0) + \varepsilon] = \\ &= \lim_{\alpha \rightarrow \infty} \alpha^{-1} [f(x_0 + \alpha g) - f(x_0) + \varepsilon]. \end{aligned} \quad (8.44)$$

Допустим, что в рассматриваемом случае утверждение леммы неверно, т. е. для некоторого $\gamma > 0$ оказалось

$$\frac{\partial_e f(x_0)}{\partial g} < \gamma^{-1} [f(x_0 + \gamma g) - f(x_0)]. \quad (8.45)$$

Из (8.34) тогда

$$\frac{\partial_e f(x_0)}{\partial g} < \frac{\partial f(x_0 + \gamma g)}{\partial g}. \quad (8.46)$$

Из (8.36)

$$f(x_0 + \alpha g) - f(x_0 + \gamma g) \geq (\alpha - \gamma) \frac{\partial f(x_0 + \gamma g)}{\partial g} \quad \forall \alpha \geq \gamma.$$

Отсюда и из (8.44) и (8.45)

$$\begin{aligned} \frac{\partial_e f(x_0)}{\partial g} &= \lim_{\alpha \rightarrow +\infty} \alpha^{-1} [f(x_0 + \alpha g) - f(x_0) + \varepsilon] \geq \\ &\geq \lim_{\alpha \rightarrow +\infty} \alpha^{-1} \left[f(x_0 + \gamma g) + (\alpha - \gamma) \frac{\partial f(x_0 + \gamma g)}{\partial g} - f(x_0) + \varepsilon \right] \geq \\ &\geq \lim_{\alpha \rightarrow +\infty} \alpha^{-1} \left[\alpha \frac{\partial f(x_0 + \gamma g)}{\partial g} + \gamma \left(\frac{\partial_e f(x_0)}{\partial g} - \frac{\partial f(x_0 + \gamma g)}{\partial g} \right) + \varepsilon \right] = \\ &= \frac{\partial f(x_0 + \gamma g)}{\partial g}, \end{aligned}$$

что противоречит (8.46). Лемма доказана.

§ 9. Некоторые свойства и неравенства для выпуклых функций

1. Пусть $f(x)$ — выпуклая конечная на E_n функция, т. е.

$$f(\alpha x_1 + (1 - \alpha)x_2) \leq \alpha f(x_1) + (1 - \alpha)f(x_2) \quad (9.1)$$

$$\forall x_1, x_2 \in E_n, \quad \forall \alpha \in [0, 1].$$

В силу леммы 4.1, для любых $x_1, x_2 \in E_n$ и $\alpha \notin [0, 1]$ будет

$$f(\alpha x_1 + (1 - \alpha)x_2) \geq \alpha f(x_1) + (1 - \alpha)f(x_2). \quad (9.2)$$

Определение. Пусть $x_0 \in E_n$. Множество

$$\mathcal{D}(x_0) = \{x \in E_n \mid f(x) \leq f(x_0)\} \quad (9.3)$$

будем называть *лебеговым множеством* функции $f(x)$ в точке x_0 .

Очевидно, что это множество непусто, ибо $x_0 \in \mathcal{D}(x_0)$, выпукло и замкнуто.

Возьмем любое $v \in \partial f(x_0)$. Тогда по определению

$$f(x) - f(x_0) \geq (v, x - x_0) \quad \forall x \in E_n.$$

Для $x \in \mathcal{D}(x_0)$ будет $(v, x - x_0) \leq 0$, т. е. гиперплоскость $P(x_0) = \{x \in E_n \mid (x - x_0, v) = 0\}$ — опорная к множеству $\mathcal{D}(x_0)$ в точке x_0 . Положим

$$\mathcal{M}^* = \{x \in E_n \mid f(x) \leq f(z) \quad \forall z \in E_n\}. \quad (9.4)$$

Множество $\mathcal{M}^*$ является множеством точек минимума функции $f(x)$ на E_n . Это множество может быть и пустым (см. пример 1 ниже).

Лемма 9.1. Если множество $\mathcal{M}^*$ непусто и ограничено, то для любого $x_0 \in E_n$ лебегово множество $\mathcal{D}(x_0)$ (см. (9.3)) является ограниченным.

Доказательство. Выше уже отмечалось, что множество $\mathcal{D}(x_0)$ непустое, замкнутое и выпуклое.

Покажем, что (рис. 11) для любого $\varepsilon > 0$

$$\psi_\varepsilon = \min_{x^* \in \mathcal{M}^*} \psi_\varepsilon(x^*) = \min_{x^* \in \mathcal{M}^*} \min_{\|g\|=1} \frac{\partial_\varepsilon f(x^*)}{\partial g} > 0. \quad (9.5)$$

Действительно, допустим противное. Тогда для некоторых $\varepsilon > 0$, $x^* \in \mathcal{M}^*$ и $g \in E_n$, $\|g\|=1$, будет

$$\frac{\partial_\varepsilon f(x^*)}{\partial g} \leq 0. \quad (9.6)$$

Так как $\partial f(x^*) \subset \partial_\varepsilon f(x^*)$ и $x^* \in \mathcal{M}^*$, то

$$\frac{\partial_\varepsilon f(x^*)}{\partial g} \geq \frac{\partial f(x^*)}{\partial g} \geq 0. \quad (9.7)$$

Из (9.6) и (9.7) заключаем, что $\frac{\partial_\varepsilon f(x^*)}{\partial g} = \frac{\partial f(x^*)}{\partial g} = 0$.
Отсюда, в силу леммы 8.4,

$$f(x^* + \alpha g) = f(x^*) + \alpha \frac{\partial f(x^*)}{\partial g} = f(x^*) \quad \forall \alpha \geq 0,$$

что противоречит ограниченности множества $\mathcal{M}^*$.

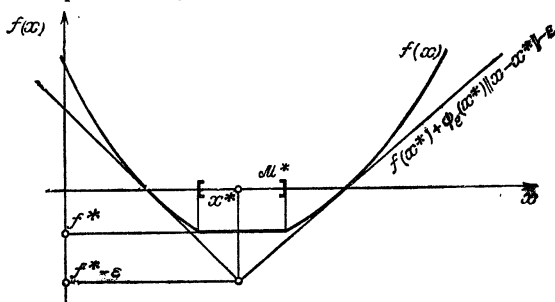


Рис. 11

Для любых $x \in E_n$, $x^* \in \mathcal{M}^*$ и $\varepsilon > 0$ имеем

$$f(x) - f(x^*) \geq \max_{v \in \partial_\varepsilon f(x^*)} (v, x - x^*) \geq \psi_\varepsilon(x^*) \|x - x^*\| - \varepsilon. \quad (9.8)$$

Пусть $\varepsilon = f(x_0) - f^*$. Тогда из (9.5) и (9.8) получаем, что

$$\mathcal{D}(x_0) \subset S_{r_0}(\mathcal{M}^*),$$

где $r_0 = 2\varepsilon\psi_\varepsilon^{-1}$. Лемма доказана.

Следствие 1. Для того чтобы множество точек минимума выпуклой функции было ограничено, необходимо, чтобы для любого $x_0 \in E_n$ лебегово множество $\mathcal{D}(x_0)$ было ограничено, и достаточно, чтобы хотя бы для одного $x_0 \in E_n$ это множество было ограничено.

Необходимость только что доказана, а достаточность очевидна.

Положим

$$\mathcal{D}_\varepsilon = \{x \in E_n \mid f(x) \leq f^* + \varepsilon\}.$$

В силу леммы 8.1 $\mathcal{D}_\varepsilon$ есть множество ε -стационарных точек функции $f(x)$ на E_n .

Следствие 2. Если множество $\mathcal{M}^*$ непусто и ограничено, то множество $\mathcal{D}_\varepsilon$ ограничено для любого фиксированного $\varepsilon > 0$.

Положим $a_\varepsilon = \max_{x \in \mathcal{D}_\varepsilon} \rho(x, \mathcal{M}^*)$, где $\rho(x, \mathcal{M}) = \min_{v \in \mathcal{M}} \|x - v\|$.

Следствие 3. Если множество $\mathcal{M}^*$ непусто и ограничено, то $a_\varepsilon \xrightarrow{\varepsilon \rightarrow +0} 0$.

Действительно, допустим противное, пусть найдутся число $b > 0$ и последовательность $\{x_k\}$ такие, что $x_k \in \mathcal{D}_{\varepsilon_k}$, $\varepsilon_k \rightarrow +0$, $\rho(x_k, \mathcal{M}^*) \geq b$. Поскольку при $\varepsilon \leq \varepsilon_0$ множества $\mathcal{D}_\varepsilon$ ограничены, то, не умаляя общности, можем считать, что $x_k \rightarrow x^*$. При этом

$$\rho(x^*, \mathcal{M}^*) \geq b. \quad (9.9)$$

С другой стороны, $x_k \in \mathcal{D}_{\varepsilon_k}$, т. е. $f(x_k) \leq f^* + \varepsilon_k$. Переходя здесь к пределу, получим $f(x^*) = f^*$, что противоречит (9.9). ■

Следствие 4. Если множество $\mathcal{M}^*$ непусто и ограничено, то для любого $a > 0$ найдется такое $\varepsilon = \varepsilon(a) > 0$, что $\mathcal{D}_\varepsilon \subset S_a(\mathcal{M}^*) = \{x \in E_n \mid \rho(x, \mathcal{M}^*) \leq a\}$.

Лемма 9.2. Если множество $\mathcal{M}^*$ непусто и ограничено, то для любых $\varepsilon \geq 0$ и $r > 0$ существует $c_{\varepsilon r} > 0$ такое, что

$$\|v\| \geq c_{\varepsilon r} > 0 \quad \forall v \in \partial_\varepsilon f(x), \quad \forall x \notin S_r(\mathcal{D}_\varepsilon). \quad (9.10)$$

Доказательство. Положим

$$G_{\varepsilon r} = \{x \in E_n \mid \rho(x, \mathcal{D}_\varepsilon) = r\}.$$

По следствию 2 из леммы 9.1 множество $\mathcal{D}_\varepsilon$ ограничено. Тогда ограничено и множество $G_{\varepsilon r}$. Так как $\mathcal{D}_\varepsilon \cap G_{\varepsilon r} = \emptyset$, то

$$f_{\varepsilon r} = \min_{x \in G_{\varepsilon r}} f(x) > f^* + \varepsilon. \quad (9.11)$$

Зафиксируем $x^* \in \mathcal{M}^*$ и возьмем любое $x \notin S_r(\mathcal{D}_\varepsilon)$. Пусть $x_1 = x_1(x) \in G_{\varepsilon r} \cap \text{co}\{x^*, x\}$. По лемме 4.1

$$\begin{aligned} f(x) - f(x^*) &\geq \|x_1 - x^*\|^{-1} \|x - x^*\| (f(x_1) - f(x^*)) \geq \\ &\geq \frac{\|x - x^*\|}{\|x_1 - x^*\|} (f_{\varepsilon r} - f^*). \end{aligned} \quad (9.12)$$

Так как $v \in \partial_e f(x)$, то $f(x^*) - f(x) \geq (v, x^* - x) - \varepsilon \geq -\|v\| \|x^* - x\| - \varepsilon$. Отсюда и из (9.12) $\|v\| \geq \|x_1 - x^*\|^{-1} \times (f_{er} - f^*) - \|x - x^*\|^{-1} \varepsilon$. Поскольку $\|x - x^*\| \geq \|x_1 - x^*\|$, то $\|v\| \geq \|x_1 - x^*\|^{-1} (f_{er} - f^* - \varepsilon)$. Но $\|x_1 - x^*\| \leq d \equiv \text{diam } S_r(\mathcal{D}_\varepsilon)$. Поэтому, учитывая (9.11), имеем

$$\|v\| \geq d^{-1} (f_{er} - f^* - \varepsilon) \equiv c_{er} > 0$$

$$\forall x \notin S_r(\mathcal{D}_\varepsilon), \quad \forall v \in \partial_e f(x). \quad \blacksquare \quad (9.13)$$

2. Если $f(x)$ — строго (или сильно) выпуклая функция на замкнутом выпуклом множестве $\Omega \subset E_n$, то из неравенства

$$f\left(\frac{x_1 + x_2}{2}\right) < \frac{1}{2} [f(x_1) + f(x_2)] \quad \forall x_1, x_2 \in \Omega, \quad x_1 \neq x_2, \quad (9.14)$$

заключаем, что функция $f(x)$ достигает своего наименьшего на Ω значения не более чем в одной точке.

Лемма 9.3. Если $f(x)$ — сильно выпуклая на Ω функция, то она достигает своего наименьшего значения на множестве Ω (причем в единственной точке).

Доказательство. Единственность отмечена выше. Выберем любую точку $x_0 \in \Omega$. Из (8.14) вытекает, что для любого $v \in \partial f(x_0)$ будет

$$\mathcal{D}(x_0) \equiv \{x \in \Omega \mid f(x) \leq f(x_0)\} \subset$$

$$\subset \{x \in \Omega \mid 0 \geq -\|v\| \cdot \|x - x_0\| + m\|x - x_0\|^2\}.$$

Отсюда заключаем, что $\mathcal{D}(x_0)$ — компакт. Так как функция $f(x)$ непрерывна, то она достигает своей нижней грани на $\mathcal{D}(x_0)$. $\blacksquare$

Замечание 1. Для строго выпуклых функций инфимум может и не достигаться. Так, функция $f(x) = e^{-x}$, заданная на E_1 , является строго выпуклой и $\inf_{x \in E_n} f(x) = 0$, однако этот инфимум не достигается.

3. Получим некоторые оценки для сильно выпуклых функций. Пусть $f(x)$ — сильно выпуклая на E_n функция с константой сильной выпуклости $m > 0$, т. е.

$$f(\alpha x_1 + (1 - \alpha)x_2) \leq$$

$$\leq \alpha f(x_1) + (1 - \alpha)f(x_2) - \alpha(1 - \alpha)m\|x_1 - x_2\|^2 \quad (9.15)$$

$$\forall x_1, x_2 \in E_n, \quad \forall \alpha \in [0, 1].$$

Лемма 9.4. Если $f(x)$ — сильно выпуклая на E_n функция с константой сильной выпуклости $m > 0$, то для произвольных

$$x, x_0 \in E_n, \quad v(x_0) \in \partial_\varepsilon f(x_0), \quad \varepsilon \geq 0, \quad \alpha \in (0, 1]$$

справедливо неравенство

$$\begin{aligned} f(x) - f(x_0) &\geq \\ &\geq (v(x_0), x - x_0) - \varepsilon \alpha^{-1} + (1 - \alpha) m \|x_0 - x\|^2. \end{aligned} \quad (9.16)$$

Доказательство. Пусть $x_0 \in E_n$ и $\alpha \in (0, 1]$ фиксированы. Возьмем любое $v(x_0) \in \partial_\varepsilon f(x_0)$. По определению ε -субградиента

$$f(z) \geq f(x_0) + (v(x_0), z - x_0) - \varepsilon \quad \forall z \in E_n.$$

Положив $z = \alpha x + (1 - \alpha)x_0$, отсюда и из (9.15) имеем

$$\begin{aligned} \alpha f(x) + (1 - \alpha) f(x_0) - \alpha(1 - \alpha) m \|x_0 - x\|^2 &\geq \\ &\geq f(\alpha x + (1 - \alpha)x_0) = f(z) = f(x_0 + \alpha(x - x_0)) \geq \\ &\geq f(x_0) + \alpha(v(x_0), x - x_0) - \varepsilon. \end{aligned}$$

Неравенство (9.16) теперь очевидно. ■

Следствие 1. Для любых $x_0 \in E_n$ и $v(x_0) \in \partial f(x_0)$ справедливо неравенство

$$f(x) - f(x_0) \geq (v(x_0), x - x_0) + m \|x - x_0\|^2 \quad \forall x \in E_n. \quad (9.17)$$

Справедливость (9.17) следует из (9.16), если положить там $\varepsilon = 0$ и перейти к пределу при $\alpha \rightarrow +0$.

Следствие 2. Для любых $x_1, x_2 \in E_n$, $v_1 \in \partial f(x_1)$, $v_2 \in \partial f(x_2)$ справедливо неравенство

$$(v_1 - v_2, x_1 - x_2) \geq 2m \|x_1 - x_2\|^2. \quad (9.18)$$

Доказательство. Для любых $x_1, x_2 \in E_n$ в силу (9.17)

$$f(x) - f(x_1) \geq (v_1, x - x_1) + m \|x - x_1\|^2 \quad \forall x \in E_n,$$

$$f(x) - f(x_2) \geq (v_2, x - x_2) + m \|x - x_2\|^2 \quad \forall x \in E_n,$$

где $v_1 \in \partial f(x_1)$, $v_2 \in \partial f(x_2)$.

Подставляя в первое неравенство $x = x_2$, а во второе $x = x_1$ и складывая получившиеся неравенства, получим (9.18). ■

Из (9.18) ясно, что

$$\|v_1 - v_2\| \geq 2m \|x_1 - x_2\|.$$

Отсюда, в частности, вытекает, что если $f(x)$ — сильно выпуклая функция, то любой вектор $v \in E_n$ может быть ее субградиентом не более чем для одной точки.

Задача 9.1. Доказать, что этот результат справедлив и для строго выпуклых функций, т. е. если $f(x)$ — строго выпуклая функция, то любой вектор может быть ее субградиентом не более чем для одной точки.

Следствие 3. Пусть $x \neq x_0$. Максимум по α выражения в правой части (9.16) достигается при $\alpha = \sqrt{em}^{-1} \|x - x_0\|^{-1}$, и тогда, если $\|x - x_0\| \geq \sqrt{em}^{-1}$, то из (9.16)

$$\begin{aligned} f(x) - f(x_0) &\geq \\ &\geq (v(x_0), x - x_0) + m \|x - x_0\|^2 - 2\sqrt{em} \|x - x_0\|^2. \end{aligned}$$

Таким образом, для сильно выпуклой функции выполнено (9.17). Оказывается, справедливо и обратное утверждение.

Лемма 9.5 Пусть на E_n задана функция $f(x)$. Если для любого $x_0 \in E_n$ найдется $v(x_0) \in E_n$ такое, что имеет место (9.17), то функция $f(x)$ является сильно выпуклой на E_n с константой сильной выпуклости m .

Доказательство. Возьмем произвольные $x_1, x_2 \in E_n$ и $\alpha \in [0, 1]$. Положим $x_0 = \alpha x_1 + (1 - \alpha)x_2$. По предположению найдется $v(x_0) \in E_n$, для которого выполнено (9.17). При $x = x_1$ из (9.17) следует

$$f(x_1) \geq f(x_0) + (1 - \alpha)(v(x_0), x_1 - x_2) + m(1 - \alpha)^2 \|x_1 - x_2\|^2. \quad (9.19)$$

Полагая теперь $x = x_2$, аналогично имеем

$$f(x_2) \geq f(x_0) - \alpha(v(x_0), x_1 - x_2) + m\alpha^2 \|x_1 - x_2\|^2. \quad (9.20)$$

Умножая (9.19) на α , а (9.20) на $(1 - \alpha)$ и складывая, получим

$$\begin{aligned} f(x_0) = f(\alpha x_1 + (1 - \alpha)x_2) &\leq \\ &\leq \alpha f(x_1) + (1 - \alpha)f(x_2) - \alpha(1 - \alpha)m \|x_1 - x_2\|^2. \end{aligned}$$

Поскольку здесь $x_1, x_2 \in E_n$ и $\alpha \in [0, 1]$ произвольные, то, сравнивая с (9.15), заключаем, что $f(x)$ — сильно выпуклая функция с константой сильной выпуклости m . Лемма доказана.

Замечание 2. Если для некоторых $x_0 \in E_n$ и $v(x_0) \in E_n$ имеет место (9.17), то ясно, что $v(x_0) \in \partial f(x_0)$.

4. Пусть $f(x)$ — сильно выпуклая функция с константой сильной выпуклости $m > 0$. Положим $f^* = \inf_{x \in E_n} f(x)$.

Как установлено в лемме 9.3, значение f^* достигается (причем в единственной точке), т. е. существует точка $x^* \in E_n$ такая, что $f(x^*) = f^*$. Возьмем любые $x_0 \in E_n$, $\varepsilon > 0$ и положим

$$d_\varepsilon = d_\varepsilon(x_0) \equiv \min_{v \in \partial_\varepsilon f(x_0)} \|v\| = \|v_\varepsilon(x_0)\|.$$

Лемма 9.6. Справедливы неравенства

$$0 \leq f(x_0) - f^* \leq \varepsilon + (d_\varepsilon^2 + 4d_\varepsilon \sqrt{\varepsilon m})(4m)^{-1}, \quad (9.21)$$

$$\|x_0 - x^*\| \leq (d_\varepsilon + 2\sqrt{\varepsilon m})(2m)^{-1}. \quad (9.22)$$

Доказательство. Установим вначале справедливость неравенства (9.21). При $d_\varepsilon = 0$ оно очевидно из леммы 8.1. Пусть $d_\varepsilon > 0$. Возьмем $v \in \partial_\varepsilon f(x_0)$. Из (9.16) для всех $x \in E_n$ и $\beta \in (0, 1]$ будет

$$\begin{aligned} f(x) &\geq f(x_0) + \\ &+ \max_{v \in \partial_\varepsilon f(x_0)} (v, x - x_0) - \varepsilon\beta^{-1} + (1 - \beta)m\|x_0 - x\|^2 \geq \\ &\geq \|x - x_0\| \min_{\|g\|=1} \max_{v \in \partial_\varepsilon f(x_0)} (v, g) - \varepsilon\beta^{-1} + (1 - \beta)m\|x - x_0\|^2. \end{aligned}$$

По следствию 2 из леммы 6.4

$$d_\varepsilon = - \min_{\|g\|=1} \max_{v \in \partial_\varepsilon f(x_0)} (v, g).$$

Поэтому

$$\begin{aligned} f(x) - f(x_0) &\geq -d_\varepsilon\|x - x_0\| + (1 - \beta)m\|x - x_0\|^2 - \varepsilon\beta^{-1} \\ &\quad \forall x \in E_n, \quad \forall \beta \in (0, 1]. \end{aligned}$$

Отсюда

$$\begin{aligned} f^* &= \inf_{x \in E_n} f(x) \geqslant \\ &\geqslant f(x_0) + \inf_{x \in E_n} [(1 - \beta)m \|x - x_0\|^2 - d_e \|x - x_0\|] - \varepsilon \beta^{-1} = \\ &= f(x_0) + \inf_{\alpha \geqslant 0} [(1 - \beta)m\alpha^2 - d_e \alpha] - \varepsilon \beta^{-1} \quad \forall \beta \in (0, 1]. \end{aligned} \quad (9.23)$$

Инфимум здесь достигается при $\alpha = 2^{-1}m^{-1}(1 - \beta)^{-1}d_e$ и равен $-d_e^2[4m(1 - \beta)]^{-1}$. Отсюда и из (9.23)

$$0 \leqslant f(x_0) - f^* \leqslant \varepsilon \beta^{-1} + d_e^2[4m(1 - \beta)]^{-1}. \quad (9.24)$$

Это неравенство справедливо при всех $\beta \in (0, 1]$. Найдем минимум функции

$$h(\beta) = \varepsilon \beta^{-1} + d_e^2[4m(1 - \beta)]^{-1}.$$

Так как $h'(\beta) = [d_e^2\beta^2 - 4m\varepsilon(1 - \beta)^2][4m\beta^2(1 - \beta)^2]^{-1}$, то $h'(\beta) = 0$ при

$$\beta_1 = 2\sqrt{m\varepsilon}(d_e + 2\sqrt{m\varepsilon})^{-1}, \quad \beta_2 = 2\sqrt{m\varepsilon}(2\sqrt{m\varepsilon} - d_e)^{-1}.$$

Поскольку $\beta_2 \notin (0, 1)$, то в $(0, 1)$ существует единственная точка β_1 , в которой $h(\beta)$ достигает экстремума. Нетрудно проверить, что $h''(\beta_1) > 0$, т. е. β_1 — точка минимума функции $h(\beta)$. При этом, как легко подсчитать, $h(\beta_1) = (4m)^{-1}(d_e + 2\sqrt{m\varepsilon})^2 = \varepsilon + (4m)^{-1}(d_e^2 + 4d_e\sqrt{m\varepsilon})$.

Отсюда и из (9.24) получаем (9.21).

Осталось доказать (9.22). Из (9.17) для точки минимума x^*

$$f(x_0) - f(x^*) \geqslant (v, x - x^*) + m\|x_0 - x^*\|^2 \quad \forall v \in \partial f(x^*).$$

Поскольку x^* — точка минимума $f(x)$ на E_n , то по необходимому и достаточному условию минимума $0 \in \partial f(x^*)$. Поэтому

$$f(x_0) - f(x^*) \geqslant m\|x_0 - x^*\|^2.$$

Отсюда и из (9.21) сразу получаем (9.22). Лемма доказана.

Следствие. При $\varepsilon = 0$ имеем

$$0 \leq f(x_0) - f^* \leq (4m)^{-1} d^2, \quad \|x_0 - x^*\| \leq (2m)^{-1} d,$$

где $d = d_0 = d_0(x_0) = \min_{v \in \partial f(x_0)} \|v\|$.

При $d_\varepsilon = 0$ (т. е. если x_0 — ε -стационарная точка) из (9.21) и (9.22) получаем

$$0 \leq f(x_0) - f^* \leq \varepsilon, \quad \|x_0 - x^*\| \leq \sqrt{\varepsilon m^{-1}}.$$

Таким образом, для сильно выпуклых функций можно оценить, насколько значение функции в точке x_0 отличается от минимального (неравенство (9.21)), а также расстояние от точки x_0 до точки минимума (неравенство (9.22)).

5. Пусть $\Omega \subset E_n$ — выпуклый компакт, $f(x)$ — выпуклая на E_n функция.

Лемма 9.7. Если функция $f(x)$ не является постоянной на Ω , то она достигает наибольшего на Ω значения только на границе множества Ω .

Доказательство. Предположим противное, пусть $x_0 \in \text{int } \Omega$,

$$f(x_0) \geq f(x) \quad \forall x \in \Omega, \quad (9.25)$$

т. е. x_0 — точка максимума $f(x)$ на Ω . Так как $x_0 \in \text{int } \Omega$, то существует $r > 0$ такое, что $S_r(x_0) \subset \Omega$. Возьмем любое $\bar{x} \in \Omega$, $\bar{x} \neq x_0$. Положим

$$\tilde{x} = x_0 - r(\bar{x} - x_0) \|\bar{x} - x_0\|^{-1}. \quad (9.26)$$

Так как $\tilde{x} \in S_r(x_0)$, то $\tilde{x} \in \Omega$, и поэтому из (9.25)

$$f(\tilde{x}) \leq f(x_0). \quad (9.27)$$

С другой стороны, для любого $v \in \partial f(x_0)$

$$f(\tilde{x}) \geq f(x_0) + (v, \tilde{x} - x_0). \quad (9.28)$$

Поэтому из (9.27) и (9.28) $(v, \tilde{x} - x_0) \leq 0$. Отсюда и из (9.26)

$$(v, \bar{x} - x_0) \geq 0. \quad (9.29)$$

Так как по определению субградиента $f(\bar{x}) \geq f(x_0) + (v, \bar{x} - x_0)$, то из (9.29) $f(\bar{x}) \geq f(x_0)$. Отсюда и из (9.25) следует

$$f(\bar{x}) = f(x_0). \quad (9.30)$$

Поскольку $\bar{x}$ — произвольная точка Ω , то из (9.30) получаем

$$f(x) = f(x_0) \quad \forall x \in \Omega,$$

что противоречит условию леммы. Лемма доказана.

З а м е ч а н и е 3. Из доказательства леммы очевидно, что если множество Ω не является ограниченным и функция $f(x)$ отлична от постоянной, то ее наибольшее значение на Ω может достигаться разве лишь на границе множества Ω (но может и не достигаться, если функция $f(x)$ не является ограниченной на Ω).

Лемма 9.8. Если функция $f(x)$ сильно выпукла на E_n , то для любого $v \in E_n$ найдется единственная точка $x \in E_n$ такая, что $v \in \partial f(x)$.

Доказательство. Возьмем произвольное $v_0 \in E_n$. Так как функция $f(x)$ сильно выпукла на E_n , то и функция $h(x) = f(x) - (v_0, x)$ тоже сильно выпукла, как сумма сильно выпуклой и линейной функций. По лемме 9.3 существует единственная точка $x_0 \in E_n$, в которой функция $h(x)$ достигает своего наименьшего на E_n значения. По условию минимума функции $h(x)$ на E_n будет $0 \in \partial h(x_0)$. Но $\partial h(x_0) = \partial f(x_0) - \{v_0\}$ (так как $h(x)$ есть сумма двух выпуклых функций, то по лемме 5.2 $\partial h(x_0) = \partial f(x_0) + \partial F(x_0)$, где $F(x_0) = -(v_0, x)$). Отсюда $v_0 \in \partial f(x_0)$. Таким образом, для любого $v_0 \in E_n$ нашлась точка $x_0 \in E_n$ такая, что $v_0 \in \partial f(x_0)$. Единственность точки x_0 вытекает из следствия 2 леммы 9.4. ■

З а м е ч а н и е 4. Условие сильной выпуклости функции $f(x)$ в лемме 9.8 существенно, что видно из следующего примера.

Пример 1. Пусть $x \in E_1$, $f(x) = e^x$. Функция $f(x)$ строго выпукла и дифференцируема. Поэтому

$$\partial f(x) = \{f'(x)\} \quad \forall x \in E_1.$$

Очевидно, что не существует точки $x_0 \in E_1$ такой, что $f'(x_0) = e^{x_0} = 0$, т. е. для $v_0 = 0$ нельзя найти соответствующей точки x_0 .

З а д а н и е 9.2. Пусть $f(x)$ — строго выпуклая на E_n функция. Доказать, что верна следующая

Лемма 9.9. Если $f(x)$ строго выпукла на E_n , то для любых $x_0 \in E_n$ и $v(x_0) \in \partial f(x_0)$ справедливо неравенство

$$f(x) - f(x_0) > (v(x_0), x - x_0) \quad \forall x \in E_n, \quad x \neq x_0. \quad (9.31)$$

Обратно, если для любого $x_0 \in E_n$ найдется $v(x_0) \in E_n$ такое, что имеет место (9.31), то функция $f(x)$ является строго выпуклой.

§ 10. Условный ε -субдифференциал

1. Определение и свойства условного субдифференциала и условного ε -субдифференциала. Пусть $\Omega \subset E_n$ — выпуклое множество. Через $\Gamma(x)$ обозначим конус возможных направлений множества Ω в точке $x \in \Omega$, являющийся замыканием множества

$$\gamma(x) = \{v = \lambda(z - x) \mid \lambda > 0, z \in \Omega\}.$$

Пусть на множестве Ω определена выпуклая конечная функция $f(x)$. Положим для фиксированных $x_0 \in \Omega$ и $\varepsilon \geq 0$

$$\partial^0 f(x_0) = \{v \in E_n \mid f(z) - f(x_0) \geq (v, z - x_0) \quad \forall z \in \Omega\},$$

$$\partial_\varepsilon^0 f(x_0) = \{v \in E_n \mid f(z) - f(x_0) \geq (v, z - x_0) - \varepsilon \quad \forall z \in \Omega\}.$$

Множества $\partial^0 f(x_0)$ и $\partial_\varepsilon^0 f(x_0)$ назовем соответственно *условным субдифференциалом* и *условным ε -субдифференциалом* функции $f(x)$ в точке $x_0 \in \Omega$ по множеству Ω .

Элемент $v \in \partial^0 f(x_0)$ назовем *условным субградиентом*, а элемент $v \in \partial_\varepsilon^0 f(x_0)$ — *условным ε -субградиентом* функции $f(x)$ в точке $x_0 \in \Omega$ по множеству Ω .

Если функция $f(x)$ определена и выпукла на E_n , то множества $\partial^{E_n} f(x_0) \equiv \partial f(x_0)$ и $\partial_\varepsilon^{E_n} f(x_0) \equiv \partial_\varepsilon f(x_0)$, как обычно, называются *субдифференциалом* и *ε -субдифференциалом* функции $f(x)$ в точке x_0 по E_n .

Очевидно, что если $\Omega_1 \supset \Omega_2$, то при $\varepsilon \geq 0$ будет

$$\partial_\varepsilon^{\Omega_1} f(x) \subset \partial_\varepsilon^{\Omega_2} f(x) \quad \forall x \in \Omega_2. \quad (10.1)$$

Отметим еще одно свойство условного ε -субдифференциала: если $\bar{\Omega}$ — замыкание выпуклого множества Ω и $f(x)$ непрерывна на $\bar{\Omega}$, то

$$\partial_\varepsilon^{\bar{\Omega}} f(x) = \partial_\varepsilon^{\Omega} f(x) \quad \forall x \in \Omega, \quad \varepsilon \geq 0.$$

Лемма 10.1. Пусть функция $f(x)$ конечна и выпукла на E_n . Если $x_0 \in \text{int } \Omega$, то для любого фиксированного

$\varepsilon \geq 0$ множество $\partial_\varepsilon^{\Omega} f(x_0)$ непусто, выпукло, замкнуто и ограничено.

Доказательство. Непустота, выпуклость и замкнутость множества $\partial_\varepsilon^{\Omega} f(x_0)$ очевидны. Доказательство ограниченности аналогично доказательству ограниченности $\partial_\varepsilon f(x_0)$ в теореме 7.1. ■

Лемма 10.2. Если функция $f(x)$ конечна и выпукла на E_n , то

$$\partial^{\Omega} f(x_0) = \partial f(x_0) - \Gamma^+(x_0), \quad (10.2)$$

где $\Gamma^+(x_0)$ — конус, сопряженный конусу возможных направлений множества Ω в точке $x_0 \in \Omega$.

Доказательство. Правую часть в (10.2) обозначим $\mathcal{D}$. Пусть $v = v_1 + v_2$, где $v_1 \in \partial f(x_0)$, $-v_2 \in \Gamma^+(x_0)$. Тогда $v \in \mathcal{D}$ и

$$\begin{aligned} f(x) - f(x_0) &\geq (v_1, x - x_0) \quad \forall x \in E_n, \\ (-v_2, x - x_0) &\geq 0 \quad \forall x \in \Omega. \end{aligned}$$

Складывая эти неравенства, получаем

$$f(x) - f(x_0) \geq (v_1 + v_2, x - x_0) \quad \forall x \in \Omega.$$

Значит, $v \in \partial^{\Omega} f(x_0)$, т. е.

$$\mathcal{D} \subset \partial^{\Omega} f(x_0). \quad (10.3)$$

Пусть теперь $v \in \partial^{\Omega} f(x_0)$. Тогда

$$f(x) - f(x_0) \geq (v, x - x_0) \quad \forall x \in \Omega. \quad (10.4)$$

Положим $h(x) = f(x) - f(x_0) - (v, x - x_0)$. Очевидно, что функция $h(x)$ выпукла на Ω , а из (10.4) следует, что $h(x) \geq 0$ для всех $x \in \Omega$, причем $h(x_0) = 0$. Значит, x_0 — точка минимума функции $h(x)$ на множестве Ω . По необходимому и достаточному условию минимума функции $h(x)$ на Ω будет (см. теорему 6.1)

$$\partial h(x_0) \cap \Gamma^+(x_0) \neq \emptyset.$$

Но $\partial h(x_0) = \partial f(x_0) - \{v\}$. Поэтому найдутся $v_0 \in \partial f(x_0)$ и $w_0 \in \Gamma^+(x_0)$ такие, что $v_0 - v = w_0$, т. е. $v = v_0 - w_0$. Отсюда

$$\partial^{\Omega} f(x_0) \subset \mathcal{D}. \quad (10.5)$$

Из (10.3) и (10.5) следует (10.2). Лемма доказана.

Следствие 1. Если $x_0 \in \text{int } \Omega$, а функция $f(x)$ конечна и выпукла на E_n , то

$$\partial^\Omega f(x_0) = \partial f(x_0). \quad (10.6)$$

Действительно, для $x_0 \in \text{int } \Omega$ имеем

$$\Gamma(x_0) = E_n, \quad \Gamma^+(x_0) = \{0\}.$$

Таким образом, *условный субдифференциал функции $f(x)$ во внутренних точках множества Ω совпадает с субдифференциалом функции $f(x)$* . Положим

$$\begin{aligned} \Omega_\delta &= S_\delta(x_0) = \\ &= \{x \mid \|x - x_0\| \leq \delta\}, \quad \delta > 0. \end{aligned}$$

Тогда по доказанному выше

$$\partial f(x_0) = \partial^\Omega f(x_0). \quad \blacksquare$$

Отсюда ясно, что *субдифференциал функции $f(x)$ определяется лишь локальными свойствами функции $f(x)$* .

Замечание 1. Вообще говоря, при $\varepsilon > 0$ может оказаться

$$\partial_\varepsilon^\Omega f(x_0) \neq \partial_\varepsilon f(x_0).$$

Это видно из следующего примера. Пусть $f(x) = (a, x) + b$, где $a, x \in E_n$, $b \in E_1$. В примере 2 § 7 было показано, что в этом случае (рис. 12)

$$\partial_\varepsilon f(x) = \partial f(x) = \{a\} \quad \forall x \in E_n, \quad \forall \varepsilon \geq 0.$$

Пусть $\Omega = \{z \in E_n \mid \|z - x_0\| \leq 1\}$. Покажем, что

$$\partial_\varepsilon^\Omega f(x_0) = \{v = a + r \mid r \in E_n, \|r\| \leq \varepsilon\}. \quad (10.7)$$

По определению $\partial_\varepsilon^\Omega f(x_0)$ для любого $v \in \partial_\varepsilon^\Omega f(x_0)$ будет

$$f(z) - f(x_0) = (a, z - x_0) \geq (v, z - x_0) - \varepsilon \quad \forall z \in \Omega,$$

т. е.

$$(v - a, z - x_0) \leq \varepsilon \quad \forall z \in \Omega, \quad (10.8)$$

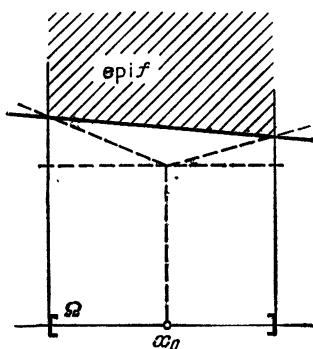


Рис. 12.

Но

$$\max_{z \in \Omega} (v - a, z - x_0) = (v - a, z_0 - x_0),$$

где $z_0 - x_0 = \|v - a\|^{-1}(v - a)$. Отсюда и из (10.8) имеем

$$\max_{z \in \Omega} (v - a, z - x_0) = \|v - a\| \leq \varepsilon \quad \forall v \in \partial_\varepsilon f(x_0).$$

Из последнего неравенства и следует (10.7).

Таким образом, при $\varepsilon > 0$ для линейной функции ε -субдифференциал и условный ε -субдифференциал могут не совпадать.

Следствие 2. Если функция $f(x)$ конечна и выпукла на E_n , то из формулы (10.2) ясно, что

$$\partial^2 f(x_0) = \partial^{\mathcal{K}_\Omega(x_0)} f(x_0),$$

где

$$\mathcal{K}_\Omega(x_0) = \{x = x_0 + v \mid v \in \Gamma(x_0)\}, \quad \Gamma(x_0) = \bar{\gamma}(x_0),$$

$$\gamma(x_0) = \{v = \lambda(x - x_0) \mid \lambda > 0, x \in \Omega\}$$

(т. е. $\mathcal{K}_\Omega(x_0)$ — это конус возможных направлений множества Ω в точке x_0 , перенесенный в точку x_0).

Следствие 3. Если $g \in \Gamma(x_0)$, то

$$\frac{\partial f(x_0)}{\partial g} = \max_{v \in \partial f(x_0)} (v, g) = \sup_{v \in \partial^2 f(x_0)} (v, g) = \max_{v \in \partial^2 f(x_0)} (v, g).$$

Если же $g \notin \Gamma(x_0)$, то $\sup_{v \in \partial^2 f(x_0)} (v, g) = +\infty$.

Действительно, пусть $v \in \partial^2 f(x_0)$, $g \in \Gamma(x_0)$. Тогда по лемме 10.2 $v = v_1 - w$, где $v_1 \in \partial f(x_0)$, $w \in \Gamma^+(x_0)$. Так как $(w, g) \geq 0$ для всех $g \in \Gamma(x_0)$, то

$$(v, g) = (v_1, g) - (w, g) \leq (v_1, g).$$

Отсюда

$$\sup_{v \in \partial^2 f(x_0)} (v, g) \leq \max_{v \in \partial f(x_0)} (v, g).$$

Поскольку $\partial^2 f(x_0) \supset \partial f(x_0)$, то справедливо неравенство

$$\max_{v \in \partial f(x_0)} (v, g) \leq \sup_{v \in \partial^2 f(x_0)} (v, g).$$

Таким образом,

$$\sup_{v \in \partial^2 f(x_0)} (v, g) = \max_{v \in \partial f(x_0)} (v, g).$$

В силу компактности множества $\partial f(x_0)$ супремум в последнем равенстве достигается, а следовательно, достигается и супремум по множеству $\partial^{\Omega} f(x_0)$.

Пусть теперь $g \notin \Gamma(x_0)$. Тогда найдется $w_0 \in \Gamma^+(x_0)$ такое, что $(g, w_0) = -a < 0$. Возьмем любое $v_0 \in \partial f(x_0)$. По лемме 10.2 для произвольного $\lambda > 0$ будет

$$v_{\lambda} = [v_0 - \lambda w_0] \in \partial^{\Omega} f(x_0).$$

Но

$$(v_{\lambda}, g) = (v_0, g) + \lambda a \xrightarrow{\lambda \rightarrow +\infty} +\infty.$$

Следствие 3 доказано полностью.

Следствие 4. Если функции $f_1(x)$ и $f_2(x)$ конечны и выпуклы на E_n , точка $x_0 \in \Omega$, то для функции $f(x) = f_1(x) + f_2(x)$ справедливо соотношение

$$\partial^{\Omega} f(x_0) = \partial^{\Omega} f_1(x_0) + \partial^{\Omega} f_2(x_0). \quad (10.9)$$

Действительно, поскольку (см. лемму 5.2)

$$\partial f(x_0) = \partial f_1(x_0) + \partial f_2(x_0),$$

то получаем

$$\begin{aligned} \partial^{\Omega} f(x_0) &= \partial f(x_0) - \Gamma^+(x_0) = \partial f_1(x_0) + \partial f_2(x_0) - \Gamma^+(x_0) = \\ &= [\partial f_1(x_0) - \Gamma^+(x_0)] + [\partial f_2(x_0) - \Gamma^+(x_0)] = \\ &= \partial^{\Omega} f_1(x_0) + \partial^{\Omega} f_2(x_0). \end{aligned}$$

Здесь использовано равенство (см. задание 3.6)

$$\Gamma^+(x_0) = \Gamma^+(x_0) + \Gamma^+(x_0),$$

которое справедливо, так как $\Gamma^+(x_0)$ — выпуклый конус. ■

Замечание 2. Из доказательства формулы (10.9) ясно, что справедливы соотношения

$$\partial^{\Omega} f(x_0) = \partial^{\Omega} f_1(x_0) + \partial f_2(x_0) = \partial f_1(x_0) + \partial^{\Omega} f_2(x_0), \quad (10.10)$$

т.е. для одной из функций $f_1(x)$ или $f_2(x)$ можно рассматривать субдифференциал по E_n .

2. Представление условного ε -субдифференциала. Пусть $\varepsilon > 0$ фиксировано. Положим для $x_0 \in \Omega$

$$\begin{aligned} B_{\varepsilon}^{\Omega} f(x_0) &= \{v \in E_n \mid \exists x \in \Omega: v \in \partial^{\Omega} f(x), f(x) - f(x_0) \geq \\ &\geq (v, x - x_0) - \varepsilon\}. \end{aligned}$$

Теорема 10.1. *Имеет место соотношение*

$$\partial_{\varepsilon}^{\Omega} f(x_0) = \overline{B_{\varepsilon}^{\Omega} f(x_0)}. \quad (10.11)$$

Доказательство. Возьмем $v \in \overline{B_{\varepsilon}^{\Omega} f(x_0)}$. Тогда

$$v = \lim_{k \rightarrow \infty} v_k, \quad v_k \in B_{\varepsilon}^{\Omega} f(x_0) \quad \forall k.$$

В силу определения множества $\overline{B_{\varepsilon}^{\Omega} f(x_0)}$ найдутся $x_k \in \Omega$ такие, что $v_k \in \partial^{\Omega} f(x_k)$,

$$f(x_k) - f(x_0) \geq (v_k, x_k - x_0) - \varepsilon.$$

Поскольку $v_k \in \partial^{\Omega} f(x_k)$, то

$$f(z) - f(x_k) \geq (v_k, z - x_k) \quad \forall z \in \Omega.$$

Складывая два последних неравенства, имеем

$$f(z) - f(x_0) \geq (v_k, z - x_0) - \varepsilon \quad \forall z \in \Omega.$$

Переходя здесь к пределу по k при любом фиксированном $z \in \Omega$, получаем

$$f(z) - f(x_0) \geq (v, z - x_0) - \varepsilon \quad \forall z \in \Omega.$$

Значит, $v \in \partial_{\varepsilon}^{\Omega} f(x_0)$. Следовательно,

$$\overline{B_{\varepsilon}^{\Omega} f(x_0)} \subset \partial_{\varepsilon}^{\Omega} f(x_0). \quad (10.12)$$

Пусть теперь $v \in \partial_{\varepsilon}^{\Omega} f(x_0)$, т. е.

$$f(z) - f(x_0) \geq (v, z - x_0) - \varepsilon \quad \forall z \in \Omega. \quad (10.13)$$

Положим $h(z) = f(z) - f(x_0) - (v, z - x_0)$. Тогда из (10.13)

$$h(z) \geq -\varepsilon \quad \forall z \in \Omega. \quad (10.14)$$

Зафиксируем любое $\delta > 0$ и построим функцию

$$h_{\delta}(z) = h(z) + \delta \max_{i \in 1:n} |z^{(i)} - x_0^{(i)}| = h(z) + \delta q(z),$$

где $z = (z^{(1)}, \dots, z^{(n)})$, $x_0 = (x_0^{(1)}, \dots, x_0^{(n)})$, $q(z) = \max_{i \in 1:n} |z^{(i)} - x_0^{(i)}|$.

Функции $h(z)$, $h_{\delta}(z)$, $q(z)$ — выпуклые на Ω . Из (10.14) следует, что $h_{\delta}(z) \rightarrow +\infty$ при $\|z\| \rightarrow +\infty$. Поэтому при любом фиксированном $\delta > 0$ функция $h_{\delta}(z)$ достигает своего наименьшего на Ω значения в некоторой точке z_{δ} .

По необходимому и достаточному условию минимума функции $h_\delta(z)$ на множестве Ω будет (см. теорему 6.1)

$$\partial h_\delta(z_\delta) \cap \Gamma^+(z_\delta) \neq \emptyset. \quad (10.15)$$

Но

$$\partial h_\delta(z) = \partial h(z) + \delta \partial q(z), \quad (10.16)$$

где (см. пример 2 § 5)

$$\partial h(z) = \partial f(z) - \{v\}, \quad \partial q(z) = \text{co} \{w_i \mid i \in R(z)\},$$

$$R(z) = \{i \in 1:n \mid |z^{(i)} - x_0^{(i)}| = q(z)\},$$

$$w_i = \begin{cases} \text{co} \{e_i, -e_i\}, & \text{если } z^{(i)} - x_0^{(i)} = 0, \\ e_i, & \text{если } z^{(i)} - x_0^{(i)} > 0, \\ -e_i, & \text{если } z^{(i)} - x_0^{(i)} < 0, \end{cases}$$

e_i — единичные орты. Из (10.15) и (10.16) следует существование $w_\delta \in \partial q(z_\delta)$ такого, что

$$v_\delta = (v - \delta w_\delta) \in [\partial f(z_\delta) - \Gamma^+(z_\delta)] = \partial^\Omega f(z_\delta). \quad (10.17)$$

Ясно, что

$$v_\delta \rightarrow v \quad \text{при } \delta \rightarrow 0. \quad (10.18)$$

Из (10.14) следует $h_\delta(z_\delta) \geq h(z_\delta) \geq -\varepsilon$. Отсюда

$$f(z_\delta) - f(x_0) + (v - \delta w_\delta, x_0 - z_\delta) \geq -\varepsilon, \quad (10.19)$$

так как (см. пример 2 § 5)

$$\max_{i \in 1:n} |z^{(i)} - x_0^{(i)}| = (w_\delta, z_\delta - x_0).$$

Из (10.17) и (10.19) ясно, что $v_\delta \in B_\delta^\Omega f(x_0)$, а из (10.18) вытекает, что $v \in \overline{B_\varepsilon^\Omega f(x_0)}$. Таким образом,

$$\partial_\varepsilon^\Omega f(x_0) \subset \overline{B_\varepsilon^\Omega f(x_0)}. \quad (10.20)$$

Из (10.12) и (10.20) получаем (10.11). Теорема доказана.

В лемме 7.2 показано, что в случае сильно выпуклой функции $f(x)$ справедливо соотношение

$$\partial_\varepsilon f(x_0) = B_\varepsilon f(x_0).$$

Аналогично показывается, что если функция $f(x)$ сильно выпукла на множестве Ω , то

$$\partial_\varepsilon^\Omega f(x_0) = B_\varepsilon^\Omega f(x_0).$$

Задание 10.1. Доказать, что если функция $f(x)$ конечна и выпукла на E_n , а множество Ω есть гиперплоскость в E_n , то справедливо равенство

$$\partial_{\varepsilon}^{\Omega} f(x_0) = \partial_{\varepsilon} f(x_0) - \Gamma^{+}(x_0) \quad \forall \varepsilon \geq 0. \quad (10.21)$$

3. Условный ε -субдифференциал суммы функций. Пусть функции $f_1(x)$ и $f_2(x)$ выпуклы на E_n . Положим $f(x) = f_1(x) + f_2(x)$. Возьмем $x_0 \in \Omega$.

Теорема 10.2. Справедливо соотношение

$$\partial_{\varepsilon}^{\Omega} f(x_0) = \bigcup_{\substack{\varepsilon_1 + \varepsilon_2 = \varepsilon \\ \varepsilon_1 \geq 0, \varepsilon_2 \geq 0}} [\partial_{\varepsilon_1}^{\Omega} f_1(x_0) + \partial_{\varepsilon_2}^{\Omega} f_2(x_0)]. \quad (10.22)$$

Доказательство. Правую часть в (10.22) обозначим $\mathcal{D}$. Включение

$$\mathcal{D} \subset \partial_{\varepsilon}^{\Omega} f(x_0) \quad (10.23)$$

очевидно. Докажем противоположное включение. Пусть $v \in \partial_{\varepsilon}^{\Omega} f(x_0)$. По теореме 10.1 существует последовательность $\{v_k\}$ такая, что

$$v_k \in B_{\varepsilon}^{\Omega} f(x_0), \quad v_k \xrightarrow[k \rightarrow \infty]{} v, \quad (10.24)$$

т. е. для каждого v_k найдется такое $x_k \in \Omega$, что

$$v_k \in \partial^{\Omega} f(x_k), \quad f(x_k) - f(x_0) + (v_k, x_0 - x_k) \geq -\varepsilon.$$

Из (10.10) следует, что $v_k = v'_k + v''_k$, где

$$v'_k \in \partial f_1(x_k), \quad v''_k \in \partial^{\Omega} f_2(x_k). \quad (10.25)$$

Поэтому из (10.24) имеем

$$[(v'_k, x_0 - x_k) + f_1(x_k) - f_1(x_0)] + \\ + [(v''_k, x_0 - x_k) + f_2(x_k) - f_2(x_0)] \geq -\varepsilon. \quad (10.26)$$

Выражения в квадратных скобках в (10.26) неположительны. Поэтому неравенство (10.26) возможно лишь тогда, когда существуют $\varepsilon_{k1} \geq 0$, $\varepsilon_{k2} \geq 0$ такие, что $\varepsilon_{k1} + \varepsilon_{k2} \leq \varepsilon$ и

$$(v'_k, x_0 - x_k) + f_1(x_k) - f_1(x_0) \geq -\varepsilon_{k1}, \\ (v''_k, x_0 - x_k) + f_2(x_k) - f_2(x_0) \geq -\varepsilon_{k2}.$$

Отсюда и из (10.25) вытекает

$$v'_k \in B_{\varepsilon_{k1}} f_1(x_0), \quad v''_k \in B_{\varepsilon_{k2}}^{\Omega} f_2(x_0).$$

Последовательность $\{v'_k\}$ ограничена, так как

$$B_{\varepsilon_{k1}} f_1(x_0) \subset \partial_{\varepsilon_{k1}} f_1(x_0) \subset \partial_{\varepsilon} f_1(x_0),$$

а множество $\partial_{\varepsilon} f_1(x_0)$ — компакт в E_n . Из (10.24) следует, что ограничена и последовательность $\{v''_k\}$. Без ограничения общности можно считать, что

$$v'_k \rightarrow v', \quad v''_k \rightarrow v'', \quad \varepsilon_{k1} \rightarrow \varepsilon' \geq 0, \quad \varepsilon_{k2} \rightarrow \varepsilon'' \geq 0,$$

при этом $\varepsilon' + \varepsilon'' \leq \varepsilon$. Поскольку $v'_k \in \partial_{\varepsilon_{k1}} f_1(x_0)$, а $v''_k \in \partial_{\varepsilon_{k2}} f_2(x_0)$, то

$$f_1(x) - f_1(x_0) \geq (v'_k, x - x_0) - \varepsilon_{k1} \quad \forall x \in E_n,$$

$$f_2(z) - f_2(x_0) \geq (v''_k, z - x_0) - \varepsilon_{k2} \quad \forall z \in \Omega.$$

Переходя в этих неравенствах к пределу по k при фиксированных $x \in E_n$, $z \in \Omega$, убеждаемся в том, что $v' \in \partial_{\varepsilon'} f_1(x_0)$, $v'' \in \partial_{\varepsilon''} f_2(x_0)$. Поэтому из (10.24) и (10.25) будет

$$v = (v' + v'') \in [\partial_{\varepsilon'} f_1(x_0) + \partial_{\varepsilon''} f_2(x_0)] \subset \\ \subset [\partial_{\varepsilon'} f_1(x_0) + \partial_{\varepsilon''} f_2(x_0)],$$

т. е. $\partial_{\varepsilon} f(x_0) \subset \mathcal{D}$. Отсюда и из (10.23) следует (10.22). Теорема доказана.

Следствие. Из доказательства теоремы 10.2 ясно, что вместо (10.22) можно написать

$$\partial_{\varepsilon} f(x_0) = \bigcup_{\substack{\varepsilon_1 + \varepsilon_2 = \varepsilon \\ \varepsilon_1 \geq 0, \varepsilon_2 \geq 0}} [\partial_{\varepsilon_1} f_1(x_0) + \partial_{\varepsilon_2} f_2(x_0)] = \\ = \bigcup_{\substack{\varepsilon_1 + \varepsilon_2 = \varepsilon \\ \varepsilon_1 \geq 0, \varepsilon_2 \geq 0}} [\partial_{\varepsilon_1} f_1(x_0) + \partial_{\varepsilon_2} f_2(x_0)], \quad (10.27)$$

т. е. для одной из функций, $f_1(x)$ или $f_2(x)$, можно рассматривать не условный, а обычный ε -субдифференциал по E_n .

4. Распространение выпуклой функции. Пусть $\Omega \subset E_n$ — выпуклое множество, $G \subset \text{int } \Omega$ — выпуклый компакт, и пусть на Ω задана выпуклая функция $f(x)$. Образует функцию

$$\mathcal{F}(x) = \sup_{z \in G} [f(z) + (v(z), x - z)], \quad (10.28)$$

где $v(z)$ — произвольный (но фиксированный для каждого z) вектор из $\partial^2 f(z)$.

Лемма 10.3. *Функция $\mathcal{F}(z)$ выпукла и конечна на E_n и совпадает на G с функцией $f(x)$, т. е.*

$$\mathcal{F}(x) = f(x) \quad \forall x \in G. \quad (10.29)$$

Доказательство. Выпуклость $\mathcal{F}(x)$ очевидна. Докажем (10.29). Пусть $x \in G$. Тогда тем более $x \in \Omega$, и по определению множества $\partial^2 f(x)$ будет

$$f(x) - f(z) \geq (v(z), x - z) \quad \forall z \in G.$$

Значит,

$$f(x) \geq \sup_{z \in G} [f(z) + (v(z), x - z)] = \mathcal{F}(x).$$

Отсюда вытекает конечность $\mathcal{F}(x)$. Но $\mathcal{F}(x) \geq f(x)$. Из предыдущих двух неравенств следует (10.29). ■

Таким образом, любая выпуклая конечная на Ω функция может быть распространена (с сохранением выпуклости и конечности) на E_n , причем на G эта функция совпадает с $f(x)$, при этом супремум в (10.28) достигается, т. е.

$$\mathcal{F}(x) = \max_{z \in \Omega} [f(z) + (v(z), x - z)].$$

Замечание 3. В лемме 10.3 предполагалось, что функция $f(x)$ задана на множестве Ω , а множество $G \subset \subset \text{int } \Omega$ замкнуто и ограничено. Это условие нельзя отбросить, что видно из следующего примера.

Пусть $x \in E_1$, $f(x) = 1 - \sqrt{1 - x^2}$, $G = [0, 1]$. Функция $f(x)$ дифференцируема на $[0, 1)$. Поскольку

$$f''(x) = (1 - x^2)^{-3/2} > 0 \quad \forall x \in [0, 1),$$

то функция $f(x)$ выпукла. Она ограничена на $[0, 1]$. Однако $f'(x)$ не является ограниченной на $[0, 1)$, поэтому функция

$$\mathcal{F}(x) = \sup_{z \in [0, 1)} [f(z) + (f'(z), x - z)]$$

уже не является конечной при $x > 1$, ибо для любого фиксированного $x > 1$ имеем $\mathcal{F}(x) = +\infty$.

§ 11. Условные производные по направлениям. Непрерывность условного ε -субдифференциального отображения

1. Пусть функция $f(x)$ выпукла и конечна на E_n , $\Omega \subset E_n$ — выпуклое множество. Пусть $x_0 \in \Omega$, $\varepsilon \geq 0$, $g \in E_n$, $\|g\| = 1$. Положим

$$\frac{\partial^\Omega f(x_0)}{\partial g} = \sup_{v \in \partial^\Omega f(x_0)} (v, g), \quad \frac{\partial_\varepsilon^\Omega f(x_0)}{\partial g} = \sup_{v \in \partial_\varepsilon^\Omega f(x_0)} (v, g).$$

Величину $\frac{\partial^\Omega f(x_0)}{\partial g}$ назовем *условной производной функции $f(x)$ в точке x_0 по направлению g по множеству Ω* , а $\frac{\partial_\varepsilon^\Omega f(x_0)}{\partial g}$ назовем *условной ε -производной функции $f(x)$ в точке x_0 по направлению g по множеству Ω* .

Из (6.5) имеем (при $C = \partial^\Omega f(x_0)$, $B = \partial f(x_0)$)

$$\frac{\partial^\Omega f(x_0)}{\partial g} = \begin{cases} \frac{\partial f(x_0)}{\partial g}, & \text{если } g \in \Gamma(x_0), \\ +\infty, & \text{если } g \notin \Gamma(x_0). \end{cases} \quad (11.1)$$

Поскольку имеет место включение $\partial^\Omega f(x_0) \subset \partial_\varepsilon^\Omega f(x_0)$, то

$$\frac{\partial_\varepsilon^\Omega f(x_0)}{\partial g} = +\infty \quad \text{при } g \notin \Gamma(x_0).$$

Лемма 11.1. Если $\mathcal{K}$ — выпуклый конус, содержащий точки g и $-g$, причем $g \in \text{int } \mathcal{K}$, то

$$\mathcal{K} = E_n. \quad (11.2)$$

Доказательство. Поскольку $g \in \text{int } \mathcal{K}$, то найдется $r > 0$ такое, что

$$x_q = g + q \in \mathcal{K} \quad \forall q \in S_r(0).$$

Так как $\mathcal{K}$ — выпуклый конус, $g_1 = -g \in \mathcal{K}$, то

$$y_q = [0,5g_1 + 0,5x_q] \in \mathcal{K} \quad \forall q \in S_r(0). \quad (11.3)$$

Но $y_q = 0,5q$, поэтому из (11.3) $S_{0,5r}(0) \subset \mathcal{K}$. Итак, шар радиуса $0,5r$ с центром в начале координат содержится в конусе $\mathcal{K}$. Отсюда и вытекает (11.2). ■

Следствие. Пусть $\bar{x}$ — граничная точка множества Ω , $g \in E_n$, $\|g\| = 1$. Если $g \in \text{int } \Gamma(\bar{x})$, то

$$g_1 = -g \notin \Gamma(\bar{x}). \quad (11.4)$$

Действительно, допустим противное. Пусть $g_1 \in \Gamma(\bar{x})$. Тогда из леммы 11.1 следует, что

$$\Gamma(\bar{x}) = E_n,$$

что невозможно, поскольку $\bar{x}$ — граничная точка. ■

Теорема 11.1. Если $x_0 \in \Omega$, $g \in \gamma(x_0)$, $\|g\| = 1$, то

$$\begin{aligned} \frac{\partial_g^2 f(x_0)}{\partial g} &= \sup_{v \in \partial_g^2 f(x_0)} (v, g) = \\ &= \inf_{\substack{\alpha > 0 \\ x_0 + \alpha g \in \Omega}} \alpha^{-1} [f(x_0 + \alpha g) - f(x_0) + \varepsilon]. \end{aligned} \quad (11.5)$$

Доказательство. Прежде всего заметим, что если инфимум в (11.5) достигается при некотором α_1 , то $\alpha_1 > 0$, ибо $h(\alpha) \xrightarrow{\alpha \rightarrow +0} +\infty$, где $h(\alpha) = \alpha^{-1} [f(x_0 + \alpha g) - f(x_0) + \varepsilon]$.

Из определения $\partial_g^2 f(x_0)$ имеем

$$f(x_\alpha) - f(x_0) \geq \alpha(v, g) - \varepsilon \quad \forall v \in \partial_g^2 f(x_0), \quad \forall x_\alpha \in \Omega,$$

где $x_\alpha = x_0 + \alpha g$. Отсюда

$$h_\Omega^* \equiv \inf_{\alpha > 0, x_\alpha \in \Omega} h(\alpha) \geq \sup_{v \in \partial_g^2 f(x_0)} (v, g). \quad (11.6)$$

Установим противоположное неравенство. Поскольку

$$\partial_g^2 f(x) = \partial_g^2 f(x) \quad \forall x \in \Omega \text{ и } h_\Omega^* = h_\Omega^* \equiv \inf_{\alpha > 0, x_\alpha \in \Omega} h(\alpha),$$

то далее без ограничения общности считаем, что множество Ω замкнуто.

Напомним (см. следствие из леммы 8.3), что если α_1 — точка локального минимума функции $h(\alpha)$ на $(0, \infty)$, то

$$h^* \equiv \inf_{\alpha > 0} h(\alpha) = h(\alpha_1). \quad (11.7)$$

Возможны два случая:

$$a) \quad h^* = h_\Omega^*; \quad (11.8)$$

$$б) \quad h^* < h_\Omega^*. \quad (11.9)$$

Рассмотрим случай а). Согласно (8.2)

$$h^* = \max_{v \in \partial_g f(x_0)} (v, g). \quad (11.10)$$

Из того, что $\partial_g f(x_0) \subset \partial_g^{\Omega} f(x_0)$, и из (11.8), (11.10) имеем

$$h_{\Omega}^* \leq \sup_{v \in \partial_g^{\Omega} f(x_0)} (v, g),$$

что вместе с (11.6) дает (11.5).

Обратимся к случаю б). Пусть

$$h_{\Omega}^* = h(\alpha_1), \quad (11.11)$$

причем $x_{\alpha} \in \Omega$ при всех $\alpha \in (0, \alpha_1)$. Тогда $x_{\alpha} \notin \Omega$ при всех $\alpha > \alpha_1$, ибо если допустить противное, то, с учетом выпуклости множества Ω , точка α_1 будет точкой локального минимума функции $h(\alpha)$ на $(0, \infty)$ и поэтому, в силу (11.7) и (11.11), будет выполнено (11.8), что противоречит (11.9).

Таким образом, x_{α_1} — граничная точка множества Ω .

Поскольку $h(\alpha_1) = \min_{\alpha \in (0, \alpha_1]} h(\alpha)$, то, в силу замечания

1 § 6, должно быть

$$h'_-(\alpha_1) = \lim_{\alpha \rightarrow -0} \alpha^{-1} [h(\alpha_1 + \alpha) - h(\alpha_1)] \geq 0.$$

Но, как нетрудно подсчитать,

$$h'_-(\alpha_1) = \alpha_1^{-1} \frac{\partial f(x_1)}{\partial g_1} + \alpha_1^{-2} [f(x_1) - f(x_0) + \varepsilon],$$

где

$$g_1 = -g, \quad x_1 = x_{\alpha_1}.$$

Следовательно, существует $v_0 \in \partial f(x_1)$ такое, что

$$\alpha_1^{-1} [(v_0, g_1) + \alpha_1^{-1} (f(x_1) - f(x_0) + \varepsilon)] \geq 0,$$

т. е.

$$\alpha_1^{-1} [f(x_1) - f(x_0) + \varepsilon] \geq (v_0, g). \quad (11.12)$$

Предположим вначале, что $x_0 \in \text{int } \Omega$. Тогда $-g \in \text{int } \gamma(x_1)$ и, в силу следствия из леммы 11.1, $g \notin \Gamma(x_1)$. Значит, найдется $w \in \Gamma^+(x_1)$ такое, что $(w, g) < 0$. Поскольку $\Gamma^+(x_1)$ — конус, то существует $\lambda \geq 0$ такое, что $w_0 = \lambda w \in \Gamma^+(x_1)$, и вместо (11.12) будет

$$f(x_1) - f(x_0) = \alpha_1 (v_0 - w_0, g) - \varepsilon. \quad (11.13)$$

Учитывая, что

$$v_1 = (v_0 - w_0) \in \partial^{\Omega} f(x_1), \quad g = \alpha_1^{-1}(x_1 - x_0),$$

из (11.13) получим $f(x_1) - f(x_0) = (v_1, x_1 - x_0) - \varepsilon$. Поэтому по определению множества $B_{\varepsilon}^{\Omega} f(x_0)$ и по теореме 10.1 будет $v_1 \in B_{\varepsilon}^{\Omega} f(x_0) \subset \partial_{\varepsilon}^{\Omega} f(x_0)$. Таким образом, из (11.11) и (11.13) имеем

$$h_{\Omega}^* = h(\alpha_1) = (v_1, g) \leq \sup_{v \in \partial_{\varepsilon}^{\Omega} f(x_0)} (v, g). \quad (11.14)$$

Пусть теперь x_0 — граничная точка множества Ω . Положим

$$\Omega_{\delta} = \text{co} \{ \Omega \cup S_{\delta}(x_0) \}, \quad h_{\Omega_{\delta}}^* = \inf_{\alpha > 0, x_{\alpha} \in \Omega_{\delta}} h(\alpha), \quad \delta > 0.$$

В этом случае (рис. 13) $\Omega_{\delta} \cap \{x = x_0 + \alpha g \mid \alpha \geq 0\} = [x_0, x_{1\delta}]$, где $x_{1\delta} = x_0 + \alpha_{1\delta} g$. Покажем, что $x_{1\delta} \xrightarrow{\delta \rightarrow +0} x_1 = x_0 + \alpha_1 g$.

Поскольку $g \notin \gamma(x_1)$, то при любом фиксированном $\alpha > \alpha_1$ точка $x_{\alpha} = x_0 + \alpha g \notin \Omega$. По теореме отделимости (теорема 1.2) найдутся $g_0 \in E_n$ и число $a > 0$ такие, что

$$(x - x_{\alpha}, g_0) \leq -a \quad \forall x \in \Omega.$$

При достаточно малых $\delta > 0$ будет

$$(x - x_{\alpha}, g_0) \leq -0,5a \quad \forall x \in \Omega \cup S_{\delta}(x_1),$$

а тогда справедливо неравенство

$$(x - x_{\alpha}, g_0) \leq -0,5a \quad \forall x \in \Omega_{\delta},$$

т. е. $x_{\alpha} \in \Omega_{\delta}$ и $x_{1\delta} \in [x_1, x_{\alpha}]$ при достаточно малых $\delta > 0$.

Итак, $x_{1\delta} \xrightarrow{\delta \rightarrow +0} x_1$, а в силу непрерывности $h(\alpha)$ и $h_{\Omega_{\delta}}^* \rightarrow h_{\Omega}^*$. Отсюда и из (11.9) вытекает, что при достаточно малых δ будет $h^* < h_{\Omega_{\delta}}^*$. Значит, снова находимся в условиях случая б), но уже для множества Ω_{δ} . Тогда вместо (11.14) соответственно имеем

$$h_{\Omega_{\delta}}^* \leq \sup_{v \in \partial_{\varepsilon}^{\Omega_{\delta}} f(x_0)} (v, g).$$

Отсюда и из того, что $\partial_{\varepsilon}^{\Omega} f(x_0) \subset \partial_{\varepsilon}^{\Omega} f(x_0)$, ибо $\Omega \subset \Omega_{\delta}$, следует

$$h_{\Omega_{\delta}}^* \leq \sup_{v \in \partial_{\varepsilon}^{\Omega} f(x_0)} (v, g).$$

Переходя здесь к пределу по $\delta \rightarrow +0$, снова получаем (11.14).

Таким образом, в случае б) независимо от того, является точка x_0 граничной или внутренней точкой множества Ω , имеет место (11.14). Из (11.14) и (11.6) следует (11.5). Теорема доказана.

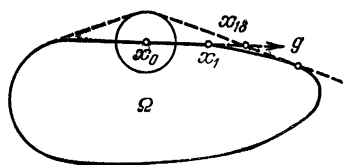


Рис. 13.

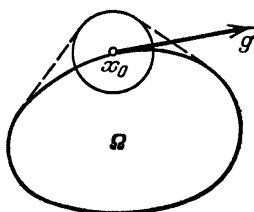


Рис. 14.

Следствие 1. Для любого $\varepsilon > 0$

$$\frac{\partial_{\varepsilon}^{\Omega} f(x_0)}{\partial g} = +\infty \quad \forall g \notin \gamma(x_0).$$

Действительно, пусть $x_0 \in \Omega$, $\varepsilon > 0$, $g \in \Gamma(x_0) \setminus \gamma(x_0)$ (рис. 14), т. е. $g \notin \gamma(x_0)$, но $g = \lim_{k \rightarrow \infty} g_k$, $g_k \in \gamma(x_0)$. Рассмотрим множества $\Omega_{\delta} = \text{co}\{\Omega \cup \bar{S}_{\delta}(x_0)\} \supset \Omega$, где $\delta > 0$. Тогда $g \in \gamma_{\delta}(x_0)$, где

$$\gamma_{\delta}(x_0) = \{v = \lambda(x - x_0) \mid \lambda > 0, x \in \Omega_{\delta}\}.$$

В силу теоремы 11.1 имеем

$$h_{\delta} = \inf_{\alpha > 0, x_{\alpha} \in \Omega_{\delta}} h(\alpha) = \sup_{v \in \partial_{\varepsilon}^{\Omega_{\delta}} f(x_0)} (v, g) \leq \sup_{v \in \partial_{\varepsilon}^{\Omega} f(x_0)} (v, g). \quad (11.15)$$

Но $h_{\delta} = \inf_{\alpha \in (0, \alpha_{\delta})} h(\alpha)$. Поскольку $g \notin \gamma(x_0)$, то $\alpha_{\delta} \xrightarrow{\delta \rightarrow +0} 0$,

поэтому $h_{\delta} \xrightarrow{\delta \rightarrow +0} +\infty$. Из (11.15) получаем $\frac{\partial_{\varepsilon}^{\Omega} f(x_0)}{\partial g} = +\infty$.

Выше было показано, что при $g \notin \Gamma(x_0)$

$$\frac{\partial_g^2 f(x_0)}{\partial g} = +\infty. \blacksquare$$

Следствие 2. Если для некоторого $g_0 \in E_n$, $\|g\| = 1$, оказалось

$$\frac{\partial_g^2 f(x_0)}{\partial g_0} < 0, \quad (11.16)$$

то

$$\inf_{\alpha > 0, x_\alpha \in \Omega} f(x_0 + \alpha g_0) < f(x_0) - \varepsilon. \quad (11.17)$$

Действительно, в силу (11.6) найдется $\alpha_1 > 0$ такое, что

$$h(\alpha_1) = \alpha_1^{-1} [f(x_0 + \alpha_1 g_0) - f(x_0) + \varepsilon] < 0, \quad (11.18)$$

причем $x_0 + \alpha_1 g_0 \in \Omega$. Поскольку $\alpha_1 > 0$, то из (11.18)

$$f(x_0 + \alpha_1 g_0) - f(x_0) + \varepsilon < 0,$$

откуда и следует (11.17). $\blacksquare$

Следствие 3. Если $f(x)$ сильно выпукла и $\frac{\partial_g^2 f(x_0)}{\partial g_0} \leq 0$, то

$$\inf_{\alpha > 0, x_0 + \alpha g_0 \in \Omega} f(x_0 + \alpha g_0) \leq f(x_0) - \varepsilon.$$

Доказательство. Случай $\frac{\partial_g^2 f(x_0)}{\partial g_0} < 0$ рассмотрен в следствии 2. Пусть теперь

$$\frac{\partial_g^2 f(x_0)}{\partial g_0} = 0.$$

Тогда найдутся последовательности $\{\alpha_k\}$, $\{\beta_k\}$ такие, что

$$\alpha_k > 0, \quad x_0 + \alpha_k g_0 \in \Omega, \quad h(\alpha_k) \leq \beta_k, \quad \beta_k \rightarrow 0,$$

где

$$h(\alpha) = \alpha^{-1} [f(x_0 + \alpha g_0) - f(x_0) + \varepsilon].$$

В силу сильной выпуклости функции $f(x)$ существует $\alpha_0 > 0$ такое, что $\alpha_k \leq \alpha_0$ для всех k . Поэтому

$$f(x_0 + \alpha_k g_0) - f(x_0) + \varepsilon \leq \alpha_k \beta_k \leq \alpha_0 \beta_k$$

Отсюда

$$\inf_{\alpha > 0, x_0 + \alpha g_0 \in \Omega} f(x_0 + \alpha g_0) - f(x_0) + \varepsilon \leq 0,$$

что и требовалось доказать. ■

Заметим, что условие сильной выпуклости существенно, что видно из следующего примера.

Пусть $f(x) \equiv \text{const}$, $\Omega = E_n$. Тогда для любого $g_0 \in E_n$, $\|g_0\| = 1$, будет

$$\inf_{\alpha > 0} \alpha^{-1} [f(x_0 + \alpha g_0) - f(x_0) + \varepsilon] = \inf_{\alpha > 0} \alpha^{-1} \varepsilon = 0,$$

но уменьшить функцию $f(x)$ нельзя.

2. Рассмотрим точечно-множественное отображение

$$\partial_{\varepsilon}^{\Omega} f: (0, \infty) \times \Omega \rightarrow \Pi(E_n),$$

где при фиксированных $x \in \Omega$ и $\varepsilon > 0$ по определению

$$\partial_{\varepsilon}^{\Omega} f(x) = \{v \in E_n \mid f(z) - f(x) \geq (v, z - x) - \varepsilon \quad \forall z \in \Omega\}.$$

Отображение $\partial_{\varepsilon}^{\Omega} f(x)$, как и отображение $\partial^{\Omega} f(x)$, не является ограниченным.

Пусть функция $f(x)$ конечна и выпукла на E_n .

Теорема 11.2. Если $\Omega \subset E_n$ — замкнутое выпуклое множество, то отображение $\partial_{\varepsilon}^{\Omega} f(x)$ непрерывно по Какутани по x и ε на множестве $(0, \infty) \times \Omega$ (т. е. полунепрерывно и сверху, и снизу).

Доказательство. Полунепрерывность сверху отображения $\partial_{\varepsilon}^{\Omega} f(x)$ следует непосредственно из определения. Докажем полунепрерывность этого отображения снизу.

Пусть даны $\varepsilon_0 > 0$, $x_0 \in \Omega$, $v_0 \in \partial_{\varepsilon_0}^{\Omega} f(x_0)$ и произвольные последовательности $\{\varepsilon_k\}$, $\{x_k\}$ такие, что $\varepsilon_k > 0$, $\varepsilon_k \rightarrow \varepsilon_0$, $x_k \in \Omega$, $x_k \rightarrow x_0$. Надо показать, что найдутся $v_k \in \partial_{\varepsilon_k}^{\Omega} f(x_k)$ такие, что $v_k \rightarrow v_0$. Допустим противное. Тогда существует $a > 0$ такое, что

$$\inf_{v \in \partial_{\varepsilon_k}^{\Omega} f(x_k)} \|v - v_0\| = \|v_k - v_0\| \geq a > 0 \quad \forall k. \quad (11.19)$$

Инфимум в (11.19) достигается, поскольку множества $\partial_{\varepsilon_k}^{\Omega} f(x_k)$ замкнуты. Без ограничения общности можно считать, что

$$v_k \rightarrow v^*, \quad g_k \equiv \|v_0 - v_k\|^{-1} (v_0 - v_k) \rightarrow g_0, \quad \|g_0\| = 1.$$

В силу (11.10)

$$(v - v_0, g_k) \leq (v_k - v_0, g_k) \leq -a \quad \forall v \in \partial_{e_k}^{\Omega} f(x_k).$$

Значит,

$$\frac{\partial_{e_k}^{\Omega} f(x_k)}{\partial g_k} = \sup_{v \in \partial_{e_k}^{\Omega} f(x_k)} (v, g_k) = (v_k, g_k) \quad (11.20)$$

и $(v_k, g_k) \leq (v_0, g_k) - a$. Отсюда и из (11.20) ясно, что найдется номер K такой, что

$$\frac{\partial_{e_k}^{\Omega} f(x_k)}{\partial g_k} = (v_k, g_k) \leq (v_0, g_0) - 0,5a \quad \forall k > K. \quad (11.21)$$

Поэтому

$$\frac{\partial_{e_0}^{\Omega} f(x_0)}{\partial g_0} \geq \frac{\partial_{e_k}^{\Omega} f(x_k)}{\partial g_k} + 0,5a \quad \forall k > K. \quad (11.22)$$

По теореме 11.1

$$\begin{aligned} \frac{\partial_{e_k}^{\Omega} f(x_k)}{\partial g_k} &= \inf_{\alpha > 0, x_k + \alpha g_k \in \Omega} \alpha^{-1} [f(x_k + \alpha g_k) - f(x_k) + \varepsilon] = \\ &= \inf_{\alpha > 0, x_k + \alpha g_k \in \Omega} h(x_k, g_k, \alpha, e_k). \end{aligned}$$

Из (11.21) ясно, что существует $\alpha_0 > 0$ такое, что

$$\frac{\partial_{e_k}^{\Omega} f(x_k)}{\partial g_k} = \inf_{\alpha \geq \alpha_0, x_k + \alpha g_k \in \Omega} h(x_k, g_k, \alpha, e_k). \quad (11.23)$$

Предположим вначале, что функция $f(x)$ сильно выпукла. Тогда при каждом k инфимум в (11.23) достигается в некоторой точке α_k , т. е.

$$\frac{\partial_{e_k}^{\Omega} f(x_k)}{\partial g_k} = h(x_k, g_k, \alpha_k, e_k), \quad (11.24)$$

причем $\alpha_k \leq \beta_0$, где $0 < \beta_0 < +\infty$, β_0 одно и то же для всех k . Можем считать, что $\alpha_k \rightarrow \alpha' \geq \alpha_0$. Тогда из (11.22) и (11.24) имеем

$$\frac{\partial_{e_k}^{\Omega} f(x_k)}{\partial g_k} \rightarrow h(x_0, g_0, \alpha', e_0) \leq \frac{\partial_{e_0}^{\Omega} f(x_0)}{\partial g_0} - 0,5a. \quad (11.25)$$

Так как $x_k + \alpha_k g_k \in \Omega$, то и $x_0 + \alpha' g_0 \in \Omega$. Поэтому

$$\frac{\partial_e^{\Omega} f(x_0)}{\partial g_0} = \inf_{\alpha \geq \alpha_0, x_0 + \alpha g_0 \in \Omega} h(x_0, g_0, \alpha, e_0) \leq h(x_0, g_0, \alpha', e_0),$$

что противоречит (11.25). Полученное противоречие и доказывает полунепрерывность снизу отображения $\partial_e^{\Omega} f(x)$ в точке $[e_0, x_0]$.

Пусть $f(x)$ — произвольная выпуклая конечная функция. Построим функцию $\mathcal{F}_{\delta}(x) = f(x) + \delta(x - x_0)^2$, где $\delta > 0$. По доказанному выше отображение $\partial_e^{\Omega} \mathcal{F}_{\delta}(x)$ полунепрерывно снизу в точке $[e_0, x_0]$ в силу сильной выпуклости функции $\mathcal{F}_{\delta}(x)$. По формуле (10.22) имеем

$$\partial_e^{\Omega} \mathcal{F}_{\delta}(x) = \bigcup_{\substack{e_1 + e_2 = e \\ e_1 \geq 0, e_2 \geq 0}} [\partial_{e_1}^{\Omega} f(x) + \partial_{e_2} f_{\delta}(x)],$$

где $f_{\delta}(x) = \delta(x - x_0)^2$. Как следует из примера 1 § 7 (см. (7.7)),

$$\partial_e f_{\delta}(x) = \{v = 2\delta(x - x_0) + r \mid r \in E_n, \|r\| \leq 2\sqrt{\epsilon\delta}\}.$$

Возьмем произвольные $v_0 \in \partial_{e_0}^{\Omega} f(x_0)$ и последовательности $\{e_k\}$ и $\{x_k\}$ такие, что $e_k \rightarrow e_0$, $x_k \rightarrow x_0$. Поскольку $0 \in \partial_{e_0} f_{\delta}(x_0)$, то $v_0 \in \partial_{e_0}^{\Omega} \mathcal{F}_{\delta}(x_0)$ для любого $\delta > 0$. Из полунепрерывности снизу отображения $\partial_e^{\Omega} \mathcal{F}_{\delta}(x)$ в точке $[e_0, x_0]$ найдется последовательность $\{v_{k\delta}\}$ такая, что

$$v_{k\delta} \in \partial_{e_k}^{\Omega} \mathcal{F}_{\delta}(x_k), \quad v_{k\delta} \rightarrow v_0.$$

При этом, в силу представления $\partial_e^{\Omega} \mathcal{F}_{\delta}(x)$, будет $v_{k\delta} = v'_{k\delta} + v''_{k\delta}$, где $v'_{k\delta} \in \partial_{e_k}^{\Omega} f(x_k)$, $v''_{k\delta} \in \partial_{e_k} f_{\delta}(x_k)$. Так как $\|v''_{k\delta}\| \xrightarrow{\delta \rightarrow 0} 0$, то можно выбрать подпоследовательности $\{\delta_k\}$ и $\{v'_{k\delta_k}\}$ такие, что $\delta_k \xrightarrow{k \rightarrow \infty} +0$, $v'_{k\delta_k} \xrightarrow{k \rightarrow \infty} v_0$. Значит, отображение $\partial_e^{\Omega} f(x)$ полунепрерывно снизу в точке $[e_0, x_0]$. В силу произвольности точки $[e_0, x_0]$ отображение $\partial_e^{\Omega} f(x)$ полунепрерывно снизу на множестве $(0, \infty) \times \Omega$. Теорема доказана.

Следствие. Если $x_0 \in \text{int } \Omega$, $e_0 > 0$, то отображение $\partial_e^{\Omega} f(x)$ непрерывно по Хаусдорфу в точке $[e_0, x_0]$.

Доказательство следует из леммы 10.1.

3. Необходимые и достаточные условия минимума в задаче с ограничениями. Пусть $\Omega \subset E_n$ — замкнутое выпуклое множество, и пусть на множестве Ω задана выпуклая функция $f(x)$. Рассмотрим задачу минимизации функции $f(x)$ на множестве Ω .

Справедливо следующее очевидное утверждение.

Лемма 11.2. *Для того чтобы в точке $x^* \in \Omega$ функция $f(x)$ достигала своего наименьшего на множестве Ω значения, необходимо и достаточно, чтобы*

$$\emptyset \subseteq \partial^2 f(x^*). \quad (11.26)$$

З а м е ч а н и е. Если функция $f(x)$ конечна и выпукла на E_n , то по лемме 10.2

$$\partial^2 f(x^*) = \partial f(x^*) - \Gamma^+(x^*),$$

и условие (11.26) эквивалентно соотношению (см. (6.16))

$$\partial f(x^*) \cap \Gamma^+(x^*) \neq \emptyset.$$

Если $\Omega = E_n$, то условие (11.26) имеет вид

$$\emptyset \subseteq \partial f(x^*).$$

Пусть $\varepsilon \geq 0$. Точка $\bar{x} \in \Omega$ называется ε -стационарной точкой функции $f(x)$ на множестве Ω , если

$$\emptyset \subseteq \partial_\varepsilon^2 f(\bar{x}). \quad (11.27)$$

Лемма 11.3. *Соотношение (11.27) эквивалентно неравенству*

$$0 \leq f(\bar{x}) - f^* \leq \varepsilon, \quad (11.28)$$

где $f^* = \inf_{x \in \Omega} f(x)$.

Доказательство. По определению соотношение (11.27) эквивалентно тому, что

$$f(x) \geq f(\bar{x}) + (\emptyset, x - \bar{x}) - \varepsilon = f(\bar{x}) - \varepsilon \quad \forall x \in \Omega,$$

что и требовалось доказать.

Пусть $x \in \Omega$. Положим

$$\psi_\varepsilon(x) = \inf_{\|g\|=1} \frac{\partial_\varepsilon^2 f(x)}{\partial g}.$$

Согласно следствию 1 теоремы 11.1

$$\psi_\varepsilon(x) = \min_{\|g\|=1, g \in \gamma(x)} \frac{\partial_\varepsilon^\Omega f(x)}{\partial g}.$$

Следствие 1. Условие (11.27) (или, что то же, условие (11.28)) эквивалентно тому, что

$$\psi_\varepsilon(\bar{x}) \geq 0.$$

Это вытекает из следствия 2 леммы 6.4.

Следствие 2. Если $0 \notin \partial_\varepsilon^\Omega f(\bar{x})$, то существует направление $g_0 \in \gamma(\bar{x})$, $\|g_0\|=1$, такое, что

$$\frac{\partial_\varepsilon^\Omega f(\bar{x})}{\partial g_0} < 0.$$

При этом (см. следствие 2 из теоремы 11.2) имеют место неравенства

$$\inf_{\alpha > 0, \bar{x} + \alpha g_0 \in \Omega} f(\bar{x} + \alpha g_0) < f(\bar{x}) - \varepsilon.$$

Пусть $x_0 \in \Omega$, $\varepsilon \geq 0$ и $\psi_\varepsilon(x_0) < 0$.

Определение. Вектор $g_0 \in E_n$, $\|g_0\|=1$, такой, что

$$\frac{\partial_\varepsilon^\Omega f(x_0)}{\partial g_0} = \inf_{\|g\|=1} \frac{\partial_\varepsilon^\Omega f(x_0)}{\partial g} = \psi_\varepsilon(x_0),$$

назовем направлением условного ε -наискорейшего (наискорейшего, если $\varepsilon = 0$) спуска функции $f(x)$ в точке $x_0 \in \Omega$.

Оно, согласно следствию из леммы 6.3, существует и единственно. Более того, если для $v(x_0) \in \partial_\varepsilon^\Omega f(x_0)$ имеем $\|v(x_0)\| = \min_{v \in \partial_\varepsilon^\Omega f(x_0)} \|v\|$, то вектор $g_0 = -\|v(x_0)\|^{-1}v(x_0)$

является направлением условного ε -наискорейшего спуска функции $f(x)$ в точке $x_0 \in \Omega$.

Отметим также, что, в силу следствия 1 из теоремы 11.1 и формулы (11.1),

$$\psi_\varepsilon(x_0) = \min_{\|g\|=1, g \in \gamma(x_0)} \frac{\partial_\varepsilon^\Omega f(x_0)}{\partial g}, \quad g_0 \in \gamma(x_0),$$

а если $\varepsilon = 0$, то

$$\psi(x_0) \equiv \psi_0(x_0) = \min_{\|g\|=1, g \in \Gamma(x_0)} \frac{\partial^\Omega f(x_0)}{\partial g}, \quad g_0 \in \Gamma(x_0).$$

Наконец, на основании следствия 2 из теоремы 11.1 заключаем, что если

$$\frac{\partial_e^{\Omega} f(x_0)}{\partial g_0} < 0,$$

то

$$\inf_{\alpha > 0, x_0 + \alpha g_0 \in \Omega} f(x_0 + \alpha g_0) < f(x_0) - \varepsilon. \quad (11.29)$$

§ 12. Задание выпуклого множества с помощью неравенств

1. Пусть $\Omega \subset E_n$ — выпуклое замкнутое множество. Для любого $x_0 \in \Omega$ рассмотрим *нормальный конус*

$$\mathcal{N}(x_0) = \{v \in E_n \mid (v, x - x_0) \leq 0 \ \forall x \in \Omega\}. \quad (12.1)$$

Если $x_0 \in \text{int } \Omega$, то ясно, что $\mathcal{N}(x_0) = \{0\}$. Если же x_0 — граничная точка Ω , то $\mathcal{N}(x_0) \setminus \{0\} \neq \emptyset$ (см. следствие 1 из теоремы 1.2). Положим

$$\mathcal{P} = \{x \in E_n \mid (v(x_0), x - x_0) \leq 0 \ \forall v(x_0) \in \mathcal{N}(x_0), \ \forall x_0 \in \Omega\}. \quad (12.2)$$

Замечание 1. Очевидно, что если $\Omega \neq E_n$, то

$$\mathcal{P} = \{x \in E_n \mid (v(x_0), x - x_0) \leq 0 \ \forall v(x_0) \in \mathcal{N}(x_0), \ \|v(x_0)\| = 1, \ \forall x_0 \in \Omega_{\text{fr}}\},$$

где Ω_{fr} — множество граничных точек Ω .

Лемма 12.1. *Имеет место соотношение*

$$\Omega = \mathcal{P}. \quad (12.3)$$

Доказательство. Пусть $x \in \Omega$. Из (12.1) ясно, что $x \in \mathcal{P}$, т. е.

$$\Omega \subset \mathcal{P}. \quad (12.4)$$

Пусть теперь $z \in \mathcal{P}$. Покажем, что $z \in \Omega$. Допустим противное. Предположим, что $z \notin \Omega$. Тогда

$$\min_{x \in \Omega} \|x - z\| = \|x_0 - z\| > 0. \quad (12.5)$$

Ясно, что x_0 — граничная точка Ω . Положим $g = \|z - x_0\|^{-1}(z - x_0)$. Из доказательства теоремы 1.2 вытекает, что

$$(x - x_0, g) \leq 0 \ \forall x \in \Omega,$$

т. е. $g \in \mathcal{N}(x_0)$. Из (12.5) $(g, z - x_0) = \|z - x_0\| > 0$, а это противоречит тому, что $z \in \mathcal{P}$. Значит, в действительности $z \in \mathcal{P}$, т. е. $\mathcal{P} \subset \Omega$. Отсюда и из (12.4) и следует утверждение леммы. ■

В определении множества $\mathcal{P}$ в (12.2) участвовали все векторы множества $\mathcal{N}(x_0)$. Установим условия, при которых для каждого x_0 достаточно взять лишь один ненулевой вектор из $\mathcal{N}(x_0)$. Пусть

$$\mathcal{P}_1 = \{x \in E_n \mid (A(x_0), x - x_0) \leq 0 \quad \forall x_0 \in \Omega_{\text{fr}}\}, \quad (12.6)$$

где Ω_{fr} — множество граничных точек Ω , $A(x_0) \in \mathcal{N}(x_0)$, $\|A(x_0)\| = 1$ (т. е. в каждой точке $x_0 \in \Omega_{\text{fr}}$ выбираем только один — все равно какой — единичный вектор из $\mathcal{N}(x_0)$).

Лемма 12.2. Если $\text{int } \Omega \neq \emptyset$, то

$$\Omega = \mathcal{P}_1. \quad (12.7)$$

Доказательство. Включение

$$\Omega \subset \mathcal{P}_1 \quad (12.8)$$

очевидно. Пусть теперь $z \in \mathcal{P}_1$. Надо показать, что $z \in \Omega$. Допустим противное. Пусть $z \notin \Omega$. Возьмем $\bar{x} \in \text{int } \Omega$. Тогда существует $r > 0$ такое, что $S_r(\bar{x}) \subset \Omega$. На отрезке $[\bar{x}, z]$ найдем граничную точку $x_1 \in \Omega_{\text{fr}}$ (она существует и единственна). Тогда $(A(x_1), x - x_1) \leq 0 \quad \forall x \in \Omega$. В частности, $(A(x_1), x - x_1) \leq 0 \quad \forall x \in S_r(\bar{x})$. Тогда

$$(A(x_1), \bar{x} + v - x_1) \leq 0 \quad \forall v \in S_r(0).$$

Значит,

$$(A(x_1), \bar{x} - x_1) \leq (-v, A(x_1)) \quad \forall v \in S_r(0). \quad (12.9)$$

Возьмем $v = rA(x_1)$. Имеем из (12.9) $(A(x_1), \bar{x} - x_1) \leq -r < 0$. С другой стороны, $z - x_1 = -\alpha(\bar{x} - x_1)$, где $\alpha > 0$. Отсюда

$$(A(x_1), z - x_1) = -\alpha(A(x_1), \bar{x} - x_1) \geq \alpha r > 0. \quad (12.10)$$

Неравенство (12.10) противоречит тому, что $z \in \mathcal{P}_1$. Из полученного противоречия следует $z \in \Omega$, т. е. $\mathcal{P}_1 \subset \Omega$, что вместе с (12.8) завершает доказательство леммы.

