Лекция 1. Выпуклые множества.

Определение 1.1. Множество A в $\mathbb{R}^n$ называется выпуклым, если вместе с любыми двумя своими точками оно содержит и весь отрезок, их соединяющий.
Т.е. $\forall\, x, y \in A$, $\forall\, \alpha, \beta \geq 0$, $\alpha + \beta = 1$, точка $\alpha x + \beta y \in A$.

Отрезок, состоящий из точек $z = \alpha x + \beta y$ указанного вида, мы будем обозначать $[x, y]$.

Отметим, что пустое множество и все пространство удовлетворяют этому определению, поэтому они являются выпуклыми. Другие примеры выпуклых множеств — точка, отрезок прямой, круг, квадрат, трапеция, треугольник, прямая, угол $\leq 180^o$ между двумя лучами на плоскости, шар, куб, тетраэдр, цилиндр, плоскость в $\mathbb{R}^3$.

Примеры невыпуклых множеств — окружность и любая ее дуга, граница квадрата, крест, сфера, угол $> 180^o$ между двумя лучами на плоскости, объединение двух непересекающихся кругов, конечное множество точек в количестве ≥ 2 .

Замечание. Понятие выпуклого множества имеет смысл в векторном пространстве любой размерности, однако в данном курсе мы для простоты ограничимся случаем конечномерного пространства.

Элементарные свойства выпуклых множеств

1) Пересечение любого числа (точнее, любого семейства), конечного или бесконечного, выпуклых множеств выпукло: если $B = \bigcap_{\mu \in \mathcal{M}} A_\mu$, где $\mathcal{M}$ — произвольное множество индексов, и каждое A_μ выпукло, то B выпукло. Это непосредственно вытекает из определения.

2) Если A выпукло, то для любого конечного числа точек $x_1, \ldots, x_k \in A$ и $\forall\, \alpha_1, \ldots, \alpha_k \geq 0$, $\sum \alpha_i = 1$, точка $x = \sum_1^k \alpha_i x_i$ (выпуклая комбинация точек $x_1, \ldots, x_k$) принадлежит A.

Доказательство. Индукция по k. Для $k = 2$ это свойство выполнено по определению выпуклого множества. Пусть для некоторого k оно доказано. Рассмотрим произвольную выпуклую комбинацию $k+1$ точки: $x = \alpha_1 x_1 + \cdots + \alpha_k x_k + \alpha_{k+1} x_{k+1}$, где все $x_i \in A$, $\alpha_i \geq 0$, и $\alpha_1 + \ldots + \alpha_{k+1} = 1$. Если среди α_i есть нулевые, то x есть выпуклая комбинация $\leq k$ из этих точек, и тогда по предположению индукции $x \in A$, ч.т.д. Поэтому далее считаем все $\alpha_i > 0$. Положим $\beta = \alpha_1 + \ldots + \alpha_k$. Тогда $\beta > 0$ и точка $y = \frac{\alpha_1}{\beta} x_1 + \ldots + \frac{\alpha_k}{\beta} x_k \in A$ по предположению индукции как выпуклая комбинация k точек. Но так как при этом $x = \beta y + \alpha_{k+1} x_{k+1}$ и $\beta + \alpha_{k+1} = 1$, т.е. x есть выпуклая комбинация точек y и x_{k+1}, то $x \in A$ по определению выпуклого множества, ч. т. д. $\square$

3) Образ и полный прообраз выпуклого множества при аффинном отображении — выпуклое множество. Пусть $f : \mathbb{R}^n \to \mathbb{R}^m$ — аффинное отображение, т.е. $\forall x_1, x_2 \in \mathbb{R}^n$, $\forall \alpha_1, \alpha_2 \in \mathbb{R}$, $\alpha_1 + \alpha_2 = 1$, выполняется равенство

$$f(\alpha_1 x_1 + \alpha_2 x_2) \;=\; \alpha_1 f(x_1) + \alpha_2 f(x_2). \tag{1.1}$$

Если A есть выпуклое множество в $\mathbb{R}^n$, то $B = f(A)$ выпукло.

Действительно, для любых $y_1, y_2 \in B$ существуют $x_1, x_2 \in A$ такие, что $f(x_1) = y_1$, $f(x_2) = y_2$. Тогда для любых $\alpha_1, \alpha_2 \ge 0$, $\alpha_1 + \alpha_2 = 1$, точка $x = \alpha_1 x_1 + \alpha_2 x_2 \in A$, и в силу (1.1) $f(x) = \alpha_1 y_1 + \alpha_2 y_2$, следовательно точка $\alpha_1 y_1 + \alpha_2 y_2 \in B$ (как образ точки $x \in A$), поэтому B выпукло.

Если B есть выпуклое множество в $\mathbb{R}^m$, то $A = f^{-1}(B)$ выпукло.

Действительно, пусть $x_1, x_2 \in A$, т.е. $f(x_1) = y_1 \in B$, $f(x_2) = y_2 \in B$. Тогда $\forall \alpha_1, \alpha_2 \ge 0$, $\alpha_1 + \alpha_2 = 1$ в силу (1.1) $f(\alpha_1 x_1 + \alpha_2 x_2) = \alpha_1 y_1 + \alpha_2 y_2 \in B$, следовательно, по определению $\alpha_1 x_1 + \alpha_2 x_2 \in A$.

3') (Следствие из свойства 3). Сдвиг на любой вектор и гомотетия с любым коэффициентом выпуклого множества выпуклы. Т.е. $\forall z \in \mathbb{R}^n$, $\forall \lambda \in \mathbb{R}$ множества $A + z$ и λA выпуклы.

Проекция выпуклого множества на любое подпространство (вдоль любого дополняющего подпространства) выпукла: если $\mathbb{R}^n = L \bigoplus L'$, то проекция A на L вдоль L' выпукла.

4) Прямое (декартово) произведение выпуклых множеств выпукло: если $A \subset \mathbb{R}^n$ и $B \subset \mathbb{R}^m$ выпуклы, то $A \times B \in \mathbb{R}^{n+m}$ выпукло.

Это вытекает непосредственно из определения: если (x_1, y_1), $(x_2, y_2) \in A \times B$, то любая выпуклая комбинация $x = \alpha_1 x_1 + \alpha_2 x_2 \in A$, и $y = \alpha_1 y_1 + \alpha_2 y_2 \in B$, а тогда $\alpha_1 (x_1, y_1) + \alpha_2 (x_2, y_2) = (x, y) \in A \times B$.

5) Алгебраическая сумма и разность $A \pm B$ двух выпуклых множеств выпуклы. Можно проверить непосредственно, а можно поступить так. Рассмотрим линейное отображение $f : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^n$, $f(x, y) = x + y$. Тогда $f(A \times B) = A + B$. По свойству 4 $A \times B$ выпукло, а тогда по свойству 3 выпукло и $A + B$. Для $A - B$ надо взять $f(x, y) = x - y$.

5') (Следствие предыдущего). Линейная комбинация $\lambda_1 A_1 + \ldots + \lambda_k A_k$ (с произвольными коэффициентами) любого конечного числа выпуклых множеств выпукла.

Упражнения. 1) Верно ли, что если $\forall x, y \in A$ полусумма $\frac{1}{2}(x + y) \in A$, то A выпукло ? (Контрпример — множество рациональных чисел).

2) Доказать, что для выпуклых множеств при любых $\alpha, \beta > 0$ справедливо равенство $\alpha A + \beta A = (\alpha + \beta) A$, в частности, $A + A = 2A$. Показать, что для невыпуклых множеств это, вообще говоря, неверно. (Т.е. построить контрпример).

Важными примерами выпуклых множеств служат *полупространства:* замкнутое $(a,x) \le c$, и открытое $(a,x) < c$ (здесь $a \in \mathbb{R}^n$, $c \in \mathbb{R}$). Их выпуклость можно проверить и непосредственно, а можно заметить, что они есть прообразы полупрямых $(-\infty, c]$ и $(-\infty, c)$ при линейном отображении $x \mapsto (a,x)$.

Особенно важную роль играют *замкнутые* полупространства. Пусть задано произвольное семейство векторов $a_\mu \in \mathbb{R}^n$ и чисел c_μ, где $\mu \in \mathcal{M}$ — произвольное множество индексов. Тогда множество

$$B = \{ x \in \mathbb{R}^n \mid (a_\mu, x) \le c_\mu \qquad \forall \mu \in \mathcal{M} \} \tag{1.2}$$

выпукло и замкнуто как пересечение семейства выпуклых замкнутых множеств (полупространств) $(a_\mu, x) \le c_\mu$.

В дальнейшем мы покажем — и в этом и состоит фундаментальная роль замкнутых полупространств — что любое выпуклое замкнутое множество представимо в таком виде.

Отметим, что если множество индексов $\mathcal{M}$ конечно, то B называется *полиэдральным (многогранным)* множеством. Примером являются квадрат, куб, треугольник, полоса, первый квадрант на плоскости. Ограниченное многогранное множество называется *многогранником.*

Операция взятия выпуклой оболочки

Из любого множества можно получить выпуклое множество путем взятия его *выпуклой оболочки* или *овыпукления.*

Определение 1.2. Выпуклой оболочкой $co\,A$ множества A называется наименьшее выпуклое множество, содержащее A.

Здесь "наименьшее" — в смысле включения, то есть $co\,A$ — это такое выпуклое множество, что если некоторое выпуклое множество $B \supset A$, то $B \supset co\,A$.

Покажем, что это определение корректно, т.е. что такое наименьшее выпуклое множество всегда существует и единственно.

Лемма 1.1. Выпуклая оболочка множества A есть пересечение всех выпуклых множеств, содержащих A.

Действительно, пусть $\mathcal{M} = \{B\}$ есть совокупность всех выпуклых множеств B, содержащих A. Так как каждое B выпукло и содержит A, то и пересечение B_0 всех $B \in \mathcal{M}$ также выпукло и содержит A. Поскольку любое $B \supset B_0$, то по определению $B_0 = co\,A$. Отсюда следует и единственность $\mathrm{co}\,A$. (Впрочем, она вытекает и из определения: если B_1 и B_2 — два наименьших выпуклых множества, содержащих A, то $B_1 \supset B_2$ и $B_2 \supset B_1$.) $\square$

Следующая лемма дает другой, в некотором смысле "более конструктивный" способ нахождения выпуклой оболочки.

Лемма 1.2. Выпуклая оболочка любого множества есть совокупность всевозможных выпуклых комбинаций его элементов:

$$\mathrm{co}\, A = \ \{z = \sum_1^k \alpha_i x_i \mid k \in \mathbb{N}, \ \ \forall\, i \ \ \alpha_i \geq 0, \ \ x_i \in A, \ \ \sum_1^k \alpha_i = 1\}$$

Обратим внимание, что здесь k пробегает всё множество натуральных чисел, т.е. надо брать выпуклые комбинации всех возможных длин.

Доказательство. Обозначим множество выпуклых комбинаций (стоящее справа) через Q. Ясно, что оно выпукло (ибо содержит выпуклые комбинации любых двух своих точек; проверить!), и содержит A. Тогда по определению $Q \supset co\, A$. С другой стороны, так как $\mathrm{co}\, A$ выпукло и содержит A, то $\mathrm{co}\, A$ содержит все выпуклые комбинации своих элементов, в частности все выпуклые комбинации элементов множества A, то есть $co\, A \supset Q$. Из полученных двух включений следует, что $co\, A = Q$. □

Свойства выпуклой оболочки

1) $\forall\, \lambda \in \mathbb{R} \quad co\,(\lambda A) = \lambda\, co\, A$. (Проверить).

2) $co\,(A + B) = co\, A + co\, B$.

Доказательство. Пусть $x_i \in A$, $y_i \in B$, $\alpha_i \geq 0$, $i = 1, \ldots, k$, $\sum \alpha_i = 1$. Тогда

$$\alpha_1(x_1 + y_1) + \ldots + \alpha_k(x_k + y_k) = \sum \alpha_i x_i + \sum \alpha_i y_i,$$

откуда $co\,(A + B) \subset co\, A + co\, B$.

Докажем обратное включение $\supset$. Пусть $\hat{x} \in co\, A$, $\hat{y} \in co\, B$, и $z = \hat{x} + \hat{y}$. Тогда $\hat{x} = \sum_1^k \alpha_i\, x_i$, где $x_i \in A$, $\alpha_i \geq 0$, $\sum_1^k \alpha_i = 1$, и $\hat{y} = \sum_1^m \beta_j\, y_j$, где $y_j \in B$, $\beta_j \geq 0$, $\sum_1^m \beta_j = 1$.

Для любого фиксированного j точка $\hat{x} + y_j = \sum_i \alpha_i(x_i + y_j) \in co\,(A + B)$, а так как последнее множество выпукло, то и выпуклая комбинация этих точек $\sum_j \beta_j(\hat{x} + y_j) = \hat{x} + \hat{y} \in co\,(A + B)$, поэтому $co\, A + co\, B \subset co\,(A + B)$, ч. т. д. □

3) *Выпуклая оболочка ограниченного множества ограничена.* Действительно, если A содержится в некотором шаре, то, поскольку этот шар является выпуклым множеством, по определению и $\mathrm{co}\, A$ содержится в нем.

Комбинации точек и операция взятия оболочки множества

Введенное понятие выпуклой оболочки аналогично известным из курса линейной алгебры понятиям аффинной и линейной оболочки множества. Отметим здесь эту аналогию.

Пусть даны точки $x_1, \ldots, x_k \in \mathbb{R}^n$. Положим $y = \lambda_1 x_1 + \ldots + \lambda_k x_k$.

Эта точка называется

а) линейной комбинацией данных точек, если $\lambda_i -$ произвольные числа (из $\mathbb{R}$);

б) аффинной комбинацией данных точек, если $\sum \lambda_i = 1$;

в) выпуклой комбинацией данных точек, если $\sum \lambda_i = 1,$ и все $\lambda_i \geq 0$.

Множество $A \subset \mathbb{R}^n$ называется:

а) линейным подпространством, б) аффинным подпространством, в) выпуклым, если оно содержит соответственно все а) линейные, б) аффинные, в) выпуклые комбинации любых своих точек.

Для любого множества $A \subset \mathbb{R}^n$ определяется его а) линейная оболочка $Lin\, A$, б) аффинная оболочка $Aff\, A$, в) выпуклая оболочка $co\, A$, как наименьшее а) линейное подпространство, б) аффинное подпространство, в) выпуклое множество, содержащее A.

Тогда справедлива следующая

Теорема 1.1. $Lin\, A$, $Aff\, A$, $\mathrm{co}\, A$ есть, соответственно,

а) множество всех линейных, аффинных, выпуклых комбинаций элементов множества A,

б) пересечение всех линейных подпространств, аффинных подпространств, выпуклых множеств, содержащих множество A.

Доказательство во всех трех случаях фактически одно и то же; оно повторяет приведенные выше доказательства лемм 1.1 и 1.2.

Выпуклые конусы.

Вернемся к выпуклым множествам. Важным частным случаем выпуклого множества является выпуклый конус.

Определение 1.3. Множество $K \subset \mathbb{R}^n$ называется конусом, если $\forall x \in K$ и $\forall \alpha > 0$ точка $\alpha x \in K$. Конус, являющийся выпуклым множеством, называется выпуклым конусом.

Отметим, что конус может не содержать нуля. Например, на плоскости $\mathbb{R}^2$ таковыми являются множества $\{(x,y) \mid x > 0\}$ и $\{(x,y) \mid x > 0,\ y > 0\}$. Вообще, любой открытый конус, не совпадающий со всем пространством, не содержит нуля.

(Покажите!) Обратим также внимание, что конус не обязательно "заострён": например, любое подпространство является очевидно конусом.

Утверждение 1.1. Конус K является выпуклым $\iff$ $\forall\, x, y \in K$ точка $x + y \in K$, (т.е. $K + K \subset K$).

Доказательство. Если конус K выпуклый, то $\frac{x+y}{2} \in K$, и по определению конуса $2\,\frac{x+y}{2} = x + y \in K$.

Обратно, пусть $K + K \subset K$. Возьмем любые точки $x, y \in K$ и покажем, что $\forall\, \alpha, \beta \geq 0, \quad \alpha + \beta > 0$ точка $\alpha x + \beta y \in K$. Действительно, при $\alpha = 0$ или $\beta = 0$ это выполнено тривиально. Если же $\alpha > 0$ и $\beta > 0$, то по определению конуса $\alpha x, \ \beta y \in K$, и тогда по условию $\alpha x + \beta y \in K$. В частности, это выполнено при $\alpha + \beta = 1$, то есть для выпуклой комбинации точек x, y. Следовательно, K — выпуклый, ч. т. д. □

Таким образом, для конуса проверка его выпуклости проще, чем для произвольного множества: не надо рассматривать всевозможные выпуклые комбинации; достаточно рассматривать лишь суммы любых двух его точек.

Отметим следующие очевидные свойства конусов:

1) Пересечение любого числа конусов — конус.
2) Если K — конус, то его выпуклая оболочка $\mathrm{co}\, K$ — выпуклый конус.

(Доказать самостоятельно).

Упражнение 3. Показать, что $\mathrm{co}\, K \supset K + K$, и обратное включение выполнено не всегда. Если конусы K_1 и K_2 выпуклы, то $\mathrm{co}\,(K_1 \cup K_2) = K_1 + K_2$.

Важным примером конуса служит полупространство $(a, x) \leq 0$ или $(a, x) < 0$, где $a \in \mathbb{R}^n$. (Доказать, что если полупространство $(a, x) \leq c$ или $(a, x) < c$ является конусом, то $c = 0$.)

Если задано произвольное семейство векторов $a_\mu \in \mathbb{R}^n, \ \ \mu \in \mathcal{M}$, то

$$K = \ \{x \mid (a_\mu, x) \leq 0, \ \ \forall\, \mu \in \mathcal{M}\} \qquad (1.3)$$

есть, очевидно, замкнутый выпуклый конус (как пересечение замкнутых выпуклых конусов). Если множество индексов $\mathcal{M}$ конечно, то такой конус K называется полиэдральным или многогранным (конечногранным).

С каждым выпуклым множеством можно связать некоторые конусы. Например, для любого множества A можно построить наименьший конус, его содержащий. Будем обозначать его $con\, A$ и называть конической оболочкой множества A. Легко сообразить, что

$$con\, A = \ \{\alpha x \mid x \in A, \ \ \alpha > 0\}. \qquad (1.4)$$

(Показать, что если A — выпуклое множество, то $con\, A$ — выпуклый конус.)

Операция перехода от множества A к конусу $con\, A$ является во многих случаях полезной. Однако нетрудно заметить, что она не взаимно однозначна: множества A_1

и A_2 могут быть разными, а конуса $con\,A_1$ и $con\,A_2$ совпадающими. (Приведите примеры).

Укажем здесь еще один способ перехода от множества к конусу, который уже является взаимно однозначным. Пусть $A \subset \mathbb{R}^n$. Рассмотрим в $\mathbb{R}^{n+1}$ множество $(A, 1)$, состоящее из элементов вида $(x, 1)$, где $x \in A$, и возьмем его коническую оболочку:

$$K = \ con\,(A, 1) = \ \{\,\alpha(x, 1) \mid \ \alpha > 0, \ \ x \in A\}. \tag{1.5}$$

(Иногда бывает удобно добавить сюда точку 0.)

Упражнение 4. Проверить, что эта операция взаимно однозначна. Доказать, что если A — выпукло, то конус K получается выпуклым.

Этот прием полезен тем, что позволяет любое утверждение относительно множеств свести к утверждению относительно конусов. А для выпуклых конусов многие утверждения формулируются проще, чем для произвольных выпуклых множеств. (Например, как уже отмечалось, для конуса вида $(a, x) \le c$ всегда $c = 0$).

С понятием выпуклого конуса связаны следующие комбинации точек.
Вектор (или точку) $\alpha_1 x_1 + \ldots + \alpha_k x_k$ будем называть *положительной комбинацией* точек $x_1, \ldots, x_k$, если все $\alpha_i > 0$. Если K — выпуклый конус, то он, очевидно, содержит все положительные комбинации любых своих элементов.

Пусть $A \subset \mathbb{R}^n$. Выпуклой конической оболочкой множества A называется наименьший выпуклый конус, содержащий A. Будем обозначать его $cocon\,A$.

Утверждение 1.2. $cocon\,A$ есть а) множество всех положительных комбинаций элементов множества A; б) пересечение всех выпуклый конусов, содержащих множество A. (Доказать самостоятельно.)

Наряду с положительными рассматриваются также *позитивные комбинации* $\alpha_1 x_1 + \ldots + \alpha_k x_k$, где все $\alpha_i \ge 0$. Позитивная комбинация называется *нетривиальной,* если $\sum \alpha_i > 0$ (т.е. хотя бы одно $\alpha_i > 0$). Утверждение 1.2 говорит, что $cocon\,A$ есть множество всех нетривиальных позитивных комбинаций элементов множества A.

Вернемся к построению выпуклой оболочки произвольного множества $A \subset \mathbb{R}^n$. Доказанная выше лемма 1.2 утверждает, что $co\,A$ состоит из всевозможных выпуклых комбинаций элементов множества A; при этом требуется перебирать все возможные длины комбинаций. Следующая замечательная теорема позволяет ограничить число элементов, участвующих в этих выпуклых комбинациях.

Теорема 1.2 (Каратеодори). Пусть $A \subset \mathbb{R}^n$ — любое множество. Тогда $co\,A$ состоит из всевозможных выпуклых комбинаций не более чем $n+1$ элемента множества A. Т.е. $\forall\, y \in co\,A$ найдутся точки $x_1, \ldots, x_k \in A$, где $k \le n+1$, и коэффициенты $\alpha_1, \ldots, \alpha_k \ge 0$, $\sum \alpha_i = 1$, такие что $y = \alpha_1 x_1 + \ldots + \alpha_k x_k$.

Как мы видим, уже в самой формулировке этой теоремы участвует размерность пространства; она используется и в доказательстве.

Доказательство. Возьмем произвольный $y \in co\, A$. По лемме 1.1 $y = \alpha_1 x_1 + \ldots + \alpha_k x_k$, где все $x_i \in A$, $\alpha_i \geq 0$ и $\sum \alpha_i = 1$. Среди всех таких представлений вектора y возьмем такое, у которого k — минимально возможное для данного y. (При этом, очевидно, все $\alpha_i > 0$.) Нам надо доказать, что $k \leq n+1$.

Допустим, что $k > n+1$. Тогда рассмотрим следующую систему линейных уравнений относительно вектора $\beta = (\beta_1, \ldots, \beta_k)$:

$$\beta_1 x_1 + \ \ldots \ + \beta_k x_k = 0, \tag{1.6}$$

$$\beta_1 + \ \ldots \ + \beta_k = 0. \tag{1.7}$$

Здесь имеется k неизвестных и $n+1$ уравнение, причем $k > n+1$, поэтому существует ненулевое решение $(\beta_1, \ldots, \beta_k)$. В силу равенства (1.7) среди этих β_i имеются $\beta_i < 0$. Для любого $\varepsilon > 0$ рассмотрим новые коэффициенты $\alpha_i' = \alpha_i + \varepsilon \beta_i$. Для них по-прежнему будем иметь равенства

$$(\alpha_1 + \varepsilon\beta_1)\, x_1 + \ \ldots \ + (\alpha_k + \varepsilon\beta_k)\, x_k \ = y, \tag{1.8}$$

$$(\alpha_1 + \varepsilon\beta_1) + \ \ldots \ + (\alpha_k + \varepsilon\beta_k) \ = \ 1. \tag{1.9}$$

Так как все $\alpha_i > 0$, то при малых ε и новые $\alpha_i' > 0$, поэтому это действительно выпуклая комбинация. Однако, поскольку некоторые $\beta_i < 0$, то при росте ε соответствующие α_i' убывают, и наступит такой момент $\varepsilon_0 > 0$, что для некоторого i_0 (не обязательно единственного) будет $\alpha_{i_0}' = \alpha_{i_0} + \varepsilon_0 \beta_{i_0} = 0$, а остальные $\alpha_i' \geq 0$.
Но тогда точка x_{i_0} не участвует в выпуклой комбинации (1.8), следовательно, число k можно уменьшить. А это противоречит предположению о минимальности k, и тем самым теорема доказана. $\square$

Из этой теоремы сразу вытекает следующий весьма полезный факт. (Отметим, что он имеет место только в конечномерном случае.)