Замечание. 2. Требование наличия внутренних точек в лемме 12.2 (это условие Слейтера) существенно, что видно из следующего примера.

Пусть $x = (x^{(1)}, x^{(2)}, x^{(3)}) \in E_3$. На плоскости $Q = \{x \in E_3 \mid x^{(3)} = 0\}$ возьмем многоугольник Ω . Это множество выпуклое, но внутренних точек в пространстве E_3 оно не имеет. В данном случае вектор $e_3 = (0, 0, 1) \in \mathcal{N}(x_0)$ для всех $x_0 \in \Omega$. Следовательно, можно взять $A(x_0) = e_3$ для всех $x_0 \in \Omega_{\text{fr}}$, поэтому

$$\mathcal{P}_1 = \{x \in E_3 \mid (e_3, x - x_0) \leq 0\} = \{x \in E_3 \mid x^{(3)} \leq 0\}.$$

Однако ясно, что $\mathcal{P}_1 \neq \Omega$.

Из лемм 12.1 и 12.2 вытекает

Следствие. *Справедливо представление*

$$\Omega = \{x \in E_n \mid h(x) \leq 0\},$$

где

$$h(x) = \begin{cases} h_1(x), & \text{если } \text{int } \Omega = \emptyset, \\ h_2(x), & \text{если } \text{int } \Omega \neq \emptyset, \end{cases}$$

$$h_1(x) = \sup_{x_0 \in \Omega_{\text{fr}}} \sup_{\substack{A(x_0) \in \mathcal{N}(x_0) \\ \|A(x_0)\| = 1}} (A(x_0), x - x_0),$$

$$h_2(x) = \sup_{x_0 \in \Omega_{\text{fr}}} (A(x_0), x - x_0).$$

В определении $h_2(x)$ для каждого $x_0 \in \Omega_{\text{fr}}$ выбирается только один (все равно какой) вектор $A(x_0)$ из $\mathcal{N}(x_0)$, причем $\|A(x_0)\| = 1$. Функция $h(x)$ определена на E_n . Очевидно, что $h(x)$ — выпуклая конечная на Ω функция, и если Ω — ограниченное множество, то $h(x)$ — выпуклая конечная на E_n функция.

2. Пусть

$$\Omega = \{x \in E_n \mid h(x) \leq 0\}, \quad (12.11)$$

где $h(x)$ — некоторая выпуклая функция. Очевидно, что множество Ω , заданное (12.11), всегда является (если оно непустое) выпуклым и, если $h(x)$ — конечная функция, то и замкнутым.

Рассмотрим множество

$$\Omega = \{x \in E_n \mid h_1(x, y) \leq 0 \quad \forall y \in \omega\}, \quad (12.12)$$

где $\omega \subset E_p$ — ограниченное замкнутое множество, а функция $h_1(x, y)$ выпукла на E_n по x при любом $y \in \omega$.

Множество Ω , заданное (12.12), тоже выпукло, причем его можно представить в виде (12.11), если положить $h(x) = \max_{y \in \omega} h_1(x, y)$. При этом функция $h(x)$ является выпуклой.

Итак, пусть Ω задано в виде (12.11), где $h(x)$ — конечная выпуклая на E_n функция. Ниже будем предполагать, что выполнено *условие Слейтера*: существует $\bar{x} \in E_n$ такое, что

$$h(\bar{x}) < 0. \quad (12.13)$$

Возьмем $x_0 \in \Omega$ и положим

$$\begin{aligned} \gamma(x_0) &= \{v \in E_n \mid v = \lambda(x - x_0), \lambda > 0, x \in \Omega\}, \\ \Gamma(x_0) &= \bar{\gamma}(x_0). \end{aligned} \quad (12.14)$$

Множество $\gamma(x_0)$ можно представить и в виде, эквивалентном (12.14):

$$\begin{aligned} \gamma(x_0) &= \{v \in E_n \mid \exists \alpha_0 = \\ &= \alpha_0(v) > 0: h(x_0 + \alpha v) \leq 0 \quad \forall \alpha \in [0, \alpha_0]\}. \end{aligned} \quad (12.15)$$

Множества $\Gamma(x_0)$ и $\Gamma^+(x_0)$ изучались в § 3.

Положим

$$B(x_0) = \begin{cases} E_n, & \text{если } h(x_0) < 0, \\ \{v \in E_n \mid (w, v) \leq 0 \quad \forall w \in \partial h(x_0)\}, & \text{если } h(x_0) = 0 \end{cases}$$

где $\partial h(x_0)$ — субдифференциал функции $h(x)$ в точке x_0 , г. е.

$$\partial h(x_0) = \{w \in E_n \mid h(x) - h(x_0) \geq (w, x - x_0) \quad \forall x \in E_n\}.$$

Очевидно, что если $h(x_0) < 0$, то $x_0 \in \text{int } \Omega$. Из (12.13) следует, что если $x_0 \in \Omega$, но $x_0 \notin \text{int } \Omega$ (т. е. $h(x_0) = 0$), то

$$0 \notin \partial h(x_0), \quad (12.16)$$

ибо x_0 не является точкой минимума $h(x)$ на E_n .

Лемма 12.3. Если выполнено *условие Слейтера* (12.13), то

$$\Gamma(x_0) = B(x_0). \quad (12.17)$$

Доказательство. Соотношение (12.17) очевидно, если $h(x_0) < 0$. Пусть теперь

$$h(x_0) = 0. \quad (12.18)$$

Вначале покажем, что $\Gamma(x_0) \subset B(x_0)$. Пусть $v \in \Gamma(x_0)$; тогда найдется последовательность $\{v_k\}$ такая, что

$$v = \lim_{k \rightarrow \infty} v_k, \quad v_k \in \gamma(x_0).$$

Фиксируем произвольное k . Тогда (см. (12.15)) существует $\alpha_0(v_k) > 0$ такое, что $h(x_0 + \alpha v_k) \leq 0 \quad \forall \alpha \in (0, \alpha_0(v_k))$. Отсюда и из определения $\partial h(x_0)$ для всех $w \in \partial h(x_0)$ будет

$$(w, v_k) \leq \alpha^{-1} [h(x_0 + \alpha v_k) - h(x_0)] \leq 0,$$

ибо в силу (12.18) $h(x_0) = 0$. Фиксируя произвольное $w \in \partial h(x_0)$ и переходя к пределу при $k \rightarrow \infty$, имеем

$$(w, v) \leq 0 \quad \forall w \in \partial h(x_0),$$

т. е. $\Gamma(x_0) \subset B(x_0)$.

Пусть теперь $v \in B(x_0)$. Вначале предположим, что

$$\max_{w \in \partial h(x_0)} (w, v) = -a < 0. \quad (12.19)$$

Тогда

$$\begin{aligned} h(x_0 + \alpha v) &= h(x_0) + \alpha \frac{\partial h(x_0)}{\partial v} + o(\alpha) = \\ &= \alpha \max_{w \in \partial h(x_0)} (w, v) + o(\alpha) = -\alpha a + o(\alpha). \end{aligned}$$

Очевидно, при достаточно малых $\alpha > 0$ будет $h(x_0 + \alpha v) < 0$, т. е. $x_0 + \alpha v \in \Omega$. Значит, $v \in \gamma(x_0) \subset \Gamma(x_0)$. Если же $\max_{w \in \partial h(x_0)} (w, v) = 0$, то возьмем вектор $v_\varepsilon = v + \varepsilon(\bar{x} - x_0)$, где $\varepsilon > 0$, $\bar{x}$ — точка, удовлетворяющая условию Слейтера (12.13). Из определения $\partial h(x_0)$ имеем

$$h(\bar{x}) - h(x_0) \geq (w, \bar{x} - x_0) \quad \forall w \in \partial h(x_0),$$

т. е. $(w, \bar{x} - x_0) \leq h(\bar{x}) - h(x_0) < 0$. Поэтому

$$(w, v_\varepsilon) = (w, v) + \varepsilon(w, \bar{x} - x_0).$$

Отсюда

$$\begin{aligned} \max_{w \in \partial h(x_0)} (w, v_\varepsilon) &\leq \max_{w \in \partial h(x_0)} (w, v) + \\ &+ \varepsilon \max_{w \in \partial h(x_0)} (w, \bar{x} - x_0) \leq \varepsilon h(\bar{x}) < 0. \end{aligned}$$

Итак, для v_ε имеет место (12.19), и по доказанному $v_\varepsilon \in \gamma(x_0)$. При $\varepsilon \rightarrow 0$ имеем $v_\varepsilon \rightarrow v$, $v_\varepsilon \in \gamma(x_0)$. Значит, $v \in \Gamma(x_0)$. Лемма доказана.

Пусть снова $x_0 \in \Omega$. Положим

$$\mathcal{K}(x_0) = \begin{cases} \{\emptyset\}, & \text{если } h(x_0) < 0, \\ \{v = \lambda w \mid \lambda \leq 0, w \in \partial h(x_0)\}, & \text{если } h(x_0) = 0. \end{cases}$$

Лемма 12.4. Если выполнено условие Слейтера (12.13), то

$$\Gamma^+(x_0) = \mathcal{K}(x_0). \quad (12.20)$$

Доказательство. Утверждение очевидно, если $h(x_0) < 0$. Рассмотрим случай

$$h(x_0) = 0. \quad (12.21)$$

Тогда по лемме 12.3

$$\Gamma(x_0) = \{v \in E_n \mid (w, v) \leq 0 \quad \forall w \in \partial h(x_0)\}.$$

Согласно (12.15) для любого $v \in \gamma(x_0)$ существует $\alpha_0(v) > 0$ такое, что

$$h(x_0 + \alpha v) \leq 0 \quad \forall \alpha \in [0, \alpha_0(v)]. \quad (12.22)$$

Для любого $w \in \partial h(x_0)$ по определению субдифференциала

$$h(x_0 + \alpha v) - h(x_0) \geq \alpha(w, v).$$

Учитывая (12.22) и (12.21), имеем $(w, v) \leq 0$. Поэтому

$$(\lambda w, v) \geq 0 \quad \forall \lambda \leq 0, \quad \forall v \in \gamma(x_0).$$

Так как $\Gamma(x_0) = \overline{\gamma(x_0)}$, то отсюда

$$(\lambda w, v) \geq 0 \quad \forall v \in \Gamma(x_0),$$

а это и значит (см. определение сопряженного конуса в § 3), что $\lambda w \in \Gamma^+(x_0)$. Итак,

$$\mathcal{K}(x_0) \subset \Gamma^+(x_0). \quad (12.23)$$

Пусть теперь $v \in \Gamma^+(x_0)$, т. е.

$$(v, g) \geq 0 \quad \forall g \in \Gamma(x_0). \quad (12.24)$$

Рассмотрим луч $\mathcal{L} = \{w = -\alpha v \mid \alpha \geq 0\}$. Если $\mathcal{L} \cap \partial h(x_0) \neq \emptyset$, то существует α_0 такое, что $w = -\alpha_0 v \in \partial h(x_0)$.

Очевидно, что $\alpha_0 > 0$, ибо в случае $\alpha_0 = 0$ имеем $0 \in \partial h(x_0)$, что противоречит (12.16). Тогда $v = -\alpha_0^{-1}w$, а это и значит, что $v \in \mathcal{H}(x_0)$, ибо $\lambda = -\alpha_0^{-1} < 0$.

Покажем, что всегда

$$\mathcal{L} \cap \partial h(x_0) \neq \emptyset. \quad (12.25)$$

Допустим противное, пусть $\mathcal{L} \cap \partial h(x_0) = \emptyset$. Как и при доказательстве леммы 6.2, получаем, что существует z , $\|z\| = 1$, такое, что

$$(z, v) = 0, \quad (z, g) = -a < 0 \quad \forall g \in \partial h(x_0). \quad (12.26)$$

В силу (12.26) (рис. 15) найдется z' такое, что

$$(z', g) \leq -0,5a \quad \forall g \in \partial h(x_0), \quad (12.27)$$

$$(z', v) < 0. \quad (12.28)$$

Из (12.27) и леммы 12.3 заключаем, что $z' \in \Gamma$. Но тогда (12.28) противоречит (12.24).

Итак, всегда имеет место (12.25), а тогда, как уже установлено, $v \in \mathcal{H}(x_0)$, т. е. $\Gamma^+(x_0) \subset \mathcal{H}(x_0)$. Отсюда и из (12.23) получаем (12.20). Лемма доказана.

3. Рассмотрим случай, когда множество Ω имеет вид

$$\Omega = \{x \in E_n \mid h(x) \leq 0\}, \quad (12.29)$$

где

$$h(x) = \max_{i \in I} h_i(x), \quad I = 1 : N, \quad (12.30)$$

функции $h_i(x)$ выпуклы и непрерывно дифференцируемы на E_n . Предположим, что выполнено условие

Слейтера (12.13): существует точка $\bar{x} \in E_n$ такая, что

$$h_i(\bar{x}) < 0 \quad \forall i \in I. \quad (12.31)$$

Очевидно, что $\bar{x} \in \text{int } \Omega$, и поэтому $\text{int } \Omega \neq \emptyset$.

Задание 12.1. Доказать, что условие Слейтера эквивалентно условию: найдутся такие точки $x_i \in \Omega$, $i \in I$, что

$$h_i(x_i) < 0 \quad \forall i \in I.$$

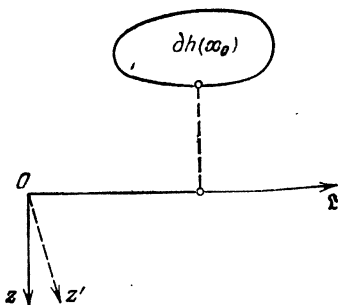


Рис. 15.

Пусть $x \in \Omega$. Положим $Q(x) = \{i \in I \mid h_i(x) = 0\}$. Очевидно, что $h_i(x) < 0 \forall i \notin Q(x)$.

Заметим, что если $Q(x) = \emptyset$, то $x \in \text{int } \Omega$. Обратное неверно, так как из того, что $x \in \text{int } \Omega$, еще не следует $Q(x) = \emptyset$.

Из лемм 5.4 и 12.3 вытекает

Лемма 12.5. Если $x_0 \in \Omega$, множество Ω задано соотношениями (12.29), (12.30) и выполнено условие Слейтера (12.31), то

$$\Gamma(x_0) = B(x_0), \quad (12.32)$$

где

$$B(x_0) = \begin{cases} E_n, & \text{если } Q(x_0) = \emptyset, \\ \{v \in E_n \mid (v, h'_i(x_0)) \leq 0 \forall i \in Q(x_0)\}, & \text{если } Q(x_0) \neq \emptyset. \end{cases} \quad (12.33)$$

Из лемм 5.4 и 12.4 вытекает

Лемма 12.6. В условиях леммы 12.5

$$\Gamma^+(x_0) = A(x_0), \quad (12.34)$$

где

$$A(x_0) = \begin{cases} \{0\}, & \text{если } Q(x_0) = \emptyset, \\ \left\{ v = - \sum_{i \in Q(x_0)} \lambda_i h'_i(x_0) \mid \lambda_i \geq 0 \forall i \in Q(x_0) \right\}, & \text{если } Q(x_0) \neq \emptyset. \end{cases} \quad (12.35)$$

§ 13. Нормальный конус.

Коническое отображение

1. Пусть $\Omega \subset E_n$ — выпуклое замкнутое множество. Для $x_0 \in \Omega$ в § 12 был введен нормальный конус:

$$\mathcal{N}(x_0) = \{v \in E_n \mid (v, x - x_0) \leq 0 \forall x \in \Omega\}. \quad (13.1)$$

Пусть теперь x_0 — произвольная точка E_n (не обязательно из Ω).

Лемма 13.1. Для любого $x_0 \in E_n$ множество $\mathcal{N}(x_0)$ является непустым, выпуклым, замкнутым конусом. Если $x_0 \notin \text{int } \Omega$, то

$$\mathcal{N}(x_0) \setminus \{0\} \neq \emptyset. \quad (13.2)$$

Если $x_0 \in \text{int } \Omega$, то $\mathcal{N}(x_0) = \{0\}$.

Доказательство. Все свойства, кроме (13.2), очевидны. Соотношение (13.2) вытекает из теоремы 1.2 (если $x_0 \notin \Omega$) или из ее следствия 1 (если x_0 — граничная точка Ω). ■

Теорема 13.1. Если x_0 не является граничной точкой множества Ω , то точечно-множественное отображение $\mathcal{N}(x)$ непрерывно по Какутани в точке x_0 .

Доказательство. Если $x \in \text{int } \Omega$, то, как уже отмечалось в п. 1 § 1, $\mathcal{N}(x) = \{0\}$ и непрерывность в точке $x_0 \in \text{int } \Omega$ очевидна. Пусть теперь $x_0 \notin \text{int } \Omega$. Необходимо установить полунепрерывность $\mathcal{N}(x)$ в точке x_0 сверху и снизу. Для любых последовательностей $\{x_k\}$ и $\{v_k\}$ таких, что $x_k \rightarrow x_0$, $v_k \rightarrow v_0$, $v_k \in \mathcal{N}(x_k)$, из (13.1) следует, что $v_0 \in \mathcal{N}(x_0)$, т. е. отображение $\mathcal{N}(x)$ полунепрерывно в точке x_0 сверху.

Возьмем произвольное $v_0 \in \mathcal{N}(x_0)$ и любую последовательность $\{x_k\}$ такую, что $x_k \rightarrow x_0$. Найдем $\min_{z \in \Omega} \|z - x_k\| = \|z_k - x_k\|$. Без ограничения общности можно считать, что $x_k \notin \Omega$ и

$$\|z_k - x_k\| \geq a > 0, \quad (13.3)$$

где $2a = \min_{z \in \Omega} \|z - x_0\|$. Как следует из замечания 3 § 1,

$$(x - x_k, z_k - x_k) \geq \|z_k - x_k\|^2 \quad \forall x \in \Omega. \quad (13.4)$$

Положим

$$\begin{aligned} \delta_k &= 2\|z_k - x_k\|^{-1} |(v_0, x_0 - x_k)|, \\ v_k &= v_0 + \delta_k \|x_k - z_k\|^{-1} (x_k - z_k). \end{aligned} \quad (13.5)$$

Ясно, что $\delta_k \rightarrow 0$. Тогда для $x \in \Omega$ имеем

$$\begin{aligned} (x - x_k, v_k) &= (x - x_k, v_0 + \delta_k \|x_k - z_k\|^{-1} (x_k - z_k)) = \\ &= (x - x_0, v_0) + (x_0 - x_k, v_0) + \\ &\quad + \delta_k (x - x_k, \|x_k - z_k\|^{-1} (x_k - z_k)). \end{aligned} \quad (13.6)$$

Поскольку $v_0 \in \mathcal{N}(x_0)$, то

$$(x - x_0, v_0) \leq 0 \quad \forall x \in \Omega. \quad (13.7)$$

Из (13.4) — (13.7) получаем

$$(x - x_k, v_k) \leq -|(v_0, x_0 - x_k)| \leq 0 \quad \forall x \in \Omega, \quad (13.8)$$

т. е. $v_k \in \mathcal{N}(x_k)$. Так как $\delta_k \rightarrow 0$, то $v_k \rightarrow v_0$, а это и значит, что отображение $\mathcal{N}(x)$ полунепрерывно снизу в точке x_0 . Итак, отображение $\mathcal{N}(x)$ полунепрерывно и сверху, и снизу в точке $x_0 \notin \Omega$. Теорема доказана.

2. Пусть $f(x)$ — выпуклая на E_n функция. В пространстве E_{n+1} рассмотрим множество (см. п. 1 § 4)

$$G = \text{epi } f = \{\bar{x} = [\beta, x] \mid \beta \in E_1, \beta \geq f(x), x \in E_n\}.$$

Это выпуклое замкнутое множество. Возьмем точку $\bar{x}_0 = [f(x_0) - \varepsilon_0, x_0]$, где $\varepsilon_0 > 0$. Ясно, что $\bar{x}_0 \notin G$. По доказанному в теореме 13.1 отображение $\mathcal{N}(\bar{x}_0) = \{\bar{v} \in E_{n+1} \mid (\bar{v}, \bar{x} - \bar{x}_0) \leq 0 \quad \forall \bar{x} \in G\}$ непрерывно по Какутани (или K -непрерывно). Положим

$$Q(\bar{x}_0) = \{v \in E_n \mid [-1, v] \in \mathcal{N}(\bar{x}_0)\}.$$

Лемма 13.2. *Имеет место соотношение*

$$Q(\bar{x}_0) = \partial_{\varepsilon_0} f(x_0). \quad (13.9)$$

Доказательство. Возьмем любое $v_0 \in \partial_{\varepsilon_0} f(x_0)$. Тогда

$$f(x) - f(x_0) \geq (v_0, x - x_0) - \varepsilon_0 \quad \forall x \in E_n.$$

Для $\beta \geq f(x)$ тем более

$$(\beta - [f(x_0) - \varepsilon_0])(-1) + (v_0, x - x_0) \leq 0 \quad \forall x \in E_n.$$

Положив $\bar{v}_0 = [-1, v_0]$, $\bar{x} = [\beta, x]$, отсюда имеем

$$(\bar{v}_0, \bar{x} - \bar{x}_0) \leq 0 \quad \forall \bar{x} \in G, \quad (13.10)$$

ибо $\bar{x} = [\beta, x] \in G$ при $\beta \geq f(x)$. Из (13.10) следует, что $\bar{v}_0 \in \mathcal{N}(\bar{x}_0)$, и потому $v_0 \in Q(\bar{x}_0)$, а значит,

$$\partial_{\varepsilon} f(x_0) \subset Q(\bar{x}_0). \quad (13.11)$$

Теперь пусть $v_0 \in Q(\bar{x}_0)$, тогда $\bar{v}_0 = [-1, v_0] \in \mathcal{N}(\bar{x}_0)$, т. е.

$$(\bar{v}_0, \bar{x} - \bar{x}_0) \leq 0 \quad \forall \bar{x} \in G. \quad (13.12)$$

Так как

$$\bar{x} = [\beta, x], \quad \beta \geq f(x), \quad \bar{x}_0 = [f(x_0) - \varepsilon_0, x_0],$$

то из (13.12)

$$[(-1)(\beta - f(x_0) + \varepsilon_0) + (v_0, x - x_0)] \leq 0 \quad \forall x \in E_n.$$

При $\beta = f(x)$ будет

$$f(x) - f(x_0) \geq (v_0, x - x_0) - \varepsilon_0 \quad \forall x \in E_n,$$

а это значит, что $v_0 \in \partial_{\varepsilon_0} f(x_0)$. Итак, $Q(x_0) \subset \partial_{\varepsilon_0} f(x_0)$. Отсюда и из (13.11) получаем (13.9). ■

Из только что доказанной леммы и теоремы 13.1 вытекает

Следствие. *Отображение $\partial_\varepsilon f(x)$ K -непрерывно по ε и x на $(0, \infty) \times E_n$.*

Действительно, полунепрерывность сверху отображения $\partial_\varepsilon f(x)$ очевидна. Покажем полунепрерывность его снизу. Пусть $x_k \rightarrow x_0$, $v_0 \in \partial_\varepsilon f(x_0) = Q(\bar{x}_0)$, $\varepsilon_k \rightarrow \varepsilon$ и $\bar{x}_k = [f(x_k) - \varepsilon_k, x_k]$. Так как $\bar{x}_k \rightarrow \bar{x}_0 \notin G$ и $\bar{v}_0 = [-1, v_0] \in \mathcal{N}(\bar{x}_0)$, то в силу полунепрерывности снизу отображения $\mathcal{N}(\bar{x})$ найдется такая последовательность $\{\bar{v}_k\}$, что $\bar{v}_k \in \mathcal{N}(\bar{x}_k)$, $\bar{v}_k \rightarrow \bar{v}_0$. Поскольку $\bar{v}_k = [\beta_k, v_k] \in \mathcal{N}(\bar{x}_k)$, то и

$$\bar{v}'_k = |\beta_k|^{-1} \bar{v}_k = [-1, |\beta_k|^{-1} v_k] \in \mathcal{N}(\bar{x}_k),$$

причем $\bar{v}'_k \rightarrow \bar{v}_0$. Тогда $v'_k = |\beta_k|^{-1} v_k \rightarrow v_0$ и $v'_k \in Q(x_k)$, что и требовалось доказать. ■

Отображение $\partial_\varepsilon f(x)$ — ограниченное, замкнутое, и по теореме 2.1 оно является и H -непрерывным, т. е. непрерывно по Хаусдорфу. H -непрерывность отображения $\partial_\varepsilon f(x)$ уже была установлена в § 8 другим способом.

Замечание. Если функция $f(x)$ определена лишь на выпуклом замкнутом множестве Ω , то, положив

$$G = \{\bar{x} = [\beta, x] \in E_{n+1} \mid x \in \Omega, \beta \geq f(x)\}$$

и рассуждая так же, как при доказательстве леммы 13.2, устанавливаем, что для $\bar{x}_0 = [f(x_0) - \varepsilon_0, x_0]$, где $\varepsilon_0 > 0$, $x_0 \in \Omega$, будет

$$Q(\bar{x}_0) = \partial_{\varepsilon_0} f(x_0). \quad (13.13)$$

Из (13.12) и теоремы 13.1 следует K -непрерывность условного ε -субдифференциала. Этот факт также был установлен другим способом в § 11.

3. Пусть $\mathcal{K}: G \rightarrow \Pi(E_n)$ — точечно-множественное отображение, заданное на $G \subset E_n$, причем при каждом $x \in G$ множество $\mathcal{K}(x)$ — выпуклый замкнутый конус с вершиной в начале координат. Такое отображение назовем *коническим*.

Обозначим $\mathcal{K}^+(x)$ сопряженный конус:

$$\mathcal{K}^+(x) = \{w \in E_n \mid (v, w) \geq 0 \forall v \in \mathcal{K}(x)\}.$$

Свойства сопряженных конусов изучались в § 3. В частности, там было установлено, что $\mathcal{K}^+(x)$ — выпуклый замкнутый конус, причем

$$(\mathcal{K}^+(x))^+ \equiv \mathcal{K}^{++}(x) = \mathcal{K}(x). \quad (13.14)$$

Очевидно, что отображение $\mathcal{K}^+(x)$ тоже является коническим.

Лемма 13.3. Если т. м. о. $\mathcal{K}(x)$ полунепрерывно сверху в точке $x_0 \in G$, то т. м. о. $\mathcal{K}^+(x)$ является полунепрерывным снизу в этой же точке.

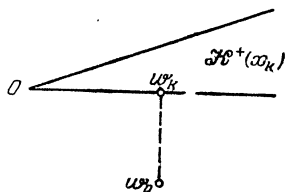


Рис. 16.

Доказательство. Пусть $w_0 \in \mathcal{K}^+(x_0)$, и пусть последовательность $\{x_k\}$ такова, что $x_k \in G$, $x_k \rightarrow x_0$. Надо доказать, что найдутся $z_k \in \mathcal{K}^+(x_k)$ такие, что $z_k \rightarrow w_0$.

Допустим, что это не так. Тогда без ограничения общности можно считать, что существует $a > 0$ такое, что

$$\min_{w \in \mathcal{K}^+(x_k)} \|w - w_0\| = \|w_k - w_0\| = a_k \geq a > 0 \quad \forall k.$$

Заметим, что $\|w_k\|$ ограничены, ибо (рис. 16) $\|w_k\| \leq \|w_0\|$. По необходимому и достаточному условию минимума (см. замечание 3 § 1)

$$(w_k - w_0, w - w_0) \geq a_k^2 \geq a^2 > 0 \quad \forall w \in \mathcal{K}^+(x_k). \quad (13.15)$$

Положим $v_k = w_k - w_0$. Как и при доказательстве теоремы 3.1 (см. (3.7)), имеем $(v_k, w_k) = 0$. Поэтому из (13.15)

$$\begin{aligned} (v_k, w) &= (v_k, w - w_0) + (v_k, w_0 - w_k) + (v_k, w_k) \geq \\ &\geq a^2 - a^2 = 0 \quad \forall w \in \mathcal{K}^+(x_k), \end{aligned}$$

т. е. в силу (13.14)

$$v_k \in (\mathcal{K}^+(x_k))^+ \equiv \mathcal{K}^{++}(x_k) = \mathcal{K}(x_k).$$

Без ограничения общности можем считать, что $v_k \rightarrow v_0$. Поскольку $\mathcal{K}(x)$ полунепрерывно сверху в точке x_0 , то

$v_0 \in \mathcal{K}(x_0)$. При $w = 0$ из (13.15) получим $(v_k, w_0) \leq \leq -a_k^2 \leq -a^2 < 0$. Переходя к пределу, имеем

$$(v_0, w_0) \leq -a^2. \quad (13.16)$$

Но $v_0 \in \mathcal{K}(x_0)$ и (13.16) противоречит тому, что $w_0 \in \mathcal{K}^+(x_0)$. Полученное противоречие и доказывает лемму.

Лемма 13.4. *Если т. м. о. $\mathcal{K}(x)$ полунепрерывно снизу в точке $x_0 \in G$, то т. м. о. $\mathcal{K}^+(x)$ является полунепрерывным сверху в точке x_0 .*

Доказательство. Допустим противное. Тогда найдутся последовательности $\{x_k\}$ и $\{w_k\}$ такие, что

$$x_k \in G, \quad x_k \rightarrow x_0, \quad w_k \in \mathcal{K}^+(x_k), \\ w_k \rightarrow w_0, \quad \text{но} \quad w_0 \notin \mathcal{K}^+(x_0).$$

Найдем

$$\min_{w \in \mathcal{K}^+(x_0)} \|w - w_0\| = \|\bar{w} - w_0\| = \|\bar{v}\|.$$

Тогда

$$\|\bar{v}\| = a > 0, \quad (\bar{v}, w - w_0) \geq a^2 > 0 \quad \forall w \in \mathcal{K}^+(x_0),$$

т. е. $\bar{v} \in (\mathcal{K}^+(x_0))^+ = \mathcal{K}(x_0)$. При этом

$$(\bar{v}, w_0) = -\|\bar{v}\|^2 = -a^2 < 0. \quad (13.17)$$

Поскольку т. м. о. $\mathcal{K}(x)$ полунепрерывно снизу в точке x_0 , то для точки $\bar{v}$ и последовательности $\{x_k\}$ найдутся $v_k \in \mathcal{K}(x_k)$ такие, что $v_k \rightarrow \bar{v}$. Тогда

$$(v_k, w) \geq 0 \quad \forall w \in \mathcal{K}^+(x_k).$$

В частности, для $w_k \in \mathcal{K}^+(x_k)$ отсюда получаем

$$(v_k, w_k) \geq 0.$$

Переходя здесь к пределу, имеем $(\bar{v}, w_0) \geq 0$, что противоречит (13.17). Лемма доказана.

Из лемм 13.3 и 13.4 вытекает

Следствие. *Если $\mathcal{K}(x)$ — коническое отображение, непрерывное по Какутани в точке x_0 , то и т. м. о. $\mathcal{K}^+(x)$ тоже является непрерывным по Какутани в точке x_0 .*

§ 14. Дифференцируемость по направлениям функции супремума

В § 5 уже рассматривался пример функции максимума конечного числа выпуклых функций и была установлена формула для производной по направлениям и вид субградиента. Ниже, в п. 1, более подробно изучается задача дифференцируемости по направлениям функции супремума дифференцируемых, не обязательно выпуклых функций, а затем в п. 2 рассматривается функция супремума выпуклых функций.

1. Пусть

$$f(x) = \sup_{y \in G} \varphi(x, y), \quad (14.1)$$

где $x \in E_n$, $y \in E_p$, $G \subset E_p$ — произвольное множество. Параметрические семейства функций $\varphi(x, y)$ и $\varphi'_x(x, y)$, где $y \in G$ — параметр, предполагаются равностепенно непрерывными по x в точке x_0 , т. е. для любого $\varepsilon > 0$ найдется $\delta > 0$ такое, что

$$|\varphi(x, y) - \varphi(x_0, y)| \leq \varepsilon,$$

$$\|\varphi'_x(x, y) - \varphi'_x(x_0, y)\| \leq \varepsilon \quad \forall x: \|x - x_0\| \leq \delta, \quad \forall y \in G.$$

Кроме того, предполагается, что

$$|\varphi(x_0, y)| \leq K_0 < \infty, \quad \|\varphi'_x(x_0, y)\| \leq K < \infty \quad \forall y \in G. \quad (14.2)$$

Легко убедиться, что $f(x)$ — непрерывная в точке x_0 функция. Возьмем $\varepsilon > 0$. Введем множества

$$R_\varepsilon(x) = \{y \in G \mid f(x) - \varphi(x, y) \leq \varepsilon\},$$

$$\mathcal{H}(x) = \{v \in E_n \mid \exists \varepsilon_i \rightarrow +0, y_i \in R_{\varepsilon_i}(x): \varphi'_x(x, y_i) \rightarrow v\}.$$

Заметим, что множество $R_\varepsilon(x)$ не обязательно замкнуто. Множество $\mathcal{H}(x_0)$ непусто, замкнуто и ограничено. Точечно-множественное отображение $\mathcal{H}(x)$ полунепрерывно сверху в точке x_0 , т. е. из того, что $v_i \in \mathcal{H}(x_i)$, $x_i \rightarrow x_0$, $v_i \rightarrow v$, следует $v \in \mathcal{H}(x_0)$.

Теорема 14.1. *Функция $f(x)$, заданная соотношением (14.1), является дифференцируемой по любому направлению $g \in E_n$ в точке x_0 , причем*

$$\frac{\partial f(x_0)}{\partial g} = \max_{v \in \mathcal{H}(x_0)} (v, g). \quad (14.3)$$

Доказательство. Положим

$$h(\alpha) = \alpha^{-1} [f(x_0 + \alpha g) - f(x_0)].$$

Пусть последовательность $\{\alpha_i\}$ такова, что $\alpha_i \rightarrow +0$,

$$h(\alpha_i) \longrightarrow \lim_{\alpha \rightarrow +0} h(\alpha). \quad (14.4)$$

Для любого $v \in \mathcal{H}(x_0)$ существуют последовательности $\{e_i\}$ и $\{y_i\}$ такие, что

$$e_i \rightarrow +0, \quad y_i \in R_{e_i}(x_0), \quad \varphi'_x(x_0, y_i) \rightarrow v. \quad (14.5)$$

Без ограничения общности можем считать, что

$$\alpha_i^{-1} e_i \xrightarrow{i \rightarrow \infty} 0. \quad (14.6)$$

Тогда

$$\begin{aligned} f(x_0 + \alpha_i g) &\geq \varphi(x_0 + \alpha_i g, y_i) = \\ &= \varphi(x_0, y_i) + \alpha_i (\varphi'_x(x_0 + \theta_i g, y_i), g), \end{aligned}$$

где $\theta_i \in (0, \alpha_i)$. В силу равностепенной непрерывности семейства функций $\varphi'_x(x, y)$ в точке x_0 имеем

$$f(x_0 + \alpha_i g) \geq \varphi(x_0, y_i) + \alpha_i (\varphi'_x(x_0, y_i), g) + o_i(\alpha_i), \quad (14.7)$$

где $\frac{o_i(\alpha_i)}{\alpha_i} \xrightarrow{i \rightarrow \infty} 0$. Поскольку $y_i \in R_{e_i}(x_0)$, то

$$f(x_0 + \alpha_i g) \geq f(x_0) - e_i + \alpha_i (\varphi'_x(x_0, y_i), g) + o_i(\alpha_i).$$

Отсюда

$$h(\alpha_i) \geq (\varphi'_x(x_0, y_i), g) + \frac{o_i(\alpha_i) - e_i}{\alpha_i}.$$

Учитывая (14.4) — (14.7), получаем $\lim_{\alpha \rightarrow +0} h(\alpha) \geq (v, g)$.

Это неравенство справедливо для всех $v \in \mathcal{H}(x_0)$. Поэтому

$$\lim_{\alpha \rightarrow +0} h(\alpha) \geq \max_{v \in \mathcal{H}(x_0)} (v, g). \quad (14.8)$$

Для любого $\varepsilon > 0$ существует $\alpha_0 > 0$ такое, что для всех $\alpha \in [0, \alpha_0]$ будет

$$f(x_0 + \alpha g) = \sup_{y \in R_\varepsilon(x_0)} \varphi(x_0 + \alpha g, y).$$

Так как

$$\varphi(x_0 + \alpha g, y) = \varphi(x_0, y) + \alpha(\varphi'_x(x_0, y), g) + o(\alpha, y),$$

где $\frac{o(\alpha, y)}{\alpha} \xrightarrow{\alpha \rightarrow +0} 0$ равномерно по $y \in G$, то

$$f(x_0 + \alpha g) \leq \sup_{y \in R_\varepsilon(x_0)} \varphi(x_0, y) + \\ + \alpha \sup_{y \in R_\varepsilon(x_0)} (\varphi'_x(x_0, y), g) + o(\alpha).$$

Но $\sup_{y \in R_\varepsilon(x_0)} \varphi(x_0, y) = f(x_0)$, поэтому

$$h(\alpha) \leq \sup_{y \in R_\varepsilon(x_0)} (\varphi'_x(x_0, y), g) + \frac{o(\alpha)}{\alpha}. \quad (14.9)$$

Пусть $\alpha_i \rightarrow +0$ и $h(\alpha_i) \rightarrow \overline{\lim}_{\alpha \rightarrow +0} h(\alpha)$. Отсюда и из (14.9) получаем

$$\overline{\lim}_{\alpha \rightarrow +0} h(\alpha) \leq \sup_{y \in R_\varepsilon(x_0)} (\varphi'_x(x_0, y), g). \quad (14.10)$$

Неравенство (14.10) справедливо для всех $\varepsilon > 0$.

Пусть $\varepsilon_k \rightarrow +0$ и $y_k \in R_{\varepsilon_k}(x_0)$ таковы, что

$$\sup_{y \in R_{\varepsilon_k}(x_0)} (\varphi'_x(x_0, y), g) \leq (\varphi'_x(x_0, y_k), g) + \delta_k,$$

где $\delta_k \xrightarrow{k \rightarrow \infty} +0$. Поскольку в силу (14.2) множество $\{\varphi'_x(x_0, y) | y \in G\}$ ограничено, то без ограничения общности можем считать, что $\varphi'_x(x_0, y_k) \rightarrow \bar{v}$. Очевидно, $\bar{v} \in \mathcal{H}(x_0)$. Из (14.10) заключаем, что

$$\overline{\lim}_{\alpha \rightarrow +0} h(\alpha) \leq (\bar{v}, g) \leq \max_{v \in \mathcal{H}(x_0)} (v, g). \quad (14.11)$$

Объединяя (14.8) и (14.11), получаем

$$\max_{v \in \mathcal{H}(x_0)} (v, g) \leq \lim_{\alpha \rightarrow +0} h(\alpha) \leq \overline{\lim}_{\alpha \rightarrow +0} h(\alpha) \leq \max_{v \in \mathcal{H}(x_0)} (v, g).$$

Отсюда следует существование $\frac{\partial f(x_0)}{\partial g} \equiv \lim_{\alpha \rightarrow +0} h(\alpha)$ и формула (14.3). Теорема доказана.

З а м е ч а н и е 1. Из доказательства ясно, что достаточно равномерной непрерывности функций φ и φ'_x по x в точке x_0 относительно $y \in R_{\varepsilon_1}(x_0)$, где $\varepsilon_1 > 0$ — любое фиксированное число.

З а м е ч а н и е 2. Ясно, что в формуле (14.3) можно написать

$$\frac{\partial f(x_0)}{\partial g} = \max_{v \in \text{co } \mathcal{H}(x_0)} (v, g).$$

2. Пусть

$$f(x) = \sup_{y \in G} \varphi(x, y),$$

где $x \in E_n$, $y \in E_p$, $G \subset E_p$ — произвольное множество, функция $\varphi(x, y)$ выпукла по x при любом фиксированном $y \in G$. Возьмем $g \in E_n$, $x_0 \in E_n$. Тогда

$$\varphi(x_0 + \alpha g, y) = \varphi(x_0, y) + \alpha \frac{\partial \varphi(x_0, y)}{\partial g} + o(\alpha, g, y), \quad (14.12)$$

где $\frac{o(\alpha, g, y)}{\alpha} \xrightarrow{\alpha \rightarrow +0} 0$,

$$\frac{\partial \varphi(x_0, y)}{\partial g} = \max_{v \in \partial \varphi(x_0, y)} (v, g), \quad (14.13)$$

а множество

$$\begin{aligned} \partial \varphi(x_0, y) = \{v \in E_n \mid \varphi(z, y) \geq \\ \geq \varphi(x_0, y) + (v, z - x_0) \quad \forall z \in E_n\} \end{aligned}$$

есть субдифференциал функции $\varphi(x, y)$ в точке $x_0 \in E_n$ при фиксированном $y \in G$.

Предположим, что множества $\partial \varphi(x_0, y)$ ограничены в совокупности, т. е.

$$\|v\| \leq K < \infty \quad \forall v \in \partial \varphi(x_0, y), \quad \forall y \in G. \quad (14.14)$$

Допустим также, что выполнено условие: в (14.12)

$\frac{o(\alpha, g, y)}{\alpha} \xrightarrow{\alpha \rightarrow +0} 0$ равномерно по $y \in R_{\varepsilon_0}(x_0)$, где $\varepsilon_0 > 0$ фиксировано.

Легко видеть, что функция $f(x)$ является выпуклой и потому (см. § 4) дифференцируемой по всем направлениям. Возьмем $\varepsilon > 0$, $\varepsilon \leq \varepsilon_0$. Введем множества

$$R_\varepsilon(x_0) = \{y \in G \mid f(x_0) - \varphi(x_0, y) \leq \varepsilon\},$$

$$\mathcal{H}(x_0) = \{v \in E_n \mid \exists \varepsilon_i \rightarrow +0, y_i \in R_{\varepsilon_i}(x_0),$$

$$v_i \in \partial \varphi(x_0, y_i): v_i \rightarrow v\}.$$

При сделанных предположениях множество $\mathcal{H}(x_0)$ не пусто, замкнуто и ограничено.

Теорема 14.2. *Функция $f(x)$ является дифференцируемой по любому направлению g в точке x_0 , причем*

$$\frac{\partial f(x_0)}{\partial g} = \max_{v \in \mathcal{H}(x_0)} (v, g). \quad (14.15)$$

Доказательство. Дифференцируемость $f(x)$ по направлениям вытекает, как уже отмечалось, из выпуклости функции $f(x)$. Нам остается лишь найти вид производной по направлению. Положим

$$h(\alpha) = \alpha^{-1} [f(x_0 + \alpha g) - f(x_0)].$$

Для любого $v \in \mathcal{H}(x_0)$ найдутся последовательности $\{\varepsilon_i\}$, $\{y_i\}$, $\{v_i\}$ такие, что $\varepsilon_i \rightarrow +0$, $y_i \in R_{\varepsilon_i}(x_0)$, $v_i \in \partial f(x_0, y_i)$, $v_i \rightarrow v$. Возьмем последовательность $\{\alpha_i\}$ такую, что $\alpha_i \rightarrow +0$. Как и в п. 1, считаем, что

$$\alpha_i^{-1} \varepsilon_i \xrightarrow{i \rightarrow \infty} 0. \quad (14.16)$$

Тогда в силу определения $\partial f(x_0, y_i)$

$$f(x_0 + \alpha_i g) \geq f(x_0 + \alpha_i g, y_i) \geq f(x_0, y_i) + \alpha_i (v_i, g). \quad (14.17)$$

Так как $y_i \in R_{\varepsilon_i}(x_0)$, то $f(x_0 + \alpha_i g) \geq f(x_0) - \varepsilon_i + \alpha_i (v_i, g)$. Поэтому $h(\alpha_i) \geq (v_i, g) - \alpha_i^{-1} \varepsilon_i$. Отсюда и из (14.16), (14.17) следует $\frac{\partial f(x_0)}{\partial g} \geq (v, g)$. Поскольку это неравенство справедливо для всех $v \in \mathcal{H}(x_0)$, то

$$\frac{\partial f(x_0)}{\partial g} \geq \max_{v \in \mathcal{H}(x_0)} (v, g). \quad (14.18)$$

Докажем неравенство противоположного знака.

Для любого $\varepsilon > 0$ при достаточно малых $\alpha > 0$ будет

$$f(x_0 + \alpha g) = \sup_{y \in R_\varepsilon(x_0)} \varphi(x_0 + \alpha g, y).$$

Из (14.12) имеем

$$f(x_0 + \alpha g) \leq \sup_{y \in R_\varepsilon(x_0)} \varphi(x_0, y) + \alpha \sup_{y \in R_\varepsilon(x_0)} \frac{\partial \varphi(x_0, y)}{\partial g} + o(\alpha).$$

Но $\sup_{y \in R_\varepsilon(x_0)} \varphi(x_0, y) = f(x_0)$, поэтому

$$h(\alpha) \leq \sup_{y \in R_\varepsilon(x_0)} \frac{\partial \varphi(x_0, y)}{\partial g} + \frac{o(\alpha)}{\alpha}.$$

Переходя к пределу при $\alpha \rightarrow +0$, получаем

$$\frac{\partial f(x_0)}{\partial g} \leq \sup_{y \in R_\varepsilon(x_0)} \frac{\partial \varphi(x_0, y)}{\partial g}. \quad (14.19)$$

Пусть $\varepsilon_k \rightarrow +0$ и $y_k \in R_{\varepsilon_k}(x_0)$ таковы, что

$$\sup_{y \in R_{\varepsilon_k}(x_0)} \frac{\partial \varphi(x_0, y)}{\partial g} \leq \frac{\partial \varphi(x_0, y_k)}{\partial g} + \delta_k,$$

где $\delta_k \xrightarrow[k \rightarrow \infty]{} +0$. Поскольку

$$\frac{\partial \varphi(x_0, y_k)}{\partial g} = \max_{v \in \partial \varphi(x_0, y_k)} (v, g) = (v_k, g_k),$$

то, без ограничения общности, в силу (14.14) считаем, что $v_k \rightarrow \bar{v}$. Очевидно, $\bar{v} \in \mathcal{H}(x_0)$. Поэтому из (14.19) заключаем, что

$$\frac{\partial f(x_0)}{\partial g} \leq (\bar{v}, g) \leq \max_{v \in \mathcal{H}(x_0)} (v, g). \quad (14.20)$$

Из (14.18) и (14.20) следует (14.15). Теорема доказана. Введем множество $\mathfrak{L}(x_0) = \text{co } \mathcal{H}(x_0)$.

Следствие 1. *Имеет место равенство*

$$\partial f(x_0) = \mathfrak{L}(x_0).$$

Действительно, из (14.15) видно, что

$$\frac{\partial f(x_0)}{\partial g} = \max_{v \in \mathfrak{L}(x_0)} (v, g). \quad (14.21)$$

С другой стороны,

$$\frac{\partial f(x_0)}{\partial g} = \max_{v \in \partial f(x_0)} (v, g), \quad (14.22)$$

где $\partial f(x_0) = \{v \in E_n \mid f(x) - f(x_0) \geq (v, x - x_0) \quad \forall x \in E_n\}$ — субдифференциал функции $f(x)$ в точке x_0 . Поскольку (14.21) и (14.22) справедливы для любых $g \in E_n$, а множества $\mathfrak{L}(x_0)$ и $\partial f(x_0)$ — выпуклые компакты, то (см. следствие из леммы 2.1) $\partial f(x_0) = \mathfrak{L}(x_0)$. ■

Следствие 2. Если $f(x) = \max_{y \in G} \varphi(x, y)$, где G — компакт, а функция $f(x, y)$ выпукла по x при любом фиксированном $y \in G$ и непрерывна по x , то

$$\partial f(x) = \text{co} \{ \varphi'_x(x, y) \mid y \in R(x) \},$$

$$R(x) = \{ y \in G \mid f(x) = \varphi(x, y) \}.$$

Пример 1. Пусть $x \in E_1$, $f(x) = \max_{i \in 1:2} f_i(x)$, где $f_1(x) = (x-1)^2$, $f_2(x) = (x+1)^2$. Возьмем точку $x_0 = 0$. Так как $f(x)$ — функция максимума, то (см. пример 2 § 5)

$$\frac{\partial f(x_0)}{\partial g} = \max_{i \in R(x_0)} (f'_i(x_0), g),$$

где $R(x_0) = \{ i \in 1:2 \mid f_i(x_0) = f(x_0) \} = 1:2$. Значит,

$$\frac{\partial f(x_0)}{\partial g} = \max \{ -2g, 2g \}.$$

Здесь $\|g\| = 1$. Но в E_1 только два таких направления:

$$g_1 = +1, \quad g_2 = -1.$$

Поэтому

$$\frac{\partial f(x_0)}{\partial g} = +2 \quad (14.23)$$

и для $g = g_1$, и для $g = g_2$.

С другой стороны, поскольку $f(x)$ — выпуклая функция, то (см. (5.11))

$$f(x) = \max_{y \in E_1} [f(y) + (v(y), x - y)], \quad (14.24)$$

где $v(y) \in \partial f(y)$ (для каждого y берется один $v(y)$ из $\partial f(y)$). Но

$$f(x) = \begin{cases} f_1(x), & \text{если } x \leq 0, \\ f_2(x), & \text{если } x \geq 0. \end{cases}$$

Поэтому

$$\partial f(x) = \begin{cases} \{2(x-1)\}, & \text{если } x < 0, \\ \{2(x+1)\}, & \text{если } x > 0, \end{cases}$$

и для $x = 0$ имеем $\partial f(0) = [-2, 2]$.

В формуле (14.24) возьмем $v(0) = 0$. Очевидно, что в (14.24) $R(0) = \{0\}$, поэтому $\max_{y \in R(0)} (v(y), g) = 0$. Отсюда ясно, что формула

$$\frac{\partial f(x_0)}{\partial g} = \max_{y \in R(x_0)} (v(y), g)$$

неверна, ибо противоречит (14.23).

В формуле (14.15) будет $\mathcal{H}(x_0) = \{-2, 2\}$.

§ 15. Дифференцируемость выпуклой функции

1. В настоящем параграфе изучается вопрос о дифференцируемости выпуклой функции, заданной на E_n . Оказывается, что множество точек, в которых эта функция имеет дифференциал, достаточно массивно. Вначале рассмотрим случай, когда $n = 1$.

Пусть функция $f(x)$ задана на прямой E_1 , $x_0 \in E_1$. Пределы

$$f'_+(x_0) = \lim_{\alpha \rightarrow +0} \alpha^{-1} [f(x_0 + \alpha) - f(x_0)]$$

и

$$f'_-(x_0) = \lim_{\alpha \rightarrow +0} \alpha^{-1} [f(x_0 + \alpha) - f(x_0)]$$

называются соответственно *левой* и *правой производными* функции $f(x)$ в точке x_0 . На прямой E_1 существует лишь два направления. Одно из них задается вектором $g_1 = +1$, другое — вектором $g_2 = -1$. Понятно, что

$$f'_+(x_0) = \lim_{\alpha \rightarrow +0} \alpha^{-1} [f(x_0 + \alpha g_1) - f(x_0)] = \frac{\partial f(x_0)}{\partial g_1},$$

$$f'_-(x_0) = \lim_{\alpha \rightarrow +0} \alpha^{-1} [f(x_0) - f(x_0 + \alpha g_2)] = - \frac{\partial f(x_0)}{\partial g_2}.$$

Если функция $f(x)$ выпукла, то, как следует из теоремы 4.2, производные по направлениям $\frac{\partial f(x)}{\partial g_1}$ и $\frac{\partial f(x)}{\partial g_2}$ существуют. Точнее, справедлива

Лемма 15.1. Если $f(x)$ — выпуклая функция, заданная на E_1 , то в каждой точке $x_0 \in E_1$ существуют левая и правая производные, причем

$$f'_+(x_0) = \frac{\partial f(x_0)}{\partial g_1}, \quad f'_-(x_0) = - \frac{\partial f(x_0)}{\partial g_2},$$

где $g_1 = +1$, $g_2 = -1$.

Приведем неравенства, связывающие левую и правую производные выпуклой функции.

Лемма 15.2. Пусть $f(x)$ — выпуклая на E_1 функция, $x_0, x_1 \in E_1$, причем $x_1 > x_0$. Тогда

$$f'_+(x_0) \leq (x_1 - x_0)^{-1} (f(x_1) - f(x_0)) \leq f'_-(x_1).$$

Доказательство. По следствию из теоремы 4.2

$$\frac{\partial f(x_0)}{\partial g_1} = \inf_{\alpha > 0} \alpha^{-1} [f(x_0 + \alpha) - f(x_0)],$$

$$\frac{\partial f(x_1)}{\partial g_2} = \inf_{\alpha > 0} \alpha^{-1} [f(x_1 - \alpha) - f(x_1)],$$

где $g_1 = +1, g_2 = -1$.

Полагая $\alpha = x_1 - x_0$, получим

$$f'_+(x_0) = \frac{\partial f(x_0)}{\partial g_1} \leq \alpha^{-1} [f(x_1) - f(x_0)] \leq -\frac{\partial f(x_1)}{\partial g_2} = f'_-(x_1),$$

что и требовалось доказать.

Лемма 15.3. Если $f(x)$ — выпуклая функция на E_1 и $x_0 \in E_1$, то

$$f'_-(x_0) \leq f'_+(x_0).$$

Доказательство. При $g_1 = +1, g_2 = -1$ имеем

$$f'_+(x_0) - f'_-(x_0) = \frac{\partial f(x_0)}{\partial g_1} + \frac{\partial f(x_0)}{\partial g_2} =$$

$$= \max_{v \in \partial f(x_0)} v + \max_{v \in \partial f(x_0)} (-v) \geq \max_{v \in \partial f(x_0)} (v + (-v)) = 0. \blacksquare$$

Вывясним теперь основные свойства односторонних производных.

Теорема 15.1. Если $f(x)$ — выпуклая на E_1 функция, то ее односторонние производные не убывают. При этом для всех $x \in E_1$ справедливы равенства

$$\lim_{y \rightarrow x+0} f'_+(y) = f'_+(x), \quad \lim_{y \rightarrow x-0} f'_-(y) = f'_-(x), \quad (15.1)$$

$$\lim_{z \rightarrow x-0} f'_+(z) = f'_-(x), \quad \lim_{z \rightarrow x+0} f'_-(z) = f'_+(x). \quad (15.2)$$

Доказательство. Пусть $y > x$. Из лемм 15.2 и 15.3 вытекают соотношения

$$f'_-(x) \leq f'_+(x) \leq (y - x)^{-1} (f(y) - f(x)) \leq f'_-(y) \leq f'_+(y). \quad (15.3)$$

Отсюда ясно, что функции $f'_+(x)$ и $f'_-(x)$ не убывают. Из непрерывности функции $f(x)$ следует

$$(y-x)^{-1}(f(y)-f(x)) = \lim_{t \rightarrow x+0} (y-t)^{-1}(f(y)-f(t)).$$

Так как по лемме 15.2

$$(y-t)^{-1}(f(y)-f(t)) \geq f'_+(t)$$

при $y > t$, то

$$(y-x)^{-1}(f(y)-f(x)) \geq \lim_{t \rightarrow x+0} f'_+(t)$$

(существование предела гарантируется монотонностью функции $f'_+(x)$). Поэтому

$$f'_+(x) = \lim_{y \rightarrow x+0} (y-x)^{-1}(f(y)-f(x)) \geq \lim_{t \rightarrow x+0} f'_+(t).$$

Заменив в последнем пределе t на y , получим окончательно

$$f'_+(x) \geq \lim_{y \rightarrow x+0} f'_+(y). \quad (15.4)$$

С другой стороны, как следует из (15.3),

$$f'_+(x) \leq f'_-(y) \leq f'_+(y),$$

и потому, переходя к пределам, которые существуют вследствие монотонности f'_+ и f'_- , имеем

$$f'_+(x) \leq \lim_{y \rightarrow x+0} f'_-(y) \leq \lim_{y \rightarrow x+0} f'_+(y). \quad (15.5)$$

Из (15.4) и (15.5) вытекают равенства (15.1). Аналогично устанавливаются и равенства (15.2). Теорема доказана.

Как известно, существование производной в некоторой точке равносильно (для функции одной переменной) тому, что правая и левая производные в этой точке совпадают (существование односторонних производных следует из леммы 15.1). Положим

$$\psi(x) = f'_+(x) - f'_-(x).$$

В силу леммы 15.3 $\psi(x) \geq 0$ при всех x . Из сказанного выше следует, что функция $f(x)$ не дифференцируема в точке x_0 тогда и только тогда, когда $\psi(x_0) > 0$.

Будем говорить, что $f(x)$ не дифференцируема в точке x_0 с точностью до ε (где $\varepsilon > 0$ фиксировано), если $\psi(x_0) \geq \varepsilon$. Геометрически это означает, что разность тангенсов углов наклона правой и левой касательных в точке x_0 не меньше ε .

Лемма 15.4. На любом ограниченном интервале (a, b) при каждом $\varepsilon > 0$ множество точек, в которых $f(x)$ не дифференцируема с точностью до ε , либо пусто, либо конечно.

Доказательство. Предположим, что указанное множество бесконечно. Тогда из его элементов можно выбрать монотонную последовательность $\{x_k\}$, сходящуюся к некоторой точке $x_0 \in [a, b]$. Определенности ради считаем, что последовательность $\{x_k\}$ убывает. Тогда, в силу теоремы 15.1,

$$\lim_{k \rightarrow \infty} (f'_+(x_k) - f'_-(x_k)) = \lim_{y \rightarrow x_0+0} f'_+(y) - \lim_{y \rightarrow x_0+0} f'_-(y) = 0,$$

что противоречит неравенству

$$f'_+(x_k) - f'_-(x_k) > \varepsilon \quad \forall k.$$

Лемма доказана.

Теорема 15.2. Выпуклая функция $f(x)$, заданная на E_1 , дифференцируема во всех точках E_1 , за исключением, быть может, не более чем счетного множества. Если функция $f(x)$ дифференцируема в некоторой точке x_0 , то и правая, и левая производные непрерывны в этой точке. В частности, если $f(x)$ дифференцируема на некотором интервале, то ее производная непрерывна там.

Доказательство. Положим

$$A_{mk} = \{x \in [-m, m] \mid \psi(x) > k^{-1}\}.$$

Тогда множество $A = \{x \in E_1 \mid \psi(x) > 0\}$, состоящее из всех точек, в которых $f(x)$ не дифференцируема, представимо в виде

$$A = \bigcup_{m=1}^{\infty} \bigcup_{k=1}^{\infty} A_{mk}.$$

В силу леммы 15.4 каждое из множеств A_{mk} конечно или пусто. Поэтому при каждом m множество $A_m = \bigcup_{k=1}^{\infty} A_{mk}$,

являющееся счетным объединением множеств A_{mk} , не более чем счетно. По тем же причинам множество

$$A = \bigcup_{m=1}^{\infty} A_m \text{ не более чем счетно.}$$

Если $f(x)$ дифференцируема в точке x_0 , то $f'_+(x_0) = f'_-(x_0)$, и потому все четыре предела, участвующие в формулировке теоремы 15.1, совпадают. Из этого, в частности, следует непрерывность левой и правой производных в точке x_0 . Если $f(x)$ дифференцируема на интервале, то ее производная совпадает с правой (или левой) производной на всем интервале и потому непрерывна. Теорема доказана.

2. Перейдем теперь к общему случаю. Пусть функция $f(x)$ задана на E_n (n — произвольное натуральное число). Как известно, функция $f(x)$, определенная на E_n , называется *дифференцируемой в точке* $x_0 \in E_n$, если существует вектор $A \in E_n$, при котором

$$\lim_{x \rightarrow x_0} \|x - x_0\|^{-1} [f(x) - f(x_0) - (A, x - x_0)] = 0$$

или, что то же самое,

$$f(x) = f(x_0) + (A, x - x_0) + o(\|x - x_0\|).$$

При этом вектор A называется *градиентом* функции $f(x)$ в точке x_0 и обозначается $f'(x_0)$. Хорошо известно, что

$$f'(x_0) = \left(\frac{\partial f(x_0)}{\partial x^{(1)}}, \dots, \frac{\partial f(x_0)}{\partial x^{(n)}} \right),$$

где

$$x = (x^{(1)}, \dots, x^{(n)}),$$

$\frac{\partial f(x_0)}{\partial x^{(i)}}$ — частная производная функции $f(x)$ по i -й переменной. Таким образом, из дифференцируемости следует существование частных производных. Обратное неверно: из существования частных производных еще не следует дифференцируемость функции. Как известно из курса анализа, наличие непрерывных частных производных уже влечет дифференцируемость. Покажем, что в случае выпуклой функции дифференцируемость равносильна существованию частных производных. Предварительно

отметим, что справедливы следующие очевидные равенства:

$$\frac{\partial f(x)}{\partial x^{(i)}} = \frac{\partial f(x)}{\partial e_i}, \quad -\frac{\partial f(x)}{\partial x^{(i)}} = \frac{\partial f(x)}{\partial (-e_i)}, \quad (15.6)$$

где e_i — i -й орт пространства E_n .

Теорема 15.3. Пусть $f(x)$ — выпуклая на E_n функция. Функция $f(x)$ дифференцируема в точке $x_0 \in E_n$ в том и только в том случае, когда она обладает частными производными $\frac{\partial f(x_0)}{\partial x^{(1)}}, \dots, \frac{\partial f(x_0)}{\partial x^{(n)}}$.

Если $f(x)$ дифференцируема в точке x_0 , то

$$\partial f(x_0) = \{f'(x_0)\}.$$

Доказательство. Если функция $f(x)$ дифференцируема, то она обладает частными производными. При этом для $g \in E_n$

$$f(x_0 + \alpha g) = f(x_0) + \alpha (f'(x_0), g) + o(\alpha).$$

Из леммы 5.3 теперь получаем

$$\partial f(x_0) = \{f'(x_0)\}.$$

Пусть существуют частные производные $\frac{\partial f(x_0)}{\partial x^{(1)}}, \dots, \frac{\partial f(x_0)}{\partial x^{(n)}}$. Покажем, что вектор $f'(x_0) = \left(\frac{\partial f(x_0)}{\partial x^{(1)}}, \dots, \frac{\partial f(x_0)}{\partial x^{(n)}} \right)$ является субградиентом функции $f(x)$ в точке x_0 . Для $v = (v^{(1)}, \dots, v^{(n)}) \in \partial f(x_0)$ имеем

$$\begin{aligned} \frac{\partial f(x_0)}{\partial x^{(i)}} &= \frac{\partial f(x_0)}{\partial e_i} = \max_{w \in \partial f(x_0)} (w, e_i) \geq \\ &\geq (v, e_i) = v^{(i)} \quad \forall i \in 1:n, \\ -\frac{\partial f(x_0)}{\partial x^{(i)}} &= \frac{\partial f(x_0)}{\partial (-e_i)} = \max_{w \in \partial f(x_0)} (w, -e_i) \geq \\ &\geq (-v, e_i) = -v^{(i)} \quad \forall i \in 1:n. \end{aligned}$$

Таким образом, $v^{(i)} = \frac{\partial f(x_0)}{\partial x^{(i)}}$, и, следовательно, $f'(x_0)$ — субградиент. Более того, $\partial f(x_0) = \{f'(x_0)\}$.

При всех $u \in E_n$ выполняется неравенство

$$f(x_0 + u) \geq f(x_0) + \sum_{i=1}^n \frac{\partial f(x_0)}{\partial x^{(i)}} u^{(i)}. \quad (15.7)$$

Здесь $u = (u^{(1)}, \dots, u^{(n)})$.

Проверим, что функция $f(x)$ дифференцируема в точке x_0 . Рассмотрим векторы $g_1, \dots, g_n, g_{n+1}, \dots, g_{2n}$, где $g_{n+i} = -e_i \quad \forall i \in 1:n, g_i = e_i \quad \forall i \in 1:n$. Положим

$$\begin{aligned} z_i &= \begin{cases} \max(u^{(i)}, 0), & i \in 1:n, \\ \max(-u^{(i)}, 0), & i \in (n+1):2n; \end{cases} \\ I_+ &= \{i \in 1:n \mid u^{(i)} \geq 0\}, \\ I_- &= \{i \in 1:n \mid u^{(i)} < 0\}. \end{aligned}$$

Тогда

$$\begin{aligned} u &= \sum_{i=1}^n u^{(i)} e_i = \sum_{i \in I_+} u^{(i)} e_i + \sum_{i \in I_-} u^{(i)} e_i = \\ &= \sum_{i=1}^n z^{(i)} g_i + \sum_{i=n+1}^{2n} z^{(i)} g_i = \sum_{i=1}^{2n} z^{(i)} g_i. \end{aligned}$$

Обозначим $\lambda(u)$ сумму $\sum_{i=1}^{2n} z^{(i)} = \sum_{i=1}^n u^{(i)}$. Из определения

сразу следует, что $\sum_{i=1}^{2n} \frac{z^{(i)}}{\lambda(u)} = 1$. Кроме того, $[\lambda(u)]^{-1} z^{(i)} \geq$

≥ 0 . Поэтому для любых векторов $v_1, \dots, v_{2n}$ будет

$$f\left(\sum_{i=1}^{2n} [\lambda(u)]^{-1} z^{(i)} v_i\right) \leq \sum_{i=1}^{2n} [\lambda(u)]^{-1} z^{(i)} f(v_i). \quad (15.8)$$

Нам понадобятся величины $o_i(\alpha)$, определяемые в точке x_0 из равенства

$$f(x_0 + \alpha g_i) = f(x_0) + \alpha \frac{\partial f(x_0)}{\partial g_i} + o_i(\alpha). \quad (15.9)$$

Так как $f(x)$ — выпуклая функция, то $o_i(\alpha) \geq 0$. Оценим теперь величину $f(x_0 + u)$, считая вектор u отличным от

нуля. Воспользуемся соотношениями (15.8) при $v_i = x_0 + \lambda(u)g_i$ и (15.9) при $\alpha = \lambda(u)$. Имеем

$$\begin{aligned}
 f(x_0 + u) &= f\left(x_0 + \sum_{i=1}^{2n} z^{(i)} g_i\right) = \\
 &= f\left(x_0 + \lambda(u) \sum_{i=1}^{2n} [\lambda(u)]^{-1} z^{(i)} g_i\right) = \\
 &= f\left(\sum_{i=1}^{2n} [\lambda(u)]^{-1} z^{(i)} (x_0 + \lambda(u) g_i)\right) \leqslant \\
 &\leqslant \sum_{i=1}^{2n} [\lambda(u)]^{-1} z^{(i)} f(x_0 + \lambda(u) g_i) = \\
 &= \sum_{i=1}^{2n} [\lambda(u)]^{-1} z^{(i)} \left(f(x_0) + \lambda(u) \frac{\partial f(x_0)}{\partial g_i} + o_i(\lambda(u))\right) = \\
 &= f(x_0) + \sum_{i=1}^{2n} z^{(i)} \frac{\partial f(x_0)}{\partial g_i} + \sum_{i=1}^{2n} o_i(\lambda(u)) [\lambda(u)]^{-1} z^{(i)}. \quad (15.10)
 \end{aligned}$$

Покажем, что последнее из трех слагаемых полученной суммы стремится к нулю при $\|u\| \rightarrow 0$ быстрее, чем $\|u\|$. Действительно,

$$\begin{aligned}
 \|u\|^{-1} \sum_{i=1}^{2n} [\lambda(u)]^{-1} z^{(i)} o_i(\lambda(u)) &= \\
 &= \sum_{i=1}^{2n} [\lambda(u)]^{-1} o_i(\lambda(u)) \|u\|^{-1} z^{(i)} \leqslant \\
 &\leqslant \sum_{i=1}^{2n} [\lambda(u)]^{-1} o_i(\lambda(u)) \xrightarrow{\|u\| \rightarrow 0} 0.
 \end{aligned}$$

Здесь воспользовались очевидным соотношением $z^{(i)} \leqslant |u^{(i)}| \leqslant \|u\|$ и неотрицательностью величины $o_i(\lambda(u))$.

Среднее слагаемое рассматриваемой суммы имеет вид

$$\sum_{i=1}^{2n} z^{(i)} \frac{\partial f(x_0)}{\partial g_i} = \sum_{i=1}^n u^{(i)} \frac{\partial f(x_0)}{\partial x^{(i)}}.$$

Это следует из равенств (15.6) и определения $z^{(i)}$. Поэтому из (15.10) вытекает

$$f(x_0 + u) \leq f(x_0) + \sum_{i=1}^n u^{(i)} \frac{\partial f(x_0)}{\partial x^{(i)}} + o(u). \quad (15.11)$$

Дифференцируемость функции $f(x)$ следует теперь из (15.7) и (15.11). Теорема доказана.

Теорема 15.4. *Выпуклая функция $f(x)$ дифференцируема в точке $x_0 \in E_n$ тогда и только тогда, когда ее субдифференциал $\partial f(x_0)$ состоит из единственной точки.*

Доказательство. Если функция $f(x)$ дифференцируема, то

$$\partial f(x_0) = \{f'(x_0)\}$$

состоит из единственной точки. Пусть $\partial f(x_0) = \{v\}$, $v = (v^{(1)}, \dots, v^{(n)})$. Тогда

$$\begin{aligned} \lim_{\alpha \rightarrow +0} \alpha^{-1} [f(x_0 + \alpha e_i) - f(x_0)] &= \\ &= \frac{\partial f(x_0)}{\partial e_i} = \max_{w \in \partial f(x_0)} (w, e_i) = (v, e_i) = v^{(i)}, \end{aligned}$$

$$\begin{aligned} \lim_{\alpha \rightarrow -0} \alpha^{-1} [f(x_0 + \alpha e_i) - f(x_0)] &= \\ &= \lim_{\beta \rightarrow +0} \beta^{-1} [f(x_0) - f(x_0 + \beta(-e_i))] = -\frac{\partial f(x_0)}{\partial(-e_i)} = \\ &= -\max_{w \in \partial f(x_0)} (w, -e_i) = -(v, -e_i) = v^{(i)}. \end{aligned}$$

Таким образом, существуют частные производные

$$\frac{\partial f(x_0)}{\partial x^{(i)}} = \lim_{\alpha \rightarrow 0} \alpha^{-1} [f(x_0 + \alpha e_i) - f(x_0)],$$

и, как следует из теоремы 15.3, функция $f(x)$ дифференцируема в точке x_0 . Теорема доказана.

Опишем структуру множества, на котором выпуклая функция $f(x)$ не дифференцируема. Обозначим это множество A и рассмотрим пересечение A_m множества A с открытым кубом

$$\mathcal{K}_m = \{x \in E_n \mid |x^{(i)}| < m \quad \forall i \in 1:n\},$$

где m — натуральное число. Обозначим A_{km} ($k \in 1:n$) множество тех точек $x \in \mathcal{K}_m$, в которых не существует

частная производная $\frac{\partial f(x)}{\partial x^{(k)}}$. Из теоремы 15.3 следует

$$A_m = A_{1m} \cup A_{2m} \cup \dots \cup A_{nm}.$$

Соотношение

$$A = \bigcup_{m=1}^{\infty} A_m = \bigcup_{m=1}^{\infty} \bigcup_{k=1}^n A_{km}$$

сводит изучение множества A к изучению множества A_{km} . Расщепим, в свою очередь, эти множества на просто устроенные части. Рассмотрим правую и левую частные производные $f'_{+k}(x_0)$ и $f'_{-k}(x_0)$ функции $f(x)$ в точке x_0 по k -й переменной. По определению

$$f'_{+k}(x_0) = \lim_{\alpha \rightarrow +0} \alpha^{-1} [f(x_0 + \alpha e_k) - f(x_0)],$$

$$f'_{-k}(x_0) = \lim_{\alpha \rightarrow -0} \alpha^{-1} [f(x_0 + \alpha e_k) - f(x_0)].$$

Если функция $h(\alpha)$ одной вещественной переменной определена соотношением $h(\alpha) = f(x_0 + \alpha e_k)$, то

$$f'_{+k}(x_0) = h'_+(0), \quad f'_{-k}(x_0) = h'_-(0).$$

Из выпуклости $f(x)$ сразу следует выпуклость $h(\alpha)$, и потому по лемме 15.3

$$f'_{-k}(x_0) \leq f'_{+k}(x_0).$$

Положим

$$\psi_k(x_0) = f'_{+k}(x_0) - f'_{-k}(x_0).$$

Тогда $\psi_k(x_0) \geq 0$, причем существование частной производной эквивалентно равенству $\psi_k(x_0) = 0$. Иными словами,

$$A_{km} = \{x \in \mathcal{H}_m \mid \psi_k(x) > 0\}.$$

Положим

$$A_{kmi} = \{x \in \mathcal{H}_m \mid \psi_k(x) \geq i^{-1}\}, \quad i \in 1 : \infty.$$

Тогда

$$A_{km} = \bigcup_{i=1}^{\infty} A_{kmi}.$$

Структуру множества A_{kmi} проще всего описать, выяснив, как устроено пересечение этого множества с каждой прямой, проходящей в направлении e_k . Для этого

рассмотрим гиперплоскость $\mathcal{H}_k = \{x \in E_n \mid (x, e_k) = 0\}$ и пересечение $\mathcal{H}_{km}$ этой гиперплоскости с кубом $\mathcal{K}_m$. Пусть

$$x = (x^{(1)}, \dots, x^{(k-1)}, 0, x^{(k+1)}, \dots, x^{(n)}) \in \mathcal{H}_{km}$$

и $\Pi_x = \{x + \lambda e_k \mid -\infty < \lambda < \infty\}$ — прямая, проходящая через x в направлении e_k . Положим $\varphi(\lambda) = f(x_\lambda)$, где

$$x_\lambda = x + \lambda e_k = (x^{(1)}, \dots, x^{(k-1)}, \lambda, x^{(k+1)}, \dots, x^{(n)}).$$

Тогда

$$\varphi'_+(\lambda) = f'_{+k}(x_\lambda), \quad \varphi'_-(\lambda) = f'_{-k}(x_\lambda),$$

и потому, как нетрудно проверить, пересечение множества A_{kmi} с прямой Π_x совпадает с совокупностью точек x_λ , для которых

$$\varphi'_+(\lambda) - \varphi'_-(\lambda) \geq i^{-1}.$$

В то же время пересечение всего куба $\mathcal{K}_m$ с прямой Π_x представляет собой интервал $\{x_\lambda \mid |\lambda| < m\}$. Из леммы 15.4 следует, что множество чисел λ из интервала $(-m, m)$, в которых функция $\varphi(x)$ не дифференцируема с точностью до $\varepsilon = i^{-1}$, конечно либо пусто. Отсюда сразу вытекает, что множество точек x_λ на интервале $\{x_\lambda \mid |\lambda| < m\}$, в которых

$$\varphi'_+(\lambda) - \varphi'_-(\lambda) = f'_{+k}(x_\lambda) - f'_{-k}(x_\lambda) \geq i^{-1},$$

также конечно или пусто. Иными словами, пересечение множества A_{kmi} с прямой Π_x для любого $x \in \mathcal{H}_{km}$ конечно или пусто.

Итак, множество A_{km} представимо в виде счетного объединения множеств A_{kmi} , пересечение каждого из которых с любой прямой, параллельной вектору e_k , конечно или пусто. Указанное свойство множеств A_{kmi} показывает, что они в некотором смысле «малы», являются разреженными, «тощими».

Совокупность множеств $\{A_{kmi} \mid k \in 1:n, m \in 1:\infty, i \in 1:\infty\}$ является счетной. Поэтому множество $A = \bigcup_{k=1}^n \bigcup_{m=1}^{\infty} \bigcup_{i=1}^{\infty} A_{kmi}$ точек, в которых функция $f(x)$ не дифференцируема, можно рассматривать как счетное объединение множеств, «малых» в указанном выше смысле.

Счетное объединение «малых» множеств и само «не слишком велико». В этом смысле совокупность точек, в которых функция $f(x)$ дифференцируема, достаточно «массивна».

Отметим еще, что если функция $f(x)$ дифференцируема на некотором открытом множестве, то ее градиент непрерывен на этом множестве.

3. Уточним приведенные выше рассуждения для читателя, знакомого с теорией меры (см., например [85]). Рассмотрим множества

$$B_{ki} = \{x \in E_n \mid f'_{+k}(x) - f'_{-k}(x) \geq i^{-1}\}, \quad k \in 1:n, \quad i \in 1:\infty.$$

Поскольку

$$\begin{aligned} f'_{+k}(x) - f'_{-k}(x) &= \frac{\partial f(x)}{\partial e_k} + \frac{\partial f(x)}{\partial (-e_k)} = \\ &= \max_{v \in \partial_+ f(x)} (v, e_k) + \max_{v \in \partial_- f(x)} (v, -e_k), \end{aligned}$$

то

$$B_{ki} = \{x \in E_n \mid \max_{v \in \partial_+ f(x)} (v, e_k) + \max_{v \in \partial_- f(x)} (v, -e_k) \geq i^{-1}\}.$$

Из этого равенства и полунепрерывности сверху отображения $\partial f(x)$ легко следует замкнутость B_{ki} . Действительно, пусть $x_j \in B_{ki}$, $x_j \rightarrow x_0$. Найдем такие векторы $v'_j \in \partial f(x_j)$ и $v''_j \in \partial f(x_j)$, чтобы выполнялись соотношения

$$\begin{aligned} \max_{v \in \partial f(x_j)} (v, e_k) &= (v'_j, e_k), \\ \max_{v \in \partial f(x_j)} (v, -e_k) &= (v''_j, -e_k). \end{aligned}$$

По следствию из теоремы 5.1 отображение $\partial f(x)$ ограничено в окрестности точки x_0 . Поэтому, не умаляя общности, считаем, что существуют пределы

$$v' = \lim_{j \rightarrow \infty} v'_j, \quad v'' = \lim_{j \rightarrow \infty} v''_j.$$

При этом $v' \in \partial f(x_0)$, $v'' \in \partial f(x_0)$. Так как

$$(v'_j, e_k) + (v''_j, -e_k) \geq i^{-1},$$

то

$$\max_{v \in \partial f(x_0)} (v, e_k) + \max_{v \in \partial f(x_0)} (v, -e_k) \geq (v', e_k) + (v'', -e_k) \geq i^{-1},$$

т. е. $x_0 \in B_{kl}$. Таким образом, B_{ki} — замкнутое и тем более измеримое множество. Равенство $A_{kmi} = B_{ki} \cap \mathcal{H}_m$ показывает, что и множество A_{kmi} измеримо. Найдем его меру. Пусть

$$A_{kmi}(x) = A_{kmi} \cap \Pi_x \quad \forall x \in \mathcal{H}_{km}.$$

Как показано выше, множество $A_{kmi}(x)$ конечно или пусто, и потому мера $\mu A_{kmi}(x)$ этого множества равна нулю. Равенство

$$\mu A_{kmi} = \int_{\mathcal{H}_{km}} \mu A_{kmi}(x) dx,$$

восстанавливающее меру A_{kmi} по мере его сечений $A_{kmi}(x)$, показывает, что и множество A_{kmi} имеет нулевую меру. Так как множество точек недифференцируемости A представимо в виде счетного объединения множеств A_{kmi} , то и $\mu A = 0$. Таким образом, доказана следующая

Теорема 15.5. Пусть $f(x)$ — выпуклая на E_n функция. Тогда $f(x)$ дифференцируема почти всюду на E_n .

4. Теперь можно получить следующий интересный результат. Пусть функция $f(x)$ задана на E_n . По теореме 15.5 функция $f(x)$ дифференцируема почти всюду на E_n . Как и выше, обозначим A множество точек, в которых функция $f(x)$ не является дифференцируемой.

Теорема 15.6. Имеет место представление

$$\partial f(x_0) = \text{co} \{z \in E_n \mid z = \lim_{k \rightarrow \infty} f'(x_k), \quad x_k \rightarrow x_0, x_k \notin A\}. \quad (15.12)$$

Доказательство. Множество в правой части (15.12) обозначим B . Включение

$$B \subset \partial f(x_0) \quad (15.13)$$

очевидно из полунепрерывности сверху отображения $\partial f(x)$, ибо для $x_k \notin A$ по теореме 15.3 будет $\partial f(x_k) = \{f'(x_k)\}$. Установим включение противоположного знака. Пусть $v \in \partial f(x_0)$. Надо показать, что $v \in B$. Допустим противное. Тогда найдутся такие число $a > 0$ и вектор $g_0 \in E_n$, $\|g_0\| = 1$, что

$$(v - z, g_0) \geq a > 0 \quad \forall z \in B. \quad (15.14)$$

По следствию из теоремы 5.1 отображение $\partial f(x)$ ограничено на любом ограниченном множестве. Поскольку для $x \notin A$ будет $\partial f(x) = \{f'(x)\}$, а мера множества A равна нулю, то найдутся последовательность $\{x_k\}$ и точка $z_0 \in B$ такие, что

$$\begin{aligned} x_k &= x_0 + \alpha_k g_k, \quad \alpha_k \rightarrow +0, \quad g_k \rightarrow g_0, \\ x_k &\notin A, \quad z_k = f'(x_k) \rightarrow z_0 \in B. \end{aligned} \quad (15.15)$$

Так как $v \in \partial f(x_0)$, то

$$f(x) - f(x_0) \geq (v, x - x_0) \quad \forall x \in E_n.$$

Отсюда и из (15.14) при достаточно больших k в силу ограниченности B (ограниченность B очевидна из (15.13))

$$\begin{aligned} h(\alpha_k) &= \alpha_k^{-1} (f(x_0 + \alpha_k g_k) - f(x_0)) \geq \\ &\geq (v, g_k) \geq (z, g_k) + 0,5a \quad \forall z \in B. \end{aligned}$$

В частности, $h(\alpha_k) \geq (z_0, g_k) + 0,5a$. Из (15.15) при достаточно больших k имеем

$$h(\alpha_k) \geq (z_k, g_k) + 0,25a. \quad (15.16)$$

Так как $\{z_k\} = \partial f(x_k)$, то

$$f(x) - f(x_k) \geq (z_k, x - x_k) \quad \forall x \in E_n.$$

Отсюда

$$\alpha_k^{-1} (f(x_0) - f(x_k)) \geq (-z_k, g_k),$$

т. е.

$$\begin{aligned} h(\alpha_k) &\geq \alpha_k^{-1} (f(x_0 + \alpha_k g_k) - f(x_0)) = \\ &= \alpha_k^{-1} (f(x_k) - f(x_0)) \leq (z_k, g_k), \end{aligned}$$

что противоречит (15.16). Итак, $v \in B$, т. е. $\partial f(x_0) \subset B$. Отсюда и из (15.16) вытекает утверждение теоремы. ■

§ 16*. Сопряженные функции

Удобным аппаратом для исследования свойств выпуклых функций и выпуклых множеств являются сопряженные функции. Многие изложенные ранее результаты можно получить на языке этих функций. Ниже приводятся первоначальные сведения из теории сопряженных функций.

Выпуклая функция f , определенная на выпуклом множестве $\Omega \subset E_n$, называется *замкнутой*, если ее надграфик

$$\text{epi } f = \{(\lambda, x) \in E_1 \times \Omega \mid \lambda \geq f(x)\}$$

является замкнутым множеством. Нетрудно проверить, что каждая непрерывная выпуклая функция, определенная на замкнутом выпуклом множестве, является замкнутой. Обратное, однако, неверно. В частности, замкнутая функция может быть определена и на незамкнутом множестве. В качестве примера приведем функцию $f(x) = \lg x$, определенную на промежутке $[0, \pi/2)$. Отметим, что

$$\lim_{x \rightarrow \pi/2 - 0} \lg x = +\infty.$$

Как показывает следующая ниже лемма, подобным свойством обладают все замкнутые функции, определенные на незамкнутом множестве.

Лемма 16.1. Пусть f — выпуклая замкнутая функция, определенная на множестве Ω , и x — предельная точка Ω , причем $x \notin \Omega$. Тогда

$$\lim_{y \rightarrow x} f(y) = +\infty.$$

Доказательство. Предположим, что лемма неверна. Тогда найдется последовательность $y_k \rightarrow x$, $y_k \in \Omega$, такая, что для некоторого числа λ выполняется неравенство

$$f(y_k) \leq \lambda \quad \forall k.$$

Это означает, что $(\lambda, y_k) \in \text{epi } f$. Так как $(\lambda, y_k) \rightarrow (\lambda, x)$ и множество $\text{epi } f$ замкнуто, то $(\lambda, x) \in \text{epi } f$. Отсюда по определению надграфика получаем $x \in \Omega$, что противоречит условию леммы. ■

Задание 16.1 Показать, что выпуклая функция f , определенная на выпуклом множестве Ω , замкнута в том и только в том случае, когда она обладает следующими свойствами: а) f полунепрерывна снизу в каждой точке $x \in \Omega$, т. е. $f(x) \leq \liminf_{y \rightarrow x} f(y)$, если $x_k \in \Omega$, $x_k \rightarrow x$; б) $\lim_{y \rightarrow x} f(y) = +\infty$, если x — предельная точка Ω , не входящая в Ω .

Напомним, что если Ω имеет непустую внутренность, то f непрерывна во внутренних точках Ω . В граничных же точках замкнутая функция может иметь разрывы.

Задание 16.2. Привести пример выпуклой незамкнутой функции, определенной на сегменте $[0, 1] \subset E_1$.

Пусть f — выпуклая функция, определенная на выпуклом множестве Ω . Множество

$$U_f = \{(\mu, l) \in E_1 \times E_n \mid (l, x) - \mu \leq f(x) \quad \forall x \in \Omega\}$$

называется *опорным множеством* функции f . Можно сказать, что элементы опорного множества определяют линейные функции $h(x) = (l, x) - \mu$, не превышающие функцию f . Отметим, что для каждого субградиента v функции f в некоторой точке $x_0 \in \Omega$ найдется число μ такое, что $(\mu, v) \in U_f$. В самом деле, по определению субградиента

$$f(x) - f(x_0) \geq (v, x) - (v, x_0),$$

т. е. при $\mu \geq (v, x_0) - f(x_0)$ выполняется $(v, x) - \mu \leq f(x)$. Это и означает справедливость включения $(\mu, v) \in U_f$.

Установим связь выпуклой замкнутой функции f и ее опорного множества U_f .

Лемма 16.2. Если f — выпуклая замкнутая функция, определенная на выпуклом множестве $\Omega \subset E_n$, то

$$f(x) = \sup_{(\mu, l) \in U_f} \{(l, x) - \mu\} \quad \forall x \in \Omega. \quad (16.1)$$

Доказательство. Для любого $\varepsilon > 0$ рассмотрим точку $(f(x) - \varepsilon, x)$ пространства $E_1 \times E_n$. Так как эта точка не входит в замкнутое множество $\text{epi } f$, то, применяя теорему отделимости, можно указать вектор $(c, h) \in E_1 \times E_n$, при котором выполняется неравенство $c(f(x) - \varepsilon) + (h, x) < c\lambda + (h, z) \quad \forall (\lambda, z) \in \text{epi } f. \quad (16.2)$

Полагая здесь $z = x$, $\lambda = f(x)$, получим $-c\varepsilon < 0$, т. е. $c > 0$. Поэтому из (16.2) при $\lambda = f(z)$ вытекает неравенство

$$f(z) \geq (-c^{-1}h, x) + f(x) - \varepsilon + c^{-1}(h, x) \quad \forall z \in \Omega.$$

Положим $-c^{-1}h = l_0$, $f(x) - \varepsilon + c^{-1}(h, x) = -\mu_0$. Тогда $(h, z) \geq (l_0, z) - \mu_0 \quad \forall z \in \Omega$,

т. е. вектор (μ_0, l_0) входит в U_f . В то же время $(l_0, x) - \mu_0 = f(x) - \varepsilon$. Поэтому

$$\sup_{(\mu, l) \in U_f} \{(l, x) - \mu\} \geq (l_0, x) - \mu_0 = f(x) - \varepsilon.$$

Ввиду произвольности ε получаем $\sup_{(\mu, l) \in U_f} (l, x) - \mu \geq \geq f(x)$. Противоположное неравенство вытекает непосредственно из определения опорного множества. Лемма доказана.

Замечание 1. Если x_0 — внутренняя точка множества Ω , то найдется такой элемент (μ_0, l_0) опорного множества, что $f(x_0) = (l_0, x_0) - \mu_0$, т. е. супремум в формуле (16.1) реализуется. Это сразу следует из формулы (5.13). Если x_0 — граничная точка Ω , то такого элемента может и не найтись.

Задание 16.3. Пусть $\Omega = \{x = (x^{(1)}, x^{(2)}) \in E_2 \mid x^{(1)} \geq 0, x^{(2)} \geq 0\}$, $f(x) = -\sqrt{x^{(1)}x^{(2)}}$. Показать, что для точки $x_0 = (1, 0)$ не существует такого элемента $(\mu_0, l_0) \in U_f$, при котором выполняется равенство $f(x_0) = (l_0, x_0) - \mu_0$.

Для дальнейшего понадобится вычислить супремум по опорному множеству при любом элементе $x \in E_n$, не обязательно входящем в Ω .

Лемма 16.3. Пусть f — выпуклая замкнутая функция, определенная на выпуклом множестве Ω . Тогда, если $x \notin \Omega$, то

$$\sup_{(\mu, l) \in U_f} \{(l, x) - \mu\} = +\infty.$$

Доказательство. Рассмотрим отдельно два случая.

1. Пусть x — предельная точка Ω . Возьмем некоторое число λ . Поскольку, как установлено в лемме 16.1, $\lim_{y \rightarrow x} f(y) = +\infty$, то найдется выпуклая окрестность V точки x , для элементов y которой ($y \in \Omega$) выполняется неравенство $f(y) > \lambda$. Возьмем точку $y_0 \in V \cap \Omega$ и рассмотрим в пространстве $E_1 \times E_n$ отрезок $\mathfrak{L}$ с концами в точках (λ, x) и (λ, y) . Покажем, что этот отрезок не пересекается с надграфиком $\text{epi } f$ функции f . Действительно, если $(v, v) \in \mathfrak{L}$, то $v = \lambda$, $v = \alpha x + (1 - \alpha)y$ при некотором $\alpha > 0$. Так как $x \in V$, $y \in V$, то и $v \in V$, поэтому $f(v) > \lambda$. Это означает, что $(\lambda, v) \notin \text{epi } f$. Так как множества $\mathfrak{L}$ и $\text{epi } f$ не пересекаются и замкнуты и, кроме того, отрезок $\mathfrak{L}$ ограничен, то, применяя теорему отделимости, найдем вектор $(c, h) \in E_1 \times E_n$, обладающий

следующими свойствами:

$$cv + (h, z) > c\lambda + (h, x) \quad \forall (v, z) \in \text{epi } f, \quad (16.3)$$

$$cv + (h, z) > c\lambda + (h, y) \quad \forall (v, z) \in \text{epi } f. \quad (16.4)$$

Полагая в (16.4) $z = y$, $v = f(y)$, получим $cf(y) > c\lambda$, откуда вытекает неравенство $c > 0$. Полагая в (16.3) $v = f(z)$, придем к неравенству

$$f(z) > (-c^{-1}h, z) + \lambda + (c^{-1}h, x) \quad \forall z \in \Omega.$$

Пусть $l_0 = -c^{-1}h$, $-\mu_0 = \lambda + c^{-1}(h, x)$. Тогда $f(z) > (l_0, z) - \mu_0$, т. е. $(\mu_0, l_0) \in U_f$. В то же время $(l_0, x) - \mu_0 = \lambda$. Таким образом,

$$\sup_{(\mu, l) \in U_f} \{(l, x) - \mu\} \geq (l_0, x) - \mu_0 = \lambda.$$

Ввиду произвольности λ получаем отсюда, что

$$\sup_{(\mu, l) \in U_f} \{(l, x) - \mu\} = +\infty.$$

2. Пусть $x \notin \bar{\Omega}$, где $\bar{\Omega}$ — замыкание Ω . Зафиксируем некоторую точку $y \in \Omega$ и рассмотрим отрезок $\{\alpha x + (1 - \alpha)y \mid \alpha \in [0, 1]\}$ с концами в точках x и y . Так как $x \notin \bar{\Omega}$, то найдется число $\alpha_0 \in [0, 1]$ такое, что точки $\alpha x + (1 - \alpha)y$ не входят в Ω при $\alpha > \alpha_0$. Возьмем теперь некоторое число λ и покажем, что найдется такое число δ , при котором отрезок $\mathcal{E}$ в пространстве $E_1 \times E_n$ с концами в точках (λ, x) и (δ, y) не пересекается с надграфиком $\text{epi } f$ функции f . В качестве δ можно взять любое число, меньшее

$$\delta_0 = \inf_{0 \leq \alpha \leq \alpha_0} (1 - \alpha)^{-1} [f(\alpha x + (1 - \alpha)y) - \alpha\lambda]$$

(заметим, что $\delta_0 > -\infty$). Действительно, если $\mathcal{E} \cap \text{epi } f \neq \emptyset$, то найдется такое $\alpha \in [0, \alpha_0]$, при котором

$$\alpha\lambda + (1 - \alpha)\delta \geq f(\alpha x + (1 - \alpha)y).$$

Отсюда следует, что $\delta \geq \delta_0$. Итак, существует число δ , при котором множества $\mathcal{E}$ и $\text{epi } f$ не пересекаются. Применяя теорему отделимости, найдем вектор $(c, h) \in E_1 \times E_n$, обладающий теми свойствами, что

$$cv + (h, z) > c\lambda + (h, x) \quad \forall (v, z) \in \text{epi } f, \quad (16.5)$$

$$cv + (h, z) > c\delta + (h, y) \quad \forall (v, z) \in \text{epi } f. \quad (16.6)$$

Полагая в (16.6) $z = y$, получим $c(v - \delta) > 0$. Так как $v - \delta > 0$ при достаточно больших v , то $c > 0$. Пусть $l_0 = -c^{-1}h$, $-\mu_0 = \lambda + (c^{-1}h, x)$. Тогда, полагая, в (16.5) $v = f(z)$, получим $f(z) > (l_0, z) - \mu_0$, т. е. $(\mu, l_0) \in U_f$. В то же время $(l_0, x) - \mu = \lambda$. Таким образом, $\sup_{(\mu, l) \in U_f} \{(l, x) - \mu\} > \lambda$, что ввиду произвольности λ и доказывает лемму.