Следствие. Выпуклая оболочка компакта, лежащего в конечномерном пространстве, есть компакт.

Доказательство. Пусть A — компакт, и пусть дана произвольная последовательность точек $y^{(k)} \in coA$, $k = 1, 2, \ldots$. Нам надо показать, что она содержит подпоследовательность, сходящуюся к некоторой точке $y^{(0)} \in \mathrm{co}\, A$. По теореме Каратеодори $\forall\, k$ имеем

$$y^{(k)} = \alpha_1^{(k)} x_1^{(k)} + \ldots + \alpha_{n+1}^{(k)} x_{n+1}^{(k)}\,,$$

где вектор $\alpha^{(k)} = (\alpha_1^{(k)}, \ldots, \alpha_{n+1}^{(k)})$ принадлежит множеству $\sum = \{\alpha \in \mathbb{R}^{n+1} \mid \alpha_i \geq 0,\ \sum \alpha_i = 1\}$, а все $x_i^{(k)} \in A$. Множество $\sum \subset \mathbb{R}^{n+1}$ замкнуто и ограничено, поэтому оно является компактом. ß

Так как $\sum$ и A — компакты, то, переходя к подпоследовательности, можно считать, что при $k \to \infty$ векторы $\alpha^{(k)} \to \alpha^{(0)} \in \sum$ и $\forall i$ точки $x_i^{(k)} \to x_i^{(0)} \in A$. Тогда $y^{(k)}$ сходится к точке

$$y^{(0)} = \alpha_1^{(0)} x_1^{(0)} + \ldots + \alpha_{n+1}^{(0)} x_{n+1}^{(0)},$$

которая лежит в $\operatorname{co} A$ как выпуклая комбинация точек $x_i^{(0)} \in A$, ч.т.д. $\square$

Можно дать и другое (в некотором смысле более простое) доказательство. Рассмотрим компакт $Q = \sum \times A^{n+1}$, на котором зададим отображение

$$\pi : Q \to \mathbb{R}^n, \qquad \pi(\alpha, x_1, \ldots, x_{n+1}) = \alpha_1 x_1 + \ldots + \alpha_{n+1} x_{n+1} .$$

Ясно, что π непрерывно, поэтому $\pi(Q)$ — компакт, а по теореме Каратеодори $\pi(Q) = \operatorname{co} A$, следовательно, $\operatorname{co} A$ — компакт. $\square$

Задачи.

1.0. Показать, что любой непустой выпуклый компакт на прямой есть отрезок $[a, b]$ при некоторых $a \le b$.

1.1. (Теорема Каратеодори для конусов.) Пусть A — конус в $\mathbb{R}^n$. Тогда $co\, A$ состоит из всевозможных нетривиальных позитивных комбинаций не более чем n элементов множества A.

1.2. Доказать с помощью этой теоремы теорему Каратеодори для произвольного множества. (Воспользоваться вышеописанным приемом перехода от множества A к конусу $con\,(A, 1)$.)

1.3*. Пусть A — связное множество в $\mathbb{R}^n$. Тогда опять в теореме Каратеодори можно заменить $n + 1$ на n.

1.4*. Привести пример компакта в бесконечномерном пространстве, выпуклая оболочка которого не компактна.

Лекция 2. Топологические свойства выпуклых множеств.

Выпуклое множество в $\mathbb{R}^n$ обладает рядом топологических свойств, т.е. свойств, связанных с понятием близости, непрерывности и т.п., которые не имеют места для произвольного множества. Установим здесь некоторые из них.

Пусть A — выпуклое множество в $\mathbb{R}^n$.

Свойство 1. Внутренность $intA$ и замыкание $\overline{A}$ множества A выпуклы.

Доказательство. Если $x_1, x_2 \in \operatorname{int} A$, то по определению $\exists \varepsilon > 0$ такое, что $x_1 + B_\varepsilon \subset A$ и $x_2 + B_\varepsilon \subset A$. Тогда для любых $\alpha_1, \alpha_2 \geq 0$, $\alpha_1 + \alpha_2 = 1$ в силу выпуклости A имеем $\alpha_1(x_1 + B_\varepsilon) + \alpha_2(x_2 + B_\varepsilon) \subset A$. Раскрывая здесь скобки и учитывая выпуклость шара B_ε (поэтому $\alpha_1 B_\varepsilon + \alpha_2 B_\varepsilon = B_\varepsilon$), получаем $(\alpha_1 x_1 + \alpha_2 x_2) + B_\varepsilon \subset A$, откуда вытекает, что $\alpha_1 x_1 + \alpha_2 x_2 \in \operatorname{int} A$, следовательно, $\operatorname{int} A$ выпукло.

Докажем выпуклость $\overline{A}$. Пусть $x_1, x_2 \in \overline{A}$, и пусть даны $\alpha_1, \alpha_2 \geq 0$, $\alpha_1 + \alpha_2 = 1$. Покажем, что $x = \alpha_1 x_1 + \alpha_2 x_2 \in \overline{A}$. По определению $\forall \varepsilon > 0$ $\exists z_1, z_2 \in B_\varepsilon$ такие, что $x_1 + z_1 \in A$ и $x_2 + z_2 \in A$. Тогда $z = \alpha_1 z_1 + \alpha_2 z_2 \in B_\varepsilon$ (в силу выпуклости B_ε) и $x + z = \alpha_1(x_1 + z_1) + \alpha_2(x_2 + z_2) \in A$ (в силу выпуклости A). Отсюда в силу произвольности $\varepsilon > 0$ следует, что $x \in A$, ч.т.д. $\square$

Свойство 2. Если $x_0 \in \operatorname{int} A$ и $x_1 \in A$, то $(x_0, x_1) \subset \operatorname{int} A$. Более того, если $\varepsilon > 0$ таково, что $B_\varepsilon(x_0) \subset A$, то $\forall \alpha \in (0, 1)$ для точки $y = \alpha x_0 + (1 - \alpha) x_1$ шар радиуса $\alpha\varepsilon$ содержится в A : $B_{\alpha\varepsilon}(y) \subset A$.

Доказательство. Пусть $B_\varepsilon(x_0) \subset A$ и $y = \alpha x_0 + (1 - \alpha) x_1$, где $\alpha \in (0, 1)$. Так как $x_1 \in A$, то в силу выпуклости A имеем $\alpha[x_0 + B_\varepsilon(0)] + (1 - \alpha) x_1 \subset A$, т.е. $(\alpha x_0 + (1 - \alpha) x_1) + \alpha B_\varepsilon(0) \subset A$, т.е. $y + B_\alpha \varepsilon(0) \in A$, ч. т. д. $\square$

Свойство 2 допускает следующее усиление.

Свойство $2'$. Если $x_0 \in \operatorname{int} A$ и $x_1 \in \overline{A}$, то по-прежнему $(x_0, x_1) \subset \operatorname{int} A$. Если при этом $B_\varepsilon(x_0) \subset A$, то $\forall \alpha \in (0, 1)$ для точки $y = \alpha x_0 + (1 - \alpha) x_1$ выполняется включение $B_{\alpha\varepsilon'}(y) \subset A$ при любом $\varepsilon' < \varepsilon$, т.е. открытый шар радиуса $\alpha\varepsilon$ с центром в y содержится в A .

Доказательство. Пусть $B_\varepsilon(x_0) \subset A$ и $y = \alpha x_0 + (1 - \alpha) x_1$, где $\alpha \in (0, 1)$. Нам надо показать, что $y \in \operatorname{int} A$. Так как $x_1 \in \overline{A}$, то $x_1 = \lim z_k$ для некоторой последовательности точек $z_k \in A$. Положим $y_k = \alpha x_0 + (1 - \alpha) z_k$. По доказанному свойству 2 $B_{\alpha\varepsilon}(y_k) \subset A$. Поскольку α, ε не зависят от k, а $y_k \to y$, то $B_{\alpha\varepsilon}(y_k) \to B_{\alpha\varepsilon}(y)$ по Хаусдорфу, а тогда для любого $\varepsilon' < \varepsilon$ при достаточно больших k будет $B_{\alpha\varepsilon'}(y) \subset B_{\alpha\varepsilon}(y_k) \subset A$, и следовательно, $y \in \operatorname{int} A$, ч. т. д.

Свойство 3. Если $\operatorname{int} A$ непусто, то $\overline{\operatorname{int} A} = \overline{A}$.

Доказательство. Так как $\operatorname{int} A \subset A$, то включение $\subset$ очевидно. Докажем обратное включение, т.е. покажем, что если $x_1 \in \overline{A}$, то $x_1 \in \overline{\operatorname{int} A}$. Зафиксируем произвольную точку $x_0 \in \operatorname{int} A$. Любая точка $x_1 \neq x_0$ является, очевидно, предельной для интервала (x_0, x_1). Если при этом $x_1 \in \overline{A}$, то по свойству 2' $(x_0, x_1) \subset \operatorname{int} A$, поэтому x_1 является предельной точкой множества $\operatorname{int} A$, т.е. $x_1 \in \overline{\operatorname{int} A}$, ч.т.д. (Для $x_1 = x_0$ это выполнено тривиальным образом.) $\Box$

Свойство 4. Пусть $x_0 \in \operatorname{int} A$, $x_1 \in \partial A = \overline{A} \setminus \operatorname{int} A$. Тогда $\forall \alpha > 1$ точка $x_\alpha = (1-\alpha)x_0 + \alpha x_1 \notin \overline{A}$.

Доказательство вытекает из того очевидного факта, что точка x_1 лежит в интервале между x_0 и x_α, поэтому, если $x_\alpha \in \overline{A}$, то по свойству 2' должно быть $x_1 \in \operatorname{int} A$, что противоречит условию. $\Box$

Вместе со свойством $2'$ это свойство дает точную границу тех $\alpha > 0$, для которых при движении вдоль луча, выходящего из точки x_0 в направлении x_1, точка $x_\alpha = x_0 + \alpha(x_1 - x_0)$ остается во множестве A .

Перейдем теперь к установлению того факта, что выпуклое множество в некотором смысле "всегда" имеет непустую внутренность. Напомним, что в предыдущей лекции было введено понятие аффинной оболочки произвольного множества A — это есть наименьшее аффинное подпространство, содержащее A. Для нее справедливы следующие простые свойства:

а) $\forall x \in \mathbb{R}^n \qquad \operatorname{Aff}(x + A) = x + \operatorname{Aff} A$.

б) $\forall x_0 \in A \qquad \operatorname{Aff} A = x_0 + \operatorname{Lin}(A - x_0)$.

в) Если $A \ni 0$, то $\operatorname{Aff} A = \operatorname{Lin} A$.

Их доказательства мы оставляем в качестве упражнения. Напомним также следующее понятие.

Определение 2.1. Точки $x_0, x_1, \ldots, x_m \in \mathbb{R}^n$ называются аффинно зависимыми, если существует нетривиальный набор чисел $\lambda_0, \lambda_1, \ldots, \lambda_m$ такой, что $\sum \lambda_i = 0$ и $\sum \lambda_i x_i = 0$.

Это эквивалентно тому, что одна из данных точек является аффинной комбинацией остальных: если, например, $\lambda_0 \neq 0$, то, считая $\lambda_0 = -1$, имеем $-x_0 + \sum_1^m \lambda_i x_i = 0$, $-1 + \sum_1^m \lambda_i = 0$, т.е. $x_0 = \sum_1^m \lambda_i x_i$, где $\sum_1^m \lambda_i = 1$. (Обратное очевидно.)

Если указанного набора чисел λ_i не существует, то точки x_i называются аффинно независимыми. (Набор, состоящий из любой одной точки x_0, по определению считается аффинно независимым.) С понятием линейной независимости это понятие связано следующим образом.

Лемма 2.1. Точки $x_0, x_1, \ldots, x_m$ аффинно независимы тогда и только тогда, когда вектора $(x_1 - x_0), \ldots, (x_m - x_0)$ линейно независимы.

Доказать самим в качестве простого упражнения.

Как известно, если вектора линейно независимы, то и любые достаточно близкие к ним вектора также будут линейно независимыми. (Этот факт вытекает, например, из того, что линейная независимость векторов означает, что в любом базисе матрица, составленная из их координат, имеет минор порядка m с ненулевым детерминантом, а детерминант матрицы есть непрерывная функция ее коэффициентов.) Поэтому из леммы 2.1 вытекает

Следствие. Если точки $x_0, x_1, \ldots, x_m$ аффинно независимы, то и любые достаточно близкие к ним точки также будут аффинно независимыми.

Определение 2.2. $m-$ мерным симплексом называется выпуклая оболочка любого набора из $m+1$ аффинно независимой точки: $\Sigma^m = co\{x_0, x_1, \ldots, x_m\}$. Точки $x_0, x_1, \ldots, x_m$ называются вершинами симплекса.

В частности, нульмерный симплекс $\Sigma^0 = \{x_0\}$ есть просто произвольная точка в $\mathbb{R}^n$, одномерный симплекс $\Sigma^1 = [x_0, x_1]$ есть отрезок, двумерный симплекс $\Sigma^2 = co\{x_0, x_1, x_2\}$ — треугольник, и т. д.

Симплекс обладает следующими простыми, но важными свойствами. Во-первых, любая точка $x \in \Sigma^m$ единственным образом представляется в виде выпуклой комбинации его вершин: $x = \alpha_0 x_0 + \ldots + \alpha_m x_m$, где $\alpha_i \geq 0$, $\alpha_0 + \ldots + \alpha_m = 1$.

Существование такого представления вытекает из свойств выпуклой оболочки (лемма 1.2), а единственность устанавливается следующим образом. Пусть имеется другое представление $x = \beta_0 x_0 + \ldots + \beta_m x_m$, где $\beta_i \geq 0$, $\beta_0 + \ldots + \beta_m = 1$. Вычитая его из первого представления, получим

$$(\alpha_0 - \beta_0)\, x_0 + \ \ldots \ + (\alpha_m - \beta_m)\, x_m = 0,$$

$$(\alpha_0 - \beta_0) + \ \ldots \ + (\alpha_0 - \beta_m) \ = \ 0,$$

откуда в силу аффинной независимости точек x_i следует, что $(\alpha_0 - \beta_0) = \ldots = (\alpha_m - \beta_m) = 0$, т.е. оба этих представления совпадают.

Коэффициенты $\alpha_0, \ldots, \alpha_m$ называются *барицентрическими координатами* точки x в симплексе Σ^m. Это название объясняется тем фактом, что если считать x_i материальными точками массы α_i, то центр масс полученной системы материальных точек будет находиться в точке $x = \alpha_0 x_0 + \ldots + \alpha_m x_m$. (Проверьте!)

Далее, из приведенного определения ясно, что аффинная оболочка симплекса Σ^m имеет размерность m. Если $m = n$, то аффинная оболочка есть все пространство $\mathbb{R}^n$, поэтому можно поставить вопрос о внутренности симплекса.

Лемма 2.2. В пространстве $\mathbb{R}^n$ любой n-мерный симплекс Σ^n имеет непустую внутренность. Она состоит из тех и только тех точек, у которых все барицентрические координаты положительны.

Для доказательства заметим, что при невырожденном аффинном преобразовании пространства внутренние точки любого множества переходят во внутренние, $n-$ мерный симплекс переходит в $n-$ мерный симплекс, а барицентрические координаты точки при этом не меняются. (Проверьте!) Поэтому достаточно рассмотреть стандартный симплекс $\Sigma^n = co\{0, e_1, \ldots, e_n\}$, порожденный началом координат и всеми базисными векторами (из которого любой другой $n-$ мерный симплекс получается некоторым невырожденным аффинным преобразованием — докажите!). Для него уже очевидно, что $\operatorname{int} \Sigma^n$ состоит из тех точек $x = (x^1, \ldots, x^n)$, у которых все координаты $x^i > 0$ и $\sum x^i < 1$ (здесь $\alpha_0 = 1 - \sum x^i > 0$), т.е. утверждение леммы выполнено. □

Таким образом, для симплекса справедливо следующее свойство: если размерность его аффинной оболочки равна n, то он содержит некоторый n -мерный шар. Для произвольного множества это свойство, конечно же, не имеет места; контрпример — два пересекающихся отрезка на плоскости ("крест"). Обратное же утверждение справедливо и для произвольного множества.

В случае $m < n$ симплекс, конечно, не будет иметь внутренности (почему?), однако он будет иметь так называемую *относительную внутренность*. Напомним, что это такое.

Пусть L есть аффинное подпространство в $\mathbb{R}^n$, и дано множество $A \subset L$.

Определение 2.3. Точка $x \in A$ называется внутренней точкой множества A относительно подпространства L, если существует такая окрестность $U(x)$ точки x, что $U(x) \cap L \subset A$. (Обычно в качестве $U(x)$ рассматривают открытый шар некоторого радиуса с центром в x.)

Множество всех таких точек в A будем обозначать $\operatorname{reint}_L A$ и называть внутренностью множества A относительно подпространства L, или просто относительной внутренностью множества A, если ясно о каком L идет речь. Если L не указано, подразумевается, что $L = \operatorname{Aff} A$.

(Понятие относительной внутренности определяется и в более общей ситуации: вместо $L \subset \mathbb{R}^n$ можно рассматривать произвольное подмножество произвольного топологического пространства; см. [88]. Однако для наших целей вполне достаточно будет приведенного определения.)

Из определения 2.3 и леммы 2.2 сразу вытекает

Лемма 2.3. В пространстве $\mathbb{R}^n$ любой $m-$ мерный симплекс $(m \leq n)$ имеет непустую относительную внутренность, и она состоит из тех и только тех точек, у которых все барицентрические координаты положительны.

Действительно, если $m < n$, то, считая без ограничения общности, что $0 \in \Sigma^m$, получим, что аффинная оболочка симплекса Σ^m совпадает с его линейной оболочкой $L = \operatorname{Lin} \Sigma^m$ (по свойству "в" аффинной оболочки). Так как при этом $dim\, L = m$, то можно считать, что L совпадает с координатным подпространством, порожденным первыми m базисными векторами $e_1, \ldots, e_m$, и можно все рассмотрения проводить в этом пространстве $\mathbb{R}^m$. Но по лемме 2.2 в пространстве $\mathbb{R}^m$ любой $m-$ мерный симплекс имеет непустую внутренность (которая задается положительностью барицентрических координат), а она же и есть относительная внутренность симплекса Σ^m, рассматриваемого в пространстве $\mathbb{R}^n$. $\square$

Таким образом, для симплекса указанное выше свойство о содержащемся шаре справедливо в следующей усиленной форме : если размерность его аффинной оболочки равна m, то он содержит некоторый $m-$ мерный шар.

Как мы уже видели, для произвольного множества это свойство очевидно неверно: $\operatorname{Aff} A$ может совпадать с $\mathbb{R}^n$, но A может не содержать ни одного n-мерного шара. Однако оно остается верным для выпуклых множеств, и в некотором смысле это максимально широкий класс множеств, для которого данное свойство справедливо.

Теорема 2.1. В конечномерном пространстве любое выпуклое множество имеет непустую внутренность относительно своей аффинной оболочки.

Доказательство. Без нарушения общности считаем, что $0 \in A \subset \mathbb{R}^n$, и что $\operatorname{Aff} A = Lin\, A = \mathbb{R}^n$. Покажем, что тогда $\operatorname{int} A$ непусто. Так как $\operatorname{Lin} A = \mathbb{R}^n$, то найдутся n линейно независимых векторов $x_1, \ldots, x_n \in A$ (доказать!). При этом точки $0, x_1, \ldots, x_n$ будут аффинно независимы, и следовательно, их выпуклая оболочка есть n-мерный симплекс Σ^n, содержащийся в A. Но по лемме 2.2 внутренность этого симплекса непуста, а тогда непуста и внутренность множества A. $\square$

Определение 2.4. Размерностью выпуклого множества называется размерность его аффинной оболочки:

$$\dim A = \dim(\operatorname{Aff} A).$$

Это определение оправдано тем, что если $\dim A = m$, т.е. $L = \operatorname{Aff} A$ имеет размерность m, то по теореме 2.2 $\exists\, x \in L$, $\varepsilon > 0$ такие, что m-мерный шарик $L \cap B_\varepsilon(x)$ содержится в A. Обратим внимание, что размерность выпуклого множества не зависит от того, в каком объемлющем пространстве мы его рассматриваем; например, круг двумерен и на плоскости, и будучи помещен в трехмерное или любое другое пространство.

Из установленных выше свойств вытекает

Свойство 5. Размерность выпуклого множества совпадает с максимальной размерностью содержащихся в нём симплексов. (Докажите.)

Продолжим изучение топологических свойств выпуклых множеств. Нам потребуется следующий простой факт.

Лемма 2.4. Пусть U — произвольное открытое множество в $\mathbb{R}^n$, и пусть $x \in U$. Тогда найдется симплекс $\Sigma \subset U$ такой, что $x \in \operatorname{int} \Sigma$. (Ясно, что этот симплекс n–мерный.)

Доказательство. Считая $x = 0$, достаточно доказать, что в $\mathbb{R}^n$ существует симплекс Σ, у которого ноль есть внутренняя точка (ибо тогда при достаточно малом $\alpha > 0$ получим $\alpha\Sigma \subset U$), а это эквивалентно тому, что существует симплекс Σ, у которого внутренность непуста. Но по лемме 2.2 таким свойством обладает любой n-мерный симплекс. Лемма доказана. $\square$

Пусть теперь, как и раньше, A — выпуклое множество.

Свойство 6. $\operatorname{int} A = \operatorname{int}(\overline{A})$. (Сравнить со свойством 3.)

Доказательство. Включение $\subset$ очевидно. Докажем $\supset$. Пусть $x \in \operatorname{int}(\overline{A})$. Тогда по лемме 2.4, примененной к $U = \operatorname{int}(\overline{A})$, найдется симплекс Σ с вершинами $a_1, \ldots, a_{n+1} \in U$, такой, что $x \in \operatorname{int}\Sigma$, т.е. $B_r(x) \subset \Sigma$ при некотором $r > 0$. Так как все $a_i \in U \subset \overline{A}$, то по определению $\forall\, \varepsilon > 0$ найдутся точки $a_i' \in A$, такие что $|\, a_i' - a_i \,| < \varepsilon$. Если ε достаточно мало, то симплекс $\Sigma' = co\,\{a_1', \ldots, a_{n+1}'\}$ отличается (по Хаусдорфу) от симплекса Σ меньше, чем на $r/2$, и тогда он содержит шар $B_{r/2}(x)$, поэтому $x \in \operatorname{int}\Sigma'$, и так как $\Sigma' \subset A$, то $x \in \operatorname{int} A$, ч. т. д. $\square$

Другой способ доказательства. Как уже отмечалось, надо доказать лишь включение $\supset$. Если $A \subset \mathbb{R}^n$ и $L = \operatorname{Aff} A \neq \mathbb{R}^n$, то $A \subset L$, значит и $\overline{A} \subset L$, а тогда A и $\overline{A}$ оба имеют пустую внутренность, и требуемое равенство доказано. Поэтому далее считаем, что Aff A $= \mathbb{R}^n$, и следовательно, $\operatorname{int} A$ непуста, т.е. $\exists\, x_0 \in \operatorname{int} A$.