Леммы 16.2 и 16.3 показывают, что удобно рассматривать выпуклые замкнутые функции, определенные на всем пространстве и принимающие, возможно, значение $+\infty$. В самом деле, для таких функций равенство $f(x) = \sup_{(\mu, l) \in U_f} \{(l, x) - \mu\}$ справедливо при всех x .

Введем соответствующие определения. Функция f , определенная на всем пространстве E_n и принимающая значения в полурастretched числовой прямой $(-\infty, +\infty]$, называется *выпуклой*, если

$$f(\alpha x + (1 - \alpha)y) \leq \alpha f(x) + (1 - \alpha)f(y)$$

для всех $x, y \in E_n$, $\alpha \in [0, 1]$. Предполагается, что $0 \cdot (+\infty) = +\infty$. Множество

$$\text{dom } f = \{x \in E_n \mid f(x) < +\infty\}$$

называется *эффективным множеством* функции f . Если $\text{dom } f \neq \emptyset$, то f называется *собственной выпуклой* функцией. Как и для функций, принимающих конечные значения, множества

$$\text{epi } f = \{(\lambda, x) \in E_1 \times E_n \mid f(x) \geq \lambda\}$$

и

$$U_f = \{(\mu, l) \in E_1 \times E_n \mid (l, x) - \mu \leq f(x) \quad \forall x \in E_n\}$$

называются соответственно *надграфиком* и *опорным множеством* функции f . Если надграфик $\text{epi } f$ замкнут, то функция f называется *замкнутой*.

Пусть f — выпуклая функция, определенная на множестве Ω и принимающая лишь конечные значения. Положим

$$\bar{f}(x) = \begin{cases} f(x), & x \in \Omega, \\ +\infty, & x \notin \Omega. \end{cases}$$

Тогда $\text{dom } \bar{f} = \Omega$, $\text{epi } \bar{f} = \text{epi } f$, $U_{\bar{f}} = U_f$. Понятно, что изучение функции f сводится к изучению функции $\bar{f}$. Так

как $\text{epi } \bar{f} = \text{epi } f$, то замкнутость функции f эквивалентна замкнутости функции $\bar{f}$.

Всюду ниже в этом параграфе, говоря о функции f , считаем, не оговаривая этого особо, что f определена на всем пространстве и принимает, возможно, значение $+\infty$, однако f не равна тождественно $+\infty$, т. е. f — собственная функция.

Леммы 16.2 и 16.3 выражают основное свойство рассматриваемых функций. Сформулируем его в виде теоремы.

Теорема 16.1. Пусть f — выпуклая замкнутая функция. Тогда для всех $x \in E_n$ выполняется равенство

$$f(x) = \sup_{(\mu, l) \in U_f} \{(l, x) - \mu\}.$$

Верно и обратное утверждение, которое приведем в следующем виде.

Теорема 16.2. Пусть функция f обладает тем свойством, что для некоторого множества $U \subset E_1 \times E_n$ и всех $x \in E_n$ выполняется равенство

$$f(x) = \sup_{(\mu, l) \in U} \{(l, x) - \mu\}.$$

Тогда функция f выпукла и замкнута.

Доказательство. 1. Функция f выпукла. Действительно, если $x, y \in E_n, \alpha \in [0, 1]$, то

$$\begin{aligned} f(\alpha x + (1 - \alpha)y) &= \\ &= \sup_{(\mu, l) \in U} \{(l, \alpha x + (1 - \alpha)y) - \mu\} = \\ &= \sup_{(\mu, l) \in U} \{\alpha[(l, x) - \mu] + (1 - \alpha)[(l, y) - \mu]\} \leq \\ &\leq \alpha \left[\sup_{(\mu, l) \in U} \{(l, x) - \mu\} \right] + \\ &+ (1 - \alpha) \left[\sup_{(\mu, l) \in U} \{(l, y) - \mu\} \right] = \alpha f(x) + (1 - \alpha)f(y). \end{aligned}$$

2. Функция f замкнута. Для каждого элемента (μ, l) множества U рассмотрим подмножество

$$\mathcal{H}_{\mu l} = \{(\lambda, x) \in E_1 \times E_n \mid (l, x) - \lambda \leq \mu\}$$

пространства $E_1 \times E_n$. Понятно, что $\mathcal{H}_{\mu l}$ — замкнутое множество (на самом деле $\mathcal{H}_{\mu l}$ — полупространство).

Покажем справедливость равенства

$$\text{epi } f = \bigcap_{(\mu, l) \in U} \mathcal{H}_{\mu l}. \quad (16.7)$$

Из этого равенства и будет следовать справедливость теоремы, ибо пересечение замкнутых множеств замкнуто.

Перейдем к доказательству (16.7). Пусть $(\lambda, x) \in \text{epi } f$ и $(\mu, l) \in U$. Тогда

$$\lambda \geq f(x) \geq (l, x) - \mu \quad \forall x \in E_n,$$

откуда вытекает включение $(\lambda, x) \in \mathcal{H}_{\mu l}$. Так как (λ, x) — произвольный элемент $\text{epi } f$, то $\text{epi } f \subset \mathcal{H}_{\mu l}$. Так как (μ, l) — произвольный элемент U , то

$$\text{epi } f \subset \bigcap_{(\mu, l) \in U} \mathcal{H}_{\mu l}.$$

Пусть теперь $(\lambda_0, x_0) \in \mathcal{H}_{\mu l}$ при всех $(\mu, l) \in U$. Тогда

$$\lambda_0 \geq \sup_{(\mu, l) \in U} \{(l, x) - \mu\} = f(x_0)$$

и, следовательно, $(\lambda_0, x_0) \in \text{epi } f$. Теорема доказана.

Задание 16.4. Пусть множество $U \subset E_1 \times E_n$ обладает тем свойством, что $f(x) = \sup_{(\mu, l) \in U} \{(l, x) - \mu\}$. Показать, что в этом случае выпуклая замкнутая оболочка U совпадает с опорным множеством U_f функции f .

Перейдем к определению сопряженной функции. Предварительно докажем следующую лемму.

Лемма 16.4. Множество U в пространстве $E_1 \times E_n$ является надграфиком некоторой выпуклой замкнутой функции f в том и только в том случае, когда U непусто, выпукло, замкнуто, с каждой своей точкой (λ, x) содержит «вертикальный» луч $\{(\lambda', x) \mid \lambda' \geq \lambda\}$ и, кроме того, не содержит вертикальных прямых $\{(\lambda, x) \mid -\infty < \lambda < +\infty\}$. При этом для всех $x \in E_n$ выполняется равенство

$$f(x) = \inf \{\lambda \mid (\lambda, x) \in U\}. \quad (16.8)$$

(Как обычно, считаем, что $\inf$ пустого множества равен $+\infty$.)

Доказательство. То обстоятельство, что надграфик $\text{epi } f$ выпуклой замкнутой функции f обладает указанными в лемме свойствами, очевидно. Пусть теперь множество U обладает этими свойствами. Определим

функцию f формулой (16.8). Так как U не содержит вертикальных прямых, то $f(x) > -\infty$ при всех x . Непосредственно проверяется, что надграфик $\text{epi } f$ совпадает с U . При этом эффективное множество $\text{dom } f$ совпадает с совокупностью таких $x \in E_n$, что прямая $\{\lambda x | \lambda \in (-\infty, +\infty)\}$ пересекает множество U . Поэтому из непустоты U следует, что $\text{dom } f \neq \emptyset$. Из выпуклости и замкнутости множества $U = \text{epi } f$ вытекает выпуклость и замкнутость функции f . Лемма доказана.

С каждой выпуклой замкнутой функцией f связано два множества — надграфик $\text{epi } f$ и опорное множество

$$U_f = \{(\mu, l) \in E_1 \times E_n \mid f(x) \geq (l, x) - \mu \quad \forall x \in E_n\}.$$

Непосредственно из определения следует, что множество U_f выпукло и замкнуто. В силу теоремы 16.1 это множество непусто. Если $(\mu, l) \in U_f$, $\mu' \geq \mu$, то и $(\mu', l) \in U_f$. Кроме того, U_f не содержит вертикальных прямых. Действительно, если $x_0 \in \text{dom } f$ и $(\mu, l) \in U_f$, то $\mu \geq (l, x_0) - f(x_0) > -\infty$. Как следует из леммы 16.4, опорное множество U_f является надграфиком некоторой выпуклой замкнутой функции. Эта функция называется сопряженной к f и обозначается f^* . Итак, сопряженная функция f^* определяется равенством

$$\text{epi } f^* = U_f. \quad (16.9)$$

Приведем явное выражение этой функции. Используя формулу (16.8), имеем

$$\begin{aligned} f^*(l) &= \inf \{ \mu \mid (\mu, l) \in U_f \} = \\ &= \inf \{ \mu \mid (l, x) - \mu \leq f(x) \quad \forall x \in E_n \} = \\ &= \inf \{ \mu \mid \mu \geq (l, x) - f(x) \quad \forall x \in E_n \}. \end{aligned}$$

Если множество чисел $\{(l, x) - f(x) \mid x \in E_n\}$ ограничено сверху, то $f^*(l) = \sup_{x \in E_n} [(l, x) - f(x)]$. В противном случае

$$\begin{aligned} \text{множество } \{ \mu \mid \mu \geq (l, x) - f(x) \quad \forall x \in E_n \} &\text{ пусто и потому} \\ \inf \{ \mu \mid \mu \geq (l, x) - f(x) \quad \forall x \in E_n \} &= \\ &= +\infty = \sup \{ (l, x) - f(x) \mid x \in E_n \}. \end{aligned}$$

Таким образом, во всех случаях справедливо равенство

$$f^*(l) = \sup_{x \in E_n} [(l, x) - f(x)]. \quad (16.10)$$

Непосредственно из (16.10) вытекает следующее утверждение (*неравенство Юнга**): для любых $x, l \in E_n$ справедливо соотношение

$$(l, x) \leq f(x) + f^*(l).$$

Так как функция f^* выпукла и замкнута, то имеет смысл говорить о сопряженной к ней функции, которая обозначается символом f^{**} и называется *второй сопряженной* к f . Одно из основных утверждений теории сопряженных функций заключено в следующей теореме, называемой обычно *теоремой Фенхеля — Мора*.

Теорема 16.3. Если f — выпуклая замкнутая функция, то $f^{**} = f$.

Доказательство. Пусть $(\mu', l') \in U_f$. Тогда $(l', x) - f(x) \leq \mu'$ при всех $x \in E_n$ и потому $f^*(l') = \sup_{x \in E_n} [(l', x) - f(x)] \leq \mu'$. Отсюда следует, что

$$f^{**}(x) = \sup_{l \in E_n} [(l, x) - f^*(l)] \geq (l', x) - f^*(l') \geq (l', x) - \mu'.$$

Привлекая теорему 16.1, имеем

$$f^{**}(x) \geq \sup_{(\mu', l') \in U_f} [(l', x) - \mu'] = f(x).$$

С другой стороны, как непосредственно следует из неравенства Юнга,

$$f^{**}(x) = \sup_{l \in E_n} [(l, x) - f^*(l)] \leq f(x).$$

Теорема доказана.

З а м е ч а н и е 2. Из доказанной теоремы и (16.9) вытекает, что

$$U_{f^*} = \text{epi } f^{**} = \text{epi } f.$$

Таким образом, если U — множество в $E_1 \times E_n$, удовлетворяющее условиям леммы 16.4, то его можно трактовать двояко: либо как надграфик некоторой выпуклой замкнутой функции f , либо как опорное множество другой выпуклой замкнутой функции f^* . При этом по множеству U определяется сопряженное ему множество V ,

*) Иногда это неравенство называют *неравенством Фенхеля*.

которое является опорным к f и одновременно надграфиком f^* . Можно сказать, что в некотором смысле теория выпуклых замкнутых функций совпадает с теорией некоторых специальных выпуклых множеств, а именно множеств, описанных в лемме 16.4.

З а м е ч а н и е 3. Пусть f — произвольная, вообще говоря, невыпуклая и незамкнутая функция, определенная на E_n и принимающая значения в $(-\infty, +\infty)$. Формула (16.10) определяет функцию f^* , которая и в этом случае называется сопряженной к f . Положим

$$U = \{(\mu, x) \in E_1 \times E_n \mid x \in \text{dom } f, \mu = f(x)\}.$$

Так как $f^*(l) = \sup_{(\mu, x) \in U} [(l, x) - \mu]$, то в силу теоремы 16.2

f^* является выпуклой замкнутой функцией и в данном случае. Легко проверить, что функция f^{**} совпадает в рассматриваемом случае с «замкнутым овыпуклением» функции f , а именно с наибольшей выпуклой функцией, не превышающей f . Иными словами, надграфик $\text{epi } f^{**}$ функции f^{**} есть выпуклая замкнутая оболочка надграфика $\text{epi } f = \{(x, \lambda) \mid \lambda \geq f(x)\}$ функции f .

Приведем несколько примеров.

1. Пусть f — постоянная функция, $f(x) = c$ для всех $x \in E_n$. Тогда

$$\begin{aligned} f^*(l) &= \sup_{x \in E_n} [(l, x) - f(x)] = \\ &= \sup_{x \in E_n} [(l, x) - c] = \begin{cases} -c, & l = 0, \\ +\infty, & l \neq 0. \end{cases} \end{aligned}$$

В частности, если $f \equiv 0$, то

$$f^*(l) = \begin{cases} 0, & l = 0, \\ +\infty, & l \neq 0. \end{cases}$$

2. Пусть f — линейная функция, $f(x) = (h, x)$ для $x \in E_n$. Тогда

$$\begin{aligned} f^*(l) &= \sup_{x \in E_n} [(l, x) - (h, x)] = \\ &= \sup_{x \in E_n} [(l - h, x)] = \begin{cases} 0, & l = h, \\ +\infty, & l \neq h. \end{cases} \end{aligned}$$

3. Пусть U — выпуклое замкнутое подмножество пространства E_n . Введем в рассмотрение следующие две функции, связанные с U :

$$\delta_U(x) = \begin{cases} 0, & x \in U, \\ +\infty, & x \notin U, \end{cases}$$

$$P_U(x) = \sup_{l \in U} (l, x).$$

Обычно δ_U называется *индикаторной функцией* множества U , а P_U — *опорной функцией* этого множества. Понятно, что δ_U и P_U — выпуклые замкнутые функции. Найдем

$$\delta_U^*(l) = \sup_{x \in E_n} [(l, x) - \delta_U(x)] = \sup_{x \in U} (l, x) = P_U(l).$$

Таким образом, $\delta_U^* = P_U$. В силу теоремы 16.3 $P_U^* = \delta_U^{**} = \delta_U$.

Задание 16.5. Рассмотрим следующие выпуклые замкнутые функции, определенные на E_1 : $f_1(x) = e^x$, $f_2(x) = p^{-1}|x|^p$ ($p > 1$),

$$f_3(x) = \begin{cases} -0,5 - \ln x, & x > 0, \\ +\infty, & x \leq 0. \end{cases}$$

Показать, что

$$f_1^*(l) = \begin{cases} l \ln l - l, & l > 0, \\ 0, & l = 0, \\ +\infty, & l < 0, \end{cases}$$

$$f_2^*(l) = q^{-1} |l|^q, \quad \text{где } p^{-1} + q^{-1} = 1,$$

$$f_3^*(l) = f_3(-l) = \begin{cases} -0,5 - \ln(-l), & l < 0, \\ +\infty, & l \geq 0. \end{cases}$$

Сопряженные функции представляют собой удобный аппарат для изучения выпуклых функций. На их языке, в частности, удобно изучать субдифференциалы. При этом основную роль играет следующее утверждение.

Теорема 16.4. Пусть f — выпуклая замкнутая функция, $x_0, l_0 \in E_n$. Следующие условия эквивалентны:

- а) $l_0 \in \partial f(x_0)$;
- б) $x_0 \in \partial f^*(l_0)$;
- в) $(l_0, x_0) = f(x_0) + f^*(l_0)$

(соотношение в) показывает, что на элементах x_0 и l_0 в неравенстве Юнга реализуется равенство).

Доказательство. Пусть $l_0 \in \partial f(x_0)$. Тогда по определению субградиента имеем

$$(l_0, x_0) - f(x_0) \geq (l, x) - f(x).$$

Отсюда следует, что

$$f^*(l) = \sup_{x \in E_n} [(l, x) - f(x)] = (l_0, x_0) - f(x_0).$$

Наоборот, равенство $f^*(l) = (l_0, x_0) - f(x_0)$ влечет соотношение $(l_0, x_0) - f(x_0) \geq (l, x) - f(x)$. Это означает эквивалентность утверждений а) и в). Рассуждая таким же образом относительно функции f^* , получим, что включение $x_0 \in \partial f^*(l_0)$ равносильно равенству

$$f^{**}(x_0) = (l_0, x_0) - f^*(l_0).$$

Поскольку в силу теоремы 16.3 $f^{**} = f$, то условия б) и в) тоже эквивалентны. Теорема доказана.

В терминах сопряженных функций удобно формулировать экстремальные свойства выпуклой замкнутой функции f . Отметим прежде всего, что

$$f^*(0) = \sup_{x \in E_n} (-f(x)) = - \inf_{x \in E_n} f(x).$$

Таким образом,

$$\inf_{x \in E_n} f(x) = -f^*(0).$$

Отсюда следует, что f ограничена снизу в том и только в том случае, когда $0 \in \text{dom } f^*$. Пусть

$$\mathcal{D} = \{x \in E_n \mid f(x) = -f^*(0)\} = \{x \in E_n \mid f(x) = \inf_{y \in E_n} f(y)\}$$

— множество минимумов функции f .

Соотношение $x \in \mathcal{D}$ равносильно тому, что для пары элементов $x, 0$ реализуется равенство в неравенстве Юнга:

$$0 = (x, 0) = f(x) + f^*(0).$$

Поэтому, как следует из теоремы 16.4, соотношение $x \in \mathcal{D}$ равносильно включению $0 \in \partial f(x)$, и мы приходим к известному условию минимума (см. § 6).

С другой стороны, в силу той же теоремы соотношения $0 \in \partial f(x)$ и $x \in \partial f^*(0)$ равносильны. Следовательно, множество минимумов $\mathcal{D}$ выражается через сопряженную функцию f^* следующим образом:

$$\mathcal{D} = \partial f^*(0).$$

Отсюда следует, в частности, что функция f достигает минимума в том и только в том случае, когда сопряженная функция f^* имеет непустой субдифференциал в нуле (это можно гарантировать, например, в том случае, когда 0 — внутренняя точка эффективного множества $\text{dom } f^*$).

Дальнейшие применения сопряженных функций основаны на теоремах, позволяющих вычислять сопряженные функции от суммы, произведения на число, максимума выпуклых замкнутых функций и т. п. При этом приходится рассматривать не обязательно замкнутые функции, а также функции, принимающие, возможно, значение $-\infty$. С помощью указанных теорем на языке сопряженных функций можно изучать практически все задачи выпуклого анализа. Проиллюстрируем это на примере условного субдифференциала.

Пусть f — выпуклая замкнутая функция и Ω — выпуклое замкнутое множество, содержащееся в $\text{dom } f$. Рассмотрим индикаторную функцию δ_Ω множества Ω

$$\delta_\Omega(x) = \begin{cases} 0, & x \in \Omega, \\ +\infty, & x \notin \Omega, \end{cases}$$

и положим $f_1 = f + \delta_\Omega$. Понятно, что

$$f_1(x) = \begin{cases} f(x), & x \in \Omega, \\ +\infty, & x \notin \Omega, \end{cases}$$

т. е. $\text{dom } f_1 = \Omega$. Вычисление условного субдифференциала функции f на множестве Ω сводится поэтому к вычислению обычного субдифференциала суммы $f_1 = f + \delta_\Omega$, которое можно провести, используя технику сопряженных функций.

Подробное изложение многих результатов выпуклого анализа на основе теории сопряженных функций содержится в книге Р. Рокафеллара [110], а также в работах [57, 76].

§ 17. Вычисление ε -субградиентов некоторых классов выпуклых функций

Вычисление всего множества $\partial_\varepsilon f(x)$, $\varepsilon > 0$, — задача, несомненно, очень трудная, ибо, в силу определения $\partial_\varepsilon f(x)$, для этого необходима информация о глобальном поведении функции $f(x)$ на всем пространстве E_n . Что же касается вычисления отдельных ε -субградиентов функции $f(x)$, то, вообще говоря, эта задача более простая, чем вычисление субградиентов функции $f(x)$, поскольку $\partial f(x) \subset \partial_\varepsilon f(x)$ и множество $\partial_\varepsilon f(x)$ может оказаться богаче, чем $\partial f(x)$. Но и она достаточно трудная, ибо опять же требует знания поведения функции $f(x)$ на всем пространстве E_n . Ниже будет показано, как вычислять ε -субградиенты сильно выпуклых гладких функций и функций максимума сильно выпуклых гладких функций.

Лемма 17.1. Пусть функция $f(x)$ сильно выпукла на E_n с константой сильной выпуклости m и пусть $\varepsilon \geq 0$. Справедливо включение

$$\partial f(x) + S_{2\sqrt{\varepsilon m}}(0) \subset \partial_\varepsilon f(x), \quad (17.1)$$

где

$$S_\delta(0) = \{v \in E_n \mid \|v\| \leq \delta\}, \quad \delta \geq 0.$$

Доказательство. Пусть $v(x_0) \in \partial f(x_0)$, $r \in E_n$. Получим условия, при которых

$$(v(x_0) + r) \in \partial_\varepsilon f(x_0). \quad (17.2)$$

В силу определения $\partial_\varepsilon f(x_0)$ для того, чтобы выполнялось (17.2), необходимо и достаточно, чтобы

$$\begin{aligned} f(z) - f(x_0) &\geq \\ &\geq (v(x_0), z - x_0) + (r, z - x_0) - \varepsilon \equiv A(z) \quad \forall z \in E_n. \end{aligned} \quad (17.3)$$

Так как функция $f(x)$ сильно выпукла на E_n , то (см. (9.24))

$$f(z) - f(x_0) \geq (v(x_0), z - x_0) + m \|z - x_0\|^2 \equiv B(z) \quad \forall z \in E_n.$$

Отсюда вытекает, что (17.3) справедливо тогда, когда

$$A(z) \leq B(z) \quad \forall z \in E_n, \quad (17.4)$$

т. е. когда

$$m \|z - x_0\|^2 \geq (r, z - x_0) - \varepsilon \quad \forall z \in E_n.$$

Перепишем последнее неравенство следующим образом:

$$h(r, z) \equiv (r, z - x_0) - m \|z - x_0\|^2 \leq \varepsilon \quad \forall z \in E_n.$$

Теперь видно, что (17.4) будет выполнено, если

$$\max_{z \in E_n} h(r, z) \leq \varepsilon. \quad (17.5)$$

Найдем $\max_{z \in E_n} h(r, z)$. Для этого приравняем нулю производную функции $h(r, z)$ по z . Имеем $r - 2m(z - x_0) = 0$. Следовательно, $(z - x_0) = r \cdot 2^{-1} m^{-1}$. Поэтому $\max_{z \in E_n} h(r, z) = \|r\|^2 \cdot 4^{-1} m^{-1}$. Отсюда ясно, что неравенство

$$\|r\| \leq 2 \sqrt{\varepsilon m}$$

влечет за собой последовательно (17.5), (17.4), (17.3) и (17.2). Таким образом, включение (17.1) доказано.

Из леммы 17.1 следует, что y сильно выпуклой функции $f(x)$ любой субградиент $v(x) \in \partial f(x)$ вместе со своей окрестностью радиуса $2 \sqrt{\varepsilon m}$ принадлежит ε -субдифференциалу $\partial_\varepsilon f(x)$.

Следствие 1. Пусть $x, g \in E_n$, $\|g\| = 1$. Тогда

$$\frac{\partial f(x)}{\partial g} + 2 \sqrt{\varepsilon m} \leq \frac{\partial_\varepsilon f(x)}{\partial g}.$$

Действительно,

$$\begin{aligned} \frac{\partial_\varepsilon f(x)}{\partial g} &= \max_{v \in \partial_\varepsilon f(x)} (v, g) \geq \max_{v \in \partial f(x) + S_{2\sqrt{\varepsilon m}}(0)} (v, g) = \\ &= \max_{v \in \partial f(x)} (v, g) + \max_{v \in S_{2\sqrt{\varepsilon m}}(0)} (v, g) = \frac{\partial f(x)}{\partial g} + 2 \sqrt{\varepsilon m}. \end{aligned}$$

Следствие 2. Пусть $f(x) = 0,5(Ax, x) + (a, x)$, где $a \in E_n$, A — $n \times n$ -матрица, причем матрица A симметрична, положительно определена и m — ее минимальное собственное число. Тогда

$$(2Ax + a + r) \in \partial_\varepsilon f(x) \quad \forall r \in S_{2\sqrt{\varepsilon m}}(0).$$

Лемма 17.2. Пусть функция $f(x)$ дифференцируема на E_n и ее производная $f'(x)$ удовлетворяет на E_n условию Липшица с константой $L < \infty$, т. е.

$$\|f'(y) - f'(x)\| \leq L \|y - x\| \quad \forall x, y \in E_n.$$

Тогда

$$f(y) \leq f(x) + (f'(x), y - x) + 0,5L \|y - x\|^2 \quad \forall x, y \in E_n.$$

Доказательство. Для гладкой функции $f(x)$ справедливо разложение

$$f(y) = f(x) + \int_0^1 (f'(x + t(y - x)), y - x) dt.$$

Отсюда

$$\begin{aligned} f(y) - f(x) - (f'(x), y - x) &= \\ &= \int_0^1 (f'(x + t(y - x)) - f'(x), y - x) dt \leq \\ &\leq \int_0^1 \|f'(x + t(y - x)) - f'(x)\| \cdot \|y - x\| dt \leq \\ &\leq L \|y - x\|^2 \int_0^1 t dt = 0,5L \|y - x\|^2. \quad \blacksquare \end{aligned}$$

Пусть функция $f(x)$ сильно выпукла и дифференцируема на E_n и ее производная $f'(x)$ удовлетворяет на E_n условию Липшица, m — константа сильной выпуклости функции $f(x)$, L — константа Липшица производной функции $f(x)$. Положим

$$v_0^{(i)} = (f_v(x_0 + he_i) - f_v(x_0)) \cdot h^{-1}, \quad i \in 1:n, \quad (17.6)$$

где e_i — i -й координатный орт, $h > 0$, $v \geq 0$ — точность вычисления функции $f(x)$, т. е.

$$f(x) - v \leq f_v(x) \leq f(x) + v.$$

Лемма 17.3. Если

$$0,5Lh + h^{-1} \cdot 2v \leq 2\sqrt{\varepsilon mn^{-1}}, \quad (17.7)$$

то вектор $v_0 = (v_0^{(1)}, \dots, v_0^{(n)})$ с координатами (17.6) является ε -субградиентом функции $f(x)$ в точке x_0 , т. е. $v_0 \in \partial_\varepsilon f(x_0)$.

Доказательство. Из леммы 17.2 следует
 $|f(x + \Delta x) - f(x) - (f'(x), \Delta x)| \leq 0,5L \|\Delta x\|^2 \quad \forall x, \Delta x \in E_n.$
 (17.8)

Пусть $f'(x_0) = \left(\frac{\partial f(x_0)}{\partial x^{(1)}}, \dots, \frac{\partial f(x_0)}{\partial x^{(n)}} \right)$ — градиент функции $f(x)$ в точке x_0 . Учитывая, что

$$\frac{\partial f(x_0)}{\partial x^{(i)}} = (f'(x_0), e_i) \quad \forall i \in 1:n,$$

оценим $\left| v_0^{(i)} - \frac{\partial f(x_0)}{\partial x^{(i)}} \right|$. Имеем (см. (17.6))

$$\begin{aligned} |(f_v(x_0 + he_i) - f_v(x_0))h^{-1} - (f'(x_0), e_i)| &\leq \\ &\leq |f(x_0 + he_i) - f(x_0) - h(f'(x_0), e_i)|h^{-1} + 2vh^{-1}. \end{aligned}$$

Отсюда, согласно (17.8), получаем

$$\left| v_0^{(i)} - \frac{\partial f(x_0)}{\partial x^{(i)}} \right| \leq 0,5Lh + 2vh^{-1}.$$

Поэтому в силу (17.7)

$$\begin{aligned} \|v_0 - f'(x_0)\| &\leq \sqrt{n(0,5Lh + 2vh^{-1})^2} = \\ &= (0,5Lh + 2vh^{-1})\sqrt{n} \leq 2\sqrt{\varepsilon n}. \end{aligned}$$

Теперь из леммы 17.1 вытекает, что $v_0 \in \partial_\varepsilon f(x_0)$. Лемма доказана.

Таким образом, в лемме 17.2 дается способ построения ε -субградиента функции $f(x)$ в точке x_0 при использовании приближенных значений функции $f(x)$ в точках x_0 и $x_0 + he_i$, $i \in 1:n$. Отметим, что конструкция, описанная в лемме 17.2, весьма часто употребляется во многих практических вычислениях, когда нахождение аналитического вида градиента функции $f(x)$ оказывается сложным.

Остановимся теперь на вопросе вычисления ε -субградиента функции максимума выпуклых функций. Пусть

$$f(x) = \max_{y \in G} \varphi(x, y), \quad (17.9)$$

где $G \subset E_p$ — замкнутое ограниченное множество, функция $\varphi(x, y)$ непрерывна вместе с $\frac{\partial \varphi(x, y)}{\partial x}$ на $E_n \times G$ и выпукла по x на E_n при каждом $y \in G$.

Введем обозначения:

$$\begin{aligned} R_\varepsilon(x) &= \{y \in G \mid f(x) - \varphi(x, y) \leq \varepsilon\}, \quad \varepsilon \geq 0, \quad R(x) = R_0(x), \\ y(x) &\in R(x), \quad y_\varepsilon(x) \in R_\varepsilon(x), \\ \mathcal{L}_{\varepsilon\mu}(x) &= \text{co} \{ \partial_\mu \varphi(x, y) \mid y \in R_\varepsilon(x) \}, \\ \varepsilon &\geq 0, \quad \mu \geq 0, \quad \mathcal{L}(x) = \mathcal{L}_{00}(x). \end{aligned}$$

Известно (см. § 5), что $\partial f(x) = \mathcal{L}(x)$. Поэтому, если экстремальная задача $\max_{y \in G} \varphi(x, y)$ является достаточно сложной, то для вычисления функции $f(x)$ требуется бесконечный процесс последовательных приближений. Тем более это касается вычисления субградиентов функции $f(x)$, поскольку для нахождения $\frac{\partial \varphi(x, y)}{\partial x}$ в общем случае также требуется бесконечный процесс последовательных приближений.

Лемма 17.4. *Справедливо включение*

$$\mathcal{L}_{\varepsilon\mu}(x) \subset \partial_{\varepsilon+\mu} f(x) \quad \forall x \in E_n, \quad \forall \varepsilon \geq 0, \quad \mu \geq 0. \quad (17.10)$$

Доказательство. Зафиксируем $\varepsilon \geq 0, \mu \geq 0, x_0 \in E_n$. Пусть $y_\varepsilon(x_0) \in R_\varepsilon(x_0)$, т. е. $\varphi(x_0, y_\varepsilon(x_0)) \geq f(x_0) - \varepsilon$ и $v_\mu(x_0, y_\varepsilon(x_0)) \in \partial_\mu \varphi(x_0, y_\varepsilon(x_0))$.

Нужно показать, что

$$f(z) - f(x_0) \geq (v_\mu(x_0, y_\varepsilon(x_0)), z - x_0) - (\varepsilon + \mu) \quad \forall z \in E_n.$$

Используя определения $f(x)$ и $\partial_\mu \varphi(x, y)$, получаем

$$\begin{aligned} f(z) - f(x_0) &\geq \varphi(z, y(z)) - \varphi(x_0, y_\varepsilon(x_0)) - \varepsilon \geq \\ &\geq \varphi(z, y_\varepsilon(x_0)) - \varphi(x_0, y_\varepsilon(x_0)) - \varepsilon \geq \\ &\geq (v_\mu(x_0, y_\varepsilon(x_0)), z - x_0) - \mu - \varepsilon \quad \forall z \in E_n. \end{aligned}$$

Лемма доказана.

Из лемм 17.2 и 17.3 вытекает следующий способ вычисления $(\varepsilon + \mu)$ -субградиентов функции $f(x)$, определяемой соотношением (17.9). Зафиксируем $\varepsilon > 0, \mu > 0$ и x_0 . Найдем $y_\varepsilon(x_0) \in R_\varepsilon(x_0)$. Для этого достаточно решить задачу максимизации функции $\varphi(x_0, y)$ по y на G с точностью ε , что в общем-то может быть осуществлено за конечное число шагов некоторой бесконечной процедуры вычисления значения $f(x_0)$.

Построим вектор $v_0 = (v_0^{(1)}, \dots, v_0^{(n)})$ с компонентами

$$v_0^{(i)} = h^{-1} [\varphi_v(x_0 + he_i, y_\varepsilon(x_0)) - \varphi_v(x_0, y_\varepsilon(x_0))], \quad i \in 1:n. \quad (17.11)$$

Теорема 17.1. *Предположим, что функция $\varphi(x, y)$ сильно выпукла по x на E_n при каждом $y \in G$ с константой сильной выпуклости m и что*

$$\left\| \frac{\partial \varphi(x, y)}{\partial x} - \frac{\partial \varphi(x', y)}{\partial x} \right\| \leq L \|x - x'\|$$

$$\forall x \in E_n, \quad \forall x' \in S_h(x), \quad \forall y \in G.$$

Если

$$0,5Lh + 2\nu h^{-1} \leq 2\sqrt{\mu mn^{-1}},$$

то для вектора v_0 с компонентами (17.11) справедливо соотношение

$$v_0 \in \partial_{\varepsilon+\mu} f(x_0). \quad (17.12)$$

Доказательство. Действительно, в силу леммы 17.3 вектор $v_0 \in \partial_\mu \varphi(x_0, y_\varepsilon(x_0))$, откуда, на основании леммы 17.4, следует (17.12).

ГЛАВА II

КВАЗИДИФФЕРЕНЦИРУЕМЫЕ ФУНКЦИИ

§ 1. Определение и примеры квазидифференцируемых функций

1. В гл. I мы познакомились с рядом негладких функций. Подробно были изучены выпуклые функции. В частности, было установлено, что если $f(x)$ — конечная выпуклая на выпуклом открытом множестве $S \subset E_n$ функция, то она дифференцируема по направлениям в любой точке $x_0 \in S$, причем

$$\frac{\partial f(x_0)}{\partial g} = \max_{v \in \partial f(x_0)} (v, g), \quad (1.1)$$

где $\partial f(x_0)$ — субдифференциал функции $f(x)$ в точке x_0 . Множество $\partial f(x_0) \subset E_n$ является выпуклым компактом.

В § 14 гл. I при некоторых предположениях была установлена дифференцируемость по направлениям функции супремума и формула (см. замечание 2 § 14 гл. I)

$$\frac{\partial f(x_0)}{\partial g} = \max_{v \in \partial f(x_0)} (v, g), \quad (1.2)$$

где $\partial f(x_0) = \text{co } \mathcal{H}(x_0) \subset E_n$ — выпуклый компакт. Это множество $\partial f(x_0)$ будем тоже называть *субдифференциалом* функции $f(x)$ в точке x_0 .

Таким образом, с каждой точкой x_0 у функции $f(x)$, являющейся выпуклой функцией или функцией супремума, связано множество $\partial f(x_0)$ — субдифференциал. При этом оказывается (для выпуклой функции это было установлено в гл. I, но справедливо также и для функции супремума), что:

1. Производная по направлениям выражается с помощью $\partial f(x_0)$ по формулам (1.1) и (1.2).

2. Необходимое условие минимума на E_n имеет вид

$$0 \in \partial f(x^*), \quad (1.3)$$

где x^* — точка минимума $f(x)$ на E_n .

3. Если $0 \notin \partial f(x_0)$, то направление

$$g(x_0) = -v(x_0) \|v(x_0)\|^{-1},$$

где $\|v(x_0)\| = \min_{v \in \partial f(x_0)} \|v\|$, является направлением наискорейшего спуска функции $f(x)$ в точке x_0 .

В следующих главах будет показано, как можно использовать направление наискорейшего спуска для построения численных методов для минимизации $f(x)$.

Отсюда видно, что понятие субдифференциала полезно для исследования функций, в частности для решения задач оптимизации.

Эти соображения привели к попыткам найти более широкий класс функций, для которых можно было бы построить нечто, аналогичное субдифференциалу. Среди этих попыток следует выделить работы Ф. Кларка [133], Б. Н. Пшеничного [104], Н. З. Шора [126]. Еще в монографии [103] Б. Н. Пшеничный предложил рассматривать функции (он назвал их квазидифференцируемыми), для которых в каждой точке x_0 существует множество $\partial f(x_0)$ такое, что производная по направлению вычисляется по формуле (1.1). Пример выпуклых функций показывает, что класс таких функций не является линейным пространством (уже функция $f_1(x) = -f(x)$ не является квазидифференцируемой по Пшеничному).

Ф. Кларком [133] было введено понятие верхней субпроизводной. Пусть $f(x)$ — непрерывная в окрестности x_0 функция. Положим

$$\frac{\partial_{Cl} f(x_0)}{\partial g} = \overline{\lim}_{\substack{x' \rightarrow x_0 \\ \alpha \rightarrow 0}} \frac{f(x' + \alpha g) - f(x')}{\alpha}.$$

Эту величину, если она существует и конечна, назовем *верхней субпроизводной* (или *производной Кларка*) функции $f(x)$ в точке x_0 по направлению g и будем говорить, что функция $f(x)$ *дифференцируема по Кларку* в точке x_0 .

Было установлено, что если функция $f(x)$ липшицева в окрестности точки x_0 , то она дифференцируема по Кларку в точке x_0 , причем существует такой выпуклый компакт $\partial_{Cl} f(x_0) \subset E_n$, что

$$\frac{\partial_{Cl} f(x_0)}{\partial g} = \max_{v \in \partial_{Cl} f(x_0)} (v, g).$$

Множество $\partial_{\text{Clf}}(x_0)$ называется *субдифференциалом Кларка*. $\Lambda(f) \subset S$ обозначим множество точек, в которых функция $f(x)$ дифференцируема. Ф. Кларк [133] показал, что

$$\partial_{\text{Clf}}(x_0) = \text{co} \{v \mid \exists \{x_k\}: x_k \rightarrow x_0, x_k \in \Lambda(f), f'(x_k) \rightarrow v\}. \quad (1.3)$$

При этом, если $S = E_n$, то для того, чтобы точка $x^* \in E_n$ была точкой минимума функции $f(x)$ на E_n , необходимо, чтобы

$$0 \in \partial_{\text{Clf}} f(x^*). \quad (1.4)$$

Субдифференциал Кларка обладает рядом важных и интересных свойств (см. обзор [165], а также работы [9, 65, 136, 140]).

Пример 1. Пусть $x = (x^{(1)}, x^{(2)}) \in E_2$, $x_2 = 0 = (0, 0)$. Рассмотрим функции

$$\begin{aligned} f_1(x) &= |x^{(1)}| + |x^{(2)}|, & f_2(x) &= |x^{(1)}| - |x^{(2)}|, \\ f_3(x) &= -|x^{(1)}| - |x^{(2)}|, & f_4(x) &= -|x^{(1)}| + |x^{(2)}|, \\ f_5(x) &= \max \{f_2(x), f_4(x)\}. \end{aligned}$$

Нетрудно показать, что для всех функций $f_i(x)$, $i \in 1:5$, справедливо следующее свойство: плоскость E_2 можно разбить на четыре части, в каждой из которых функция $f_i(x)$ совпадает с одной из функций:

$$x^{(1)} + x^{(2)}, \quad x^{(1)} - x^{(2)}, \quad -x^{(1)} + x^{(2)}, \quad -x^{(1)} - x^{(2)}.$$

Так, для функции $f_1(x)$ будет

$$f_1(x) = \begin{cases} x^{(1)} + x^{(2)}, & \text{если } x^{(1)} \geq 0, \quad x^{(2)} \geq 0, \\ x^{(1)} - x^{(2)}, & \text{если } x^{(1)} \geq 0, \quad x^{(2)} \leq 0, \\ -x^{(1)} + x^{(2)}, & \text{если } x^{(1)} \leq 0, \quad x^{(2)} \geq 0, \\ -x^{(1)} - x^{(2)}, & \text{если } x^{(1)} \leq 0, \quad x^{(2)} \leq 0. \end{cases}$$

Поэтому из (1.3) заключаем, что

$$\partial_{\text{Clf}} f_i(x_0) = \text{co} \{(1, 1), (1, -1), (-1, 1), (-1, -1)\} \quad \forall i \in 1:5.$$

В точке $x_0 = 0$ для всех функций $f_i(x)$, $i \in 1:5$, вполне необходимо условие минимума (1.4), хотя точка x_0 является точкой минимума только для функций $f_1(x)$ и $f_5(x)$.

Отметим, кроме того, что субдифференциал Кларка не годится ни для вычисления производных по направлениям (ибо производная Кларка не всегда равна производной по направлению), ни для нахождения направления наискорейшего спуска. Так, даже для вогнутой на E_n функции (см. п. 4 § 5 гл. I) оказывается, что $\partial_{\text{Cl}} f(x_0) = \overline{\partial f(x_0)}$, но

$$\frac{\partial f(x_0)}{\partial g} = \min_{v \in \overline{\partial f(x_0)}} (v, g) = \min_{v \in \partial_{\text{Cl}} f(x_0)} (v, g),$$

т. е. уже не имеет места формула вида (1.1).

В настоящей главе изучен класс функций не такой обширный, как класс липшицевых функций, однако достаточно богатый и обладающий «хорошими» свойствами.

2. Пусть на открытом множестве $S \subset E_n$ задана конечная функция $f(x)$. Будем говорить, что функция $f(x)$ квазидифференцируема в точке $x_0 \in S$, если она дифференцируема в точке x_0 по любому направлению $g \in E_n$ и если существуют выпуклые компакты $\underline{\partial f(x_0)} \subset E_n$ и $\overline{\partial f(x_0)} \subset E_n$ такие, что

$$\begin{aligned} \frac{\partial f(x_0)}{\partial g} &= \lim_{\alpha \rightarrow +0} \alpha^{-1} [f(x_0 + \alpha g) - f(x_0)] = \\ &= \max_{v \in \underline{\partial f(x_0)}} (v, g) + \min_{w \in \overline{\partial f(x_0)}} (w, g). \end{aligned} \quad (1.5)$$

Пару множеств

$$\mathcal{D}f(x_0) = [\underline{\partial f(x_0)}, \overline{\partial f(x_0)}]$$

назовем квазидифференциалом (КВД) функции $f(x)$ в точке x_0 , а множества $\underline{\partial f(x_0)}$ и $\overline{\partial f(x_0)}$ — соответственно субдифференциалом и супердифференциалом функции $f(x)$ в точке x_0 .

Из формулы (1.5) видно, что КВД определяется неоднозначно: пара множеств

$$A(x_0) = [\underline{\partial f(x_0)} + B, \overline{\partial f(x_0)} - B],$$

где $B \subset E_n$ — произвольный выпуклый компакт, тоже является квазидифференциалом.

Если среди квазидифференциалов функции $f(x)$ в точке x_0 есть элемент вида

$$\mathcal{D}f(x_0) = [\underline{\partial f(x_0)}, \{0\}] = [\underline{\partial f(x_0)}, 0],$$

то функцию $f(x)$ будем называть *субдифференцируемой* в точке x_0 . Аналогично, если среди квазидифференциалов функции $f(x)$ в точке x_0 есть элемент вида

$$\mathcal{D}f(x_0) = [\{0\}, \overline{\partial f(x_0)}] = [0, \overline{\partial f(x_0)}],$$

то функцию $f(x)$ будем называть *супердифференцируемой* в точке x_0 .

Пусть $\Omega \subset S$. Если функция $f(x)$ квазидифференцируема в каждой точке $x_0 \in \Omega$, то она называется *квазидифференцируемой* на множестве Ω .

Приведем простейшие примеры квазидифференцируемых функций.

Пример 2. Пусть функция $f(x)$ непрерывно дифференцируема на $S \subset E_n$ и $x_0 \in S$. Поскольку

$$\frac{\partial f(x_0)}{\partial g} = (f'(x_0), g),$$

то ясно, что $f(x)$ является квазидифференцируемой функцией на S , причем в качестве КВД функции $f(x)$ в точке x_0 можно взять пару множеств

$$\mathcal{D}f(x_0) = [f'(x_0), 0],$$

т. е.

$$\underline{\partial f(x_0)} = \{f'(x_0)\}, \quad \overline{\partial f(x_0)} = \{0\}.$$

Очевидно, что пара множеств

$$[0, f'(x_0)]$$

тоже является КВД функции $f(x)$ в точке x_0 .

Таким образом, функция $f(x)$, дифференцируемая в точке x_0 , является и субдифференцируемой, и супердифференцируемой в этой точке.

Пример 3. Пусть $f(x)$ — конечная выпуклая на выпуклом открытом множестве $S \subset E_n$ функция, $x_0 \in S$. Тогда $f(x)$ дифференцируема по направлениям. По теореме 5.2 гл. I

$$\frac{\partial f(x_0)}{\partial g} = \max_{v \in \partial f(x_0)} (v, g),$$

где $\partial f(x_0) = \{v \in E_n \mid f(x) - f(x_0) \geq (v, x - x_0) \quad \forall x \in S\}$.

Отсюда ясно, что функция $f(x)$ является квазидифференцируемой в точке x_0 , причем в качестве КВД можно взять

$$\mathcal{D}f(x_0) = [\partial f(x_0), \mathbb{O}],$$

т. е. $\underline{\partial f(x_0)} = \partial f(x_0)$, $\overline{\partial f(x_0)} = \{\mathbb{O}\}$.

Поскольку $\overline{\partial f(x_0)} = \{\mathbb{O}\}$, то выпуклая функция субдифференцируема в точке x_0 (и на S).

Пример 4. Пусть $f(x)$ — конечная вогнутая на выпуклом открытом множестве $S \subset E_n$ функция и $x_0 \in S$. Функция $f(x)$ дифференцируема в точке x_0 по направлениям. В п. 4 § 5 гл. I было показано, что

$$\frac{\partial f(x_0)}{\partial g} = \min_{w \in \overline{\partial f(x_0)}} (w, g),$$

где

$$\overline{\partial f(x_0)} = \{w \in E_n \mid f(x) - f(x_0) \leq (w, x - x_0) \quad \forall x \in S\}.$$

Отсюда очевидно, что функция $f(x)$ квазидифференцируема в точке x_0 , причем в качестве КВД можно взять

$$\mathcal{D}f(x_0) = [\mathbb{O}, \overline{\partial f(x_0)}].$$

Так как здесь $\underline{\partial f(x_0)} = \{\mathbb{O}\}$, то вогнутая функция супердифференцируема в точке x_0 (и на S).

Пример 5. Пусть

$$f(x) = \sup_{y \in G} \varphi(x, y),$$

где $x \in E_n$, $y \in E_p$, $G \subset E_p$ — произвольное множество. Семейства функций $\varphi(x, y)$ и $\varphi'_x(x, y)$, где $y \in G$, равномерно непрерывны по x в точке x_0 . Также предполагается, что

$$|\varphi(x_0, y)| \leq K < \infty, \quad |\varphi'_x(x_0, y)| \leq K < \infty \quad \forall y \in G.$$

Как показано в теореме 14.1 гл. I, функция $f(x)$ субдифференцируема, а тогда она тем более квазидифференцируема. В качестве ее КВД можно взять пару множеств

$$\mathcal{D}f(x_0) = [\text{co } \mathcal{H}(x_0), \{\mathbb{O}\}],$$

где $\mathcal{H}(x_0)$ определено в п. 1 § 14 гл. I.

§ 2. Свойства квазидифференцируемых функций.

Основные формулы

квазидифференциального исчисления

1. Вначале установим простейшие свойства квазидифференцируемых функций. Пусть по-прежнему все рассматриваемые функции заданы на открытом множестве $S \subset E_n$.

1. Если $f_1(x)$ — квазидифференцируемая в точке $x_0 \in S$ функция, то и функция $f(x) = cf_1(x)$, где $c \geq 0$, тоже квазидифференцируема в точке x_0 .

Действительно, так как

$$\begin{aligned} \frac{\partial f(x_0)}{\partial g} &= c \frac{\partial f_1(x_0)}{\partial g} = \\ &= c \left[\max_{v \in \underline{\partial f_1(x_0)}} (v, g) + \min_{w \in \overline{\partial f_1(x_0)}} (w, g) \right] = \\ &= \max_{v \in [c \underline{\partial f_1(x_0)}]} (v, g) + \min_{w \in [c \overline{\partial f_1(x_0)}]} (w, g), \end{aligned}$$

то ясно, что функция $f(x)$ квазидифференцируема в точке x_0 , причем

$$\mathcal{D}f(x_0) = [\underline{\partial f(x_0)}, \overline{\partial f(x_0)}],$$

где $\underline{\partial f(x_0)} = c \underline{\partial f_1(x_0)}$, $\overline{\partial f(x_0)} = c \overline{\partial f_1(x_0)}$. ■

2. Если $f_1(x)$ — квазидифференцируемая в точке $x_0 \in S$ функция, то и функция $f(x) = -f_1(x)$ тоже квазидифференцируема в точке x_0 .

Действительно, имеем

$$\begin{aligned} \frac{\partial f(x_0)}{\partial g} &= - \frac{\partial f_1(x_0)}{\partial g} = \\ &= - \left[\max_{v \in \underline{\partial f_1(x_0)}} (v, g) + \min_{w \in \overline{\partial f_1(x_0)}} (w, g) \right]. \quad (2.1) \end{aligned}$$

Поскольку

$$\begin{aligned} - \max_{v \in \underline{\partial f_1(x_0)}} (v, g) &= \min_{v \in \underline{\partial f_1(x_0)}} (-v, g) = \min_{w \in [-\underline{\partial f_1(x_0)}]} (w, g), \\ - \min_{w \in \overline{\partial f_1(x_0)}} (w, g) &= \max_{v \in \overline{\partial f_1(x_0)}} (-v, g) = \max_{v \in [-\overline{\partial f_1(x_0)}]} (v, g), \end{aligned}$$

то из (2.1)

$$\frac{\partial f(x_0)}{\partial g} = \max_{v \in [-\overline{\partial f_1(x_0)}]} (v, g) + \min_{w \in [-\underline{\partial f_1(x_0)}]} (w, g).$$

Таким образом, функция $f(x)$ квазидифференцируема в точке x_0 , причём

$$\mathcal{D}f(x_0) = [\underline{\partial f}(x_0), \overline{\partial f}(x_0)],$$

где

$$\underline{\partial f}(x_0) = -\overline{\partial f_1}(x_0), \quad \overline{\partial f}(x_0) = -\underline{\partial f_1}(x_0). \blacksquare$$

3. Если функции $f_1(x)$ и $f_2(x)$ квазидифференцируемы в точке $x_0 \in S$, то и функция $f(x) = f_1(x) + f_2(x)$ тоже квазидифференцируема в точке x_0 .

Действительно, поскольку

$$\begin{aligned} \frac{\partial f(x_0)}{\partial g} &= \frac{\partial f_1(x_0)}{\partial g} + \frac{\partial f_2(x_0)}{\partial g} = \\ &= \max_{v \in \underline{\partial f_1}(x_0)} (v, g) + \min_{w \in \underline{\partial f_1}(x_0)} (w, g) + \max_{v \in \underline{\partial f_2}(x_0)} (v, g) + \\ &\quad + \min_{w \in \underline{\partial f_2}(x_0)} (w, g) = \max_{v \in [\underline{\partial f_1}(x_0) + \underline{\partial f_2}(x_0)]} (v, g) + \\ &\quad + \min_{w \in [\overline{\partial f_1}(x_0) + \overline{\partial f_2}(x_0)]} (w, g), \end{aligned}$$

то ясно, что $f(x)$ — квазидифференцируемая в точке x_0 функция, причём

$$\mathcal{D}f(x_0) = [\underline{\partial f}(x_0), \overline{\partial f}(x_0)],$$

где

$$\underline{\partial f}(x_0) = [\underline{\partial f_1}(x_0) + \underline{\partial f_2}(x_0)], \quad \overline{\partial f}(x_0) = [\overline{\partial f_1}(x_0) + \overline{\partial f_2}(x_0)].$$

4. Если функции $f_1(x), \dots, f_k(x)$ квазидифференцируемы в точке $x_0 \in S$, то и функция $f(x) = \sum_{i=1}^k c_i f_i(x)$, где $c_i \in E_1$, тоже квазидифференцируема в точке x_0 .

Это очевидным образом следует из свойств 1—3. Нетрудно выписать и формулы для субдифференциала и супердифференциала функции $f(x)$.

Из свойств 1—4 теперь следует

Лемма 2.1. Множество квазидифференцируемых функций является линейным пространством.

2. Пусть $f_i(x)$, $i \in I \equiv 1 : N$, — квазидифференцируемые в точке $x_0 \in S$ функции. Образует функцию

$$f(x) = \max_{i \in I} f_i(x). \quad (2.2)$$

Прежде всего заметим, что если функции $f_i(x)$ дифференцируемы в точке x_0 по направлению $g \in E_n$, то и функция $f(x)$ тоже является дифференцируемой в точке x_0 по направлению g , причем

$$\frac{\partial f(x_0)}{\partial g} = \max_{i \in R(x_0)} \frac{\partial f_i(x_0)}{\partial g}, \quad (2.3)$$

$$\partial_e R(x) = \{i \in I \mid f_i(x) = f(x)\}.$$

Действительно, поскольку

$$f_i(x_0 + \alpha g) = f_i(x_0) + \alpha \frac{\partial f_i(x_0)}{\partial g} + o_i(\alpha), \quad (2.4)$$

где $\frac{o_i(\alpha)}{\alpha} \xrightarrow{\alpha \rightarrow +0} 0$, то при достаточно малых $\alpha > 0$

$$f(x_0 + \alpha g) = \max_{i \in I} f_i(x_0 + \alpha g) = \max_{i \in R(x_0)} f_i(x_0 + \alpha g).$$

Отсюда и из (2.4)

$$f(x_0 + \alpha g) - f(x_0) = \max_{i \in R(x_0)} \left[\alpha \frac{\partial f_i(x_0)}{\partial g} + o_i(\alpha) \right].$$

Теперь имеем

$$\underline{o}(\alpha) \leq f(x_0 + \alpha g) - f(x_0) - \alpha \max_{i \in R(x_0)} \frac{\partial f_i(x_0)}{\partial g} \leq \bar{o}(\alpha), \quad (2.5)$$

где

$$\underline{o}(\alpha) = \min_{i \in I} o_i(\alpha), \quad \bar{o}(\alpha) = \max_{i \in I} o_i(\alpha),$$

$$\frac{\underline{o}(\alpha)}{\alpha} \xrightarrow{\alpha \rightarrow +0} 0, \quad \frac{\bar{o}(\alpha)}{\alpha} \xrightarrow{\alpha \rightarrow +0} 0.$$

Положим $h(\alpha) = \alpha^{-1} [f(x_0 + \alpha g) - f(x_0)]$. Из (2.5)

$$\left| h(\alpha) - \max_{i \in R(x_0)} \frac{\partial f_i(x_0)}{\partial g} \right| \xrightarrow{\alpha \rightarrow +0} 0,$$

а это и означает, что справедливо (2.3). ■

Лемма 2.2. Если функции $f_i(x)$, $i \in I \equiv 1 : N$, квазидифференцируемы в точке $x_0 \in S$, то функция

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

является квазидифференцируемой в точке x_0 , причем

$$\mathcal{D}f(x_0) = [\underline{\partial}f(x_0), \overline{\partial}f(x_0)],$$

где

$$\underline{\partial f}(x_0) = \text{co} \left\{ \frac{\partial f_k(x_0)}{\partial g} - \sum_{\substack{i \in R(x_0) \\ i \neq k}} \overline{\partial f_i(x_0)} \mid k \in R(x_0) \right\}, \quad (2.6)$$

$$\overline{\partial f}(x_0) = \sum_{k \in R(x_0)} \overline{\partial f_k(x_0)}. \quad (2.7)$$

Доказательство. Функции $f_i(x)$ квазидифференцируемы в точке x_0 , т. е.

$$\frac{\partial f_i(x_0)}{\partial g} = \max_{v \in \underline{\partial f_i}(x_0)} (v, g) + \min_{w \in \overline{\partial f_i}(x_0)} (w, g). \quad (2.8)$$

Обозначим

$$A_i = \underline{\partial f_i}(x_0), \quad B_i = \overline{\partial f_i}(x_0). \quad (2.9)$$

Без ограничения общности можем считать, что $R(x_0) = 1:m$, где $m \leq N$. Из (2.3) и (2.8) имеем

$$\frac{\partial f(x_0)}{\partial g} = \max_{i \in 1:m} \left[\max_{v \in A_i} (v, g) + \min_{w \in B_i} (w, g) \right]. \quad (2.10)$$

Рассмотрим вначале случай $m = 2$. Пусть a, b, c, d, r — вещественные числа. Поскольку

$$\max \{a + b, c + d\} = \max \{a + b - r, c + d - r\} + r,$$

то, полагая $r = b + d$, имеем

$$\max \{a + b, c + d\} = \max \{a - d, c - b\} + (b + d).$$

Отсюда и из (2.10)

$$\begin{aligned} H_2 &= \max_{i \in 1:2} \left[\max_{v \in A_i} (v, g) + \min_{w \in B_i} (w, g) \right] = \\ &= \max \left\{ \max_{v \in A_1} (v, g) - \min_{w \in B_2} (w, g), \max_{v \in A_2} (v, g) - \right. \\ &\quad \left. - \min_{w \in B_1} (w, g) \right\} + \left[\min_{w \in B_1} (w, g) + \min_{w \in B_2} (w, g) \right]. \quad (2.11) \end{aligned}$$

Так как

$$- \min_{w \in B_i} (w, g) = \max_{v \in B_i} (-w, g) = \max_{v \in [-B_i]} (v, g), \quad i \in 1:2,$$

то из (2.11)

$$\begin{aligned} H_2 &= \max \left\{ \max_{v \in A_1} (v, g) + \max_{v \in [-B_2]} (v, g), \max_{v \in A_2} (v, g) + \right. \\ &\quad \left. + \max_{v \in [-B_1]} (v, g) \right\} + \min_{w \in B_1} (w, g) + \min_{w \in B_2} (w, g). \quad (2.12) \end{aligned}$$

Но

$$\begin{aligned}\max_{v \in A} (v, g) + \max_{v \in B} (v, g) &= \max_{v \in [A+B]} (v, g), \\ \min_{w \in A} (w, g) + \min_{w \in B} (w, g) &= \min_{w \in [A+B]} (w, g).\end{aligned}$$

Поэтому из (2.12)

$$H_2 = \max \left\{ \max_{v \in [A_1 - B_2]} (v, g); \max_{v \in [A_2 - B_1]} (v, g) \right\} + \min_{w \in [B_1 + B_2]} (w, g).$$

Поскольку

$$\max \left\{ \max_{v \in A} (v, g), \max_{v \in B} (v, g) \right\} = \max_{v \in \text{co}\{A \cup B\}} (v, g),$$

то

$$H_2 = \max_{v \in G_{21}} (v, g) + \min_{w \in G_{22}} (w, g), \quad (2.13)$$

где $G_{21} = \text{co}\{A_1 - B_2, A_2 - B_1\}$, $G_{22} = B_1 + B_2$.

Далее доказываем по индукции. Пусть уже установлено, что

$$\begin{aligned}H_l &= \max_{i \in 1:l} \left[\max_{v \in A_i} (v, g) + \min_{w \in B_i} (w, g) \right] = \\ &= \max_{v \in G_{l1}} (v, g) + \min_{w \in G_{l2}} (w, g),\end{aligned} \quad (2.14)$$

где

$$G_{l1} = \text{co} \left\{ A_k - \sum_{i \in 1:l, i \neq k} B_i \mid k \in 1:l \right\}, \quad (2.15)$$

$$G_{l2} = \sum_{i=1}^l B_i. \quad (2.16)$$

Найдем

$$\begin{aligned}H_{l+1} &= \max_{i \in 1:(l+1)} \left[\max_{v \in A_i} (v, g) + \min_{w \in B_i} (w, g) \right] = \\ &= \max \left\{ H_l, \max_{v \in A_{l+1}} (v, g) + \min_{w \in B_{l+1}} (w, g) \right\} = \\ &= \max \left\{ \max_{v \in G_{l1}} (v, g) + \min_{w \in G_{l2}} (w, g), \max_{v \in A_{l+1}} (v, g) + \right. \\ &\quad \left. + \min_{w \in B_{l+1}} (w, g) \right\}.\end{aligned}$$

Рассуждая, как при выводе (2.13), получим

$$H_{l+1} = \max_{v \in A} (v, g) + \min_{w \in B} (w, g),$$

где $A = \text{co}\{G_{l1} - B_{l+1}, A_{l+1} - G_{l2}\}$, $B = G_{l2} + B_{l+1}$.
Осталось заметить, что

$$A = \text{co}\left\{A_k - \sum_{\substack{i \in I: (l+1) \\ i \neq k}} B_i \mid k \in 1 : (l+1)\right\} \equiv G_{l+1,1},$$

$$B = \sum_{i=1}^{l+1} B_i \equiv G_{l+1,2}.$$

Итак, доказано, что

$$H_m = \max_{v \in G_{m1}} (v, g) + \min_{w \in G_{m2}} (w, g).$$

Справедливость леммы теперь вытекает из (2.15), (2.16), (2.9) и (2.10). ■

Лемма 2.3. Если функции $f_i(x)$, $i \in I \equiv 1 : N$, квазидифференцируемы в точке $x_0 \in S$, то функция

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

является квазидифференцируемой в точке x_0 , причем

$$\mathcal{D}f(x_0) = [\underline{\partial}f(x_0), \overline{\partial}f(x_0)],$$

где

$$\begin{aligned} \underline{\partial}f(x_0) &= \sum_{k \in Q(x_0)} \underline{\partial}f_k(x_0), \\ \overline{\partial}f(x_0) &= \text{co}\left\{\overline{\partial}f_k(x_0) - \sum_{\substack{i \in Q(x_0) \\ i \neq k}} \underline{\partial}f_i(x_0) \mid k \in Q(x_0)\right\}, \end{aligned} \quad (2.17)$$

$$Q(x) = \{i \in I \mid f_i(x) = f(x)\},$$

$$\text{co}\{A_k \mid k \in Q\} \equiv \text{co}\{\cup \{A_k \mid k \in Q\}\}.$$

Доказательство. Поскольку $f(x) = -\max_{i \in I} h_i(x)$, где $h_i(x) = -f_i(x)$, то, используя лемму 2.2 и свойство 2 п. 1, сразу получаем утверждение леммы. ■

Из лемм 2.2 и 2.3 вытекает

Следствие. Если функции

$$f_{ij}(x) \quad (i \in I \equiv 1:N, \quad j \in J_i \equiv 1:N_i)$$

квазидифференцируемы в точке $x_0 \in S$, то и функция

$$f(x) = \max_{i \in I} \min_{j \in J_i} f_{ij}(x) \quad (2.18)$$

тоже квазидифференцируема в точке x_0 .

Аналогично устанавливается, что функция последовательного максимина вида

$$f(x) = \max_i \min_j \max_k \dots \min_l f_{ijkl \dots l}(x)$$

является квазидифференцируемой, если $f_{ijkl \dots l}(x)$ — квазидифференцируемые функции.

Задание 2.1. Получить вид квазидифференциала функции (2.18).

Пример 1. Пусть $f(x)$ — квазидифференцируемая на S функция. Тогда и функция

$$F(x) = \text{sat } f(x) \equiv \begin{cases} +1, & \text{если } f(x) > 1, \\ f(x), & \text{если } -1 \leq f(x) \leq 1, \\ -1, & \text{если } f(x) < -1, \end{cases} \quad (2.19)$$

является квазидифференцируемой на S .

Действительно,

$$F(x) = \max \{f_1(x), -1\},$$

где $f_1(x) = \min \{f(x), +1\}$.

Функция $f_1(x)$ квазидифференцируема по лемме 2.3, а функция $F(x)$ теперь квазидифференцируема по лемме 2.2.

Задание 2.2. Найти квазидифференциал функции (2.19).

3. Пусть функция $f_1(x)$ квазидифференцируема в точке $x_0 \in S$ и непрерывна в окрестности точки x_0 .

Лемма 2.4. Если $f_1(x_0) \neq 0$, то функция $f(x) = f_1^{-1}(x)$ квазидифференцируема в точке x_0 , причем $\mathcal{D}f(x_0) = [\underline{\partial}f(x_0), \overline{\partial}f(x_0)]$, где

$$\underline{\partial}f(x_0) = -\overline{f_1^{-2}(x_0)} \overline{\partial f_1(x_0)}, \quad \overline{\partial}f(x_0) = -f_1^{-2}(x_0) \underline{\partial}f_1(x_0).$$

Здесь и ниже $f_1^{-1}(x) = (f_1(x))^{-1}$, $f_1^{-2}(x) = (f_1(x))^{-2}$.

Доказательство. Заметим, что

$$\begin{aligned}\frac{\partial f(x_0)}{\partial g} &= \lim_{\alpha \rightarrow +0} \alpha^{-1} [f(x_0 + \alpha g) - f(x_0)] = \\ &= \lim_{\alpha \rightarrow +0} \alpha^{-1} [f_1^{-1}(x_0 + \alpha g) - f_1^{-1}(x_0)] = \\ &= \lim_{\alpha \rightarrow +0} [f_1^{-1}(x_0) f_1^{-1}(x_0 + \alpha g) \alpha^{-1} (f_1(x_0) - \\ &\quad - f_1(x_0 + \alpha g))] = -f_1^{-2}(x_0) \frac{\partial f_1(x_0)}{\partial g}.\end{aligned}$$

Теперь осталось воспользоваться свойствами 1 и 2 п. 1. ■

Лемма 2.5. Если функции $f_1(x)$ и $f_2(x)$ квазидифференцируемы в точке $x_0 \in S$, то и функция $f(x) = f_1(x)f_2(x)$ тоже является квазидифференцируемой в этой точке, причем

$$\mathcal{D}f(x_0) = [\underline{\partial}f(x_0), \overline{\partial}f(x_0)],$$

где

$$\underline{\partial}f(x_0) = \begin{cases} f_1 \underline{\partial}f_2 + f_2 \underline{\partial}f_1, & \text{если } f_1 \geq 0, f_2 \geq 0, \\ f_1 \overline{\partial}f_2 + f_2 \underline{\partial}f_1, & \text{если } f_1 \leq 0, f_2 \geq 0, \\ f_1 \overline{\partial}f_2 + f_2 \overline{\partial}f_1, & \text{если } f_1 \leq 0, f_2 \leq 0, \\ f_1 \underline{\partial}f_2 + f_2 \overline{\partial}f_1, & \text{если } f_1 \geq 0, f_2 \leq 0, \end{cases}$$

$$\overline{\partial}f(x_0) = \begin{cases} f_1 \overline{\partial}f_2 + f_2 \overline{\partial}f_1, & \text{если } f_1 \geq 0, f_2 \geq 0, \\ f_1 \underline{\partial}f_2 + f_2 \overline{\partial}f_1, & \text{если } f_1 \leq 0, f_2 \geq 0, \\ f_1 \underline{\partial}f_2 + f_2 \underline{\partial}f_1, & \text{если } f_1 \leq 0, f_2 \leq 0, \\ f_1 \overline{\partial}f_2 + f_2 \underline{\partial}f_1, & \text{если } f_1 \geq 0, f_2 \leq 0. \end{cases}$$

Здесь везде значения функций $f_1(x)$ и $f_2(x)$ и их суб- и супердифференциалов вычислены в точке x_0 .

Доказательство. Очевидно, что

$$\frac{\partial f(x_0)}{\partial g} = f_1(x_0) \frac{\partial f_2(x_0)}{\partial g} + f_2(x_0) \frac{\partial f_1(x_0)}{\partial g}. \quad (2.20)$$

Поскольку

$$\frac{\partial f_i(x_0)}{\partial g} = \max_{v \in \underline{\partial}f_i(x_0)} (v, g) + \min_{w \in \overline{\partial}f_i(x_0)} (w, g), \quad i \in 1:2,$$

то из (2.20)

$$\begin{aligned} \frac{\partial f(x_0)}{\partial g} = & f_1(x_0) \max_{v \in \underline{\partial f_1(x_0)}} (v, g) + f_1(x_0) \min_{w \in \overline{\partial f_1(x_0)}} (w, g) + \\ & + f_2(x_0) \max_{v \in \underline{\partial f_2(x_0)}} (v, g) + f_2(x_0) \min_{w \in \overline{\partial f_2(x_0)}} (w, g). \end{aligned}$$

Осталось воспользоваться свойствами 1 и 2 п. 1. ■

4. Пусть $F(y_1, y_2, \dots, y_s)$ — непрерывно дифференцируемая на $\Omega \subset E_s$ функция, и пусть $f_i(x)$, $i \in 1:s$, — квазидифференцируемые в точке x_0 функции. Предположим, что

$$y_0 = (y_0^{(1)}, \dots, y_0^{(s)}) = (f_1(x_0), \dots, f_s(x_0)) \in \Omega.$$

Лемма 2.6. *Функция $f(x) = F(f_1(x), \dots, f_s(x))$ квазидифференцируема в точке x_0 .*

Доказательство очевидно, если учесть, что

$$\begin{aligned} \frac{\partial f(x_0)}{\partial g} = & \frac{\partial F(y_0)}{\partial y^{(1)}} \frac{\partial f_1(x_0)}{\partial g} + \\ & + \frac{\partial F(y_0)}{\partial y^{(2)}} \frac{\partial f_2(x_0)}{\partial g} + \dots + \frac{\partial F(y_0)}{\partial y^{(s)}} \frac{\partial f_s(x_0)}{\partial g}. \end{aligned}$$

Далее рассуждаем, как при доказательстве леммы 2.5. ■

Таким образом, множество квазидифференцируемых функций представляет собой линейное пространство, замкнутое относительно всех «алгебраических» операций, операций взятия поточечного максимума, минимума, последовательного максимина.

Замечание. Пусть $\mathcal{D} = [A, B]$ — пара множеств из E_n . Определим операции умножения на вещественное число λ и сложения:

$$\lambda \mathcal{D} = \begin{cases} [\lambda A, \lambda B], & \text{если } \lambda \geq 0, \\ [\lambda B, \lambda A], & \text{если } \lambda < 0. \end{cases}$$

Если $\mathcal{D}_1 = [A_1, B_1]$, $\mathcal{D}_2 = [A_2, B_2]$, то $\mathcal{D}_1 + \mathcal{D}_2 = [A_1 + A_2, B_1 + B_2]$.

Тогда формулы для квазидифференциала линейной комбинации, произведения и частного квазидифференцируемых функций можно записать в более компактном

виде:

$$\mathcal{D} \left[\sum_{i=1}^N \lambda_i f_i(x_0) \right] = \sum_{i=1}^N \lambda_i \mathcal{D} f_i(x_0),$$

$$\mathcal{D}[f_1(x_0) f_2(x_0)] = f_1(x_0) \mathcal{D} f_2(x_0) + f_2(x_0) \mathcal{D} f_1(x_0),$$

$$\mathcal{D} \left[\frac{f_1(x_0)}{f_2(x_0)} \right] = \frac{1}{f_2^2(x_0)} [f_2(x_0) \mathcal{D} f_1(x_0) - f_1(x_0) \mathcal{D} f_2(x_0)]$$

(в последней формуле предполагается, что $f_2(x_0) \neq 0$).

Эти выражения являются обобщением обычных формул дифференциального исчисления.

§ 3. Примеры вычисления квазидифференциалов

Приводимые ниже простейшие примеры вычисления квазидифференциалов играют чисто иллюстративную роль, и подготовленный читатель может считать их упражнениями, которые ему следует выполнить самостоятельно.

Пусть

$$x = (x^{(1)}, x^{(2)}) \in E_2, \quad x_0 = (0, 0) = \mathbb{O},$$

$$A_1 = (1, 0), \quad A_2 = (-1, 0),$$

$$B_1 = (0, 1), \quad B_2 = (0, -1),$$

$$C_1 = (1, 1), \quad C_2 = (-1, -1).$$

Пример 1. Пусть $f(x) = f(x^{(1)}, x^{(2)}) = |x^{(1)}|$. Имеем

$$|x^{(1)}| = \max \{x^{(1)}, -x^{(1)}\} = \max \{f_1(x), f_2(x)\},$$

где $f_1(x) = x^{(1)}$, $f_2(x) = -x^{(1)}$. Функции $f_1(x)$ и $f_2(x)$ непрерывно дифференцируемы, и потому (см. пример 2 § 1)

$$\mathcal{D} f_1(x_0) = [\underline{\partial f_1(x_0)}, \overline{\partial f_1(x_0)}],$$

$$\text{где } \underline{\partial f_1(x_0)} = \{A_1\}, \quad \overline{\partial f_1(x_0)} = \{\mathbb{O}\},$$

$$\mathcal{D} f_2(x_0) = [\underline{\partial f_2(x_0)}, \overline{\partial f_2(x_0)}],$$

$$\text{где } \underline{\partial f_2(x_0)} = \{A_2\}, \quad \overline{\partial f_2(x_0)} = \{\mathbb{O}\}.$$

По лемме 2.2 имеем (рис. 17)

$$\mathcal{D} f(x_0) = [\underline{\partial f(x_0)}, \overline{\partial f(x_0)}],$$

$$\text{где } \underline{\partial f(x_0)} = \text{co} \{A_1, A_2\}, \quad \overline{\partial f(x_0)} = \{\mathbb{O}\}.$$

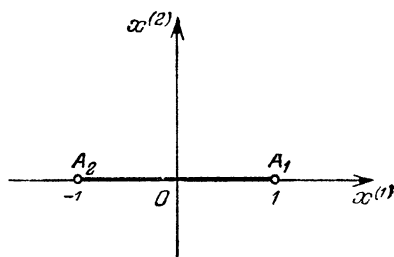


Рис. 17.

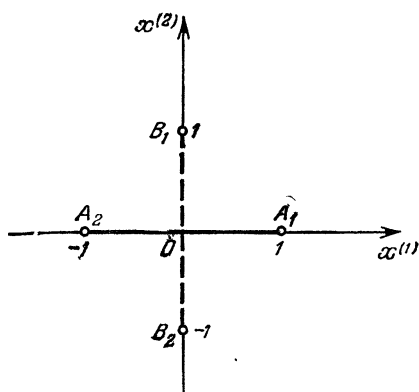


Рис. 18.

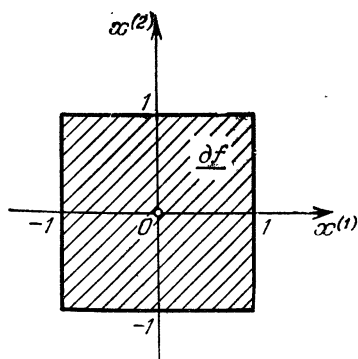


Рис. 19.

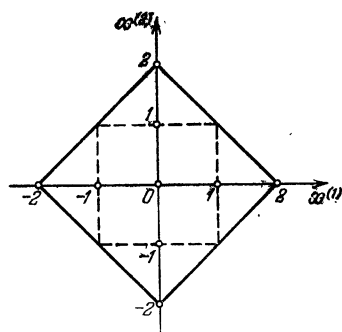


Рис. 20.

Пример 2. Пусть $f(x) = -|x_1| = -f_1(x)$. Из предыдущего примера и свойства 2 п. 1 § 2 $\mathcal{D}f(x_0) = [\underline{\partial f(x_0)}, \overline{\partial f(x_0)}]$, где $\underline{\partial f(x_0)} = \{\emptyset\}$, $\overline{\partial f(x_0)} = -\text{co}\{A_1, A_2\}$. В данном случае (рис. 17)

$$-\text{co}\{A_1, A_2\} = \text{co}\{A_1, A_2\}.$$

Пример 3. Пусть $f(x) = |x^{(2)}|$. Как и в примере 1, получаем

$$\mathcal{D}f(x_0) = [\underline{\partial f(x_0)}, \overline{\partial f(x_0)}],$$

где $\underline{\partial f(x_0)} = \text{co}\{B_1, B_2\}$, $\overline{\partial f(x_0)} = \{\emptyset\}$.

Пример 4. Пусть $f(x) = -|x^{(2)}|$. Как в примере 2, имеем

$$\mathcal{D}f(x_0) = [\underline{\partial f(x_0)}, \overline{\partial f(x_0)}],$$

где $\underline{\partial f(x_0)} = \{\emptyset\}$, $\overline{\partial f(x_0)} = \text{co}\{B_1, B_2\}$.

Пример 5. Пусть $f(x) = |x^{(1)}| - |x^{(2)}|$. По свойству 3 п. 1 § 2 из примеров 1 и 4 имеем

$$\mathcal{D}f(x_0) = [\underline{\partial f(x_0)}, \overline{\partial f(x_0)}],$$

где (рис. 18) $\underline{\partial f(x_0)} = \text{co}\{A_1, A_2\}$, $\overline{\partial f(x_0)} = \text{co}\{B_1, B_2\}$.

Пример 6. Пусть $f(x) = |x^{(1)}| + |x^{(2)}|$. По свойству 3 из п. 1 § 2 имеем

$$\mathcal{D}f(x_0) = [\underline{\partial f(x_0)}, \overline{\partial f(x_0)}],$$

где (рис. 19) $\underline{\partial f(x_0)} = \text{co}\{A_1, A_2\} + \text{co}\{B_1, B_2\}$, $\overline{\partial f(x_0)} = \{\emptyset\}$.

Пример 7. Пусть $f_1(x) = |x^{(1)}| - |x^{(2)}|$, $f_2(x) = |x^{(2)}| - |x^{(1)}|$. Положим $f(x) = \max\{f_1(x), f_2(x)\} = ||x^{(1)}| - |x^{(2)}||$.

Из примера 5

$$\mathcal{D}f_1(x_0) = [\underline{\partial f_1(x_0)}, \overline{\partial f_1(x_0)}],$$

где $\underline{\partial f_1(x_0)} = \text{co}\{A_1, A_2\}$, $\overline{\partial f_1(x_0)} = \text{co}\{B_1, B_2\}$.

Из свойства 2 п. 1 § 2

$$\mathcal{D}f_2(x_0) = [\underline{\partial f_2(x_0)}, \overline{\partial f_2(x_0)}],$$

где $\underline{\partial f_2(x_0)} = -\overline{\partial f_1(x_0)}$, $\overline{\partial f_2(x_0)} = -\underline{\partial f_1(x_0)}$.

В нашем случае (рис. 18)

$$\underline{\partial f_2(x_0)} = -\operatorname{co}\{B_1, B_2\} = \operatorname{co}\{B_1, B_2\},$$

$$\overline{\partial f_2(x_0)} = -\operatorname{co}\{A_1, A_2\} = \operatorname{co}\{A_1, A_2\}.$$

Применяя лемму 2.2, получим $\mathcal{D}f(x_0) = [\underline{\partial f(x_0)}, \overline{\partial f(x_0)}]$, где

$$\begin{aligned}\underline{\partial f(x_0)} &= \operatorname{co}\{\underline{\partial f_1} - \overline{\partial f_2}, \underline{\partial f_2} - \overline{\partial f_1}\} = \\ &= \operatorname{co}\{\operatorname{co}\{A_1, A_2\} + \operatorname{co}\{A_1, A_2\}, \operatorname{co}\{B_1, B_2\} + \operatorname{co}\{B_1, B_2\}\}, \\ \overline{\partial f(x_0)} &= \overline{\partial f_1} + \overline{\partial f_2} = \operatorname{co}\{B_1, B_2\} + \operatorname{co}\{A_1, A_2\}.\end{aligned}$$

Множества $\underline{\partial f(x_0)}$ и $\overline{\partial f(x_0)}$ изображены на рис. 20, где сплошной жирной линией обозначена граница $\underline{\partial f(x_0)}$, а пунктирной линией — граница $\overline{\partial f(x_0)}$.

Пример 8. Пусть $f_1(x) = |x^{(1)}| + |x^{(2)}|$, $f_2(x) = |x^{(1)}| - |x^{(2)}|$. Положим

$$f(x) = \max\{f_1(x), f_2(x)\}.$$

Из примеров 5 и 6 по лемме 2.2 $\mathcal{D}f(x_0) = [\underline{\partial f(x_0)}, \overline{\partial f(x_0)}]$, где (рис. 21)

$$\begin{aligned}\underline{\partial f(x_0)} &= \operatorname{co}\{\underline{\partial f_1} - \overline{\partial f_2}, \underline{\partial f_2} - \overline{\partial f_1}\} = \\ &= \operatorname{co}\{\operatorname{co}\{A_1, A_2\}, \operatorname{co}\{A_1, A_2\} + \operatorname{co}\{B_1, B_2\} - \operatorname{co}\{B_1, B_2\}\}, \\ \overline{\partial f(x_0)} &= \operatorname{co}\{B_1, B_2\}.\end{aligned}$$

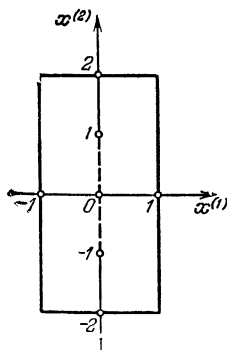


Рис. 21.

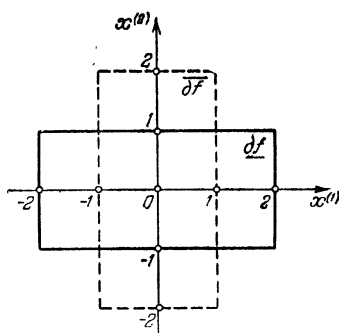


Рис. 22.

Пример 9. Пусть $f_1(x) = |x^{(1)}| + |x^{(2)}|$, $f_2(x) = |x^{(1)}| - |x^{(2)}|$. Положим

$$f(x) = \min \{f_1(x), f_2(x)\}.$$

Из примеров 5 и 6 и леммы 2.3 $\mathcal{D}f(x_0) = [\underline{\partial}f(x_0), \overline{\partial}f(x_0)]$, где

$$\begin{aligned} \underline{\partial}f(x_0) &= \underline{\partial}f_1 + \underline{\partial}f_2 = \text{co} \{A_1, A_2\} + \text{co} \{A_1, A_2\} + \text{co} \{B_1, B_2\}, \\ \overline{\partial}f(x_0) &= \text{co} \{\overline{\partial}f_1 - \underline{\partial}f_2, \overline{\partial}f_2 - \underline{\partial}f_1\} = \\ &= \text{co} \{\text{co} \{B_1, B_2\} - \text{co} \{A_1, A_2\} - \text{co} \{B_1, B_2\}, \text{co} \{A_1, A_2\}\} = \\ &= \text{co} \{\text{co} \{B_1, B_2\} + \text{co} \{A_1, A_2\} + \text{co} \{B_1, B_2\}, \text{co} \{A_1, A_2\}\} \end{aligned}$$

(последнее равенство справедливо в силу того, что

$$\text{co} \{A_1, A_2\} = -\text{co} \{A_1, A_2\},$$

$$\text{co} \{B_1, B_2\} = -\text{co} \{B_1, B_2\}).$$

Множества $\underline{\partial}f(x_0)$ и $\overline{\partial}f(x_0)$ изображены на рис. 22.

Пример 10. Пусть $f(x) = ||x^{(1)}| + x^{(2)}|$. Имеем $f(x) = \max \{f_1(x), f_2(x)\}$, где

$$f_1(x) = |x^{(1)}| + x^{(2)}, \quad f_2(x) = -|x^{(1)}| - x^{(2)}.$$

Так как для функции $f_3(x) = x^{(2)}$ будет $\mathcal{D}f_3(x_0) = [\underline{\partial}f_3(x_0), \overline{\partial}f_3(x_0)]$, где $\underline{\partial}f_3(x_0) = \{B_1\}$, $\overline{\partial}f_3(x_0) = \{\emptyset\}$, то из примеров 1 и 2

$$\mathcal{D}f_1(x_0) = [\underline{\partial}f_1(x_0), \overline{\partial}f_1(x_0)],$$

где

$$\underline{\partial}f_1(x_0) = \text{co} \{A_1, A_2\} + B_1, \quad \overline{\partial}f_1(x_0) = \{\emptyset\},$$

$$\mathcal{D}f_2(x_0) = [\underline{\partial}f_2(x_0), \overline{\partial}f_2(x_0)],$$

где

$$\underline{\partial}f_2(x_0) = \{\emptyset\}, \quad \overline{\partial}f_2(x_0) = -\text{co} \{A_1, A_2\} - B_1.$$

По лемме 2.2 $\mathcal{D}f(x_0) = [\underline{\partial}f(x_0), \overline{\partial}f(x_0)]$, где (рис. 23)

$$\begin{aligned} \underline{\partial}f(x_0) &= \text{co} \{\underline{\partial}f_1 - \overline{\partial}f_2, \underline{\partial}f_2 - \overline{\partial}f_1\} = \\ &= \text{co} \{\text{co} \{A_1, A_2\} + B_1 + \text{co} \{A_1, A_2\} + B_1, \emptyset\}, \\ \overline{\partial}f(x_0) &= \overline{\partial}f_1 + \overline{\partial}f_2 = -\text{co} \{A_1, A_2\} - B_1 = \text{co} \{A_1, A_2\} - B_1. \end{aligned}$$

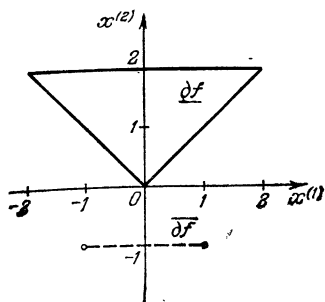


Рис. 23.

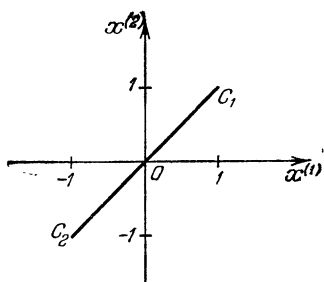


Рис. 24.

Пример 11. Пусть $f(x) = |x^{(1)} + x^{(2)}|$. Так как $f(x) = \max\{f_1(x), f_2(x)\}$, где $f_1(x) = x^{(1)} + x^{(2)}$, $f_2(x) = -x^{(1)} - x^{(2)}$, то (рис. 24)

$$\mathcal{D}f(x_0) = [\underline{\partial f}(x_0), \overline{\partial f}(x_0)],$$

$$\underline{\partial f}(x_0) = \text{co}\{C_1, C_2\}, \quad \overline{\partial f}(x_0) = \{\emptyset\}.$$

Пример 12. Пусть $f_1(x) = 1 + |x^{(1)}|$, $f_2(x) = 1 + |x^{(1)} + x^{(2)}|$. Положим $f(x) = f_1(x) \cdot f_2(x)$.

Если c — константа, то $\mathcal{D}[f(x) + c] = \mathcal{D}f(x)$. Поэтому

$$\mathcal{D}f_1(x_0) = [\underline{\partial f_1}(x_0), \overline{\partial f_1}(x_0)],$$

$$\underline{\partial f_1}(x_0) = \text{co}\{A_1, A_2\}, \quad \overline{\partial f_1}(x_0) = \{\emptyset\},$$

$$\mathcal{D}f_2(x_0) = [\underline{\partial f_2}(x_0), \overline{\partial f_2}(x_0)],$$

$$\underline{\partial f_2}(x_0) = \text{co}\{C_1, C_2\}, \quad \overline{\partial f_2}(x_0) = \{\emptyset\}.$$

Отсюда и по лемме 2.5 $\mathcal{D}f(x_0) = [\underline{\partial f}(x_0), \overline{\partial f}(x_0)]$, где (рис. 25)

$$\begin{aligned} \underline{\partial f}(x_0) &= f_1(x_0) \underline{\partial f_2}(x_0) + f_2(x_0) \underline{\partial f_1}(x_0) = \\ &= \text{co}\{A_1, A_2\} + \text{co}\{C_1, C_2\}, \\ \overline{\partial f}(x_0) &= \{\emptyset\}. \end{aligned}$$

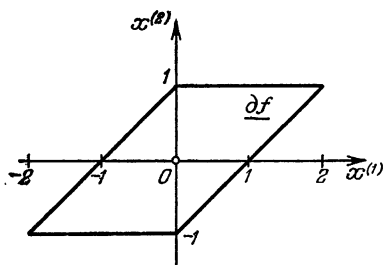


Рис. 25.

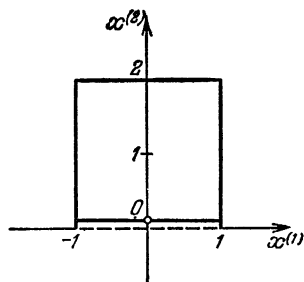


Рис. 26.

Пример 13. Пусть $f_1(x) = |x^{(1)}| + x^{(2)} + 1$, $f_2(x) = [f_3(x)]^{-1}$, $f_3(x) = |x^{(1)}| - |x^{(2)}| + 1$. Положим $f(x) = f_1(x) \cdot f_2(x)$.

Так как $\mathcal{D}f_1(x) = \mathcal{D}f_4(x)$, где $f_4(x) = |x^{(1)}| + x^{(2)}$, то (см. пример 10)

$$\mathcal{D}f_1(x_0) = [\underline{\partial f_1(x_0)}, \overline{\partial f_1(x_0)}],$$

$$\underline{\partial f_1(x_0)} = \text{co} \{A_1, A_2\} + B_1, \quad \overline{\partial f_1(x_0)} = \{0\}.$$

Поскольку $\mathcal{D}f_3(x) = \mathcal{D}f_5(x)$, где $f_5(x) = |x^{(1)}| - |x^{(2)}|$, то (см. пример 5)

$$\mathcal{D}f_3(x_0) = [\underline{\partial f_3(x_0)}, \overline{\partial f_3(x_0)}],$$

$$\underline{\partial f_3(x_0)} = \text{co} \{A_1, A_2\}, \quad \overline{\partial f_3(x_0)} = \text{co} \{B_1, B_2\}.$$

По лемме 2.4

$$\mathcal{D}f_2(x_0) = [\underline{\partial f_2(x_0)}, \overline{\partial f_2(x_0)}],$$

где

$$\underline{\partial f_2(x_0)} = -\overline{\partial f_3(x_0)} = -\text{co} \{B_1, B_2\} = \text{co} \{B_1, B_2\},$$

$$\overline{\partial f_2(x_0)} = -\underline{\partial f_3(x_0)} = -\text{co} \{A_1, A_2\} = \text{co} \{A_1, A_2\}.$$

По лемме 2.5 окончательно получаем (рис. 26)

$$\mathcal{D}f(x_0) = [\underline{\partial f(x_0)}, \overline{\partial f(x_0)}],$$

где

$$\underline{\partial f(x_0)} = f_1(x_0) \underline{\partial f_2(x_0)} + f_2(x_0) \underline{\partial f_1(x_0)} = \\ = \text{co}\{B_1, B_2\} + \text{co}\{A_1, A_2\} + B_1,$$

$$\overline{\partial f(x_0)} = f_1(x_0) \overline{\partial f_2(x_0)} + f_2(x_0) \overline{\partial f_1(x_0)} = \\ = \text{co}\{A_1, A_2\} + \{0\} = \text{co}\{A_1, A_2\}.$$

Пример 14. Пусть $f_1(x) = 1 + |x^{(1)}| + x^{(2)}$. Образует функцию $f(x) = [f_1(x)]^2$. В предыдущем примере уже отмечалось, что $\mathcal{D}f_1(x_0) = [\underline{\partial f_1(x_0)}, \overline{\partial f_1(x_0)}]$, где $\underline{\partial f_1(x_0)} = \text{co}\{A_1, A_2\} + B_1$, $\overline{\partial f_1(x_0)} = \{0\}$.