Возьмем произвольный $x_1 \in \operatorname{int}(\overline{A})$. Тогда $x_1 \in \overline{A}$ и более того, некоторая его окрестность $\mathcal{O}(x_1) \subset \overline{A}$. Допустим, что $x_1 \notin \operatorname{int} A$. Тогда по определению $x_1 \in \partial A = \overline{A} \setminus \operatorname{int} A$, и по свойству 4 $\forall\, \alpha > 1$ точка $x_\alpha = (1-\alpha)x_0 + \alpha x_1 \notin \overline{A}$. Но с другой стороны, $x_\alpha \to x_1$ при $\alpha \to 1+0$, поэтому $x_\alpha \in \mathcal{O}(x_1) \subset \overline{A}$ при α достаточно близком к 1, следовательно $x_\alpha \in \overline{A}$, противоречие. $\square$

Следствие. Для выпуклого множества справедливо равеннство $\partial A = \partial(\overline{A})$.

Действительно, пользуясь определением границы множества и свойством 6, получаем: $\partial(\overline{A}) = \overline{A} \setminus \operatorname{int}(\overline{A}) = \overline{A} \setminus \operatorname{int} A = \partial A$, ч. т. д.

Свойство 7. Пусть A выпукло и $\operatorname{int} A \neq \emptyset$. Тогда $\operatorname{int} A$ есть наименьшее выпуклое множество C такое, что $\overline{A} \subset \overline{C}$ (при этом для $C = \operatorname{int} A$ выполнено равенство $\overline{A} = \overline{C}$).

Доказательство. По свойству 1 множество $C = \operatorname{int} A$ выпукло, и по свойству 3 для него $\overline{C} = \overline{A}$. Покажем теперь, что если выпуклое множество C таково, что $\overline{A} \subset \overline{C}$, то $\operatorname{int} A \subset C$. Действительно, пусть $\overline{C} \supset \overline{A}$. Тогда $\overline{C} \supset A \supset \operatorname{int} A$, а это открытое множество, следовательно, $\operatorname{int}(\overline{C}) \supset \operatorname{int} A$. Но по свойству 6 $\operatorname{int}(\overline{C}) = \operatorname{int} C$, и тогда $\operatorname{int} A \subset \operatorname{int} C \subset C$, ч. т. д. $\square$

Задачи. 1) Показать, что для произвольных (не выпуклых) множеств свойства 1–7, вообще говоря, не имеют места.

2) Пусть A — выпукло, $x_0 \in \operatorname{int} A$. Тогда для любого аффинного подпространства $L \ni x_0$ справедливо равенство $reint\,(A \cap L) = (\operatorname{int} A) \cap L$. Показать, что условие $x_0 \in \operatorname{int} A$ здесь существенно.

3) Пусть A — выпукло, $\operatorname{int} A$ непусто, L – аффинное подпространство, π – проекция на L вдоль некоторого его дополнения. Тогда $reint\,\Pi(A) = \Pi\,(\operatorname{int} A)$. Показать, что условие $\operatorname{int} A \neq \emptyset$ существенно.

4) Пусть A, B – выпуклы, $int\,A \cap int\,B$ непусто. Тогда $\overline{A \cap B} = \overline{A} \cap \overline{B}$.

Лекция 3. Теоремы об отделимости выпуклых множеств

Будем рассматривать на пространстве $\mathbb{R}^n$ линейные функционалы. Нам будет удобно считать, что в $\mathbb{R}^n$ задана евклидова структура, т.е. скалярное произведение; тогда линейные функционалы можно отождествить с векторами нашего пространства, и в любом ортонормированном базисе значение функционала p на векторе x есть скалярное произведение $(p,x) = \sum p_i x_i$, где p_i, x_i — координаты векторов p, x в данном базисе.

Пусть A и $B -$ два произвольных непустых множества в $\mathbb{R}^n$. Введем следующее понятие, которое является основным для всей теории выпуклого анализа.

Определение 3.1. Множества A и B называются отделимыми друг от друга, если существует ненулевой функционал $p \in \mathbb{R}^n$ такой, что

$$\sup\,(p,A) \;\le\; \inf\,(p,B), \tag{3.1}$$

или, что то же самое,

$$\forall\, x \in A,\;\; \forall\, y \in B \qquad (p,x) \le (p,y), \tag{3.2}$$

или, что также эквивалентно: $\exists\, c$ такое, что

$$(p,A) \le c \le (p,B). \tag{3.3}$$

(Мы пишем для краткости (p,A) для обозначения множества значений (p,x), когда x пробегает множество A.)

Геометрически это означает, что множества A и B лежат по разные стороны от гиперплоскости $(p,x) = c$, или точнее, лежат в разных замкнутых полупространствах, на которые гиперплоскость $(p,x) = c$ разбивает все пространство $\mathbb{R}^n$ (рис. 3.1).

Определение 3.2. Множества A и B называются строго отделимыми друг от друга, если существует функционал $p \in \mathbb{R}^n$ такой, что

$$\sup\,(p,A) \;<\; \inf\,(p,B), \tag{3.4}$$

или: $\exists\, \delta > 0$, при котором

$$\forall\, x \in A,\; \forall\, y \in B \qquad (p,x) \le (p,y) - \delta, \tag{3.5}$$

или: $\exists\, c' < c''$, для которых

$$(p,A) \le c' < c'' \le (p,B). \tag{3.6}$$

(Ясно, что при выполнении любого из этих свойств $p \neq 0$.)

Строгая отделимость означает, что между данными множествами можно вставить "пластинку" $c' \le (p,x) \le c''$ ненулевой толщины (рис. 3.2).

Упражнение 3.1. Проверить, что свойства (3.1)–(3.3) действительно эквивалентны между собой, и также эквивалентны между собой свойства (3.4)–(3.6).

Упражнение 3.2. Показать, что если A и $B -$ конусы, и они отделимы друг от друга, то в неравенстве (3.3) константа $c = 0$, т.е. отделяющая их гиперплоскость проходит через 0. Строго отделимыми конусы быть не могут.

Во все указанные условия оба множества A и B входят симметрично (проверить!), но часто говорят, что A отделимо от B (подразумевая, что и B отделимо от A). Например, если A состоит из одной точки x_0, то удобно говорить, что точка x_0 отделима от множества B.

Прежде чем двигаться дальше, установим следующее простое свойство, которым обладают замкнутые множества в конечномерном пространстве.

Лемма 3.1. Пусть F — произвольное непустое замкнутое множество в $\mathbb{R}^n$, и $||\cdot||$ — произвольная норма. Тогда $\forall x_0 \in \mathbb{R}^n$ во множестве F существует хотя бы одна точка y, ближайшая к x_0. Другими словами, существует точка, в которой функция $\rho(x) = ||x - x_0||$ достигает своего минимума на множестве F.

На этом свойстве основаны все теоремы об отделимости выпуклых множеств в конечномерном пространстве.

Доказательство. Возьмем замкнутый шар $B_r(x_0)$ достаточно большого радиуса, так чтобы он имел непустое пересечение с множеством F; обозначим это пересечение через Q. Ясно, что ближайшую точку надо искать только среди точек Q, ибо $\inf_F \rho = \inf_Q \rho$. Но множество Q замкнуто и ограничено, т.е. компакт, а функция ρ непрерывна, поэтому на Q она достигает своего минимума. □

Задача 3.1. Верна ли эта лемма в бесконечномерных пространствах? Где в приведенном доказательстве использовалась конечномерность пространства? Показать, что в пространствах $L_1[a,b]$, $C[a,b]$ эта лемма неверна.

Перейдем теперь непосредственно к теме данной лекции.

Теорема 3.1 (о строгой отделимости точки от замкнутого выпуклого множества). Пусть имеется непустое замкнутое выпуклое множество $A \subset \mathbb{R}^n$ и точка $x_0 \notin A$. Тогда существует функционал (т.е. вектор) $p \in \mathbb{R}^n$, строго отделяющий x_0 от A, т.е. такой, что

$$\inf\,(p, A) > (p, x_0).$$

Другими словами, $\exists\,\delta > 0$ такое, что

$$\forall\, x \in A \qquad (p,x) \ge (p,x_0) + \delta\,.$$

Доказательство. Пусть $y \in A$ есть ближайшая к x_0 точка из A относительно евклидовой нормы $|\cdot|$ (она существует по лемме 3.1). Это означает по определению, что $\forall\, x \in A \quad |x - x_0|^2 \geq |y - x_0|^2$. Положим $p = y - x_0$. (Отметим сразу, что $p \neq 0$, ибо $x_0 \notin A$.) Зафиксируем любую точку $x \in A$ и положим $\bar{x} = x - y$. Тогда точка $x_\varepsilon = y + \varepsilon \bar{x} \in A$ при любом $\varepsilon \in [0,1]$ (ибо $x_\varepsilon = \varepsilon x + (1-\varepsilon) y$ есть выпуклая комбинация точек x, y), поэтому $(x_\varepsilon - x_0)^2 \geq |p|^2$, а так как

$$x_\varepsilon - x_0 = \ (y - x_0) + \varepsilon \bar{x} \ = p + \varepsilon \bar{x},$$

то последнее неравенство означает, что $\forall\, \varepsilon \in [0,1]$ выполнено $(p + \varepsilon \bar{x})^2 \geq p^2$, т.е.

$$p^2 + \varepsilon (p, \bar{x}) + \varepsilon^2 |\bar{x}|^2 \ \geq \ p^2\,.$$

Тогда $\forall\, \varepsilon > 0$ имеем $(p, \bar{x}) + \varepsilon |\bar{x}|^2 \geq 0$, откуда в пределе при $\varepsilon \to 0+$ получаем $(p, \bar{x}) \geq 0$, т.е. $(p,\, x - y) \geq 0$.

Итак, мы получили, что $(p, x) \geq (p, y) \quad \forall\, x \in A$, или, так как $y = x_0 + p$, что $(p, x) \geq (p, x_0) + |p|^2$. Поскольку $\delta = |h|^2 > 0$ (ибо $p \neq 0$), то это и означает, что точка x_0 строго отделима от A. □

Иллюстрацией теоремы 3.1 служит рис. 3.4, который является ключевым для всего выпуклого анализа. По сути дела, он выражает ту основную идею (и пожалуй, единственную!), на которой строится эта теория (причем не только в конечномерном случае).

Из теоремы 3.1 вытекает следующее важное свойство.

Следствие 1. Любое замкнутое выпуклое множество есть пересечение всех замкнутых полупространств, его содержащих.

Доказательство. Пусть A – замкнутое выпуклое множество. Обозначим через B пересечение всех замкнутых полупространств, содержащих A. Поскольку каждое такое полупространство содержит A, то и $B \supset A$.

Докажем обратное: $A \supset B$, или, что эквивалентно, если $x_0 \notin A$, то $x_0 \notin B$.

Так как $x_0 \notin A$, то по теореме 3.1 найдется вектор $p \in \mathbb{R}^n$, строго отделяющий точку x_0 от множества A :

$$\sup\,(p, A) = \ c \ < (p, x_0).$$

Но тогда полупространство $(p, x) \leq c$ содержит A, т.е. входит в наше семейство, и не содержит x_0, поэтому $x_0 \notin B$, ч.т.д. □

Для произвольного множества A рассмотрение вышеуказанного множества B приводит к следующему свойству. Сначала дадим

Определение 3.3. Выпуклой замкнутой оболочкой множества $A \subset \mathbb{R}^n$ называют наименьшее выпуклое замкнутое множество, содержащее A, т.е. пересечение всех выпуклых замкнутых множеств, содержащих A. Обозначается: $\overline{co}\, A$.