По лемме 2.5 $\mathcal{D}f(x_0) = [\underline{\partial f(x_0)}, \overline{\partial f(x_0)}]$, где (рис. 27)

$$\underline{\partial f(x_0)} = f_1(x_0) \underline{\partial f_1(x_0)} + f_1(x_0) \underline{\partial f_1(x_0)} = \\ = \underline{\partial f_1(x_0)} + \underline{\partial f_1(x_0)} = \text{co}\{A_1, A_2\} + B_1 + \text{co}\{A_1, A_2\} + B_1, \\ \overline{\partial f(x_0)} = f_1(x_0) \overline{\partial f_1(x_0)} + f_1(x_0) \overline{\partial f_1(x_0)} = \{0\}.$$

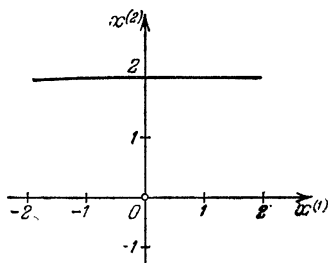


Рис. 27.

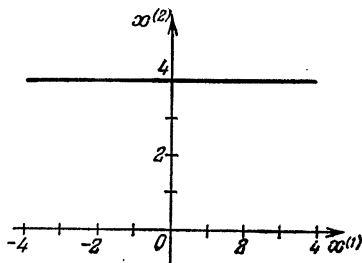


Рис. 28.

Пример 15. Пусть $f_1(x) = 2 + |x^{(1)}| + x^{(2)}$. Положим $f(x) = [f_1(x)]^2$. Поскольку (см. пример 13) $\mathcal{D}f_1(x_0) = [\underline{\partial f_1(x_0)}, \overline{\partial f_1(x_0)}]$, где $\underline{\partial f_1(x_0)} = \text{co}\{A_1, A_2\} + B_1$, $\overline{\partial f_1(x_0)} = \{0\}$, то по лемме 2.5 $\mathcal{D}f(x_0) = [\underline{\partial f(x_0)}, \overline{\partial f(x_0)}]$. Здесь (рис. 28)

$$\underline{\partial f(x_0)} = f_1(x_0) \underline{\partial f_1(x_0)} + f_1(x_0) \underline{\partial f_1(x_0)} = \\ = 2\underline{\partial f_1(x_0)} + 2\underline{\partial f_1(x_0)} = 2\text{co}\{A_1, A_2\} + 2B_1 + \\ + 2\text{co}\{A_1, A_2\} + 2B_1, \\ \overline{\partial f(x_0)} = f_1(x_0) \overline{\partial f_1(x_0)} + f_1(x_0) \overline{\partial f_1(x_0)} = \{0\}.$$

Пример 16. Пусть $f(x) = \|x\|$, где $\|x\| = \sqrt{(x^{(1)})^2 + (x^{(2)})^2}$. Очевидно, что $\frac{\partial f(x_0)}{\partial g} = \lim_{\alpha \rightarrow +0} \alpha^{-1} [f(x_0 + \alpha g) - f(x_0)] = \alpha \cdot \alpha^{-1} \|g\| = \|g\|$. Значит, $\frac{\partial f(x_0)}{\partial g} = \max_{v \in S_1(\emptyset)} (v, g)$, где $S_1(\emptyset) = \{v \in E_2 \mid \|v\| \leq 1\}$. Поэтому функция $f(x)$ является квазидифференцируемой (даже субдифференцируемой), причем $\mathcal{D}f(x_0) = [\underline{\partial f(x_0)}, \overline{\partial f(x_0)}]$, где $\underline{\partial f(x_0)} = S_1(\emptyset)$, $\overline{\partial f(x_0)} = \{\emptyset\}$.

Замечание 1. Если $x = (x^{(1)}, \dots, x^{(n)}) \in E_n$, $f(x) = \|x\|$, где $\|x\| = \sqrt{\sum (x^{(i)})^2}$, то аналогично показывается, что

$$\mathcal{D}f(x_0) = [\underline{\partial f(x_0)}, \overline{\partial f(x_0)}],$$

где

$$\underline{\partial f(x_0)} = S_1(\emptyset) = \{v \in E_n \mid \|v\| \leq 1\}, \quad \overline{\partial f(x_0)} = \{\emptyset\}.$$

Замечание 2. Везде выше для каждой функции был построен один из возможных квазидифференциалов. Функция в примере 8 совпадает с функцией в примере 6, а функция в примере 9 — с функцией в примере 5, хотя квазидифференциалы у совпадающих функций получены разные.

Задача 3.1*. Пусть $A \subset E_n$ — выпуклый компакт. Определим отображение $\mathcal{H}_A: E_n \rightarrow \Pi(E_n)$:

$$\mathcal{H}_A(g) = \{v \in A \mid (v, g) = \max_{z \in A} (z, g)\}.$$

Ясно, что $\mathcal{H}_A(\emptyset) = A$ и что достаточно рассматривать $\mathcal{H}_A$ на единичной сфере $C_1 = \{g \in E_n \mid \|g\| = 1\}$, ибо

$$\mathcal{H}_A(\lambda g) = \mathcal{H}_A(g) \quad \forall \lambda > 0.$$

Можно показать, что почти для всех $g \in C_1$ множество $\mathcal{H}_A(g)$ состоит из единственной точки.

Пусть $f(x)$ — квазидифференцируемая в точке x_0 функция, и пусть $\mathcal{D}f(x_0) = [\underline{\partial f(x_0)}, \overline{\partial f(x_0)}]$ — ее квазидифференциал в точке x_0 . Введем отображение $Q: C_1 \rightarrow \Pi(E_n)$ следующим образом:

$$Q(g) = \mathcal{H}_{\underline{\partial f(x_0)}}(g) - \mathcal{H}_{\overline{\partial f(x_0)}}(g).$$

Множество $Q(g)$ почти для всех $g \in C_1$ тоже состоит из единственной точки. Через $\bar{C}_1$ обозначим множество тех $g \in C_1$, для которых $Q(g)$ состоит из единственной точки.

Доказать или опровергнуть следующее утверждение:

Если функция $f(x)$ липшицева и квазидифференцируема в окрестности точки x_0 , то для субдифференциала Кларка (см. (2.3)) этой функции имеет место представление

$$\partial_{\text{Cl}} f(x_0) = \text{co} \{g(g) \mid g \in \bar{C}_1\}. \quad (3.1)$$

Проверить, что во всех рассмотренных в настоящем параграфе примерах соотношение (3.1) выполняется.

§ 4*. Квазидифференцируемость выпукло-вогнутой функции

1. Пусть функция $f(x) \equiv f(y, z)$ задана и конечна на $S = S_1 \times S_2 \subset E_n$, где $x = [y, z] \in E_n$, $n = n_1 + n_2$, $S_1 \subset E_{n_1}$ и $S_2 \subset E_{n_2}$ — выпуклые открытые множества в соответствующих пространствах. Функция $f(x)$ называется *выпукло-вогнутой* на S , если функция $f(y, z)$ выпукла по y на S_1 при каждом фиксированном $z \in S_2$ и вогнута по z на S_2 при любом фиксированном $y \in S_1$. Положим

$$\begin{aligned} \partial f_y(y_0, z_0) = \{v \in E_{n_1} \mid f(y, z_0) - f(y_0, z_0) \geq \\ \geq (v, y - y_0) \quad \forall y \in S_1\}, \end{aligned} \quad (4.1)$$

$$\begin{aligned} \partial f_z(y_0, z_0) = \{w \in E_{n_2} \mid f(y_0, z) - f(y_0, z_0) \leq \\ \leq (w, z - z_0) \quad \forall z \in S_2\}. \end{aligned} \quad (4.2)$$

Множество $\partial f_y(y_0, z_0)$ является субдифференциалом выпуклой функции $h_1(y) = f(y, z_0)$ в точке y_0 , а множество $\partial f_z(y_0, z_0)$ — супердифференциалом вогнутой функции $h_2(z) = f(y_0, z)$ в точке z_0 . Эти множества непусты, выпуклы, замкнуты и ограничены.

Установим непрерывность выпукло-вогнутой функции $f(x)$ по совокупности переменных (непрерывность $f(x)$ по каждой из переменных y и z следует из свойств выпуклых и вогнутых функций). Вначале докажем лемму, представляющую самостоятельный интерес.

Лемма 4.1. Если функция $f(y, z)$ выпукла по y на S_1 при любом фиксированном $z \in S_2$ и непрерывна по z на S_2 при каждом фиксированном $y \in S_1$, то $f(y, z)$ непрерывна по совокупности переменных.

Доказательство. Пусть $x_0 = [y_0, z_0] \in S$ и $\Omega = \Omega_1 \times \Omega_2$, где $\Omega_1 \subset S_1$, $\Omega_2 \subset S_2$ — выпуклые компакты такие, что $y_0 \in \text{int } \Omega_1$, $z_0 \in \text{int } \Omega_2$. Вначале установим, что

функция $f(x)$ ограничена на Ω . Для этого рассмотрим функцию $\varphi(y) = \sup_{z \in \Omega_2} f(y, z)$. При каждом фиксированном $y \in S_1$ функция $f(y, z)$ непрерывна по z на Ω_2 и, поскольку Ω_2 — компакт, то и ограничена там. Поэтому $\varphi(y)$ — конечная функция, причем можно написать $\varphi(y) = \max_{z \in \Omega_2} f(y, z)$. Нетрудно теперь проверить, что $\varphi(y)$ — выпуклая на S_1 функция. Действительно, поскольку $f(y, z)$ — выпуклая по y функция, то

$$\begin{aligned} \varphi(\alpha y_1 + (1 - \alpha) y_2) &= \max_{z \in \Omega_2} f(\alpha y_1 + (1 - \alpha) y_2, z) \leq \\ &\leq \max_{z \in \Omega_2} [\alpha f(y_1, z) + (1 - \alpha) f(y_2, z)] \leq \\ &\leq \alpha \max_{z \in \Omega_2} f(y_1, z) + (1 - \alpha) \max_{z \in \Omega_2} f(y_2, z) = \\ &= \alpha \varphi(y_1) + (1 - \alpha) \varphi(y_2) \quad \forall \alpha \in [0, 1], \end{aligned}$$

а это и означает выпуклость $\varphi(y)$ на S_1 . Но конечная выпуклая функция (см. § 4 гл. I) является непрерывной на S_1 , а тогда она и ограничена на любом компакте $\Omega' \subset S_1$. Отсюда заключаем, что $f(y, z)$ — ограниченная на $\Omega' \times \Omega_2$ функция. Зафиксируем любое $\Omega' \subset S_1$ такое, что

$$\Omega_1 \subset \text{int } \Omega'. \quad (4.3)$$

Покажем теперь, что множество субдифференциалов $\partial f_y(y, z)$ ограничено на $\Omega = \Omega_1 \times \Omega_2$. Допустим противное, пусть найдется такая последовательность $\{v_k\}$, что

$$v_k \in \partial f_y(y_k, z_k), \quad \|v_k\| \rightarrow +\infty, \quad x_k = [y_k, z_k] \in \Omega. \quad (4.4)$$

По определению (см. (4.1))

$$f(y, z_k) - f(y_k, z_k) \geq (v_k, y - y_k) \quad \forall y \in S_1. \quad (4.5)$$

В силу (4.3) найдется $\beta > 0$ такое, что

$$y'_k = y_k + \beta v_k \|v_k\|^{-1} \in \Omega' \subset S_1 \quad \forall k.$$

Из (4.5)

$$f(y'_k, z_k) \geq f(y_k, z_k) + \beta \|v_k\|. \quad (4.6)$$

Так как значения $f(y_k, z_k)$ ограничены на $\Omega' \times \Omega_2$, то из (4.4) и (4.6) $f(y'_k, z_k) \rightarrow +\infty$, что противоречит ограни-

ценности $f(y, z)$ на $\Omega' \times \Omega_2$. Итак, найдется такое $L < \infty$, что

$$\|v\| \leq L \quad \forall v \in \partial f_y(y, z), \quad \forall [y, z] \in \Omega. \quad (4.7)$$

Пусть теперь $x' \rightarrow x_0$, $x' \in \Omega$, $x' = (y', z')$, $x_0 = (y_0, z_0)$. Имеем

$$\begin{aligned} |f(x') - f(x_0)| &= |f(y', z') - f(y_0, z_0)| = \\ &= |f(y', z') - f(y_0, z') + f(y_0, z') - f(y_0, z_0)| \leq \\ &\leq |f(y', z') - f(y_0, z')| + |f(y_0, z') - f(y_0, z_0)|. \end{aligned} \quad (4.8)$$

Из (4.5) и (4.7)

$$f(y', z') - f(y_0, z') \geq (v_0, y' - y_0) \geq -L\|y' - y_0\|, \quad (4.9)$$

где $v_0 \in \partial f_y(y_0, z')$. Аналогично

$$f(y_0, z') - f(y', z') \geq (v', y_0 - y') \geq -L\|y' - y_0\|, \quad (4.10)$$

где $v' \in \partial f_y(y', z')$. Из (4.9) и (4.10)

$$|f(y', z') - f(y_0, z')| \leq L\|y' - y_0\|. \quad (4.11)$$

Функция $f(y_0, z)$ непрерывна по z . Поэтому

$$f(y_0, z') \xrightarrow{z' \rightarrow z_0} f(y_0, z_0).$$

Отсюда, из (4.11) и (4.8) имеем $f(x') \rightarrow f(x_0)$, т. е. функция $f(x)$ непрерывна по совокупности переменных. ■

Теорема 4.1. Если $f(y, z)$ — выпукло-вогнутая на S функция, то она и непрерывна на S по совокупности переменных.

Доказательство. Функция $f(y, z)$ вогнута по z на S_2 при любом фиксированном $y \in S_1$, а потому она и непрерывна по z при фиксированном y . Утверждение теоремы теперь следует из леммы 4.1.

Следствие. Выпукло-вогнутая функция $f(y, z)$ является липшицевой на любом множестве $\Omega = \Omega_1 \times \Omega_2$, где $\Omega_1 \subset S_1$, $\Omega_2 \subset S_2$, Ω_1 и Ω_2 — выпуклые компакты.

Действительно, как и при доказательстве леммы 4.1, можно установить, что множество супердифференциалов $\partial f_z(y, z)$ ограничено на Ω , т. е. существует $L_1 < \infty$ такое, что

$$\|w\| \leq L_1 \quad \forall w \in \partial f_z(y, z), \quad \forall [y, z] \in \Omega. \quad (4.12)$$

Из (4.8), (4.11) и (4.12) теперь заключаем, что функция $f(y, z)$ является липшицевой по совокупности переменных на множестве $\Omega = \Omega_1 \times \Omega_2$. ■

Лемма 4.2. Пусть функция $f(y, z)$ выпукла по y на S_1 при любом фиксированном $z \in S_2$ и непрерывна по z на S_2 при каждом фиксированном $y \in S$, $x_0 = [y_0, z_0] \in S = S_1 \times S_2$. Тогда для любого $\varepsilon > 0$ найдется $\delta > 0$ (зависящее от x_0) такое, что

$$\partial f_y(y, z) \subset \partial f_y(y_0, z_0) + S_{1\varepsilon} \quad \forall [y, z] \in S_\delta(y_0, z_0), \quad (4.13)$$

где

$$S_{1\varepsilon} = \{v \in E_{n_1} \mid \|v\| \leq \varepsilon\},$$

$$S_\delta(y_0, z_0) = \{[y, z] \in S \mid \|y - y_0\| + \|z - z_0\| \leq \delta\}.$$

Доказательство. Предположим противное. Тогда найдутся $\varepsilon_0 > 0$ и последовательность $\{v_k\}$ такие, что

$$\begin{aligned} v_k &\in \partial f_y(y_k, z_k), \quad [y_k, z_k] \rightarrow [y_0, z_0] = x_0, \\ \rho(v_k, \partial f_y(y_0, z_0)) &= \min_{v \in \partial f_y(y_0, z_0)} \|v_k - v\| \geq \varepsilon_0. \end{aligned} \quad (4.14)$$

Поскольку множества $\partial f_y(y, z)$ ограничены в совокупности (это было установлено при доказательстве леммы 4.1), то без ограничения общности можем считать, что

$$v_k \rightarrow v_0. \quad (4.15)$$

Из (4.14) очевидно, что

$$v_0 \notin \partial f_y(y_0, z_0). \quad (4.16)$$

Так как $v_k \in \partial f_y(y_k, z_k)$, то

$$f(y, z_k) - f(y_k, z_k) \geq (v_k, y - y_k) \quad \forall y \in S_1.$$

Отсюда

$$\begin{aligned} f(y, z_0) - f(y_0, z_0) &\geq (v_k, y - y_0) + (v_k, y_0 - y_k) + \\ &+ f(y, z_0) - f(y, z_k) + f(y_k, z_k) - f(y_0, z_0) \quad \forall y \in S_1. \end{aligned} \quad (4.17)$$

Зафиксируем любое $y \in S_1$. По лемме 4.1 функция $f(y, z)$ непрерывна по совокупности переменных. Поэтому из (4.17) и (4.15) при $k \rightarrow \infty$ имеем

$$f(y, z_0) - f(y_0, z_0) \geq (v_0, y - y_0).$$

Так как полученное неравенство справедливо для каждого $y \in S_1$, то отсюда и из (4.1) $v_0 \in \partial f_y(y_0, z_0)$, что противоречит (4.16). Лемма доказана.

Следствие 1. Пусть $f(y, z)$ — выпукло-вогнутая на $S = S_1 \times S_2$ функция, $x_0 = [y_0, z_0] \in S$. Тогда для любого $\varepsilon > 0$ найдется $\delta > 0$ такое, что при всех $[y, z] \in S_\delta(y_0, z_0)$

$$\begin{aligned}\partial f_y(y, z) &\subset \partial f_y(y_0, z_0) + S_{1\varepsilon}, \\ \partial f_z(y, z) &\subset \partial f_z(y_0, z_0) + S_{2\varepsilon},\end{aligned}\quad (4.18)$$

где

$$\begin{aligned}S_{1\varepsilon} &= \{v \in E_{n_1} \mid \|v\| \leq \varepsilon\}, \quad S_{2\varepsilon} = \{w \in E_{n_2} \mid \|w\| \leq \varepsilon\}, \\ S_\delta(y_0, z_0) &= \{[y, z] \in S \mid \|y - y_0\| + \|z - z_0\| \leq \delta\}.\end{aligned}$$

Доказательство аналогично только что проведенному.

Следствие 2. Отображения $\partial f_y(y, z)$ и $\partial f_z(y, z)$ полунепрерывны сверху на S .

2. Найдем теперь производную по направлениям выпукло-вогнутой функции. Пусть

$$x_0 = [y_0, z_0] \in S = S_1 \times S_2, \quad g = [g_1, g_2] \in E_{n_1} \times E_{n_2}.$$

Теорема 4.2. Функция $f(x) \equiv f(y, z)$ дифференцируема в точке x_0 по любому направлению $g = [g_1, g_2] \in E_{n_1} \times E_{n_2}$, причем

$$\frac{\partial f(x_0)}{\partial g} = \max_{v \in \partial f_y(y_0, z_0)} (v, g_1) + \min_{w \in \partial f_z(y_0, z_0)} (w, g_2). \quad (4.19)$$

Доказательство. Правую часть в (4.19) обозначим B . Положим

$$h(\alpha) = \alpha^{-1} [f(y_0 + \alpha g_1, z_0 + \alpha g_2) - f(y_0, z_0)]. \quad (4.20)$$

Надо установить существование предела

$$\lim_{\alpha \rightarrow +0} h(\alpha) \equiv \frac{\partial f(x_0)}{\partial g} \quad (4.21)$$

и найти его вид. Имеем

$$\begin{aligned}f(y_0 + \alpha g_1, z_0 + \alpha g_2) - f(y_0, z_0) &= \\ &= [f(y_0 + \alpha g_1, z_0 + \alpha g_2) - f(y_0 + \alpha g_1, z_0)] + \\ &\quad + [f(y_0 + \alpha g_1, z_0) - f(y_0, z_0)] \equiv A_1 + A_2.\end{aligned}\quad (4.22)$$

Так как функция $f(y, z_0)$ — выпуклая по y , то

$$A_2 \equiv f(y_0 + \alpha g_1, z_0) - f(y_0, z_0) = \alpha \frac{\partial f(y_0, z_0)}{\partial g_1} + o_1(\alpha), \quad (4.23)$$

где

$$\begin{aligned} \alpha_1(\alpha) &\geq 0, \quad \frac{\partial f(y_0, z_0)}{\partial g_1} = \\ &= \lim_{\alpha \rightarrow +0} \alpha^{-1} [f(y_0 + \alpha g_1, z_0) - f(y_0, z_0)] = \max_{v \in \partial f_y(y_0, z_0)} (v, g_1). \end{aligned}$$

Поскольку $f(y, z)$ — вогнутая по z функция, то из соотношения (5.45) гл. I

$$\begin{aligned} f(y_0 + \alpha g_1, z_0 + \alpha g_2) &= \\ &= f(y_0 + \alpha g_1, z_0) + \int_0^\alpha \frac{\partial f(y_0 + \alpha g_1, z_0 + \tau g_2)}{\partial g_2} d\tau, \end{aligned} \quad (4.24)$$

где

$$\begin{aligned} \frac{\partial f(y, z)}{\partial g_2} &= \lim_{\alpha \rightarrow +0} \alpha^{-1} [f(y, z + \alpha g_2) - f(y, z)] = \\ &= \min_{w \in \partial f_z(y, z)} (w, g_2). \end{aligned} \quad (4.25)$$

Зафиксируем $\varepsilon > 0$. Из (4.18) следует, что найдется такое $\alpha(\varepsilon) > 0$, при котором

$$\begin{aligned} \partial f_z(y_0 + \alpha g_1, z_0 + \tau g_2) &\subset \partial f_z(y_0, z_0) + S_{2\varepsilon} \\ \forall \tau \in [0, \alpha(\varepsilon)], \quad \forall \alpha \in [0, \alpha(\varepsilon)]. \end{aligned} \quad (4.26)$$

Если $G \subset G_1 + G_2$, то $\min_{w \in G_1} (w, g_2) \geq \min_{w \in G_1 + G_2} (w, g_2) = \min_{w \in G_1} (w, g_2) + \min_{w \in G_2} (w, g_2)$. Поэтому из (4.24) и (4.26)

$$\begin{aligned} A_1 &= f(y_0 + \alpha g_1, z_0 + \alpha g_2) - f(y_0 + \alpha g_1, z_0) \geq \\ &\geq \alpha \min_{w \in \partial f_z(y_0, z_0)} (w, g_2) - \alpha \varepsilon \|g_2\|. \end{aligned}$$

Поскольку $\min_{w \in S_{2\varepsilon}} (w, g_2) = -\varepsilon \|g_2\|$, то из (4.22) и (4.23)

$$\begin{aligned} h(\alpha) &\geq \max_{v \in \partial f_y(y_0, z_0)} (v, g_1) + \min_{w \in \partial f_z(y_0, z_0)} (w, g_2) - \varepsilon \|g_2\| = \\ &= B - \varepsilon \|g_2\|. \end{aligned}$$

Значит,

$$\lim_{\alpha \rightarrow +0} h(\alpha) \geq B - \varepsilon \|g_2\|.$$

В силу произвольности ε отсюда

$$\lim_{\alpha \rightarrow +0} h(\alpha) \geq B. \quad (4.27)$$

С другой стороны,

$$\begin{aligned} f(y_0 + \alpha g_1, z_0 + \alpha g_2) - f(y_0, z_0) &= \\ &= [f(y_0 + \alpha g_1, z_0 + \alpha g_2) - f(y_0, z_0 + \alpha g_2)] + \\ &\quad + [f(y_0, z_0 + \alpha g_2) - f(y_0, z_0)] \equiv A_3 + A_4. \end{aligned} \quad (4.28)$$

Так как $f(y_0, z)$ — вогнутая функция, то

$$A_4 \equiv f(y_0, z_0 + \alpha g_2) - f(y_0, z_0) = \alpha \frac{\partial f(y_0, z_0)}{\partial g_2} + o_2(\alpha), \quad (4.29)$$

где $o_2(\alpha) \leq 0$,

$$\begin{aligned} \frac{\partial f(y_0, z_0)}{\partial g_2} &= \lim_{\alpha \rightarrow +0} \alpha^{-1} [f(y_0, z_0 + \alpha g_2) - f(y_0, z_0)] = \\ &= \min_{w \in \partial f_z(y_0, z_0)} (w, g_2). \end{aligned}$$

Функция $f(y, z_0 + \alpha g_2)$ выпукла по y , поэтому из формулы (5.26) гл. I

$$\begin{aligned} A_3 &\equiv f[(y_0 + \alpha g_1, z_0 + \alpha g_2) - f(y_0, z_0 + \alpha g_2)] = \\ &= \int_0^\alpha \frac{\partial f(y_0 + \tau g_1, z_0 + \alpha g_2)}{\partial g_1} d\tau. \end{aligned}$$

При этом

$$\frac{\partial f(y_0 + \tau g_1, z_0 + \alpha g_2)}{\partial g_1} = \max_{v \in \partial f_y(y_0 + \tau g_1, z_0 + \alpha g_2)} (v, g_1). \quad (4.30)$$

Возьмем $\varepsilon > 0$. По лемме 4.2 найдется такое $\alpha(\varepsilon) > 0$, что

$$\begin{aligned} \partial f_y(y_0 + \tau g_1, z_0 + \alpha g_2) &\subset \partial f_y(y_0, z_0) + S_{1\varepsilon} \\ \forall \tau \in [0, \alpha(\varepsilon)], \quad \forall \alpha \in [0, \alpha(\varepsilon)]. \end{aligned}$$

Так как для $G \subset G_1 + G_2$ будет

$$\max_{v \in G} (v, g_1) \leq \max_{v \in G_1 + G_2} (v, g_1) = \max_{v \in G_1} (v, g_1) + \max_{v \in G_2} (v, g_1),$$

то из (4.30)

$$A_3 \leq \alpha \max_{v \in \partial f_y(y_0, z_0)} (v, g_1) + \alpha \varepsilon \|g_1\|,$$

поскольку

$$\max_{v \in S_{1\varepsilon}} (v, g_1) = \varepsilon \|g_1\|.$$

Отсюда и из (4.29)

$$h(\alpha) \leq \max_{v \in \partial f_y(y_0, z_0)} (v, g_1) + \min_{w \in \partial f_z(y_0, z_0)} (w, g_2) + \varepsilon \|g_2\| = \\ = B + \varepsilon \|g_2\|.$$

Поэтому

$$\overline{\lim}_{\alpha \rightarrow +0} h(\alpha) \leq B + \varepsilon \|g_2\|.$$

В силу произвольности ε имеем

$$\overline{\lim}_{\alpha \rightarrow +0} h(\alpha) \leq B. \quad (4.31)$$

Сопоставляя (4.27) и (4.31), заключаем, что существует

$$\lim_{\alpha \rightarrow +0} h(\alpha) \quad \text{и} \quad \lim_{\alpha \rightarrow +0} h(\alpha) = \frac{\partial f(x_0)}{\partial g} = B,$$

что и требовалось доказать. ■

Следствие. *Выпукло-вогнутая на S функция квазидифференцируема там, причем*

$$\mathcal{D}f(x_0) = [\underline{\partial f}(x_0), \overline{\partial f}(x_0)],$$

где $\underline{\partial f}(x_0) = [\partial f_y(y_0, z_0), 0_n]$, $\overline{\partial f}(x_0) = [0_n, \partial f_z(y_0, z_0)]$.

Это следует из формулы для производной по направлениям (формула (4.19)).

§ 5. Необходимые условия оптимальности квазидифференцируемой функции на E_n

1. Пусть $f(x)$ — квазидифференцируемая на E_n функция. Установим вначале необходимые условия минимума и максимума.

Теорема 5.1. *Для того чтобы квазидифференцируемая на E_n функция $f(x)$ достигала в точке $x^* \in E_n$ своего наименьшего на E_n значения, необходимо, чтобы*

$$-\overline{\partial f}(x^*) \subset \underline{\partial f}(x^*). \quad (5.1)$$

Доказательство. Если функция $f(x)$ дифференцируема по направлениям в точке x^* , а x^* — точка минимума функции $f(x)$ на E_n , то (см. лемму 6.1 гл. I)

$$\frac{\partial f(x^*)}{\partial g} \geq 0 \quad \forall g \in E_n. \quad (5.2)$$

Поскольку в рассматриваемом случае

$$\frac{\partial f(x^*)}{\partial g} = \max_{v \in \underline{\partial f(x^*)}} (v, g) + \min_{w \in \overline{\partial f(x^*)}} (w, g), \quad (5.3)$$

то из (5.2) и (5.3) имеем

$$\inf_{g \in S_1} [\max_{v \in \underline{\partial f(x^*)}} (v, g) + \min_{w \in \overline{\partial f(x^*)}} (w, g)] = 0, \quad (5.4)$$

где (здесь и далее в этом параграфе)

$$S_1 = \{g \in E_n \mid \|g\| \leq 1\}.$$

Из (5.4)

$$\begin{aligned} \min_{g \in S_1} \min_{w \in \overline{\partial f(x^*)}} \max_{v \in \underline{\partial f(x^*)}} (v + w, g) = \\ = \min_{g \in S_1} \min_{w \in \overline{\partial f(x^*)}} \max_{v \in [\underline{\partial f(x^*)} + w]} (v, g) = 0. \end{aligned} \quad (5.5)$$

Так как множества S_1 и $\overline{\partial f(x^*)}$ — выпуклые компакты, то из (5.5)

$$\min_{w \in \overline{\partial f(x^*)}} \min_{g \in S_1} \max_{v \in [\underline{\partial f(x^*)} + w]} (v, g) = 0,$$

т. е.

$$\min_{g \in S_1} \max_{v \in [\underline{\partial f(x^*)} + w]} (v, g) = 0 \quad \forall w \in \overline{\partial f(x^*)}. \quad (5.6)$$

По следствию из леммы 6.1 гл. I и по следствию из теоремы 6.1 гл. I получаем, что условие (5.6) эквивалентно условию

$$0 \in \underline{\partial f(x^*)} + w \quad \forall w \in \overline{\partial f(x^*)}.$$

Но это и означает, что

$$-w \in \underline{\partial f(x^*)} \quad \forall w \in \overline{\partial f(x^*)},$$

т. е. $-\overline{\partial f(x^*)} \subset \underline{\partial f(x^*)}$, что и требовалось доказать. ■

Аналогично доказывается

Теорема 5.2. Для того чтобы квазидифференцируемая на E_n функция $f(x)$ достигала в точке $x^{**} \in E_n$ своего наибольшего на E_n значения, необходимо, чтобы

$$-\underline{\partial f(x^{**})} \subset \overline{\partial f(x^{**})}. \quad (5.7)$$

Определение. Точку $x^* \in E_n$, удовлетворяющую (5.1), будем называть *инф-стационарной точкой* функции

$f(x)$ на E_n , а точку $x^{**} \in E_n$, удовлетворяющую (5.7), будем называть *sup-стационарной* точкой функции $f(x)$ на E_n .

Очевидно, что если для некоторой точки $x_0 \in E_n$ оказалось, что $-\partial f(x_0) = \partial f(x_0)$, то точка x_0 является одновременно и *inf-стационарной*, и *sup-стационарной*.

2. Пусть x^* — *inf-стационарная* точка функции $f(x)$ на E_n , т. е. имеет место (5.1). По определению

$$f(x^* + \alpha g) = f(x^*) + \alpha \frac{\partial f(x^*)}{\partial g} + o(\alpha, g), \quad (5.8)$$

где

$$\frac{o(\alpha, g)}{\alpha} \xrightarrow{\alpha \rightarrow +0} 0. \quad (5.9)$$

Теорема 5.3. Если x^* — *inf-стационарная* точка функции $f(x)$ на E_n и при этом

$$-\partial f(x^*) \subset \text{int } \partial f(x^*), \quad (5.10)$$

а соотношение (5.9) имеет место равномерно по $g \in S_1$, то точка x^* — точка строгого локального минимума функции $f(x)$.

Доказательство. Из (5.10) следует, что найдется $r > 0$ такое, что

$$S_r \subset [\partial f(x^*) + w] \quad \forall w \in \overline{\partial f(x^*)},$$

где $S_r = \{z \in E_n \mid \|z\| \leq r\}$.

Тогда

$$\max_{v \in [\partial f(x^*) + w]} (v, g) \geq r \quad \forall w \in \overline{\partial f(x^*)}, \quad \forall g: \|g\| = 1.$$

Из (5.5)

$$\min_{\substack{g \in E_n \\ \|g\|=1}} \min_{w \in \overline{\partial f(x^*)}} \max_{v \in [\partial f(x^*) + w]} (v, g) \geq r,$$

т. е.

$$\frac{\partial f(x^*)}{\partial g} \geq r \quad \forall g \in E_n: \|g\| = 1. \quad (5.11)$$

Поскольку (5.9) имеет место равномерно по $g \in S_1$, то из (5.8) и (5.11) найдется $\alpha_0 > 0$ такое, что

$$f(x^* + \alpha g) \geq f(x^*) + \alpha r 2^{-1} \\ \forall \alpha \in [0, \alpha_0], \quad \forall g \in E_n: \|g\| = 1,$$

а это и означает, что x^* — точка строгого локального минимума функции $f(x)$. ■

Аналогично доказывается

Теорема 5.4. Если x^{**} — $\sup$ -стационарная точка функции $f(x)$ на E_n , при этом

$$-\partial f(x^{**}) \subset \text{int } \overline{\partial f(x^{**})} \quad (5.12)$$

и соотношение (5.9) имеет место равномерно по $g \in S_1$, то точка x^{**} — точка строгого локального максимума функции $f(x)$.

В качестве примеров рассмотрим некоторые функции, изученные в § 3. Точка $x_0 = (0, 0)$ является $\inf$ -стационарной (а в действительности и точкой минимума) для функций, описанных в примерах 1, 3, 6, 7, 8, 10, 11, 12, 13 § 3. Точка x_0 является $\sup$ -стационарной для функций, описанных в примерах 2, 4. Точка x_0 является точкой строгого локального минимума для функций, описанных в примерах 6 и 8.

3. Пусть точка x_0 не является $\inf$ -стационарной, т. е. условие (5.1) не выполнено. Тогда не выполнено и условие (5.2), т. е.

$$\inf_{g \in S_1} \frac{\partial f(x_0)}{\partial g} = -a < 0. \quad (5.13)$$

Отсюда и из (5.3)

$$\begin{aligned} \min_{g \in S_1} \min_{w \in \overline{\partial f(x_0)}} \max_{v \in \underline{\partial f(x_0)}} (v + w, g) = \\ = \min_{w \in \overline{\partial f(x_0)}} \min_{g \in S_1} \max_{v \in \underline{\partial f(x_0)}} (v + w, g) = -a. \end{aligned} \quad (5.14)$$

Как и при доказательстве теоремы 6.2 гл. I, показывается, что

$$\min_{g \in S_1} \max_{v \in \underline{\partial f(x_0)}} (v + w, g) = - \min_{v \in \underline{\partial f(x_0)}} \|v + w\|.$$

Отсюда, из (5.13) и (5.14)

$$a = - \max_{w \in \overline{\partial f(x_0)}} \min_{v \in \underline{\partial f(x_0)}} \|v + w\|. \quad (5.15)$$

Выберем $w_0 \in \overline{\partial f(x_0)}$ и $v_0 \in \underline{\partial f(x_0)}$ так, чтобы

$$\max_{w \in \overline{\partial f(x_0)}} \min_{v \in \underline{\partial f(x_0)}} \|v + w\| = \min_{v \in \underline{\partial f(x_0)}} \|v + w_0\| = \|v_0 + w_0\|.$$

Тогда направление $g_0 = -(v_0 + w_0) \|v_0 + w_0\|^{-1}$ является направлением наискорейшего спуска функции $f(x)$ в точке x_0 . Оно может оказаться не единственным.

Аналогично показывается, что если точка x_0 не является суп-стационарной, т. е. $-\underline{\partial f(x_0)} \not\subset \overline{\partial f(x_0)}$, то

$$a \equiv \sup_{g \in S_1} \frac{\partial f(x_0)}{\partial g} = \max_{v \in \underline{\partial f(x_0)}} \min_{w \in \overline{\partial f(x_0)}} \|v + w\|. \quad (5.16)$$

Если $a > 0$, то направление $g_1 = (v_1 + w_1) \|v_1 + w_1\|^{-1}$, где $\min_{w \in \overline{\partial f(x_0)}} \|v_1 + w\| = \max_{v \in \underline{\partial f(x_0)}} \min_{w \in \overline{\partial f(x_0)}} \|v + w\|$, $\|v_1 + w_1\| = \min_{w \in \overline{\partial f(x_0)}} \|v_1 + w\|$, является направлением наискорейшего подъема функции $f(x)$ в точке x_0 . Оно тоже может оказаться не единственным.

Найдем расстояние между множествами $\underline{\partial f(x_0)}$ и $-\overline{\partial f(x_0)}$ в метрике Хаусдорфа (см. (2.1) гл. I):

$$\begin{aligned} \rho(\underline{\partial f(x_0)}, -\overline{\partial f(x_0)}) &\equiv \\ &\equiv \max_{w \in \overline{\partial f(x_0)}} \min_{v \in \underline{\partial f(x_0)}} \|v + w\| + \max_{v \in \underline{\partial f(x_0)}} \min_{w \in \overline{\partial f(x_0)}} \|v + w\|. \end{aligned}$$

Первое слагаемое здесь есть уклонение множества $-\overline{\partial f(x_0)}$ от множества $\underline{\partial f(x_0)}$ (в смысле Хаусдорфа), а второе слагаемое есть уклонение множества $\underline{\partial f(x_0)}$ от множества $-\overline{\partial f(x_0)}$.

Поэтому из (5.15) и (5.16) получаем, что сумма абсолютных величин скоростей наискорейшего спуска и наискорейшего подъема функции $f(x)$ в точке x_0 равна расстоянию между множествами $\underline{\partial f(x_0)}$ и $-\overline{\partial f(x_0)}$ в метрике Хаусдорфа.

Пример 1. Пусть $x = (x^{(1)}, x^{(2)}) \in E_2$, $x_0 = (0, 0)$. Возьмем функцию

$$f(x) = \min \{ |x^{(1)}| + |x^{(2)}|, |x^{(1)}| - |x^{(2)}| \}.$$

Эта функция рассматривалась в примере 9 § 3. Там было показано, что $\underline{\partial f(x_0)}$ есть (рис. 29) прямоугольник с вершинами $(2, 1)$, $(-2, 1)$, $(-2, -1)$, $(2, -1)$, а множество $\overline{\partial f(x_0)}$ есть прямоугольник с вершинами $(1, 2)$,

$(-1, 2)$, $(-1, -2)$, $(1, -2)$. Очевидно, что в силу симметричности $\overline{\partial f}(x_0) = -\partial f(x_0)$. Применяя формулы (5.15) и (5.16), найдем, что существует два направления

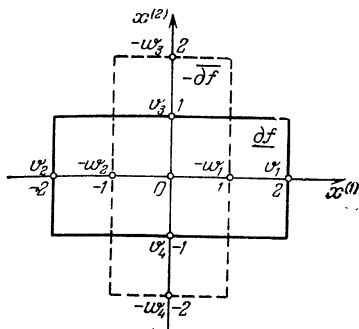


Рис. 29.

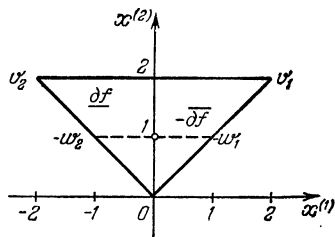


Рис. 30.

наискорейшего спуска $g_0 = (0, 1)$ и $g'_0 = (0, 1)$ и два направления наискорейшего подъема $g_1 = (1, 0)$ и $g'_1 = (-1, 0)$. Очевидно также, что

$$\frac{\partial f(x_0)}{\partial g_0} = \frac{\partial f(x_0)}{\partial g'_0} = -1, \quad \frac{\partial f(x_0)}{\partial g_1} = \frac{\partial f(x_0)}{\partial g'_1} = 1.$$

Пример 2. Пусть теперь

$$x = (x^{(1)}, x^{(2)}) \in E_2, \quad x_0 = (0, 0),$$

$$f(x) = ||x^{(1)}| + x^{(2)}|.$$

Эта функция рассматривалась в примере 10 § 3. Там было показано, что (рис. 30)

$$\underline{\partial f}(x_0) = \text{co} \{(0, 0), (2, 2), (-2, 2)\},$$

$$\overline{\partial f}(x_0) = \text{co} \{(1, -1), (-1, -1)\}.$$

Тогда $-\overline{\partial f}(x_0) = \text{co} \{(1, 1), (-1, 1)\}$.

Точка x_0 является inf-стационарной, т. е. $\inf_{g \in E_n} \frac{\partial f(x_0)}{\partial g} = 0$. Очевидно, что существует два направления наискорейшего подъема

$$g_1 = \left(\frac{\sqrt{2}}{2}, \frac{\sqrt{2}}{2} \right), \quad g'_1 = \left(-\frac{\sqrt{2}}{2}, \frac{\sqrt{2}}{2} \right)$$

и при этом

$$\sup_{g \in S_1} \frac{\partial f(x_0)}{\partial g} = \frac{\partial f(x_0)}{\partial g_1} = \frac{\partial f(x_0)}{\partial g'_1} = \sqrt{2}.$$

Таким образом, проверка необходимых условий минимума сводится к разысканию уклонения множества $-\bar{\partial}f(x_0)$ от множества $\underline{\partial}f(x_0)$ (в смысле Хаусдорфа), а проверка необходимых условий максимума сводится к разысканию уклонения множества $\underline{\partial}f(x_0)$ от множества $-\bar{\partial}f(x_0)$. При решении этих задач одновременно находятся направления наискорейшего спуска или наискорейшего подъема (в зависимости от того, какое условие проверялось — минимума или максимума).

§ 6. Квазидифференцируемые множества

1. Пусть $\Omega \subset E_n$ — замкнутое множество, $x_0 \in \Omega$. Вектор $v \in E_n$ назовем *возможным в широком смысле направлением* множества Ω в точке x_0 , если существуют число $\lambda \geq 0$ и последовательность $\{x_i\}$ такие, что

$$\begin{aligned} x_i \in \Omega, \quad x_i \neq x_0, \quad x_i \rightarrow x_0, \\ v_i = (x_i - x_0) \|x_i - x_0\|^{-1} \rightarrow v, \quad v = \lambda v_0. \end{aligned} \quad (6.1)$$

Очевидно, что множество возможных в широком смысле направлений образует замкнутый конус с вершиной в начале координат. Обозначим его через $\Gamma(x_0)$ и будем называть *конусом возможных в широком смысле направлений* множества Ω в точке x_0 .

Рассмотрим случай, когда Ω имеет вид

$$\Omega = \{x \in E_n \mid h(x) \leq 0\},$$

где $h(x)$ — непрерывная квазидифференцируемая на E_n функция. Предположим, что $\Omega \neq \emptyset$. Будем говорить, что множество Ω имеет *квазидифференцируемую границу*, или просто называть его *квазидифференцируемым*. Очевидно, что Ω — замкнутое множество. Пусть $x_0 \in \Omega$. Положим

$$\begin{aligned} \gamma(x_0) &= \left\{ g \in E_n \mid \frac{\partial h(x_0)}{\partial g} < 0 \right\}, \\ \gamma_1(x_0) &= \left\{ g \in E_n \mid \frac{\partial h(x_0)}{\partial g} \leq 0 \right\}. \end{aligned}$$

Очевидно, что если $h(x_0) = 0$, то

$$\gamma(x_0) \subset \Gamma(x_0). \quad (6.2)$$

Пусть $h(x_0) = 0$. Будем говорить, что в точке x_0 выполнено условие невырожденности, если

$$\bar{\gamma}(x_0) = \gamma_1(x_0). \quad (6.3)$$

Здесь $\bar{\gamma}$ — замыкание множества γ . Нетрудно построить пример, показывающий, что условие (6.3) выполнено не всегда.

Лемма 6.1. Если $h(x_0) < 0$, то $\Gamma(x_0) = E_n$. Если же $h(x_0) = 0$ и в точке x_0 выполнено условие невырожденности (6.3), а функция $h(x)$ является липшицевой в окрестности точки x_0 , то

$$\Gamma(x_0) = \gamma_1(x_0). \quad (6.4)$$

Доказательство. Если $h(x_0) < 0$, то $\Gamma(x_0) = E_n$, и все доказано. Пусть теперь $h(x_0) = 0$. Из (6.2) и (6.3) вытекает

$$\gamma_1(x_0) = \bar{\gamma}(x_0) \subset \Gamma(x_0). \quad (6.5)$$

Осталось установить обратное включение. Возьмем $g \in \Gamma(x_0)$. Тогда существуют число $\lambda \geq 0$ и последовательность $\{x_i\}$ такие, что

$$\begin{aligned} x_i &\in \Omega, \quad x_i \neq x_0, \quad x_i \rightarrow x_0, \\ g_i &= (x_i - x_0) \|x_i - x_0\|^{-1} \rightarrow g, \quad g = \lambda g_0. \end{aligned} \quad (6.6)$$

Если $\lambda = 0$, то $g = 0$ и, очевидно, $g = 0 \in \gamma_1(x_0)$. Пусть теперь $\lambda > 0$. Тогда

$$\frac{\partial h(x_0)}{\partial g} = \lim_{\alpha \rightarrow +0} \alpha^{-1} [h(x_0 + \alpha g) - h(x_0)] = \lim_{\alpha \rightarrow +0} H(\alpha),$$

где $H(\alpha) = \alpha^{-1} [h(x_0 + \alpha g) - h(x_0)]$. Возьмем

$$\alpha_i = \lambda^{-1} \|x_i - x_0\|. \quad (6.7)$$

Ясно, что $\alpha_i \xrightarrow{i \rightarrow \infty} 0$. Имеем

$$\begin{aligned} H(\alpha_i) &= \alpha_i^{-1} [h(x_0 + \alpha_i g) - h(x_0)] = \\ &= \alpha_i^{-1} [h(x_0 + \alpha_i \lambda g_0) - h(x_0 + \alpha_i \lambda g_i)] + \\ &+ h(x_0 + \alpha_i \lambda g_i) - h(x_0) = \alpha_i^{-1} [h(x_0 + \alpha_i \lambda g_0) - \\ &- h(x_0 + \alpha_i \lambda g_i)] + \alpha_i^{-1} [h(x_0 + \alpha_i \lambda g_i) - h(x_0)]. \end{aligned} \quad (6.8)$$

Функция $h(x)$ липшицева в окрестности точки x_0 , т. е. существуют числа $\delta > 0$ и $L < \infty$ такие, что

$$|h(x) - h(x')| \leq L \|x - x'\| \quad \forall x, x' \in S_\delta(x_0).$$

Значит, при достаточно больших i

$$|h(x_0 + \alpha_i \lambda g_0) - h(x_0 + \alpha_i \lambda g_i)| \leq L \alpha_i \lambda \|g_i - g_0\|. \quad (6.9)$$

Так как из (6.7) $x_0 + \alpha_i \lambda g_i = x_i \in \Omega$, то $h(x_0 + \alpha_i \lambda g_i) \leq h(x_0) = 0$. Отсюда, из (6.8) и (6.9)

$$H(\alpha_i) \leq \alpha_i^{-1} [h(x_0 + \alpha_i \lambda g_0) - h(x_0 + \alpha_i \lambda g_i)] \rightarrow 0,$$

т. е. $\frac{\partial h(x_0)}{\partial g} = \lim_{i \rightarrow \infty} H(\alpha_i) \leq 0$, а это значит, что $g \in \gamma_1(x_0)$.

Таким образом,

$$\Gamma(x_0) \subset \gamma_1(x_0). \quad (6.10)$$

Утверждение леммы теперь вытекает из (6.5) и (6.10). ■

Лемма 6.2. *Справедливо соотношение*

$$\gamma_1(x_0) = \bigcup_{w \in \partial h(x_0)} [-\mathcal{K}^+(\partial h(x_0) + w)]. \quad (6.11)$$

Доказательство. По определению

$$\gamma_1(x_0) = \left\{ g \in E_n \mid \frac{\partial h(x_0)}{\partial g} \leq 0 \right\}.$$

Так как

$$\begin{aligned} \frac{\partial h(x_0)}{\partial g} &= \max_{v \in \partial h(x_0)} (v, g) + \min_{w \in \partial h(x_0)} (w, g) = \\ &= \min_{w \in \partial h(x_0)} \left[\max_{v \in [\partial h(x_0) + w]} (v, g) \right], \end{aligned}$$

то

$$\gamma_1(x_0) = \bigcup_{w \in \partial h(x_0)} \Gamma_w, \quad (6.12)$$

где

$$\Gamma_w = \{ g \in E_n \mid (v, g) \leq 0 \quad \forall v \in [\partial h(x_0) + w] \}.$$

По лемме 3.9 гл. I

$$\begin{aligned} \Gamma_w^+ &= \{ q \in E_n \mid (q, g) \geq 0 \quad \forall g \in \Gamma_w \} = \\ &= -\overline{\mathcal{K}}(\text{co}(\partial h(x_0) + w)) = -\overline{\mathcal{K}}(\partial h(x_0) + w). \end{aligned}$$

Поскольку Γ_w — замкнутый выпуклый конус, то (см. лемму 3.3 гл. I)

$$\Gamma_w^{++} = \Gamma_w, \quad \text{т. е.} \quad \Gamma_w = -\mathcal{K}^+(\overline{\partial h(x_0)} + w).$$

Отсюда и из (6.12) вытекает (6.11). ■

Лемма 6.3. *Имеет место соотношение*

$$\gamma(x_0) = E_n \setminus \bigcup_{v \in \underline{\partial h(x_0)}} \mathcal{K}^+(\overline{\partial h(x_0)} + v). \quad (6.13)$$

Доказательство. Обозначим $\mathcal{D}$ множество в правой части (6.13). Возьмем любое $g \in \gamma(x_0)$; тогда $\frac{\partial h(x_0)}{\partial g} < 0$, т. е. $\max_{v \in \underline{\partial h(x_0)}} \min_{w \in [\overline{\partial h(x_0)} + v]} (w, g) < 0$. Значит,

$$\min_{w \in [\overline{\partial h(x_0)} + v]} (w, g) < 0 \quad \forall v \in \underline{\partial h(x_0)}.$$

Поэтому

$$g \notin \mathcal{K}^+(\overline{\partial h(x_0)} + v) \quad \forall v \in \underline{\partial h(x_0)}.$$

Отсюда

$$g \notin \bigcup_{v \in \underline{\partial h(x_0)}} \mathcal{K}^+(\overline{\partial h(x_0)} + v),$$

т. е.

$$g \in E_n \setminus \bigcup_{v \in \underline{\partial h(x_0)}} \mathcal{K}^+(\overline{\partial h(x_0)} + v) \equiv \mathcal{D}.$$

Поскольку g — произвольный вектор из $\gamma(x_0)$, то

$$\gamma(x_0) \subset \mathcal{D}. \quad (6.14)$$

Возьмем теперь $g \in \mathcal{D}$, т. е. $g \notin \bigcup_{v \in \underline{\partial h(x_0)}} \mathcal{K}^+(\overline{\partial h(x_0)} + v)$.

Тогда

$$g \notin \mathcal{K}^+(\overline{\partial h(x_0)} + v) \quad \forall v \in \underline{\partial h(x_0)}.$$

Отсюда

$$\min_{w \in [\overline{\partial h(x_0)} + v]} (w, g) < 0 \quad \forall v \in \underline{\partial h(x_0)}. \quad (6.15)$$

Множество $\underline{\partial h(x_0)}$ — компакт, поэтому из (6.15)

$$\max_{v \in \underline{\partial h(x_0)}} \min_{w \in [\overline{\partial h(x_0)} + v]} (w, g) < 0,$$

т. е. $\frac{\partial h(x_0)}{\partial g} < 0$, а это и значит, что $g \in \gamma(x_0)$. Итак $\mathcal{D} \subset \gamma(x_0)$. Отсюда и из (6.14) получаем (6.13). ■

2. Проиллюстрируем введенные в п. 1 определения множеств $\Gamma(x_0)$, $\gamma(x_0)$ и $\gamma_1(x_0)$.

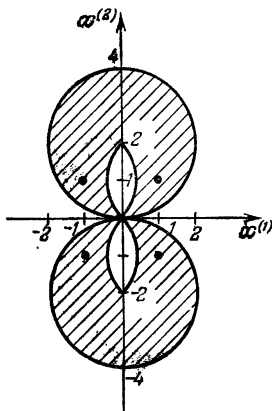


Рис. 31.

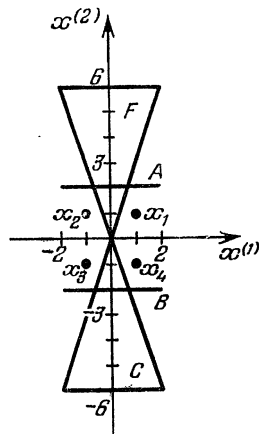


Рис. 32.

Пример 1. Пусть $x = (x^{(1)}, x^{(2)}) \in E_2$. Возьмем точки

$$\begin{aligned} x_1 &= (1, 1), & x_2 &= (-1, 1), & x_3 &= (-1, -1), \\ x_4 &= (1, -1), & x_5 &= (0, 2), & x_6 &= (0, -2) \end{aligned}$$

и построим функции

$$\begin{aligned} h_1(x) &= (x - x_1)^2 - 2, & h_2(x) &= (x - x_2)^2 - 2, \\ h_3(x) &= (x - x_3)^2 - 2, & h_4(x) &= (x - x_4)^2 - 2, \\ h_5(x) &= (x - x_5)^2 - 4, & h_6(x) &= (x - x_6)^2 - 4, \\ h_7(x) &= \max\{h_1(x), h_2(x)\}, & h_8(x) &= \max\{h_3(x), h_4(x)\}, \\ h_9(x) &= -h_7(x), & h_{10}(x) &= -h_8(x), \\ h_{11}(x) &= \max\{h_5(x), h_9(x)\}, & h_{12}(x) &= \max\{h_6(x), h_{10}(x)\}, \\ h(x) &= \min\{h_{11}(x), h_{12}(x)\}. \end{aligned}$$

Образуем множество

$$\Omega = \{x \in E_2 \mid h(x) \leq 0\}.$$

Это множество изображено на рис. 31. Возьмем точку $x_0 = (0, 0)$ и найдем $\mathcal{D}h(x_0)$, конусы $\gamma(x_0)$, $\gamma_1(x_0)$, $\Gamma(x_0)$. Прежде всего, заметим, что

$$\begin{aligned}\mathcal{D}h_1(x_0) &= [-2x_1, 0], & \mathcal{D}h_2(x) &= [-2x_2, 0], \\ \mathcal{D}h_3(x_0) &= [-2x_3, 0], & \mathcal{D}h_4(x_0) &= [-2x_4, 0], \\ \mathcal{D}h_5(x_0) &= [-2x_5, 0], & \mathcal{D}h_6(x_0) &= [-2x_6, 0]\end{aligned}$$

(точнее говоря, надо писать $\mathcal{D}h_1(x_0) = [\{-2x_1\}, \{0\}]$ и т. д.). По лемме 2.2

$$\begin{aligned}\mathcal{D}h_7(x_0) &= [\text{co}\{-2x_1, -2x_2\}, 0] = [-\text{co}\{2x_1, 2x_2\}, 0], \\ \mathcal{D}h_8(x_0) &= [\text{co}\{-2x_3, -2x_4\}, 0] = [-\text{co}\{2x_3, 2x_4\}, 0].\end{aligned}$$

Так как $h_9(x) = -h_7(x)$, $h_{10}(x) = -h_8(x)$, то

$$\mathcal{D}h_9(x_0) = [0, A], \quad \mathcal{D}h_{10}(x_0) = [0, B],$$

где

$$A = \text{co}\{2x_1, 2x_2\}, \quad B = \text{co}\{2x_3, 2x_4\}.$$

Снова по лемме 2.2

$$\mathcal{D}h_{11}(x_0) = [C, D], \quad \mathcal{D}h_{12}(x_0) = [F, G],$$

где

$$\begin{aligned}C &= \text{co}\{-2x_5 - A, 0\}, & D &= A, \\ F &= \text{co}\{-2x_6 - B, 0\}, & G &= B.\end{aligned}$$

Множества A, B, C, F изображены на рис. 32 (множества A и B — отрезки, множества C и F — треугольники). Теперь по лемме 2.3 имеем

$$\mathcal{D}h(x_0) = [\underline{\partial h(x_0)}, \overline{\partial h(x_0)}],$$

где

$$\begin{aligned}\underline{\partial h(x_0)} &= \underline{\partial h_{11}(x_0)} + \underline{\partial h_{12}(x_0)} = C + F, \\ \overline{\partial h(x_0)} &= \text{co}\{\overline{\partial h_{11}} - \underline{\partial h_{12}}, \overline{\partial h_{12}} - \underline{\partial h_{11}}\} = \\ &= \text{co}\{D - F, G - C\} = \text{co}\{A - F, B - C\}.\end{aligned}$$

Множества $\underline{\partial h(x_0)}$ и $\overline{\partial h(x_0)}$ изображены на рис. 33 (множество $\underline{\partial h(x_0)}$ ограничено сплошной жирной линией, а $\overline{\partial h(x_0)}$ — пунктирной). Так как

$$\frac{\partial h(x_0)}{\partial g} = \max_{v \in \underline{\partial h(x_0)}} (v, g) + \min_{w \in \overline{\partial h(x_0)}} (w, g),$$

ТО

$$\begin{aligned} \gamma(x_0) = & \{v = \lambda g \mid \lambda \geq 0, g = (1, \alpha), \alpha \in \{(0, 1)\} \cup \{(0, -1)\}\} \cup \\ & \cup \{v = \lambda g \mid \lambda \geq 0, g = (-1, \alpha), \alpha \in \{(0, 1)\} \cup \{(0, -1)\}\}, \\ \gamma_1(x_0) = & \{v = \lambda g \mid \lambda \geq 0, g = (1, \alpha), \alpha \in [-1, 1]\} \cup \\ & \cup \{v = \lambda g \mid \lambda \geq 0, g = (-1, \alpha), \alpha \in [-1, 1]\}. \end{aligned}$$

Очевидно, что $\gamma_1(x_0) = \overline{\gamma(x_0)}$, т. е. в точке x_0 выполнено условие невырожденности, а тогда по лемме 6.1 $\Gamma(x_0) = \gamma_1(x_0)$.

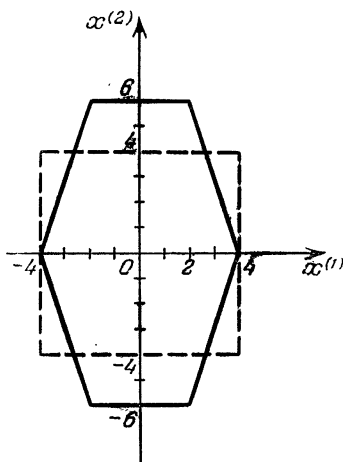


Рис. 33.

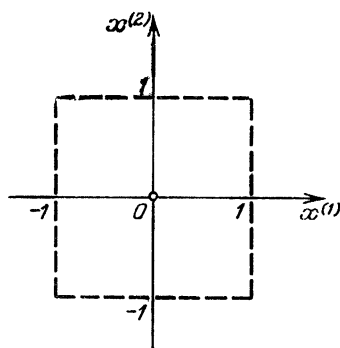


Рис. 34.

Пример 2. Пусть $x = (x^{(1)}, x^{(2)}) \in E_2$, $x_0 = (0, 0)$,

$$\Omega = \{x \in E_2 \mid h(x) \leq 0\},$$

где $h(x) = (x^{(1)} + 1)^2 + (x^{(2)} + 1)^2 - |x^{(1)}| - |x^{(2)}|$. Ясно, что $x_0 \in \Omega$. Поскольку $h(x) = h_1(x) + h_2(x)$, где

$$h_1(x) = (x^{(1)} + 1)^2 + (x^{(2)} + 1)^2,$$

$$h_2(x) = -|x^{(1)}| - |x^{(2)}|,$$

функция $h_1(x)$ гладкая, а функция $h_2(x)$ вогнутая, то можно взять (рис. 34)

$$\partial h(x_0) = \{(2, 2)\},$$

$$\overline{\partial h(x_0)} = \text{co} \{(1, 1), (-1, 1), (-1, -1), (1, -1)\}.$$

Задание 6.1. Пусть

$$\begin{aligned} x = (x^{(1)}, x^{(2)}) \in E_2, \quad x_0 = (0, 0), \quad x_1 = (1, 1), \quad x_2 = (-1, 1), \\ x_3 = (-1, -1), \quad x_4 = (1, -1), \quad x_5 = (2, 2), \quad x_6 = (-2, 2), \\ x_7 = (-2, -2), \quad x_8 = (2, -2). \end{aligned}$$

Образует множества

$$\begin{aligned} S_1 &= \{x \in E_2 \mid (x - x_1)^2 \leq 2\}, \quad S_2 = \{x \mid (x - x_2)^2 \leq 2\} \\ S_3 &= \{x \mid (x - x_3)^2 \leq 2\}, \quad S_4 = \{x \mid (x - x_4)^2 \leq 2\}, \\ S_5 &= \{x \mid (x - x_5)^2 \leq 8\}, \quad S_6 = \{x \mid (x - x_6)^2 \leq 8\}, \\ S_7 &= \{x \mid (x - x_7)^2 \leq 8\}, \quad S_8 = \{x \mid (x - x_8)^2 \leq 8\} \end{aligned}$$

Построим множество $\Omega = \Omega_1 \cup \Omega_2$, где $\Omega_1 = \overline{\Omega_3 \setminus \Omega_4}$, $\Omega_2 = \overline{\Omega_5 \setminus \Omega_6}$, $\Omega_3 = S_5 \cap S_6$, $\Omega_4 = S_1 \cap S_2$, $\Omega_5 = S_7 \cap S_8$, $\Omega_6 = S_3 \cap S_4$. Множество Ω изображено на рис. 37.

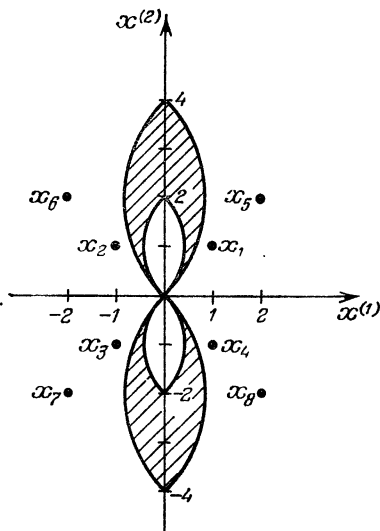


Рис. 37

Найти функцию $h(x)$ такую, что

$$\Omega = \{x \in E_2 \mid h(x) \leq 0\};$$

построить $\mathcal{D}h(x_0)$, $\gamma(x_0)$, $\gamma_1(x_0)$, $\Gamma(x_0)$. Проверить, выполнено ли в точке x_0 условие невырожденности.

§ 7. Необходимые условия оптимальности квазидифференцируемой функции на квазидифференцируемом множестве

1. Пусть на открытом множестве $S \subset E_n$ задана конечная непрерывная функция $f(x)$, и пусть $\Omega \subset S$ — замкнутое множество.

Лемма 7.1. Пусть функция $f(x)$ липшицева в окрестности точки $x^* \in \Omega$ и дифференцируема в точке x^* по направлениям. Для того чтобы функция $f(x)$ достигала в точке x^* своего наименьшего на Ω значения, необходимо, чтобы

$$\frac{\partial f(x^*)}{\partial g} \geq 0 \quad \forall g \in \Gamma(x^*), \quad (7.1)$$

где $\Gamma(x^*)$ — конус возможных в широком смысле направлений множества Ω в точке x^* (см. § 6).

Доказательство. Допустим противное; тогда существует $g_0 \in \Gamma(x^*)$ такое, что

$$\frac{\partial f(x^*)}{\partial g_0} = -a < 0. \quad (7.2)$$

Из (6.1) следует, что найдутся число $\lambda > 0$ и последовательность $\{x_i\}$ такие, что

$$\begin{aligned} x_i &\in \Omega, \quad x_i \neq x^*, \quad x_i \rightarrow x^*, \\ v_i &= (x_i - x^*) \|x_i - x^*\|^{-1} \rightarrow v_0, \quad g_0 = \lambda v_0. \end{aligned}$$

Поскольку

$$f(x_i) = f(x^* + x_i - x^*) = f(x^* + \|x_i - x^*\| v_i) = f(x^* + \alpha_i v_i),$$

где $\alpha_i = \|x_i - x^*\|$, то по лемме 5.6 гл. I

$$\begin{aligned} \lim_{i \rightarrow \infty} \alpha_i^{-1} [f(x_i) - f(x^*)] &= \\ &= \lim_{i \rightarrow \infty} \alpha_i^{-1} [f(x^* + \alpha_i v_i) - f(x^*)] = \frac{\partial f(x^*)}{\partial v_0}. \end{aligned}$$

Отсюда

$$f(x_i) = f(x_0) + \alpha_i \frac{\partial f(x^*)}{\partial v_0} + o(\alpha_i). \quad (7.3)$$

Но

$$\begin{aligned} \frac{\partial f(x^*)}{\partial v_0} &= \lim_{\alpha \rightarrow +0} \alpha^{-1} [f(x^* + \alpha v_0) - f(x^*)] = \\ &= \lambda^{-1} \lim_{\alpha \rightarrow +0} \lambda \alpha^{-1} [f(x^* + \lambda \alpha^{-1} (\lambda v_0)) - f(x^*)] = \lambda^{-1} \frac{\partial f(x^*)}{\partial g_0}. \end{aligned}$$

Из (7.2) и (7.3) теперь получаем, что при достаточно больших i будет $f(x_i) < f(x^*)$, что противоречит предположению о том, что x^* — точка минимума функции $f(x)$ на множестве Ω , ибо $x_i \in \Omega$. Лемма доказана.

Пусть теперь $\Omega \subset E_n$ — квазидифференцируемое множество, т. е.

$$\Omega = \{x \in E_n \mid h(x) \leq 0\},$$

где $h(x)$ — непрерывная квазидифференцируемая на E_n функция. По-прежнему предполагаем, что $\Omega \subset S$. Везде далее в этом параграфе функция $f(x)$ квазидифференцируема на S .

Теорема 7.1. Пусть функция $f(x)$ липшицева в окрестности точки $x^* \in \Omega$. Если $h(x^*) = 0$, то предполагается, что функция $h(x)$ липшицева в окрестности точки x^* , а в точке x^* выполнено условие невырожденности (6.3). Для того чтобы функция $f(x)$ достигала в точке x^* своего наименьшего на Ω значения, необходимо, чтобы

$$-\overline{\partial f(x^*)} \subset \underline{\partial f(x^*)}, \text{ если } h(x^*) < 0, \quad (7.4)$$

и для всех $w \in \overline{\partial f(x^*)}$, $w' \in \overline{\partial h(x^*)}$

$$(\underline{\partial f(x^*)} + w) \cap [-\mathcal{H}(\overline{\partial h(x^*)} + w')] \neq \emptyset, \text{ если } h(x^*) = 0. \quad (7.5)$$

Здесь, как обычно, $\mathcal{H}(A)$ — коническая оболочка множества A .

Доказательство. В случае $h(x^*) < 0$ условие (7.4) следует из теоремы 5.1. Пусть теперь $h(x^*) = 0$. По лемме 7.1.

$$\frac{\partial f(x^*)}{\partial g} = \max_{v \in \underline{\partial f(x^*)}} (v, g) + \min_{w \in \overline{\partial f(x^*)}} (w, g) \geq 0 \quad (7.6)$$

$$\forall g \in \Gamma(x^*).$$

Конус $\Gamma(x^*)$ по лемме 6.1 есть

$$\begin{aligned} \Gamma(x^*) = \gamma_1(x^*) &= \left\{ g \in E_n \mid \frac{\partial h(x^*)}{\partial g} = \right. \\ &= \left. \max_{v' \in \underline{\partial h(x^*)}} (v', g) + \min_{w' \in \overline{\partial h(x^*)}} (w', g) \leq 0 \right\}. \end{aligned} \quad (7.7)$$

Из (7.6) и (7.7)

$$\min_{w \in \underline{\partial f}(x^*)} \max_{v \in \underline{\partial f}(x^*) + w} (v, g) \geq 0 \quad \forall g \in \Gamma(x^*), \quad (7.8)$$

$$\Gamma(x^*) = \left\{ g \in E_n \mid \min_{w' \in \underline{\partial h}(x^*)} \max_{v' \in \underline{\partial h}(x^*) + w'} (v', g) \leq 0 \right\}. \quad (7.9)$$

Из (7.9)

$$\Gamma(x^*) = \bigcup_{w' \in \underline{\partial h}(x^*)} \Gamma_{w'}, \quad (7.10)$$

где $\Gamma_{w'} = \left\{ g \in E_n \mid \max_{v' \in \underline{\partial h}(x^*) + w'} (v', g) \leq 0 \right\}$. Из (7.8) и (7.10) получаем для любых $w \in \underline{\partial f}(x^*)$ и $w' \in \underline{\partial h}(x^*)$

$$\max_{v \in \underline{\partial f}(x^*) + w} (v, g) \geq 0 \quad \forall g \in \Gamma_{w'}. \quad (7.11)$$

Положим

$$F_w(g) = \max_{v \in \underline{\partial f}(x^*) + w} (v, g), \quad H_{w'}(g) = \max_{v' \in \underline{\partial h}(x^*) + w'} (v', g).$$

Тогда $\Gamma_{w'} = \{g \in E_n \mid H_{w'}(g) \leq 0\}$. Функции $F_w(g)$ и $H_{w'}(g)$ выпуклые. Найдем

$$\psi_{w, w'} = \min_{g \in \Gamma_{w'}} F_w(g). \quad (7.12)$$

Из (7.11) следует, что $\psi_{w, w'} = 0$, т. е. $g^* = 0$ — решение задачи (7.12). По теореме 6.1 гл. I должно быть

$$\partial F_w(g^*) \cap \mathcal{F}_{w'}^+(g^*) \neq \emptyset, \quad (7.13)$$

где $\mathcal{F}_{w'}^+(g^*)$ — конус, сопряженный конусу возможных направлений множества $\Gamma_{w'}$. Функция $F_w(g)$ — функция максимума, поэтому по следствию 2 из теоремы 14.2 гл. I

$$\partial F_w(0) = \underline{\partial f}(x^*) + w. \quad (7.14)$$

Поскольку конус возможных направлений множества $\Gamma_{w'}$ в точке $g^* = 0$ имеет вид

$$\mathcal{F}_{w'} = \{q \in E_n \mid (q, v') \leq 0 \quad \forall v' \in \overline{\partial h}(x^*) + w'\},$$

то по лемме 3.9 гл. I

$$\mathcal{F}_{w'}^+(0) = -\bar{\mathcal{H}}(\text{co}(\underline{\partial h}(x^*) + w')) = -\bar{\mathcal{H}}(\underline{\partial h}(x^*) + w'). \quad (7.15)$$

Из (7.13), (7.14) и (7.15) получаем

$$\begin{aligned} &(\underline{\partial f(x^*)} + w) \cap [-\bar{\mathcal{K}}(\underline{\partial h(x^*)} + w')] \neq \emptyset \\ &\forall w \in \overline{\partial f(x^*)}, \quad \forall w' \in \overline{\partial h(x^*)}, \end{aligned}$$

а это и есть (7.5). Теорема доказана.

Теорема 7.2. Условие (7.5) эквивалентно условию

$$-\overline{\partial f(x^*)} \subset \bigcap_{w \in \overline{\partial h(x^*)}} [\underline{\partial f(x^*)} + \bar{\mathcal{K}}(\underline{\partial h(x^*)} + w)]. \quad (7.16)$$

Доказательство. Пусть выполнено (7.5). Возьмем любые $w \in \overline{\partial f(x^*)}$ и $w' \in \overline{\partial h(x^*)}$. Из (7.5) следует существование вектора $v \in \underline{\partial f(x^*)}$ такого, что

$$v + w \in [-\bar{\mathcal{K}}(\underline{\partial h(x^*)} + w')].$$

Отсюда

$$-w \in [v + \bar{\mathcal{K}}(\underline{\partial h(x^*)} + w')] \subset \underline{\partial f(x^*)} + \bar{\mathcal{K}}(\underline{\partial h(x^*)} + w').$$

В силу произвольности $w' \in \overline{\partial h(x^*)}$ имеем

$$-w \in \bigcap_{w' \in \overline{\partial h(x^*)}} [\underline{\partial f(x^*)} + \bar{\mathcal{K}}(\underline{\partial h(x^*)} + w')].$$

Так как вектор $w \in \overline{\partial f(x^*)}$ произволен, то отсюда получаем включение (7.16). Пусть теперь имеет место (7.16). Тогда для любых $w \in \overline{\partial f(x^*)}$ и $w' \in \overline{\partial h(x^*)}$ будет

$$-w \in [\underline{\partial f(x^*)} + \bar{\mathcal{K}}(\underline{\partial h(x^*)} + w')].$$

Отсюда

$$0 \in [\underline{\partial f(x^*)} + w + \bar{\mathcal{K}}(\underline{\partial h(x^*)} + w')],$$

т. е.

$$[\underline{\partial f(x^*)} + w] \cap [-\bar{\mathcal{K}}(\underline{\partial h(x^*)} + w')] \neq \emptyset,$$

а это и есть (7.5). Теорема доказана.

Следствие. Пусть $h(x^*) = 0$, в точке x^* функция $h(x)$ субдифференцируема и выполнено условие невырожденности. Тогда

$$-\overline{\partial f(x^*)} \subset \underline{\partial f(x^*)} - \Gamma^+(x^*).$$

Это вытекает из (7.16) и того, что

$$\overline{\partial h(x^*)} = \{0\}, \quad \mathcal{K}(\overline{\partial h(x^*)}) = -\Gamma^+(x^*),$$

где $\Gamma(x^*)$ — конус возможных направлений множества Ω в точке x^* .

Аналогично устанавливаются необходимые условия максимума.

Теорема 7.3. Пусть функция $f(x)$ липшицева в окрестности точки $x^{**} \in \Omega$. Если $h(x^{**}) = 0$, то предположим, что функция $h(x)$ тоже липшицева в окрестности точки x^{**} и в точке x^{**} выполнено условие невырожденности (6.3). Для того чтобы функция $f(x)$ достигала в точке x^{**} своего наибольшего на Ω значения, необходимо, чтобы

$$-\overline{\partial f(x^{**})} \subset \overline{\partial f(x^{**})}, \quad (7.17)$$

$$\text{если } h(x^{**}) < 0,$$

$$-\overline{\partial f(x^{**})} \subset \bigcap_{w \in \overline{\partial h(x^{**})}} [\overline{\partial f(x^{**})} - \mathcal{K}(\overline{\partial h(x^{**})} + w)], \quad (7.18)$$

$$\text{если } h(x^{**}) = 0.$$

Задание 7.1. Доказать теорему 7.3 и показать, что если $h(x)$ — выпуклая функция, то из (7.18) вытекает условие

$$-\overline{\partial f(x^{**})} \subset \overline{\partial f(x^{**})} + \Gamma^+(x^{**}).$$

2. Пусть $x_0 \in \Omega$. Предположим, что функция $f(x)$ липшицева в окрестности точки x_0 . Если $h(x_0) = 0$, то будем также предполагать, что функция $h(x)$ липшицева в окрестности точки x_0 и в точке x_0 выполнено условие невырожденности (6.3).

Определение. Точку $x_0 \in \Omega$, для которой выполнено соотношение (7.4), если $h(x_0) < 0$, и соотношение (7.5), если $h(x_0) = 0$, назовем *inf-стационарной точкой* функции $f(x)$ на множестве Ω . Точку $x_0 \in \Omega$, для которой выполнено условие (7.17), если $h(x_0) < 0$, и (7.18), если $h(x_0) = 0$, назовем *sup-стационарной точкой* функции $f(x)$ на множестве Ω .

Рассмотрим подробнее случай, когда $h(x_0) = 0$. Пусть точка x_0 не является inf-стационарной, т. е. не выпол-

нено условие (7.16) (или, что то же, (7.5)). Найдем

$$\begin{aligned} \rho(x_0) &= \max_{\substack{w \in \underline{\partial f}(x_0) \\ w' \in \underline{\partial h}(x_0)}} \min_{\substack{z \in [\underline{\partial f}(x_0) + w] \\ z' \in \mathcal{H}(\underline{\partial h}(x_0) + w')}} \|z + z'\| = \\ &= \max_{\substack{w \in \underline{\partial f}(x_0) \\ w' \in \underline{\partial h}(x_0)}} d(w, w') = d(w_0, w'_0), \quad (7.19) \end{aligned}$$

где

$$d(w, w') = \min_{\substack{z \in [\underline{\partial f}(x_0) + w] \\ z' \in \mathcal{H}(\underline{\partial h}(x_0) + w')}} \|z + z'\|.$$

Поскольку (7.5) не имеет места, то $\rho(x_0) > 0$.

Пусть $d(w_0, w'_0) = \|z_0 + z'_0\|$ (при этом $\|z_0 + z'_0\| = \rho(x_0)$). Положим

$$g_0 = -(z_0 + z'_0) \|z_0 + z'_0\|^{-1}.$$

Лемма 7.2. Если $h(x_0) = 0$ и в точке x_0 выполнено условие невырожденности (6.3), то направление g_0 есть направление наискорейшего спуска функции $f(x)$ в точке x_0 по множеству Ω , а величина $\rho(x_0) = \|z_0 + z'_0\|$ есть скорость наискорейшего спуска, т. е.

$$\frac{\partial f(x_0)}{\partial g_0} = \min_{\substack{g \in \Gamma(x_0) \\ \|g\|=1}} \frac{\partial f(x_0)}{\partial g} = -\rho(x_0). \quad (7.20)$$

Доказательство. Как и при доказательстве леммы 6.4 гл. I, устанавливается, что

$$g_0 \in [-\mathcal{H}^+(\underline{\partial h}(x_0) + w'_0)],$$

а тогда (см. (6.4) и (6.11)) $g_0 \in \Gamma(x_0)$. Осталось доказать (7.20). Допустим, что (7.20) не выполнено. Тогда существует $g_1 \in \Gamma(x_0)$, $\|g_1\| = 1$, такое, что

$$\frac{\partial f(x_0)}{\partial g_1} < -\rho(x_0). \quad (7.21)$$

Так как

$$\frac{\partial f(x_0)}{\partial g_1} = \min_{w \in \underline{\partial f}(x_0)} \max_{v \in [\underline{\partial f}(x_0) + w]} (v, g_1),$$

то найдется вектор $w_1 \in \overline{\partial f(x_0)}$ такой, что

$$\frac{\partial f(x_0)}{\partial g_1} = \max_{v \in [\partial f(x_0) + w_1]} (v, g_1). \quad (7.22)$$

Так как $g_1 \in \Gamma(x_0)$, то (см. (6.4) и (6.11)) существует такое $w'_1 \in \overline{\partial h(x_0)}$, что

$$g_1 \in [-\mathcal{K}^+(\partial h(x_0) + w'_1)].$$

Найдем

$$d(w_1, w'_1) = \min_{z \in [\partial f(x_0) + w_1]} \|z + z'\| = \|z_1 + z'_1\|,$$

$$z' \in \mathcal{K}(\partial h(x_0) + w'_1)$$

Как и при доказательстве леммы 6.4 гл. I, показывается, что

$$-\|z_1 + z'_1\| = \min_{g \in [-\mathcal{K}^+(\partial h(x_0) + w'_1)]} \max_{v \in [\partial f(x_0) + w_1]} (v, g) \leq$$

$$\leq \max_{v \in [\partial f(x_0) + w_1]} (v, g_1) = \frac{\partial f(x_0)}{\partial g_1}.$$

Отсюда и из (7.21) имеем

$$d(w_1, w'_1) \geq -\frac{\partial f(x_0)}{\partial g_1} > \rho(x_0),$$

что противоречит (7.19). Лемма доказана.

З а м е ч а н и е. Направление наискорейшего спуска может быть и не единственным.

З а д а н и е 7.2. Доказать результат, аналогичный лемме 7.2, для направлений наискорейшего подъема.

П р и м е р 1. Пусть $x = (x^{(1)}, x^{(2)}) \in E_2$, $x_0 = (0, 0)$, $f(x) = |x^{(1)}| - |x^{(2)}| + x^{(2)}$, $h(x) = -0,5|x^{(1)}| - x^{(2)}$. Тогда

$$f(x) = f_1(x) + f_2(x) + f_3(x), \quad h(x) = h_1(x) + h_2(x),$$

где $f_1(x) = |x^{(1)}|$, $f_2(x) = -|x^{(2)}|$, $f_3(x) = x^{(2)}$,

$$h_1(x) = -0,5|x^{(1)}|, \quad h_2(x) = -x^{(2)}.$$

Образует множество (рис. 38)

$$\Omega = \{x \in E_2 \mid h(x) \leq 0\}.$$

Поскольку (см. примеры 1—4 § 3)

$$\underline{\partial f_1(x)} = \text{co} \{A_1; A_2\}, \quad \overline{\partial f_1(x_0)} = \{\emptyset\},$$

$$\underline{\partial f_2(x_0)} = \{\emptyset\}, \quad \overline{\partial f_2(x_0)} = \text{co} \{B_1; B_2\},$$

$$\underline{\partial f_3(x_0)} = \{B_1\}, \quad \overline{\partial f_3(x_0)} = \{\emptyset\},$$

$$\underline{\partial h_1(x_0)} = \{\emptyset\}, \quad \overline{\partial h_1(x_0)} = \text{co} \{0,5A_1; 0,5A_1\},$$

$$\underline{\partial h_2(x_0)} = \{B_2\}, \quad \overline{\partial h_2(x_0)} = \{\emptyset\},$$

где $A_1 = (1, 0)$, $A_2 = (-1, 0)$, $B_1 = (0, 1)$, $B_2 = (0, -1)$, то

$$\underline{\partial f(x_0)} = \text{co} \{A_1, A_2\} + \{B_1\}, \quad \overline{\partial f(x_0)} = \text{co} \{B_1, B_2\},$$

$$\underline{\partial h(x_0)} = \{B_2\}, \quad \overline{\partial h(x_0)} = \text{co} \{0,5A_1; 0,5A_2\}$$

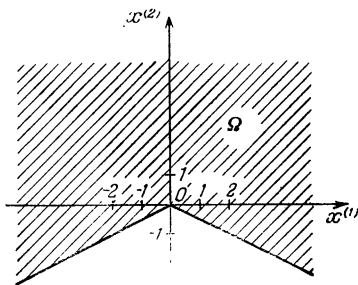


Рис. 38.

Прежде всего проверим, что в точке x_0 выполнено условие невырожденности. Имеем

$$\frac{\partial h(x_0)}{\partial g} = (B_2, g) + \min \{(0,5A_1; g), (0,5A_2; g)\}.$$

Найдем те $g \in E_2$, $\|g\| = 1$, для которых $\frac{\partial h(x_0)}{\partial g} = 0$. Очевидно, что

$$\frac{\partial h(x_0)}{\partial g_1} = \frac{\partial h(x_0)}{\partial g_2} = 0,$$

где

$$g_1 = \left(\frac{2}{\sqrt{5}}, \frac{-1}{\sqrt{5}} \right), \quad g_2 = \left(\frac{-2}{\sqrt{5}}, \frac{-1}{\sqrt{5}} \right).$$

Заметим, что

$$\gamma(x_0) = \beta_1 \cup \beta_2 \cup \beta_3,$$

где

$$\beta_1 = \{g \in E_2 \mid g = \lambda(-2, z), \lambda > 0, z \in (-1, 0]\},$$

$$\beta_2 = \{g \in E_2 \mid g = \lambda(y, 1), \lambda > 0, y \in [-1, 1]\},$$

$$\beta_3 = \{g \in E_2 \mid g = \lambda(2, z), \lambda > 0, z \in (-1, 0]\}.$$

Очевидно (рис. 39),

$$\gamma_1(x_0) = \gamma(x_0) \cup \{g = \lambda g_1 \mid \lambda \geq 0\} \cup \{g = \lambda g_2 \mid \lambda \geq 0\} = \overline{\gamma(x_0)},$$

т. е. в точке x_0 выполнено условие невырожденности (6.3).

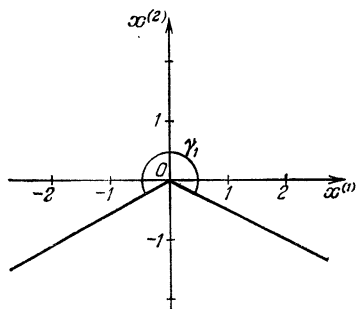


Рис. 39.

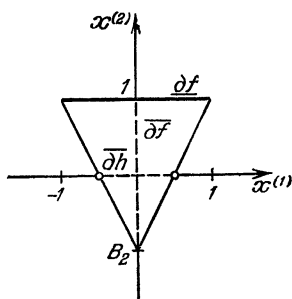


Рис. 40.

Из рис. 40 ясно, что

$$\begin{aligned} \mathcal{H} &\equiv \bigcap_{w \in \overline{\partial h(x_0)}} [\underline{\partial f(x_0)} + \overline{\mathcal{H}}(\underline{\partial h(x_0)} + w)] = \\ &= \text{co}\{(-1, 1), (1, 1), (0, -1)\}. \end{aligned}$$

Отсюда $-\overline{\partial f(x_0)} \subset \mathcal{H}$, т. е. условие (7.16) выполнено; поэтому точка x_0 является inf-стационарной точкой функции $f(x) = |x^{(1)}| - |x^{(2)}| + x^{(2)}$ на множестве $\Omega = \{x \mid -0,5|x^{(1)}| - x^{(2)} \leq 0\}$. В действительности точка x_0 является точкой минимума функции $f(x)$ на множестве Ω .

Пример 2. Пусть снова $x = (x^{(1)}, x^{(2)}) \in E_2$, $x_0 = (0, 0)$, функция $f(x)$ та же, что и в примере 1. Возь-

нем множество (рис. 41) $\Omega = \{x \in E_2 | h(x) \leq 0\}$, где $h(x) = -|x^{(1)}| - x^{(2)}$. Имеем (см. пример 1)

$$\underline{\partial f}(x_0) = \text{co}\{A_1, A_2\} + \{B_1\}, \quad \overline{\partial f}(x_0) = \text{co}\{B_1, B_2\}.$$

Легко видеть, что

$$\underline{\partial h}(x_0) = \{B_2\}, \quad \overline{\partial h}(x_0) = \text{co}\{A_1, A_2\}.$$

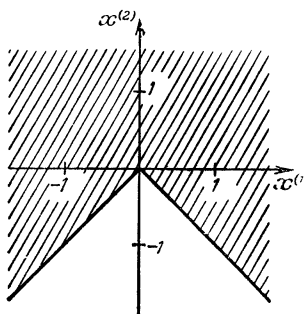


Рис. 41.

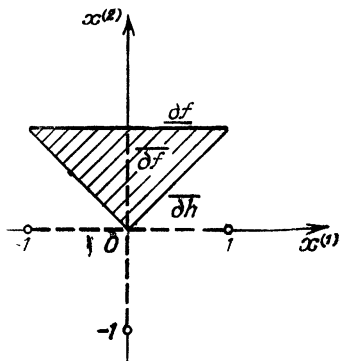


Рис. 42.

Как и в примере 1, проверяется условие невырожденности в точке x_0 . Множество $\mathcal{H}$ изображено на рис. 42:

$$\begin{aligned} \mathcal{H} &\equiv \bigcap_{w \in \overline{\partial h}(x_0)} [\underline{\partial f}(x_0) + \overline{\mathcal{H}}(\underline{\partial h}(x_0) + w)] = \\ &= \text{co}\{(-1, 1), (1, 1), (0, 0)\}. \end{aligned}$$

Очевидно, что

$$-\overline{\partial f}(x_0) = -\text{co}\{B_1, B_2\} = \text{co}\{B_1, B_2\} \not\subset \mathcal{H}.$$

Найдем направления наискорейшего спуска. Имеем для $w \in \overline{\partial f}(x_0)$, $w' \in \overline{\partial h}(x_0)$

$$\begin{aligned} \underline{\partial f}(x_0) + w &= \text{co}\{A_1 + B_1 + w, A_2 + B_1 + w\} \equiv \mathfrak{A}_w, \\ -\overline{\mathcal{H}}(\underline{\partial h}(x_0) + w') &= \\ &= \{z = \lambda(-B_2 - w') = \lambda(B_1 - w') | \lambda \geq 0\} \equiv \mathfrak{B}_{w'}. \end{aligned}$$

Множества $\mathfrak{A}_w$ есть отрезки, параллельные отрезку $\text{co}\{A_1, A_2\}$ и заполняющие прямоугольник с вершинами

$(-1, 0)$, $(-1, 2)$, $(1, 2)$, $(1, 0)$. Множества $\mathfrak{B}_{w'}$ есть лучи с центром в начале координат, проходящие через точки отрезка, соединяющего точки $(-1, 1)$ и $(1, 1)$. Найдем $\rho(x_0)$ (см. (7.19)):

$$\rho(x_0) = \max_{\substack{w \in \frac{\partial f(x_0)}{\partial h(x_0)} \\ w' \in \frac{\partial h(x_0)}{\partial h(x_0)}}} d(w, w') = d(w_0, w'_0) = d(w_1, w'_1),$$

где

$$w_0 = (0, 1), \quad w'_0 = A_1, \quad w_1 = w_0 = (0, 1), \quad w'_1 = A_2.$$

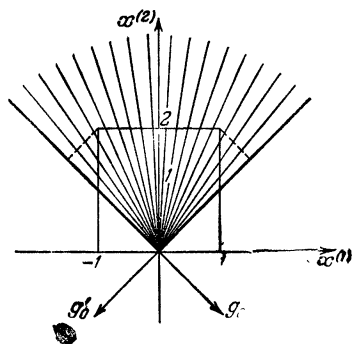


Рис. 43.

Имеем два направления наискорейшего спуска (рис. 43):

$$g_0 = \left(\frac{\sqrt{2}}{2}, -\frac{\sqrt{2}}{2} \right), \quad g'_0 = \left(-\frac{\sqrt{2}}{2}, -\frac{\sqrt{2}}{2} \right).$$

При этом

$$\frac{\partial f(x_0)}{\partial g_0} = \max_{v \in \frac{\partial f(x_0)}{\partial h(x_0)}} (v, g_0) + \min_{w \in \frac{\partial f(x_0)}{\partial h(x_0)}} (w, g_0) = -\frac{\sqrt{2}}{2}.$$

$$\begin{aligned} \text{Аналогично } \frac{\partial f(x_0)}{\partial g'_0} &= -\frac{\sqrt{2}}{2}. \text{ Очевидно, что } \frac{\partial f(x_0)}{\partial g_0} = \\ &= -\rho(x_0) = -\frac{\sqrt{2}}{2}. \end{aligned}$$

3. Установим достаточные условия локального минимума функции $f(x)$ на множестве Ω .

Пусть $x_0 \in \Omega$ — inf-стационарная точка функции $f(x)$ на множестве Ω . Теорема 5.3 устанавливает достаточ-

ные условия минимума, если $h(x_0) < 0$. Рассмотрим случай, когда $h(x_0) = 0$. Предполагаем, что в точке x_0 выполнено условие невырожденности (6.3) и функции $f(x)$ и $h(x)$ липшицевы в окрестности точки x_0 . По теореме 7.1 для любых $w \in \underline{\partial} f(x_0)$ и $w' \in \underline{\partial} h(x_0)$ имеет место (7.5), т. е.

$$\emptyset \in [\underline{\partial} f(x_0) + w + \overline{\mathcal{K}}(\underline{\partial} h(x_0) + w')] = \mathfrak{L}(w, w'). \quad (7.23)$$

Через $r(w, w')$ обозначим радиус максимальной сферы с центром в начале координат, вписанной в $\mathfrak{L}(w, w')$. В силу (7.23) $r(w, w') \geq 0$. Положим

$$r = \min_{w \in \underline{\partial} f(x_0), w' \in \underline{\partial} h(x_0)} r(w, w').$$

Напомним, что

$$f(x_0 + \alpha g) = f(x_0) + \alpha \frac{\partial f(x_0)}{\partial g} + o(\alpha, g). \quad (7.24)$$

Теорема 7.4. *Предположим, что в (7.24) $\frac{o(\alpha, g)}{\alpha} \rightarrow 0$ равномерно по $g \in E_n$, $\|g\| = 1$. Если $r > 0$, то точка x_0 является точкой локального минимума, причем*

$$\min_{g \in \Gamma(x_0), \|g\|=1} \frac{\partial f(x_0)}{\partial g} = r.$$

Доказательство очевидно.

Пример 3. Пусть

$$x = (x^{(1)}, x^{(2)}) \in E_2, \quad x_0 = (0, 0),$$

$$f(x) = |x^{(1)}| - 0,5|x^{(2)}| + x^{(2)},$$

и пусть множество Ω задано так:

$$\Omega = \{x \in E_2 \mid h(x) \leq 0\},$$

где $h(x) = -0,25|x^{(1)}| - x^{(2)}$.

Как и в примерах 1 и 2, получаем (рис. 44), что

$$\underline{\partial} f(x_0) = \text{co} \{A_1, A_2\} + \{B_1\} = \text{co} \{(1; 1), (-1; 1)\},$$

$$\underline{\partial} f(x_0) = \text{co} \{0,5B_1; 0,5B_2\} = \text{co} \{(0; 0,5), (0; -0,5)\},$$

$$\underline{\partial} h(x_0) = \{B_2\},$$

$$\overline{\partial} h(x_0) = \text{co} \{0,25A_1; 0,25A_2\} = \text{co} \{(0,25; 0), (-0,25; 0)\}.$$

Возьмем $w \in \overline{\partial f(x_0)}$ и $w' \in \overline{\partial h(x_0)}$. Тогда

$$\overline{\partial f(x_0)} + w = \text{co} \{A_1 + B_1 + w, A_2 + B_1 + w\} = \mathfrak{A}_w,$$

$$\mathcal{K}(\overline{\partial h(x_0)} + w') = \{z = \lambda(B_2 + w') \mid \lambda \geq 0\} = \mathfrak{B}_{w'}.$$

Множество $\mathfrak{L}(w, w')$ (см. (7.23)) изображено на рис. 45

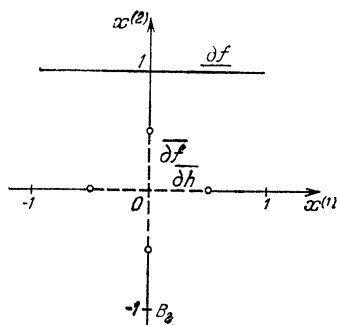


Рис. 44.

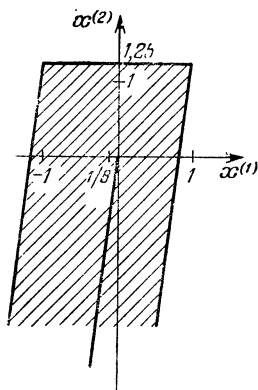


Рис. 45.

для $w = (0; 0,25)$, $w' = (-0,125, 0)$. Лучи l_1 , l_2 , l_3 параллельны. Очевидно, что

$$0 \in \mathfrak{L}(w, w') \quad \forall w \in \overline{\partial f(x_0)}, \quad \forall w' \in \overline{\partial h(x_0)}.$$

Нетрудно заметить, что

$$r = \min_{w \in \overline{\partial f(x_0)}, w' \in \overline{\partial h(x_0)}} r(w, w').$$

Минимум здесь достигается при $w_0 = (0; 0,5)$, $w'_0 = (0,25; 0)$ и при $w_1 = (0; 0,5)$, $w'_1 = (-0,25; 0)$. Вычисляя $r = r(w_0, w'_0)$, найдем $r = 34^{-1} \sqrt{425} \approx 0,6063$.

§ 8*. Функция расстояния до множества

В настоящем параграфе изучается дифференцируемость по направлениям функции, являющейся евклидовым расстоянием от точки до замкнутого множества $\Omega \subset E_n$. В п. 1 рассматривается случай, когда Ω — выпуклое множество, в п. 2 — случай, когда Ω задано с по-

мощью квазидифференцируемой функции. Такие функции интересны сами по себе и естественным образом возникают при решении практических задач.

1. Пусть $\Omega \subset E_n$ — выпуклое замкнутое множество. Образует функцию

$$f(x) = \min_{y \in \Omega} \|x - y\| = \|x - y(x)\|.$$

Очевидно, что если $x \in \Omega$, то $f(x) = 0$. Заметим прежде всего, что $f(x)$ — выпуклая на E_n функция. Действительно, возьмем точки $x_1 \in E_n$ и $x_2 \in E_n$. Пусть

$$f(x_1) = \|x_1 - y_1\|, \quad f(x_2) = \|x_2 - y_2\|,$$

где $y_1 \in \Omega$, $y_2 \in \Omega$. При $\alpha \in [0, 1]$ в силу выпуклости Ω будет

$$y_\alpha = \alpha y_1 + (1 - \alpha) y_2 \in \Omega.$$

Поэтому

$$\begin{aligned} f(\alpha x_1 + (1 - \alpha) x_2) &\leq \|\alpha x_1 + (1 - \alpha) x_2 - y_\alpha\| = \\ &= \|\alpha(x_1 - y_1) + (1 - \alpha)(x_2 - y_2)\| \leq \\ &\leq \alpha \|x_1 - y_1\| + (1 - \alpha) \|x_2 - y_2\| = \alpha f(x_1) + (1 - \alpha) f(x_2), \end{aligned}$$

а это и означает, что $f(x)$ — выпуклая функция.

Наша цель — найти субдифференциал этой функции.

Пусть $x_0 \in E_n$. Если $x_0 \in \text{int } \Omega$, то

$$\frac{\partial f(x_0)}{\partial g} = \lim_{\alpha \rightarrow +0} \alpha^{-1} [f(x_0 + \alpha g) - f(x_0)] = 0$$

$$\forall g \in E_n.$$

Следовательно,

$$\partial f(x_0) = \{0\}. \quad (8.1)$$

Если $x_0 \notin \Omega$, то

$$f(x) = \min_{y \in \Omega} \|y - x\| = \min_{y \in \Omega} \sqrt{(y - x)^2} = \min_{y \in \Omega} \varphi(x, y).$$

Функция $\varphi(x, y) = \sqrt{(y - x)^2}$ непрерывно дифференцируема по x в окрестности точки x_0 . Так как

$$\min_{y \in \Omega} \varphi(x, y) = - \max_{y \in \Omega} (-\varphi(x, y)),$$

то выполнены все условия теоремы 13.1 гл. I и поэтому

$$\frac{\partial f(x_0)}{\partial g} = \min_{y \in Q(x_0)} (\varphi'_x(x_0, y), g),$$

где $Q(x_0) = \{y \in \Omega \mid \varphi(x_0, y) = f(x_0)\}$. Но в данном случае множество $Q(x_0)$ состоит из единственной точки $y(x_0)$. Поэтому

$$\frac{\partial f(x_0)}{\partial g} = (v, g), \quad (8.2)$$

где

$$v = \frac{1}{\sqrt{(y(x_0) - x_0)^2}} (x_0 - y(x_0)) = \frac{x_0 - y(x_0)}{f(x_0)}, \quad (8.3)$$

т. е. функция $f(x)$ непрерывно дифференцируема в окрестности точки x_0 .

Осталось рассмотреть случай, когда x_0 — граничная точка множества Ω . Имеем

$$\frac{\partial f(x_0)}{\partial g} = \lim_{\alpha \rightarrow +0} H(\alpha),$$

где $H(\alpha) = \alpha^{-1}[f(x_0 + \alpha g) - f(x_0)]$. Так как $x_0 \in \Omega$, то $f(x_0) = 0$, т. е.

$$H(\alpha) = \alpha^{-1} f(x_0 + \alpha g) = \alpha^{-1} \min_{y \in \Omega} \|x_0 + \alpha g - y\|.$$

Для любого $y \in \Omega$

$$\|x_0 + \alpha g - y\| = \|x_0 + \alpha g - (x_0 + (y - x_0))\| = \alpha \|g - \alpha^{-1}(y - x_0)\|.$$

Поскольку $\alpha > 0$, то $v(y, \alpha) = \alpha^{-1}(y - x_0) \in \gamma(x_0) = \{v = \lambda(x - x_0) \mid x \in \Omega\}$. Поэтому

$$\|x_0 + \alpha g - y\| = \alpha \|g - v(y, \alpha)\| \geq \alpha \min_{v \in \Gamma(x_0)} \|g - v\| = \alpha P,$$

где $P = \min_{v \in \Gamma(x_0)} \|g - v\|$, $\Gamma(x_0) = \bar{\gamma}(x_0)$ — конус возможных направлений множества Ω в точке x_0 . Так как P не зависит от y , то $\min_{y \in \Omega} \|x_0 + \alpha g - y\| \geq \alpha P$, т. е.

$$H(\alpha) \geq P. \quad (8.4)$$

Возьмем теперь любое $v \in \gamma(x_0)$. При достаточно малых $\alpha > 0$ будет $x_0 + \alpha v \in \Omega$. Поэтому

$$f(x_0 + \alpha g) = \min_{y \in \Omega} \|x_0 + \alpha g - y\| \leq$$

$$\leq \|x_0 + \alpha g - x_0 - \alpha v\| = \alpha \|g - v\|.$$

Отсюда

$$H(\alpha) \leq \|g - v\|.$$

Поскольку это неравенство справедливо для любого $v \in \gamma(x_0)$, то

$$H(\alpha) \leq \inf_{v \in \gamma(x_0)} \|g - v\| = \min_{v \in \Gamma(x_0)} \|g - v\| = P. \quad (8.5)$$

Так как P не зависит от α , то из (8.4) и (8.5) имеем

$$\frac{\partial f(x_0)}{\partial g} = P = \min_{v \in \Gamma(x_0)} \|g - v\|.$$

По следствию 3 из леммы 6.4 гл. I имеем (поменяв ролями Γ и Γ^+)

$$\min_{v \in \Gamma(x_0)} \|g - v\| = \max_{\substack{w \in [-\Gamma^+(x_0)] \\ \|w\| \leq 1}} (w, g),$$

т. е.

$$\frac{\partial f(x_0)}{\partial g} = \max_{\substack{w \in [-\Gamma^+(x_0)] \\ \|w\| \leq 1}} (w, g). \quad (8.6)$$

Значит,

$$\begin{aligned} \partial f(x_0) &= \{w \in E_n \mid \|w\| \leq 1, w \in [-\Gamma^+(x_0)]\} = \\ &= (-\Gamma^+(x_0)) \cap S_1, \end{aligned}$$

где

$$S_1 = \{w \in E_n \mid \|w\| \leq 1\}.$$

Так как в случае $h(x_0) < 0$ будет $\Gamma^+(x_0) = \{\emptyset\}$, то из (8.1), (8.2) и (8.6) окончательно получим, что справедлива

Лемма 8.1. *Функция $f(x)$ дифференцируема по направлениям на E_n , причем*

$$\frac{\partial f(x_0)}{\partial g} = \max_{v \in \partial f(x_0)} (v, g), \quad (8.7)$$

где

$$\partial f(x_0) = \begin{cases} \left\{ \frac{y(x_0) - x_0}{f(x_0)} \right\}, & \text{если } x_0 \notin \Omega, \\ (-\Gamma^+(x_0) \cap S_1), & \text{если } x_0 \in \Omega. \end{cases} \quad (8.8)$$

З а д а н и е 8.1. Найти субдифференциал функции

$$f(x) = \min_{y \in \Omega} \|x - y\|_m,$$

где $\|x - y\|_m = \max_{i \in 1:n} \{|x^{(i)} - y^{(i)}|\}$, $x = (x^{(1)}, \dots, x^{(n)})$, $y = (y^{(1)}, \dots, y^{(n)})$, $\Omega \subset E_n$ — выпуклое замкнутое множество.

2. Пусть $\Omega \subset E_n$ — замкнутое множество (уже не обязательно выпуклое). Рассмотрим снова функцию

$$f(x) = \min_{y \in \Omega} \|x - y\|. \quad (8.9)$$

Положим

$$Q(x_0) = \{y \in \Omega \mid \|x - y\| = f(x)\}. \quad (8.10)$$

Ясно, что множество $Q(x)$ может состоять более чем из одной точки. Функция $f(x)$ уже не обязательно выпуклая.

Пусть $x_0 \in E_n$. Если $x_0 \in \text{int } \Omega$, то

$$\frac{\partial f(x_0)}{\partial g} = 0 \quad \forall g \in E_n. \quad (8.11)$$

Если $x_0 \notin \Omega$, то $f(x_0) = \min_{y \in \Omega} \|y - x\|$. Как и в п. 1,

$$\frac{\partial f(x_0)}{\partial g} = \min_{y \in Q(x_0)} (\varphi'_x(x_0, y), g),$$

где $\varphi(x, y) = \sqrt{(y - x)^2}$. Значит,

$$\frac{\partial f(x_0)}{\partial g} = \min_{v \in A} (v, g), \quad (8.12)$$

где

$$A = \text{co} \left\{ \frac{x_0 - y}{f(x_0)} \mid y \in Q(x_0) \right\}.$$

Пусть теперь x_0 — граничная точка множества Ω . Предположим, что множество Ω имеет вид

$$\Omega = \{x \in E_n \mid h(x) \leq 0\},$$

где $h(x)$ — квазидифференцируемая на E_n функция. Предположим также, что в окрестности точки x_0 функция $h(x)$ липшицева и в точке x_0 выполнено условие не-

вырожденности (6.3). Поскольку x_0 — граничная точка множества Ω , то $h(x_0) = 0$. По лемме 6.1

$$\Gamma(x_0) = \gamma_1(x_0) = \bar{\gamma}(x_0),$$

где

$$\gamma(x_0) = \left\{ g \in E_n \mid \frac{\partial h(x_0)}{\partial g} < 0 \right\}. \quad (8.13)$$

Положим

$$\begin{aligned} H(\alpha) &= \alpha^{-1} [f(x_0 + \alpha g) - f(x_0)] = \\ &= \alpha^{-1} f(x_0 + \alpha g) = \alpha^{-1} \min_{y \in \Omega} \|x_0 + \alpha g - y\|. \end{aligned}$$

Пусть последовательность $\{\alpha_k\}$ такова, что

$$\alpha_k \rightarrow +0, \quad H(\alpha_k) \rightarrow \lim_{\alpha \rightarrow +0} H(\alpha) \quad (8.14)$$

(такая последовательность существует), и пусть

$$\begin{aligned} f(x_0 + \alpha_k g) &\equiv \min_{y \in \Omega} \|x_0 + \alpha_k g - y\| = \|x_0 + \alpha_k g - y_k\| = \\ &= \alpha_k \left\| g - \frac{y_k - x_0}{\alpha_k} \right\| = \alpha_k \left\| g - \frac{y_k - x_0}{\|y_k - x_0\|} \frac{\|y_k - x_0\|}{\alpha_k} \right\| = \\ &= \alpha_k \|g - \lambda_k z_k\|, \quad (8.15) \end{aligned}$$

где

$$z_k = \frac{y_k - x_0}{\|y_k - x_0\|}, \quad \lambda_k = \frac{\|y_k - x_0\|}{\alpha_k}.$$

Поскольку

$$f(x_0 + \alpha_k g) \geq 0, \quad \|z_k\| = 1,$$

то в (8.15) λ_k ограничено. Можно считать, что $\lambda_k \rightarrow \lambda \geq 0$, $z_k \rightarrow z_0$. Из (8.14) тогда

$$\lim_{\alpha \rightarrow +0} H(\alpha) = \lim_{k \rightarrow \infty} H(\alpha_k) = \|g - \lambda z_0\|.$$

Так как здесь $z_0 \in \Gamma(x_0)$, $z_1 = \lambda z_0 \in \Gamma(x_0)$, то

$$\lim_{\alpha \rightarrow +0} H(\alpha) = \|g - z_1\| \geq \min_{z \in \Gamma(x_0)} \|g - z\| \equiv P \quad (8.16)$$

(выше неявно предполагалось, что $y_k \neq x_0$, в случае $y_k = x_0$ соотношение (8.16) очевидно).

Возьмем теперь последовательность $\{\alpha_k\}$ такую, что

$$\alpha_k \rightarrow +0, \quad H(\alpha_k) \rightarrow \overline{\lim}_{\alpha \rightarrow +0} H(\alpha). \quad (8.17)$$

Зафиксируем любое $z \in \gamma(x_0)$. Тогда из (8.13) следует, что при достаточно малых $\alpha > 0$ будет $x_\alpha = x_0 + \alpha z \in \Omega$; поэтому

$$\begin{aligned} f(x_0 + \alpha_k g) &= \min_{y \in \Omega} \|x_0 + \alpha_k g - y\| \leq \\ &\leq \|x_0 + \alpha_k g - x_{\alpha_k}\| = \alpha_k \|g - z\|. \end{aligned}$$

Отсюда и из (8.17)

$$\overline{\lim}_{\alpha \rightarrow +0} H(\alpha) = \lim_{k \rightarrow \infty} H(\alpha_k) \leq \|g - z\|.$$

Это неравенство справедливо для всех $z \in \gamma(x_0)$. Поэтому

$$\overline{\lim}_{\alpha \rightarrow +0} H(\alpha) \leq \inf_{z \in \gamma(x_0)} \|g - z\|.$$

Так как $\Gamma(x_0) = \bar{\gamma}(x_0)$, то отсюда

$$\overline{\lim}_{\alpha \rightarrow +0} H(\alpha) \leq \min_{z \in \Gamma(x_0)} \|g - z\| = P. \quad (8.18)$$

Из (8.16) и (8.18) получаем, что существует $\lim_{\alpha \rightarrow +0} H(\alpha)$, причем

$$\frac{\partial f(x_0)}{\partial g} = \lim_{\alpha \rightarrow +0} H(\alpha) = P = \min_{z \in \Gamma(x_0)} \|g - z\|, \quad (8.19)$$

т.е. функция $f(x)$ дифференцируема в точке x_0 по направлениям. Так как $\Gamma(x_0) = \gamma_1(x_0)$, то

$$\begin{aligned} \Gamma(x_0) &= \left\{ z \in E_n \mid \max_{v \in \underline{\partial h}(x_0)} (v, z) + \min_{w \in \underline{\partial h}(x_0)} (w, z) \leq 0 \right\} = \\ &= \left\{ z \in E_n \mid \min_{w \in \underline{\partial h}(x_0)} \left[\max_{v \in [\underline{\partial h}(x_0) + w]} (v, z) \right] \leq 0 \right\} = \\ &= \bigcup_{w \in \underline{\partial h}(x_0)} \left\{ \max_{v \in [\underline{\partial h}(x_0) + w]} (v, z) \leq 0 \right\} = \\ &= \bigcup_{w \in \underline{\partial h}(x_0)} [-\mathcal{K}^+(\underline{\partial h}(x_0) + w)], \end{aligned}$$

где $\mathcal{K}(A)$ — коническая оболочка множества A , $\mathcal{K}^+$ — конус, сопряженный конусу $\mathcal{K}$. Из (8.19)

$$\frac{\partial f(x_0)}{\partial g} = \min_{w \in \partial h(x)} \min_{z \in [-\mathcal{K}^+(\partial h(x_0) + w)]} \|g - z\|.$$

По следствию 3 из леммы 6.4 гл. I

$$\min_{z \in [-\mathcal{K}^+(\partial h(x_0) + w)]} \|g - z\| = \max_{\substack{v \in \mathcal{K}(\partial h(x_0) + w) \\ \|v\| \leq 1}} (v, g).$$

Окончательно имеем

$$\frac{\partial f(x_0)}{\partial g} = \min_{w \in \partial h(x)} \max_{\substack{v \in \mathcal{K}(\partial h(x_0) + w) \\ \|v\| \leq 1}} (v, g). \quad (8.20)$$

Из (8.11), (8.12) и (8.20) вытекает

Лемма 8.2. *Функция $f(x)$, заданная соотношением (8.9), дифференцируема по направлениям для всех $x_0 \in \text{int } \Omega$, для всех $x_0 \notin \Omega$ и в тех граничных точках множества Ω , в окрестности которых функция $h(x)$ липшицева и в которых выполнено условие невырожденности (6.3). При этом производные по направлениям вычисляются соответственно по формулам (8.11), (8.12) и (8.20).*

Замечание. Из (8.11) и (8.12) заключаем, что $f(x)$ квазидифференцируема для $x_0 \in \text{int } \Omega$ и $x_0 \notin \Omega$. Формула (8.20) показывает, что в граничных точках функция $f(x)$ уже не является квазидифференцируемой (по крайней мере из (8.20) нельзя сделать вывод о квазидифференцируемости функции $f(x)$). Приведем пример, показывающий, что условие невырожденности в граничной точке x_0 существенно для дифференцируемости в точке x_0 по направлениям.

Пример 1. Пусть

$$x = (x^{(1)}, x^{(2)}) \in E_2, \quad \Omega = \{0\} \cup \bigcup_{k=0}^{\infty} \mathcal{D}_k,$$

где $\mathcal{D}_k$ — выпуклая оболочка точек A_k, B_k, C_k ,

$$A_k = (2^{-k}, 0), \quad B_k = (2^{1-k}, 0), \quad C_k = (2^{-k}, 2^{-k}).$$

Множество Ω изображено на рис. 46. Положим $x_0 = (0, 0)$, $g_0 = (0, 1)$. Ясно, что для точки x_0 конус воз-

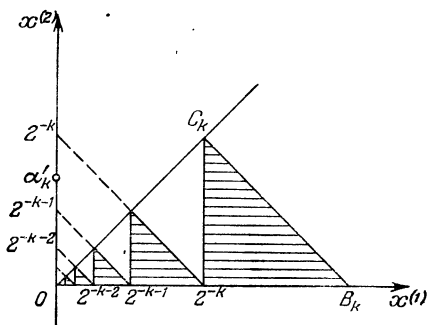


Рис. 46.

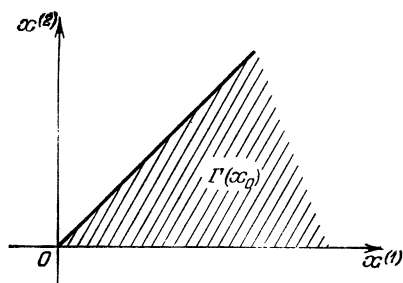


Рис. 47.

можных в широком смысле направлений $\Gamma(x_0)$ есть (рис. 47)

$$\Gamma(x_0) = \{v = \lambda z \mid \lambda \geq 0, z = (1, \alpha), \alpha \in [0, 1]\}.$$

Рассмотрим последовательности чисел $\{\alpha_k\}$, где $\alpha_k = 2^{-k}$, и точек $x_k = x_0 + \alpha_k g = (0, 2^{-k})$. Очевидно, что $f(x_k) = \min_{y \in \Omega} \|x_k - y\| = \|x_k - C_{k-1}\| =$

$$= \sqrt{\left(\frac{1}{2^{k+1}}\right)^2 + \left(\frac{1}{2^{k+1}} - \frac{1}{2^k}\right)^2} = \sqrt{2} \cdot 2^{-(k+1)},$$

т. е.

$$\begin{aligned} H(\alpha_k) &= \alpha_k^{-1} [f(x_0 + \alpha_k g) - f(x_0)] = \\ &= \sqrt{2} \cdot 2^k \cdot 2^{-(k+1)} = \frac{\sqrt{2}}{2} \quad \forall k. \end{aligned}$$

Пусть теперь

$$\alpha'_k = \frac{1}{2} \left(\frac{1}{2^k} + \frac{1}{2^{k+1}} \right) = \frac{3}{2^{k+2}}.$$

Для точек $x'_k = x_0 + \alpha'_k g_0 = \left(0, \frac{3}{2^{k+2}} \right)$ будет

$$\begin{aligned} f(x'_k) &= \min_{y \in \Omega} \|x'_k - y\| = \|x'_k - C_{k-2}\| = \\ &= \sqrt{\left(\frac{1}{2^{k+2}} \right)^2 + \left(\frac{1}{2^{k+2}} - \frac{3}{2^{k+2}} \right)^2} = \frac{\sqrt{5}}{2^{k+2}}, \end{aligned}$$

т. е.

$$H(\alpha'_k) = (\alpha'_k)^{-1} [f(x_0 + \alpha'_k g_0) - f(x_0)] = \sqrt{5} \cdot 3^{-1} \quad \forall k.$$

Таким образом, пределы функции $H(\alpha)$ по последовательностям $\{\alpha_k\}$ и $\{\alpha'_k\}$ не совпадают (они равны со-

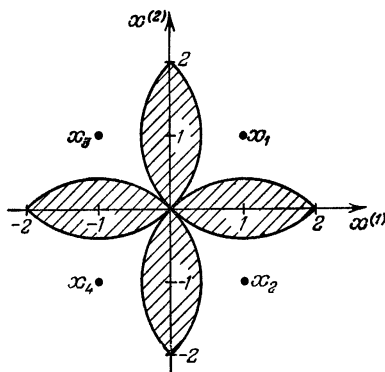


Рис. 48.

ответственно $\sqrt{2} \cdot 2^{-1}$ и $\sqrt{5} \cdot 3^{-1}$), т. е. функция $f(x) = \min_{y \in \Omega} \|x - y\|$ в точке $x_0 = (0, 0)$ не является дифференцируемой по направлению g_0 .

Задача 8.2. Пусть $x = (x^{(1)}, x^{(2)}) \in E_2$. Рассмотрим функции

$$r_1(x) = (x - x_1)^2 - 2, \quad r_2(x) = (x - x_2)^2 - 2,$$

$$r_3(x) = (x - x_3)^2 - 2, \quad r_4(x) = (x - x_4)^2 - 2,$$

где $x_1 = (1, 1)$, $x_2 = (1, -1)$, $x_3 = (-1, 1)$, $x_4 = (-1, -1)$.

Положим

$$\begin{aligned} h_1(x) &= \max \{r_1(x), r_2(x)\}, & h_2(x) &= \max \{r_1(x), r_4(x)\}, \\ h_3(x) &= \max \{r_3(x), r_4(x)\}, & h_4(x) &= \max \{r_2(x), r_3(x)\}. \end{aligned}$$

Пусть

$$h(x) = \min_{i \in 1:4} h_i(x).$$

Образуем множество $\Omega = \{x \in E_2 | h(x) \leq 0\}$ и функцию $f(x) = \min_{y \in \Omega} \|x - y\|$. Множество Ω изображено на рис. 48. Для точки

$x_0 = (0, 0)$ найти $\frac{\partial f(x_0)}{\partial g}$ (надо воспользоваться формулой (8.20), для этого найти $\frac{\partial h(x_0)}{\partial g}$ и $\overline{\partial h(x_0)}$). Множество Ω также называют пропеллером Карлсона — Пшеничного.

§ 9. Неявные функции

Пусть $f(z) = f(x, y)$ — функция двух переменных, $z = (x, y) \in E_2$, и пусть $z_0 = (x_0, y_0)$ — решение уравнения $f(z) = 0$, т. е.

$$f(x_0, y_0) = 0. \quad (9.1)$$

Предположим, что функция $f(z)$ квазидифференцируема в окрестности точки z_0 . Зафиксируем любое $g \in E_1$, $\|g\| = 1$ (таких g в E_1 может быть только два: $g_1 = +1$ и $g_2 = -1$), и рассмотрим функцию $F(\alpha, y) = f(x_0 + \alpha g, y)$, где $\alpha \geq 0$, $y \in E_1$.

Требуется найти условия, при которых существуют число $\alpha_0 > 0$ и непрерывная функция $y(\alpha)$, заданная на $[0, \alpha_0]$, такие, что

$$F(\alpha, y(\alpha)) = f(x_0 + \alpha g, y(\alpha)) = 0 \quad \forall \alpha \in [0, \alpha_0]. \quad (9.2)$$

Так как $f(z)$ — квазидифференцируемая в окрестности точки z_0 функция, то

$$f(z_0 + \alpha \eta) = f(z_0) + \alpha \left[\max_{v \in \underline{\partial} f(z_0)} (v, \eta) + \min_{w \in \overline{\partial} f(z_0)} (w, \eta) \right] + o(\alpha). \quad (9.3)$$

Здесь $\eta = (g, q)$. Допустим, что существует $q_0 \in E_1$ такое, что

$$\frac{\partial f(z_0)}{\partial \eta_0} = \max_{v \in \underline{\partial} f(z_0)} (v, \eta_0) + \min_{w \in \overline{\partial} f(z_0)} (w, \eta_0) = 0, \quad (9.4)$$

где $\eta_0 = (g, q_0)$. Предположим, что функция $o(\alpha) =$

$= o(\alpha, q)$ в (9.3) такова, что $\frac{o(\alpha, q)}{\alpha} \xrightarrow{\alpha \rightarrow +0} 0$ равномерно по q в некоторой окрестности точки $q_0 \in E_1$. Положим $v = (v_1, v_2)$, $w = (w_1, w_2)$ и рассмотрим функцию

$$h(q) = \max_{v \in \partial f(z_0)} (v_1 g + v_2 q) + \min_{w \in \partial f(z_1)} (w_1 g + w_2 q). \quad (9.5)$$

Из (9.4) следует, что

$$h(q_0) = 0. \quad (9.6)$$

Теорема 9.1. Пусть $h(q_0) = 0$. Если функция $h(q)$ строго монотонна в окрестности точки q_0 , то существует функция $y(\alpha)$, заданная и непрерывная при $\alpha \in [0, \alpha_0]$, где $\alpha_0 > 0$, и такая, что выполнено (9.2). При этом $y(\alpha)$ имеет производную справа в точке $\alpha = 0$ и $y'_+(0) = q_0$.

Доказательство аналогично доказательству теоремы о неявных функциях в гладком случае и представляется читателю.

Замечание 1. Факт существования неявной функции устанавливается рассмотрением функции (9.5), свойства которой определяются лишь квазидифференциалом функции $f(z)$ в точке z_0 . В частности, может оказаться, что q_0 , удовлетворяющее (9.6), неединственно.

Пример 1. Пусть $f(z) = |x| - |y|$, $z_0 = (0, 0)$. Очевидно, что $f(z_0) = 0$. Эта функция изучалась в примере 5 § 3. Возьмем $g = +1$. Из рис. 36 ясно, что уравнение (9.5) имеет два решения $q_0 = +1$ и $\bar{q}_0 = -1$. Ясно также, что функция $h(q)$ в окрестности точки q_0 строго убывает, а в окрестности точки $\bar{q}_0$ — строго возрастает, поэтому выполнены условия теоремы и, значит, существует два решения уравнения $|\alpha| - |y| = 0$ (при $\alpha > 0$) $y(\alpha)$ и $\bar{y}(\alpha)$, причем $y'_+(0) = +1$, $\bar{y}'_+(0) = -1$. Аналогично из рис. 36 ясно, что при $g = -1$ уравнение (9.5) тоже имеет два решения: $q_0 = +1$ и $\bar{q}_0 = -1$, а функция $h(q)$ строго монотонна в окрестности этих точек.

Пример 2. Пусть $f(z) = ||x| + y|$, $z_0 = (0, 0)$. Ясно, что $f(z_0) = 0$. Эта функция изучалась в примере 10 § 3. Из рис. 41 ясно, что при $g = +1$ уравнение (9.5) имеет решение $q_0 = -1$, однако функция $h(q)$ не является строго монотонной. Тем не менее решение уравнения $||\alpha| + y| = 0$ существует, причем $y'_+(0) = -1$.

Задание 9.1. Изучить случай, когда $f(z)$ является суб- или супердифференцируемой функцией в точке z_0 , для которой $f(z_0) = 0$.

З а м е ч а н и е 2. Рассмотрим систему уравнений

$$f_i(z) \equiv f_i(x, y) = 0 \quad \forall i \in I \equiv 1:n. \quad (9.7)$$

Здесь $x \in E_m$, $y \in E_n$. Пусть известно решение $z_0 = [x_0, y_0]$ системы (9.7): $f_i(z_0) = 0 \quad \forall i \in I$. Возьмем $g \in E_m$ и рассмотрим систему

$$f_i(x_0 + \alpha g, y) = 0 \quad \forall i \in I.$$

Функции $f_i(z)$ предполагаются непрерывными и квазидифференцируемыми в точке z_0 . Наша задача — выяснить, существуют ли число $\alpha_0 > 0$ и непрерывная вектор-функция $y(\alpha)$ такие, что

$$y(0) = y_0, \quad f_i(x_0 + \alpha g, y(\alpha)) = 0 \quad \forall \alpha \in [0, \alpha_0], \quad \forall i \in I. \quad (9.8)$$

Из квазидифференцируемости функций $f_i(z)$ имеем

$$\begin{aligned} f_i(x_0 + \alpha g, y_0 + \alpha q) &= \\ &= f_i(x_0, y_0) + \alpha \left[\max_{v_i \in \partial f_i(z_0)} ((v_{1i}, g) + (v_{2i}, q)) + \right. \\ &\quad \left. + \min_{w_i \in \partial f_i(z_0)} ((w_{1i}, g) + (w_{2i}, q)) \right] + o_i(\alpha) \quad \forall i \in I, \end{aligned} \quad (9.9)$$

где $v_i = [v_{1i}, v_{2i}] \in E_m \times E_n$, $w_i = [w_{1i}, w_{2i}] \in E_m \times E_n$.

Рассмотрим систему

$$\begin{aligned} F_i(q) &\equiv \max_{v_i \in \partial f_i(z_0)} ((v_{1i}, g) + (v_{2i}, q)) + \\ &\quad + \min_{w_i \in \partial f_i(z_0)} ((w_{1i}, g) + (w_{2i}, q)) = 0 \quad \forall i \in I. \end{aligned} \quad (9.10)$$

Здесь неизвестным является вектор $q \in E_n$.

Из (9.9) следует, что вопрос о существовании неявной функции нескольких переменных сводится к изучению свойств решений системы (9.10). Эти системы, следуя Беллману и Калабе, будем называть *квазилинейными*.

§ 1. Необходимые и достаточные условия минимума выпуклой функции на E_n

Пусть $f(x)$ — выпуклая на E_n функция. Напомним (§ 6 гл. I), что, для того чтобы в точке x^* выпуклая функция $f(x)$ достигала своего наименьшего на E_n значения, необходимо и достаточно, чтобы

$$0 \in \partial f(x^*).$$

Пусть $\varepsilon \geq 0$. Точка $x_\varepsilon \in E_n$ называется ε -стационарной точкой функции $f(x)$, если

$$0 \in \partial_\varepsilon f(x_\varepsilon)$$

или, что то же самое,

$$0 \leq f(x_\varepsilon) - f^* \leq \varepsilon, \quad (1.1)$$

где $f^* = \inf_{x \in E_n} f(x)$.

Напомним также, что если $0 \notin \partial_\varepsilon f(x_0)$ и $\frac{\partial_\varepsilon f(x_0)}{\partial g_0} = \max_{v \in \partial_\varepsilon f(x_0)} (v, g_0) < 0$, где $g_0 \in E_n$, то (см. п. 3 § 8 гл. I)

$$f(x_0) - \inf_{\alpha > 0} f(x_0 + \alpha g_0) > \varepsilon. \quad (1.2)$$

Введем обозначения:

$$\psi_\varepsilon(x) = \min_{\|g\|=1} \max_{v \in \partial_\varepsilon f(x)} (v, g);$$

$$\rho_\varepsilon(x) = \min_{v \in [\partial_\varepsilon f(x)]_{fr}} \|v\|,$$

$$\rho(x) = \rho_0(x), \quad d_\varepsilon(x) = \min_{v \in \partial_\varepsilon f(x)} \|v\|, \quad d(x) = d_0(x).$$

Ясно, что если $d_\varepsilon(x) > 0$, то $\rho_\varepsilon(x) = d_\varepsilon(x)$, а если $d_\varepsilon(x) = 0$, то $\rho_\varepsilon(x)$ есть радиус наибольшего шара с центром в начале координат, содержащегося в $\partial_\varepsilon f(x)$.

Связь между $\psi_\varepsilon(x)$, $\rho_\varepsilon(x)$ и $d_\varepsilon(x)$ устанавливает Лемма 1.1. Имеет место следующее соотношение:

$$\psi_\varepsilon(x) = \begin{cases} -\rho_\varepsilon(x) = -d_\varepsilon(x), & \text{если } d_\varepsilon(x) > 0, \\ \rho_\varepsilon(x), & \text{если } d_\varepsilon(x) = 0. \end{cases}$$

Доказательство получаем, применяя лемму 6.4 гл. I для случая $B = \partial_\varepsilon f(x)$, $\Gamma^+ = \{\emptyset\}$.

Лемма 1.2. При любых $x, y \in E_n$ и $\varepsilon \geq 0$ справедливо неравенство

$$f(y) - f(x) \geq \psi_\varepsilon(x) \|y - x\| - \varepsilon. \quad (1.3)$$

Доказательство. В силу определения $\partial_\varepsilon f(x)$ и $\psi_\varepsilon(x)$ имеем

$$f(y) - f(x) \geq \max_{v \in \partial_\varepsilon f(x)} (v, y - x) - \varepsilon \geq \|y - x\| \psi_\varepsilon(x) - \varepsilon. \quad \blacksquare$$

Лемма 1.3. Пусть $\varepsilon > 0$, x_ε — ε -стационарная точка функции $f(x)$ на E_n и $\psi_\varepsilon(x_\varepsilon) > 0$.

Тогда $f^* > -\infty$, множество $\mathcal{M}^* = \{x \in E_n \mid f(x) = f^*\}$ есть непустой выпуклый компакт,

$$f(x_\varepsilon) - f^* < \varepsilon, \quad (1.4)$$

и если $f^* < f(x_\varepsilon)$, то

$$(f(x_\varepsilon) - f^*) (\rho(x_\varepsilon))^{-1} \leq \|x^* - x_\varepsilon\| \leq (f^* - f(x_\varepsilon) + \varepsilon) (\rho_\varepsilon(x_\varepsilon))^{-1} \quad (1.5)$$

где x^* — точка минимума функции $f(x)$ на E_n , $\rho(x) = \rho_0(x)$.

Доказательство. Так как $\psi_\varepsilon(x_\varepsilon) > 0$, то $0 \in \partial_\varepsilon f(x_\varepsilon)$ и из (1.1)

$$f(x_\varepsilon) - f^* \leq \varepsilon; \quad (1.6)$$

следовательно, $f^* > -\infty$.

В силу (1.3) для любого $x \in E_n$ будет

$$f(x) - f(x_\varepsilon) \geq \psi_\varepsilon(x_\varepsilon) \|x - x_\varepsilon\| - \varepsilon.$$

Поэтому

$$\|x - x_\varepsilon\| \leq (f(x) - f(x_\varepsilon) + \varepsilon) (\psi_\varepsilon(x_\varepsilon))^{-1}. \quad (1.7)$$

Из (1.7), в частности, следует, что лебегово множество $\mathcal{D}(x_e) = \{x \in E_n | f(x) \leq f(x_e)\}$ ограничено, более того,

$$\mathcal{D}(x_e) \subset S_\delta(x_e), \quad \text{где } \delta = \varepsilon(\psi_\varepsilon(x_e))^{-1}.$$

Тем более непусто и ограничено множество $\mathcal{M}^* \subset \mathcal{D}(x_e)$. Замкнутость и выпуклость множества $\mathcal{M}^*$ очевидны.

Допустим теперь, что

$$f(x_e) - f^* = \varepsilon. \quad (1.8)$$

Полагая в (1.7) $x = x^*$ и учитывая (1.8), получаем $x_e = x^*$, а это противоречит (1.8). Таким образом, равенство (1.8) невозможно. Теперь (1.4) вытекает из (1.6).

Далее, из леммы 1.1 следует, что $\psi_\varepsilon(x_e) = \rho_\varepsilon(x_e)$, ибо $d_\varepsilon(x_e) = 0$, и $\psi(x_e) = -\rho(x_e)$, ибо $d(x_e) > 0$. Теперь неравенство в правой части (1.5) получаем, полагая в (1.7) $x = x^*$, а неравенство в левой части (1.5) следует из того, что согласно (1.3)

$$\begin{aligned} f^* - f(x_e) &= f(x^*) - f(x_e) \geq \\ &\geq \psi(x_e) \|x^* - x_e\| = -\rho(x_e) \|x^* - x_e\|. \quad \blacksquare \end{aligned}$$

З а м е ч а н и е. Условие $f^* < f(x_e)$ требуется только для справедливости левостороннего неравенства в (1.5).

§ 2. Минимизация гладкой функции

Ряд методов негладкой оптимизации, обсуждаемых ниже, в идейном плане тесно связан с методами, разработанными для минимизации гладких функций.

Пусть $f(x)$ — непрерывно дифференцируемая на E_n функция, не обязательно выпуклая. Для того чтобы $f(x)$ достигала в точке $x^* \in E_n$ своего наименьшего на E_n значения, необходимо, а если $f(x)$ — выпуклая на E_n функция, то и достаточно, чтобы

$$f'(x^*) = 0. \quad (2.1)$$

Точка x^* , удовлетворяющая (2.1), называется стационарной точкой функции $f(x)$ на E_n . Если $f'(x_0) \neq 0$, то

направление $g(x_0) = -f'(x_0)/\|f'(x_0)\|$ является направлением наискорейшего спуска функции $f(x)$ в точке x_0 ,

$$\text{т. е. } \frac{\partial f(x_0)}{\partial g(x_0)} = \min_{\|g\|=1} \frac{\partial f(x_0)}{\partial g}.$$

Напомним, что в этом случае $\frac{\partial f(x_0)}{\partial g} = (f'(x_0), g)$.

Многие методы оптимизации используют условие (2.1).

1. Метод непрерывного спуска. Рассмотрим систему обыкновенных дифференциальных уравнений

$$\dot{x}(t) = -A(x(t))f'(x), \quad (2.2)$$

$$x(0) = x_0, \quad (2.3)$$

где $A(x)$ — непрерывная на $\mathcal{D}(x_0) \equiv \{x \in E_n | f(x) \leq f(x_0)\}$ матричная функция размерности $n \times n$, причем существует $m > 0$ такое, что

$$(A(x)z, z) \geq m\|z\|^2, \quad \forall z \in E_n, \quad \forall x \in \mathcal{D}(x_0). \quad (2.4)$$

При наложенных условиях решение $x(t, x_0)$ системы (2.2) — (2.3) существует на $[0, \infty)$. Отметим также, что

$$x(t, x_0) \in \mathcal{D}(x_0) \quad \forall t \geq 0. \quad (2.5)$$

Лемма 2.1. Если множество $\mathcal{D}(x_0)$ ограничено, то всякая предельная точка решения $x(t, x_0)$ является стационарной точкой $f(x)$ на E_n .

В качестве дискретных аналогов непрерывного метода (2.2) можно получить метод наискорейшего спуска, метод наискорейшего спуска с постоянным шагом и метод Ньютона (см. [16], [59], [61], [105], [114]).

2. Метод наискорейшего спуска. В качестве начального приближения выберем произвольное $x_0 \in E_n$. Пусть уже найдено $x_k \in E_n$. Если $f'(x_k) = 0$, то точка x_k — стационарная, и процесс прекращается. Если же $f'(x_k) \neq 0$, то положим $g_k = f'(x_k)$ и рассмотрим луч

$$\{x \in E_n | x = x_{k\alpha} \equiv x_k - \alpha g_k, \alpha \geq 0\}.$$

Найдем $\min_{\alpha \geq 0} f(x_{k\alpha}) = f(x_{k\alpha_k})$ (предполагаем, что $\inf_{\alpha \geq 0} f(x_{k\alpha})$

достигается). Положим $x_{k+1} = x_{k\alpha_k}$.

В результате строим последовательность $\{x_k\}$. Если эта последовательность состоит из конечного числа то-

чек, то последняя полученная точка является стационарной по построению. В противном случае справедлива

Теорема 2.1. Если множество $\mathcal{D}(x_0)$ ограничено, то всякая предельная точка последовательности $\{x_k\}$ является стационарной.

Оказывается, что при весьма естественных предположениях скорость сходимости метода наискорейшего спуска — геометрическая.

Замечание. В качестве g_k можно выбирать любой вектор такой, что

$$(g_k, f'(x_k)) \geq \|g_k\| \cdot \|f'(x_k)\| \cdot \theta,$$

где $g_k \neq 0$, $\theta \in (0, 1]$ фиксировано.

3. Метод наискорейшего спуска с постоянным шагом. Пусть $f'(x)$ удовлетворяет на E_n условию Липшица: существует $L < \infty$ такое, что $\|f'(x) - f'(z)\| \leq L\|x - z\| \forall x, z \in E_n$.

Зафиксируем $\lambda \in (0, \frac{2}{L})$. Выберем $x_0 \in E_n$. Пусть уже найдено $x_k \in \mathcal{D}(x_0)$. Если $f'(x_k) = 0$, то точка x_k — стационарная, и процесс прекращается. Если же $f'(x_k) \neq 0$, то полагаем $x_{k+1} = x_k - \lambda f'(x_k)$. Для последовательности $\{x_k\}$ справедлива теорема 2.1.

§ 3. Метод наискорейшего спуска

1. Пусть $f(x)$ — выпуклая на E_n функция. В § 8 гл. I было введено понятие ε -производной выпуклой функции $f(x)$ в точке x_0 по направлению g :

$$\frac{\partial_\varepsilon f(x_0)}{\partial g} = \max_{v \in \partial_\varepsilon f(x_0)} (v, g).$$

Если $0 \in \partial_\varepsilon f(x_0)$, то точка x_0 — ε -стационарная, и тогда

$$0 \leq f(x_0) - f^* \leq \varepsilon,$$

где $f^* = \inf_{x \in E_n} f(x)$.

Если $0 \notin \partial_\varepsilon f(x_0)$, то вектор $g_\varepsilon(x_0) = -v_\varepsilon(x_0) \|v_\varepsilon(x_0)\|^{-1}$, где $v_\varepsilon(x_0) \in \partial_\varepsilon f(x_0)$, $\|v_\varepsilon(x_0)\| = \min_{v \in \partial_\varepsilon f(x_0)} \|v\| = d_\varepsilon(x_0)$, яв-

ляется направлением ε -наискорейшего спуска, т. е.

$$\frac{\partial f(x_0)}{\partial g_\varepsilon(x_0)} = \min_{\|g\|=1} \frac{\partial f(x_0)}{\partial g}.$$

Направление $g_\varepsilon(x_0)$ единственно.

В § 8 гл. I было также показано, что если для некоторого $g \in E_n$ оказалось $\frac{\partial f(x_0)}{\partial g} < 0$, то

$$\inf_{\alpha > 0} f(x_0 + \alpha g) < f(x_0) - \varepsilon, \quad (3.1)$$

т. е. в направлении g функция $f(x)$ может быть уменьшена не меньше чем на ε .

2. ε -алгоритм. Зафиксируем $\varepsilon > 0$ и выберем произвольное начальное приближение $x_0 \in E_n$. Пусть уже найдено k -е приближение $x_k \in E_n$. Если $0 \in \partial_\varepsilon f(x_k)$, то x_k является ε -стационарной точкой функции $f(x)$ на E_n . Если же $0 \notin \partial_\varepsilon f(x_k)$, то найдем такой вектор $g_k \in E_n$, $\|g_k\| = 1$, что

$$\frac{\partial f(x_k)}{\partial g_k} = \max_{v \in \partial_\varepsilon f(x_k)} (v, g_k) < 0. \quad (3.2)$$

После этого на луче $\{x_k(\alpha) = x_k + \alpha g_k | \alpha \geq 0\}$ найдем точку $x_{k+1} = x_k(\alpha_k)$, в которой

$$f(x_{k+1}) \leq f(x_k) - \varepsilon. \quad (3.3)$$

Из (3.1) следует, что такая точка x_{k+1} обязательно найдется.

На основании (3.3) заключаем, что либо за конечное число шагов будет получена ε -стационарная точка функции $f(x)$ на E_n (это произойдет, если $f^* > -\infty$), либо $f^* = -\infty$.

Поясним, как в случае $0 \notin \partial_\varepsilon f(x_k)$ находить вектор g_k , удовлетворяющий (3.2). Очевидно, что в качестве g_k может быть взят вектор $g_\varepsilon(x_k)$ — направление ε -наискорейшего спуска функции $f(x)$ в точке x_k . Действительно, пусть

$$0 \notin \partial_\varepsilon f(x_k), \quad v_\varepsilon \in \partial_\varepsilon f(x_k), \quad \|v_\varepsilon(x_k)\| = d_\varepsilon(x_k), \\ g_\varepsilon(x_k) = -v_\varepsilon(x_k) \|v_\varepsilon(x_k)\|^{-1}.$$

Тогда

$$\frac{\partial_e f(x_k)}{\partial g_e(x_k)} = -\|v_e(x_k)\| < 0.$$

Если в качестве g_k брать $g_e(x_k)$ — направление ε -наискорейшего спуска функции $f(x)$ в точке x_k , то получающийся при этом метод назовем *методом ε -наискорейшего спуска*.

Для нахождения $v_e(x_k)$ требуется решить следующую задачу: найти $v_e(x_k) \in \partial_e f(x_k)$ такой, что

$$\|v_e(x_k)\| = \min_{v \in \partial_e f(x_k)} \|v\|. \quad (3.4)$$

Если в процессе решения задачи (3.4) найден такой вектор $v_k \in \partial_e f(x_k)$, что $v_k \neq 0$,

$$(v, v_k) \geq \theta (v_k, v_k) \quad \forall v \in \partial_e f(x_k), \quad (3.5)$$

где $\theta \in (0, 1]$, то вектор $g_k = -v_k \|v_k\|^{-1}$ будет удовлетворять (3.2).

З а м е ч а н и е 1. Если вместо направления g_k такого, что $\frac{\partial_e f(x_k)}{\partial g_k} < 0$, брать направление, для которого

$\frac{\partial f(x_k)}{\partial g_k} < 0$, то метод может не сходиться (среди всех предельных точек последовательности $\{x_k\}$ может не оказаться точки минимума функции $f(x)$ на E_n). Это объясняется тем, что в этом случае нет оценки типа (3.1) — гарантированного убывания функции $f(x)$ на каждом шаге.

3. В изложенной схеме ε -алгоритма для построения точки x_{k+1} , удовлетворяющей (3.3), в общем случае необходимо прибегать к одномерной минимизации функции $f(x_k + \alpha g_k)$ по $\alpha \geq 0$, в процессе которой может потребоваться вычислять функцию $f(x)$ в достаточно большом числе точек. Если $g_k = -v_k \|v_k\|^{-1}$, где v_k удовлетворяет (3.5), то можно выбирать длину шага из точки x_k в направлении g_k , вычисляя функцию $f(x)$ на луче $\{x_k(\alpha) | \alpha \geq 0\}$ лишь один раз. Одним из способов такого выбора чисел α_k является следующий:

$$\alpha_k = \varepsilon \left[\frac{\partial_e f(x_k)}{\partial g_k} - \frac{\partial f(x_k)}{\partial g_k} \right]^{-1}.$$

Если числа α_k выбирать таким образом, то в силу леммы 8.6 гл. I при каждом k будут выполняться неравенства

$$f(x_{k+1}) \leq f(x_k) + \alpha_k \frac{\partial_e f(x_k)}{\partial g_k}. \quad (3.6)$$

Если лебегово множество $\mathcal{D}(x_0) = \{x \in E_n | f(x) \leq f(x_0)\}$ ограничено, то любая предельная точка последовательности $\{x_{k+1} = x_k + \alpha_k g_k\}$ будет ε -стационарной точкой функции $f(x)$ на E_n . Действительно, если допустить, что существует подпоследовательность $\{x_{k_s}\}$ такая, что $x_{k_s} \rightarrow x^*$, но $0 \notin \partial_e f(x^*)$ (тогда $d_e(x^*) > 0$), то при достаточно больших k_s , в силу полунепрерывности сверху $\partial_e f(x)$, окажется (см. (3.5))

$$\frac{\partial_e f(x_{k_s})}{\partial g_{k_s}} \leq -\theta \|v_{k_s}\| \leq -\theta \|v_e(x_{k_s})\| \leq -\frac{1}{2} \theta d_e(x^*).$$

Отсюда, из (3.6) и вида α_k имеем

$$f(x_{k_s+1}) \leq f(x_{k_s}) - \frac{1}{2} \theta \varepsilon b^{-1} d_e(x^*),$$

где $b = \max_{x \in \mathcal{D}(x_0)} \max_{v \in \partial_e f(x)} \|v\|$.

Следовательно, $f(x_k) \xrightarrow{k \rightarrow \infty} -\infty$, а это противоречит ограниченности $f(x)$ на $\mathcal{D}(x_0)$. ■

4. Покажем теперь, как применять ε -алгоритм для нахождения точек минимума функции $f(x)$ на E_n .

Выберем произвольным образом точку $x_0 \in E_n$ и число $\varepsilon_0 > 0$. Применяя ε -алгоритм п. 2 с $\varepsilon = \varepsilon_0$ из начального приближения $x_{00} = x_0$, либо выясняем, что $f^* = -\infty$, либо за конечное число шагов t_0 находим точку $x_1 = x_{0t_0}$, которая является ε_0 -стационарной точкой функции $f(x)$ на E_n . Далее рассматриваем случай, когда $f^* > -\infty$. Положим $x_{10} = x_1$ и выберем $\varepsilon_1 < \varepsilon_0$. Из точки x_{10} , двигаясь с помощью ε -алгоритма с $\varepsilon = \varepsilon_1$, за конечное число шагов t_1 найдем точку $x_2 = x_{1t_1}$, которая будет ε_1 -стационарной точкой функции $f(x)$ на E_n .

Продолжая аналогично, построим последовательность $\{x_k\}$, каждая точка x_k которой является ε_{k-1} -стационарной точкой функции $f(x)$ на E_n .

Теорема 3.1. Пусть $f^* > -\infty$, $\varepsilon_k \rightarrow 0$. Тогда

$$\lim_{k \rightarrow \infty} f(x_k) = f^*, \quad (3.7)$$

и если множество $\mathcal{D}(x_0)$ ограничено, то любая предельная точка последовательности $\{x_k\}$ является точкой минимума функции $f(x)$ на E_n .

Доказательство. Так как при каждом k точка x_k есть ε_{k-1} -стационарная точка функции $f(x)$ на E_n , то

$$0 \leq f(x_k) - f^* \leq \varepsilon_{k-1}. \quad (3.8)$$

Из (3.8) и того, что $\varepsilon_k \rightarrow 0$, вытекает (3.7). Если $\mathcal{D}(x_0)$ ограничено, то, поскольку $x_k \in \mathcal{D}(x_0)$ при каждом k , последовательность $\{x_k\}$ имеет предельные точки. Пусть $x_{k_s} \rightarrow x^*$. Тогда, в силу (3.7), $f(x^*) = f^*$. Теорема доказана.

5. Опишем еще одну модификацию ε -алгоритма для отыскания точек минимума функции $f(x)$ на всем пространстве.

Выберем произвольным образом начальное приближение $x_0 \in E_n$. Если $0 \in \partial f(x_0)$, то x_0 — точка минимума. Пусть $0 \notin \partial f(x_0)$. Найдем такое число $\varepsilon_0 > 0$, что $0 \in \partial_{\varepsilon_0} f(x_0)$, т. е. $f(x_0) - f^* \leq \varepsilon_0$. В качестве ε_0 можно взять, например, $f(x_0) - f^*$.

Заметим, что если при любом $\varepsilon > 0$ будет

$$0 \notin \partial_{\varepsilon} f(x_0), \quad (3.9)$$

то $f^* = -\infty$. Действительно, условие $0 \in \partial_{\varepsilon} f(x_0)$ эквивалентно условию $f(x_0) - f^* \leq \varepsilon$. Поэтому (3.9) равносильно тому, что $f(x_0) - f^* > \varepsilon$. Поскольку это неравенство справедливо для всех $\varepsilon > 0$, то $f^* = -\infty$.

Предположим, что $f^* > -\infty$.

Будем перебирать числовую последовательность $\varepsilon_{0q} = 2^{-q}\varepsilon_0$, $q = 1, 2, \dots$, до тех пор, пока не окажется, что при некоторых $q_0 \geq 1$ и $\varepsilon_1 = \varepsilon_{0q_0}$ $0 \notin \partial_{\varepsilon_1} f(x_0)$, причем $\varepsilon_1 \leq 2^{-1}\varepsilon_0$ и

$$0 \in \partial_{2\varepsilon_1} f(x_0), \quad f(x_0) - f^* \leq 2\varepsilon_1. \quad (3.10)$$

Теперь найдем такой вектор $g_0 \in E_n$, $\|g_0\| = 1$, что $\frac{\partial_{\varepsilon_1} f(x_0)}{\partial g_0} < 0$, и на луче $\{x_0(\alpha) = x_0 + \alpha g_0 \mid \alpha \geq 0\}$ найдем такую точку $x_1 = x_0(\alpha_0)$, что

$$f(x_1) - f(x_0) \leq -\varepsilon_1. \quad (3.11)$$

Согласно (3.1) такая точка x_1 обязательно найдется. Из (3.10) и (3.11) вытекает, что

$$f(x_1) \leq f(x_0) - \varepsilon_1 \leq f^* + \varepsilon_1.$$

Следовательно, $0 \in \partial_{\varepsilon_1} f(x_1)$ и точка x_1 является ε_1 -стационарной точкой функции $f(x)$ на E_n .

Продолжая аналогично, построим последовательность $\{x_k\}$, которая либо конечна (тогда найдена точка минимума), либо каждая ее точка x_k есть ε_k -стационарная точка функции $f(x)$ на E_n , где $\varepsilon_k = 2^{-q_{k-1}} \varepsilon_{k-1}$,

$$0 \leq f(x_k) - f^* < 2^{-q_{k-1}} \varepsilon_{k-1} \leq 2^{-k} \varepsilon_0 \quad \forall k. \quad (3.12)$$

В этом случае справедлива следующая

Теорема 3.2. Пусть $f^* > -\infty$. Тогда

$$\lim_{k \rightarrow \infty} f(x_k) = f^*,$$

причем скорость сходимости геометрическая, и если множество $\mathcal{D}(x_0)$ ограничено, то любая предельная точка последовательности $\{x_k\}$ является точкой минимума функции $f(x)$ на E_n .

Доказательство аналогично доказательству теоремы 3.1. Неравенство (3.12) означает, что скорость сходимости геометрическая.

6. Для нахождения f^* можно поступать также следующим образом. Выберем произвольно начальное приближение $x_0 \in E_n$. Зафиксируем положительные числа a и b , $a < b$.

Предположим, что множество $\mathcal{D}(x_0)$ ограничено и $\mathcal{D}$ — его диаметр.

Пусть уже найдено k -е приближение $x_k \in \mathcal{D}(x_0)$. Возьмем такое $\varepsilon_k > 0$, что

$$\mathcal{D}ad_{\varepsilon_k}(x_k) \leq \varepsilon_k \leq \mathcal{D}bd_{\varepsilon_k}(x_k). \quad (3.13)$$

В силу непрерывности $\partial_{\varepsilon} f(x_k)$ по $\varepsilon > 0$ и того, что $d_{\varepsilon}(x_k) \xrightarrow{\varepsilon \rightarrow 0} 0$, такое ε_k существует.

Найдем $g_k \in E_n$, $\|g_k\| = 1$, такой, что $\frac{\partial_{\varepsilon_k} f(x_k)}{\partial g_k} < 0$, и на луче $\{x_k(\alpha) = x_k + \alpha g_k \mid \alpha \geq 0\}$ найдем точку $x_{k+1} = x_k(\alpha_k)$ такую, что

$$f(x_{k+1}) \leq f(x_k) - \varepsilon_k. \quad (3.14)$$

Продолжая аналогично, построим последовательность $\{x_k\}$. Теорема 3.2 справедлива и в данном случае. Кроме того,

$$f(x_{k+1}) - f^* \leq (1 - a(1+b)^{-1})(f(x_k) - f^*) \quad \forall k. \quad (3.15)$$

Действительно, в силу лемм 1.1 и 1.2 имеем

$$f(x_k) - f^* \leq \varepsilon_k + \mathcal{D}d_{\varepsilon_k}(x_k). \quad (3.16)$$

Объединяя (3.13), (3.14) и (3.16), получаем (3.15).

З а м е ч а н и е 2. Если функция $f(x)$ сильно выпукла или если она строго выпукла и множество $\mathcal{D}(x_0)$ ограничено, то последовательности $\{x_k\}$ из теорем 3.1 и 3.2 сходятся к единственной точке минимума функции $f(x)$ на E_n .

7. Рассмотренные в этом параграфе модификации ε -алгоритма, как и сам ε -алгоритм, представляют в основном теоретический интерес, ибо на каждой итерации необходимо знать все соответствующее множество $\partial_{\varepsilon} f(x)$. К сожалению, класс функций, для которых ε -субдифференциалы эффективно вычисляются, недостаточно широк.

Таким образом, ε -субдифференциалы, с одной стороны, очень удобны и обладают рядом важных и полезных свойств (непрерывность $\partial_{\varepsilon} f(x)$, гарантированная оценка (3.1) уменьшения функции на ε , если $0 \notin \partial_{\varepsilon} f(x)$, возможность получения геометрической скорости сходимости последовательности $\{f(x_k)\}$ к f^* без предположения о сильной выпуклости функции $f(x)$), которых не имеют их аналоги в нелинейном программировании и минимаксных задачах. С другой стороны, ε -субдифференциалы трудно вычисляемы, ибо являются глобальной характеристикой функции, что делает методы, изложенные в этом параграфе, практически малоприменимыми.

П р и м е р. Пусть $f(x) = (Ax, x) + (a, x)$, где A — симметричная положительно определенная $n \times n$ -матрица, $a \in E_n$.

Тогда

$$\partial_{\varepsilon} f(x) = \{2A(x+v) + a \mid (Av, v) \leq \varepsilon\},$$

т. е. даже для такой «простой» функции, как квадратичная, для нахождения $\partial_{\varepsilon} f(x)$ необходимо знать все множество решений нелинейной системы $(Av, v) \leq \varepsilon$.

Далее, метод наискорейшего спуска в данном случае состоит в построении последовательности $\{x_k\}$ по формуле

$$x_{k+1} = x_k - (f'(x_k), f'(x_k)) (2Af'(x_k), f'(x_k))^{-1} f'(x_k).$$

А для нахождения только направления ε -наискорейшего спуска функции $(Ax, x) + (a, x)$ в ε -алгоритме на каждом шаге требуется решить задачу

$$\min_{(Av, v) \leq \varepsilon} \|2A(x + v) + a\|.$$

В результате оказалось, что вспомогательная задача, которую необходимо решать при применении метода ε -наискорейшего спуска, сложнее, по крайней мере формально, исходной задачи минимизации.

§ 4. Субградиентный метод для минимизации выпуклой функции

1. Субградиентный метод. Пусть $f(x)$ — выпуклая конечная на E_n функция. Выберем произвольные точку $x_0 \in E_n$ и последовательность чисел $\{\lambda_k\}$ такую, что

$$\lambda_k \rightarrow +0, \quad \sum_{k=0}^{\infty} \lambda_k = +\infty. \quad (4.1)$$

Пусть уже найдено $x_k \in E_n$. Возьмем любое $v_k \in \partial f(x_k)$. Если $v_k = 0$, то $0 \in \partial f(x_k)$ и точка x_k — точка минимума $f(x)$ на E_n , и процесс прекращается. В противном случае положим

$$x_{k+1} = x_k - \lambda_k v_k \|v_k\|^{-1} \equiv x_k + z_k, \quad (4.2)$$

где

$$z_k = -\lambda_k v_k \|v_k\|^{-1}, \quad \|z_k\| = \lambda_k.$$

Таким образом строим последовательность $\{x_k\}$. Если эта последовательность состоит из конечного числа точек, то по построению последняя полученная точка — точка минимума функции $f(x)$ на E_n . Рассмотрим случай, когда последовательность $\{x_k\}$ бесконечна. Как и выше,

$$\mathcal{M}^* = \{x \in E_n \mid f(x) = f^*\},$$

где

$$f^* = \inf_{x \in E_n} f(x), \quad \rho(x, \mathcal{M}^*) = \min_{z \in \mathcal{M}^*} \|z - x\|$$

(множество $\mathcal{M}^*$ может оказаться пустым).

Теорема 4.1. Если множество $\mathcal{M}^*$ непусто и ограничено, то при $k \rightarrow \infty$

$$\rho(x_k, \mathcal{M}^*) \rightarrow 0, \quad f(x_k) \rightarrow f^*. \quad (4.3)$$

Доказательство. Покажем вначале, что

$$\lim_{k \rightarrow \infty} \rho(x_k, \mathcal{M}^*) = 0. \quad (4.4)$$

Допустим противное, пусть существуют $a > 0$ и $K_1 < \infty$ такие, что

$$\rho(x_k, \mathcal{M}^*) \geq 2a > 0 \quad \forall k \geq K_1. \quad (4.5)$$

Поскольку $\mathcal{M}^*$ — ограниченное множество, то, согласно следствию 4 из леммы 9.1 гл. I, найдется $c > 0$ такое, что

$$\mathcal{D}_c = \{x \mid f(x) \leq f^* + c\} \subset S_a(\mathcal{M}^*), \quad (4.6)$$

где

$$S_a(\mathcal{M}^*) = \{x \mid \rho(x, \mathcal{M}^*) \leq a\}.$$

Существует такое $r > 0$, что

$$S_r(\bar{x}) = \{x \mid \|x - \bar{x}\| \leq r\} \subset \mathcal{D}_c \quad \forall \bar{x} \in \mathcal{M}^*. \quad (4.7)$$

Зафиксируем любое $\bar{x} \in \mathcal{M}^*$ (рис. 49). Из определения $\partial f(x_k)$ имеем

$$f(x) - f(x_k) \geq (v_k, x - x_k) \quad \forall x \in E_n. \quad (4.8)$$

Из (4.5) и (4.6) получаем $f(x_k) > f^* + c \quad \forall k > K_1$, и поэтому из (4.8)

$$\begin{aligned} (v_k, x - x_k) &< 0 \quad \forall x \in \mathcal{D}_c, \\ (z_k, x_k - x) &< 0 \quad \forall x \in \mathcal{D}_c. \end{aligned} \quad (4.9)$$

Возьмем $\tilde{x}_k = \bar{x} - rz_k \|z_k\|^{-1}$. Из (4.7) следует $\tilde{x}_k \in \mathcal{D}_c$, и поэтому из (4.9)

$$(z_k, x_k - \tilde{x}_k) = (-\lambda_k v_k \|v_k\|^{-1}, x_k - \bar{x} + rz_k \|z_k\|^{-1}) < 0,$$

т. е.

$$(z_k, x_k - \bar{x}) < -r\lambda_k.$$

Далее,

$$\begin{aligned}\|x_{k+1} - \bar{x}\|^2 &= \|x_k + z_k - \bar{x}\|^2 = \\ &= \|x_k - \bar{x}\|^2 + 2(z_k, x_k - \bar{x}) + z_k^2 < \|x_k - \bar{x}\|^2 - 2r\lambda_k + \lambda_k^2.\end{aligned}$$

Поскольку $\lambda_k \rightarrow +0$, то при достаточно больших k

$$\|x_{k+1} - \bar{x}\|^2 < \|x_k - \bar{x}\|^2 - r\lambda_k.$$

Продолжая аналогично, получим

$$\|x_{k+s} - \bar{x}\|^2 < \|x_k - \bar{x}\|^2 - r \sum_{i=0}^{s-1} \lambda_{k+i}. \quad (4.10)$$

В силу (4.1) отсюда следует $\|x_k - \bar{x}\| \rightarrow -\infty$, что невозможно. Полученное противоречие и доказывает (4.4).

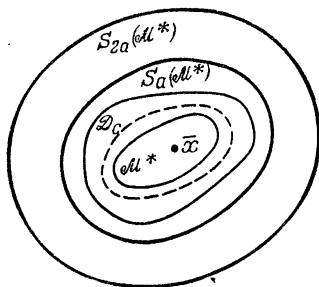


Рис. 49.

Зафиксируем $\delta > 0$. Как и выше, найдем $r_\delta > 0$ такое, что

$$\|x_{k+1} - \bar{x}\|^2 < \|x_k - \bar{x}\|^2 - r_\delta \lambda_k \quad \forall \bar{x} \in \mathcal{M}^* \quad (4.11)$$

для всех достаточно больших k и таких, что $\rho(x_k, \mathcal{M}^*) > \delta$. Из (4.11)

$$\rho^2(x_{k+1}, \mathcal{M}^*) < \rho^2(x_k, \mathcal{M}^*) - r_\delta \lambda_k. \quad (4.12)$$

В силу (4.4) существуют k такие, что

$$\rho(x_k, \mathcal{M}^*) < \delta. \quad (4.13)$$

Поскольку $x_{k+1} = x_k + z_k$, $\|z_k\| = \lambda_k$, то, учитывая (4.1), можно считать, что $\rho(x_{k+1}, \mathcal{M}^*) < 2\delta$ для k , удовлетво-

ряющих (4.13). Если при этом оказалось $\rho(x_{k+1}, \mathcal{M}^*) > \delta$, то в силу (4.12)

$$\rho(x_{k+2}, \mathcal{M}^*) < \rho(x_{k+1}, \mathcal{M}^*) < 2\delta.$$

Далее рассуждаем аналогично. Итак, если $\rho(x_{k+i}, \mathcal{M}^*) < \delta$, то $\rho(x_{k+i+1}, \mathcal{M}^*) < 2\delta$; если же $\delta < \rho(x_{k+i}, \mathcal{M}^*) < 2\delta$, то $\rho(x_{k+i+1}, \mathcal{M}^*) < \rho(x_{k+i}, \mathcal{M}^*) < 2\delta$. Окончательно для всех $i \geq 1$ имеем $\rho(x_{k+i}, \mathcal{M}^*) < 2\delta$. В силу произвольности δ отсюда следует $\rho(x_k, \mathcal{M}^*) \rightarrow 0$, а тогда $f(x_k) \rightarrow f^*$, что и завершает доказательство теоремы.

Следствие. Если $\text{int } \mathcal{M}^* \neq \emptyset$, то последовательность $\{x_k\}$ конечна, т. е. для некоторого k будет $v_k = 0$.

Доказательство. Допустим противное. Предположим, что $v_k \neq 0$ для всех k ; тогда последовательность $\{x_k\}$ бесконечна. Возьмем любое $\bar{x} \in \text{int } \mathcal{M}^*$. Найдется $r > 0$ такое, что $S_r(\bar{x}) \subset \mathcal{M}^*$. Построим точку $\tilde{x}_k = \bar{x} - rz_k \|z_k\|^{-1}$. Ясно, что $\tilde{x}_k \in S_r(\bar{x}) \subset \mathcal{M}^*$, и потому $f(\tilde{x}_k) = f^*$. По определению субградиента

$$f(\tilde{x}_k) - f(x_k) \geq (v_k, \tilde{x}_k - x_k).$$

Так как $f(x_k) \geq f^* = f(\tilde{x}_k)$, то отсюда

$$(v_k, \tilde{x}_k - x_k) \leq f(\tilde{x}_k) - f(x_k) \leq 0,$$

т. е.

$$(v_k, \bar{x} - x_k) \leq r(v_k, z_k \|z_k\|^{-1}).$$

Так как $z_k = -\lambda_k v_k \|v_k\|^{-1}$, то отсюда $(z_k, x_k - \bar{x}) \leq -r\lambda_k$. Рассуждая далее, как при доказательстве теоремы, получаем

$$\|x_{k+1} - \bar{x}\|^2 \leq \|x_k - \bar{x}\|^2 - r\lambda_k. \quad (4.14)$$

Значит, $\|x_{k+s} - \bar{x}\| \xrightarrow{s \rightarrow \infty} -\infty$, что невозможно. Полученное противоречие и доказывает конечность последовательности $\{x_k\}$. ■

Замечание 1. После получения неравенства (4.14) доказательство можно завершить и иначе: из (4.14) вытекает, что расстояние до точки $\bar{x}$ убывает, и из (4.14) заключаем, что за конечное число шагов найдем точку x_{k+s} такую, что $x_{k+s} \in \text{int } \mathcal{M}^*$. Но если $x \in \text{int } \mathcal{M}^*$, то (см. задание 6.2 гл. I) $\partial f(x) = \{0\}$. Значит, $v_{k+s} = 0$, и процесс прекратится, т. е. последовательность $\{x_k\}$ конечна.

З а м е ч а н и е 2. Описанный субградиентный метод часто называют *методом обобщенного градиентного спуска* (ОГС-методом).

2. ε -субградиентный метод. Множество ε -субградиентов (ε -субдифференциал), вообще говоря, богаче множества $\partial f(x)$. Поэтому найти элемент из $\partial_\varepsilon f(x)$ может оказаться проще, чем найти элемент из $\partial f(x)$. В связи с этим естественным образом возникает следующий метод для нахождения ε -стационарных точек.

Зафиксируем $\varepsilon > 0$. Пусть числовая последовательность $\{\lambda_k\}$ удовлетворяет (4.1). Возьмем произвольное $x_0 \in E_n$. Предположим, что уже найдено $x_k \in E_n$. Возьмем любое $v_k \in \partial_\varepsilon f(x_k)$.

Если $v_k = 0$, то $0 \in \partial_\varepsilon f(x_k)$, и точка x_k — ε -стационарная точка функции $f(x)$ на E_n , и, как показано в п. 3 § 8 гл. I, $0 \leq f(x_k) - f^* \leq \varepsilon$, где $f^* = \inf_{x \in E_n} f(x)$.

Если же $v_k \neq 0$, то полагаем

$$x_{k+1} = x_k - \lambda_k v_k \|v_k\|^{-1} = x_k + z_k. \quad (4.15)$$

Теорема 4.2. Если множество $\mathcal{M}^*$ непусто и ограничено, то

$$\lim_{k \rightarrow \infty} \rho(x_k, \mathcal{D}_\varepsilon) = 0, \quad \rho(x_k, S_{a_\varepsilon}(\mathcal{M}^*)) \xrightarrow{k \rightarrow \infty} 0, \quad (4.16)$$

где

$$\mathcal{D}_\varepsilon = \{x \mid f(x) \leq f^* + \varepsilon\}, \quad a_\varepsilon = \max_{x \in Q_\varepsilon} \rho(x, \mathcal{M}^*),$$

$$Q_\varepsilon = \{x \mid f(x) = f^* + \varepsilon\}.$$

Доказательство. Вначале покажем, что

$$\lim_{k \rightarrow \infty} \rho(x_k, \mathcal{D}_\varepsilon) = 0. \quad (4.17)$$

Допустим противное, пусть существуют $a > 0$ и $K_1 < \infty$ такие, что

$$\rho(x_k, \mathcal{D}_\varepsilon) \geq 2a > 0 \quad \forall k > K_1. \quad (4.18)$$

Так как $\mathcal{M}^*$ — ограниченное множество, то и множество $\mathcal{D}_\varepsilon$ тоже ограничено, и потому найдется $c > 0$ такое, что

$$\mathcal{D}_{\varepsilon+c} = \{x \mid f(x) \leq f^* + c + \varepsilon\} \subset S_a(\mathcal{D}_\varepsilon). \quad (4.19)$$

Тем более $\mathcal{D}_c \subset S_a(\mathcal{D}_\varepsilon)$. Существует $r > 0$ такое, что

$$S_r(\bar{x}) = \{x \mid \|x - \bar{x}\| \leq r\} \subset \mathcal{D}_c \quad \forall \bar{x} \in \mathcal{M}^*. \quad (4.20)$$

Зафиксируем любое $\bar{x} \in \mathcal{M}^*$. Из определения $\partial_\varepsilon f(x_k)$ имеем

$$f(x) - f(x_k) \geq (v_k, x - x_k) - \varepsilon \quad \forall x \in E_n. \quad (4.21)$$

Из (4.18) и (4.19) следует $f(x_k) > f^* + \varepsilon + c$, и потому из (4.21)

$$(v_k, x - x_k) \leq f(x) - f(x_k) + \varepsilon < f(x) - f^* - c. \quad (4.22)$$

Для $x \in \mathcal{D}_c$ будет $f(x) \leq f^* + c$. Теперь из (4.22) получаем

$$(v_k, x - x_k) < 0 \quad \forall x \in \mathcal{D}_c. \quad (4.23)$$

Возьмем $\tilde{x}_k = \bar{x} - rz_k \|z_k\|^{-1}$. Из (4.20) $\tilde{x}_k \in \mathcal{D}_c$. Значит, из (4.23) и (4.15)

$$(z_k, x_k - \tilde{x}_k) < 0.$$

Отсюда

$$(z_k, x_k - \bar{x}) < -r\lambda_k.$$

Повторяя далее дословно рассуждения, проведенные при доказательстве теоремы 4.1, устанавливаем, что $\|x_k - \bar{x}\| \rightarrow -\infty$, что невозможно. Итак, (4.17) доказано.

Выше фактически показано, что для любого $\delta > 0$ будет

$$\rho(x_{k+1}, \mathcal{M}^*) < \rho(x_k, \mathcal{M}^*)$$

при достаточно больших k таких, что $\rho(x_k, \mathcal{D}_\varepsilon) > \delta$. Рассуждая теперь так же, как при доказательстве второй части теоремы 4.1, получаем, что $\rho(x_k, S_{a_\varepsilon}(\mathcal{M}^*)) \rightarrow 0$. ■

Итак, существует подпоследовательность $\{k_s\}$ такая, что $x_{k_s} \rightarrow x^*$, где $f(x^*) \leq f^* + \varepsilon$. Далее, для любого $\delta > 0$ найдется $K_\delta < \infty$ такое, что

$$\rho(x_k, \mathcal{M}^*) \leq a_\varepsilon + \delta \quad \forall k \geq K_\delta.$$

Ранее (см. следствие 3 из леммы 9.1 гл. I) уже отмечалось, что $a_\varepsilon \xrightarrow{\varepsilon \rightarrow +0} 0$.

3. ε_k -субградиентный метод. Наконец, опишем еще один метод последовательных приближений, в котором

на каждом шаге используются ε_k -субградиенты, причем $\varepsilon_k \xrightarrow{k \rightarrow \infty} 0$.

Пусть снова последовательность $\{\lambda_k\}$ удовлетворяет условию (4.1). Выберем такую последовательность $\{\varepsilon_k\}$, что

$$\varepsilon_k \rightarrow +0. \quad (4.24)$$

Возьмем произвольную точку $x_0 \in E_n$. Пусть уже найдена точка $x_k \in E_n$. Выберем любое $v_k \in \partial_{\varepsilon_k} f(x_k)$ и положим $x_{k+1} = x_k + z_k$, где

$$z_k = \begin{cases} 0, & \text{если } v_k = 0, \\ -\lambda_k v_k \|v_k\|^{-1}, & \text{если } v_k \neq 0. \end{cases} \quad (4.25)$$

Теорема 4.3. Если множество $\mathcal{M}^*$ непусто и ограничено, то при $k \rightarrow \infty$

$$\rho(x_k, \mathcal{M}^*) \rightarrow 0, \quad f(x_k) \rightarrow f^*. \quad (4.26)$$

Доказательство. Вначале установим, что

$$\lim_{k \rightarrow \infty} \rho(x_k, \mathcal{M}^*) = 0. \quad (4.27)$$

Допустим противное. Тогда найдутся $a > 0$ и $K_1 < \infty$, такие, что

$$\rho(x_k, \mathcal{M}^*) \geq 2a > 0 \quad \forall k \geq K_1. \quad (4.28)$$

Без ограничения общности можно считать, что $v_k \neq 0$, ибо если таких k , для которых $v_k = 0$, бесконечно много, то можно найти подпоследовательность $\{x_{k_s}\}$, для которой $\rho(x_{k_s}, \mathcal{M}^*) \rightarrow 0$. (Пользуемся ограниченностью $\mathcal{D}_{\varepsilon_0}$.)

Как и при доказательстве теоремы 4.1, существует $c > 0$ такое, что

$$\mathcal{D}_c = \{x \mid f(x) \leq f^* + c\} \subset S_a(\mathcal{M}^*).$$

При этом можем считать, что

$$\inf_{x \notin S_{2a}(\mathcal{M}^*)} f(x) - f^* = d > c > 0. \quad (4.29)$$

Существует $r > 0$ такое, что

$$S_r(\bar{x}) = \{x \mid \|x - \bar{x}\| \leq r\} \subset \mathcal{D}_c \quad \forall \bar{x} \in \mathcal{M}^*. \quad (4.30)$$

Зафиксируем любое $\bar{x} \in \mathcal{M}^*$. Тогда

$$f(x) - f(x_k) \geq (v_k, x - x_k) - \varepsilon_k \quad \forall x \in E_n. \quad (4.31)$$

Из (4.28) и (4.30) имеем $f(x_k) > f^* + d$, и потому из (4.31) при достаточно больших k таких, что $e_k < d - c$, будет

$$(v_k, x - x_k) \leq f(x) - f(x_k) + e_k < f^* + c - f^* - d + e_k < 0.$$

Теперь, дословно повторяя доказательство теоремы 4.1, получаем (4.26). ■

4. Субградиентные методы с почти полной релаксацией. Все три метода, описанные в этом параграфе, не являются релаксационными, т. е. значение функции не обязательно убывает при переходе из точки x_k в точку x_{k+1} . Отличительной особенностью субградиентных методов является их чрезвычайная простота, ибо при их использовании на каждом шаге требуется лишь найти какой-нибудь субградиент (или ϵ -субградиент). Однако существенным недостатком этих методов является очень медленная сходимость. Для преодоления этого недостатка можно использовать различные идеи, некоторые из которых будут обсуждаться ниже. Опишем, например, следующий вариант субградиентного метода. Пусть по-прежнему последовательность $\{\lambda_k\}$ удовлетворяет условию (4.1). Последовательность $\{x_k\}$ строим так:

$$x_{k+1} = x_k - \alpha_k v_k \|v_k\|^{-1} \equiv x_k + \alpha_k g_k,$$

где $\alpha_k \in [\beta\lambda_k, \gamma\lambda_k]$, числа $\beta > 0$ и $\gamma > \beta$ фиксированы. Правило выбора α_k из отрезка $[\beta\lambda_k, \gamma\lambda_k]$ может быть любым, например:

1) α_k — произвольное число из $[\beta\lambda_k, \gamma\lambda_k]$;

2) $f(x_k + \alpha_k g_k) = \min_{\alpha \in [\beta\lambda_k, \gamma\lambda_k]} f(x_k + \alpha g_k)$;

3) $f(x_k + \alpha_k g_k) = \min_{i \in 1:N} f(x_k + \xi_{ik} g_k)$,

где

$$\xi_{ik} = \beta_i \lambda_k, \quad \beta \leq \beta_i, \quad \beta_i = 2^{i-1} \beta_1, \quad \dots, \quad \beta_N = 2^{N-1} \beta_1 \leq \gamma.$$

При использовании последних двух правил субградиентный метод становится методом с почти полной релаксацией, если взять маленькое β и большое γ , например: $\beta = 10^{-10}$, $\gamma = 10^{10}$. Сходимость при этом сохра-

няется, ибо последовательность $\{\alpha_k\}$ по-прежнему удовлетворяет условию

$$\alpha_k \rightarrow +0, \quad \sum_{k=0}^{\infty} \alpha_k = +\infty.$$

Аналогичный подход можно применить к ε -субградиентному методу и к ε_k -субградиентному методу.

5. Субградиентный метод с постоянным шагом. Выберем произвольные $x_0 \in E_n$ и $\lambda > 0$. Пусть уже построено x_k . Найдем любое $v_k \in \partial f(x_k)$ и положим

$$x_{k+1} = x_k - \lambda v_k \|v_k\|^{-1}. \quad (4.32)$$

Если оказалось $v_k = 0$, то процесс прекращаем, ибо x_k — точка минимума.

В результате строим последовательность $\{x_k\}$. Если эта последовательность конечна, то по построению последняя полученная точка — точка минимума. Пусть последовательность $\{x_k\}$ бесконечна, а множество $\mathcal{M}^*$ ограничено.

Положим

$$r = 2^{-1}\lambda, \quad S_r(\mathcal{M}^*) = \{x \mid \rho(x, \mathcal{M}^*) \leq r\},$$

$$\max_{x \in S_r(\mathcal{M}^*)} f(x) = f^* + c_r.$$

Заметим, что максимум здесь достигается на границе (см. лемму 9.7 гл. I) и что $c_r > 0$,

$$\mathcal{D}_{c_r} = \{x \mid f(x) \leq f^* + c_r\}, \quad r_1 = \max_{x \in \mathcal{D}_{c_r}} \rho(x, \mathcal{M}^*).$$

Теорема 4.4 Если последовательность $\{x_k\}$ бесконечна, то

$$\rho(x_k, S_{\lambda+r_1}(\mathcal{M}^*)) \xrightarrow{k \rightarrow \infty} 0,$$

при этом

$$\lim_{k \rightarrow \infty} \rho(x_k, \mathcal{D}_{c_r}) = 0. \quad (4.33)$$

Доказательство. Вначале установим (4.33). Допустим противное. Тогда существуют $a > 0$ и $K_1 < \infty$ такие, что

$$\rho(x_k, \mathcal{D}_{c_r}) \geq a \quad \forall k \geq K_1. \quad (4.34)$$

Найдутся $c' > c_r$ и $r' > r$ такие, что (рис. 50)

$$\mathcal{D}_{c'} \equiv \{x \mid f(x) \leq f^* + c'\} \subset S_a(\mathcal{D}_{c_r}), \quad (4.35)$$

$$S_{r'}(\mathcal{M}^*) \subset \mathcal{D}_{c'}.$$

Возьмем любое $\bar{x} \in \mathcal{M}^*$. Ясно, что

$$S_{r'}(\bar{x}) = \{x \mid \|x - \bar{x}\| \leq r'\} \subset \mathcal{D}_{c'}. \quad (4.36)$$

Из определения $\partial f(x_k)$ имеем

$$f(x) - f(x_k) \geq (v_k, x - x_k) \quad \forall x \in E_n. \quad (4.37)$$

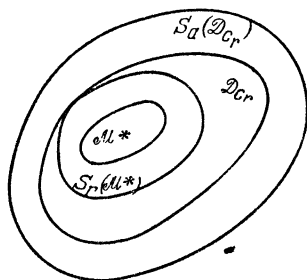


Рис. 50.

Так как из (4.34) и (4.35) $f(x_k) > f^* + c'$, то из (4.37) при $k \geq K_1$ будет

$$(v_k, x - x_k) < 0 \quad \forall x \in \mathcal{D}_{c'}. \quad (4.38)$$

Положим $\tilde{x}_k = \bar{x} + r' v_k \|v_k\|^{-1}$. Ясно, что $\tilde{x}_k \in \mathcal{D}_{c'}$, и поэтому из (4.38)

$$(-\lambda v_k \|v_k\|^{-1}, x_k - \tilde{x}_k) < 0.$$

Отсюда

$$(-\lambda v_k \|v_k\|^{-1}, x_k - \bar{x}) + (\lambda v_k \|v_k\|^{-1}, r' v_k \|v_k\|^{-1}) < 0,$$

т. е.

$$(-\lambda v_k \|v_k\|^{-1}, x_k - \bar{x}) < -\lambda r'.$$

Далее,

$$\begin{aligned} \|x_{k+1} - \bar{x}\|^2 &= \|x_k - \lambda v_k \|v_k\|^{-1} - \bar{x}\|^2 = \\ &= \|x_k - \bar{x}\|^2 + 2(-\lambda v_k \|v_k\|^{-1}, x_k - \bar{x}) + \lambda^2 < \\ &< \|x_k - \bar{x}\|^2 + \lambda[-2r' + \lambda]. \end{aligned} \quad (4.39)$$

Так как $r' > r = \lambda/2$, то из (4.39) следует $\|x_k - \bar{x}\| \rightarrow -\infty$, что невозможно. Полученное противоречие и доказывает (4.33).

Возьмем любое $\delta > 0$. Как и выше, найдется $K_\delta < \infty$ такое, что для $k > K_\delta$ и таких, что $\rho(x_k, \mathcal{D}_{c_r}) > \delta$, будет выполнено (4.39). Из (4.33) следует, что существуют k такие, что $\rho(x_k, \mathcal{D}_{c_r}) \leq \delta$. Поскольку $x_{k+1} = x_k - \lambda v_k \|v_k\|^{-1}$, то $\|x_{k+1} - x_k\| = \lambda$, и поэтому $x_{k+1} \in S_{\lambda+\delta}(\mathcal{D}_{c_r})$. Если при этом $\rho(x_{k+1}, \mathcal{D}_{c_r}) > \delta$, то в силу (4.39) при достаточно больших k (при $k \geq K_\delta$) будет

$$\rho(x_{k+2}, \mathcal{M}^*) < \rho(x_{k+1}, \mathcal{M}^*) \leq \lambda + r_1 + \delta.$$

Если же $\rho(x_{k+1}, \mathcal{D}_{c_r}) \leq \delta$, то снова $\rho(x_{k+2}, \mathcal{D}_{c_r}) \leq \lambda + \delta$, т. е. $v_{k+2} \in S_{\lambda+\delta}(\mathcal{D}_{c_r})$. Далее продолжаем аналогично. В силу произвольности $\delta > 0$ отсюда вытекает

$$\rho(x_k, S_{\lambda+r_1}(\mathcal{M}^*)) \xrightarrow[k \rightarrow \infty]{} 0.$$

Теорема доказана.

З а м е ч а н и е 3. Из формулы (5.26) гл. I следует

$$f(x+z) = f(x) + \int_0^1 \max_{v \in \partial f(x+\tau z)} (v, z) d\tau.$$

Отсюда, если $\|z\| \leq r$, то $c_r \leq rB = \lambda B/2$, где

$$B = \max_{v \in S_r(\mathcal{M}^*)} \max_{v \in \partial f(x)} \|v\|.$$

Очевидно, что за счет λ можно добиться сколь угодно малого c_r .

З а м е ч а н и е 4. С практической точки зрения удобно использовать следующий вариант описанного метода. Пусть $A > 0$ фиксировано. Строим последовательность $\{x_k\}$ по формуле

$$x_{k+1} = \begin{cases} x_k - \lambda v_k \|v_k\|^{-1}, & \text{если } \|v_k\| > A, \\ x_k - \lambda v_k, & \text{если } \|v_k\| \leq A. \end{cases} \quad (4.40)$$

Для гладкого случая (см. § 2) этот метод обеспечивает сходимость последовательности $\{x_k\}$ к точке минимума.

Задание 5.1. Исследовать ε -субградиентный метод с постоянным шагом (последовательность $\{x_k\}$ строится по формуле (4.32) или (4.40), а $v_k \in \partial_\varepsilon f(x_k)$).

6. Рассмотрим снова субградиентный метод, описанный в п. 1. Пусть последовательность $\{\lambda_k\}$ удовлетворяет (4.1) и, кроме того,

$$\sum_{k=0}^{\infty} \lambda_k^2 = A < \infty. \quad (4.41)$$

Теорема 4.5. Если множество $\mathcal{M}^*$ непусто, то последовательность $\{x_k\}$ сходится к точке множества $\mathcal{M}^*$.

Доказательство. Возьмем любое $x^* \in \mathcal{M}^*$. Как и в п. 1, имеем

$$\|x_{k+1} - x^*\|^2 = \|x_k - x^*\|^2 + \frac{2\lambda_k}{\|v_k\|} (v_k, x^* - x_k) + \lambda_k^2. \quad (4.42)$$

Так как $v_k \in \partial f(x_k)$ и $x^* \in \mathcal{M}^*$, то

$$0 \geq f(x^*) - f(x_k) \geq (v_k, x^* - x_k). \quad (4.43)$$

Отсюда

$$(v_k, x^* - x_k) \leq 0 \quad \forall k. \quad (4.44)$$

Поэтому из (4.42)

$$\|x_k - x^*\|^2 \leq \|x_0 - x^*\|^2 + \sum_{i=0}^{k-1} \lambda_i^2.$$

Из (4.41) следует ограниченность последовательности $\{x_k\}$. Значит, и последовательность $\{v_k\}$ тоже ограничена, т. е. существует число $L < \infty$ такое, что

$$\|v_k\| \leq L \quad \forall k. \quad (4.45)$$

Покажем теперь, что существует такая подпоследовательность $\{x_{k_s}\}$, что

$$a_{k_s} = (v_{k_s}, x^* - x_{k_s}) \rightarrow 0. \quad (4.46)$$

Допустим противное. Тогда из (4.44) вытекает, что найдутся $b > 0$ и $K < \infty$ такие, что

$$a_k \leq -b \quad \forall k > K. \quad (4.47)$$

Из (4.42), (4.45) и (4.47) получаем $\|x_k - x^*\|^2 \rightarrow -\infty$, что невозможно. Итак, существует подпоследователь-

ность $\{x_{k_s}\}$, для которой выполнено (4.46). Из (4.43), (4.44) и (4.46) следует

$$f(x_{k_s}) \rightarrow f(x^*). \quad (4.48)$$

Так как последовательность $\{x_k\}$ ограничена, то без ограничения общности можем считать, что $x_{k_s} \rightarrow \bar{x}$. Из (4.48) заключаем, что $\bar{x} \in \mathcal{M}^*$. Покажем теперь, что вся последовательность $\{x_k\}$ сходится к точке $\bar{x}$. В (4.42) и (4.44) можно вместо x^* взять точку $\bar{x}$. Для любого $\varepsilon > 0$ найдется $K_1 < \infty$ такое, что

$$\|x_{k_s} - \bar{x}\|^2 < \frac{\varepsilon}{2}, \quad \sum_{i=k_s}^{\infty} \lambda_i^2 < \frac{\varepsilon}{2} \quad \forall k_s > K_1.$$

Из (4.42) и (4.44) для таких k_s

$$\|x_{k_s+N} - \bar{x}\|^2 \leq \|x_{k_s} - \bar{x}\|^2 + \sum_{i=k_s}^{k_s+N-1} \lambda_i^2 < \varepsilon \quad \forall N \geq 1,$$

а это и означает, что $x_k \rightarrow \bar{x}$.

З а м е ч а н и е 5. Таким образом, при выполнении дополнительного условия (4.41) последовательность $\{x_k\}$, построенная с помощью субградиентного метода, имеет единственную предельную точку, и эта точка — из множества $\mathcal{M}^*$. При этом множество $\mathcal{M}^*$ может быть и неограниченным. Примером последовательности $\{\lambda_k\}$, удовлетворяющей (4.1) и (4.41), может служить последовательность $\left\{\frac{1}{k}\right\}$. Условие (4.41) может быть использовано и в методах пп. 2 и 3 настоящего параграфа.