Лемма 3.2. $\overline{co}\,A = \overline{coA} \neq co\,(\overline{A})$. Последнее неравенство выполняется в виде включения $\supset$, которое в общем случае не обратимо. Для ограниченных множеств $\overline{coA} = co\,(\overline{A})$.

(Доказать самим в качестве упражнения.)

Следствие 2. $\overline{co}\,A$ есть пересечение всех замкнутых полупространств, содержащих A.

Доказательство. Пусть, как и выше, B есть пересечение этих полупространств. Поскольку каждое такое полупространство есть частный случай выпуклого замкнутого множества, содержащего A, то по определению оно содержит и $\overline{co}\,A$, а поэтому $B \supset \overline{co}\,A$. Обратно: каждое выпуклое замкнутое множество F, содержащее A, само есть пересечение замкнутых полупространств, содержащих F, а так как каждое из этих полупространств содержит A, то пересечение всех F есть пересечение некоторых полупространств, содержащих A:

$$\bigcap F = \bigcap \{\text{некоторых полупространств, содержащих } A\},$$

и поэтому

$$\bigcap F \supset \bigcap \{\text{всех полупространств, содержащих } A\},$$

то есть $\overline{co}\,A \supset B$, ч.т.д. (см. рис. 3.5). □

Перейдем к следующей теореме отделимости.

Теорема 3.2 (об отделимости точки от выпуклого множества). Пусть $A \subset \mathbb{R}^n$ – произвольное непустое выпуклое множество, и $x_0 \notin A$. Тогда существует вектор $p \neq 0$, отделяющий (вообще говоря, нестрого) точку x_0 от множества A, т.е. $\forall\, x \in A$ будет выполнено $(p, x) \leq (p, x_0)$.

Доказательство. Если $x_0 \notin \overline{A}$, то требуемая отделимость вытекает из теоремы 3.1 (точка x_0 отделяется от замкнутого выпуклого множества $\overline{A}$).

Пусть $x_0 \in \overline{A}$. Так как $x_0 \notin A$, то $x_0 \in \partial A$. (Напомним, что $\partial A = \overline{A} \setminus \operatorname{int} A$.) Но, как мы знаем, $\partial A = \partial \overline{A}$ (см. лекцию 2, свойство 5), т.е. $x_0 \in \partial \overline{A}$. Тогда существует последовательность точек $x_k \notin \overline{A}$, $x_k \to x_0$. По теореме 3.1 каждую точку x_k можно отделить (и даже строго) от замкнутого множества $\overline{A}$, т.е. $\forall\, k$ найдется вектор $p_k \neq 0$ и число $\delta_k > 0$ такие, что $\forall\, x \in A$

$$(p_k,\, x) \;\leq\; (p_k,\, x_k) - \delta_k \;\leq\; (p_k, x_k). \tag{3.8}$$

Без нарушения общности можно считать, что $|p_k| = 1$, а тогда, в силу компактности единичной сферы в $\mathbb{R}^n$, что вектора p_k сходятся к некоторому единичному вектору p_0. При этом $\forall\, x \in A$ из неравенства (3.8), отбрасывая его среднюю часть, в пределе получаем $(p_0, x) \leq (p_0, x_0)$, ч.т.д. □

Замечание. Указанную последовательность $x_k \notin \overline{A}$ можно было бы построить и не используя свойство $\partial A = \partial\overline{A}$. Считая, что $dim\, A = n$ (в противном случае утверждение теоремы тривиально), надо взять любую точку $\hat{x} \in int\, A$. Тогда по свойству 4 выпуклых множеств последовательность $x_k = -\frac{1}{k}\,\hat{x} + (1+\frac{1}{k})\,x_0 \notin \overline{A}$.

Теорема 3.2 также имеет важное следствие.

Определение 3.4. Вектор p называется *опорным* к выпуклому множеству A в точке $x_0 \in \overline{A}$, если $(p, A) \le (p, x_0)$, или, эквивалентно, $(p, \overline{A}) \le (p, x_0)$ (почему?), т.е. если функционал p достигает в точке x_0 своего максимума по множеству $\overline{A}$:

$$x_0 \in \ Argmax\,(p,x) \mid x \in \overline{A}\ .$$

Отметим, что всегда имеется тривиальный опорный вектор $p = 0$. Множество всех опорных векторов к множеству A в точке x_0 является, очевидно, конусом, который называется *конусом внешних нормалей* и обозначается $N_A(x_0)$ (рис. 3.6). Из приведенного определения следует, что множество A и его замыкание $\overline{A}$ имеют в любой точке $x_0 \in \overline{A}$ одни и те же опорные векторы: $N_A(x_0) \ = \ N_{\overline{A}}(x_0)$.

Упражнение 3.3. Доказать, что конус $N_A(x_0)$ выпуклый и замкнутый.

Следствие из теоремы 3.2. В любой граничной точке выпуклого множества существует нетривиальный опорный вектор:

$$x_0 \in \partial A \implies \exists\, p \in N_A(x_0), \quad p \ne 0.$$

Доказательство. Если $dim A < n$, т.е. $\mathrm{Aff}\, A \ne \mathbb{R}^n$, то просто $\exists\, p \ne 0$ такой, что $(p, A) = const$. По сути, это тривиальный опорный, так как его ортогональная проекция на $Lin(A - x)$ при любом $x \in A$ — нулевая.

Если же $dim A = n$, то рассмотрим множество $A_0 = int A$. Как мы уже знаем, оно выпукло и непусто. По условию $x_0 \notin A_0$, следовательно, по теореме 3.2 $\exists\, p \ne 0$ такой, что

$$(p, A_0) \le (p, x_0),$$

а тогда и

$$(p, \overline{A}_0) \le (p, x_0),$$

а поскольку $\overline{A}_0 = \overline{A} \supset A$, то $(p, A) \le (p, x_0)$, ч. т. д. □

Упражнение 3.4. Доказать, что конус внешних нормалей не меняется при сдвиге: $\forall\, z \in \mathbb{R}^n \quad N_{A+z}\,(x+z) \ = \ N_A(x)$. Как он меняется при других аффинных преобразованиях?

Из теорем 3.1, 3.2 легко вытекают теоремы об отделимости двух выпуклых множеств.

Теорема 3.3 (об отделимости двух выпуклых множеств).
Любые два непустых непересекающихся выпуклых множества в $\mathbb{R}^n$ отделимы (вообще говоря, нестрого): если A, B — непустые выпуклые множества, и $A \cap B = \emptyset$, то $\exists\, p \neq 0$ такой, что $\forall\, x \in A, \forall\, y \in B \quad (p, x) \leq (p, y)$ (рис. 3.7).

Доказательство основано на следующем простом приеме. Рассмотрим алгебраическую разность $C = A - B$. Она, как мы знаем, выпукла, и из непересечения A и B очевидно следует, что $0 \notin C$. Тогда по теореме 3.2 $\exists\, p \neq 0$, отделяющий точку 0 от множества C, т.е. $(p, C) \leq 0$. А по определению множества C это значит, что $\forall\, x \in A, \forall\, y \in B \quad (p, x - y) \leq 0$, ч. т. д. □

Теорема 3.4 (о строгой отделимости двух выпуклых замкнутых множеств, одно из которых компакт). Два непересекающихся выпуклых замкнутых множества, одно из которых компакт, строго отделимы: если A, B — выпуклые замкнутые множества, $A \cap B = \emptyset$ и B — компакт, то $\exists\, p \in \mathbb{R}^n$ такой, что $\sup\,(p, A) < \inf\,(p, B)$, или, что эквивалентно, $\exists\, \delta > 0$ такое, что

$$\forall\, x \in A, \;\; \forall\, y \in B \qquad (p, x) \leq (p, y) - \delta. \tag{3.9}$$

Доказательство опирается на следующее свойство алгебраической суммы двух множеств, которое имеет и самостоятельный интерес.

Лемма 3.3. Пусть A — замкнутое множество, B — компакт. Тогда их алгебраическая сумма $C = A + B$ (а также алгебраическая разность $A - B$) замкнута.

Доказательство. Пусть $c_k \in C$ и $c_k \to c_0$. Требуется доказать, что $c_0 \in C$. По условию $c_k = a_k + b_k$, где $a_k \in A$, $b_k \in B$. Так как B —компакт, то найдется подпоследовательность номеров $k_m \to \infty$ такая, что элементы b_{k_m} сходятся к некоторому $b_0 \in B$. Тогда $a_{k_m} = c_{k_m} - b_{k_m} \to c_0 - b_0$, и в силу того, что A замкнуто, $a_0 = c_0 - b_0 \in A$. Таким образом, $c_0 = a_0 + b_0$, где $a_0 \in A$, $b_0 \in B$. Следовательно, $c_0 \in A + B = C$, ч. т. д. □

Доказательство теоремы 3.4. Опять рассматриваем выпуклое множество $C = A - B$. По лемме 3.3 оно замкнуто, и по условию не содержит точки 0. Тогда по теореме 3.1 найдутся вектор p и число $\delta > 0$ такие, что $(p, C) \leq (p, 0) - \delta$, то есть $\forall\, x \in A$, $y \in B$ имеем $(p, x - y) \leq \delta$, а это и есть требуемое свойство (3.9). □

Замечание. Требование компактности хотя бы одного из множеств в теореме 3.4 существенно. Пример: ось x и надграфик функции $1/x$ на плоскости не являются строго отделимыми. (Проверить!)

Лекция 4. Некоторые геометрические свойства выпуклых множеств

Крайние точки. Пусть A — выпуклое множество. Среди его граничных точек есть точки, обладающие следующим важным свойством.

Определение 4.1. Точка $x \in A$ называется *крайней точкой* множества A, если она не является серединой никакого отрезка ненулевой длины, лежащего в A, то есть если не существует точек $a_1 \neq a_2$ из A таких, что $x = \frac{a_1+a_2}{2}$.

Например, если A состоит из одной точки, то она является в нем крайней, ибо не может быть полусуммой двух различных точек из A.

Ясно, что внутренняя точка множества не может быть крайней: любая крайняя точка является граничной.

Множество всех крайних точек выпуклого множества A обозначается $\mathrm{ex}A$.

Укажем ряд простых свойств крайних точек, доказательства которых мы оставим читателю в качестве упражнения.

Свойство 1. Если $a_1 \neq a_2 \in A$, $0 < \gamma < 1$, то $x = \gamma\, a_1 + (1-\gamma)\, a_2$ не крайняя.

Свойство 2. Крайняя точка не может быть выпуклой комбинацией других точек множества: если $x \in \mathrm{ex}\, A$ и $x = \sum_{i=1}^{k} \gamma_i\, \alpha_i$, где все $x_i \in A$, $\gamma_i > 0$, и $\sum_{i=1}^{k} \gamma_i = 1$, то все $a_i = x$. (Доказать индукцией по k.)

Отсюда вытекает

Свойство 3. Пусть A - произвольное множество. Тогда $\mathrm{ex}\,(co\, A) \subset A$.

Действительно, любая точка $x \in co\, A$ есть выпуклая комбинация некоторых точек $a_i \in A$. Если при этом $x \in \mathrm{ex}(co\, A)$, то по свойству 1 все $a_i = x$, и поэтому $x \in A$. □

Свойство 4. Крайние точки любого симплекса — это его вершины.

Легко заметить, что не каждое выпуклое множество имеет крайние точки. Например, любое открытое выпуклое множество, аффинное подпространство, полупространство не имеют крайних точек.

Упражнение 1. Верно ли, что если $A' \subset A$, то $\mathrm{ex}A' \subset \mathrm{ex}A$?
Привести контрпример.