§ 5. Многошаговый субградиентный метод

При минимизации гладких функций весьма эффективными являются многошаговые методы, т. е. методы, в которых при нахождении очередного $(k+1)$ -го приближения используется информация не только о значении функции и ее производных на данном k -м шаге, но и на одном или нескольких предыдущих шагах. К таким методам относится, например, метод сопряженных градиентов. Можно надеяться, что и в негладких задачах

многошаговые методы будут эффективными. Если в гладких задачах градиенты в близких точках не очень отличаются друг от друга в силу непрерывной дифференцируемости функции, то в негладких выпуклых задачах по какому-нибудь субградиенту в одной точке еще трудно судить обо всем субдифференциале, даже в этой точке. Поэтому знание субградиентов в нескольких соседних точках может дать большую информацию о субдифференциале, а потому и о направлении спуска. Ниже описывается многошаговый субградиентный метод.

1. Пусть $f(x)$ — конечная выпуклая на E_n функция. В § 5 гл. I было показано, что

$$f(x) = \max_{z \in E_n} [f(z) + (v(z), x - z)], \quad (5.1)$$

где $v(z)$ — любой фиксированный элемент из $\partial f(z)$.

Пусть дан набор точек $\tau = \{y_1, \dots, y_p\}$, $y_i \in E_n$, где $p \geq 1$. Положим

$$\varphi(x) = \max_{y \in \tau} [f(y) + (v(y), x - y)].$$

Отсюда

$$\varphi(x_0 + z) = \max_{y \in \tau} [f(y) + (v(y), x_0 - y) + (v(y), z)].$$

Найдем

$$\min_{\|z\|=\alpha} \varphi(x_0 + z) = \varphi(x_0 + z_0(\alpha)).$$

Кривую $\{x | x = x_0 + z_0(\alpha), \alpha \geq 0\}$ назовем *аппроксимирующей кривой* функции $f(x)$, построенной в точке x_0 по множеству τ . Если $p = 1$ и $\tau = \{x_0\}$, то аппроксимирующая кривая — луч $\{x | x = x_0 - \alpha v(x_0), \alpha \geq 0\}$. Опишем следующий метод последовательных приближений. Выберем число $N > n + 2$ и возрастающую последовательность чисел $\{\lambda_k\}$ такую, что

$$\lambda_k \rightarrow +0, \quad \sum_{k=0}^{\infty} \lambda_k = +\infty. \quad (5.2)$$

Возьмем любую точку $x_0 \in E_n$ и зафиксируем произвольное $A > f(x_0)$. Положим $\sigma_0 = \{x_0\}$. Пусть уже найдены точка $x_k \in E_n$ и множество σ_k , причем

$$\sigma_k = \{x_{j_{1k}}, \dots, x_{j_{s_k k}}\}, \quad j_{1k} < j_{2k} < \dots < j_{s_k k}, \quad (5.3)$$

$$j_{1k} \geq k - 1 - N, \quad |\sigma_k| = s_k \leq n + 1.$$

Напомним, что $|\sigma|$ — число элементов в множестве σ . Возьмем произвольное $v(x_k) \in \partial f(x_k)$. Если $v(x_k) = 0$, то x_k — точка минимума функции $f(x)$ на E_n , и процесс прекращается. Рассмотрим случай $v(x_k) \neq 0$. Если $f(x_k) > A$, то полагаем

$$\sigma'_k = \{x_k\}. \quad (5.4)$$

Если же $f(x_k) \leq A$, то возьмем

$$\sigma'_k = \sigma_k \cup \{x_k\}. \quad (5.5)$$

В силу (5.3)

$$|\sigma'_k| \leq n + 2. \quad (5.6)$$

Найдем

$$\min_{\|z\| \leq \lambda_k} \varphi_k(z) = \varphi_k(z_k), \quad (5.7)$$

где

$$\varphi_k(z) = \max_{y \in \sigma'_k} B(x_k + z, y), \quad (5.8)$$

$$B(x, y) = f(y) + (x - y, v(y)).$$

Заметим, что в случае (5.4)

$$z_k = -\lambda_k v_k \|v_k\|^{-1}, \quad (5.9)$$

где $v_k = v(x_k)$. Введем множества

$$R_k = \{y \in \sigma'_k \mid B(x_k + z_k, y) = \varphi_k(z_k)\}, \quad \mathcal{L}_k = \text{co} \bigcup_{y \in R_k} v(y).$$

Так как $R_k \subset \sigma'_k$, то из (5.6) имеем $|R_k| \leq n + 2$. По необходимому и достаточному условию минимума функции $\varphi_k(z)$ на множестве $\{z \mid \|z\|^2 \leq \lambda_k^2\}$ будет (см. § 6 гл. I):

1. Если $\|z_k\| < \lambda_k$, то

$$0 \in \mathcal{L}_k. \quad (5.10)$$

Если $z_k = 0$, то x_k — точка минимума, и процесс прекращается.

2. Если $\|z_k\| = \lambda_k$, то существует $\gamma_k \geq 0$ такое, что

$$-2\gamma_k z_k \in \mathcal{L}_k. \quad (5.11)$$

По теореме Каратеодори (теорема 1.1 гл. I) существует множество $R'_k \subset R_k$ такое, что $|R'_k| \leq n+1$ и

$$0 \in \text{co} \bigcup_{y \in R'_k} v(y) = \mathcal{L}'_k \quad \text{в случае (5.10),}$$

$$-2\gamma_k z_k \in \mathcal{L}'_k \quad \text{в случае (5.11).}$$

В частности, если $|R_k| < n+2$, то можно взять $R'_k = R_k$. Положим теперь

$$\sigma_{k+1} = \begin{cases} R'_k, & \text{если } x_{k-N-1} \notin R'_k, \\ R'_k \setminus \{x_{k-N-1}\}, & \text{если } x_{k-N-1} \in R'_k. \end{cases}$$

Таким образом, базис σ_{k+1} есть некоторое подмножество множества точек $\{x_k, x_{k-1}, \dots, x_{k-N}\}$. Пусть

$$f^* = \inf_{x \in E_n} f(x), \quad \mathcal{M}^* = \{z \in E_n \mid f(z) = f^*\},$$

$$\rho(x, \mathcal{M}^*) = \min_{z \in \mathcal{M}^*} \|z - x\|.$$

Теорема 5.1. Если множество $\mathcal{M}^*$ непусто и ограничено, то

$$\rho(x_k, \mathcal{M}^*) \rightarrow 0, \quad f(x_k) \rightarrow f^*. \quad (5.12)$$

Доказательство. Из ограниченности $\mathcal{M}^*$ вытекает (см. § 9 гл. I) ограниченность множества

$$\mathcal{D}_{A_1} = \{x \in E_n \mid f(x) \leq f^* + A_1\}, \quad \text{где } A_1 = A - f^*.$$

Кроме того (см. леммы 9.1 и 9.2 гл. I), существуют $b > 0$ и $B < \infty$ такие, что

$$\|v\| \geq b \quad \forall v \in \partial f(x), \quad x \notin \mathcal{D}_{A_1}, \quad (5.13)$$

$$\|v\| \leq B \quad \forall v \in \partial f(x), \quad x \in \mathcal{D}_{A_1}.$$

Докажем вначале, что

$$\lim_{k \rightarrow \infty} \rho(x_k, \mathcal{M}^*) = 0. \quad (5.14)$$

Допустим противное. Покажем, что тогда найдется $d > 0$ такое, что

$$\min_{v \in \mathcal{L}'_k} \|v\| = \|v_k\| \geq d > 0 \quad \forall k. \quad (5.15)$$

Действительно, предположим, что это не так. Тогда существует такая подпоследовательность индексов $\{k_s\}$, что

$$\min_{v \in \mathcal{L}_{k_s}} \|v\| = \|v_{k_s}\| \xrightarrow{s \rightarrow \infty} 0, \quad (5.16)$$

где $v_{k_s} \in \mathcal{L}_{k_s}$.

Без ограничения общности можно тогда считать, что $x_{k_s} \rightarrow x^* \in \mathcal{D}_{A_1}$, ибо если допустить противное, то в силу (5.13) получим противоречие с (5.16).

Так как $|R'_{k_s}| \leq n+2$, то можно также считать, что при всех k_s

$$|R'_{k_s}| = m, \quad v_{k_s} = \sum_{i=1:m} \alpha_{k_s i} v_{k_s i}$$

и

$$v_{k_s i} \rightarrow v_i^*, \quad y_{k_s i} \rightarrow y_i^*, \quad \alpha_{k_s i} \rightarrow \alpha_i^* \quad \forall i \in 1:m,$$

где $m \leq n+2$, $\sum_{i=1:m} \alpha_{k_s i} = 1$.

Из полунепрерывности сверху отображения $\partial f(x)$ следует, что $v_i^* \in \partial f(y_i^*)$. Так как $|R'_{k_s}| = m$, то

$$\begin{aligned} \varphi_{k_s}(z_{k_s}) &\equiv \max_{y \in \sigma'_k} B(x_{k_s+1}, y) = \\ &= f(x_{k_s}) + (v_{k_s i}, x_{k_s} - y_{k_s i}) + (v_{k_s i}, z_{k_s}) \quad (5.17) \\ &\quad \forall i \in 1:m. \end{aligned}$$

Поскольку $x_{k_s} \in \sigma'_k$, то из определения $\varphi_{k_s}(z)$ вытекает

$$\varphi_{k_s}(x_{k_s+1}) \geq f(x_{k_s}) + (v(x_{k_s}), z_{k_s}). \quad (5.18)$$

Кроме того,

$$f(x_{k_s} + z_{k_s}) \geq \varphi_{k_s}(z_{k_s}), \quad (5.19)$$

ибо

$$\begin{aligned} f(x+z) &= \max_{y \in E_n} B(x+z, y) \geq \\ &\geq \max_{y \in \sigma_k} B(x+z, y) = \varphi_k(x+z) \quad \forall x, z \in E_n, \quad \forall k. \end{aligned}$$

Переходя в (5.17) — (5.19) к пределу при $k_s \rightarrow \infty$, получим

$$f(x^*) \geq \max_{i \in 1:m} [f(y_i^*) + (v_i^*, x^* - y_i^*)] \geq f(x^*),$$

т. е.

$$f(x^*) = \max_{i \in 1:m} [f(y_i^*) + (v_i^*, x^* - y_i^*)]. \quad (5.20)$$

При этом из (5.16) вытекает

$$0 \in \text{co} \{v_i^* | i \in 1:m\}, \quad (5.21)$$

а из (5.17)

$$f(x^*) = f(y_i^*) + (v_i^*, x^* - y_i^*) \quad \forall i \in 1:m. \quad (5.22)$$

Соотношения (5.21) и (5.22) означают, что функция

$$\varphi^*(x) = \max_{i \in 1:m} [f(y_i^*) + (v_i^*, x - y_i^*)]$$

в точке x^* достигает своего наименьшего на E_n значения, которое в силу (5.20) равно $f(x^*)$. Но

$$f(x) = \max_{y \in E_n} [f(y) + (v(y), x - y)] \geq \varphi^*(x),$$

где $v(y) \in \partial f(y)$. Поэтому

$$f^* = \inf_{x \in E_n} f(x) \geq \min_{x \in E_n} \varphi^*(x) = f(x^*),$$

что и означает справедливость (5.14).

Итак, если имеется такая подпоследовательность индексов $\{k_s\}$, что выполнено (5.16), то выполняется и (5.14).

Таким образом, для некоторого $d > 0$ выполнено (5.15). По предположению (5.14) не имеет места, т. е. существует $a > 0$ такое, что

$$\rho(x_k, \mathcal{M}^*) \geq 2a \quad \forall k. \quad (5.23)$$

Из (5.11) и (5.15) вытекает, что

$$\begin{aligned} z_k &= -\alpha_k \sum_{y \in R_k} \alpha'_k(y) y, \quad \sum_{y \in R_k} \alpha'_k(y) = 1, \\ \alpha_k &> 0, \quad \alpha'_k(y) \geq 0, \quad y \in R_k. \end{aligned} \quad (5.24)$$

Так как $v_k \in \mathcal{P}_k$ и $\|v_k\| \geq d$, а $\|z_k\| = \lambda_k$, то $\alpha_k \leq d^{-1}\lambda_k$.
Отсюда

$$\max_{y \in R_k} \alpha_k(y) \xrightarrow{k \rightarrow \infty} 0, \quad (5.25)$$

где $\alpha_k(y) = \alpha_k \alpha'_k(y)$.

В силу ограниченности $\mathcal{M}^*$ найдется $c > 0$ такое, что

$$\mathcal{D}_c = \{x \mid f(x) \leq f^* + c\} \subset S_a(\mathcal{M}^*),$$

где

$$S_a(\mathcal{M}^*) = \{x \mid \rho(x, \mathcal{M}^*) \leq a\}.$$

Возьмем произвольное $\bar{x} \in \mathcal{M}^*$. Очевидно, что $\bar{x} \in \text{int } \mathcal{D}_c$, т. е. существует $r > 0$ такое, что

$$S_r(\bar{x}) \subset \mathcal{D}_c.$$

Из определения $\partial f(x_k)$ имеем

$$f(x) - f(x_k) \geq (v(x_k), x - x_k) \quad \forall x \in E_n. \quad (5.26)$$

Поскольку в силу (5.23) $f(x_k) > f^* + c$, то из (5.26)

$$(v(x_k), x - x_k) < 0 \quad \forall x \in \mathcal{D}_c. \quad (5.27)$$

Если $f(x_k) > A$, то по построению $z_k = -\lambda_k v(x_k) \|v(x_k)\|^{-1}$, и потому из (5.27)

$$(z_k, x_k - x) < 0 \quad \forall x \in \mathcal{D}_c. \quad (5.28)$$

Рассмотрим случай $f(x_k) \leq A$. Возьмем $y \in \sigma'_k$. Из (5.13) имеем $\|v(y)\| \leq B$. Существует $i \geq k - N$ такое, что $y = x_i$. Так как

$$\begin{aligned} (v(y), y - x_k) &= (v(y), x_i - x_k) = \\ &= (v(y), (x_i - x_{i+1}) + (x_{i+1} - x_{i+2}) + \dots + (x_{k-1} - x_k)), \end{aligned}$$

то

$$(v(y), y - x_k) \leq B \sum_{j=i+1}^k \|x_{j-1} - x_j\| \leq BN\lambda_{kN},$$

где $\lambda_{kN} = \max_{i \in (k-N): k} \lambda_i$. Заметим, что $\lambda_{kN} \rightarrow 0$. Отсюда и из (5.27) для $y \in \sigma'_k$ имеем

$$\begin{aligned} (v(y), x - x_k) &= \\ &= (v(y), x - y) + (v(y), y - x_k) \leq BN\lambda_{kN} \quad \forall x \in \mathcal{D}_c. \end{aligned} \quad (5.29)$$

Из (5.24) и (5.29)

$$\begin{aligned} (z_k, x_k - x) &= \\ &= \sum_{y \in \sigma_k} \alpha_k(y) (v(y), x - x_k) \leq BN \lambda_{kN} \alpha_k = \lambda_{kN} \gamma_k \quad (5.30) \\ &\quad \forall x \in \mathcal{D}_c. \end{aligned}$$

Здесь $\gamma_k = BN \alpha_k$. Из (5.24) получаем $\gamma_k \rightarrow 0$.

Неравенство (5.30) установлено для случая $f(x_k) \leq A$. Для случая $f(x_k) > A$ имеет место (5.28). Поэтому можно считать, что неравенство (5.30) справедливо для всех k .

Возьмем $\tilde{x}_k = \bar{x} - rz_k \|z_k\|^{-1}$. Тогда из (5.30)

$$(z_k, x_k - \bar{x} + rz_k \|z_k\|^{-1}) < \gamma_k \lambda_{kN},$$

т. е.

$$(z_k, x_k - \bar{x}) \leq [-r\lambda_k + \gamma_k \lambda_{kN}].$$

Но

$$\begin{aligned} \|x_{k+1} - \bar{x}\|^2 &= \|x_k + z_k - \bar{x}\|^2 = \\ &= \|x_k - \bar{x}\|^2 + 2(z_k, x_k - \bar{x}) + z_k^2 \leq \\ &\leq \|x_k - \bar{x}\|^2 + [-2r\lambda_k + 2\gamma_k \lambda_{kN}] + \lambda_k^2. \end{aligned}$$

Поскольку $\lambda_k \rightarrow +0$ и $\gamma_k \rightarrow 0$, то при достаточно больших k

$$\|x_{k+1} - \bar{x}\|^2 \leq \|x_k - \bar{x}\|^2 - r\lambda_k.$$

Продолжая аналогично, получим

$$\|x_{k+s} - \bar{x}\|^2 \leq \|x_k - \bar{x}\|^2 - r \sum_{i=0}^{s-1} \lambda_{k+i}.$$

В силу (5.2) $\|x_k - \bar{x}\| \rightarrow -\infty$, что невозможно. Итак, (5.14) установлено.

Возьмем произвольное $\delta > 0$. Рассуждая, как и выше, заключаем, что найдется такое $r_\delta > 0$, что для всех $\bar{x} \in \mathcal{M}^*$ будет

$$\|x_{k+1} - \bar{x}\|^2 \leq \|x_k - \bar{x}\|^2 - r_\delta \lambda_k$$

при всех достаточно больших k и таких, что $\rho(x_k, \mathcal{M}^*) > \delta$. Отсюда

$$\rho^2(x_{k+1}, \mathcal{M}^*) \leq \rho^2(x_k, \mathcal{M}^*) - r_\delta \lambda_k. \quad (5.31)$$

В силу (5.14) существуют k такие, что $\rho(x_k, \mathcal{M}^*) < \delta$. Так как $x_{k+1} = x_k + z_k$, $\|z_k\| \leq \lambda_k$, то можно считать, что

$\rho(x_{k+1}, \mathcal{M}^*) < 2\delta$. Если оказалось при этом $\rho(x_{k+1}, \mathcal{M}^*) > \delta$, то в силу (5.31)

$$\rho(x_{k+2}, \mathcal{M}^*) < \rho(x_{k+1}, \mathcal{M}^*) < 2\delta.$$

Далее рассуждаем аналогично. Итак, если $\rho(x_{k+i}, \mathcal{M}^*) < \delta$, то $\rho(x_{k+i+1}, \mathcal{M}^*) < 2\delta$; если же $\delta < \rho(x_{k+i}, \mathcal{M}^*) < 2\delta$, то

$$\rho(x_{k+i+1}, \mathcal{M}^*) < \rho(x_{k+i}, \mathcal{M}^*) < 2\delta.$$

Окончательно получаем, что $\rho(x_{k+i}, \mathcal{M}^*) < 2\delta$ при $i \geq 1$. В силу произвольности δ отсюда следует, что $\rho(x_k, \mathcal{M}^*) \rightarrow 0$, а тогда и $f(x_k) \rightarrow f^*$, чем и завершается доказательство теоремы.

Следствие. Как и при доказательстве следствия из теоремы 4.1, можно показать, что *если $\text{int } \mathcal{M}^* \neq \emptyset$, то последовательность $\{x_k\}$ конечна.*

Замечание. Задача (5.7) является задачей линейного программирования, если вместо евклидовой нормы в (5.7) взять m -норму: $\|z\| = \max_{l=1:n} |z^{(l)}|$. Здесь $z = (z^{(1)}, \dots, z^{(n)})$.

2. Многошаговый ε -субградиентный метод. Как и в п. 2 § 4, в описанном выше методе можно выбирать $v(x_k) \in \partial_\varepsilon f(x_k)$; где $\varepsilon > 0$ фиксировано. В результате строим последовательность $\{x_k\}$. Если эта последовательность конечна, то по построению последняя полученная точка является ε -стационарной, ибо по построению процесс прекращается, лишь когда $v(x_k) = 0$, т. е. $0 \in \partial_\varepsilon f(x_k)$. Если же последовательность $\{x_k\}$ бесконечна, то справедлива

Теорема 5.2. *Если множество $\mathcal{M}^*$ непусто и ограничено, то*

$$\lim_{k \rightarrow \infty} \rho(x_k, \mathcal{D}_\varepsilon) \rightarrow 0, \quad \rho(x_k, S_{a_\varepsilon}(\mathcal{M}^*)) \xrightarrow{k \rightarrow \infty} 0,$$

где

$$\mathcal{D}_\varepsilon = \{x \in E_n \mid f(x) \leq f^* + \varepsilon\},$$

$$a_\varepsilon = \max_{x \in Q_\varepsilon} \rho(x, \mathcal{M}^*),$$

$$Q_\varepsilon = \{x \mid f(x) = f^* + \varepsilon\}.$$

Доказательство аналогично доказательству теоремы 4.2.

§ 6. Релаксационный субградиентный метод

Метод, излагаемый в этом параграфе, идейно близок, с одной стороны, к методу ε -наискорейшего спуска для минимизации произвольной выпуклой функции $f(x)$, а с другой — к методу сопряженных градиентов для минимизации непрерывно дифференцируемой выпуклой функции. Причем если в методе ε -наискорейшего спуска для нахождения направления движения требуется знать все множество $\partial_\varepsilon f(x)$, то в релаксационном субградиентном методе для построения направления спуска используется некоторое конечное число точек из $\partial_\varepsilon f(x)$, аппроксимирующих это множество.

1. Принципиальная схема. В этом пункте предполагается, что функция $f(x)$ строго выпукла на E_n и что могут быть точно выполнены операции вычисления самой функции $f(x)$, ее субградиентов и операция нахождения минимума функции $f(x)$ на луче.

Положим $d(x) = \min_{v \in \partial f(x)} \|v\|$. Сначала опишем релаксационный субградиентный метод для нахождения таких точек $x_0 \in E_n$, что $d(x_0) \leq \varepsilon$, где $\varepsilon > 0$, а потом покажем, как его применять для нахождения точек минимума функции $f(x)$ на E_n .

Зафиксируем положительные числа ε_0 , δ и целое число $m_0 \geq n$.

Выберем произвольно начальное приближение $x_0 \in E_n$. Предположим, что множество $\mathcal{D}(x_0) = \{x \in E_n | f(x) \leq f(x_0)\}$ ограничено.

Пусть уже найдено k -е приближение $x_k \in E_n$. Переход к $(k+1)$ -му приближению x_{k+1} , который назовем k -м циклом, осуществляется следующим образом.

Положим $x_{k0} = x_k$ и вычислим произвольный вектор $v_{k0} \in \partial f(x_{k0})$. Если $\|\bar{v}_{k,-1}\| \leq \varepsilon_0$, где $\bar{v}_{k,-1} = v_{k,0}$, то полагаем $t_k = 0$, $x_{k+1} = x_{kt_k}$, и на этом k -й цикл заканчивается (через t_k будем обозначать число шагов на k -м цикле). Если $\|\bar{v}_{k,-1}\| > \varepsilon_0$, то на луче $\{x_{k0}(\alpha) = x_{k0} - \alpha v_{k0} | \alpha \geq 0\}$ находим точку $x_{k1} = x_{k0}(\alpha_{k0})$ такую, что

$$f(x_{k1}) = \min_{\alpha \geq 0} f(x_{k0}(\alpha)).$$

Обратим внимание на то, что возможен случай, когда $x_{k1} = x_{k0}$. Это будет, если $-v_{k0}$ не является направлением спуска $f(x)$ в x_{k0} .

Если $\|x_{k1} - x_{k0}\| \geq \delta$ или $f(x_{k0}) - f(x_{k1}) \geq \delta$, то полагаем $t_k = 1$, $x_{k+1} = x_{kt_k}$, и на этом k -й цикл заканчивается. В противном случае вычислим такой вектор $\bar{v}_{k0} \in \partial f(x_{k1})$, что

$$(v_{k0}, \bar{v}_{k0}) \leq 0.$$

Такой вектор $\bar{v}_{k0} \in \partial f(x_{k1})$ обязательно найдется, ибо, в силу необходимого условия минимума функции $f(x_{k0}(\alpha))$ по α , должно быть

$$\frac{\partial f(x_{k1})}{\partial (-v_{k0})} = \max_{v \in \partial f(x_{k1})} (v, -v_{k0}) \geq 0.$$

Если $\|\bar{v}_{k0}\| \leq \varepsilon_0$, то полагаем $t_k = 1$, $x_{k+1} = x_{kt_k}$, и на этом k -й цикл заканчивается. Если же $\|\bar{v}_{k0}\| > \varepsilon_0$, то найдем такой вектор v_{k1} , что

$$v_{k1} = \beta_{k0} \bar{v}_{k0} + (1 - \beta_{k0}) v_{k0}, \quad \beta_{k0} \in (0, 1),$$

и

$$\|v_{k1}\| = \min_{\beta \in [0, 1]} \|\beta \bar{v}_{k0} + (1 - \beta) v_{k0}\|,$$

т.е. v_{k1} есть ближайшая к началу координат точка отрезка $[v_{k0}, \bar{v}_{k0}]$.

Нетрудно убедиться в том, что

$$v_{k1} = v_{k0} - (v_{k0}, \bar{v}_{k0} - v_{k0}) \|\bar{v}_{k0} - v_{k0}\|^{-2} (\bar{v}_{k0} - v_{k0}).$$

Теперь на луче $\{x_{k1}(\alpha) = x_{k1} - \alpha v_{k1} | \alpha \geq 0\}$ найдем такую точку $x_{k2} = x_{k1}(\alpha_{k1})$, что

$$f(x_{k2}) = \min_{\alpha \geq 0} f(x_{k1}(\alpha)).$$

Далее продолжаем аналогично до тех пор, пока не будет сделано такое число шагов t_k , что выполнится по крайней мере одно из следующих условий:

$$\begin{aligned} \text{a)} \quad & t_k = m_0; \\ \text{b)} \quad & \|x_{k, t_k-1} - x_{kt_k}\| \geq \delta; \end{aligned} \tag{6.1}$$

$$\text{c)} \quad f(x_{k, t_k-1}) - f(x_{kt_k}) \geq \delta; \tag{6.2}$$

$$\text{d)} \quad \|\bar{v}_{k, t_k-1}\| \leq \varepsilon_0.$$

Теперь за $(k+1)$ -приближение x_{k+1} принимаем точку x_{kt_k} . На этом k -й цикл заканчивается, причем если выполнено условие d), то заканчивается и весь процесс последовательных приближений, ибо $d(x_{k+1}) \leq \varepsilon_0$.

Заметим, что по построению

$$(v_{kt}, \bar{v}_{kt}) \leq 0 \quad \forall t \in 0:l_k, \quad (6.3)$$

где $l_k = t_k - 1$.

Доказательство сходимости релаксационного субградиентного метода опирается на неравенства (6.3) и следующую лемму.

Лемма 6.1. Пусть функция $f(x)$ строго выпукла на E_n , множество $\mathcal{D}(x_0)$ ограничено, а последовательность $\{x_k\}_{k=0}^\infty$ такова, что

$$f(x_{k+1}) = \min_{\alpha \in [0, 1]} f(x_k + \alpha(x_{k+1} - x_k)). \quad (6.4)$$

Тогда $\lim_{k \rightarrow \infty} \|x_{k+1} - x_k\| = 0$.

Доказательство. Допустим противное. Тогда найдутся такие число $h > 0$ и подпоследовательность $\{x_{k_s}\}$, что

$$\|x_{k_s+1} - x_{k_s}\| \geq h \quad \forall s.$$

Так как $f(x_{k+1}) \leq f(x_k)$, то $x_k \in \mathcal{D}(x_0)$ при каждом k . В силу ограниченности $\mathcal{D}(x_0)$, не умаяя общности, можно считать, что $x_{k_s} \rightarrow \bar{x}$, $x_{k_s+1} \rightarrow \bar{x}$, причем $\|\bar{x} - \bar{x}\| \geq h$. Поскольку $f(x_{k+1}) \leq f(x_k)$ и $f(x)$ ограничена снизу, то

$$f(\bar{x}) = f(\bar{x}). \quad (6.5)$$

Так как функция $f(x)$ строго выпукла, то

$$f(2^{-1}\bar{x} + 2^{-1}\bar{x}) < f(\bar{x}).$$

В силу (6.4) имеем

$$\begin{aligned} f(x_{k_s+1}) &= \min_{\alpha \in [0, 1]} f(x_{k_s} + \alpha(x_{k_s+1} - x_{k_s})) \leq \\ &\leq f(2^{-1}x_{k_s} + 2^{-1}x_{k_s+1}). \end{aligned}$$

Переходя в этом неравенстве к пределу, получаем

$$f(\bar{x}) \leq f(2^{-1}\bar{x} + 2^{-1}\bar{x}) < f(\bar{x}),$$

а это противоречит (6.5). Лемма доказана.

Отметим, что лемма 6.1 справедлива и для непрерывной строго квазивыпуклой функции $f(x)$.

Перейдем теперь к обоснованию сходимости релаксационного субградиентного метода. Если последовательность $\{x_k\}$ конечна, то для последней полученной точки x_k по построению будем иметь $d(x_k) \leq \varepsilon_0$. В противном случае справедлива

Теорема 6.1. Пусть функция $f(x)$ строго выпукла на E_n , множество $\mathcal{D}(x_0)$ ограничено.

Тогда, если x^* — предельная точка последовательности $\{x_k\}$, то

$$d(x^*) \leq \max\{\varepsilon_0, bq^{m_0-1}\} = d_0,$$

где $b = \max_{x \in \mathcal{D}(x_0)} \max_{v \in \partial f(x)} \|v\|$, $q = [1 + (\varepsilon_0 b^{-1})^2]^{-1/2}$.

В частности, если $m_0 + 1 > [\ln q]^{-1} \ln(\varepsilon_0 b^{-1})$, то $d(x^*) \leq \varepsilon_0$.

Доказательство. Существование предельных точек последовательности $\{x_k\}$ следует из ограниченности $\mathcal{D}(x_0)$ и того, что $x_k \in \mathcal{D}(x_0)$. Допустим, что утверждение теоремы не верно: предположим, что $x_{k_s} \rightarrow x^*$, но $d(x^*) = \bar{d} > d_0$.

Так как функция $f(x)$ строго выпукла, множество $\mathcal{D}(x_0)$ ограничено и на каждом шаге выполнено (6.4), то по лемме 6.1 при достаточно больших k окончание k -го цикла невозможно вследствие (6.1), а в силу равномерной непрерывности функции $f(x)$ на $\mathcal{D}(x_0)$, и вследствие (6.2). Поэтому при достаточно больших k , скажем при $k \geq K$, будет

$$t_k = m_0. \quad (6.6)$$

Положим $\varepsilon = 2^{-1}(\bar{d} - d_0)$. Поскольку $d(x^*) = \bar{d} > d_0$, то, в силу полунепрерывности сверху точечно-множественного отображения $\partial f(x)$, найдется такое $\delta > 0$, что

$$\partial f(x) \subset S_\varepsilon(\partial f(x^*)) \quad \forall x \in S_\delta(x^*), \quad (6.7)$$

т. е. при x , достаточно близких к x^* , все множества $\partial f(x)$ будут содержаться в некоторой выпуклой окрестности множества $\partial f(x^*)$, расстояние от которой до начала координат больше d_0 (на рис. 51 эта окрестность показана пунктиром). Поскольку эта окрестность выпукла, то вместе с любыми двумя точками на этой окрестности ей будет, очевидно, принадлежать и весь отрезок, соединяющий эти точки.

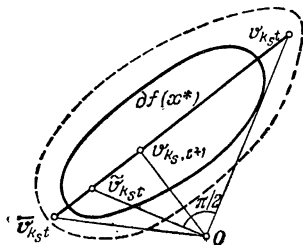


Рис. 51.

Таким образом, из (6.6) и (6.7) следует, что при достаточно больших k_s будем иметь

$$\|v_{k_s t}\| > d_0 \geq \varepsilon_0 \quad \forall t \in 0 : (m_0 - 1). \quad (6.8)$$

Из (6.8) и определения числа b вытекает, что (рис. 51)

$$\|v_{k_s, t+1}\| = \|v_{k_s t}\| \frac{\|\tilde{v}_{k_s t}\|}{\sqrt{\|\tilde{v}_{k_s t}\|^2 + \|v_{k_s t}\|^2}} \leq \|v_{k_s t}\| (\sqrt{1 + (\varepsilon_0 b^{-1})^2})^{-1},$$

где $\tilde{v}_{k_s t}$ — точка, лежащая на отрезке $[\tilde{v}_{k_s t}, v_{k_s t}]$, такая, что угол между $\tilde{v}_{k_s t}$ и $v_{k_s t}$ равен $\pi/2$. Поэтому

$$\|v_{k_s, t+1}\| \leq \|v_{k_s 0}\| q^{t+1} \leq b q^{t+1} \quad \forall t \in 0 : (m_0 - 1).$$

Отсюда и из (6.6) заключаем, что

$$\|v_{k_s, m_0-1}\| = \|v_{k_s l_{k_s}}\| \leq b q^{m_0-1}; \quad (6.9)$$

в частности, если $m_0 > m$, то $\|v_{k_s, m_0-1}\| \leq \varepsilon_0$, где $m = 1 + \ln(\varepsilon_0 b^{-1}) (\ln q)^{-1}$. Но (6.9) противоречит (6.8). Теорема доказана.

З а м е ч а н и е 1. Проверка условий (6.1) и (6.2) производится для того, чтобы избежать лишних вычислений в начале процесса последовательных приближений.

З а м е ч а н и е 2. Если оказалось, что $x_{kt} = x_{k0}$ при всех $t \in 1:t_k$ и $t_k = m_0$, то, как следует из доказательства теоремы 6.1, $d(x_k) \leq d_0$, и в таком случае процесс последовательных приближений можно закончить.

Назовем k -й цикл *полным*, если $t_k = m_0$. Из леммы 6.1 вытекает, что при достаточно больших k будет $t_k = m_0$, т. е. все циклы будут полными, и, следовательно, k -й цикл будет заключаться в построении точек $x_{k0}, x_{k1}, \dots, x_{km_0}$, субградиентов $\bar{v}_k, -1, \bar{v}_{k0}, \dots, \bar{v}_k, m_0-2, \bar{v}_{kt} \in \partial f(x_{k, t+1})$ и векторов $v_{k1}, v_{k2}, \dots, v_k, m_0-1$ таких, что при каждом $t \in 0:(m_0-2)$ вектор $v_{k, t+1}$ есть ближайшая к началу координат точка отрезка $[v_{kt}, \bar{v}_{kt}]$, причем $(v_{kt}, \bar{v}_{kt}) \leq 0$ и при каждом $t \in 0:(m_0-1)$ должно быть $f(x_{k, t+1}) = \min_{\alpha \geq 0} f(x_{kt}(\alpha))$.

С помощью релаксационного субградиентного метода строим последовательность $\{x_k\}$. Для любой ее предельной точки x^* по теореме 6.1 будет $d(x^*) \leq d_0$ и согласно лемме 1.2

$$0 \leq f(x^*) - f^* \leq \mathcal{D}d_0, \quad (6.10)$$

где $\mathcal{D}$ — диаметр множества $\mathcal{D}(x_0)$, $f^* = \inf_{x \in E_n} f(x)$.

Из (6.10) следует, что для любого $\bar{d} > d_0$ за конечное число шагов будет получена точка x_k такая, что

$$0 \leq f(x_k) - f^* \leq \mathcal{D}\bar{d}.$$

Изложенный метод можно применить и для нахождения точек минимума функции $f(x)$ на E_n . Схема метода при этом сохраняется. Изменяются лишь условия а) и d) на следующие:

$$\text{а) } t_k = m_k, \quad \text{d) } \|v_{kt_k}\| \leq \varepsilon_k,$$

причем переход к $(k+1)$ -му циклу происходит после выполнения любого из условий а) — d).

Теорема 6.2. Пусть функция $f(x)$ строго выпукла, множество $\mathcal{D}(x_0)$ ограничено и $\varepsilon_k \rightarrow 0$, $m_k \rightarrow \infty$.

Тогда любая предельная точка последовательности $\{x_k\}$ является точкой минимума функции $f(x)$ на E_n .

Доказательство. Допустим противное: предположим, что $x_{k_s} \rightarrow x^*$ и для некоторого $d_0 > 0$ оказалось $d(x^*) > d_0$.

Так как отображение $\partial f(x)$ полунепрерывно сверху, то при достаточно больших k_s все множества $\partial f(x_{k_s})$ будут содержаться в некоторой выпуклой окрестности множества $\partial f(x^*)$, удаленной от начала координат больше чем на d_0 .

Рассмотрим теперь такие k_s , что $e_{k_s} < d_0$, $t_{k_s} = m_{k_s}$ (см. аналогичную часть доказательства теоремы 6.1) и

$$m_{k_s} > 1 + (\ln q)^{-1} \ln(d_0 b^{-1}),$$

где

$$b = \max_{x \in \mathcal{D}(x_0)} \max_{v \in \partial f(x)} \|v\|, \quad q = (1 + d_0 b^{-1})^{-1/2}.$$

Тогда при достаточно больших k_s , с одной стороны,

$$\|v_{k_s t}\| > d_0 \quad \forall t \in 0 : (m_{k_s} - 1),$$

а с другой (см. доказательство теоремы 6.1),

$$\|v_{k_s, m_{k_s}-1}\| \leq b q^{m_{k_s}-1} \leq d_0.$$

Получили противоречие. Теорема доказана.

2. Связь релаксационного субградиентного метода с методом сопряженных градиентов. В этом пункте будем считать, что длина всех циклов релаксационного субградиентного метода одинакова и равна $m \geq n$, т. е. условия b), c), d) не проверяются (это будет, если $\varepsilon_0 = 0$, $\delta = \infty$).

Кратко релаксационный субградиентный метод можно описать следующим образом:

$$x_0 \in E_n, \quad v_0 \in \partial f(x_0), \quad \beta_0 = 1, \quad p_0 = v_0$$

$$x_{k+1} = x_k - \alpha_k p_k,$$

$$f(x_{k+1}) = \min_{\alpha \geq 0} f(x_k - \alpha p_k),$$

$$p_k = p_{k-1} + \beta_k (v_k - p_{k-1}), \quad v_k \in \partial f(x_k), \quad (6.11)$$

$$(v_k, p_{k-1}) \leq 0,$$

$$\beta_k = \begin{cases} 1, & \text{если } k \text{ кратно } m, \\ -(p_{k-1}, v_k - p_{k-1}) \|v_k - p_{k-1}\|^{-2} & \text{в противном случае.} \end{cases}$$

Нетрудно заметить, что такая запись релаксационного субградиентного метода формально напоминает метод сопряженных градиентов. И на самом деле между этими двумя методами имеется тесная связь: релаксационный субградиентный метод превращается в метод сопряженных градиентов, если он применяется для минимизации гладкой функции.

Методы сопряженных градиентов с восстановлением, так же как и релаксационный субградиентный метод, носят циклический характер (длина цикла — число m' — есть число шагов, через которое производится спуск в направлении антиградиента функции $f(x)$, $m' \geq n$).

Напомним, что метод сопряженных градиентов Флетчера — Ривса для минимизации непрерывно дифференцируемой на E_n функции $f(x)$ заключается в построении последовательности $\{x'_k\}$ по следующим правилам:

$$x'_0 \in E_n, \quad v'_0 = f'(x'_0), \quad \beta'_0 = 0, \quad p'_0 = v'_0$$

$$x'_{k+1} = x'_k - \alpha'_k p'_k,$$

$$f(x'_{k+1}) = \min_{\alpha \geq 0} f(x'_k - \alpha p'_k),$$

$$p'_k = v'_k + \beta'_k p'_{k-1},$$

$$v'_k = f'(x'_k) = \left(\frac{\partial f(x'_k)}{\partial x^{(1)}}, \dots, \frac{\partial f(x'_k)}{\partial x^{(n)}} \right),$$

$$\beta'_k = \begin{cases} 0, & \text{если } k \text{ кратно } m', \\ \|v'_k\|^2 \cdot \|v'_{k-1}\|^{-2} & \text{в противном случае.} \end{cases}$$

Теорема 6.3. Пусть множество $\mathcal{D}(x_0) = \{x \in E_n | f(x) \leq f(x_0)\}$ ограничено и функция $f(x)$ выпукла и непрерывно дифференцируема на E_n .

Тогда, если релаксационный субградиентный метод и метод сопряженных градиентов Флетчера — Ривса начаты из одной и той же точки ($x_0 = x'_0$), длина циклов в обоих методах одинакова ($m = m'$), то при любом k

$$x_k = x'_k, \quad (6.12)$$

$$p_k = \beta_k p'_k, \quad (6.13)$$

т. е. оба метода генерируют одни и те же последовательные приближения.

Доказательство проведем по индукции. При $k=0$ равенства (6.12), (6.13) справедливы по условию теоремы и построению методов, ибо для выпуклой дифференцируемой функции $f(x)$ имеем $\partial f(x) = \{f'(x)\}$.

Предположим, что равенства (6.12), (6.13) выполнены при любом $k \in 0:l$, где $l \geq 0$, и покажем, что они выполняются и при $k = l+1$.

Поскольку $x_l = x'_l$ и $p_l = \beta_l p'_l$, то

$$\{x_l - \alpha p_l \mid \alpha \geq 0\} = \{x'_l - \alpha p'_l \mid \alpha \geq 0\}$$

и, следовательно, $x_{l+1} = x'_{l+1}$ в силу условий выбора точек x_{l+1} и x'_{l+1} как доставляющих минимум функции $f(x)$ на лучах $\{x_l - \alpha p_l \mid \alpha \geq 0\}$ и $\{x'_l - \alpha p'_l \mid \alpha \geq 0\}$ соответственно.

Если $l+1$ кратно m , то $\beta_{l+1} = 1$, $\beta'_{l+1} = 0$ и $p_{l+1} = \beta_{l+1} p'_{l+1} = f'(x_{l+1})$.

Рассмотрим случай, когда $l+1$ не кратно m . Если функция $f(x)$ выпуклая и гладкая, то для направлений в релаксационном субградиентном методе вместо (6.11) будет

$$(v_k, p_{k-1}) = 0 \quad \forall k, \quad (6.14)$$

и, кроме того, если k не кратно m , то

$$p_k = \|p_k\|^2 \|v_k\|^{-2} (v_k + \|v_k\|^2 \|p_{k-1}\|^{-2} p_{k-1}). \quad (6.15)$$

Действительно, поскольку

$$p_k = \beta_k v_k + (1 - \beta_k) p_{k-1}, \quad (p_k, p_{k-1}) = (p_k, v_k) = \|p_k\|^2, \quad (6.16)$$

то в силу (6.14) имеем

$$\begin{aligned} (p_k, v_k) &= \beta_k (v_k, v_k) + (1 - \beta_k) (p_{k-1}, v_k) = \beta_k \|v_k\|^2, \\ (p_k, p_{k-1}) &= \beta_k (v_k, p_{k-1}) + (1 - \beta_k) (p_{k-1}, p_{k-1}) = \\ &= (1 - \beta_k) \|p_{k-1}\|^2, \end{aligned}$$

откуда

$$\beta_k = \|p_k\|^2 \|v_k\|^{-2}, \quad 1 - \beta_k = \|p_k\|^2 \|p_{k-1}\|^{-2}.$$

Теперь (6.15) очевидным образом следует из полученного представления чисел β_k и $1 - \beta_k$ и (6.16).

Воспользуемся равенством (6.15) при $k = l + 1$ и тем, что

$$p_l = \|p_l\|^2 \|v_l\|^{-2} p'_l, \quad v_{l+1} = v'_{l+1}, \quad v_l = v'_l.$$

Получим

$$\begin{aligned} p_{l+1} &= \|p_{l+1}\|^2 \|v_{l+1}\|^{-2} (v_{l+1} + \|v_{l+1}\|^2 \|p_l\|^{-2} p_l) = \\ &= \|p_{l+1}\|^2 \|v_{l+1}\|^{-2} (v_{l+1} + \|v_{l+1}\|^2 \|p_l\|^{-2} \|p_l\|^2 \|v_l\|^{-2} p'_l) = \\ &= \beta_{l+1} (v'_{l+1} + \|v'_{l+1}\|^2 \|v'_l\|^{-2} p'_l) = \beta_{l+1} p'_{l+1}. \end{aligned}$$

Итак, $p_{l+1} = \beta_{l+1} p'_{l+1}$ (рис. 52). Теорема доказана.

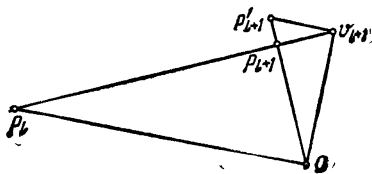


Рис. 52.

З а м е ч а н и е 3. Прерывание цикла по выполнении условий, эквивалентных условиям (6.1), (6.2), полезно и при применении метода сопряженных градиентов, ибо далеко от минимума градиентный метод работает достаточно хорошо.

3. В п. 1 был рассмотрен релаксационный субградиентный метод с полной релаксацией, т. е. метод, в котором на каждом шаге минимум в направлении движения отыскивается точно. Понятно, что в большинстве случаев осуществление полной релаксации невозможно, ибо для этого требуется вспомогательная бесконечная итерационная процедура. Поэтому ближайшей нашей целью будет построение такого варианта релаксационного субградиентного метода, в котором не требуется точного решения задач одномерной минимизации. Зарядимся и от ограничительного условия строгой выпуклости целевой функции $f(x)$.

При доказательстве теоремы о сходимости релаксационного субградиентного метода существенную роль

играла лемма 6.1; поэтому первое, что необходимо сделать, — это получить некоторый аналог леммы 6.1.

Лемма 6.2. Пусть функция $f(x)$ определена на E_n и последовательность $\{x_k\}$ такова, что

$$x_{k+1} = x_k + \alpha_k g_k, \quad f(x_{k+1}) \leq f(x_k) - \delta_2 \alpha_k \|v_k\|, \\ g_k = -v_k \|v_k\|^{-1}, \quad \|v_k\| \geq \varepsilon_0,$$

где $\varepsilon_0 > 0$, $\delta_2 > 0$ — фиксированные числа.

Если последовательность $\{f(x_k)\}$ ограничена снизу, то последовательность $\{x_k\}$ — сходящаяся.

Доказательство. В силу условий леммы имеем

$$f(x_{k+1}) \leq f(x_0) - \delta_2 \varepsilon_0 \sum_{i=1}^k \|x_{i-1} - x_i\|,$$

откуда

$$\sum_{i=1}^k \|x_{i-1} - x_i\| \leq (\varepsilon_0 \delta_2)^{-1} (f(x_0) - \inf_k f(x_k)) < \infty \quad \forall k,$$

а это означает, что последовательность $\{x_k\}$ сходится.

Таким образом, установлено, как выбирать длину шага так, чтобы норма разности между соседними приближениями стремилась к нулю, более того, так, чтобы последовательность $\{x_k\}$ сходилась. ■

Другим важным требованием, обеспечивающим сходимость релаксационного субградиентного метода, являлось получение векторов v_{kt} , $\bar{v}_{kt}$, удовлетворяющих (6.3), т. е. таких, что угол между ними не меньше прямого. Если последовательность $\{x_k\}$ строить так, как это требуется в лемме 6.2, то точка x_{k+1} не обязательно будет близка к точке, в которой функция $f(x)$ достигает минимума на луче $\{x_k + \alpha g_k | \alpha \geq 0\}$. Поэтому необходимо ослабить условие (6.3).

Лемма 6.3. Пусть $f(x)$ — выпуклая на E_n функция. Если при некоторых $\alpha > 0$ и $\delta_1 > 0$

$$f(x + \alpha g) \geq f(x) - \delta_1 \alpha \|v\|, \quad (6.17)$$

где $g = -v \|v\|^{-1}$, $v \in E_n$, то

$$(v, \bar{v}) \leq \delta_1 \|v\|^2 \quad \forall \bar{v} \in \partial f(x + \alpha g). \quad (6.18)$$

Доказательство. Действительно, для любого $\bar{v} \in \partial f(x + \alpha g)$ имеем с учетом (6.17)

$$f(x) \geq f(x + \alpha g) - \alpha (g, \bar{v}) \geq f(x) - \delta_1 \alpha \|v\| - \alpha (g, \bar{v}),$$

т. е. $-(g, \bar{v}) \leq \delta_1 \|v\|$, откуда следует (6.18). ■

Замечание 4. Совершенно аналогично показывается, что если $f(x + \alpha g) \geq f(x)$, то $(v, \bar{v}) \leq 0 \quad \forall \bar{v} \in \partial f(x + \alpha g)$.

Перейдем теперь к описанию релаксационного субградиентного метода для минимизации выпуклой (не обязательно строго) функции $f(x)$ на E_n , в котором вспомогательные задачи одномерной минимизации можно решать приближенно.

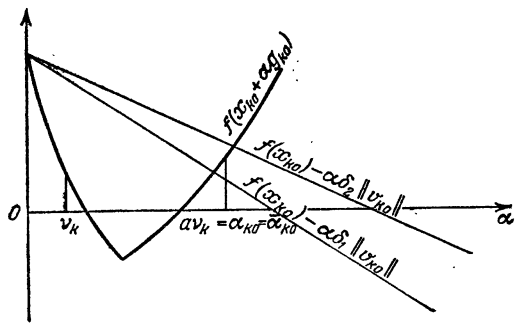


Рис. 53.

Зафиксируем положительные числа $a, \varepsilon_0, \delta, \delta_1, \delta_2$, $0 < \delta_2 < \delta_1 < 1$.

Выберем произвольно начальное приближение $x_0 \in E_n$, число $v_0 > 0$ и целое число $m_0 > 1$.

Пусть уже найдено k -е приближение $x_k \in E_n$. Опишем k -й цикл метода, т. е. процесс перехода от x_k к x_{k+1} . Выберем число $v_k > 0$ и целое число $m_k > 1$ — максимальную длину k -го цикла.

Положим $x_{k0} = x_k$ и вычислим любой субградиент $v_{k0} \in \partial f(x_{k0})$. Если $\|v_{k0}\| \leq \varepsilon_0$, то полагаем $t_k = 0$, $x_{k+1} = x_{k+t_k}$, и на этом k -й цикл заканчивается. В противном случае найдем на луче $\{x_{k0}(\alpha) = x_{k0} + \alpha g_{k0} | \alpha \geq 0\}$, где $g_{k0} = -v_{k0} \|v_{k0}\|^{-1}$, точки $x_{k1} = x_{k0}(\alpha_{k0})$ и $\bar{x}_{k1} = x_{k0}(\bar{\alpha}_{k0})$ такие, что (рис. 53)

$$f(x_{k1}) \leq f(x_{k0}) - \delta_2 \alpha_{k0} \|v_{k0}\|, \quad (6.19)$$

$$f(\bar{x}_{k1}) \geq \min \{f(x_{k1}), f(x_{k0}) - \delta_1 \bar{\alpha}_{k0} \|v_{k0}\|\}, \quad (6.20)$$

причем

$$\bar{\alpha}_{k0} \geq \alpha_{k0}, \quad v_k \leq \bar{\alpha}_{k0} \leq v_k + (a+1)\alpha_{k0}. \quad (6.21)$$

Опять обратим внимание на то, что x_{k1} может оказаться равным x_{k0} . Заметим также, что не исключен случай, когда $\bar{x}_{k1} = x_{k1}$.

Если $\|x_{k0} - x_{k1}\| \geq \delta$ или $f(x_{k0}) - f(x_{k1}) \geq \delta$, то полагаем $t_k = 1$, $x_{k+1} = x_{kt_k}$, и на этом k -й цикл заканчивается. В противном случае вычисляем любой (и в этом также существенное отличие от метода с полной релаксацией) субградиент $\bar{v}_{k0} \in \partial f(\bar{x}_{k1})$.

В силу (6.20) и леммы 6.3 должно быть $(v_{k0}, \bar{v}_{k0}) \leq \leq \delta_1 \|v_{k0}\|^2$. Более того, если меньшим в (6.20) является число $f(x_{k1})$, то по замечанию 4 $(v_{k0}, \bar{v}_{k0}) \leq 0$.

Теперь находим вектор v_{k1} — ближайшую к началу координат точку отрезка $[v_{k0}, \bar{v}_{k0}]$, т. е.

$$v_{k1} = \beta_{k0} \bar{v}_{k0} + (1 - \beta_{k0}) v_{k0}, \quad \beta_{k0} \in [0, 1],$$

$$\|v_{k1}\| = \min_{\beta \in [0, 1]} \|\beta \bar{v}_{k0} + (1 - \beta) v_{k0}\|.$$

Если $\|v_{k1}\| \leq \varepsilon_0$, то полагаем $t_k = 1$, $x_{k+1} = x_{kt_k}$, и на этом k -й цикл заканчивается. Если же $\|v_{k1}\| > \varepsilon_0$, то, имея точку x_{k1} и направление $g_{k1} = -v_{k1} \|v_{k1}\|^{-1}$, выполняя второй шаг k -го цикла, который, как и все последующие шаги k -го цикла, делается по тем же правилам, что и первый. В результате будет сделано такое число шагов $t_k \leq m_k$, что выполнится по крайней мере одно из условий:

- a) $t_k = m_k$; b) $\|x_{k, t_k-1} - x_{kt_k}\| \geq \delta$;
c) $f(x_{k, t_k-1}) - f(x_{kt_k}) \geq \delta$; d) $\|v_{kt_k}\| \leq \varepsilon_0$.

После этого полагаем $x_{k+1} = x_{kt_k}$, и на этом k -й цикл заканчивается, причем если $x_{kt} = x_{k0} \quad \forall t \in 0:t_k$ и цикл прерывается вследствие выполнения условия d), то заканчивается и весь процесс последовательных приближений — получена такая точка x_k , что $d(x_k) \leq \|v_{kt_k}\| \leq \leq \varepsilon_0$.

Отметим, что по построению на основании леммы 6.3 должно быть

$$(v_{kt}, \bar{v}_{kt}) \leq \delta_1 \|v_{kt}\|^2 \quad \forall t. \quad (6.22)$$

Теорема 6.4. Пусть

$$v_k \rightarrow +0, \quad m_k \rightarrow \infty. \quad (6.23)$$

Тогда, если последовательность $\{f(x_k)\}$ ограничена снизу, то последовательность x_k сходится к такой точке x^* , что

$$d(x^*) = \min_{v \in \partial f(x^*)} \|v\| \leq \varepsilon_0. \quad (6.24)$$

Доказательство. Поскольку последовательность $\{f(x_k)\}$ ограничена снизу, то согласно лемме 6.2

$$x_k \rightarrow x^*, \quad x_{kt} \rightarrow x^* \quad \forall t \in 0:t_k. \quad (6.25)$$

Из (6.23) и (6.25) вытекает, что при достаточно больших k , скажем при $k \geq K$, окончание k -го цикла вследствие выполнения условий б), с) невозможно.

Согласно (6.21) имеем при всех $t \in 0:(t_k - 1)$

$$\|\bar{x}_{k,t+1} - x_{kt}\| \leq v_k + a \|x_{k,t+1} - x_{kt}\|,$$

откуда, в силу (6.23) и (6.25), следует

$$\bar{x}_{kt} \rightarrow x^* \quad \forall t \in 1:t_k, \quad (6.26)$$

ибо $\|\bar{x}_{kt} - x^*\| \leq \|\bar{x}_{kt} - x_{kt}\| + \|x_{kt} - x^*\|$.

Таким образом, все точки x_{kt} , $t \in 0:t_k$, $\bar{x}_{kt}$, $t \in 1:t_k$, $\forall k$ принадлежат некоторому компакту, который обозначим X . Поскольку отображение $\partial f(x)$ ограничено на X , то

$$b = \max_{x \in X} \max_{v \in \partial f(x)} \|v\| < \infty.$$

Следовательно,

$$\begin{aligned} \|\bar{v}_{kt}\| &\leq b, \\ \|v_{kt}\| &\leq b \quad \forall t, \quad \forall k. \end{aligned} \quad (6.27)$$

Возьмем целое число $m > 1 + (\ln \bar{q})^{-1} \cdot \ln(\varepsilon_0 b^{-1})$, где $\bar{q}^2 = 1 - (1 - \delta_1)^2 [1 - 2\delta_1 + (b\varepsilon_0^{-1})^2]^{-1}$. Не ограничивая общности, считаем, что $m_k > m$ при всех $k \geq K$.

Покажем, что тогда $t_k \leq m < m_k$.

Действительно, допустим, что $t_k > m$. Тогда с учетом (6.22), (6.27) и того, что $\|v_{kt}\| > \varepsilon_0$ при всех $t \in 0 : (t_k - 1)$, имеем

$$\begin{aligned} \|v_{k, t+1}\|^2 &= \min_{\beta \in [0, 1]} \|\beta \bar{v}_{kt} + (1 - \beta) v_{kt}\|^2 = \\ &= \|v_{kt}\|^2 \min_{\beta \in [0, 1]} \left[\beta^2 \frac{\|\bar{v}_{kt}\|^2}{\|v_{kt}\|^2} + \right. \\ &\quad \left. + 2\beta(1 - \beta) \frac{(v_{kt}, \bar{v}_{kt})}{\|v_{kt}\|^2} + (1 - \beta)^2 \right] \leq \\ &\leq \|v_{kt}\|^2 \min_{\beta \in [0, 1]} [\beta^2 (be_0^{-1})^2 + 2\beta(1 - \beta) \delta_1 + (1 - \beta)^2] = \\ &= \|v_{kt}\|^2 \left(1 - \frac{(1 - \delta_1)^2}{(be_0^{-1})^2 - 2\delta_1 + 1} \right) = \|v_{kt}\|^2 \bar{q}^2. \end{aligned}$$

Отсюда вытекает, что

$$\|v_{kt}\| \leq \|v_{k0}\| \bar{q}^t \leq b \bar{q}^t \quad \forall t \in 0 : (t_k - 1),$$

и, значит, $\|v_{k, m-1}\| \leq b \bar{q}^{m-1} \leq \varepsilon_0$. Но $\|v_{kt}\| > \varepsilon_0 \quad \forall t \in 0 : (t_k - 1)$.

Следовательно, $t_k \leq m$.

Итак, если последовательность $\{f(x_k)\}$ ограничена снизу, то все циклы конечны и длина их не превосходит числа m , причем при всех $k \geq K$ окончание k -го цикла происходит только вследствие выполнения условия д):

$$\|v_{kt_k}\| \leq \varepsilon_0. \quad (6.28)$$

Предположим теперь, что (6.24) не верно. Тогда

$$\varepsilon_0 < \min_{v \in \partial f(x^*)} \|v\| = \bar{\varepsilon}.$$

Так как отображение $\partial f(x)$ полунепрерывно сверху, то найдется такое $\delta > 0$, что

$$\partial f(x) \subset S_\varepsilon(\partial f(x^*)) \equiv \mathcal{D} \quad \forall x \in S_\delta(x^*),$$

где $\varepsilon = 2^{-1}(\bar{\varepsilon} - \varepsilon_0)$.

Таким образом,

$$\|v\| > \varepsilon_0 \quad \forall v \in \mathcal{D}.$$

Но в силу (6.25) и (6.26) при достаточно больших k будет $v_{kt_k} \in \mathcal{D}$, и поэтому $\|v_{kt_k}\| > \varepsilon_0$, что противоречит (6.28). Теорема доказана.

Метод, изложенный в настоящем пункте, можно применить и для нахождения точек минимума функции $f(x)$ на E_n . Для этого вместо условия d) $\|v_{kt_k}\| \leq \varepsilon_j$ следует проверять условие d) $\|v_{kt_k}\| \leq \varepsilon_k$, причем $\varepsilon_k \rightarrow 0$. В этом случае справедлива

Теорема 6.5. Пусть

$$\varepsilon_k \rightarrow 0, \quad \gamma_k \rightarrow 0, \quad m_k \rightarrow \infty.$$

Тогда либо

$$f(x_k) \xrightarrow[k \rightarrow \infty]{} -\infty,$$

либо при некотором k и соответствующем t

$$f(x_{kt} + \alpha g_{kt}) \xrightarrow[\alpha \rightarrow \infty]{} -\infty,$$

либо любая предельная точка последовательности $\{x_k\}$ является точкой минимума функции $f(x)$ на E_n .

Задание 6.1. Доказать теорему 6.5.

Замечание 5. В методе с полной релаксацией в точке $x_{k, t+1}$ вычисляется $\bar{v}_{kt} \in \partial f(x_{k, t+1})$ такой, что $(v_{kt}, \bar{v}_{kt}) \leq 0$, причем такой субградиент существует. В методе настоящего пункта в точке $\bar{x}_{k, t+1}$ вычисляется любой $\bar{v}_{kt} \in \partial f(\bar{x}_{k, t+1})$, причем будет $(v_{kt}, \bar{v}_{kt}) \leq \delta_1 \|v_{kt}\|^2$. Поскольку выпуклая функция почти всюду дифференцируема, т. е. почти всюду $\partial f(x)$ состоит из единственной точки — градиента функции $f(x)$, то может показаться, что конструкция «любой $v \in \partial f(x)$ » немногим предпочтительней конструкции «существует $v \in \partial f(x)$ ». Но если функция $f(x)$ не является гладкой, то зачастую в процессе точной одномерной минимизации происходит попадание именно на то множество меры нуль, на котором $\partial f(x)$ — выпуклый компакт с непустой внутренностью, и поэтому в конструкции «существует $v \in \partial f(x)$ » необходимо действительно организовывать поиск «того, что существует». Это соображение указывает еще одно преимущество по простоте реализации метода настоящего пункта перед методом с полной релаксацией.

§ 7. Релаксационный ε -субградиентный метод

В § 6 предполагалось, что могут быть точно вычислены целевая функция $f(x)$ и ее субградиенты. При расчетах на ЭВМ, однако, такая возможность отсутствует,

ибо, во-первых, ЭВМ оперирует только с рациональными числами, а во-вторых, для вычисления самой целевой функции и тем более ее субградиентов может потребоваться бесконечный итерационный процесс. Примером такой функции является функция максимума $f(x) = \max_{y \in G} \varphi(x, y)$ (см. § 17 гл. I).

Поэтому важно построить такой вариант релаксационного субградиентного метода, в котором допускалось бы приближенное вычисление целевой функции и ее субградиентов, т. е. в котором вместо целевой функции $f(x)$ использовалась бы некоторая функция $f_\varepsilon(x)$ такая, что

$$f(x) - \varepsilon \leq f_\varepsilon(x) \leq f(x) + \varepsilon, \quad (7.1)$$

а вместо субградиентов $v \in \partial f(x) - \varepsilon$ -субградиенты $v_\varepsilon \in \partial_\varepsilon f(x)$.

Такой вариант релаксационного субградиентного метода и построен в настоящем параграфе.

1. Сначала получим аналог леммы 6.3 в рассматриваемом случае. Возьмем положительные числа $\delta_1, c, v, \varepsilon, \varepsilon_0$ и δ_2 , причем $\delta_1 \in (0, 1)$, $\delta_2 < \delta_1$, $\varepsilon \leq (c+2)^{-1}(\delta_1 - \delta_2)\varepsilon_0 v$.

Положим $x_1 = x_0 + \alpha_0 g_0$, где $g_0 = -v_0 \|v_0\|^{-1}$, $v_0 \in E_n$.

Лемма 7.1. Если $\|v_0\| \geq \varepsilon_0$, $\alpha_0 \geq v$ и

$$\mu \equiv f_\varepsilon(x_1) - [f_\varepsilon(x_0) - \delta_2 \alpha_0 \|v_0\|] \geq 0, \quad (7.2)$$

то

$$(v_0, \bar{v}_0) \leq \delta_1 \|v_0\|^2 \quad \forall \bar{v}_0 \in \partial_{c\varepsilon+\mu} f(x_1). \quad (7.3)$$

Доказательство. Согласно (7.2) имеем

$$\begin{aligned} f_\varepsilon(x_1) &= f_\varepsilon(x_0) - \delta_1 \alpha_0 \|v_0\| + (\delta_1 - \delta_2) \alpha_0 \|v_0\| + \mu \geq \\ &\geq f_\varepsilon(x_0) - \delta_1 \alpha_0 \|v_0\| + (c+2)\varepsilon + \mu. \end{aligned}$$

Отсюда, в силу (7.1), получаем

$$\begin{aligned} f(x_1) &\geq f_\varepsilon(x_1) - \varepsilon \geq f_\varepsilon(x_0) - \delta_1 \alpha_0 \|v_0\| + (c+1)\varepsilon + \mu \geq \\ &\geq f(x_0) - \delta_1 \alpha_0 \|v_0\| + c\varepsilon + \mu. \end{aligned}$$

Теперь справедливость (7.3) вытекает из того, что для любого $\bar{v}_0 \in \partial_{c\varepsilon+\mu} f(x_1)$ должно быть

$$f(x_0) \geq f(x_1) - \alpha_0 (g_0, \bar{v}_0) - (c\varepsilon + \mu).$$

Лемма 7.1 показывает, какой должна быть точность вычисления функции $f(x)$ для того, чтобы из (7.2) следовало (7.3).

2. Переходим к описанию релаксационного ε -субградиентного метода.

Зафиксируем положительные числа $a, c, v, \varepsilon, \delta, \varepsilon_0, \delta_i, i \in 1:3$, такие, что $\delta_1 \in (0, 1)$, $0 < \delta_3 < \delta_2 < \delta_1$, $\varepsilon \leq \leq (c+2)^{-1}(\delta_1 - \delta_2)\varepsilon_0 v$. Выберем произвольно начальное приближение $x_0 \in E_n$ и целое число m_0 . Пусть уже найдено k -е приближение $x_k \in E_n$. Переход к $(k+1)$ -му приближению x_{k+1} (k -й цикл) осуществляется следующим образом.

Выберем целое число m_k — максимальную длину k -го цикла.

Положим $x_{k_0} = x_k$ и вычислим $f_\varepsilon(x_{k_0})$ и произвольный ε -субградиент $v_{k_0} \in \partial_{\varepsilon} f(x_{k_0})$. Если $\|v_{k_0}\| \leq \varepsilon_0$, то полагаем $t_k = 0$, $x_{k+1} = x_{kt_k}$, и k -й цикл заканчивается. Если же $\|v_{k_0}\| > \varepsilon_0$, то на луче $\{x_{k_0}(\alpha) = x_{k_0} + \alpha g_{k_0} | \alpha \geq 0\}$, где $g_{k_0} = -v_{k_0} \|v_{k_0}\|^{-1}$, найдем точки $x_{k_1} = x_{k_0}(\alpha_{k_0})$ и $\bar{x}_{k_1} = x_{k_0}(\bar{\alpha}_{k_0})$ такие, что (при $t = 0$)

$$\bar{\alpha}_{kt} \geq \alpha_{kt}, \quad v \leq \bar{\alpha}_{kt} \leq v + a\alpha_{kt}, \quad (7.4)$$

причем либо $\alpha_{kt} = 0$, либо $\alpha_{kt} \geq v$, и

$$f_\varepsilon(x_{k,t+1}) \leq f_\varepsilon(x_{kt}) - \delta_3 \alpha_{kt} \|v_{kt}\|, \quad (7.5)$$

$$f_\varepsilon(\bar{x}_{k,t+1}) - \min \{f_\varepsilon(x_{k,t+1}), f_\varepsilon(x_{kt}) - \delta_2 \bar{\alpha}_{kt} \|v_{kt}\|\} \geq 0. \quad (7.6)$$

Если же (7.5) выполняется при любом $\alpha_{kt} \geq 0$, то $f_\varepsilon(x_{kt}(\alpha)) \xrightarrow{\alpha \rightarrow \infty} -\infty$, следовательно, $f(x_{kt}(\alpha)) \xrightarrow{\alpha \rightarrow \infty} -\infty$, и процесс последовательных приближений заканчивается.

Если $\|x_{k_0} - x_{k_1}\| \geq \delta$ либо $f_\varepsilon(x_{k_0}) - f_\varepsilon(x_{k_1}) \geq \delta$, то полагаем $t_k = 1$, $x_{k+1} = x_{kt_k}$, т. е. k -й цикл закончен. В противном случае вычислим произвольный ε -субградиент $\bar{v}_{k_0} \in \partial_{\varepsilon} f(\bar{x}_{k_1})$. Так как $\|v_{k_0}\| > \varepsilon_0$, $v \leq \leq \bar{\alpha}_{k_0}$, $\varepsilon \leq (c+2)^{-1}(\delta_1 - \delta_2)\varepsilon_0 v$ и выполнено (7.6), то согласно лемме 7.1

$$(v_{k_0}, \bar{v}_{k_0}) \leq \delta_1 \|v_{k_0}\|^2.$$

После этого находим вектор v_{k_1} — ближайшую к началу координат точку отрезка $[v_{k_0}, \bar{v}_{k_0}]$.

Если $\|v_{k1}\| \leq \varepsilon_0$, то полагаем $t_k = 1$, $x_{k+1} = x_{kt_k}$, и на этом k -й цикл заканчивается. Если же $\|v_{k1}\| > \varepsilon_0$, то на луче $\{x_{k1}(\alpha) = x_{k1} + \alpha g_{k1} | \alpha \geq 0\}$, где $g_{k1} = -v_{k1}/\|v_{k1}\|^{-1}$, находим такие точки $x_{k2} = x_{k1}(\alpha_{k1})$ и $\bar{x}_{k2} = x_{k1}(\bar{\alpha}_{k1})$, что при $t=1$ выполняются соотношения (7.4)–(7.6).

Далее продолжаем аналогично до тех пор, пока не будет сделано такое число шагов t_k , что выполнится по крайней мере одно из следующих условий:

- а) $t_k = m_k$; б) $\|x_{k, t_k-1} - x_{kt_k}\| \geq \delta$;
 в) $f_\varepsilon(x_{k, t_k-1}) - f_\varepsilon(x_{kt_k}) \geq \delta$; д) $\|v_{kt_k}\| \leq \varepsilon_0$.

Теперь полагаем $x_{k+1} = x_{kt_k}$, и на этом k -й цикл заканчивается, причем если k -й цикл прервался вследствие выполнения условия д) и $x_{kt} = x_{k0}$ при всех $t \in 1:t_k$, то весь процесс последовательных приближений также заканчивается.

В результате будет построена последовательность $\{x_k\}$.

Положим $\bar{v}_{k,-1} = v_{k0}$,

$$l_k = \begin{cases} t_k - 1, & \text{если } k\text{-й цикл прерван вследствие д),} \\ t_k - 2 & \text{в противном случае.} \end{cases}$$

Отметим, что по построению на основании леммы 7.1 будет

$$(v_{kt}, \bar{v}_{kt}) \leq \delta_1 \|v_{kt}\|^2 \quad \forall t \in 0:l_k.$$

Теорема 7.1. Пусть

$$m_k \rightarrow \infty. \quad (7.7)$$

Тогда либо

$$f(x_k) \xrightarrow[k \rightarrow \infty]{} -\infty,$$

либо при некотором k и соответствующем t

$$f(x_{kt} + \alpha g_{kt}) \xrightarrow[\alpha \rightarrow \infty]{} -\infty,$$

либо процесс последовательных приближений заканчивается в такой точке x_k , что

$$f(z) - f(x_k) \geq -\varepsilon_0 \|z - x_k\| - (c+2)\varepsilon - (1+\delta_2)vb_0 \quad (7.8) \\ \forall z \in E_n,$$

где $b_0 = \max_{t \in -1:l_k} \|v_{kt}\|$.

Доказательство. Рассмотрим сначала случай, когда последовательность $\{x_k\}$ бесконечна. Покажем, что тогда найдется такая подпоследовательность $\{x_{k_s}\}$, что

$$x_{k_s+1} \neq x_{k_s} \quad \forall s. \quad (7.9)$$

Допустим, что это не так. Тогда, начиная с некоторого K , при всех $k \geq K$ должно быть

$$x_k = x_{k0} = \dots = x_{kt_k} = x_{k+1}, \quad (7.10)$$

и, следовательно,

$$\varepsilon_0 < \|v_{kt}\| \leq b \quad \forall t \in 1:l_k, \quad \|\bar{v}_{kt}\| \leq b \quad \forall t \in -1:l_k, \quad (7.11)$$

где $b \leq \max_{v \in \partial_{ce} f(x_k)} \|v\| < \infty$.

Так как последовательность $\{x_k\}$ бесконечна, то из (7.10), (7.11) вытекает, что при любом $k \geq K$ будет $t_k = m_k$.

В силу (7.7), не ограничивая общности, можно считать, что $m_k > m$, где целое число m таково, что $m > 1 + \ln^{-1} \tilde{q} \cdot \ln(\varepsilon_0 b^{-1})$,

$$\tilde{q} = [1 - (1 - \delta_1)^2 (1 - 2\delta_1 + (b\varepsilon_0^{-1})^2)^{-1}]^{1/2}.$$

Тогда так же, как и при доказательстве теоремы 6.4, показывается, что $\|v_{k, m-1}\| \leq \varepsilon_0$, а это противоречит (7.11).

Итак, если последовательность $\{x_k\}$ бесконечна, то найдется такая подпоследовательность $\{x_{k_s}\}$, для которой выполняется (7.9). Тогда на основании (7.4) и (7.5) заключаем, что

$$f_\varepsilon(x_{k_s+1}) \leq f_\varepsilon(x_{k_s}) - \delta_3 \nu \varepsilon_0 \quad \forall s,$$

откуда и из того, что $f_\varepsilon(x_{k+1}) \leq f_\varepsilon(x_k)$, вытекает $f_\varepsilon(x_k) \xrightarrow[k \rightarrow \infty]{} -\infty$, а это, в силу (7.1), означает, что $f(x_k) \xrightarrow[k \rightarrow \infty]{} -\infty$.

Обратимся теперь к случаю, когда последовательность $\{x_k\}$ конечна. Предположим, что процесс заканчивается в точке $x_{k+1} = x_k$. Тогда выполняется (7.10) и $\|v_{kt_k}\| \leq \varepsilon_0$.

Так как $\bar{v}_{k,-1} = v_{k0} \in \partial_{ce} f(x_{k0})$, то

$$f(z) - f(x_k) \geq (\bar{v}_{k,-1}, z - x_k) - ce \quad \forall z \in E_n. \quad (7.12)$$

Согласно (7.4) и (7.10), имеем

$$x_k - \bar{x}_{kt} = -v g_{k,t-1} = v v_{k,t-1} \|v_{k,t-1}\|^{-1} \quad \forall t \in 1:t_k. \quad (7.13)$$

Поскольку при любых $t \in 1:t_k$ и $z \in E_n$

$$\begin{aligned} f(z) - f(\bar{x}_{kt}) &\geq (\bar{v}_{k,t-1}, z - \bar{x}_{kt}) - ce = \\ &= (\bar{v}_{k,t-1}, z - x_k) - ce + (\bar{v}_{k,t-1}, x_k - \bar{x}_{kt}), \end{aligned}$$

то, с учетом (7.13) и того, что $\|\bar{v}_{k,t-1}\| \leq b_k$, получим

$$f(z) - f(\bar{x}_{kt}) \geq (\bar{v}_{k,t-1}, z - x_k) - ce - v b_k \quad (7.14)$$

$$\forall z \in E_n, \quad \forall t \in 1:t_k.$$

В силу (7.10), (7.6) и (7.13) будет

$$\begin{aligned} f(\bar{x}_{kt}) - f(x_k) &\geq \\ &\geq -2e - \delta_2 v \|v_{k,t-1}\| \geq -2e - \delta_2 v b_k \quad \forall t \in 1:t_k, \end{aligned}$$

откуда и из (7.14), (7.12) следует

$$f(z) - f(x_k) \geq (\bar{v}_{k,t-1}, z - x_k) - (c+2)e - (1+\delta_2) v b_k \quad (7.15)$$

$$\forall z \in E_n, \quad \forall t \in 0:t_k.$$

Очевидно, что $v_{kt_k} \in \text{co} \{\bar{v}_{k,t-1} | t \in 0:t_k\}$, т. е.

$$\begin{aligned} v_{kt_k} &= \sum_{t=-1}^{l_k} \lambda_t \bar{v}_{kt}, \quad \sum_{t=-1}^{l_k} \lambda_t = 1, \\ \lambda_t &\geq 0 \quad \forall t \in -1:l_k. \end{aligned} \quad (7.16)$$

Умножим теперь каждое из неравенств (7.15) на соответствующий коэффициент λ_t из представления (7.16) вектора v_{kt_k} и сложим полученные при этом $t_k + 1$ неравенств. Тогда

$$\begin{aligned} f(z) - f(x_k) &\geq (v_{kt_k}, z - x_k) - (c+2)e - (1+\delta_2) v b_k \geq \\ &\geq -\|v_{kt_k}\| \cdot \|z - x_k\| - (c+2)e - (1+\delta_2) v b_k \quad \forall z \in E_n, \end{aligned}$$

откуда с учетом того, что $\|v_{kt_k}\| \leq e_0$, следует (7.8). Теорема доказана.

3. Теорема 7.1 и, в частности, оценка (7.8) показывают, как можно применить метод, изложенный в п. 2,

который для краткости назовем ε -алгоритмом, для нахождения точек минимума функции $f(x)$ на E_n .

Возьмем последовательности $\{\varepsilon_{0k}\}$, $\{\varepsilon_k\}$ и $\{v_k\}$ положительных чисел, стремящиеся к нулю. Выберем произвольным образом начальное приближение $\tilde{x}_0 \in E_n$. Применим ε -алгоритм при $\varepsilon_0 = \varepsilon_{01}$, $\varepsilon = \varepsilon_1$, $v = v_1$, начиная из точки $x_0 = \tilde{x}_0$. В результате его работы либо установим, что $f^* = \inf_{x \in E_n} f(x) = -\infty$, либо найдем точку x_k ,

удовлетворяющую (7.8). Положим $\tilde{x}_1 = x_k$. Применим ε -алгоритм, начиная из точки $x_0 = \tilde{x}_1$, при $\varepsilon_0 = \varepsilon_{02}$, $\varepsilon = \varepsilon_2$, $v = v_2$ и либо установим, что $f^* = -\infty$, либо найдем точку x_k , удовлетворяющую (7.8), но уже при измененных ε_0 , ε , v . Положим $\tilde{x}_2 = x_k$.

Продолжая далее аналогично, либо установим, что $f^* = -\infty$, либо построим бесконечную последовательность $\{\tilde{x}_k\}$, каждая точка которой удовлетворяет, в силу (7.8), соотношению

$$\begin{aligned} f(z) - f(\tilde{x}_k) &\geq \\ &\geq -\varepsilon_{0k} \|z - \tilde{x}_k\| - (c + 2)\varepsilon_k - (1 + \delta_2)v_k b_{k-1} \quad \forall z \in E_{i_*}. \end{aligned}$$

Отсюда ясно, что если множество $\mathcal{D}(\tilde{x}_0) = \{x \in E_n | f(x) \leq f(\tilde{x}_0)\}$ ограничено, то $f(\tilde{x}_k) \rightarrow f^*$, причем любая предельная точка последовательности $\{\tilde{x}_k\}$ будет точкой минимума функции $f(x)$ на E_n .

4. Во всех рассмотренных вариантах релаксационного субградиентного метода направление движения определяется из решения простейшей задачи квадратичного программирования — нахождения ближайшей к началу координат точки отрезка $[v_{kt}, \bar{v}_{kt}]$. Если размерность пространства E_n не слишком велика и позволяют возможности используемой ЭВМ, то эффективность релаксационного субградиентного метода значительно повышается, если ввести в процесс последовательных приближений более сложную вспомогательную задачу выбора направления движения, заключающуюся в нахождении ближайшей к началу координат точки многогранника, натянутого на множество $\mathcal{H}_{kt} = \{\bar{v}_{k,-1}, \bar{v}_{k0}, \dots, \bar{v}_{k,t-1}\}$ или его некоторую часть.

В этом случае векторы $v_{k,t+1}$, $t \in 1:t_k$, строятся следующим образом: выбирается некоторое подмножество

$\overline{\mathcal{H}}_{kt}$ множества $\mathcal{H}_{kt}$, и находится вектор $v_{k,t+1} \in \mathcal{L}_{kt}$, где $\mathcal{L}_{kt} = \text{co}\{v_{kt}, \bar{v}_{kt}, \overline{\mathcal{H}}_{kt}\}$, такой, что

$$\|v_{k,t+1}\| = \min_{v \in \mathcal{L}_{kt}} \|v\|. \quad (7.17)$$

При такой модификации релаксационный субградиентный метод становится близким к методу ε -наискорейшего спуска и в идейном плане, и по эффективности.

Поскольку $[v_{kt}, \bar{v}_{kt}] \subset \mathcal{L}_{kt}$, то все теоремы, доказанные в §§ 6, 7, остаются справедливыми, если находить вектор $v_{k,t+1}$ из условия (7.17).

Зная размеры исходной задачи минимизации функции $f(x)$ на E_n , можно подсчитать максимально допустимый на используемой ЭВМ размер задачи квадратичного программирования (7.17). Чем больше векторов из $\mathcal{H}_{kt}$ включается в $\overline{\mathcal{H}}_{kt}$, тем более точно множество $\mathcal{L}_{kt}$ будет аппроксимировать множество $\partial_e f(x_{k,t+1})$ и тем реальнее значительно уменьшить функцию $f(x)$, двигаясь из точки $x_{k,t+1}$ в направлении

$$g_{k,t+1} = -v_{k,t+1} \|v_{k,t+1}\|^{-1}.$$

5. В заключение остановимся на вспомогательной задаче одномерного поиска точек $x_{k,t+1} = x_{kt}(\alpha_{kt})$ и $\bar{x}_{k,t+1} = x_{kt}(\bar{\alpha}_{kt})$, удовлетворяющих (7.4)–(7.6). Очевидно, поиск этих точек можно производить многими способами. Покажем здесь, как это делается по методу золотого сечения. Отметим, что обычно метод золотого сечения применяют для одномерной минимизации на отрезке. Ниже приводится вариант метода золотого сечения для поиска на луче.

Введём обозначения: $x(\alpha) = x_{kt}(\alpha)$, $x = x_{kt}$, $x_+ = x_{k,t+1}$, $\bar{x}_+ = \bar{x}_{k,t+1}$, $f(\alpha) = f_\varepsilon(x(\alpha))$, $\xi_s = f(\alpha_s) - f(\alpha_{s-1}) + \delta_3 \alpha_s \|v_{kt}\|$, $\mu_s = f(\alpha_s) - \min\{f(\alpha_{s-1}), f(\alpha_0) - \delta_2 \alpha_s \|v_{kt}\|\}$.

Положим $\alpha_0 = 0$, $\alpha_1 = v$, $q = 2^{-1}(1 + \sqrt{5}) > 1$.

Вычислим $f(\alpha_0)$, $f(\alpha_1)$ и μ_1 . Если $\mu_1 \geq 0$, то полагаем $x_+ = x$, $\bar{x}_+ = x(\alpha_1)$, иначе вычислим $f(\alpha_2)$, μ_2 и ξ_2 ,

где $\alpha_2 = v(1 + q) = v \sum_{i=0}^1 q^i$.

Пусть уже найдены $f(\alpha_s)$, μ_s и ξ_s , где $\alpha_s = v \sum_{i=0}^{s-1} q^i$.

Если $\mu_s < 0$, то полагаем $\alpha_{s+1} = v \sum_{i=0}^s q^i$ и вычисляем $f(\alpha_{s+1})$, μ_{s+1} , ξ_{s+1} .

Пусть $\mu_s \geq 0$. Если $\xi_{s-1} \leq 0$ и $\alpha_s \leq v + a\alpha_{s-1}$ или $\xi_s \leq 0$, то полагаем $\bar{x}_+ = x(\alpha_s)$,

$$x_+ = \begin{cases} \bar{x}_+, & \text{если } \xi_s \leq 0, \\ x(\alpha_{s-1}), & \text{если } \xi_s > 0, \xi_{s-1} \leq 0, \alpha_s \leq v + a\alpha_{s-1}. \end{cases}$$

В противном случае (по-прежнему рассматриваем случай $\mu_s \geq 0$) продолжаем поиск на отрезке $[\alpha_{s-2}, \alpha_s]$ методом золотого сечения. Заметим, что α_{s-1} есть золотое сечение отрезка $[\alpha_{s-2}, \alpha_s]$.

Действительно, как известно, золотым сечением отрезка называется такое деление его на части, что отношение длины меньшей части к длине большей части равно отношению длины большей части к длине всего отрезка. Имеем

$$\alpha_{s-2} = v \sum_{i=0}^{s-3} q^i, \quad \alpha_{s-1} = v \sum_{i=0}^{s-2} q^i, \quad \alpha_s = v \sum_{i=0}^{s-1} q^i,$$

$$\frac{\alpha_s - \alpha_{s-1}}{\alpha_s - \alpha_{s-2}} = \frac{q^{s-1}}{q^{s-1} + q^{s-2}} = \frac{q}{1+q},$$

$$\frac{\alpha_{s-1} - \alpha_{s-2}}{\alpha_s - \alpha_{s-1}} = \frac{q^{s-2}}{q^{s-1}} = \frac{1}{q},$$

а одним из корней уравнения $\frac{q}{1+q} = \frac{1}{q}$ является число $2^{-1}(1 + \sqrt{5})$.

На отрезке $[\alpha_{s-2}, \alpha_s]$ поиск продолжается следующим образом. Положим $\alpha_{s+1} = \alpha_s$, $\alpha_s = \alpha_{s-2} + \alpha_{s+1} - \alpha_{s-1}$ и вычислим $f(\alpha_s)$, μ_s и ξ_s .

Случай $\mu_s \geq 0$ рассматривается так же, как и выше.

Пусть $\mu_s < 0$. Тогда, если $\alpha_{s+1} \leq v + a\alpha_s$, то полагаем $x_+ = x(\alpha_s)$, $\bar{x}_+ = x(\alpha_{s+1})$, в противном же случае полагаем $\alpha_{s-2} = \alpha_{s-1}$, $\alpha_{s-1} = \alpha_s$, $\alpha_s = \alpha_{s+1}$ и продолжаем поиск на отрезке $[\alpha_{s-2}, \alpha_s]$.

Ясно, что такой процесс либо конечен, либо $f(\alpha_s) \xrightarrow{\alpha_s \rightarrow \infty} -\infty$.

З а м е ч а н и е 1. Если k -й цикл прерван по выполнении условия а) и $x_{kt} = x_k \forall t \in 0 : t_k$, то на самом деле это означает, что число m_k выбрано неудачно, и, увеличив его, следует продолжить k -й цикл.

З а м е ч а н и е 2. С вычислительной точки зрения наиболее удобен случай, когда $\bar{x}_{kt} = x_{kt}$, поскольку за следующую точку принимается точка x_{kt} и субградиент $\bar{v}_{k, t-1}$ вычисляется также в точке x_{kt} и нет необходимости хранить в памяти ЭВМ лишнюю точку. Если $\alpha_{k, t-1} \geq \geq v$, то всегда можно добиться того, чтобы $\bar{x}_{kt} = x_{kt}$, причем затратив на это конечное число вычислений функции. Но специально на это много усилий затрачивать не стоит.

§ 8. Метод Келли

1. Метод Келли. Пусть $f(x)$ — выпуклая конечная на E_n функция. Тогда (см. (5.11) гл. I)

$$f(x) = \max_{z \in E_n} [f(z) + (v(z), x - z)] \equiv \max_{z \in E_n} B(z, x), \quad (8.1)$$

где $v(z) \in \partial f(z)$ (для каждой точки z выбирается один произвольный вектор из $\partial f(z)$). Требуется найти

$$\min_{x \in \Omega} f(x),$$

где $\Omega \subset E_n$ — замкнутое ограниченное множество. Выберем произвольный набор σ_0 конечного числа точек из E_n . Пусть уже найдено множество σ_k . Образует функцию

$$f_k(x) = \max_{z \in \sigma_k} B(z, x)$$

и найдем

$$\min_{x \in \Omega} f_k(x) = f_k(x_k). \quad (8.2)$$

Если

$$f_k(x_k) = f(x_k), \quad (8.3)$$

то точка x_k — точка минимума функции $f(x)$ на Ω . Дей-

ствительно, $f_k(x) \leq f(x)$, и потому

$$f_k(x_k) = \min_{x \in \Omega} f_k(x) \leq \min_{x \in \Omega} f(x) = f^* \leq f(x_k). \quad (8.4)$$

Отсюда и из (8.3) $f^* = f(x_k)$.

Если же $f_k(x_k) < f(x_k)$, то полагаем $\sigma_{k+1} = \sigma_k \cup \{x_k\}$ и далее поступаем аналогично. Заметим, что $x_k \in \Omega$,

$$\begin{aligned} f_{k+1}(x_{k+1}) &\geq f_k(x_k), \\ B(x_k, x_k) &= f(x_k). \end{aligned} \quad (8.5)$$

В результате строим последовательность множеств $\{\sigma_k\}$ и точек $\{x_k\}$.

Теорема 8.1. *Всякая предельная точка последовательности $\{x_k\}$ является точкой минимума функции $f(x)$ на Ω .*

Доказательство. Существование предельных точек вытекает из ограниченности и замкнутости множества Ω . Пусть $x_{k_s} \rightarrow x^* \in \Omega$. Надо показать, что

$$f(x^*) = f^*. \quad (8.6)$$

Заметим вначале, что $f(x_{k_s}) \rightarrow f(x^*)$ по непрерывности функции $f(x)$. Зафиксируем любое k_{s_0} . Тогда при $k_s > k_{s_0}$ в силу (8.4)

$$\begin{aligned} f^* &\geq f_{k_s}(x_{k_s}) \geq B(x_{k_{s_0}}, x_{k_s}) = \\ &= f(x_{k_{s_0}}) + (v_{k_{s_0}}, x_{k_s} - x_{k_{s_0}}). \end{aligned} \quad (8.7)$$

Учитывая ограниченность $v(z)$ на Ω (отображение $df(z)$ ограничено на любом ограниченном множестве) и переходя к пределу по k_{s_0} , $k_s \rightarrow \infty$, при $k_s > k_{s_0}$, получаем из (8.7)

$$f^* \geq f(x^*). \quad (8.8)$$

Очевидно, что

$$f(x^*) \geq \inf_{x \in \Omega} f(x) = f^*. \quad (8.9)$$

Из (8.8) и (8.9) следует (8.6). Теорема доказана.

Замечание 1. Если оказалось $f(x_k) - f_k(x_k) \leq \varepsilon$, то из неравенства $f_k(x) \leq f(x)$ имеем

$$\min_{x \in \Omega} f_k(x) = f_k(x_k) \leq \min_{x \in \Omega} f(x) = f^*.$$

Отсюда $f(x_k) - \varepsilon \leq f^*$, т. е. точка x_k — ε -стационарная точка функции $f(x)$ на Ω .

З а м е ч а н и е 2. Рассмотрим теперь задачу минимизации функции $f(x)$ на всем пространстве E_n . Предположим, что множество

$$\mathcal{M}^* = \{x \in E_n \mid f(x) = \inf_{x \in E_n} f(x)\}$$

непусто и ограничено. Выберем компакт $\Omega \subset E_n$ такой, что $\mathcal{M}^* \subset \text{int } \Omega$ (естественно выбрать Ω попроще, например, заданное линейными неравенствами). Применяем изложенный выше метод Келли. При этом из теоремы 8.1 следует, что при больших k будет

$$\min_{x \in \Omega} f_k(x) = \min_{x \in E_n} f_k(x),$$

т. е. множество Ω играет вспомогательную роль и используется лишь в начале процесса. Если заранее множество Ω неизвестно (но известно, что $\mathcal{M}^*$ ограничено), то можно взять произвольное множество $\Omega_0 \subset E_n$ и сделать несколько шагов; если на некотором шаге k_0 окажется, что

$$x_{k_0} \in \text{int } \Omega_0, \quad (8.10)$$

то дальше при $k > k_0$ можно, в силу (8.4), вместо задачи (8.2) решать задачу

$$\min_{x \in E_n} f_k(x).$$

Если же включение (8.10) долго не имеет места, то можно попробовать взять новое $\Omega'_0 \supset \Omega_0$.

З а м е ч а н и е 3. Положим

$$R_k = \{z \in \sigma_k \mid f_k(x_k) = B(z, x_k)\}.$$

Легко заметить, что в этом случае задача (8.2) эквивалентна задаче

$$\min_{x \in \Omega} \bar{f}_k(x),$$

где

$$\bar{f}_k(x) = \max_{z \in R_k} B(z, x).$$

Предположим, что $f(x)$ — сильно выпуклая функция, и пусть x^* — точка минимума функции $f(x)$ на Ω . По теореме 8.1 последовательность $\{x_k\}$ сходится к точке x^* (у сильно выпуклой функции существует единственная точка минимума). Покажем, что любая точка $\bar{x} \neq x^*$ может встретиться в R_k лишь для конечного числа k . Действительно, допустим противное. Пусть $\bar{x} \in R_{k_s}$, $k_s \rightarrow \infty$. Имеем

$$B(\bar{x}, x_{k_s}) = f(\bar{x}) + (v(\bar{x}), x_{k_s} - \bar{x}) = f_{k_s}(x_{k_s}).$$

Так как $x_{k_s} \rightarrow x^*$, то из (8.9)

$$f(\bar{x}) + (v(\bar{x}), x^* - \bar{x}) = f(x^*).$$

Но функция $f(x)$ — сильно выпуклая, и поэтому существует $m > 0$ такое, что (см. (9.17) гл. I)

$$f(x^*) = f(x + (x^* - \bar{x})) \geq f(\bar{x}) + (v(\bar{x}), x^* - \bar{x}) + m\|x^* - \bar{x}\|^2.$$

Поскольку $\bar{x} \neq x^*$, то отсюда $f(x^*) > f(\bar{x}) + (v(\bar{x}), x^* - \bar{x})$. Получили противоречие с предыдущим равенством.

Итак, всякая точка $x_k \neq x^*$ является «существенной» лишь конечное число раз. Этот факт может быть использован при практической реализации метода Келли.