Задача 4.1. Пусть A — выпукло и замкнуто. Будет ли замкнутым $\mathrm{ex}A$?
Рассмотреть в $\mathbb{R}^3$ пример: A есть выпуклая оболочка окружности, лежащей в горизонтальной плоскости (x, y) и проходящей через 0, и вертикального отрезка $[-1, 1]$ на оси z (рис. 4.1). Показать, что здесь $\mathrm{ex}A$ состоит из окружности с выколотым нулём и концов вертикального отрезка.

Задача 4.2*. Доказать, что всегда $\mathrm{ex}A$ есть множество типа G_δ, т.е. является пересечением не более чем счётного числа открытых множеств.

Установим, что если выпуклое множество непусто и компактно, то оно всегда имеет крайние точки.

Пусть дан непустой выпуклый компакт A и произвольный вектор p.
Введём множество

$$A_p = \ Arg \max_{x \in A} (p, x) \, ,$$

которое называется также *гранью* множества A.

Ясно, что $A_p -$ непустой выпуклый компакт. (Почему?) Заметим, что если $\dim A = n$ и $p \neq 0$, то $\dim A_p < n$, ибо A_p содержится в гиперплоскости $(p, x) = c$, где $c = \max(p, A)$.

Лемма 4.1. $\operatorname{ex} A_p \subset \operatorname{ex} A$.

Доказательство. Пусть $\max(p, A) = c$. Тогда $A_p = \{x \in A \,|\, (p, x) = c\}$.

Пусть $x_0 \in \operatorname{ex} A_p$, но $x_0 \notin \operatorname{ex} A$, т.е. имеются различные $a_1, a_2 \in A$ такие, что $x_0 = (a_1 + a_2)/2$. Обе точки a_1 и a_2 не могут лежать в A_p, так как $x_0 \in \operatorname{ex} A_p$. Поэтому считаем, что $a_1 \notin A_p$. Тогда $(p, a_1) < c$, и поскольку всегда $(p, a_2) \leq c$, то получаем $(p, x_0) < c$, противоречие, ибо $x_0 \in A_p$. □

Лемма 4.2. Если A — непустой выпуклый компакт в конечномерном пространстве, то $\operatorname{ex} A$ непусто.

Доказательство. Индукция по $m = \dim A$. Для $m = 1$ утверждение очевидно (A – отрезок, $\operatorname{ex} A$ – его концы).

Пусть утверждение выполняется для всех компактов размерности $< m$; рассмотрим случай $\dim A = m$. Переходя к аффинной оболочке A, считаем $A \subset \mathbb{R}^m$. Возьмём любой ненулевой вектор $p \in \mathbb{R}^m$. Тогда A_p есть непустой выпуклый компакт и, как было показано выше, $\dim A_p < m$. По предположению индукции $\operatorname{ex} A_p$ непусто, а по лемме 4.1 $\operatorname{ex} A_p \subset \operatorname{ex} A$, следовательно, и $\operatorname{ex} A$ непусто. □

Из доказанных лемм 4.1 и 4.2 вытекает важное свойство крайних точек: среди точек максимума линейной функции (p, x) на выпуклом компакте обязательно есть крайние точки (ибо $\operatorname{ex} A_p$ непусто). Отсюда следует, что при поиске указанного максимума достаточно ограничиться рассмотрением лишь крайних точек компакта. Этот факт лежит в основе весьма эффективного *симплекс-метода* численного решения задач т.н. *линейного программирования* – поиске максимума линейной функции на компакте, заданном конечным числом линейных равенств и неравенств.

Еще одно важное свойство крайних точек состоит в том, что выпуклый компакт восстанавливается по своим крайним точкам. А именно, справедлива

Теорема 4.1 (Г. Минковский). Пусть A — выпуклый компакт в $\mathbb{R}^n$.
Тогда $A = \operatorname{co}(\operatorname{ex} A)$, т.е. выпуклый компакт в $\mathbb{R}^n$ есть выпуклая оболочка множества своих крайних точек.

Доказательство. Ясно, что компакт A можно считать состоящим более чем из одной точки. Включение $\supset$ очевидно. Обратное включение $\subset$ докажем индукцией по n. При $n = 1$ утверждение очевидно. (Отрезок есть выпуклая оболочка своих концов.) Допустим, что оно справедливо для всех пространств размерности меньше некоторого n, и рассмотрим ситуацию в $\mathbb{R}^n$. Считаем, что $\dim A = n$ (иначе можно уменьшить размерность пространства); при этом $\operatorname{int} A$ непуста.

Возьмём произвольную точку $x_0 \in A$. Если $x_0 \in \partial A$, то по следствию 1 из теоремы 3.2 $\exists\, p \neq 0$ такой, что $(p, A) \leq (p, x_0) = c$ (нетривиальный опорный). Тогда $x_0 \in A_p$ и $\dim A_p < n$. По предположению индукции $x_0 \in \operatorname{co}(\operatorname{ex} A_p)$, а по лемме 4.1 $ex\, A_p \subset ex\, A$, следовательно, $x_0 \in \operatorname{co}(\operatorname{ex} A)$, ч.т.д.

Пусть теперь $x_0 \in \operatorname{int} A$. Проведём через точку x_0 любую прямую. Так как A есть компакт, то эта прямая пересекает ∂A в каких-то двух точках x_1, x_2, и точка x_0 есть их выпуклая комбинация: $x_0 \in [x_1, x_2]$. Но по доказанному выше $x_1, x_2 \in \operatorname{co} E$, где $E = \operatorname{ex} A$, поэтому (учитывая, что $\operatorname{co} E$ — выпуклое множество) и $x_0 \in \operatorname{co} E$, ч.т.д. □

Отметим ещё одно характерное свойство крайних точек. Если из компакта *вырезать* крайнюю точку (т.е. не просто выколоть, а удалить вместе с некоторой окрестностью), и оставшуюся часть овыпуклить, то полученный в итоге компакт не будет содержать данной точки:

Свойство 5. Пусть A — выпуклый компакт, x — его крайняя точка, $\mathcal{U}$ — окрестность x. Тогда $A \setminus \mathcal{U}$ — компакт, и $x \notin \operatorname{co}(A \setminus \mathcal{U})$.

(Ясно, что если точка не крайняя, то она заведомо не обладает этим свойством.)

Доказательство. Поскольку $\mathcal{U}$ открыто, то $A \setminus \mathcal{U}$ есть замкнутое подмножество компакта, и значит, само есть компакт. А тогда по теореме Минковского его выпуклая оболочка $\operatorname{co}(A \setminus \mathcal{U})$ также есть компакт.

Если $x \in \operatorname{co}(A \setminus \mathcal{U})$, то x есть выпуклая комбинация конечного числа точек $a_i \in A \setminus \mathcal{U}$. Так как все $a_i \in A$ и $x -$ крайняя в A, то по свойству 2 все $a_i = x$, а так как все $a_i \notin \mathcal{U}$, то $x \notin \mathcal{U}$, а это противоречит условию $x \in \mathcal{U}$.

Задача 4.3. Вопрос: существенна ли здесь компактность? Т.е. пусть A замкнуто и выпукло. Можно ли утверждать: а) что $co\,(A \setminus \mathcal{U})$ замкнуто, и б) что $x \notin \overline{co}\,(A \setminus \mathcal{U})$?

Выступающие точки. Наряду с понятием крайней точки имеется также близкое к нему понятие выступающей точки.

Определение 4.2. Пусть A — выпуклое множество. Точка $x \in A$ называется *выступающей*, если существует линейный функционал p, такой что $\{x\} = \operatorname{Argmax}(p, A)$, т.е. x есть единственная точка максимума функционала p на множестве A.

Множество всех выступающих точек множества A будем обозначать $S(A)$.

Пример 4.1. Пусть $A = B_1(0)$ — единичный шар в евклидовой метрике с центром в нуле. Тогда $S(A)$ есть сфера $\partial B_1(0)$. Здесь достаточно доказать включение $\supset$ (почему?). Действительно, если $x \in \partial B_1(0)$, т.е. $|x| = 1$, то линейный функционал $p = x$ будет достигать максимума именно в этой точке , поэтому $x-$ выступающая.

Отсюда вытекает, что и вообще для любого шара $B_r(x)$ в евклидовой метрике множество его выступающих точек есть сфера $\partial B_r(x)$. (Вопрос: будет ли это верно для шаров в других метриках?)

Пример 4.2 (рис. 4.2). Множество A есть прямоугольник с приклеенными по бокам полукругами (раздвижной круглый стол). Множество крайних точек $E(A)$ состоит из обеих полуокружностей, включая концы. Однако $S(A)$ состоит из обеих полуокружностей без их концов a, b, c, d. (Почему ?)

Обратим внимание, что определение выступающей точки (в отличие от крайней) выглядит на первый взгляд вполне конструктивно — это точки, выделяемые линейными функционалами. Однако, если взять произвольный вектор $p \neq 0$, то, вообще говоря, нельзя быть уверенным, что его максимум на данном компакте A достигается в единственной точке. Поэтому возникает вопрос: нельзя ли всё-таки указать какой-нибудь конструктивный способ нахождения выступающих точек? Например, можно ли хотя бы утверждать, что любой выпуклый компакт имеет такие точки ? (Индукция по размерности A здесь не проходит, так как нет свойства $S(A_p) \subset S(A)$, которое было для крайних точек; это видно уже из примера 4.2).

Отметим пока одно простое, но полезное

Свойство 6. Пусть $A \subset B$ — выпуклые множества, и точка $x_0 \in A$ является выступающей для множества B. Тогда она является выступающей и для множества A.

Доказательство очевидно (см. рис. 4.3): если линейный функционал p достигает максимума по множеству B в единственной точке x_0, то и на более узком множестве A он также достигает максимума в той же единственной точке x_0, так как сама эта точка принадлежит A . □

Это свойство и рассмотренный выше пример 4.1 наталкивают на следующий способ нахождения выступающих точек в компактных множествах: надо взять любую сферу большого радиуса и коснуться её изнутри данным множеством A. Любая точка касания будет выступающей.

Лемма 4.3. Пусть A — выпуклый компакт, а z — произвольная точка в $\mathbb{R}^n$. Пусть a_0 есть точка в A, наиболее удалённая от z, точнее — любая точка, в которой достигается $\max|a - z|$ по $a \in A$. Тогда a_0 — выступающая.

Доказательство. Положим $r = \max|A - z|$. Тогда $A \subset B_r(z)$, и по условию $|a_0 - z| = r$, т.е. $a_0 \in \partial B_r(z)$. Тогда a_0 является выступающей для шара $B_r(z)$, и по свойству 6 она выступающая и для A. □

Таким образом, хотя определение выступающей точки даётся через максимизацию линейного функционала, т.е. с помощью гиперплоскости, конструктивным образом выступающую точку удобно находить с помощью сферы (или с помощью границы любого другого строго выпуклого множества).

Задача 4.4. Всякую ли выступающую точку выпуклого компакта A можно получить указанным образом (с помощью сферы)? Т.е. для всякой ли $a_0 \in S(A)$ найдётся $z \in \mathbb{R}^n$, для которой a_0 является наиболее удалённой? Рассмотреть пример, когда A есть надграфик функции $y = x^4$, а точка $a_0 = (0, 0)$.

Анализ примера 4.2 приводит к естественному вопросу: как в общем случае соотносятся между собой множества выступающих и крайних точек? Оказывается, между ними имеется следующая связь.

Теорема 4.2. $S(A) \subset E(A) \subset \overline{S(A)}$, т.е. всякая выступающая точка является также и крайней, а всякая крайняя точка является пределом выступающих (но сама может и не быть выступающей). Оба включения здесь, вообще говоря, не обратимы.