2. ε -метод Келли. Для каждого $z \in E_n$ выберем какой-нибудь вектор $v(z) \in \partial_\varepsilon f(z)$ и положим

$$F(x) = \sup_{z \in E_n} [f(z) + (v(z), x - z)]. \quad (8.11)$$

Из определения $\partial_\varepsilon f(z)$ имеем

$$f(x) - f(z) \geq (v(z), x - z) - \varepsilon \quad \forall x \in E_n.$$

Отсюда

$$f(x) \leq \sup_{z \in E_n} [f(z) + (v(z), x - z)] = F(x) \leq f(x) + \varepsilon. \quad (8.12)$$

Функция $F(x)$ тоже является выпуклой, как супремум выпуклых функций. Пусть $\Omega \subset E_n$ — произвольное замкнутое ограниченное множество. Требуется найти $\min_{x \in \Omega} f(x) = f^*$. Напомним, что если $x^* \in \Omega$ и $f^* \leq f(x^*) \leq f^* + \varepsilon$, то точка x^* называется ε -стационарной точкой

функции $f(x)$ на Ω . При этом оказывается (см. лемму 8.1 гл. I), что если $x^* \in \text{int } \Omega$, то

$$0 \in \partial_{\varepsilon} f(x^*).$$

Напомним, что множество $\bigcup_{z \in \Omega} \partial_{\varepsilon} f(z)$ ограничено.

Опишем обобщение изложенного в п. 1 метода Келли для нахождения ε -стационарных точек. Пусть σ_0 — произвольный конечный набор точек из E_n , и пусть уже найдено множество σ_k . Образует функцию

$$f_k(x) = \max_{z \in \sigma_k} B(z, x),$$

где

$$B(z, x) = f(z) + (v(z), x - z), \quad v(z) \in \partial_{\varepsilon} f(z),$$

и найдем

$$f_k(x_k) = \min_{x \in \Omega} f_k(x). \quad (8.13)$$

Если $f_k(x_k) \geq f(x_k)$, то из (8.12)

$$f^* \leq f(x_k) \leq f_k(x_k) = \min_{x \in \Omega} f_k(x) \leq \min_{x \in \Omega} f(x) + \varepsilon = f^* + \varepsilon. \quad (8.14)$$

Значит, $f^* \leq f(x_k) \leq f^* + \varepsilon$, т. е. x_k — ε -стационарная точка функции $f(x)$ на Ω , и процесс прекращается.

Если же $f_k(x_k) < f(x_k)$, то полагаем $\sigma_{k+1} = \sigma_k \cup \{x_k\}$ и далее продолжаем аналогично. В результате строим последовательности множеств $\{\sigma_k\}$ и точек $\{x_k\}$ таких, что

$$x_k \in \Omega, \quad f_k(x_k) \leq f_{k+1}(x_{k+1}). \quad (8.15)$$

Отметим еще, что если при некоторых k и $\mu > 0$ окажется

$$f(x_k) - f_k(x_k) \leq \mu, \quad (8.16)$$

то из (8.14) получим

$$f_k(x_k) \leq f^* + \varepsilon. \quad (8.17)$$

Учитывая (8.16), имеем

$$f(x_k) - \mu \leq f_k(x_k) \leq f^* + \varepsilon,$$

т. е. $f^* \leq f(x_k) \leq f^* + \mu + \varepsilon$, а это значит, что x_k является $(\varepsilon + \mu)$ -стационарной точкой функции $f(x)$ на Ω .

Теорема 8.2. *Всякая предельная точка последовательности $\{x_k\}$ является ε -стационарной точкой функции $f(x)$ на Ω .*

Доказательство. Существование предельных точек следует из ограниченности множества Ω . Пусть $x_{k_s} \rightarrow \bar{x} \in \Omega$. Так как $f_{k_s}(x_{k_s}) \geq B(x_{k_{s_0}}, x_{k_s}) = f(x_{k_{s_0}}) + (v(x_{k_{s_0}}), x_{k_s} - x_{k_{s_0}})$ при $k_{s_0} < k_s$, то, переходя к пределу при $k_s, k_{s_0} \rightarrow \infty$, при условии $k_s > k_{s_0}$ и учитывая ограниченность v_{k_s} и (8.14), получим

$$f^* + \varepsilon \geq f(\bar{x}) \geq f^*, \quad (8.18)$$

а следовательно, $\bar{x}$ — ε -стационарная точка $f(x)$ на Ω . ■

Замечание 4. Если множество Ω задано линейными неравенствами, то задача (8.13) есть задача линейного программирования.

Замечание 5. При $\varepsilon = 0$ получим «чистый» метод Келли.

Замечание 6. В пп. 1, 2 выпуклость множества Ω не предполагалась. Требуется лишь уметь решать задачу вида (8.2).

§ 9. Минимизация функции супремума

Пусть X — линейное нормированное пространство, $G \subset E_p$ — ограниченное множество, $\Omega \subset X$. На $\Omega \times G$ задана непрерывная функция $\varphi(x, y)$, причем

$$-\infty < \inf_{x \in \Omega} \sup_{y \in G} \varphi(x, y) = f^* < +\infty.$$

Положим

$$f(x) = \sup_{y \in G} \varphi(x, y). \quad (9.1)$$

Требуется найти $\inf_{x \in \Omega} f(x)$.

Предположим, что существует конечное множество $\sigma_0 \in G$ такое, что

$$\inf_{x \in \Omega} \max_{y \in \sigma_0} \varphi(x, y) > -\infty. \quad (9.2)$$

Пусть функция $\varphi(x, y)$ удовлетворяет условию Липшица по y , т. е. существует $L < \infty$ такое, что

$$|\varphi(x, y') - \varphi(x, y'')| \leq L \|y' - y''\| \quad (9.3)$$

$$\forall x \in \Omega, \quad y', y'' \in G.$$

Лемма 9.1. Если $\{x_k\}$ ($x_k \in \Omega \forall k$) — произвольная последовательность, а последовательность $\{y_k\}$ ($y_k \in G \forall k$) такова, что

$$\varphi(x_k, y_{k+1}) \geq \max_{0 \leq i \leq k} \varphi(x_k, y_i), \quad (9.4)$$

то

$$\lim_{k \rightarrow \infty} [\varphi(x_k, y_{k+1}) - \max_{0 \leq i \leq k} \varphi(x_k, y_i)] = 0. \quad (9.5)$$

Доказательство. Предположим, что (9.5) не имеет места; тогда из (9.4) следует, что найдутся последовательность индексов $\{k_s\}$ и число $\varepsilon_0 > 0$ такие, что

$$\varphi(x_{k_s}, y_{k_s+1}) - \max_{0 \leq i \leq k_s} \varphi(x_{k_s}, y_i) \geq \varepsilon_0.$$

Отсюда и из (9.3), (9.4)

$$\varepsilon_0 \leq \varphi(x_{k_s}, y_{k_s+1}) - \varphi(x_{k_s}, y_i) \leq L \|y_{k_s+1} - y_i\| \quad \forall i \in 0 : k_s,$$

т. е.

$$\|y_{k_s+1} - y_i\| \geq L^{-1} \varepsilon_0 = \varepsilon \quad \forall i \in 0 : k_s,$$

что противоречит ограниченности множества G (всякое ограниченное множество конечномерного пространства можно покрыть лишь конечным числом шаров заданного радиуса ε). Лемма доказана.

Выберем и зафиксируем $\varepsilon \geq 0$, $\delta \geq 0$.

Возьмем произвольное множество $\sigma_0 \in G$, удовлетворяющее (9.2). Пусть уже найдено $\sigma_k \in G$. Построим функцию

$$\varphi_k(x) = \max_{y \in \sigma_k} \varphi(x, y)$$

и найдем точку $x_k \in \Omega$ такую, что

$$\varphi_k(x_k) \leq \inf_{y \in \Omega} \varphi_k(x) + \delta, \quad (9.6)$$

т. е. инфимум функции $\varphi_k(x)$ находим с δ -точностью. Теперь найдем точку $y_{k+1} \in G$ такую, что

$$\varphi(x_k, y_{k+1}) \geq f(x_k) - \varepsilon, \quad (9.7)$$

$$\varphi(x_k, y_{k+1}) \geq \varphi_k(x_k), \quad (9.8)$$

т. е. задачу (9.1) решаем с ε -точностью.

Полагаем $\sigma_{k+1} = \sigma_k \cup \{y_{k+1}\}$ и далее продолжаем аналогично.

Теорема 9.1. *Имеют место соотношения*

$$0 \leq \lim_{k \rightarrow \infty} f(x_k) - f^* \leq \varepsilon + \delta,$$

$$0 \leq \overline{\lim}_{k \rightarrow \infty} f(x_k) - f^* \leq \varepsilon + \delta.$$

Доказательство. Из (9.6) и (9.7)

$$f^* \leq f(x_k) \leq \varphi(x_k, y_{k+1}) + \varepsilon,$$

$$f^* = \inf_{x \in \Omega} \sup_{y \in G} \varphi(x, y) \geq$$

$$\geq \inf_{x \in \Omega} \max_{y \in \sigma_k} \varphi(x, y) = \inf_{x \in \Omega} \varphi_k(x) \geq \varphi_k(x_k) - \delta.$$

Отсюда

$$\begin{aligned} 0 \leq f(x_k) - f^* &\leq \varphi(x_k, y_{k+1}) + \varepsilon - f^* \leq \\ &\leq \varphi(x_k, y_{k+1}) - \varphi_k(x_k) + \varepsilon + \delta. \end{aligned} \quad (9.9)$$

Из леммы 9.1 следует, что $\varphi(x_k, y_{k+1}) - \varphi_k(x_k) \rightarrow 0$. Поэтому из (9.9) вытекает, что для любого $\mu > 0$ найдется такое $K < \infty$, что

$$0 \leq f(x_k) - f^* \leq \varepsilon + \delta + \mu \quad \forall k > K. \blacksquare$$

Следствие. Если вместо (9.6) и (9.7) точки x_k и y_{k+1} удовлетворяют неравенствам

$$\varphi_k(x_k) \leq \inf_{x \in \Omega} \varphi_k(x) + \delta_k,$$

$$\varphi(x_k, y_{k+1}) \geq f(x_k) - \varepsilon_k,$$

где $\delta_k \rightarrow +0$, $\varepsilon_k \rightarrow +0$, то

$$\lim_{k \rightarrow \infty} f(x_k) = f^*$$

и любая предельная точка последовательности $\{x_k\}$, если таковые существуют, является точкой минимума функции $f(x)$ на множестве Ω .

Замечание 1. Выше функция $\varphi(x, y)$ не предполагалась выпуклой по x , множество Ω тоже не обязательно выпукло.

Замечание 2. Поскольку любая выпуклая функция есть функция максимума (формула (5.11) гл. I),

то изложенный выше метод можно применить и для нахождения минимума выпуклой функции, в результате получится метод Келли (см. § 8).

§ 10. Минимизация выпуклой функции максимума. Метод экстремального базиса

1. Пусть

$$f(x) = \max_{y \in G} \varphi(x, y),$$

где G — компактное множество, функция $\varphi(x, y)$ непрерывна вместе с $\varphi'_x(x, y)$ по x и y на $E_n \times G$ и сильно выпукла по x с константой $m > 0$ при каждом фиксированном $y \in G$, т. е.

$$\varphi(x + \Delta, y) \geq \varphi(x, y) + (\varphi'_x(x, y), \Delta) + m \|\Delta\|^2 \quad (10.1)$$

$$\forall [x, y] \in E_n \times G, \quad \forall \Delta \in E_n.$$

Возьмем любое $x_0 \in E_n$. Положим

$$b = \min_{y \in G} \varphi(x_0, y), \quad K_1 = \max_{y \in G} \|\varphi'_x(x_0, y)\|.$$

Из (4.1) ясно, что

$$\varphi(x, y) \geq b - K_1 \|x - x_0\| + m \|x - x_0\|^2 \quad \forall [x, y] \in E_n \times G.$$

Поэтому

$$\varphi(x, y) > f(x_0), \quad (10.2)$$

если

$$\|x - x_0\| > z, \quad \text{где } z = \left[\frac{1}{2} K_1 + \left(\frac{1}{4} K_1^2 + m (f(x_0) - b) \right)^{\frac{1}{2}} \right] m^{-1}.$$

Положим $K = \max_{x \in S_z(x_0), y \in G} \|\varphi'_x(x, y)\|$. Требуется найти $\min_{x \in E_n} f(x)$.

В § 1 было показано, что для того, чтобы точка $x^* \in E_n$ была точкой минимума функции $f(x)$, необходимо и достаточно, чтобы

$$0 \in \mathcal{L}(x^*), \quad (10.3)$$

где

$$\mathcal{L}(x) = \text{co } \mathcal{H}(x), \quad \mathcal{H}(x) = \{\varphi'_x(x, y) \mid y \in R(x)\},$$

$$R(x) = \{y \in G \mid \varphi(x, y) = f(x)\}.$$

С каждой точкой $x_0 \in E_n$ связано множество $\mathcal{L}(x_0)$. Если $0 \notin \mathcal{L}(x_0)$, то направление

$$g(x_0) = -z(x_0) \|z(x_0)\|^{-1},$$

где $\|z(x_0)\| = \min_{z \in \mathcal{L}(x_0)} \|z\|$, является направлением наискорейшего спуска функции $f(x)$ в точке x_0 . Умение вычислять направление наискорейшего спуска или близкие к нему направления дает возможность разработать численные методы для минимизации (см. § 3). Однако при этом необходимо учитывать все точки множества $R(x)$ и производить минимизацию функции $f(x)$ на луче (в направлении спуска). На практике множество G , если оно содержит бесконечное количество точек, дискретизируют (заменяют конечным множеством точек), но более точная аппроксимация $f(x)$ требует учета большого количества точек множества G . Было бы желательно разработать метод минимизации функции $f(x)$, при реализации которого можно было бы обойтись сравнительно небольшим количеством точек множества G .

Опишем вначале «принципиальный алгоритм» (по терминологии Э. Полака [96]). Выберем произвольный набор из $n+2$ точек множества G :

$$\sigma_0 = \{y_{01}, \dots, y_{0, n+2}\}, \quad y_{0i} \in G \quad \forall i \in 1:(n+2).$$

Набор σ_0 назовем базисом.

Пусть уже построены точки $x_0, \dots, x_{k-1}$ и базис

$$\sigma_k = \{y_{k1}, \dots, y_{k, n+2}\}, \quad y_{ki} \in G \quad \forall i \in 1:(n+2)$$

Определим функцию

$$f_k(x) = \max_{i \in 1:(n+2)} \varphi(x, y_{ki})$$

и найдем точку x_k такую, что

$$f_k(x_k) = \min_{x \in E_n} f_k(x). \quad (10.4)$$

Решение задачи (10.4) проще, чем минимизация функции $f(x)$, ибо для нахождения функции $f_k(x)$ требуется вычислить значения функции $f(x, y)$ в $n+2$ точках. По необходимому условию минимума функции $f_k(x)$

$$0 \in \mathcal{L}_k, \quad (10.5)$$

где

$$\mathcal{L}_k = \text{co } \mathcal{H}_k, \quad \mathcal{H}_k = \{\varphi'_x(x_k, y_{ki}) \mid i \in R_k\},$$

$$R_k = \{i \mid \varphi(x_k, y_{ki}) = f_k(x_k)\}.$$

Так как по теореме Каратеодори (теорема 1.1 гл. I) любая точка множества $\mathcal{L}_k$ может быть представлена в виде выпуклой комбинации не более чем $n+1$ точек множества $\mathcal{H}_k$, то существует по крайней мере один индекс $i_k \in 1:(n+2)$, без которого «можно обойтись», т. е. либо $i_k \notin R_k$, либо начало координат в (10.5) можно сформировать и без вектора $\varphi'_x(x_k, y_{ki_k})$. Пусть $\bar{y}_k \in G$ — такая точка, что

$$\varphi(x_k, \bar{y}_k) = f(x_k) = \max_{y \in G} \varphi(x_k, y), \quad (10.6)$$

т. е. $\bar{y}_k \in R(x_k)$. Построим теперь новый базис $\sigma_{k+1} = \{y_{k+1,1}, \dots, y_{k+1,n+2}\}$ так:

$$y_{k+1,i} = \begin{cases} y_{ki}, & i \notin i_k, \\ \bar{y}_{ki}, & i = i_k. \end{cases}$$

Базис σ_{k+1} отличается от базиса σ_k одной точкой и тоже содержит $n+2$ точки. Если $f(x_k) = f_k(x_k)$, то x_k — точка минимума функции $f(x)$, ибо в этом случае $\varphi(x_k, y_{ki}) = f(x_k)$, т. е. $y_{ki} \in R(x_k)$, и в силу (10.5) выполнено необходимое и достаточное условие (10.3). Если же $f(x_k) > f_k(x_k)$, то, исходя из базиса σ_{k+1} , строим функцию $f_{k+1}(x)$ и далее продолжаем аналогично.

В результате имеем последовательность точек $\{x_k\}$. Если эта последовательность конечна, то последняя полученная точка по построению является точкой минимума функции $f(x)$. Если же последовательность $\{x_k\}$ бесконечна, то справедлива

Теорема 10.1. *Последовательность точек $\{x_k\}$ сходится к точке минимума функции $f(x)$.*

Доказательство. Отметим прежде всего, что последовательность $\{x_k\}$ ограничена, ибо из (10.2) следует, что $f_k(x) > f(x_0) \geq f_k(x_0)$, если $\|x - x_0\| > z$. Поэтому

$$\|x_k - x_0\| \leq z, \quad (10.7)$$

поскольку $f_k(x_k) = \min_{x \in E_n} f_k(x) \leq f_k(x_0)$,

Обозначим $\Delta_k = x_{k+1} - x_k$. Из (10.1) имеем

$$\begin{aligned} \varphi(x_{k+1}, y_{ki}) &= \varphi(x_k + \Delta_k, y_{ki}) \geq \\ &\geq \varphi(x_k, y_{ki}) + (\varphi'_x(x_k, y_{ki}), \Delta_k) + m \|\Delta_k\|^2. \end{aligned} \quad (10.8)$$

Для $i \in R_k$ будет $\varphi(x_k, y_{ki}) = f_k(x_k)$.

В силу (10.5) существует такое $y_{ki} \in \sigma_k$, что

$$(\varphi'_x(x_k, y_{ki}), \Delta_k) \geq 0. \quad (10.9)$$

Из (10.8) для такого i имеем

$$\varphi(x_{k+1}, y_{ki}) \geq f_k(x_k) + m\Delta_k^2,$$

т.е.

$$f_{k+1}(x_{k+1}) = \max_i \varphi(x_{k+1}, y_{k+1, i}) \geq f_k(x_k) + m\Delta_k^2. \quad (10.10)$$

Отсюда следует, что

$$\Delta_k \xrightarrow[k \rightarrow \infty]{} 0. \quad (10.11)$$

Действительно, допустим противное. Пусть существуют число $a > 0$ и подпоследовательность $\{k_s\}$ такие, что $\|\Delta_{k_s}\| \geq a$. Тогда из (10.10) имеем $f_k(x_k) \rightarrow +\infty$, но $f(x_k) \geq f_k(x_k)$, поэтому $f(x_k) \rightarrow +\infty$, что невозможно, так как $\{x_k\}$ — ограниченная последовательность, а $f(x)$ — непрерывная функция.

Итак, (10.11) доказано.

Из (10.6) имеем

$$f(x_k) = f_{k+1}(x_k). \quad (10.12)$$

Так как

$$\begin{aligned} \varphi(x_{k+1}, y) &= \varphi(x_k + \Delta_k, y) = \\ &= \varphi(x_k, y) + (\varphi'_x(x_k + \xi_k \Delta_k, y), \Delta_k), \end{aligned}$$

где $\xi_k \in (0, 1)$, то

$$\begin{aligned} |f(x_{k+1}) - f(x_k)| &\leq K \|\Delta_k\|, \\ |f_{k+1}(x_{k+1}) - f_{k+1}(x_k)| &= K \|\Delta_k\| \end{aligned}$$

и из (10.12) получаем

$$\begin{aligned} |f_{k+1}(x_{k+1}) - f(x_{k+1})| &\leq \\ &\leq |f_{k+1}(x_{k+1}) - f_{k+1}(x_k)| + |f_{k+1}(x_k) - f(x_{k+1})| = \\ &= |f_{k+1}(x_{k+1}) - f_{k+1}(x_k)| + |f(x_k) - f(x_{k+1})| \leq 2K \|\Delta_k\|. \end{aligned} \quad (10.13)$$

Из (10.5) получим $\mathbb{O} \in \text{co} \{ \varphi'_x(x_{k+1}, y_{k+1}, i) \mid i \in R_k \}$, где $R_k = \{ i \mid \varphi(x_{k+1}, y_{k+1}, i) = f_{k+1}(x_{k+1}) \}$. Отсюда и из (10.13)

$$\mathbb{O} \in \mathcal{L}_{\varepsilon_k}(x_{k+1}),$$

где $\varepsilon_k = 2K \|\Delta_k\|$, $\mathcal{L}_\varepsilon(x) = \text{co } \mathcal{H}_\varepsilon(x)$,

$$\mathcal{H}_\varepsilon(x) = \{ \varphi'_x(x, y) \mid y \in G: f(x) - \varphi(x, y) \leq \varepsilon \}.$$

Значит, x_{k+1} является ε_k -стационарной (см. § 1) точкой функции $f(x)$ на E_n .

Так как $\varepsilon_k \rightarrow 0$, то последовательность ε_k -стационарных точек сходится к единственной точке минимума функции $f(x)$ на E_n . ■

Выше предполагалось, что функция $\varphi(x, y)$ сильно выпукла по x . Это условие существенно, что видно из следующего примера.

Пример 1. Пусть $x \in E_1$, $G = 1:4$,

$$\varphi(x, 1) \equiv 1, \quad \varphi(x, 2) = x^2, \quad \varphi(x, 3) = 2 - x, \quad \varphi(x, 4) = 2 + x.$$

Положим

$$f(x) = \max_{y \in G} \varphi(x, y) = \max_{i=1:4} \varphi(x, i).$$

Так как $n = 1$, то базис должен состоять из трех точек. Положим $\sigma_0 = \{1, 2, 4\}$. Ясно, что

$$\min_{x \in E_1} f_0(x) = \min_{x \in E_1} \max \{1, x^2, 2 + x\} = f_0(-1) = 1,$$

т. е. $x_0 = -1$. При этом

$$R_0 = \{1, 2, 4\},$$

$$\mathcal{H}_0 = \{ \varphi'_x(-1, 1), \varphi'_x(-1, 2), \varphi'_x(-1, 4) \} = \{0, -2, 1\}.$$

В качестве i_0 можно взять $i_0 = 4$. Поскольку

$$f(x_0) = \max_{y \in 1:4} \varphi(x_0, y) = \varphi(x_0, 3) = 3,$$

то $\bar{y}_0 = 3$, поэтому $\sigma_1 = \{1, 2, 3\}$.

Теперь имеем $f_1(x) = \max \{1, x^2, 2 - x\}$, и поэтому $\min_{x \in E_1} f_1(x) = f_1(+1) = 1$, т. е. $x_1 = 1$. При этом $R_0 = \{1, 2, 3\}$, $\mathcal{H}_0 = \{0, 2, -1\}$. Можно положить $i_1 = 3$. Так

как $f(x_1) = \varphi(x_1, 4) = 3$, то $\bar{y}_1 = 4$. Снова имеем $\sigma_2 = \{1, 2, 4\} = \sigma_0$. Далее аналогично получаем

$$\sigma_3 = \sigma_1, \dots, \sigma_{2k} = \sigma_0, \quad \sigma_{2k+1} = \sigma_1.$$

Таким образом, случилось заикливание.

Очевидно, что $\min f(x) = f(0) = 2$.

Заикливание произошло вследствие нарушения условия сильной выпуклости функций $\varphi(x, y)$ (функции $\varphi(x, 1)$, $\varphi(x, 3)$, $\varphi(x, 4)$ не являются сильно выпуклыми).

2. Вспомогательная задача (10.4) в изложенном выше методе, вообще говоря, не решается за конечное число шагов, поэтому практически метод п. 1 неосуществим. Ниже будет изложено еще три метода, два из которых позволяют реально осуществить изложенный выше метод последовательных приближений. Доказательство сходимости можно найти в [20], [33].

Вначале опишем ε -метод. Выберем последовательность $\{\varepsilon_k\}$ такую, что

$$\varepsilon_k \rightarrow +0, \quad \sum_{k=0}^{\infty} \varepsilon_k < \infty. \quad (10.14)$$

Возьмем произвольную точку $x_0 \in E_n$ и произвольный $(n+2)$ -точечный базис

$$\sigma_0 = \{y_{01}, \dots, y_{0, n+2}\}, \quad y_{0i} \in G \quad \forall i \in 1:(n+2).$$

Пусть уже построены точка $x_k \in E_n$ и базис

$$\sigma_k = \{y_{k1}, \dots, y_{k, n+2}\}, \quad y_{ki} \in G \quad \forall i \in 1:(n+2).$$

Положим

$$f_k(x) = \max_{i \in 1:(n+2)} \varphi(x, y_{ki}) \quad (10.15)$$

и найдем точку x_{k+1} такую, что

$$0 \in \mathcal{L}_{k\varepsilon_k}, \quad (10.16)$$

где

$$\mathcal{L}_{k\varepsilon_k} = \text{co } \mathcal{H}_{k\varepsilon_k}, \quad \mathcal{H}_{k\varepsilon_k} = \{\varphi'_x(x_{k+1}, y_{ki}) \mid i \in R_{k\varepsilon_k}\},$$

$$R_{k\varepsilon_k} = \{i \in 1:(n+2) \mid \varphi(x_{k+1}, y_{ki}) \geq f_k(x_{k+1}) - \varepsilon_k\}.$$

Точка x_{k+1} является ε_k -стационарной точкой функции $f_k(x)$.

Теперь строим новый базис σ_{k+1} так же, как в п. 1. Если $f(x_{k+1}) = f_k(x_{k+1})$ и при этом имеет место (10.5), то точка x_{k+1} — точка минимума функции $f(x)$, и процесс прекращается. В противном случае, исходя из точки x_{k+1} и базиса σ_{k+1} , строим функцию $f_{k+1}(x)$ и находим e_{k+1} -стационарную точку этой функции. Если построенная таким образом последовательность точек $\{x_k\}$ конечна, то последняя полученная точка является точкой минимума функции $f(x)$. В противном случае имеет место

Теорема 10.2. *Последовательность точек $\{x_k\}$ сходится к точке минимума функции $f(x)$.*

3. К сожалению, ε -метод, изложенный в п. 2, еще не вполне устраивает нас, ибо нахождение ε -стационарных точек в ряде случаев требует осуществления бесконечной процедуры. От этого недостатка свободен излагаемый ниже ρ -метод.

Пусть последовательность $\{\rho_k\}$ такова, что

$$\rho_k \rightarrow +0, \quad \sum_{k=0}^{\infty} \rho_k < \infty. \quad (10.17)$$

Пусть уже найдены точка x_k и базис σ_k и определена функция $f_k(x)$ (см. (10.15)). Найдем точку x_{k+1} такую, что

$$\min_{z \in \mathcal{L}_k} \|z\| \leq \rho_k.$$

Здесь $\mathcal{L}_k$ — множество, определенное после соотношения (10.5). Новый базис σ_{k+1} строим, как и в п. 1. В результате имеем последовательность $\{x_k\}$.

Теорема 10.3. *Последовательность точек $\{x_k\}$ сходится к точке минимума функции $f(x)$.*

4. Наконец, изложим (ε, ρ) -метод, являющийся комбинацией ε -метода и ρ -метода.

Пусть последовательности $\{\varepsilon_k\}$ и $\{\rho_k\}$ удовлетворяют соотношениям (10.14) и (10.17). В качестве x_{k+1} выбираем точку, для которой

$$\min_{z \in \mathcal{L}_{\varepsilon_k}} \|z\| \leq \rho_k, \quad (10.18)$$

где $\mathcal{L}_{\varepsilon_k}$ — множество, определенное после включения (10.16). Смена базиса происходит, как и выше. Нахождение точки x_{k+1} , удовлетворяющей неравенству (10.18),

производится тоже в конечное число шагов. Для получающейся таким образом последовательности также справедлива теорема о сходимости.

Замечание 1. Изложенные выше методы обобщаются на случай, когда минимизация производится на множестве, заданном неравенствами ([20], [33]).

Замечание 2. Поскольку любая выпуклая функция может быть представлена в виде функции максимума своих опорных функций (т.е. всякая выпуклая функция есть функция максимума линейных функций, а всякое выпуклое множество представимо в виде пересечения полупространств), то описанный выше метод экстремального базиса можно применять для решения задачи минимизации выпуклой функции на выпуклом множестве, причем на каждом шаге необходимо решать задачу линейного программирования небольшой размерности. Для устранения возможного заикливания, вызываемого тем, что линейные функции не являются строго выпуклыми, могут быть разработаны специальные методы (см. [20]).

Итак, в отличие от метода Келли (см. § 8), количество ограничений во вспомогательной задаче линейного программирования не увеличивается.

§ 1. Необходимые и достаточные условия минимума выпуклой функции на выпуклом множестве

1. Пусть $S \subset E_n$ — открытое множество, $\Omega \subset S$ — выпуклое замкнутое множество и на S определена выпуклая функция $f(x)$.

Требуется найти такую точку $x^* \in \Omega$, что

$$f^* \equiv f(x^*) = \inf_{x \in \Omega} f(x). \quad (1.1)$$

В исследовании задачи (1.1) важную роль играют необходимые и достаточные условия минимума функции $f(x)$ на множестве Ω . Напомним их (см. §§ 6, 10 гл. I).

Положим $\gamma(x_0) = \{v \equiv \lambda(z - x_0) \mid \lambda > 0, z \in \Omega\}$.

Множество $\Gamma(x_0) = \gamma(x_0)$ называется *конусом возможных направлений* множества Ω в точке x_0 . Обозначим $\Gamma^+(x_0)$ конус, сопряженный конусу $\Gamma(x_0)$.

Множество $\gamma(x_0)$ есть выпуклый конус, а $\Gamma(x_0)$ и $\Gamma^+(x_0)$ — выпуклые замкнутые конусы, $\emptyset \equiv \gamma(x_0) \cap \Gamma^+(x_0)$ (см. § 3 гл. I).

В § 6 гл. I была доказана следующая

Теорема 1.1. *Для того чтобы в точке $x^* \in \Omega$ функция $f(x)$ достигала своего наименьшего на множестве Ω значения, необходимо и достаточно, чтобы*

$$\partial f(x^*) \cap \Gamma^+(x^*) \neq \emptyset. \quad (1.2)$$

Если воспользоваться понятием условного субдифференциала функции $f(x)$ на множестве Ω в точке x_0 :

$$\partial^\Omega f(x_0) = \{v \in E_n \mid f(z) - f(x_0) \geq (v, z - x_0) \quad \forall z \in \Omega\},$$

то, учитывая соотношение (см. лемму 10.2 гл. I)

$$\partial^\Omega f(x_0) = \partial f(x_0) - \Gamma^+(x_0), \quad (1.3)$$

условие (1.2) можно записать в следующей форме:

$$\emptyset \in \partial^{\Omega} f(x^*). \quad (1.4)$$

Пусть $x_0 \in \Omega$. Положим

$$\begin{aligned} \varphi(x_0) &= \min_{\|g\|=1} \frac{\partial^{\Omega} f(x_0)}{\partial g}, \quad \psi(x_0) = \min_{\|g\|=1, g \in \Gamma(x_0)} \frac{\partial f(x_0)}{\partial g}, \\ d(x_0) &= \min_{v \in \partial^{\Omega} f(x_0)} \|v\|, \quad \rho(x_0) = \min_{v \in [\partial^{\Omega} f(x_0)]_{\text{fr}}} \|v\|. \end{aligned}$$

Если $\emptyset \in \partial^{\Omega} f(x_0)$, то $\rho(x_0)$ есть радиус наибольшего шара с центром в начале координат, содержащегося в $\partial^{\Omega} f(x_0)$; если же $\emptyset \notin \partial^{\Omega} f(x_0)$, то $\rho(x_0) = \min_{v \in \partial^{\Omega} f(x_0)} \|v\| = d(x_0)$, т. е. $\rho(x_0)$ — расстояние от множества $\partial^{\Omega} f(x_0)$ до начала координат.

Условие (1.2) эквивалентно тому, что (см. лемму 6.3 гл. I)

$$\psi(x^*) \geq 0.$$

Лемма 1.1. Пусть $x_0 \in \Omega$. Справедливы соотношения

$$\psi(x_0) = \varphi(x_0), \quad (1.5)$$

$$\psi(x_0) = \begin{cases} -\rho(x_0) = -d(x_0), & \text{если } d(x_0) > 0, \\ \rho(x_0), & \text{если } d(x_0) = 0. \end{cases} \quad (1.6)$$

Доказательство. Применим леммы 6.2 и 6.4 гл. I для множеств $B = \partial f(x_0)$ и $\Gamma^+ = \Gamma^+(x_0)$. Получим требуемое. ■

Возьмем $x_0 \in \Omega$. Если $\psi(x_0) < 0$, $\|v(x_0)\| = \rho(x_0)$, $v(x_0) \in \partial^{\Omega} f(x_0)$, то вектор $g_0 \equiv g(x_0) = -v(x_0) \|v(x_0)\|^{-1}$ есть направление наискорейшего спуска функции $f(x)$ на множестве Ω в точке x_0 и из (1.5) следует (см. § 6 гл. I)

$$-\rho(x_0) = \frac{\partial f(x_0)}{\partial g_0} = \min_{\|g\|=1, g \in \Gamma(x_0)} \frac{\partial f(x_0)}{\partial g} = \min_{\|g\|=1} \frac{\partial^{\Omega} f(x_0)}{\partial g}.$$

Введем обозначения:

$$T_{\eta}(x) = \{v \in -\Gamma^+(x) \mid \|v\| = \eta\},$$

$$\mathcal{L}_{\eta}(x) = \text{co} \{ \partial f(x) \cup T_{\eta}(x) \},$$

где $\eta > 0$ — произвольное фиксированное число.

Заметим, что если $\Gamma^+(x) = \{0\}$, то $\mathcal{L}_\eta(x) = \partial f(x)$.

Лемма 1.2. Для того чтобы выполнялось условие (1.2), необходимо и, если множество Ω имеет внутренние точки, достаточно, чтобы

$$0 \in \mathcal{L}_\eta(x^*). \quad (1.7)$$

Доказательство. Необходимость. Пусть выполнено (1.2). Тогда существуют $v \in \partial f(x^*)$ и $w \in \Gamma^+(x^*)$ такие, что $v = w$. Если $w = 0$, то и $v = 0$, а значит, (1.7) выполнено. Если же $w \neq 0$, то вектор $g = -\eta w \|w\|^{-1} \in T_\eta(x^*)$, и поскольку $v - w = 0$, то

$$\eta w \|w\|^{-1} + g = 0. \quad (1.8)$$

Из (1.8) для $\alpha = \eta \|w\|^{-1} (1 + \eta \|w\|^{-1})^{-1}$ получаем

$$v_\alpha = \alpha v + (1 - \alpha)g = 0.$$

Так как $\alpha \in (0, 1]$, $v \in \partial f(x^*)$, $g \in T_\eta(x^*)$, то $v_\alpha \in \mathcal{L}_\eta(x^*)$, и поэтому (1.7) справедливо.

Достаточность. Если $T_\eta(x^*) = \emptyset$, то из (1.7) следует, что $0 \in \partial f(x^*)$. С другой стороны, $0 \in \Gamma^+(x^*)$, а значит, (1.2) выполнено.

Пусть теперь $T_\eta(x^*) \neq \emptyset$. Из (1.7) вытекает, что найдутся такие $v \in \partial f(x^*)$, $g \in \text{co } T_\eta(x^*)$ и число $\alpha \in [0, 1]$, что

$$\alpha v + (1 - \alpha)g = 0. \quad (1.9)$$

Так как множество Ω имеет внутренние точки, то $0 \notin \text{co } T_\eta(x^*)$. Поэтому в представлении (1.9) число $\alpha > 0$. Тогда в силу (1.9) имеем

$$v = (\alpha - 1)\alpha^{-1}g.$$

Но $(\alpha - 1)\alpha^{-1}g \in \Gamma^+(x^*)$, ибо $\alpha^{-1}(\alpha - 1) < 0$. Следовательно, (1.2) выполняется и в данном случае. Лемма 1.2 доказана.

Пусть $x_0 \in \Omega$ и $\psi(x_0) < 0$. Тогда

$$f(x_0) - f(x) \leq d(x_0) \|x - x_0\| \quad \forall x \in \Omega. \quad (1.10)$$

Положим $d_\eta(x_0) = \min_{v \in \mathcal{L}_\eta(x_0)} \|v\|$.

Для получения оценки $f(x_0) - f^*$ через $d_\eta(x_0)$ установим связь между $d(x_0)$ и $d_\eta(x_0)$.

Пусть $A, B \subset E_n$ — выпуклые компакты. Положим

$$\mathcal{L} = \text{co}\{A \cup B\}, \quad \tilde{d} = \min_{v \in \mathcal{L}} \|v\|,$$

$$\tilde{\rho} = \min_{v \in \mathcal{L}_{\text{fr}}} \|v\|, \quad \tilde{\psi} = \min_{\|g\|=1} \max_{v \in \mathcal{L}} (v, g).$$

Пусть $0 \notin A$. Положим $\Gamma^+ = -\mathcal{K}(A)$, где $\mathcal{K}(A) = \{\lambda v \mid \lambda \geq 0, v \in A\}$. Множество Γ^+ в силу леммы 3.5 гл. I будет замкнутым выпуклым конусом.

Лемма 1.3. Пусть

$$\|v\| \geq a > 0 \quad \forall v \in A, \quad \|v\| \leq b \quad \forall v \in B. \quad (1.11)$$

Тогда

$$\tilde{d} \leq d \leq (1 + 2ba^{-1})\tilde{d}, \quad (1.12)$$

$$\text{где } d = \min_{v \in B - \Gamma^+} \|v\|.$$

Доказательство. Возьмем произвольное $\lambda \in (0, 1)$. Если $\tilde{d} > \lambda a$, то, поскольку $0 \in \Gamma^+$, $B \subset B - \Gamma^+$ и выполнено (1.11), будем иметь

$$d = \min_{v \in B - \Gamma^+} \|v\| \leq \min_{v \in B} \|v\| \leq b \leq b(\lambda a)^{-1} \tilde{d},$$

т.е.

$$d \leq b(\lambda a)^{-1} \tilde{d}. \quad (1.13)$$

Рассмотрим теперь случай, когда $\tilde{d} \leq \lambda a$. Пусть $v_0 \in \mathcal{L}$ и $\|v_0\| = \tilde{d}$. Вектор v_0 можно представить в виде

$$v_0 = \alpha v_1 + (1 - \alpha) v_2, \quad \alpha \in [0, 1], \quad v_1 \in A, \quad v_2 \in B.$$

В силу (1.11) имеем

$$\begin{aligned} \lambda a \geq \tilde{d} &= \|\alpha v_1 + (1 - \alpha) v_2\| \geq \\ &\geq \alpha \|v_1\| - (1 - \alpha) \|v_2\| \geq \alpha a - (1 - \alpha) b, \end{aligned}$$

откуда

$$\alpha \leq (\lambda a + b)(a + b)^{-1} < 1,$$

ибо $\lambda \in (0, 1)$.

Так как $\alpha \in [0, 1)$, то $\alpha(1 - \alpha)^{-1}v_1 \in -\Gamma^+$, и поэтому

$$\begin{aligned} \tilde{d} = \|\alpha v_1 + (1 - \alpha) v_2\| &= (1 - \alpha) \|\alpha(1 - \alpha)^{-1}v_1 + v_2\| \geq \\ &\geq (1 - \alpha) d \geq a(1 - \lambda)(a + b)^{-1} d. \end{aligned}$$

Следовательно, если $\tilde{d} \leq \lambda a$, то

$$d \leq (a + b) [(1 - \lambda) a]^{-1} \tilde{d}. \quad (1.14)$$

Из (1.13), (1.14) получаем

$$d \leq \min_{\lambda \in (0, 1)} \max \{ (a + b) [(1 - \lambda) a]^{-1}, b (\lambda a)^{-1} \} \tilde{d} = \\ = (1 + 2ba^{-1}) \tilde{d}.$$

Докажем теперь неравенство в левой части (1.12). Пусть $d = \|v_1 + v_2\|$, $v_1 \in -\Gamma^+$, $v_2 \in B$. Так как $\Gamma^+ = -\mathcal{K}(A)$, то найдутся такие $w_1 \in A$ и $\lambda_1 \geq 0$, что $v_1 = \lambda_1 w_1$. Тогда

$$d = (1 + \lambda_1) \| \lambda_1 (1 + \lambda_1)^{-1} w_1 + (1 + \lambda_1)^{-1} v_2 \| \geq \\ \geq (1 + \lambda_1) \tilde{d} \geq \tilde{d}.$$

Лемма 1.3 доказана.

Предположим теперь, что множество Ω ограничено (следовательно, $b = \max_{x \in \Omega} \max_{v \in \partial f(x)} \|v\| < \infty$).

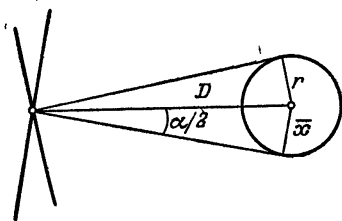


Рис. 54.

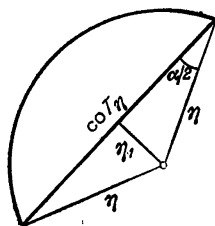


Рис. 55.

Пусть D — диаметр множества Ω , r — радиус наибольшего шара, содержащегося в Ω , $r > 0$.

Тогда для любой точки $x_0 \in \Omega$ сопряженный конус $\Gamma^+(x_0)$ содержится в телесном угле $2\pi - \alpha$, где $\sin(\alpha/2) = D^{-1}r$, и поэтому (рис. 54 и 55)

$$\min_{x \in \Omega} \min_{v \in \text{co } T_\eta(x)} \|v\| \geq \eta r D^{-1}.$$

Отсюда, из (1.10) и леммы 1.3 вытекает

$$0 \leq f(x_0) - f^* \leq D(1 + 2bD(\eta r)^{-1}) d_\eta(x_0). \quad (1.15)$$

2. Рассмотрим случай, когда множество Ω имеет вид

$$\Omega = \{x \in E_n \mid h(x) \leq 0\},$$

где $h(x)$ — выпуклая на E_n функция.

Предположим, что выполнено условие Слейтера, т. е. что для некоторого $\bar{x} \in E_n$ будет $h(\bar{x}) < 0$.

Пусть $x_0 \in \Omega$. В этом случае (см. § 12 гл. I)

$$\Gamma^+(x_0) = \begin{cases} \{0\}, & \text{если } h(x_0) < 0, \\ -\mathcal{K}(\partial h(x_0)), & \text{если } h(x_0) = 0, \end{cases}$$

где $\mathcal{K}(A) = \{w = \lambda v \mid \lambda \geq 0, v \in A\}$.

В частности, если $h(x) = \max_{j \in J} h_j(x)$, $J = 1:m$, $h_j(x)$ — выпуклые непрерывно дифференцируемые на E_n функции, то

$$\partial h(x_0) = \text{co} \{h'_j(x_0) \mid j \in Q(x_0)\},$$

где $Q(x_0) = \{j \in J \mid h_j(x_0) = h(x_0)\}$, и поэтому

$$\begin{aligned} \Gamma^+(x_0) &= \\ &= \begin{cases} \{0\}, & \text{если } h(x_0) < 0, \\ \{v = - \sum_{j \in Q(x_0)} \lambda_j h'_j(x_0) \mid \lambda_j \geq 0 \forall j \in Q(x_0)\}, & \text{если } h(x_0) = 0. \end{cases} \end{aligned}$$

Для любого $x_0 \in \Omega$ введем множество

$$\mathcal{L}(x_0) = \begin{cases} \partial f(x_0), & \text{если } h(x_0) < 0, \\ \text{co} \{\partial f(x_0) \cup \partial h(x_0)\}, & \text{если } h(x_0) = 0. \end{cases}$$

Теорема 1.2. Пусть выполнено условие Слейтера. Для того чтобы в точке $x^* \in \Omega$ функция $f(x)$ достигала своего минимального на множестве Ω значения, необходимо и достаточно, чтобы

$$0 \in \mathcal{L}(x^*). \quad (1.16)$$

Доказательство. Необходимость. Требуется показать, что из (1.2) вытекает (1.16). Пусть выполнено (1.2). Это значит, что нашлись такие $v \in \partial f(x^*)$ и $w \in \Gamma^+(x^*)$, что

$$v - w = 0. \quad (1.17)$$

Если $h(x^*) < 0$, то $\Gamma^+(x^*) = \{0\}$, $w = 0$ и из (1.17) следует, что $v = 0$, т. е. $0 \in \partial f(x^*) \subset \mathcal{L}(x^*)$.

Пусть $h(x^*) = 0$. Поскольку выполнено условие Слейтера, то $\Gamma^+(x^*) = -\mathcal{K}(\partial h(x^*))$. Поэтому найдутся та-

кие $\lambda \geq 0$ и $z \in \partial h(x^*)$, что $w = -\lambda z$, и из (1.17) получим $v + \lambda z = 0$. Следовательно, $(1 + \lambda)^{-1}v + \lambda(1 + \lambda)^{-1}z = 0$. Но $(1 + \lambda)^{-1}v + \lambda(1 + \lambda)^{-1}z \in \mathcal{L}(x^*)$; значит, $0 \in \mathcal{L}(x^*)$, что и требовалось доказать.

Достаточность. Нужно показать, что (1.2) является следствием (1.16). Если $h(x^*) < 0$, то это очевидно.

Пусть $h(x^*) = 0$. Включение (1.16) означает, что найдутся такие $v \in \partial f(x^*)$, $w \in \partial h(x^*)$ и число $\alpha \in [0, 1]$, что

$$\alpha v + (1 - \alpha)w = 0. \quad (1.18)$$

Число α в (1.18) не может быть равно 0, ибо, в силу условия Слейтера, $0 \notin \partial h(x^*)$.

Поделив (1.18) на α , получим

$$v = (\alpha - 1)\alpha^{-1}w,$$

но $(\alpha - 1)\alpha^{-1}w \in \Gamma^+(x^*)$, и, следовательно, $\partial f(x^*) \cap \Gamma^+(x^*) \neq \emptyset$.

Теорема 1.2 доказана.

Пусть $x_0 \in \Omega$. Положим

$$\tilde{\psi}(x_0) = \min_{\|g\|=1} \max_{v \in \mathcal{L}(x_0)} (v, g), \quad \tilde{d}(x_0) = \min_{v \in \mathcal{L}(x_0)} \|v\|,$$

$$\bar{\rho}(x_0) = \min_{v \in [\mathcal{L}(x_0)]_{\text{fr}}} \|v\|.$$

Если $0 \in \mathcal{L}(x_0)$, то $\bar{\rho}(x_0)$ есть радиус максимального шара с центром в 0, содержащегося в $\mathcal{L}(x_0)$, а если $0 \notin \mathcal{L}(x_0)$, то $\bar{\rho}(x_0) = \tilde{d}(x_0)$, т. е. $\bar{\rho}(x_0)$ — расстояние от множества $\mathcal{L}(x_0)$ до начала координат.

Применяя леммы 6.2—6.4 из гл. I для множеств $B = \mathcal{L}(x_0)$, $\Gamma^+ = \{0\}$, получаем, что условие (1.16) эквивалентно тому, что $\tilde{\psi}(x^*) \geq 0$, а также, что

$$\tilde{\psi}(x_0) = \begin{cases} -\bar{\rho}(x_0) = -\tilde{d}(x_0), & \text{если } \tilde{d}(x_0) > 0, \\ \bar{\rho}(x_0), & \text{если } \tilde{d}(x_0) = 0. \end{cases}$$

Пусть $\tilde{\psi}(x_0) < 0$ и $\tilde{v}(x_0) \in \mathcal{L}(x_0)$, $\|\tilde{v}(x_0)\| = \bar{\rho}(x_0)$. Вектор $\tilde{g}_0 = \tilde{g}(x_0) = -\tilde{v}(x_0) \|\tilde{v}(x_0)\|^{-1}$ называется направлением квазинискорейшего спуска функции $f(x)$ на множестве Ω в точке x_0 ([36, стр. 201]).

Пусть $h(x_0) = 0$ и $\tilde{d}(x_0) > 0$. Обратим особое внимание на то, что если $g(x_0)$ и $\tilde{g}(x_0)$ есть направления наи-

скорейшего и квазинайскорейшего спуска функции $f(x)$ на множестве Ω в точке x_0 , то может оказаться, что

$$x_0 + \alpha g(x_0) \notin \Omega \quad \forall \alpha > 0,$$

но всегда найдется такое число $\alpha_0 > 0$, что

$$x_0 + \alpha \tilde{g}(x_0) \in \Omega \quad \forall \alpha \in [0, \alpha_0].$$

Это объясняется тем, что $\frac{\partial h(x_0)}{\partial g(x_0)} \leq 0$, а $\frac{\partial h(x_0)}{\partial \tilde{g}(x_0)} < 0$.

Таким образом, направление $\tilde{g}(x_0)$ является направлением убывания функции $f(x)$ в точке x_0 , и, кроме того, в этом направлении можно двигаться из точки x_0 , оставаясь при этом в множестве Ω . Действительно, справедливо разложение

$$f(x_0 + \alpha \tilde{g}(x_0)) = f(x_0) + \alpha \frac{\partial f(x_0)}{\partial \tilde{g}(x_0)} + o(\alpha), \quad \frac{o(\alpha)}{\alpha} \xrightarrow{\alpha \rightarrow +0} 0.$$

При достаточно малых α будет $\frac{o(\alpha)}{\alpha} \leq \frac{1}{2} \tilde{\rho}(x_0)$. Поэтому при таких α

$$f(x_0 + \alpha \tilde{g}(x_0)) \leq f(x_0) - \frac{1}{2} \alpha \tilde{\rho}(x_0).$$

Аналогично показывается, что при достаточно малых α

$$h(x_0 + \alpha \tilde{g}(x_0)) \leq -\frac{1}{2} \alpha \tilde{\rho}(x_0).$$

Отметим также, что $\tilde{\rho}(x_0) \leq \rho(x_0)$. В самом деле, пусть

$$\rho(x_0) = \|v_0 + \lambda_0 w_0\|,$$

где

$$v_0 \in \partial f(x_0), \quad w_0 \in \partial h(x_0), \quad \lambda_0 \geq 0.$$

Если $\lambda_0 = 0$, то $\tilde{\rho}(x_0) = \rho(x_0)$. Если же $\lambda_0 > 0$, то

$$\begin{aligned} \rho(x_0) &= (1 + \lambda_0) \|v_0 (1 + \lambda_0)^{-1} + \lambda_0 w_0 (1 + \lambda_0)^{-1}\| \geq \\ &\geq (1 + \lambda_0) \tilde{\rho}(x_0). \end{aligned}$$

Из того, что $\tilde{\rho}(x_0) \leq \rho(x_0)$, следует $\frac{\partial f(x_0)}{\partial g(x_0)} \leq \frac{\partial f(x_0)}{\partial \tilde{g}(x_0)}$.

Более того, если $\tilde{\rho}(x_0) = \rho(x_0)$, то $g(x_0) = \tilde{g}(x_0)$, а если $\tilde{\rho}(x_0) < \rho(x_0)$, то $g(x_0) \neq \tilde{g}(x_0)$ и $\frac{\partial f(x_0)}{\partial g(x_0)} < \frac{\partial f(x_0)}{\partial \tilde{g}(x_0)}$.

Предположим теперь, что

$$\max_{v \in \partial f(x_0)} \|v\| \leq b, \quad \min_{v \in \partial h(x)} \|v\| \geq a.$$

Тогда из (1.10) и леммы 1.3 вытекает

$$0 \leq f(x_0) - f(x) \leq \|x - x_0\| (1 + 2ab^{-1}) \tilde{d}(x_0) \quad \forall x \in \Omega. \quad (1.19)$$

Задание 1.1. Пусть $\Omega_1 = \{x \in E_n \mid h_1(x) \leq 0\}$, где $h_1(x) = \max_{j \in J_1} h_{1j}(x)$, $h_{1j}(x) = (A_j, x) + b_j$, $\|A_j\| > 0 \quad \forall j \in J_1 \equiv 1 : N$.

Доказать, что для того, чтобы в точке $x^* \in \Omega \cap \Omega_1$ функции $f(x)$ достигала своего наименьшего на множестве $\Omega \cap \Omega_1$ значения, необходимо и достаточно, чтобы

$$\mathcal{L}(x^*) \cap \Gamma_1^+(x^*) \neq \emptyset,$$

где

$$\mathcal{L}(x) = \begin{cases} \partial f(x), & \text{если } h(x) < 0, \\ \text{co} \{ \partial f(x) \cup \partial h(x) \}, & \text{если } h(x) = 0, \\ \{0\}, & \text{если } h_1(x) < 0, \end{cases}$$

$$\Gamma_1^+(x) = \begin{cases} \{v = - \sum_{j \in Q_1(x)} \lambda_j A_j \mid \lambda_j \geq 0 \quad \forall j \in Q_1(x)\}, & \text{если } h_1(x) = 0, \\ Q_1(x) = \{j \in J_1 \mid h_{1j}(x) = h_1(x)\}. \end{cases}$$

§ 2. ε -стационарные точки

Пусть выпуклая функция $f(x)$ задана на некотором выпуклом открытом множестве, содержащем выпуклое замкнутое множество Ω .

1. Возьмем $x_0 \in \Omega$ и $\varepsilon \geq 0$. Положим

$$\psi_\varepsilon^\Omega(x_0) = \min_{\|g\|=1} \sup_{v \in \partial_\varepsilon^\Omega f(x_0)} (v, g),$$

$$d_\varepsilon^\Omega(x_0) = \min_{v \in \partial_\varepsilon^\Omega f(x_0)} \|v\|,$$

$$\rho_\varepsilon^\Omega(x_0) = \min_{v \in [\partial_\varepsilon^\Omega f(x_0)]_{\text{fr}}} \|v\|, \quad \rho^\Omega(x_0) = \rho_0^\Omega(x_0),$$

$$d^\Omega(x_0) = d_0^\Omega(x_0), \quad \psi^\Omega(x_0) = \psi_0^\Omega(x_0), \quad f^* = \inf_{x \in \Omega} f(x).$$

Ясно, что если $0 \notin \partial_\varepsilon^\Omega f(x_0)$, то $\rho_\varepsilon^\Omega(x_0) = d_\varepsilon^\Omega(x_0)$, а если $0 \in \partial_\varepsilon^\Omega f(x_0)$, то $\rho_\varepsilon^\Omega(x_0)$ есть радиус наибольшего шара с центром в начале координат, содержащегося в $\partial_\varepsilon^\Omega f(x_0)$.

Применив следствие 1 из леммы 6.4 гл. I для множества $B = \partial_\varepsilon^\Omega f(x_0)$, приходим к выводу, что

$$\psi_\varepsilon^\Omega(x_0) = \begin{cases} -\rho_\varepsilon^\Omega(x_0) = -d_\varepsilon^\Omega(x_0), & \text{если } d_\varepsilon^\Omega(x_0) > 0, \\ \rho_\varepsilon^\Omega(x_0), & \text{если } d_\varepsilon^\Omega(x_0) = 0. \end{cases} \quad (2.1)$$

Точку $x_0 \in \Omega$, в которой $\psi_\varepsilon^\Omega(x_0) \geq 0$ или, что то же самое, $0 \leq f(x_0) - f^* \leq \varepsilon$, назовем ε -стационарной точкой функции $f(x)$ на множестве Ω .

Очевидными обобщениями лемм 1.2 и 1.3 гл. III являются

Лемма 2.1. Для любого $x_0 \in \Omega$ имеет место оценка

$$f(y) - f(x_0) \geq \|y - x_0\| \psi_\varepsilon^\Omega(x_0) - \varepsilon \quad \forall y \in \Omega. \quad (2.2)$$

Лемма 2.2. Пусть $x_\varepsilon \in \Omega$, $\varepsilon > 0$ и $\psi_\varepsilon^\Omega(x_\varepsilon) > 0$. Тогда $f^* > -\infty$, множество $\mathcal{D}^* = \{x \in \Omega \mid f(x) = f^*\}$ есть непустой выпуклый компакт,

$$0 \leq f(x_\varepsilon) - f^* < \varepsilon$$

и, если $f(x_\varepsilon) > f^*$, то

$$\frac{f(x_\varepsilon) - f^*}{\rho_\varepsilon^\Omega(x_\varepsilon)} \leq \|x_\varepsilon - x^*\| \leq \frac{f^* - f(x_\varepsilon) + \varepsilon}{\rho_\varepsilon^\Omega(x_\varepsilon)}, \quad (2.3)$$

где x^* — точка минимума функции $f(x)$ на множестве Ω .

Пусть функция $f(x)$ сильно выпукла на Ω , m — константа ее сильной выпуклости. Так же, как и в лемме 9.6 гл. I, показывается, что для любого $x_0 \in \Omega$

$$f(x_0) - f(x^*) \leq (4m)^{-1} (d_\varepsilon^\Omega(x_0) + 2\sqrt{\varepsilon m})^2, \quad (2.4)$$

$$\|x_0 - x^*\| \leq (2m)^{-1} (d_\varepsilon^\Omega(x_0) + 2\sqrt{\varepsilon m}), \quad (2.5)$$

$$f(x_0) - f(x^*) \leq (4m)^{-1} [d_\varepsilon^\Omega(x_0)]^2, \quad (2.6)$$

$$\|x_0 - x^*\| \leq (2m)^{-1} d_\varepsilon^\Omega(x_0). \quad (2.7)$$

2. Пусть $f(x)$ и $h(x)$ — выпуклые на E_n функции. Зафиксируем $x_0 \in E_n$. Допустим, что существует такая точка $\bar{x} \in E_n$, что

$$h(\bar{x}) \leq h(x_0) - a, \quad (2.8)$$

где $a > 0$.

Лемма 2.3. Если при некоторых $\varepsilon \geq 0$ и $\mu \in [0, a)$ оказалось, что

$$\emptyset \in \text{co} \{ \partial_\varepsilon f(x_0) \cup \partial_\mu h(x_0) \}, \quad (2.9)$$

то

$$f(x_0) \leq f_{\mu-h(x_0)}^* + \varepsilon, \quad (2.10)$$

где

$$f_{\mu-h(x_0)}^* = \inf_{x \in \Omega_{\mu-h(x_0)}} f(x),$$

$$\Omega_{\mu-h(x_0)} = \{x \in E_n \mid h(x) \leq h(x_0) - \mu\}.$$

Доказательство. Из (2.9) вытекает существование $v_0 \in \partial_\varepsilon f(x_0)$, $w_0 \in \partial_\mu h(x_0)$, $\alpha_1 \in [0, 1]$ таких, что

$$\alpha_1 v_0 + (1 - \alpha_1) w_0 = \emptyset. \quad (2.11)$$

При этом $\alpha_1 > 0$, ибо в противном случае $w_0 = \emptyset$, что невозможно в силу (2.8) (если $w_0 = \emptyset$, то это противоречит (2.8)). Поэтому (2.11) можно переписать в виде

$$v_0 + \alpha_0 w_0 = \emptyset, \quad (2.12)$$

где $\alpha_0 = (1 - \alpha_1) \alpha_1^{-1}$.

В силу определения $\partial_\varepsilon f(x_0)$ и $\partial_\mu h(x_0)$ имеем

$$f(x) - f(x_0) \geq (v_0, x - x_0) - \varepsilon \quad \forall x \in E_n,$$

$$h(x) - h(x_0) \geq (w_0, x - x_0) - \mu \quad \forall x \in E_n.$$

Отсюда и из (2.12)

$$\begin{aligned} f(x) - f(x_0) + \alpha_0 (h(x) - h(x_0)) &\geq (v_0 + \alpha_0 w_0, x - x_0) - \\ &- (\varepsilon + \alpha_0 \mu) = -(\varepsilon + \alpha_0 \mu) \quad \forall x \in E_n. \end{aligned} \quad (2.13)$$

Для $x \in \Omega_{\mu-h(x_0)}$ будет $h(x) \leq h(x_0) - \mu$. Тогда из (2.13)

$$f(x) - f(x_0) \geq -\varepsilon \quad \forall x \in \Omega_{\mu-h(x_0)},$$

а это и означает, что

$$f(x_0) \leq f_{\mu-h(x_0)}^* + \varepsilon. \quad \blacksquare$$

Следствие. *Предположим, что существует такая точка $\bar{x} \in E_n$, что $f(\bar{x}) \leq f(x_0) - b$, причем $b > \varepsilon$.*

Если выполнено (2.9), то

$$h(x_0) \leq h_{\varepsilon-f(x_0)}^* + \mu,$$

где $h_{\varepsilon-f(x_0)}^* = \inf \{h(x) \mid f(x) \leq f(x_0) - \varepsilon\}$.

§ 3. Метод условного градиента

В этом параграфе будем предполагать, что $\Omega \subset E_n$ — произвольный выпуклый компакт.

1. Сначала получим необходимые и достаточные условия минимума выпуклой функции $f(x)$ на множестве Ω , которые используются в методе условного градиента.

Введем функции

$$\Phi_\varepsilon(x) = \min_{z \in \Omega} \max_{v \in \partial_\varepsilon f(x)} (v, z - x),$$

где $\varepsilon \geq 0$ — произвольное число,

$$\Phi(x) = \Phi_0(x).$$

Заметим, что

$$\Phi_\varepsilon(x) \leq 0 \quad \forall x \in \Omega, \quad \forall \varepsilon \geq 0. \quad (3.1)$$

Теорема 3.1. *Для того чтобы в точке $x^* \in \Omega$ функция $f(x)$ достигала своего наименьшего на множестве Ω значения, необходимо и достаточно, чтобы*

$$\Phi(x^*) = \min_{z \in \Omega} \max_{v \in \partial f(x^*)} (v, z - x^*) = 0.$$

Доказательство. **Необходимость.** Допустим противное; тогда найдется такая точка $z^* \in \Omega$, что (поскольку в силу (3.1) $\Phi(z) \leq 0 \quad \forall z \in \Omega$)

$$\max_{v \in \partial f(x^*)} (v, z^* - x^*) < 0.$$

Положим $g = z^* - x^*$. В силу полунепрерывности сверху отображения $\partial_\varepsilon f(x)$ по ε на $[0, \infty)$, найдется такое $\varepsilon_0 > 0$, что

$$\frac{\partial_{\varepsilon_0} f(x^*)}{\partial g} = \max_{v \in \partial_{\varepsilon_0} f(x^*)} (v, g) < 0.$$

Тогда согласно (8.19) гл. I

$$f(x^*) - \inf_{\alpha > 0} f(x^* + \alpha g) > \varepsilon_0.$$

Следовательно, при достаточно малых $\alpha \in (0, 1]$ будет $x^* + \alpha g \in \Omega$, $f(x^* + \alpha g) < f(x^*)$.

Получили противоречие, ибо x^* есть по предположению точка минимума функции $f(x)$ на множестве Ω .

Достаточность. Пусть $\Phi(x^*) = 0$. Согласно определению $\partial f(x^*)$ для любого $z \in \Omega$ имеем

$$f(z) \geq f(x^*) + \max_{v \in \partial f(x^*)} (v, z - x^*) \geq f(x^*) + \Phi(x^*) = f(x^*),$$

что и требовалось. ■

Положим $f^* = \min_{x \in \Omega} f(x)$.

Лемма 3.1. Для любых $x_0 \in \Omega$ и $\varepsilon \geq 0$ справедлива оценка

$$0 \leq f(x_0) - f^* \leq \varepsilon - \Phi_\varepsilon(x_0). \quad (3.2)$$

Доказательство. Из определения $\partial_\varepsilon f(x_0)$ имеем

$$f(z) \geq f(x_0) + \max_{v \in \partial_\varepsilon f(x_0)} (v, z - x_0) - \varepsilon,$$

$$\min_{z \in \Omega} f(z) \geq f(x_0) + \min_{z \in \Omega} \max_{v \in \partial_\varepsilon f(x_0)} (v, z - x_0) - \varepsilon,$$

откуда и следует (3.2). ■

Следствие. Если $\Phi_\varepsilon(x_0) = 0$, то из (3.2)

$$0 \leq f(x_0) - f^* \leq \varepsilon. \quad (3.3)$$

2. Опишем теперь метод условного градиента для нахождения точек, в которых выполнено (3.3), т. е. для нахождения ε -стационарных точек функции $f(x)$ на множестве Ω .

Зафиксируем число $\varepsilon > 0$. Выберем произвольным образом начальное приближение $x_0 \in \Omega$.

Пусть уже найдена точка $x_k \in \Omega$. Если $\Phi_\varepsilon(x_k) = 0$, то в точке x_k выполняется (3.3), и процесс на этом заканчивается.

Если же $\Phi_\varepsilon(x_k) = \max_{v \in \partial_\varepsilon f(x_k)} (v, z_k - x_k) < 0$, то за $(k+1)$ -е приближение принимаем такую точку $x_{k+1} = x_k + \alpha_k(z_k - x_k)$, что

$$f(x_{k+1}) = \min_{\alpha \in [0, 1]} f(x_k + \alpha(z_k - x_k)).$$

Отметим, что $x_{k+1} \in \Omega$ и $f(x_{k+1}) < f(x_k)$.

Далее процесс последовательных приближений продолжается аналогично.

Если последовательность $\{x_k\}$ конечна, то последняя ее точка по построению является ε -стационарной. Если же последовательность $\{x_k\}$ бесконечна, то справедлива следующая

Теорема 3.2. В любой предельной точке x^* последовательности $\{x_k\}$ будет $\Phi_\varepsilon(x^*) = 0$ и

$$0 \leq f(x^*) - f^* \leq \varepsilon.$$

Доказательство. Допустим противное, предположим, что $x_{k_s} \rightarrow x^*$, но $\Phi_\varepsilon(x^*) = -a < 0$.

Тогда, в силу полунепрерывности сверху отображения $\partial_\varepsilon f(x)$ по x , при достаточно больших k_s окажется

$$\frac{\partial_\varepsilon f(x_{k_s})}{\partial g_{k_s}} = \max_{v \in \partial_\varepsilon f(x_{k_s})} (v, g_{k_s}) = \Phi_\varepsilon(x_{k_s}) \leq -\frac{1}{2} a < 0, \quad (3.4)$$

где $g_{k_s} = z_{k_s} - x_{k_s}$.

Возможны два случая:

$$1) \min_{\alpha \geq 0} f(x_k(\alpha)) = \min_{\alpha \in [0, 1]} f(x_k(\alpha));$$

$$2) \inf_{\alpha \geq 0} f(x_k(\alpha)) < \min_{\alpha \in [0, 1]} f(x_k(\alpha))$$

(здесь $x_k(\alpha) = x_k + \alpha(z_k - x_k)$).

В случае 1) согласно (8.19) гл. I имеем

$$f(x_{k+1}) \leq f(x_k) - \varepsilon.$$

Рассмотрим случай 2). На основании леммы 8.6 гл. I при всех

$$\alpha \in \left[0, \varepsilon \left(\frac{\partial_\varepsilon f(x_k)}{\partial g_k} - \frac{\partial f(x_k)}{\partial g_k} \right)^{-1} \right]$$

будет

$$f(x_k + \alpha g_k) \leq f(x_k) + \alpha \frac{\partial_\varepsilon f(x_k)}{\partial g_k}, \quad (3.5)$$

где $g_k = z_k - x_k$.

Тем более (3.5) будет выполнено при

$$\alpha = -\varepsilon \left(\frac{\partial f(x_k)}{\partial g_k} \right)^{-1} = -\varepsilon (\Phi(x_k))^{-1}.$$

Следовательно, в случае 2)

$$f(x_{k+1}) \leq f(x_k) + \min\{1, -\varepsilon(\Phi(x_k))^{-1}\} \Phi_\varepsilon(x_k).$$

Положим $b = \max_{x \in \Omega} |\Phi(x)|$. Число b конечно в силу того, что отображение $\partial f(x)$ ограничено на Ω . Тогда

$$f(x_{k+1}) \leq f(x_k) + \beta \Phi_\varepsilon(x_k),$$

где

$$\beta = \min\{1, \varepsilon b^{-1}\}.$$

Таким образом, в обоих случаях

$$f(x_{k+1}) \leq f(x_k) + \max\{-\varepsilon, \beta \Phi_\varepsilon(x_k)\} \quad \forall k. \quad (3.6)$$

Из (3.4) и (3.6) заключаем, что при достаточно больших k_s

$$f(x_{k_s+1}) \leq f(x_{k_s}) - \min\{\varepsilon, \frac{1}{2}\beta\alpha\},$$

откуда вытекает, что $f(x_k) \rightarrow -\infty$, а это противоречит ограниченности функции $f(x)$ на множестве Ω . ■

Отметим, что применять метод условного градиента целесообразно тогда, когда множество Ω имеет достаточно простую структуру, например задается линейными неравенствами, и можно эффективно вычислять ε -субдифференциал функции $f(x)$.

З а м е ч а н и е 1. Вместо того, чтобы находить точку x_{k+1} из условия минимума функции $f(x)$ на отрезке $[x_k, z_k]$, можно поступать следующим образом: при $i = 0, 1, 2, \dots$ проверять выполнение неравенства

$$f(x_k + 2^{-i}(z_k - x_k)) \leq f(x_k) - 2^{-i}\Phi_\varepsilon(x_k). \quad (3.7)$$

Пусть при $i = i_k$ первый раз выполнилось (3.7), тогда полагаем

$$x_{k+1} = x_k + 2^{-i_k} (z_k - x_k) \equiv x_k + \alpha_k (z_k - x_k).$$

На основании леммы 8.6 гл. I число i_k конечно, более того,

$$\alpha_k \geq \min\{1, \varepsilon [2(\Phi_\varepsilon(x_k) - \Phi(x_k))]^{-1}\}.$$

Задание 3.1. Доказать теорему 3.2 для метода условного градиента с только что описанным вариантом одномерного поиска.

3. Укажем теперь модификацию метода условного градиента для нахождения точек минимума функции $f(x)$ на множестве Ω .

Зафиксируем некоторые $\varepsilon_0 > 0$ и $\delta_0 > 0$. В качестве начального приближения возьмем произвольную точку $x_0 \equiv x_{00} \in \Omega$. Применяя метод условного градиента с $\varepsilon = \varepsilon_0$, в конечное число шагов (это следует из теоремы 3.2) найдем точку $x_{0t_0} \in \Omega$, в которой

$$\Phi_{\varepsilon_0}(x_{0t_0}) \geq -\delta_0.$$

Положим $\varepsilon_1 = \varepsilon_0/2$, $\delta_1 = \delta_0/2$, $x_1 = x_{10} = x_{0t_0}$.

Взяв в качестве начального приближения x_1 , снова применим метод условного градиента, но уже с $\varepsilon = \varepsilon_1$. В конечное число шагов получим точку x_{1t_1} такую, что $\Phi_{\varepsilon_1}(x_{1t_1}) \geq -\delta_1$.

Продолжая описанный процесс, строим последовательность $\{x_k\}$, для которой

$$\Phi_{\varepsilon_k}(x_{kt_1}) \geq -\delta_k, \quad (3.8)$$

где $\varepsilon_k = \varepsilon_0/2^k$, $\delta_k = \delta_0/2^k$, $k = 0, 1, 2, \dots$

Очевидно, что $f(x_{k+1}) \leq f(x_k)$ и $x_k \in \Omega$ при всех k .

Теорема 3.3. Любая предельная точка последовательности $\{x_k\}$ является точкой минимума функции $f(x)$ на множестве Ω .

Доказательство. Допустим противное, предположим, что $x_{k_s} \rightarrow x^*$, но $\Phi(x^*) = -a < 0$. В силу непрерывности сверху отображения $\partial_\varepsilon f(x)$ по x и ε на $E_n \times [0, \infty)$, найдется такое $\varepsilon > 0$, что при достаточно больших k_s будет

$$\Phi_\varepsilon(x_{k_s}) \leq \frac{1}{2}a. \quad (3.9)$$

Рассмотрим теперь такие k_s , что выполнено (3.9) и, кроме того,

$$\varepsilon_{k_s} \leq \varepsilon, \quad \delta_{k_s} < a/2.$$

Тогда из (3.9) следует

$$\Phi_{\varepsilon_{k_s}}(x_{k_s+1}) \leq -\frac{1}{2} a,$$

что противоречит (3.8). ■

З а м е ч а н и е 2. Можно не требовать ограниченности множества Ω , если вместо функции $\Phi_\varepsilon(x)$ использовать функцию

$$\tilde{\Phi}_\varepsilon(x) = \min_{z \in \Omega, \|z-x\| \leq 1} \max_{v \in \partial_\varepsilon f(x)} (v, z-x).$$

§ 4. Методы наискорейшего спуска для минимизации выпуклых функций

1. В настоящем параграфе рассмотрены методы, связанные с направлением наискорейшего (или близкого к нему) спуска. Вначале обсуждаются непрерывные методы минимизации, или методы дифференциального спуска.

Хотя непрерывные методы минимизации не могут быть непосредственно реализованы на ЭВМ, но на их основе могут быть построены дискретные алгоритмы (один из них изложен ниже, в п. 3).

Пусть функция $f(x)$ конечна и выпукла на E_n . Тогда, как доказано в лемме 5.5 гл. I, функция $f(x)$ является липшицевой на любом выпуклом ограниченном множестве $G \subset E_n$.

Пусть вектор-функция $x(t)$ непрерывно дифференцируема на $[0, T]$, где $T > 0$. Тогда

$$x(t) = x_0 + \int_0^t \dot{x}(\tau) d\tau,$$

где $x_0 = x(0)$, $\dot{x}(\tau)$ — непрерывная на $[0, T]$ вектор-функция. Из (5.26) гл. I для $t \in [0, T]$ будет

$$f(x(t)) = f(x_0) + \int_0^t \frac{\partial f(x(\tau))}{\partial \dot{x}(\tau)} d\tau. \quad (4.1)$$

Здесь интеграл понимается в смысле Лебега.

Пусть $\Omega \subset E_n$ — замкнутое выпуклое множество. Как было установлено в теореме 11.2 гл. I, отображение $\partial_\varepsilon^\Omega f(x)$ непрерывно по ε и x на $(0, \infty) \times \Omega$ по Какутани. Обозначим $v_\varepsilon(x)$ ближайшую к началу координат точку множества $\partial_\varepsilon^\Omega f(x)$, т. е.

$$\|v_\varepsilon(x)\| = \min_{v \in \partial_\varepsilon^\Omega f(x)} \|v\|. \quad (4.2)$$

Лемма 4.1. *Вектор-функция $v_\varepsilon(x)$ непрерывна по ε и x на множестве $(0, \infty) \times \Omega$.*

Доказательство. Пусть $\varepsilon_0 > 0$, $x_0 \in \Omega$. Так как

$$\partial_\varepsilon f(x) \subset \partial_{\varepsilon_0}^\Omega f(x)$$

для любых $\varepsilon > 0$ и $x \in \Omega$, то

$$0 \leq \min_{v \in \partial_\varepsilon^\Omega f(x)} \|v\| \leq \min_{v \in \partial_{\varepsilon_0}^\Omega f(x)} \|v\|.$$

Функция $h(x, \varepsilon) = \min_{v \in \partial_\varepsilon^\Omega f(x)} \|v\|$ ограничена в окрестности точки $[\varepsilon_0, x_0]$. Поэтому и функция

$$H(x, \varepsilon) = \min_{v \in \partial_\varepsilon^\Omega f(x)} \|v\|$$

при $x \in S_\delta(x_0) \cap \Omega$, $\varepsilon \in S_\delta(\varepsilon_0) = \{\varepsilon > 0 \mid |\varepsilon - \varepsilon_0| \leq \delta\}$, где $\delta > 0$, ограничена. Обозначим $v_{\varepsilon_0}(x_0) = v_0$. Пусть $\varepsilon_k \rightarrow \varepsilon_0$, $x_k \rightarrow x_0$. В силу полунепрерывности снизу отображения $\partial_\varepsilon^\Omega f(x)$ существует последовательность $\{v_k\}$ такая, что $v_k \in \partial_{\varepsilon_k}^\Omega f(x_k)$, $v_k \rightarrow v_0$. Значит,

$$\|v_{\varepsilon_k}(x_k)\| = \min_{v \in \partial_{\varepsilon_k}^\Omega f(x_k)} \|v\| \leq \|v_k\|. \quad (4.3)$$

Последовательность $\{v_{\varepsilon_k}(x_k)\}$ ограничена. Поэтому можно считать, что

$$v_{\varepsilon_k}(x_k) \rightarrow \bar{v}. \quad (4.4)$$

Так как отображение $\partial_\varepsilon^\Omega f(x)$ полунепрерывно сверху, то $\bar{v} \in \partial_{\varepsilon_0}^\Omega f(x_0)$. Из (4.2) — (4.4) следует

$$\|\bar{v}\| \leq \|v_0\|. \quad (4.5)$$

Поэтому из (4.2) и (4.5) заключаем, что $\bar{v} = v_0$. Лемма доказана.

Рассмотрим систему дифференциальных уравнений

$$\begin{aligned}\dot{x}(t) &= -v_\varepsilon(t)(x(t)), \\ x(0) &= x_0 \in \Omega,\end{aligned}\quad (4.6)$$

где $v_\varepsilon(t)$ — непрерывная на $[0, \infty)$ функция, причем $v_\varepsilon(t) \xrightarrow{t \rightarrow \infty} +0$. По лемме 4.1 вектор-функция $v_\varepsilon(x)$ непрерывна по ε и x на множестве $(0, \infty) \times \Omega$. По теореме Пеано решение $x(t, x_0)$ системы (4.6) существует, причем $x(t, x_0)$ — непрерывно дифференцируемая функция на $[0, \infty)$. Снова обозначим

$$\|v_\varepsilon(x)\| = \min_{v \in \partial_\varepsilon^\Omega f(x)} \|v\|.$$

Пусть $x \in \Omega$. Если $0 \notin \partial_\varepsilon^\Omega f(x)$, то

$$\begin{aligned}\frac{\partial_\varepsilon^\Omega f(x)}{\partial \dot{x}} &= \sup_{v \in \partial_\varepsilon^\Omega f(x)} (v, \dot{x}) = - \min_{v \in \partial_\varepsilon^\Omega f(x)} (v, -\dot{x}) = \\ &= - \min_{v \in \partial_\varepsilon^\Omega f(x)} (v, v_\varepsilon(x)) = -\|v_\varepsilon(x)\|^2 < 0.\end{aligned}\quad (4.7)$$

Отсюда ясно, что если $x \in \Omega$ не является ε -стационарной точкой функции f на множестве Ω , то

1) направление $\dot{x} = -v_\varepsilon(x) \in \gamma(x) \equiv \{v = \lambda(z - x) \mid \lambda > 0, z \in \Omega\}$, причем $g_\varepsilon = -v_\varepsilon(x) \|v_\varepsilon(x)\|^{-1}$ есть направление ε -наискорейшего спуска функции f в точке $x \in \Omega$ (по множеству Ω). Поэтому существует такое $T = T(x) > 0$, что

$$x + tg_\varepsilon \in \Omega \quad \forall t \in [0, T];$$

2) справедливо неравенство

$$\inf_{\alpha > 0, x + \alpha g_\varepsilon \in \Omega} f(x + \alpha g_\varepsilon) < f(x) - \varepsilon, \quad (4.8)$$

т. е. в направлении g_ε функцию f можно уменьшить не меньше чем на ε , не покидая множество Ω .

Если для некоторой точки $\bar{x} \in \Omega$ оказалось $v_\varepsilon(\bar{x}) = 0$, т. е. $0 \in \partial_\varepsilon^\Omega f(\bar{x})$, то $\bar{x}$ — ε -стационарная точка функции f на множестве Ω .

Положим

$$\mathcal{D}(x_0) = \{x \in \Omega \mid f(x) \leq f(x_0)\}.$$

Теорема 4.1. Если множество $\mathcal{D}(x_0)$ ограничено, то любая предельная точка решения $x(t, x_0)$ системы (4.6) является точкой минимума функции $f(x)$ на множестве Ω .

Доказательство. Поскольку

$$x(t) \equiv x(t, x_0) = x_0 + \int_0^t \dot{x}(\tau) d\tau,$$

то из (4.1) и (4.7) имеем

$$\begin{aligned} f(x(\tau)) &= f(x_0) + \int_0^\tau \frac{\partial f(x(\tau))}{\partial \dot{x}(\tau)} d\tau \leq \\ &\leq f(x_0) + \int_0^\tau \frac{\partial_{\varepsilon}^{\Omega} f(x(\tau))}{\partial \dot{x}(\tau)} d\tau, \end{aligned} \quad (4.9)$$

$$\text{так как } \frac{\partial f(x(\tau))}{\partial \dot{x}(\tau)} = \max_{v \in \partial f(x(\tau))} (v, \dot{x}(\tau)) \leq \sup_{v \in \partial_{\varepsilon}^{\Omega} f(x(\tau))} (v, \dot{x}(\tau)).$$

Из (4.2) и (4.6) следует

$$\sup_{v \in \partial_{\varepsilon}^{\Omega} f(x(\tau))} (v, \dot{x}(\tau)) = -\|v_{\varepsilon(\tau)}(x(\tau))\|^2. \quad (4.10)$$

Поэтому

$$f(x(\tau)) \leq f(x_0) - \int_0^\tau \|v_{\varepsilon(\tau)}(x(\tau))\|^2 d\tau.$$

Отсюда ясно, что $f(x(t)) \leq f(x_0) \quad \forall t \geq 0$, т. е.

$$x(t) \in \mathcal{D}(x_0) \quad \forall t \geq 0.$$

В силу ограниченности множества $\mathcal{D}(x_0)$ предельные точки решения $x(t)$ существуют. Пусть x^* — предельная точка. Это значит, что найдется последовательность $\{t_k\}$ такая, что

$$t_k \rightarrow \infty, \quad x(t_k) \rightarrow x^*.$$

Ясно, что $x^* \in \Omega$. Надо показать, что x^* — точка минимума функции $f(x)$ на Ω , т. е. $0 \in \partial^\Omega f(x^*)$. Допустим противное, пусть $\|v_0(x^*)\| = a > 0$. Из полунепрерывности сверху отображения $\partial^\Omega f(x)$ следует, что найдутся $\varepsilon_0 > 0$ и $\delta > 0$ такие, что $\|v_\varepsilon(x)\| \geq a/2 \quad \forall \varepsilon \in [0, \varepsilon_0], \quad \forall x \in S_\delta(x^*)$. Значит, существует $K_1 < \infty$ такое, что

$$\|v_\varepsilon(t_k)(x(t_k))\| \geq a/2 \quad \forall k \geq K_1.$$

Кроме того, в силу непрерывности $\varepsilon(t)$ и $x(t)$ существуют число $\Delta > 0$ и номер $K_2 \geq K_1$ такие, что

$$\|v_{\varepsilon(t)}(x(t))\| \geq a/4 \quad \forall t \in [t_k - \Delta, t_k + \Delta], \quad \forall k \geq K_2.$$

Без ограничения общности можно считать, что

$$t_{k+1} - t_k > 2\Delta \quad \forall k \geq K_2.$$

Из (4.9) имеем

$$f(x(t)) \leq f(x_0) - \int_0^t \|v_{\varepsilon(\tau)}(x(\tau))\|^2 d\tau \leq f(x_0) - a^2 \Delta m(t)/8, \quad (4.11)$$

где $m(t)$ — такой номер, что $t_{K_2+m(t)} \leq t < t_{K_2+m(t)+1}$. Ясно, что $m(t) \xrightarrow[t \rightarrow \infty]{} \infty$. Поэтому из (4.11) следует $f(x(t)) \xrightarrow[t \rightarrow \infty]{} -\infty$, что противоречит ограниченности непрерывной функции $f(x)$ на компакте $\mathcal{D}(x_0)$. Теорема доказана.

2. Рассмотрим систему

$$\begin{aligned} \dot{x}(t) &= -g_\varepsilon(x(t)), \\ x(0) &= x_0 \in \Omega, \end{aligned} \quad (4.12)$$

где

$$g_\varepsilon(x(t)) = \begin{cases} v_\varepsilon(x) \|v_\varepsilon(x)\|^{-1}, & \text{если } v_\varepsilon(x) \neq 0, \\ 0, & \text{если } v_\varepsilon(x) = 0, \end{cases}$$

$\varepsilon > 0$ фиксировано. Функция в правой части системы (4.12) непрерывна, пока $v_\varepsilon(x) \neq 0$, и по теореме Пеано решение системы существует. Как и выше, можно доказать, что $x(t) \in \mathcal{D}(x_0) \quad \forall t > 0$ и что любая предельная точка решения является ε -стационарной. Более того, справедлива

Теорема 4.2. Если множество $\mathcal{D}(x_0)$ ограничено, а функция $f(x)$ сильно выпукла на E_n , то найдется такое $T < \infty$, что $v_\varepsilon(x(T)) = \emptyset$, т. е. система приходит в ε -стационарную точку за конечное время.

Доказательство. Пусть $f(x)$ — сильно выпуклая на E_n функция с константой сильной выпуклости $m > 0$. По следствию 1 из леммы 17.1 гл. I будет

$$\frac{\partial f(x)}{\partial g} \leq \frac{\partial_\varepsilon f(x)}{\partial g} - 2 \|g\| \sqrt{\varepsilon m}. \quad (4.13)$$

Так как $\partial_\varepsilon f(x) \subset \partial_\varepsilon^2 f(x)$, то из (4.13)

$$\frac{\partial f(x)}{\partial g} \leq \frac{\partial_\varepsilon^2 f(x)}{\partial g} - 2 \|g\| \sqrt{\varepsilon m}. \quad (4.14)$$

Если $v_\varepsilon(x(t)) \neq \emptyset$ (из определения $g_\varepsilon(x(t))$ и системы (4.12) тогда $v_\varepsilon(x(\tau)) \neq \emptyset$ для всех $\tau \in [0, t]$), то из (4.9), (4.10) и (4.14) имеем

$$f(x(t)) \leq f(x_0) - \int_0^t \|v_{\varepsilon(\tau)}(x(\tau))\| d\tau - 2t \sqrt{\varepsilon m}. \quad (4.15)$$

Предположим, что утверждение теоремы неверно; тогда при любом $t > 0$ будет $v_\varepsilon(x(t)) \neq \emptyset$, и в силу (4.15) $f(x(t)) \rightarrow -\infty$, что противоречит ограниченности множества $\mathcal{D}(x_0)$. Теорема доказана.

Замечание 1. Из (4.12) следует, что если $v_\varepsilon(x(T)) = \emptyset$, то $g_\varepsilon(x(t)) = \emptyset$ при $t > T$, и потому $x(t) = x(T)$.

Замечание 2. В системе (4.6) правую часть тоже можно проноормировать.

3. Теперь можно описать следующий дискретный метод минимизации выпуклой функции $f(x)$ на выпуклом замкнутом множестве Ω .

Выберем произвольное $x_0 \in \Omega$. Если $\emptyset \in \partial^2 f(x_0)$, то x_0 — точка минимума функции $f(x)$ на множестве Ω , и процесс прекращается. Если $\emptyset \notin \partial^2 f(x_0)$, то найдем такое число $\varepsilon_0 > 0$, что $\emptyset \in \partial_{\varepsilon_0}^2 f(x_0)$. Если $\emptyset \notin \partial_\varepsilon^2 f(x_0)$ при любом $\varepsilon > 0$, то $f^* \equiv \inf_{x \in \Omega} f(x) = -\infty$. Поэтому в дальнейшем предполагается, что $f^* > -\infty$.

Перебирая числовую последовательность $\varepsilon_{0q} = 2^{-q}$, $q = 1, 2, \dots$, найдем такое $q_0 \geq 1$, что $0 \notin \partial_{\varepsilon_1}^\Omega f(x_0)$, $0 \in \partial_{2\varepsilon_1}^\Omega f(x_0)$, где $\varepsilon_1 = \varepsilon_{0q_0} = \varepsilon_0/2^{q_0}$. Тогда $\varepsilon_1 \leq \frac{1}{2} \varepsilon_0$ и

$$f(x_0) - f^* \leq 2\varepsilon_1. \quad (4.16)$$

Теперь найдем такой вектор $g_0 \in \gamma(x_0)$, $\|g_0\| = 1$, что $\frac{\partial_{\varepsilon_1}^\Omega f(x_0)}{\partial g_0} < 0$, и на множестве $\{x_0(\alpha) = x_0 + \alpha g_0 \mid \alpha \geq 0\} \cap \Omega$ найдем такую точку $x_1 = x_0(\alpha_0)$, что

$$f(x_1) - f(x_0) \leq -\varepsilon_1. \quad (4.17)$$

Из (4.8) следует, что такая точка обязательно найдется. Из (4.16) и (4.17) имеем $f(x_1) \leq f(x_0) - \varepsilon_1 \leq f^* + \varepsilon_1 = f^* + \varepsilon_0/2^{q_k}$. Значит, $0 \in \partial_{\varepsilon_1}^\Omega f(x_1)$, т. е. точка x_1 является ε_1 -стационарной точкой функции $f(x)$ на множестве Ω .

Продолжая аналогично, построим последовательность $\{x_k\}$, которая либо конечна, либо каждая ее точка является ε_k -стационарной точкой функции $f(x)$ на Ω , где $\varepsilon_{k+1} = \varepsilon_k/2^{q_k}$, $q_k \geq 1$, т. е.

$$f(x_{k+1}) - f^* \leq \varepsilon_k/2^{q_k} \leq \varepsilon_0/2^k \quad \forall k. \quad (4.18)$$

Из (4.18) вытекает

Теорема 4.3. Пусть $f^* > -\infty$. Тогда $\lim_{k \rightarrow \infty} f(x_k) = f^*$, причем скорость сходимости $f(x)$ к f^* геометрическая, и если множество $M^* = \{x \in \Omega \mid f(x) = f^*\}$ ограничено, то любая предельная точка последовательности $\{x_k\}$ является точкой минимума функции $f(x)$ на множестве Ω .

§ 5. (ε, μ) -субградиентный метод при наличии ограничений

1. Пусть $G \subset E_n$ — выпуклое замкнутое множество, имеющее внутренние точки.

Возьмем произвольную точку $x_0 \in E_n$. Найдем такую точку $y_0 \equiv y(x_0)$, что

$$\min_{y \in G} \|y - x_0\| = \|y(x_0) - x_0\|. \quad (5.1)$$

Точка $y_0 \equiv y(x_0)$ (она единственна для любого $x_0 \in E_n$, ибо норма в (5.1) евклидова) называется проекцией точки x_0 на G . Отметим, что если для некоторых $x_0 \in G$ и $v_0 \in E_n$ оказалось

$$y(x_0 + v_0) = x_0,$$

то

$$(x - x_0, v_0) \leq 0 \quad \forall x \in G.$$

Поэтому $v_1 = -v_0 \in \Gamma^+(x_0)$. Отсюда, в частности, ясно, что если $v_0 \neq \emptyset$, то x_0 — граничная точка множества G (ибо для внутренних точек будет $\Gamma^+(x_0) = \{\emptyset\}$).

2. Пусть на E_n заданы выпуклые функции $f(x)$ и $h(x)$. Положим

$$\Omega_1 = \{x \in E_n \mid h(x) \leq 0\}, \quad \Omega = G \cap \Omega_1.$$

Требуется найти $\min_{x \in \Omega} f(x) \equiv f^*$.

Возьмем числа $\varepsilon > 0$, $\mu > 0$ и последовательность $\{\lambda_k\}$ такую, что

$$\lambda_k \xrightarrow{k \rightarrow \infty} +0, \quad \sum_{k=0}^{\infty} \lambda_k = \infty. \quad (5.2)$$

Предположим, что существует такая точка $\bar{x} \in \text{int } G$, в которой

$$h(\bar{x}) < -\mu. \quad (5.3)$$

Это, в частности, означает, что $\text{int } \Omega_\mu \neq \emptyset$, где $\Omega_\mu = \{x \in G \mid h(x) \leq -\mu\}$.

Предположим также, что $f^* > -\infty$.

Опишем следующий метод последовательных приближений. Возьмем произвольное $x_0 \in G$. Пусть уже найдено $x_k \in G$.

Если $h(x_k) \leq 0$, то вычислим произвольное $v_k \in \partial_\varepsilon f(x_k)$. Если $v_k = \emptyset$, то

$$0 \leq f(x_k) - \inf_{x \in E_n} f(x) \leq \varepsilon, \quad (5.4)$$

и в этом случае процесс прекращается.

Если $h(x_k) > 0$, то вычислим произвольное $w_k \in \partial_\mu h(x_k)$. В силу (5.3) и леммы 9.2 гл. I обязательно $w_k \neq \emptyset$.

Положим

$$g_k = \begin{cases} -v_k \|v_k\|^{-1}, & \text{если } h(x_k) \leq 0, \\ -w_k \|w_k\|^{-1}, & \text{если } h(x_k) > 0, \end{cases} \quad (5.5)$$

и найдем (см. (5.1))

$$x_{k+1} = y(x_k + \lambda_k g_k). \quad (5.6)$$

Отметим, что если $G = E_n$ или $x_k + \lambda_k g_k \in G$, то $x_{k+1} = x_k + \lambda_k g_k$.

Далее продолжаем аналогично. В результате строим последовательность $\{x_k\}$. Если эта последовательность конечна, то по построению последняя полученная точка x_k принадлежит Ω и удовлетворяет (5.4). Рассмотрим теперь случай, когда последовательность $\{x_k\}$ содержит бесконечное число точек. Образует множество

$$\mathcal{D}_c = \{x \in \Omega \mid f(x) \leq f_\mu^* + \varepsilon + c\}, \quad (5.7)$$

где $f_\mu^* = \inf_{x \in \Omega_\mu} f(x)$.

Поскольку $f^* > -\infty$, то множество $\mathcal{D}_c$ непусто при любом $c \geq 0$.

Теорема 5.1. *При любом $c > 0$ последовательность $\{x_k\}$ содержит бесконечное число точек, принадлежащих множеству $\mathcal{D}_c$.*

Доказательство. Допустим противное. Тогда найдутся такие $c > 0$ и $K < \infty$, что

$$x_k \notin \mathcal{D}_c \quad \forall k \geq K. \quad (5.8)$$

В силу (5.3) $\text{int} \{x \in \Omega_\mu \mid f(x) \leq f_\mu^* + c\} \neq \emptyset$. Поэтому существуют точка $\tilde{x} \in \Omega_\mu$ и число $r > 0$ такие, что

$$x \in G, \quad f(x) \leq f_\mu^* + c, \quad h(x) \leq -\mu \quad \forall x \in S_r(\tilde{x}). \quad (5.9)$$

Если $h(x_k) \leq 0$, то согласно (5.5) $g_k = -v_k \|v_k\|^{-1}$, где $v_k \in \partial_\varepsilon f(x_k)$. Следовательно,

$$f(x) - f(x_k) \geq (v_k, x - x_k) - \varepsilon \quad \forall x \in E_n. \quad (5.10)$$

Из (5.9), (5.7) и того, что $h(x_k) \leq 0$, вытекает, что $f(x_k) > f_\mu^* + \varepsilon + c$. Отсюда и из (5.9), (5.10) имеем

$$\begin{aligned} (v_k, x - x_k) &\leq f(x) + \varepsilon - f(x_k) < \\ &< f_\mu^* + c + \varepsilon - (f_\mu^* + \varepsilon + c) = 0 \quad \forall x \in S_r(\tilde{x}). \end{aligned}$$

Поэтому

$$(g_k, x_k - x) < 0 \quad \forall x \in S_r(\tilde{x}). \quad (5.11)$$

Если $h(x_k) > 0$, то согласно (5.5) $g_k = -w_k \|w_k\|^{-1}$, где $w_k \in \partial_\mu h(x_k)$. Тогда

$$h(x) - h(x_k) \geq (w_k, x - x_k) - \mu \quad \forall x \in E_n$$

и с учетом (5.9)

$$(w_k, x - x_k) \leq h(x) + \mu - h(x_k) < -\mu - 0 + \mu = 0 \\ \forall x \in S_r(\tilde{x}).$$

Поэтому снова имеет место (5.11).

Возьмем точку $x_{1k} = \tilde{x} - r g_k$. Очевидно, что $x_{1k} \in S_r(\tilde{x})$. Из (5.11) вытекает $(g_k, x_k - x_{1k}) < 0$, т. е.

$$(g_k, x_k - \tilde{x}) < -r \quad \forall k. \quad (5.12)$$

В силу (5.1) и (5.6)

$$\|x_{k+1} - \tilde{x}\|^2 = \|y(x_k + \lambda_k g_k) - \tilde{x}\|^2 \leq \|x_k + \lambda_k g_k - \tilde{x}\|^2 \leq \\ \leq \|x_k - \tilde{x}\|^2 + 2\lambda_k (g_k, x_k - \tilde{x}) + \lambda_k^2.$$

Отсюда и из (5.12) следует, что при достаточно больших k будет

$$\|x_{k+1} - \tilde{x}\|^2 \leq \|x_k - \tilde{x}\|^2 - r\lambda_k.$$

Тогда из (5.2) получаем $\|x_k - \tilde{x}\| \xrightarrow[k \rightarrow \infty]{} -\infty$, что невозможно. Теорема доказана.

Следствие 1. Существует такая подпоследовательность $\{x_{k_s}\}$, что

$$x_{k_s} \in G, \quad h(x_{k_s}) \leq 0, \quad \lim_{k_s \rightarrow \infty} f(x_{k_s}) \leq f_u^* + \varepsilon.$$

Доказательство получаем, устремляя в (5.7) $\varepsilon \rightarrow 0$.

Следствие 2. Предположим, что множество

$$\mathcal{D}^* = \{x \in \Omega_\mu \mid f(x) \leq f_u^*\}$$

ограничено. Тогда

$$\rho(x_k, S_a(\mathcal{D}^*)) \xrightarrow[k \rightarrow \infty]{} 0,$$

где $a = \max_{x \in \mathcal{D}} \rho(x, \mathcal{D}^*)$, $\mathcal{D} = \{x \in \Omega \mid f(x) \leq f_u^* + \varepsilon\}$.

Доказательство основывается на тех же рассуждениях, что и доказательство соответствующей части теоремы 4.1 гл. III.

§ 6. Субградиентный метод с постоянным шагом

Пусть $f(x)$ и $h(x)$ — выпуклые на E_n функции и $\Omega = \{x \in E_n \mid h(x) \leq 0\}$. Предположим, что при некотором $\mu_0 > 0$

$$\Omega_{\mu_0} \neq \emptyset, \quad (6.1)$$

где $\Omega_{\mu_0} = \{x \in E_n \mid h(x) \leq -\mu_0\}$.

Рассмотрим следующий метод последовательных приближений для минимизации функции $f(x)$ на множестве Ω .

Зафиксируем числа $\varepsilon \geq 0$, $\lambda > 0$, $\beta > 0$ и $\mu \in [0, \mu_0]$. Выберем произвольное $x_0 \in E_n$. Пусть уже получено $x_k \in E_n$.

Если $h(x_k) \leq \beta$, то вычислим произвольное $v_k \in \partial_\varepsilon f(x_k)$. Если $v_k = 0$ и $h(x_k) \leq 0$, то процесс прекращается. При этом точка $x_k \in \Omega$ и является ε -стационарной точкой функции $f(x)$ на E_n .

Если $h(x_k) > \beta$, то вычислим произвольное $w_k \in \partial_\mu h(x_k)$. Так как $\mu < \mu_0$, то, в силу (6.1) и леммы 9.2 гл. I, $w_k \neq 0$.

Положим

$$g_k = \begin{cases} -v_k \|v_k\|^{-1}, & \text{если } h(x_k) \leq \beta, \\ -w_k \|w_k\|^{-1}, & \text{если } h(x_k) > \beta, \end{cases} \quad (6.2)$$

и возьмем

$$x_{k+1} = x_k + \lambda g_k. \quad (6.3)$$

Далее продолжаем аналогично. В результате строим последовательность $\{x_k\}$. Если эта последовательность конечна, то последняя полученная точка по построению принадлежит множеству Ω и является ε -стационарной точкой функции $f(x)$ на E_n . Рассмотрим случай, когда последовательность $\{x_k\}$ содержит бесконечное число точек,

Положим

$$d_{\mu}^* = \inf_{x \in \Omega_{\mu}} f(x), \quad \mathcal{D}^* = \{x \in \Omega_{\mu} \mid f(x) = f_{\mu}^*\},$$

$$d_{\lambda/2} = \sup_{x \in S_{\lambda/2}(\Omega_{\mu})} h(x) + \mu, \quad b_{\gamma} = \inf_{x \notin S_{\gamma}(\Omega)} h(x),$$

$$\mathcal{D}_c = \{x \in \Omega_{-b_{\gamma}} \mid f(x) \leq c\}, \quad c_{\lambda} = \sup_{x \in S_{\lambda/2}(\mathcal{D}^*)} f(x),$$

$$r_1 = \sup_{x \in \mathcal{D}_{\varepsilon+c_{\lambda}}} \rho(x, \mathcal{D}^*), \quad \Omega_c = \{x = E_n \mid h(x) \leq c\}.$$

Предположим, что множество Ω ограничено. Тогда $\max\{c_{\lambda}, b_{\gamma}, d_{\lambda/2}\} < \infty$.

Теорема 6.1. Пусть $d_{\lambda/2} < \beta \leq b_{\gamma}$. Тогда

$$\lim_{k \rightarrow \infty} \rho(x_k, \mathcal{D}_{\varepsilon+c_{\lambda}}) = 0, \quad (6.4)$$

$$\lim_{k \rightarrow \infty} \rho(x_k, S_{r_1+\lambda}(\mathcal{D}^*)) = 0. \quad (6.5)$$

Доказательство. Установим вначале справедливость (6.4). Допустим противное. Это означает, что найдутся такие $a > 0$ и $K < \infty$, что

$$x_k \notin S_a(\mathcal{D}_{\varepsilon+c_{\lambda}}) \quad \forall k > K. \quad (6.6)$$

Ясно, что существует такое $c > c_{\lambda}$, для которого

$$\mathcal{D}_{\varepsilon+c} \subset S_a(\mathcal{D}_{\varepsilon+c_{\lambda}}). \quad (6.7)$$

Нетрудно увидеть, что

$$S_{\lambda/2}(\mathcal{D}^*) \subset \mathcal{D}_c. \quad (6.8)$$

Выберем такое $r > \lambda/2$, при котором

$$S_r(\bar{x}) \subset \mathcal{D}_c \cap S_r(\Omega_{\mu}), \quad (6.9)$$

$$d_r = \max_{x \in S_r(\Omega_{\mu})} h(x) + \mu \leq \beta, \quad (6.10)$$

где $\bar{x} \in \mathcal{D}^*$. Из (6.8) и того, что $d_{\lambda/2} < \beta$, следует, что такое r обязательно найдется.

Если $h(x_k) \leq \beta$, то в силу (6.2) $g_k = -v_k \|v_k\|^{-1}$, где $v_k \in \partial_{\varepsilon} f(x_k)$. Поэтому

$$f(x) - f(x_k) \geq (v_k, x - x_k) - \varepsilon \quad \forall x \in E_n. \quad (6.11)$$

Если $h(x_k) > \beta$, то в силу (6.2) $g_k = -w_k \|w_k\|^{-1}$, где $w_k \in \partial_\mu h(x_k)$. Поэтому

$$h(x) - h(x_k) \geq (w_k, x - x_k) - \mu \quad \forall x \in E_n. \quad (6.12)$$

Пусть $x_k \in \Omega_{-b\gamma}$. Тогда согласно (6.6) и (6.7) $f(x_k) > \varepsilon + c$, и если $h(x_k) \leq \beta$, то с учетом (6.9) для любого $x \in S_r(\bar{x})$ будет $f(x) - f(x_k) < -\varepsilon$. Отсюда и из (6.11) $(v_k, x - x_k) < 0$ и

$$(g_k, x_k - x) < 0 \quad \forall x \in S_r(\bar{x}), \quad \forall k \geq K. \quad (6.13)$$

Если же $x_k \in \Omega_{-b\gamma}$ и $h(x_k) > \beta$, то с учетом (6.10) для любого $x \in S_r(\bar{x})$ будет $h(x) - h(x_k) < d_r - \mu - \beta \leq -\mu$. Значит, согласно (6.12) $(w_k, x - x_k) < 0$, и снова имеет место (6.13).

Наконец, если $x_k \notin \Omega_{-b\gamma}$, то $h(x_k) > b_\lambda \geq \beta$, и опять будет выполнено (6.13).

Итак, во всех случаях имеет место (6.13).

Положим $\bar{x}_k = \bar{x} - r g_k$. Ясно, что $\bar{x}_k \in S_r(\bar{x})$. Поэтому из (6.13) вытекает $(g_k, x_k - \bar{x}_k) < 0$, т. е.

$$(g_k, x_k - \bar{x}) < -r \quad \forall k \geq K. \quad (6.14)$$

Согласно (6.3)

$$\begin{aligned} \|x_{k+1} - \bar{x}\|^2 &= \|x_k + \lambda g_k - \bar{x}\|^2 = \\ &= \|x_k - \bar{x}\|^2 + 2\lambda (g_k, x_k - \bar{x}) + \lambda^2. \end{aligned}$$

Отсюда и из (6.14) следует, что

$$\|x_{k+1} - \bar{x}\|^2 < \|x_k - \bar{x}\|^2 - 2\lambda r + \lambda^2. \quad (6.15)$$

В силу (6.15) $\|x_k - \bar{x}\| \xrightarrow[k \rightarrow \infty]{} -\infty$, а это невозможно. Полученное противоречие доказывает справедливость (6.4).

Покажем теперь, что справедливо (6.5). Возьмем любое $\delta > 0$. Согласно (6.4) существует бесконечное число таких k , что

$$\rho(x_k, \mathcal{D}_{\varepsilon+c_\lambda}) \leq \delta.$$

Для таких k , поскольку $\|x_{k+1} - x_k\| = \lambda$, будет

$$x_{k+1} \in S_{\lambda+\delta}(\mathcal{D}_{\varepsilon+c_\lambda}) \subset S_{r_1+\lambda+\delta}(\mathcal{D}^*).$$

Если при этом $\rho(x_{k+1}, \mathcal{D}_{\varepsilon+c_\lambda}) > \delta$, то для k , больших некоторого K_δ , окажется выполненным (6.15), и поэтому

$$x_{k+2} \in S_{r_1+\lambda+\delta}(\mathcal{D}^*). \quad (6.16)$$

Если же $\rho(x_{k+1}, \mathcal{D}_{\varepsilon+c\lambda}) < \delta$, то снова будем иметь (6.16), поскольку $\|x_{k+2} - x_{k+1}\| = \lambda$.

Рассуждая далее аналогично, приходим к выводу, что

$$x_k \in S_{r_1+\lambda+\delta}(\mathcal{D}^*) \forall k \geq K_\delta.$$

В силу произвольности δ отсюда вытекает (6.5). ■

Задание 6.1. Пусть $G \subset E_n$ — выпуклое замкнутое множество. Описать субградиентный метод с постоянным шагом и проектированием на множество G для минимизации функции $f(x)$ на множестве $G \cap \{x \in E_n | h(x) \leq 0\}$. Доказать теорему 6.1 для такого метода.

§ 7. Модифицированный (ε, μ) -субградиентный метод при наличии ограничений

Пусть $f(x)$ и $h(x)$ — выпуклые на E_n функции. Требуется найти $f^* = \inf_{x \in \Omega} f(x)$, где $\Omega = \{x \in E_n | h(x) \leq 0\}$.

Предположим, что существуют такие $\bar{x} \in E_n$ и $\gamma > 0$, что

$$h(\bar{x}) = -\gamma < 0. \quad (7.1)$$

Зафиксируем числа $\varepsilon \geq 0$, $d_1 > 0$, $\mu \in [0, \gamma)$, $d_2 \in (0, \gamma - \mu)$ и возьмем последовательность положительных чисел $\{\lambda_k\}$ такую, что

$$\lambda_k \rightarrow 0, \quad \sum_{k=0}^{\infty} \lambda_k = \infty. \quad (7.2)$$

В качестве x_0 возьмем произвольную точку E_n . Пусть уже найдена точка $x_k \in E_n$.

Если $h(x_k) \leq d_1$, то вычислим произвольное $v_k \in \partial_\varepsilon f(x_k)$, и если $v_k = 0$ и $h(x_k) \leq 0$, то на этом процесс прекращается, ибо точка $x_k \in \Omega$ и является ε -стационарной точкой функции $f(x)$ на E_n (и тем более на Ω).

Если $h(x_k) \geq -d_2$, то вычислим произвольное $w_k \in \partial_\mu h(x_k)$. В силу (7.1) и того, что $d_2 + \mu < \gamma$, должно быть $w_k \neq 0$.

Положим

$$z_k = \begin{cases} -\lambda_k v_k \|v_k\|^{-1}, & \text{если } h(x_k) < -d_2, \\ -\lambda_k w_k \|w_k\|^{-1}, & \text{если } h(x_k) > d_1. \end{cases}$$

Если же $-d_2 \leq h(x_k) \leq d_1$, то z_k определим так, чтобы

$$\|z_k\| \leq \lambda_k, \quad \min_{\|z\| \leq \lambda_k} \Phi_k(z) = \Phi_k(z_k), \quad (7.3)$$

где $\Phi_k(z) = \max\{(v_k, z), h(x_k) + (w_k, z)\}$.

Заметим, что если $v_k = 0$, то $z_k = -\lambda_k w_k \|w_k\|^{-1}$.

Если оказалось, что $\|z_k\| < \lambda_k$ и $h(x_k) \leq 0$, то процесс прекращается. В этом случае $\Phi_k(z_k) = \min_{z \in E_n} \Phi_k(z)$ и по

необходимому условию минимума функции $\Phi_k(z)$ на E_n будет $0 \in \text{co}\{v_k, w_k\}$. Тогда $0 \in \text{co}\{\partial_\varepsilon f(x_k) \cup \partial_\mu h(x_k)\}$ и в силу леммы 2.3

$$f(x_k) \leq f_\mu^* + \varepsilon. \quad (7.4)$$

Ясно, что если $\|z_k\| < \lambda_k$ и $h(x_k) > 0$, то $z_k \neq 0$.
Теперь полагаем

$$x_{k+1} = x_k + z_k. \quad (7.5)$$

Далее продолжаем аналогично. В результате строим последовательность $\{x_k\}$. Если она конечна, то по построению последняя полученная точка принадлежит множеству Ω и либо является ε -стационарной точкой функции $f(x)$ на E_n , либо удовлетворяет (7.4). Рассмотрим случай, когда последовательность $\{x_k\}$ содержит бесконечное число точек.

Для любых $c_1 \geq 0$, $c_2 \in [0, d_2]$ положим

$$\Omega_{\mu+c_2} = \{x \in E_n \mid h(x) \leq -(\mu + c_2)\},$$

$$\Omega_{-c_2} = \{x \in E_n \mid h(x) \leq c_2\},$$

$$f_{\mu+c_2}^* = \inf_{x \in \Omega_{\mu+c_2}} f(x),$$

$$\mathcal{D}_{c_1, c_2} = \{x \in \Omega_{-c_2} \mid f(x) \leq f_{\mu+c_2}^* + \varepsilon + c_1\},$$

$$\mathcal{D}_{c_2} = \{x \in \Omega_{-c_2} \mid f(x) \leq f_{\mu+c_2}^* + \varepsilon\} = \mathcal{D}_{0, c_2}.$$

Предположим, что $f^* > -\infty$; тогда множества $\mathcal{D}_{c_1, c_2}$ и $\mathcal{D}_{c_2}$ непусты. Отметим, что так как $c_2 \leq d_2$, то $\text{int } \Omega_{\mu+c_2} \neq \emptyset$.

Предположим также, что

$$\sup_{x \in Q} \sup_{v \in \partial_g f(x) \cup \partial_\mu h(x)} \|v\| < \infty, \quad (7.6)$$

где $Q = \{x \in E_n \mid -d_2 \leq h(x) \leq d_1\}$.

Теорема 7.1. Если найдется такая подпоследовательность индексов $\{k_s\}$, что $\|z_{k_s}\| < \lambda_{k_s}$, то при любом $c_2 \in (0, d_2]$ множеству $\mathcal{D}_{c_2}$ принадлежит бесконечное число точек последовательности $\{x_k\}$. В противном случае при любых $c_1 > 0$, $c_2 \in (0, d_2]$ бесконечное число точек последовательности $\{x_k\}$ принадлежит множеству $\mathcal{D}_{c_1, c_2}$.

Доказательство. Пусть нашлась такая подпоследовательность индексов $\{k_s\}$, что $\|z_{k_s}\| < \lambda_{k_s}$. В этом случае из (7.3) вытекает

$$\Phi_{k_s}(z_{k_s}) = \min_{z \in E_n} \Phi_{k_s}(z)$$

и по необходимому условию минимума функции $\Phi_{k_s}(z)$ на E_n будет

$$0 \in \text{co}\{v_{k_s}, w_{k_s}\} \subset \text{co}\{\partial_g f(x_{k_s}) \cup \partial_\mu h(x_{k_s})\}.$$

Тогда в силу леммы 2.3

$$f(x_{k_s}) \leq f_{\mu-h}^*(x_{k_s}) + \varepsilon. \quad (7.7)$$

На основании (7.2), (7.6) и того, что $\|z_{k_s}\| < \lambda_{k_s}$, заключаем

$$h(x_{k_s}) \xrightarrow{s \rightarrow \infty} 0.$$

Поэтому при достаточно больших k_s будет

$$-h(x_{k_s}) \leq c_2, \quad (7.8)$$

$$\Omega_{\mu+c_2} \subset \Omega_{\mu-h}(x_{k_s}), \quad f_{\mu-h}^*(x_{k_s}) \leq f_{\mu+c_2}^*. \quad (7.9)$$

Из (7.7) и (7.9) получаем

$$f(x_{k_s}) \leq f_{\mu+c_2}^* + \varepsilon. \quad (7.10)$$

Итак, при достаточно больших k_s будут выполнены (7.8) и (7.10), и для рассматриваемого случая утверждение теоремы доказано.

Далее, не ограничивая общности, можем считать, что $\|z_k\| = \lambda_k$ при всех k . Тогда, в силу необходимого условия минимума функции $\Phi_k(z)$ на множестве $\{z \in E_n \mid \|z\| \leq \lambda_k\}$, при каждом k

$$z_k = \lambda_k g_k, \quad g_k = -(\alpha_k v_k + \beta_k w_k), \\ \|g_k\| = 1, \quad \alpha_k \geq 0, \quad \beta_k \geq 0. \quad (7.11)$$

Допустим, что утверждение теоремы неверно и нашлись такие $c_1 > 0$, $c_2 \in [0, d_2]$ и $K < \infty$, что

$$x_k \notin \mathcal{D}_{c_1, c_2} \quad \forall k \geq K. \quad (7.12)$$

Можно считать, что $c_2 \leq d_1$. Рассмотрим множество

$$\mathcal{D}'_{c_1, c_2} = \{x \in \Omega_{\mu+c_2} \mid f(x) \leq f_{\mu+c_2}^* + c_1\}.$$

Так как $c_2 \leq d_2$ и $d_2 + \mu < \gamma$, то и $c_2 + \mu < \gamma$. Поэтому $\text{int } \mathcal{D}'_{c_1, c_2} \neq \emptyset$ и найдутся такие число $r > 0$ и точка $\tilde{x} \in \Omega_{\mu+c_2}$, что $S_r(\tilde{x}) \subset \mathcal{D}'_{c_1, c_2}$. Это значит, что

$$h(x) \leq -(\mu + c_2), \quad f(x) \leq f_{\mu+c_2}^* + c_1 \quad \forall x \in S_r(\tilde{x}). \quad (7.13)$$

Если $h(x_k) \leq c_2$, то в силу (7.12), (7.13) для любого $x \in S_r(\tilde{x})$ будет $f(x) - f(x_k) < f_{\mu+c_2}^* + c_1 - (f_{\mu+c_2}^* + \varepsilon + c_1) = -\varepsilon$. Поэтому

$$(v, x - x_k) < 0 \quad \forall x \in S_r(\tilde{x}), \quad \forall k \geq K, \quad \forall v \in \partial_\varepsilon f(x_k). \quad (7.14)$$

Если $h(x_k) > -c_2$, то в силу (7.13) для любого $x \in S_r(\tilde{x})$ будет $h(x) - h(x_k) < -(\mu + c_2) - (-c_2) = -\mu$. Поэтому

$$(w, x - x_k) < 0 \quad \forall x \in S_r(\tilde{x}), \quad \forall k \geq K, \quad \forall w \in \partial_\mu h(x_k). \quad (7.15)$$

Если $h(x_k) < -c_2$, то нетрудно видеть, что при достаточно больших k , скажем при $k \geq K_1 \geq K$, должно быть $z_k = -\lambda_k v_k \|v_k\|^{-1}$, где $v_k \in \partial_\varepsilon f(x_k)$. Тогда из (7.14) имеем $(v_k, x - x_k) < 0$ и из (7.11)

$$(g_k, x_k - x) < 0 \quad \forall x \in S_r(\tilde{x}). \quad (7.16)$$

Если $h(x_k) > c_2$, то при достаточно больших k , скажем при $k \geq K_2 \geq K_1$, должно быть $z_k = -\lambda_k w_k \|w_k\|^{-1} = \lambda_k g_k$, где $w_k \in \partial_\mu h(x_k)$. Тогда из (7.15) имеем

$(w_k, x - x_k) < 0$, и, следовательно, (7.16) выполнено и в этом случае.

Пусть, наконец, $-c_2 \leq h(x_k) \leq c_2$. Тогда согласно (7.11)

$$g_k = -(\alpha_k v_k + \beta_k w_k), \quad \alpha_k \geq 0, \quad \beta_k \geq 0, \quad \|g_k\| = 1.$$

Подставим v_k в (7.14), а w_k в (7.15) и умножим (7.14) на α_k , а (7.15) на β_k . Сложив полученные неравенства, снова убеждаемся в справедливости (7.16).

Итак, при всех $k \geq K_2$ имеет место (7.16).

Пусть $\tilde{x}_k = \tilde{x} - r g_k$. Ясно, что $\tilde{x}_k \in S_r(\tilde{x})$. Тогда из (7.16) $(g_k, x_k - \tilde{x}_k) < 0$. Отсюда

$$(g_k, x_k - \tilde{x}) < -r \quad \forall k \geq K_2. \quad (7.17)$$

В силу (7.5) и (7.11)

$$\|x_{k+1} - \tilde{x}\|^2 = \|x_k - \tilde{x}\|^2 + 2\lambda_k (g_k, x_k - \tilde{x}) + \lambda_k^2.$$

Отсюда с учетом (7.17) имеем

$$\|x_{k+1} - \tilde{x}\|^2 < \|x_k - \tilde{x}\|^2 - (2r - \lambda_k) \lambda_k.$$

Из (7.2) теперь следует $\|x_k - \tilde{x}\| \xrightarrow{k \rightarrow \infty} -\infty$, что невозможно. ■

Следствие. Если множество $\mathcal{D} = \{x \in \Omega \mid f(x) \leq f_\mu^ + \varepsilon\}$ ограничено, то существует предельная точка последовательности $\{x_k\}$, принадлежащая множеству $\mathcal{D}$.*

Действительно, в этом случае множество $\mathcal{D}_{c_1, c_2}$ ограничено при любых $c_1 \geq 0$, $c_2 \in [0, d_2)$. Поэтому в силу теоремы 7.1 предельные точки последовательности $\{x_k\}$ существуют, а то, что они принадлежат множеству $\mathcal{D}$, вытекает из полунепрерывности сверху отображения $\mathcal{D}_{c_1, c_2}$ по $c_1 \geq 0$, $c_2 \in [0, d_2)$.

Задание 7.1. Построить (ε, μ) -субградиентный метод с проектированием для минимизации функции $f(x)$ на множестве $G \cap \Omega$, где $G \subset E_n$ — произвольный компакт такой, что $\text{int } \{x \in G \mid h(x) \leq -\gamma < 0\}$ не пусто. Доказать теорему о сходимости.

§ 8. Метод негладких штрафных функций

Все методы минимизации выпуклой функции на выпуклом множестве, рассмотренные в предыдущих параграфах настоящей главы, решали непосредственно исход-

ную задачу. Можно, однако, свести исходную задачу условной минимизации к некоторой задаче безусловной минимизации либо к задаче минимизации на множестве простой структуры. Один из способов состоит в использовании методов штрафных функций. В этом параграфе будет показано, как задача минимизации выпуклой функции $f(x)$ на множестве $\Omega = \{x \in E_n | h(x) \leq 0\}$ может быть сведена к решению конечного числа некоторых задач минимизации на всем пространстве E_n .

1. Пусть $f(x)$ и $h(x)$ — выпуклые конечные на E_n функции. Требуется найти

$$f^* = \inf_{x \in \Omega} f(x), \quad (8.1)$$

где $\Omega = \{x \in E_n | h(x) \leq 0\}$.

Предположим, что задача (8.1) разрешима, т. е. что существует точка $x^* \in \Omega$ такая, что $f(x^*) = f^*$, и, кроме того, что множество $\mathcal{M}_0(x^*) = \{x \in E_n | f(x) \leq f^*\}$ ограничено. В силу непрерывности $f(x)$ множество $\mathcal{M}_0(x^*)$ замкнуто.

Положим

$$h^+(x) = \max \{h(x), 0\},$$

$$F(x, \lambda) = f(x) + \lambda h^+(x),$$

$$\Phi(\lambda) = \min_{x \in E_n} F(x, \lambda),$$

$$R(\lambda) = \{x \in E_n | F(x, \lambda) = \Phi(\lambda)\}.$$

Заметим, что $\inf_{x \in E_n} F(x, \lambda)$ достигается при каждом фиксированном $\lambda \geq 0$, поскольку множество $\mathcal{M}_0(x^*)$ ограничено.

Пусть

$$\partial_{x, \varepsilon} F(x, \lambda) = \{v \in E_n | F(z, \lambda) - F(x, \lambda) \geq \geq (v, z - x) - \varepsilon \quad \forall z \in E_n\}.$$

$$\partial_x F(x, \lambda) = \partial_{x, 0} F(x, \lambda), \quad \partial_\lambda F(x, \lambda) = \{h^+(x)\}.$$

В силу леммы 5.2 гл. I

$$\partial_x F(x, \lambda) = \begin{cases} \partial f(x), & \text{если } h(x) < 0, \\ \partial f(x) + \lambda \partial h^+(x), & \text{если } h(x) \geq 0. \end{cases}$$

Здесь использован тот факт, что если $h(x) < 0$, то $\partial h^+(x) = \{\emptyset\}$. Отметим также, что

$$\partial h(x) = \partial h^+(x), \quad \text{если } h(x) > 0,$$

$$\partial h(x) \subset \partial h^+(x), \quad \text{если } h(x) = 0.$$

Справедливы следующие утверждения.

1. Функция $\Phi(\lambda)$ вогнута на $[0, \infty)$.

Это следует из того, что функция $F(x, \lambda)$ линейна по λ , и, следовательно, вогнута, а операция взятия минимума сохраняет вогнутость. ■

$$2. \quad \Phi(\lambda) \leq \min_{x \in \Omega} f(x) \quad \forall \lambda \geq 0. \quad (8.2)$$

Действительно, при любом $\lambda \geq 0$ имеем

$$\Phi(\lambda) = \min_{x \in E_n} F(x, \lambda) \leq \min_{x \in \Omega} F(x, \lambda) = \min_{x \in \Omega} f(x). \quad \blacksquare$$

Положим $\Phi^* = \max_{\lambda \geq 0} \Phi(\lambda)$.

3. Пусть $x_0 \in R(\lambda_0)$, $\lambda_0 \geq 0$. Если $h^+(x_0) = 0$, то

$$f(x_0) = f^* = \Phi(\lambda_0) = \Phi^*. \quad (8.3)$$

В самом деле, поскольку $\Phi(\lambda_0) = f(x_0) + \lambda_0 h^+(x_0) = f(x_0)$, то

$$f(x_0) = \min_{x \in E_n} F(x, \lambda_0) \leq \min_{x \in \Omega} F(x, \lambda_0) = \min_{x \in \Omega} f(x) \leq f(x_0).$$

Отсюда и из (8.2)

$$\Phi^* = \max_{\lambda \geq 0} \Phi(\lambda) \leq f^* = f(x_0) = \Phi(\lambda_0).$$

А это означает, что (8.3) справедливо. ■

4. Пусть $\Phi(\lambda_0) = \Phi^*$, $\lambda_0 \geq 0$. Тогда $\Phi(\lambda) = \Phi(\lambda^*)$ $\forall \lambda \geq \lambda_0$.

Действительно, если $\lambda_1 \geq \lambda_0$, то для любого $x \in E_n$, в силу того что $h^+(x) \geq 0$, должно быть $F(x, \lambda_1) \geq F(x, \lambda_0)$.

Поэтому

$$\Phi^* \geq \Phi(\lambda_1) \geq \Phi(\lambda_0) = \Phi^*. \quad \blacksquare$$

5. Для производной функции $\Phi(\lambda)$ по любому направлению g справедлива формула

$$\frac{\partial \Phi(\lambda)}{\partial g} = \min_{x \in R(\lambda)} (h^+(x), g). \quad (8.4)$$

Действительно (см. § 5 гл. I), так как $F'_\lambda(x, \lambda) = h^+(x)$ и

$$\min_{x \in E_n} F(x, \lambda) = - \max_{x \in E_n} [-F(x, \lambda)],$$

то

$$\frac{\partial \Phi(\lambda)}{\partial g} = - \max_{x \in R(\lambda)} (-h^+(x), g) = \min_{x \in R(\lambda)} (h^+(x), g).$$

На основании формулы (8.4) заключаем, что если при каждом $\lambda > 0$ множество $R(\lambda)$ состоит из одной точки ($R(\lambda) = \{x(\lambda)\}$), то

$$\frac{\partial \Phi(\lambda)}{\partial g} = (h^+(x(\lambda)), g) \quad \forall g,$$

и в этом случае функция $\Phi(\lambda)$ непрерывно дифференцируема по $\lambda > 0$. В частности, $R(\lambda) = \{x(\lambda)\}$, если функция $f(x)$ строго выпукла на E_n . ■

6. Пусть $\Phi(\lambda_0) = \Phi^*$, $\lambda_0 \geq 0$. Тогда для некоторого $x_0 \in R(\lambda_0)$ будет

$$x_0 \in \Omega, \quad f(x_0) = f^*,$$

а при любом $\lambda_1 > \lambda_0$

$$x_1 \in \Omega, \quad f(x_1) = f^* \quad \forall x_1 \in R(\lambda_1).$$

В силу свойства 3 достаточно показать, что

$$\min_{x \in R(\lambda_0)} h^+(x) = 0, \quad \max_{x \in R(\lambda_1)} h^+(x) = 0.$$

Так как согласно свойству 4 $\Phi(\lambda_0 + \alpha) = \Phi(\lambda_0)$ при любом $\alpha \geq 0$, то

$$\Phi'_+(\lambda_0) = \lim_{\alpha \rightarrow +0} \frac{\Phi(\lambda_0 + \alpha) - \Phi(\lambda_0)}{\alpha} = 0,$$

а по свойству 5 $\Phi'_+(\lambda_0) = \min_{x \in R(\lambda_0)} h^+(x)$. Таким образом,

$$\min_{x \in R(\lambda_0)} h^+(x) = 0.$$

Пусть теперь $\lambda_1 > \lambda_0$. Тогда, опять же используя свойства 4 и 5, получаем

$$\begin{aligned} 0 = \Phi'_-(\lambda_1) &= \lim_{\alpha \rightarrow +0} \frac{\Phi(\lambda_1 - \alpha) - \Phi(\lambda_1)}{\alpha} = \\ &= \min_{x \in R(\lambda_1)} [-h^+(x)] = - \max_{x \in R(\lambda_1)} h^+(x). \quad \blacksquare \end{aligned}$$

Достаточные условия существования $\lambda^* \geq 0$ такого, что $\Phi(\lambda^*) = \Phi^*$, сформулированы в следующем свойстве.

7. *Предположим, что задача (8.1) разрешима и выполнено условие Слейтера: при некотором $\bar{x} \in E_n$*

$$h(\bar{x}) = -\gamma < 0. \quad (8.5)$$

Если $x^ \in \Omega$ есть точка минимума функции $f(x)$ на множестве Ω и $\mathbb{O} = v^* + \lambda^* w^*$, где $\lambda^* \geq 0$, $v^* \in \partial f(x^*)$, $w^* \in \partial h(x^*)$, то*

$$x^* \in R(\lambda^*), \quad \Phi(\lambda^*) = \Phi^*$$

и для любых $\lambda_1 > \lambda^$ и $x_1 \in R(\lambda_1)$ будет*

$$x_1 \in \Omega, \quad f(x_1) = f^*. \quad (8.6)$$

Доказательство. Отметим прежде всего, что $\lambda^* < \infty$, ибо выполнено (8.5) и $h^+(x^*) = 0$.

Равенство $\Phi(\lambda^*) = \Phi^*$ будет, в силу свойства 3, выполнено, если показать, что $x^* \in R(\lambda^*)$.

Если $\lambda^* = 0$, то, с одной стороны, $\mathbb{O} = v^* \in \partial f(x^*)$ и поэтому x^* — точка минимума функции $f(x)$ на E_n , а с другой, $F(x, \lambda^*) = f(x)$ и, следовательно, $R(\lambda^*) = \{x \in E_n \mid f(x) = \min_{z \in E_n} f(z)\}$ и $x^* \in R(\lambda^*)$.

Пусть $\lambda^* > 0$. Тогда $h(x^*) = 0$, а значит,

$$\partial h(x^*) \subset \partial h^+(x^*). \quad (8.7)$$

Условие $\mathbb{O} = v^* + \lambda^* w^*$ означает, что $\mathbb{O} \in \partial f(x^*) + \lambda^* \partial h(x^*)$. Отсюда и из (8.7) имеем

$$\mathbb{O} \in \partial_x F(x^*, \lambda^*) = \partial f(x^*) + \lambda^* \partial h^+(x^*),$$

т. е. x^* — точка минимума функции $F(x, \lambda^*)$ на E_n и $x^* \in R(\lambda^*)$.

В силу свойства 6 при любых $\lambda_1 > \lambda^*$ и $x_1 \in R(\lambda_1)$ выполняется (8.6). ■

8. Пусть $0 \in \partial_{x, \varepsilon} F(x_0, \lambda_0)$, $\varepsilon \geq 0$, $\lambda_0 \geq 0$. Тогда

$$\Phi^* \leq F(x_0, \lambda_0) + (h^+(x_0), \lambda^* - \lambda_0) + \varepsilon, \quad (8.8)$$

где λ^* таково, что $\Phi(\lambda^*) = \Phi^* = \max_{\lambda \geq 0} \Phi(\lambda)$.

В самом деле,

$$\Phi(\lambda^*) \leq \Phi(\lambda_0) + (\beta, \lambda^* - \lambda_0) + \varepsilon \quad \forall \beta \in \partial_\varepsilon \Phi(\lambda_0). \quad (8.9)$$

Но (см. § 17 гл. I)

$$\text{co} \{h^+(x) \mid x \in R_\varepsilon(\lambda_0)\} \subset \partial_\varepsilon \Phi(\lambda_0).$$

Так как $0 \in \partial_{x, \varepsilon} F(x_0, \lambda_0)$, то

$$0 \leq F(x_0, \lambda_0) - \min_{x \in \Omega} F(x, \lambda_0) \leq \varepsilon.$$

Это значит, что $0 \leq F(x_0, \lambda_0) - \Phi(\lambda_0) \leq \varepsilon$, т. е. $x_0 \in R_\varepsilon(\lambda_0)$. Тогда $h^+(x_0) \in \partial_\varepsilon \Phi(\lambda_0)$. Теперь (8.8) следует из (8.9). ■

Следствие. Если $0 \in \partial_{x, \varepsilon} F(x_0, \lambda_0)$ и $x_0 \in \Omega$, то

$$\Phi(\lambda^*) \leq F(x_0, \lambda_0) + \varepsilon.$$

9. Если $\lambda_0 \geq \lambda^*$ и $0 \in \partial_{x, \varepsilon} F(x_0, \lambda_0)$, то

$$f(x_0) + \lambda_0 h^+(x_0) - \min_{x \in \Omega} f(x) \leq \varepsilon; \quad (8.10)$$

в частности, для $x_0 \in \Omega$ имеем

$$0 \leq f(x_0) - f^* \leq \varepsilon.$$

Действительно, из того, что $\lambda_0 \geq \lambda^*$, вытекает, что $\Phi(\lambda_0) = \Phi(\lambda^*)$, а поскольку $0 \in \partial_{x, \varepsilon} F(x_0, \lambda_0)$, то

$$f(x_0) + \lambda_0 h^+(x_0) - f^* = F(x_0, \lambda_0) - \Phi(\lambda_0) \leq \varepsilon. \quad \blacksquare$$

2. Перейдем теперь непосредственно к самому методу штрафных функций.

Опишем «принципиальный» алгоритм.

Возьмем произвольное $\lambda_0 > 0$. Каким-либо из методов гл. III найдем точку $x_0 = x(\lambda_0) \in R(\lambda_0)$, т. е. такую, что

$$F(x_0, \lambda_0) = \min_{x \in E_n} F(x, \lambda_0).$$

Если $h(x_0) = 0$, то в силу свойства 3 точка x_0 есть точка минимума функции $f(x)$ на множестве Ω . Если же $h(x_0) > 0$, то положим $\lambda_1 = \lambda_0 + 2\max\{\lambda_0, h^+(x_0)\}$ и найдем точку x_1 , в которой

$$F(x_1, \lambda_1) = \min_{x \in E_n} F(x, \lambda_1).$$

Далее продолжаем аналогично.

Если задача (8.1) разрешима и выполнено условие Слейтера (8.2), то в конечное число шагов будут найдены число $\lambda_k > 0$ и точка $x_k \in E_n$ такие, что $x_k \in \Omega$ (т. е. $h(x_k) \leq 0$) и

$$\Phi(\lambda_k) = \Phi^*, \quad f(x_k) = f^*.$$

Докажем это утверждение. Пусть x^* — точка минимума $f(x)$ на Ω .

Если $h(x^*) < 0$, то $0 \in \partial f(x^*) = \partial_x F(x^*, 0)$ и, значит, $x^* \in R(0)$. В силу свойства 3 $\Phi(0) = \Phi^*$.

Отсюда для точки x_0 , найденной на первом же шаге, будет выполнено

$$x_0 \in \Omega, \quad f(x_0) = f^* = \min_{x \in E_n} f(x)$$

(по свойству 5).

Пусть теперь $h(x^*) = 0$. Тогда в силу необходимого условия минимума существуют $\lambda^* \geq 0$, $v^* \in \partial f(x^*)$, $w^* \in \partial h(x^*)$ такие, что $0 = v^* + \lambda^* w^*$. По свойству 6

$$\Phi(\lambda^*) = \Phi^* = \max_{\lambda \geq 0} \Phi(\lambda),$$

и для первого λ_k , большего λ^* , будет выполнено

$$x_k \in \Omega, \quad f(x_k) = f^*.$$

Более того, пусть $\delta = \lambda^* - \lambda_0$. Поскольку $\lambda_k \geq \lambda_0 + 2k\lambda_0$, то для того, чтобы получить $\lambda_k \geq \lambda^*$, необходимо будет сделать не более $\delta(2\lambda_0)^{-1}$ шагов. ■

З а м е ч а н и е. Все результаты этого параграфа справедливы и для случая, когда

$$\Omega = \{x \in E_n \mid h_j(x) \leq 0 \quad \forall j \in J\},$$

где $h_j(x)$, $j \in J \equiv 1:m$, — выпуклые на E_n функции.

При этом под λ следует понимать вектор $\lambda = (\lambda_1, \lambda_2, \dots, \lambda_m)$ и по определению

$$\{\lambda \geq 0\} = \{\lambda \mid \lambda_j \geq 0 \quad \forall j \in 1:m\},$$

$$\{\lambda > \lambda^*\} = \{\lambda \mid \lambda_j > \lambda_j^* \quad \forall j \in 1:m\},$$

$$h(x) = \max_{j \in J} h_j(x), \quad \lambda h^+(x) = \sum_{j=1}^m \lambda_j h_j^+(x),$$

$$(h^+(x), g) = \sum_{j=1}^m h_j^+(x) g_j.$$

§ 9. Метод Келли для минимизации на выпуклом множестве

Пусть $f(x)$ и $h(x)$ — выпуклые на E_n функции. Требуется найти

$$f^* = \inf_{x \in \Omega} f(x),$$

где $\Omega = \{x \in E_n \mid h(x) \leq 0\}$.

Предположим, что множество Ω ограничено и удовлетворяет условию Слейтера: существует точка $\bar{x} \in E_n$ такая, что

$$h(\bar{x}) < 0. \quad (9.1)$$

Пусть $S \subset E_n$ — произвольный компакт, содержащий Ω .

Зафиксируем $\varepsilon \geq 0$ и $\mu \in [0, -h(\bar{x})]$. Выберем произвольную точку $x_0 \in S$ и положим $\sigma_1 = \{x_0\}$. Вычислим любые $v(x_0) \in \partial_\varepsilon f(x_0)$ и $w(x_0) \in \partial_\mu h(x_0)$. Положим

$$\varphi_1(x) = \max_{z \in \sigma_1} B_f(x, z) = f(x_0) + (v(x_0), x - x_0),$$

$$\begin{aligned} \omega_1 &= \{x \in S \mid B_h(x, z) \leq 0 \quad \forall z \in \sigma_1\} = \\ &= \{x \in S \mid h(x_0) + (w(x_0), x - x_0) \leq 0\}, \end{aligned}$$

где

$$B_f(x, z) = f(z) + (v(z), x - z), \quad v(z) \in \partial_\varepsilon f(z),$$

$$B_h(x, z) = h(z) + (w(z), x - z), \quad w(z) \in \partial_\mu h(z).$$

Пусть уже построены множества σ_{k-1} , ω_{k-1} и функция $\varphi_{k-1}(x)$, и пусть

$$x_{k-1} \in \omega_{k-1}, \quad \varphi_{k-1}(x_{k-1}) = \min_{x \in \omega_{k-1}} \varphi_{k-1}(x).$$

Вычислим произвольные $v(x_{k-1}) \in \partial_{\varepsilon} f(x_{k-1})$ и $w(x_{k-1}) \in \partial_{\mu} h(x_{k-1})$.

Положим

$$\begin{aligned}\sigma_k &= \sigma_{k-1} \cup \{x_{k-1}\}, \\ \varphi_k(x) &= \max_{z \in \sigma_k} B_f(x, z),\end{aligned}\quad (9.2)$$

$$\omega_k = \{x \in S \mid B_h(x, z) \leq 0 \quad \forall z \in \sigma_k\}.$$

и найдем точку $x_k \in \omega_k$, в которой

$$\varphi(x_k) = \min_{x \in \omega_k} \varphi_k(x). \quad (9.3)$$

Из (9.2) вытекает, что

$$\sigma_{k-1} \subset \sigma_k, \quad \omega_k \subset \omega_{k-1}. \quad (9.4)$$

Тогда из (9.3) и (9.4) имеем

$$\begin{aligned}\varphi_{k-1}(x_{k-1}) &= \min_{x \in \omega_{k-1}} \max_{z \in \sigma_{k-1}} B_f(x, z) \leq \\ &\leq \min_{x \in \omega_k} \max_{z \in \sigma_k} B_f(x, z) = \varphi_k(x_k).\end{aligned}\quad (9.5)$$

Если оказалось, что $x_k \in \Omega$ и

$$\varphi_k(x_k) \geq f(x_k), \quad (9.6)$$

то процесс прекращается. В этом случае

$$f(x_k) \leq f_{\mu}^* + \varepsilon, \quad (9.7)$$

где $f_{\mu}^* = \min_{x \in \Omega_{\mu}} f(x)$, $\Omega_{\mu} = \{x \in E_n \mid h(x) \leq -\mu\}$.

Действительно, по определению $\partial_{\varepsilon} f(z)$

$$\begin{aligned}f(x) - f(z) &\geq (v(z), x - z) - \varepsilon \\ \forall x, z &\in E_n, \quad \forall v(z) \in \partial_{\varepsilon} f(z).\end{aligned}$$

Поэтому

$$f(x) + \varepsilon \geq \sup_{z \in E_n} B_f(x, z) \geq \max_{z \in \sigma_k} B_f(x, z) = \varphi_k(x). \quad (9.8)$$

Из определения $\partial_{\mu} h(z)$ имеем

$$\begin{aligned}h(x) - h(z) &\geq (w(z), x - z) - \mu \\ \forall x, z &\in E_n, \quad \forall w(z) \in \partial_{\mu} h(z).\end{aligned}$$

Следовательно,

$$\begin{aligned}\Omega_\mu &= \{x \in S \mid h(x) \leq -\mu\} \subset \\ &\subset \{x \in S \mid B_h(x, z) \leq 0 \quad \forall z \in E_n\} \subset \\ &\subset \{x \in S \mid B_h(x, z) \leq 0 \quad \forall z \in \sigma_k\} = \omega_k.\end{aligned}$$

Отсюда и из (9.8) заключаем, что

$$\begin{aligned}f_\mu^* + \varepsilon &= \min_{x \in \Omega_\mu} f(x) + \varepsilon \geq \min_{x \in \Omega_\mu} \varphi_k(x) \geq \\ &\geq \min_{x \in \omega_k} \varphi_k(x) = \varphi_k(x_k).\end{aligned}\quad (9.9)$$

Теперь (9.7) вытекает из (9.6) и (9.9).

Если (9.7) не имеет места, то строим σ_{k+1} , ω_{k+1} и $\varphi_{k+1}(x)$ и продолжаем аналогично. В результате получаем последовательность $\{x_k\}$. Если она конечна, то последняя ее точка по построению принадлежит множеству Ω и удовлетворяет (9.7). В противном случае справедлива

Теорема 9.1. Если x^* — предельная точка последовательности $\{x_k\}$, то $x^* \in \Omega$ и

$$f(x^*) \leq f_\mu^* + \varepsilon. \quad (9.10)$$

Доказательство. Существование предельных точек последовательности $\{x_k\}$ следует из ограниченности множества S .

Пусть $x_{k_s} \xrightarrow{s \rightarrow \infty} x^*$. В силу (9.2) и того, что $x_{k_{s_0}} \in \sigma_{k_s}$ при $k_s > k_{s_0}$, будет

$$h(x_{k_{s_0}}) + (w(x_{k_{s_0}}), x_{k_s} - x_{k_{s_0}}) \leq 0 \quad \forall k_s > k_{s_0}. \quad (9.11)$$

Поскольку $\partial_\mu h(x)$ ограничено на S , то, переходя в (9.11) к пределу по k_s , $k_{s_0} \rightarrow \infty$ при условии, что $k_s > k_{s_0}$, получим $h(x^*) \leq 0$, т. е. $x^* \in \Omega$.

Последовательность $\{\varphi_k(x_k)\}$ монотонно возрастает и, в силу (9.9), ограничена сверху числом $f_\mu^* + \varepsilon$. При $k_s > k_{s_0}$ будем иметь

$$f_\mu^* + \varepsilon \geq \varphi_{k_s}(x_{k_s}) \geq f(x_{k_{s_0}}) + (v_{k_{s_0}}, x_{k_s} - x_{k_{s_0}}). \quad (9.12)$$

Поскольку $\partial_\varepsilon f(x)$ ограничено на S , то, переходя в (9.12) к пределу по k_s , $k_{s_0} \rightarrow \infty$ при условии, что $k_s > k_{s_0}$, получим (9.10). ■

§ 10. Релаксационный субградиентный метод при наличии ограничений

В этом параграфе часть результатов § 6 гл. III распространяется на задачу минимизации выпуклой функции на выпуклом множестве.

1. Минимизация на произвольном выпуклом замкнутом множестве. Пусть $\Omega \subset E_n$ — выпуклое замкнутое множество и $f(x)$ — выпуклая на некотором выпуклом открытом множестве, содержащем Ω , функция. Требуется найти $\inf_{x \in \Omega} f(x) = f^*$.

Предположим, что множество Ω имеет внутренние точки. Тогда (см. лемму 1.2), для того чтобы в точке x^* функция $f(x)$ достигала своего наименьшего на множестве Ω значения, необходимо и достаточно, чтобы

$$0 \in \mathcal{L}_\eta(x^*),$$

где $\mathcal{L}_\eta(x) = \text{co} \{ \partial f(x) \cup T_\eta(x) \}$, $T_\eta(x) = \{ v \in -\Gamma^+(x) \mid \|v\| = \eta \}$, $\eta > 0$ — произвольное фиксированное число.

Нетрудно убедиться в том, что отображение $T_\eta(x)$ полунепрерывно сверху на Ω . Поэтому, в силу лемм 2.3 и 2.4 гл. I, полунепрерывно сверху отображение $\mathcal{L}_\eta(x)$. Кроме того, оно ограничено на любом ограниченном множестве из Ω , ибо этим свойством обладают $\partial f(x)$ и $T_\eta(x)$.

Положим $d(x) = \min_{v \in \mathcal{L}_\eta(x)} \|v\|$. Опишем релаксационный субградиентный метод для рассматриваемого случая.

Зафиксируем положительные числа $\eta, \varepsilon_0, \delta$ и целое число m_0 . В качестве начального приближения возьмем произвольную точку $x_0 \in \Omega$.

Предположим, что множество $\mathcal{D}(x_0) = \{x \in \Omega \mid f(x) \leq f(x_0)\}$ ограничено.

Пусть уже найдено k -е приближение $x_k \in \Omega$. Покажем, как выполняется k -й цикл — построение $(k+1)$ -го приближения $x_{k+1} \in \Omega$.

Положим $x_{k0} = x_k$ и возьмем произвольный вектор $v_{k0} \in \mathcal{L}_\eta(x_{k0})$. Если $\|\bar{v}_{k,-1}\| \leq \varepsilon_0$, где $\bar{v}_{k,-1} = v_{k0}$, то полагаем $t_k = 0$, $x_{k+1} = x_{kt_k}$, и на этом k -й цикл заканчи-

вается (через t_k будем, как и раньше в §§ 6 и 7 гл. III, обозначать число шагов на k -м цикле).

Если же $\|\bar{v}_{k,-1}\| > \varepsilon_0$, то на луче $\{x_{k0}(\alpha) = x_{k0} - \alpha v_{k0} | \alpha \geq 0\}$ найдем такую точку $x_{k0}(\alpha_{k0})$, в которой

$$f(x_{k0}(\alpha_{k0})) = \min_{x_{k0}(\alpha) \in \Omega, \alpha \geq 0} f(x_{k0}(\alpha)) \quad (10.1)$$

(может оказаться, что $\alpha_{k0} = 0$).

Положим теперь $x_{k1} = x_{k0}$. Очевидно, что $f(x_{k1}) \leq f(x_{k0})$ и $x_{k1} \in \Omega$. Если $\|x_{k0} - x_{k1}\| \geq \delta$ или $f(x_{k0}) - f(x_{k1}) \geq \delta$, то полагаем $t_k = 1$, $x_{k+1} = x_{kt_k}$, и на этом k -й цикл заканчивается. В противном случае возьмем такой вектор $\bar{v}_{k0} \in \mathcal{L}_\eta(x_{k1})$, что

$$(v_{k0}, \bar{v}_{k0}) \leq 0. \quad (10.2)$$

В силу (10.1) такой вектор $\bar{v}_{k0} \in \mathcal{L}_\eta(x_{k1})$ обязательно найдется. Действительно, если в точке $x_{k1} = x_{k0}(\alpha_{k0})$ выполнено (10.1), то это может произойти по двум причинам:

а) либо $\frac{\partial f(x_{k1})}{\partial (-v_{k0})} = \max_{v \in \partial f(x_{k1})} (v, -v_{k0}) \geq 0$, и тогда

найдется субградиент $\bar{v}_{k0} \in \partial f(x_{k1})$, для которого имеет место (10.2);

б) либо $x_{k1} - \beta v_{k0} \notin \Omega$ при любом $\beta > 0$, т. е. $-v_{k0} \notin \gamma(x_{k1}) = \{v = \lambda(z - x_{k1}) | \lambda > 0, z \in \Omega\}$, а тогда найдется такой вектор $\bar{v}_{k0} \in -\Gamma^+(x_{k1})$, $\|\bar{v}_{k0}\| = \eta$, что имеет место (10.2).

Если $\|\bar{v}_{k0}\| \leq \varepsilon_0$, то полагаем $t_k = 1$, $x_{k+1} = x_{kt_k}$, и на этом k -й цикл заканчивается. Если же $\|\bar{v}_{k0}\| > \varepsilon_0$, то найдем вектор v_{k1} такой, что (см. п. 1 § 6 гл. I)

$$v_{k1} = \beta_{k0} \bar{v}_{k0} + (1 - \beta_{k0}) v_{k0}, \quad \beta_{k0} \in (0, 1),$$

$$\|v_{k1}\| = \min_{\beta \in [0, 1]} \|\beta \bar{v}_{k0} + (1 - \beta) v_{k0}\|.$$

Теперь на пересечении луча $\{x_{k1}(\alpha) = x_{k1} - \alpha v_{k1} | \alpha \geq 0\}$ и множества Ω найдем такую точку $x_{k2} = x_{k1}(\alpha_{k1})$, в которой

$$f(x_{k2}) = \min_{x_{k1}(\alpha) \in \Omega, \alpha \geq 0} f(x_{k1}(\alpha)).$$

Далее продолжаем аналогично до тех пор, пока не будет сделано такое число шагов $t_k \leq m_0$, что выполнится по крайней мере одно из условий:

$$a) \quad t_k = m_0; \quad (10.3)$$

$$b) \quad \sum_{i=1}^{t_k} \|x_{k, t-1} - x_{kt}\| \geq \delta; \quad (10.4)$$

$$c) \quad f(x_{k0}) - f(x_{kt_k}) \geq \delta; \quad (10.5)$$

$$d) \quad \|\bar{v}_k, t_k-1\| \leq \varepsilon_0. \quad (10.6)$$

После этого полагаем $x_{k+1} = x_{kt_k}$, и k -й цикл заканчивается, причем если выполнено (10.6), то заканчивается и весь процесс последовательных приближений.

Отметим, что

$$(v_{kt}, \bar{v}_{kt}) \leq 0 \quad \forall t \in 0: (t_k - 1). \quad (10.7)$$

В результате построим последовательность $\{x_k\}$. Если она конечна, то для последней полученной точки x_k по построению будет $d(x_k) \leq \varepsilon_0$. В противном случае справедлива

Теорема 10.1. Пусть функция $f(x)$ строго выпукла на множестве Ω , имеющем внутренние точки, множество $\mathcal{D}(x_0)$ ограничено.

Тогда, если точка x^* является предельной точкой последовательности $\{x_k\}$, то $x^* \in \Omega$,

$$d(x^*) \leq \max\{\varepsilon_0, bq^{m_0-1}\} \equiv d_0, \quad (10.8)$$

где $q = (1 + (\varepsilon_0 b^{-1})^2)^{-1/2}$, $b = \max\{\eta, \max_{x \in \mathcal{D}(x_0)} \max_{v \in \partial f(x)} \|v\|\}$.

В частности, если $m_0 > 1 + (\ln q)^{-1} \ln(\varepsilon_0 b^{-1})$, то $d(x^*) \leq \varepsilon_0$.

Доказательство проводится аналогично доказательству теоремы 6.1 гл. I, с учетом того, что точечно-множественное отображение $\mathcal{P}_\eta(x) = \text{co}\{\partial f(x) \cup T_\eta(x)\}$ полунепрерывно сверху на Ω и выполнены неравенства (10.7). ■

Пусть D — диаметр множества $\mathcal{D}(x_0)$, а r — радиус наибольшего шара, содержащегося в Ω . Тогда в силу (1.15) и (10.8) для любой предельной точки x^* последовательности $\{x_k\}$ будет справедлива оценка

$$f(x^*) - f^* \leq D(1 + 2bD(\eta r)^{-1})d_0.$$

Изложенный метод можно применить и для нахождения точек минимума функции $f(x)$ на множестве Ω . Схема метода при этом остается прежней, только вместо условий (10.3) и (10.6) проверяются условия $t_k = m_k$ и $\|v_{kt_k}\| \leq \varepsilon_k$ соответственно, причем $\varepsilon_k \rightarrow 0$, $m_k \rightarrow \infty$.

Теорема 10.2. Пусть функция $f(x)$ строго выпукла на множестве Ω , имеющем внутренние точки, множество $\mathcal{D}(x_0)$ ограничено и $\varepsilon_k \rightarrow 0$, $m_k \rightarrow \infty$.

Тогда любая предельная точка последовательности $\{x_k\}$ является точкой минимума функции $f(x)$ на множестве Ω .

Задание 10.1. Доказать теоремы 10.1 и 10.2.

2. Пусть множество Ω имеет вид

$$\Omega = \{x \in E_n \mid h(x) \leq 0\},$$

где $h(x)$ — выпуклая на E_n функция.

Предположим, что выполнено условие Слейтера: существуют такие $\gamma > 0$ и $\bar{x} \in E_n$, что

$$h(\bar{x}) = -\gamma < 0.$$

Тогда множество Ω имеет внутренние точки.

Используя специфику задания множества Ω , можно конкретизировать метод п. 1

Положим

$$\mathcal{L}(x) = \begin{cases} \partial f(x), & \text{если } h(x) < 0, \\ \text{co} \{ \partial f(x) \cup \partial h(x) \}, & \text{если } h(x) = 0, \\ \partial h(x), & \text{если } h(x) > 0. \end{cases}$$

Для того чтобы в точке $x^* \in \Omega$ функция $f(x)$ достигала своего наименьшего на множестве Ω значения, необходимо и достаточно, чтобы (см. теорему 1.2)

$$0 \in \mathcal{L}(x^*).$$

Последовательные приближения строятся аналогично тому, как это делалось в п. 1 только вместо векторов $\bar{v}_{kt} \in \mathcal{L}_\eta(x_k, t+1) = \text{co} \{ \partial f(x_k, t+1) \cup T_\eta(x_k, t+1) \}$ берем векторы $\bar{v}_{kt} \in \mathcal{L}(x_k, t+1)$.

БИБЛИОГРАФИЧЕСКИЙ КОММЕНТАРИЙ

Глава I

К §§ 1, 3. Более подробно с материалом можно ознакомиться по монографиям В. Ф. Демьянова, В. Н. Малоземова [36], И. И. Еремина, Н. Н. Астафьева [49], А. Д. Иоффе, В. М. Тихомирова [57], С. Карлина [60], В. Л. Макарова, А. М. Рубинова [78], Б. Н. Пшеничного [103], Р. Т. Рокафеллара [110].

К § 2. Изложение основано на работе [78, гл. I]. Свойства точечно-множественных отображений изучались в [57], [99], [111], [137], [144], [146], [162], [165].

К §§ 4, 5. Свойства выпуклых функций и их субдифференциалов можно найти в [56], [57], [67], [68], [69], [70], [71], [77], [110], [111], [119], [166].

К § 6. Лемма 6.4 имеется, по существу, в [80]. Необходимым условиям экстремума посвящена обширная литература [6], [22], [24], [26], [32], [39], [45], [57], [73], [74], [82], [102], [147], [158], [159], [163].

К § 7. Леммы 7.1, 7.2 опубликованы в [43], [44]. Теорема 7.3 является частным случаем соответствующей теоремы из работы С. С. Кутателадзе [66]. Пример 4 предложен В. М. Тихомировым.

К § 8. Соотношение (8.2) установлено в [110, стр. 237], где и введено понятие ε -субдифференциала. На непрерывность ε -субдифференциала по x при фиксированном $\varepsilon > 0$ первым обратил внимание Е. А. Нурминский [90]. Теорема 8.2 опубликована в [43], [44]. Лемма 8.1 и оценка (8.19) получены в [132].

К § 9. Оценки (9.21) и (9.22) обобщают соответствующие оценки из [37].

К §§ 10, 11. Основные результаты имеются в [43], [44].

К § 13. Изложение пп. 1, 2 следует работе [100].

К § 14. Результаты изложены в [42]. Близкие вопросы изучались в работах [5], [27], [62], [116], [143].

К § 15. Принятое здесь изложение близко к [110]. Дифференцируемость выпуклой функции одной переменной рассмотрена в [8], [121]. Случай многих переменных изучался в [109]. Существование почти везде второго дифференциала выпуклой функции установлено в [2].

К § 16. Изложение следует монографии С. С. Кутателадзе, А. М. Рубинова [68]. Доказательство леммы 16.3, по-видимому, новое. Сопряженные выпуклые функции и другие аспекты двойственности изучались в работах [57], [110], [128], [138], [156], [157], [169].

§ 17, за исключением леммы 17.2, публикуется впервые. При $\mu = 0$ включение (17.10) отмечалось в [92].

Глава II

Материал этой главы является новым. Квазидифференцируемые функции были введены в [38], [41]. Там же, а также в [40], [101] указаны основные свойства этих функций. В работе Б. Н. Пшеничного [104] введено понятие верхней выпуклой аппроксимации (в. в. а.) для исследования на минимум весьма широкого класса функций (в частности, липшицевых). Между квазидифференциалом и в. в. а. имеет место следующая связь: если $\mathcal{D}f(x_0) = [\underline{\partial}f(x_0), \overline{\partial}f(x_0)]$, то для любого $\omega \in \underline{\partial}f(x_0)$ множество $\{\omega + \underline{\partial}f(x_0)\}$ есть в. в. а. функции $f(x)$ в точке x_0 .

Теоремы 5.1—5.4, 7.2, 7.3 и лемма 6.3 установлены Л. Н. Поляковой. Ею же получены необходимые условия экстремума квазидифференцируемой функции на множестве, заданном квазидифференцируемым равенством. §§ 6, 7 написаны на основе [40]. Лемма 8.1 установлена Б. Н. Пшеничным.

Недавно А. М. Рубинов, используя один результат А. Я. Заславского, доказал теорему о квазидифференцируемости композиции квазидифференцируемых функций и нашел вид ее квазидифференциала.

Для квазидифференцируемых функций нетрудно установить связь с «шатрами» В. Г. Болтянского [6].

Глава III

Различным вопросам математического программирования посвящены работы [16], [22], [49], [61], [84], [93], [55], [105], [112], [114], [129], [145], [164], [172].

К § 1. Лемма 1.3 новая.

К § 2. Непрерывные аналоги итерационных процессов, по-видимому, впервые рассмотрены в [23]. Подробно с методами минимизации гладких функций можно ознакомиться по монографиям М. Аоки [3], Ф. П. Васильева [16], Л. В. Канторовича, Г. П. Акилова [59], В. Г. Карманова [61], Э. Полака [96], Б. Н. Пшеничного, Ю. М. Данилина [105].

К § 3. Методы наискорейшего спуска для минимизации функции максимума (не обязательно выпуклой) разработаны в [36]. Там же показано, что при $\varepsilon = 0$ метод наискорейшего спуска может расходиться. Относительная свобода в выборе направления спуска при использовании ε -субдифференциалов была замечена в работе [132], на основе которой написан п. 5.

К § 4. Субградиентный метод (первоначальный вариант с постоянным ненормированным шагом) разработан в [125]. Близкая идея использована в [47]. Сходимость к множеству точек минимума при выборе шага (4.1) обоснована в работе [50]. Следствие из теоремы 4.1 замечено в [124]. ε -субградиентный метод для достаточно широкого класса негладких функций (включая выпуклые) предложен в [92] — это, по-видимому, одна из первых попыток создания конструктивных методов минимизации. Идея п. 4 высказана М. З. Хенкиным.

Обобщение субградиентного метода на задачи стохастического программирования систематически изучено в монографии Ю. М. Ер-

мольева [51]; см. также монографии [30], [91], [127]. В [30] разработана схема применения субградиентного метода для минимизации локально липшицевых функций на основе операции сглаживания. Дальнейшему развитию и совершенствованию субградиентного метода посвящены работы [4], [28], [58], [124], [127].

К § 5. Изложение основано на работе [34].

К § 6. Впервые метод этого параграфа в несколько отличных формулировках был предложен в [149], [150] и [170]. В [171] показано, что при точной одномерной минимизации метод [170] превращается в метод сопряженных градиентов, если целевая функция квадратичная. Для произвольной гладкой функции эта связь установлена в [11].

К § 7. Результаты этого параграфа ранее не публиковались.

К § 8. Первоначально метод Келли был предложен в [148] для минимизации линейной формы на выпуклом множестве. Достаточно обширное исследование различных модификаций метода Келли проведено В. П. Булатовым [7]. Здесь изложение следует работе Н. К. Храмовой [122].

К § 9. Метод этого параграфа разработан в [1] (см. также [39]).

К § 10. Изложение основано на работе [33]. Идея метода тесно связана с работами [1], [108], [148].

Глава IV

К § 1. Лемма 1.2 установлена в [12], а лемма 1.3 — в [20, стр. 49]. Теорема 1.2 есть очевидное обобщение теоремы 3.1 гл. V из [36].

К § 2. Лемма 2.3 получена в [13].

К § 3. Метод § 3 является обобщением метода из [36, § 6, гл. IV].

К § 4. Результаты пп. 1, 2 принадлежат В. К. Шомесовой.

К § 5. При $\varepsilon = \mu = 0$ метод § 5 предложен в [97]. В [52] метод из [97] обобщен на случай квазивыпуклых функций. Приводимый здесь вариант при $G = E_n$ опубликован в [13].

К § 7. Изложение основано на работе [13].

К § 8. Ключевым в этом параграфе является свойство 6, установленное в [48]. Методу штрафных функций посвящены монографии Ю. Б. Гермейера [25], Б. С. Разумихина [106], В. В. Федорова [120].

К § 10. Метод п. 1 разработан в [12] и является, по существу, обобщением методов из [149]—[151] и [170], [171] на задачи с ограничениями. В [155] предлагается подобный метод для минимизации полугладких [154] функций.

§§ 6 и 9 публикуются впервые.

ЛИТЕРАТУРА

1. Акилов Г. П., Рубинов А. М. Метод последовательных приближений для разыскания полинома наилучшего приближения. — ДАН СССР, 1964, т. 157, № 3, с. 503—505.
2. Александров А. Д. Существование почти везде второго дифференциала выпуклой функции и некоторые связанные с ним свойства выпуклых поверхностей. — Уч. записки ЛГУ. Сер. матем., 1936, вып. 6, с. 3—35.
3. Аоки М. Введение в методы оптимизации. — М.: Наука, 1977.
4. Баженов Л. Г. Об условии сходимости метода минимизации почти-дифференцируемых функций. — Кибернетика, 1972, № 4, с. 71—72.
5. Береснев В. В. Необходимые условия экстремума в выпуклой задаче максимина на связанных множествах. — Кибернетика, 1972, № 2, с. 87—91.
6. Болтянский В. Г. Метод шатров в теории экстремальных задач. — УМН, 1975, т. 30, вып. 3, с. 3—55.
7. Булатов В. П. Методы погружения в задачах оптимизации. — Новосибирск: Наука, 1977.
8. Бурбаки Н. Функции действительного переменного. — М.: Наука, 1965.
9. Варга Дж. Оптимальное управление дифференциальными и функциональными уравнениями. — М.: Наука, 1977.
10. Васильев Л. В. Выравнивание максимумов. Теорема о квадратичной скорости сходимости. — В кн.: [20], с. 34—40.
11. Васильев Л. В. О связи релаксационного метода обобщенного градиента с методом сопряженных градиентов. — В кн.: [123], ч. I, с. 45—49.
12. Васильев Л. В., Демьянов В. Ф. Релаксационный метод для минимизации выпуклой функции на выпуклом множестве. — В кн.: Численные методы нелинейного программирования: Тезисы II Всесоюзного семинара. — Харьков, 1976, с. 100—102.
13. Васильев Л. В., Демьянов В. Ф. Метод (ε, μ, τ) -обобщенного градиентного спуска при наличии ограничений. — Вестн. ЛГУ, 1979, № 19, с. 19—23.
14. Васильев Л. В., Демьянов В. Ф., Лисина С. А. Минимизация выпуклой функции с помощью ε -субградиентов. — В кн.: Управление динамическими системами. — Л.: Изд-во ЛГУ, 1978, с. 3—22.
15. Васильев Л. В., Тарасов В. Н. Об одном методе решения задач на минимакс. — В кн.: Оптимизация. — Новосибирск, 1977, вып. 9 (36), с. 53—57.

16. Васильев Ф. П. Лекции по методам решения экстремальных задач. — М.: Изд-во МГУ, 1974.
17. Войтон Е. Ф. О методах решения некоторых экстремальных задач синтеза электрических цепей. — В кн.: Исследование операций (модели, системы, решения). — М.: Изд-во ВЦ АН СССР, 1976, с. 16—23.
18. Войтон Е. Ф. Одна задача со связанными ограничениями. — В кн.: [20], с. 106—114.
19. Войтон Е. Ф., Полякова Л. Н. Об одном классе задач при связанных ограничениях. — В кн.: Оптимизация. — Новосибирск, 1973, вып. 10, с. 41—46.
20. Вопросы теории и элементы программного обеспечения минимаксных задач. — Л.: Изд-во ЛГУ, 1977/Под ред. В. Ф. Демьянова, В. Н. Малоземова.
21. Воробьев Н. Н. Современное состояние теории игр. — УМН, 1970, т. 25, вып. 2 (152), с. 81—140.
22. Габасов Р., Кириллова Ф. М. Методы оптимизации. — Минск: Изд-во БГУ, 1975.
23. Гавурин М. К. Нелинейные функциональные уравнения и непрерывные аналоги итеративных методов. — Изв. вузов. Математика, 1958, № 5, с. 18—31.
24. Гамкрелидзе Р. В. Необходимые условия первого порядка и аксиоматика в экстремальных задачах. — Тр. МИАН СССР им. В. А. Стеклова, 1971, т. 112, с. 152—180.
25. Гермейер Ю. Б. Введение в теорию исследования операций. — М.: Наука, 1971.
26. Гирсанов И. В. Лекции по математической теории экстремальных задач. — М.: Изд-во МГУ, 1970.
27. Гольштейн Е. Г. Теория двойственности в математическом программировании и ее приложения. — М.: Наука, 1971.
28. Гольштейн Е. Г. Обобщенный метод отыскания седловых точек. — Экономика и матем. методы, 1972, т. 8, вып. 4, с. 569—579.
29. Горелик В. А. Приближенное нахождение максимина с ограничениями, связывающими переменные. — ЖВМ и МФ, 1972, т. 12, № 2, с. 510—517.
30. Гупал А. М. Стохастические методы решения негладких экстремальных задач. — Киев: Наукова думка, 1979.
31. Данскин Дж. М. Теория максимина и ее приложения к задачам распределения вооружения. — М.: Советское радио, 1970.
32. Демьянов В. Ф. Минимакс: дифференцируемость по направлениям — Л.: Изд-во ЛГУ, 1974.
33. Демьянов В. Ф. К методу экстремального базиса. — ДАН СССР, 1976, т. 229, № 2, с. 272—275.
34. Демьянов В. Ф. Многошаговый метод обобщенного градиентного спуска. — ЖВМ и МФ, 1978, т. 18, № 5, с. 1112—1118.
35. Демьянов В. Ф. Модифицированный метод обобщенного градиента при наличии ограничений. — Вестн. ЛГУ, 1978, № 19, с. 25—29.
36. Демьянов В. Ф., Малоземов В. Н. Введение в минимакс. — М.: Наука, 1972.
37. Демьянов В. Ф., Певный А. Б. Некоторые оценки в минимаксных задачах. — Кибернетика, 1972, № 1, с. 107—112.

38. Демьянов В. Ф., Полякова Л. Н., Рубинов А. М. Об одном обобщении понятия субдифференциала. — В кн.: Всесоюзная конференция «Динамическое управление»: Тезисы докладов. — Свердловск: 1979, с. 79—84.
39. Демьянов В. Ф., Рубинов А. М. Приближенные методы решения экстремальных задач. — Л.: Изд-во ЛГУ, 1968.
40. Демьянов В. Ф., Полякова Л. Н. Условия минимума квазидифференцируемой функции на квазидифференцируемом множестве. — ЖВМ и МФ, 1980, т. 20, № 4, с. 849—856.
41. Демьянов В. Ф., Рубинов А. М. О квазидифференцируемых функционалах. — ДАН СССР, 1980, т. 250, № 1, с. 21—25.
42. Демьянов В. Ф., Шомесова В. К. Дифференцируемость по направлениям функции супремума. — Вестн. ЛГУ, 1978, № 7, с. 15—20.
43. Демьянов В. Ф., Шомесова В. К. Условные субдифференциалы выпуклых функций. — ДАН СССР, 1978, т. 242, № 4, с. 753—756.
44. Демьянов В. Ф., Шомесова В. К. Условные субградиенты и условные субдифференциалы выпуклых функций. — В кн.: Современное состояние теории исследования операций/Под ред. Н. Н. Моисеева. — М.: Наука, 1979, с. 311—335.
45. Дубовицкий А. Я., Милютин А. А. Задачи на экстремум при наличии ограничений. — ЖВМ и МФ, 1965, т. 5, № 3, с. 395—453.
46. Евтушенко Ю. Г. Численный метод отыскания наилучших гарантированных оценок. — ЖВМ и МФ, 1972, т. 12, № 1, с. 89—104.
47. Еремин И. И. Итеративный метод для решения чебышевских приближений несовместных линейных неравенств. — ДАН СССР, 1962, т. 143, № 6, с. 1254—1256.
48. Еремин И. И. О методе штрафов в выпуклом программировании. — Кибернетика, 1967, № 4, с. 63—67.
49. Еремин И. И., Астафьев Н. Н. Введение в теорию линейного и выпуклого программирования. — М.: Наука, 1976.
50. Ермольев Ю. М. Методы решения нелинейных экстремальных задач. — Кибернетика, 1966, № 4, с. 1—17.
51. Ермольев Ю. М. Методы стохастического программирования. — М.: Наука, 1976.
52. Заботин Я. И., Кораблев А. И., Хабибуллин Р. Ф. О минимизации квазивыпуклых функционалов. — Изв. вузов. Математика, 1972, № 10, с. 27—33.
53. Заботин Я. И., Крейнин М. И. О сходимости методов отыскания минимакса. — Изв. вузов. Математика, 1977, № 10 (185), с. 56—64.
54. Зангвилл У. И. Нелинейное программирование. Единый подход. — М.: Советское радио, 1973.
55. Зойтендейк Г. Методы возможных направлений. — М.: ИЛ, 1963.
56. Иоффе А. Д., Левин В. Л. Субдифференциалы выпуклых функций. — Тр. Моск. матем. об-ва, 1972, т. 26, с. 3—73.

57. Иоффе А. Д., Тихомиров В. М. Теория экстремальных задач. — М.: Наука, 1974.

58. Итеративные методы в теории игр и программировании/Под ред. В. З. Беленького и В. А. Волконского. — М.: Наука, 1974.

59. Канторович Л. В., Акилов Г. П. Функциональный анализ. — М.: Наука, 1977.

60. Карлин С. Математические методы в теории игр, программировании и экономике. — М.: Мир, 1964.

61. Карманов В. Г. Математическое программирование. — М.: Наука, 1975.

62. Кирин Н. Е. Методы последовательных оценок в задачах оптимизации управляемых систем. — Л.: Изд-во ЛГУ, 1975.

63. Красовский Н. Н., Субботин А. И. Позиционные дифференциальные игры. — М.: Наука, 1974.

64. Куржанский А. Б. Управление и наблюдение в условиях неопределенности. — М.: Наука, 1977.

65. Кусраев А. Г. О необходимых условиях экстремума для негладких векторнозначных отображений. — ДАН СССР, 1978, т. 242, № 1, с. 44—47.

66. Кутателадзе С. С. Выпуклое ϵ -программирование. — ДАН СССР, 1979, т. 245, № 5, с. 1048—1050.

67. Кутателадзе С. С. Выпуклые операторы. — УМН, 1979, т. 34, вып. 1 (205), с. 167—196.

68. Кутателадзе С. С., Рубинов А. М. Двойственность Минковского и ее приложения. — Новосибирск: Наука, 1976.

69. Лебедев В. Н., Тынянский Н. Т. Теория выпукло-вогнутых игр. — ДАН СССР, 1967, т. 174, вып. 6, с. 1264—1267.

70. Левин В. Л. О субдифференциалах выпуклых функционалов. — УМН, 1970, т. 25, вып. 4 (154), с. 183—184.

71. Левин В. Л. Субдифференциалы выпуклых отображений и сложных функций. — Сибирский матем. ж., 1972, т. XIII, № 6, с. 1295—1303.

72. Левитин Е. С. Об одном методе минимизации для негладких экстремальных задач. — ЖВМ и МФ, 1969, т. 9, № 4, с. 783—807.

73. Левитин Е. С., Милютин А. А., Осмоловский Н. П. Об условиях локального минимума в задачах с ограничениями. — В кн.: Математическая экономика и функциональный анализ. — М.: Наука, 1974, с. 139—202.

74. Левитин Е. С., Милютин А. А., Осмоловский Н. П. Условия высших порядков локального минимума в задачах с ограничениями. — УМН, 1978, т. 33, вып. 6 (204), с. 85—148.

75. Лисина С. А. Минимизация ϵ -выпуклой функции. — В кн.: [123], с. 122—124.

76. Лоран П.-Ж. Аппроксимация и оптимизация. — М.: Мир, 1975.

77. Лудерер Б. Задачи на минимакс и выпуклый анализ. — Вестн. МГУ, 1975, № 6, с. 25—31.

78. Макаров В. Л., Рубинов А. М. Математическая теория экономической динамики и равновесия. — М.: Наука, 1973.

79. Малоземов В. Н. О выравнивании максимумов. — ЖВМ и МФ, 1976, т. 16, № 3, с. 781—784.

80. Малоземов В. Н. О достаточных условиях локального минимакса. — Вестн. ЛГУ, 1976, № 7, с. 55—59.

81. Малоземов В. Н., Певный А. Б. Выравнивание максимумов. — В кн.: [20], с. 27—34.

82. Моисеев Н. Н. Элементы теории оптимальных систем. — М.: Наука, 1975.

83. Моцкус Й. Б. Многоэкстремальные задачи в проектировании. — М.: Наука, 1975.

84. Мухачева Э. А., Рубинштейн Г. Ш. Математическое программирование. — Новосибирск: Наука, 1977.

85. Натансон И. П. Теория функций вещественной переменной. — М.: Наука, 1974.

86. Никайдо Х. Выпуклые структуры и математическая экономика. — М.: Мир, 1972.

87. Норкин В. И. Нелокальные алгоритмы минимизации недифференцируемых функций. — Кибернетика, 1978, № 5, с. 57—60.

88. Нурминский Е. А. К дифференцируемости многозначных отображений. — Кибернетика, 1978, № 5, с. 44—48.

89. Нурминский Е. А. Квазиградиентный метод решения задачи нелинейного программирования. — Кибернетика, 1973, № 1, с. 122—125.

90. Нурминский Е. А. О непрерывности ε -субградиентных отображений. — Кибернетика, 1977, № 5, с. 148—149.

91. Нурминский Е. А. Численные методы решения детерминированных и стохастических минимаксных задач. — Киев: Наукова думка, 1979.

92. Нурминский Е. А., Желиховский А. А. ε -квазиградиентный метод решения негладких экстремальных задач. — Кибернетика, 1977, № 1, с. 109—113.

93. Ортега Дж., Рейнболдт В. Итерационные методы решения нелинейных систем уравнений со многими неизвестными. — М.: Мир, 1975.

94. Певный А. Б. Дифференцирование функции максимина. — ЖВМ и МФ, 1971, т. 11, № 2, с. 510—514.

95. Подиновский В. В. Об относительной важности критериев в многокритериальных задачах принятия решений. — В кн.: Многокритериальные задачи принятия решений. — М.: Машиностроение, 1978, с. 48—92.

96. Полак Э. Численные методы оптимизации. Единый подход. — М.: Мир, 1974.

97. Поляк Б. Т. Один общий метод решения экстремальных задач. — ДАН СССР, 1967, т. 174, № 1, с. 33—36.

98. Поляк Б. Т. Минимизация негладких функционалов. — ЖВМ и МФ, 1969, т. 9, № 3, с. 509—521.

99. Поляк Р. А. К отысканию неподвижной точки одного класса точечно-множественных отображений. — ДАН СССР, 1978, т. 242, № 6, с. 1252—1255.

100. Полякова Л. Н. Непрерывность нормальных конусов. — Вестн. ЛГУ, 1979, № 7, с. 112—113.

101. Полякова Л. Н. Минимизация одного класса недифференцируемых функций. — В кн.: [123], с. 145—148.

102. Понтрягин Л. С. и др. Математическая теория оптимальных процессов. — М.: Наука, 1961.
103. Пшеничный Б. Н. Необходимые условия экстремума. — М.: Наука, 1969.
104. Пшеничный Б. Н. О необходимых условиях экстремума для негладких функций. — Кибернетика, 1977, № 6, с. 92—96.
105. Пшеничный Б. Н., Данилин Ю. М. Численные методы в экстремальных задачах. — М.: Наука, 1975.
106. Разумихин Б. С. Физические модели и методы теории равновесия в программировании и экономике. — М.: Наука, 1975.
107. Растригин Л. А. Статистические методы поиска экстремума. — М.: Наука, 1968.
108. Ремез Е. Я. Основы численных методов чебышевского приближения. — Киев: Наукова думка, 1969.
109. Решетняк Ю. Г. Обобщенные производные и дифференцируемость почти всюду. — Матем. сб., 1978, т. 75, с. 323—334.
110. Рокафеллар Р. Выпуклый анализ. — М.: Мир, 1973.
111. Рубинов А. М. Сублинейные операторы и их приложения. — УМН, 1977, т. 33, вып. 4 (196), с. 113—174.
112. Рубинштейн Г. Ш. Конечномерные модели оптимизации. — Новосибирск: Изд-во Новосибирск. ун-та, 1970.
113. Рудин У. Основы математического анализа. — М.: Мир, 1966.
114. Сеа Ж. Оптимизация. Теория и алгоритмы. — М.: Мир, 1973.
115. Смирнов В. И. Курс высшей математики, т. V. — М.: Физматгиз, 1959.
116. Сотсков А. И. Необходимые условия минимума для одного класса негладких задач. — ДАН СССР, 1969, т. 189, № 2, с. 261—264.
117. Стронгин Р. Г. Численные методы в многоэкстремальных задачах. — М.: Наука, 1978.
118. Тарасов В. Н. Ускорение сходимости при решении нелинейных минимаксных задач. — В кн.: [20], с. 76—81.
119. Тынянский Н. Т. Сопряженные вогнуто-выпуклые функции в линейных топологических пространствах и их седловые точки. — Матем. сб., 1969, т. 78 (120), № 4, с. 512—541.
120. Федоров В. В. Численные методы максимина. — М.: Наука, 1979.
121. Харди Г. Г., Литтльвуд Д. Е., Полиа Г. Неравенства. — М.: ИЛ, 1948.
122. Храмова Н. К. Некоторые обобщения метода секущих плоскостей. — Вестн. ЛГУ, 1979, № 7, с. 116—117.
123. Численные методы нелинейного программирования: Тезисы III Всесоюзного семинара. — Харьков, 1979.
124. Шепилов М. А. О методе обобщенного градиента отыскания абсолютного минимума выпуклой функции. — Кибернетика, 1976, № 4, с. 52—57.
125. Шор Н. З. Применение метода градиентного спуска для решения сетевой транспортной задачи. — В кн.: Материалы научн. семинара по теоретическим и прикладным вопросам кибернетики и

исследования операций. — Киев: Институт кибернетики АН УССР, 1962, вып 1, с. 9—17.

126. Шор Н. З. О классе почти-дифференцируемых функций и одном методе минимизации функций этого класса. — Кибернетика, 1972, № 4, с. 65—70.

127. Шор Н. З. Методы минимизации недифференцируемых функций и их приложения. — Киев: Наукова думка, 1979.

128. Auslender A. Méthodes et théorèmes de dualité. — C. R. Ac. Sci. Paris, 1968, v. 267, p. 1—4.

129. Auslender A. Optimization. Méthodes numériques. — Paris: Masson, 1976.

130. Auslender A. Differentiable stability in non convex and nondifferentiable programming. — Math. Progr. Study, 1979, v. 10, p. 29—41.

131. Bertsekas D. P. Nondifferentiable optimization via approximation. — In: [153], p. 1—25.

132. Bertsekas D. P., Mitter S. K. A descent numerical method for optimization problems with nondifferentiable cost functionals. — SIAM J. Control, 1973, v. 11, N 4, p. 637—652.

133. Clarke F. H. Generalized gradients and applications. — Trans. Amer. Math. Soc., 1975, v. 205, p. 247—262.

134. Clarke F. H. A new approach to Lagrange multipliers. — Mathematics of Oper. Research, 1976, v. 1, N 2, p. 165—174.

135. Cullum Jane, Donath W. E., Wolfe P. The minimization of certain nondifferentiable sums of eigenvalues of symmetric matrices. — In: [153], p. 35—55.

136. Ekeland Ivar. Nonconvex minimization problems. — Bull. Amer. Math. Soc., 1979, v. 1, N 3, p. 443—474.

137. Ekeland I., Valadier M. Representation of set-valued functions. — J. Math. Anal. Appl., 1971, v. 35, p. 621—629.

138. Fenchel W. On conjugate convex functions. — Canad. J. Math., 1949, v. 1, p. 73—77.

139. Goldstein A. A. Optimization with corners. — In: Nonlinear programming. — New York: Academic Press, Inc., 1975, v. 2, p. 215—230.

140. Goldstein A. A. Optimization of Lipschitz continuous functions. — Math. Programming, 1977, v. 13, p. 14—22.

141. Halkin H. Mathematical Programming without differentiability. — In: Calculus of variations and control theory/Ed. D. L. Rusel. — New York: Academic Press, 1976, p. 279—288.

142. Hiriart-Urruty J. B. On optimality conditions in nondifferentiable programming. — Math. Programming, 1978, v. 14, p. 73—86.

143. Hogan W. W. Directional derivatives for extremal values functions. — West. Manag. Sci. Inst., Working paper. — Los Angeles, 1971, N 177.

144. Hogan W. W. Point-to-set maps in mathematical programming. — SIAM Review, 1973, v. 15, 591—603.

145. Huard P. Resolution of mathematical programming problems with nonlinear constraints by the method of centres. — In: Nonlinear programming/Ed. J. Abadie. — Amsterdam: North-Holland, 1967, p. 206—219.

146. Huard P. Optimization algorithms and point-to-set maps. — *Math. Programming*, 1975, v. 8, p. 308—331.

147. Ioffe A. D. Necessary and sufficient conditions for a local minimum. 1—3. — *SIAM J. Control and Optimization*, 1979, v. 17, N 2, p. 245—288.

148. Kelley J. F. The cutting-plane method for solving convex problems. — *SIAM J. Appl. Math.*, 1960, v. 8, N 4, p. 703—712.

149. Lemarechal C. Note on an extension of Davidon methods to nondifferentiable functions. — *Math. Programming*, 1974, v. 7, N 3, p. 384—387.

150. Lemarechal C. An algorithm for minimizing convex functions. — In: *Proc. IFIP Congress-74*. — Amsterdam, 1974, p. 552—556.

151. Lemarechal C. On extension of Davidon methods to nondifferentiable problems. — In: [153], p. 95—109.

152. Mangasarian O. L. Pseudo-convex functions. — *SIAM J. Control*, 1965, v. 3, N 3, p. 281—290.

153. *Mathematical Programming Study 3. Non-differentiable optimization*/Eds. M. L. Balinski, P. Wolfe. — Amsterdam: North-Holland, 1975.

154. Mifflin R. Semismooth and semiconvex functions in constrained optimization. — *SIAM J. Control and Optimization*, 1977, v. 15, N 6, p. 959—972.

155. Mifflin R. An algorithm for constrained optimization with semismooth functions. RR-77-3, IIASA. — Laxenburg, Austria, 1977.

156. Moreau J. — J. — Fonctionselles sous-différentiables. — *C. R. Ac. Sci. Paris*, 1963, v. 257, p. 4117—4119.

157. Moreau J. — J. Convexity and duality. — In: *Functional analysis and Optimization*/Ed. E. R. Caianello. — New York: Academic Press, 1966, p. 145—169.

158. Neustadt L. W. A general theory of extremals. — *J. Comput. System. Sci.*, 1969, v. 3, p. 57—92.

159. Neustadt L. W. *Optimization: A theory of necessary conditions*. — Princeton, N. J.: Princeton Univ. Press, 1976.

160. *Nonsmooth optimization Proceedings of a IIASA Workshop. 28 March—8 April 1977*/Eds. C. Lemarechal, R. Mifflin — Oxford: Pergamon Press, 1978.

161. Osborne M. R., Watson G. A. An algorithm for minimax approximation in the non-linear case. — *Comput. J.*, 1969, v. 12, N 1, p. 63—68.

162. Robinson S. M., Meyer R. R. Lower semicontinuity of multivalued linearization mappings. — *SIAM J. on Control*, 1973, v. 11, N 3, p. 525—533.

163. Robinson S. M. First-order conditions for general non-linear optimization. — *SIAM J. on Appl. Math.*, 1976, v. 30, p. 597—603.

164. Rockafellar R. T. Augmented Lagrangian multiplier rule and duality in nonconvex programming. — *SIAM J. Control and Optimization*, 1974, v. 12, N 2, p. 268—285.

165. Rockafellar R. T. The theory of subgradients and its applications to problems of optimization. — *Lecture Notes, Univ. of Montreal*. — Montreal, 1978.

166. Valadier M. Sous-différentiels d'une borne supérieure et d'une somme continue de fonctions convexes. — C. R. Ac. Sci. Paris, 1969, t. 268, p. 39—42.

167. Warga J. Necessary conditions without differentiability assumptions in optimal control. — J. Diff. Equations, 1975, v. 18, p. 41—62.

168. Warga J. Derivative containers, inverse functions and controllability. — Calculus of Variations and Control Theory, 1976, p. 13—45.

169. Weiss E. A. Konjugierte Funktionen. — Arch. Math., 1969, Bd. XX, S. 538—545.

170. Wolfe P. Note on a method of conjugate subgradients for minimizing nondifferentiable functions. — Math. Programming, 1974, v. 7, N 3, p. 380—383.

171. Wolfe P. A method of conjugate subgradients for minimizing nondifferentiable functions. — In: [153], p. 145—173.

172. Wolfe P. Convergence theory in nonlinear programming. — In: Integrer and nonlinear programming/Ed. J. Abadie. — Amsterdam: North-Holland, 1970, p. 1—36.

2 р. 10 к.

2р. 10к.