Доказательство первого включения просто. Пусть x – выступающая, и $x = (a_1 + a_2)/2$, где $a_1, a_2 \in A$. Для соответствующего функционала p имеем $(p, a_1) \leq (p, x)$ и $(p, a_2) \leq (p, x)$, и из предыдущего равенства следует, что оба этих неравенства обращаются в равенства: $(p, a_1) = (p, a_2) = (p, x)$. А отсюда в силу единственности точки максимума функционала p вытекает, что $a_1 = a_2 = x$. Таким образом, точка x не может быть полусуммой двух других точек из A, и тогда по определению она крайняя.

Второе включение вытекает из следующей теоремы, аналогичной теореме Минковского.

Теорема 4.3 (С.Страшевич). Пусть A — выпуклый компакт. Тогда

$$A \;=\; \overline{\mathrm{co}\, S(A)} \;=\; \mathrm{co}\, \overline{S(A)}, \tag{4.1}$$

т.е. выпуклый компакт есть выпуклая замкнутая оболочка множества своих выступающих точек.

Если эта теорема верна, то $A = \mathrm{co}\, \overline{S(A)}$, и тогда по свойству 3 крайних точек имеем $\mathrm{ex}A \subset \overline{S(A)}$, а это и есть второе включение теоремы 4.2.

Таким образом, осталось доказать теорему 4.3. Для этого нам потребуется следующий "нелинейный вариант" теоремы об отделимости (точнее, о строгой отделимости).

Лемма 4.4. Пусть дан выпуклый компакт C и точка $x_0 \notin C$. Тогда найдётся точка $z \in \mathbb{R}^n$ и число $r > 0$ такие, что шар $B_r(z)$ содержит C, но не содержит x_0. Другими словами, точку x_0 можно строго отделить от C не только гиперплоскостью, но и некоторой сферой, содержащей внутри себя множество C.

Доказательство. По теореме 2.1 о строгой отделимости $\exists\, p \neq 0$ такой, что $(p, x_0) < \min\,(p, C)$. Без нарушения общности (нормируя вектор p и сдвигая, если надо, множество C) можно считать, что $|p| = 1$, $(p, x_0) < 0$, $(p, C) \geq 1$.

Для любого $r > 0$ возьмём точку $z = rp$ и шар $B_r(rp)$ вокруг неё. Он задаётся неравенством $(x - rp)^2 \leq r^2$, т.е. $|x|^2 - 2r(p, x) \leq 0$, или

$$|x|^2 \;\leq\; 2r(p, x). \tag{4.2}$$

Ясно, что он не содержит x_0 (в силу того, что $(p, x_0) < 0$). Покажем, что при достаточно больших r этот шар содержит множество C.

Действительно, так как C ограничено, то $\forall\, x \in C$ левая часть (4.1) не превосходит $\max |C|^2$, а так как $(p, C) \geq 1$, то правая часть всегда $\geq 2r$. При достаточно больших r получим $\max|C|^2 \le 2r$, и следовательно, неравенство (4.2) будет выполнено. Итак, $B_r(rp) \supset C$, и тем самым лемма доказана. $\square$

Доказательство теоремы 4.3. В силу ограниченности $S(A)$ последнее равенство в (4.1) выполнено по лемме 3.2, поэтому достаточно доказать включение $A \subset \overline{\mathrm{co}\, S(A)}$, так как обратное включение очевидно.

От противного: допустим, что $\exists\, x_0 \in A$ такая, что $x_0 \notin C = \overline{\mathrm{co}\ S(A)}$. Так как C есть выпуклый компакт, то по лемме 4.4 найдётся точка $z \in \mathbb{R}^n$ и число $r > 0$ такие, что шар $B_r(z)$ содержит C, но не содержит x_0, т.е. $|x_0 - z| > r$. Поскольку $x_0 \in A$, то отсюда следует, что $\max|A - z| > r$, и поэтому любая точка a_0 этого максимума (т.е. любая наиболее удалённая от z точка множества A) лежит вне шара $B_r(z)$ и тем более вне C, т.е. $a_0 \notin C$. Но по лемме 4.3 любая наиболее удалённая точка является выступающей в A, т.е. $a_0 \in S(A)$, и тогда по определению $a_0 \in C$, противоречие. Теорема доказана. $\square$

На этом мы закончим наше краткое знакомство с геометрическими свойствами выпуклых компактов, и рассмотрим одно свойство *неограниченных* выпуклых множеств.

Рецессивные направления. Пусть A — произвольное множество.

Определение 4.3. Вектор h называется *рецессивным* (вектором или направлением) для множества A, если для любой точки $x \in A$ луч, исходящий из этой точки в направлении h, целиком содержится в A: $x + \mathbb{R}_+ h \subset A$.

Множество всех рецессивных направлений множества A обозначим $\mathrm{Rec}\ A$.

Упражнение 2. Показать, что $\mathrm{Rec}\ A$ для любого замкнутого выпуклого множества A есть замкнутый выпуклый конус. Существенны ли здесь замкнутость и выпуклость множества?

Ясно, что нулевой вектор $h = 0$ всегда является рецессивным, но это малоинтересный, тривиальный случай. Для ограниченных множеств других рецессивных направлений очевидно и не может быть. Естественно возникает вопрос: всегда ли существуют нетривиальные рецессивные направления для неограниченных множеств?

Примеры. Рассмотрим на плоскости (x, y) следующие множества: октант $\mathbb{R}^2_+$, полосу $|x| \le 1$, надграфик параболы $y \ge x^2$, множество $xy \ge 1$ в положительном октанте. Их множества рецессивных направлений будут соответственно: октант $\mathbb{R}^2_+$, ось y, полуось $y \ge 0$, октант $\mathbb{R}^2_+$.

Показать, что для выпуклого конуса K всегда $\mathrm{Rec}\ K = K$. Рассмотреть пример, когда K есть открытый октант $\mathbb{R}^2_{++}$ с добавлением начала координат.

Пусть A есть выпуклая оболочка оси x_1 и точки (0,1). (Как она выглядит?) Нетрудно убедиться, что $\mathrm{Rec}\ A = \{0\}$. Таким образом, не всякое неограниченное выпуклое множество имеет нетривиальные рецессивные направления.

Теорема 4.4. Пусть A — неограниченное выпуклое *замкнутое* множество. Тогда оно имеет хотя бы одно рецессивное направление $h \ne 0$.

Доказательство. Зафиксируем любую точку $x_0 \in A$. Так как A неограничено, то существует последовательность $x_k \in A$ такая, что $|x_k| \to \infty$. Рассмотрим единичные векторы $h_k = \frac{x_k - x_0}{|x_k - x_0|}$. В силу компактности единичной сферы можно считать (переходя, если необходимо, к подпоследовательности), что h_k сходятся к некоторому единичному же вектору h. При этом $x_k = x_0 + \alpha_k h_k$, где $\alpha_k = |x_k - x_0| \to \infty$. Так как точки $x_0, x_k \in A$, то и отрезок $[x_0, x_k] \subset A$, а так как для любого фиксированного $\beta > 0$ точка $y_k = x_0 + \beta h_k$ при достаточно больших k лежит на этом отрезке, то она также принадлежит A. Но тогда, поскольку A замкнуто, и предел этих точек $y = x_0 + \beta h \in A$. Это верно $\forall\, \beta > 0$. Итак, мы получили, что единичный вектор h обладает тем свойством, что луч $x_0 + \mathbb{R}_+ h \subset A$.

Покажем теперь, что и для любой другой точки $x \in A$ также будет $x + \mathbb{R}_+ h \subset A$. Это утверждение удобно сформулировать в виде отдельной леммы, которая и завершает доказательство теоремы 4.4.

Лемма 4.5. Пусть A — выпуклое замкнутое множество, $h \neq 0$ и $x_0 + \mathbb{R}_+ h \subset A$. Тогда $\forall\, x \in A$ по-прежнему $x + \mathbb{R}_+ h \subset A$.

Доказательство. Если точка x лежит на прямой $x_0 + \mathbb{R}h$, то доказательство тривиально, поэтому будем считать, что $x \notin x_0 + \mathbb{R}h$, и рассмотрим двумерную картину в плоскости $x,\, x_0 + \mathbb{R}h$. Так как точка x и луч $x_0 + \mathbb{R}_+ h$ содержатся в A, то и их выпуклая оболочка также содержится в A. Эта выпуклая оболочка есть, как легко видеть, полоса $[x_0, x] + \mathbb{R}_+ h$ плюс ещё сама точка x. (Показать строго!) В силу замкнутости A тогда и замкнутая полоса $[x_0, x] + \mathbb{R}_+ h$ содержится A, в частности $x + \mathbb{R}_+ h \subset A$, ч.т.д. □

Следующая лемма является в некотором смысле обобщением леммы 4.5.

Лемма 4.6. Пусть A — выпуклый компакт, содержащий 0, и K — выпуклый замкнутый конус. Тогда

$$\overline{co\ (A \cup K)} \;=\; A + K. \tag{4.3}$$

Доказательство. Множество $co\,(A \cup K)$ состоит из всех точек вида $y = \lambda a + (1-\lambda)x$, где $a \in A$, $x \in K$, $0 \le \lambda \le 1$. Так как A и K содержат 0, то $\lambda a \in A$ и $(1-\lambda)x \in K$, поэтому $y \in A + K$ и следовательно, $co\,(A \cup K) \subset A + K$. Но множество $A + K$ замкнуто (как сумма компакта и замкнутого множества), поэтому включение $\subset$ в (4.3) доказано.

Докажем включение $\supset$. Пусть $y = a + x$, где $a \in A$, $x \in K$. Для каждого $\varepsilon \in (0,1)$ положим $y_\varepsilon = (1-\varepsilon)a + x = (1-\varepsilon)a + \varepsilon(x/\varepsilon)$. Так как $x/\varepsilon \in K$, то $y_\varepsilon \in co\ (A \cup K)$, а тогда $\lim y_\varepsilon = a + x \in \overline{co\ (A \cup K)}$. □

Приведенные выше примеры показывают, что условие замкнутости в теореме 4.4 существенно. Тем не менее, справедлива следующая

Теорема 4.5. Пусть A — неограниченное выпуклое *открытое* множество. Тогда оно также имеет рецессивное направление $h \neq 0$. Более того, $Rec\,A \;=\; Rec\,(\overline{A})$.

Доказательство. Рассмотрим неограниченное выпуклое замкнутое множество $B = \overline{A}$, и пусть h есть любое его ненулевое рецессивное направление, существующее по теореме 4.4. Покажем, что h рецессивно и для A. По определению $\forall x \in B$ (в частности, $\forall x \in A$) имеем $x + \mathbb{R}_+ h \subset B$, т.е. $x + rh \in B \quad \forall r \geq 0$.

Итак, для любой точки $x \in A = \text{int } B$ точка $x + rh \in B$. По свойству 2 выпуклых множеств (см. лекцию 1) тогда $[x, x + rh) \subset \text{int } B = A$, и так как это справедливо $\forall r \geq 0$, то и весь луч $x + \mathbb{R}_+ h \subset A$. В силу произвольности точки $x \in A$, вектор h по определению является рецессивным для множества A. Теорема доказана. □

Касательные направления. Пусть дано произвольное множество A и точка $x_0 \in \overline{A}$.

Определение 4.4. Вектор (или направление) h называется касательным к множеству A в точке x_0, если расстояние $\rho(x_0 + \varepsilon h, A) = o(\varepsilon)$ при $\varepsilon \to 0+$.

Обратим внимание, что при этом вектор $-h$ может не быть касательным, т.е. h является односторонним касательным вектором.

Из определения сразу видно, что h одновременно является касательным к A и к $\overline{A}$, поэтому при нахождении касательных векторов множество A можно считать замкнутым.

Теорема 4.6. Множество всех касательных векторов к любому множеству A в любой точке $x_0 \in \overline{A}$ есть замкнутый конус; он называется конусом касательных направлений (или просто касательным конусом) и обозначается $K_A(x_0)$.

(*Указание:* показать, что дополнение к касательному конусу открыто.)

Задача 4.5. Пусть A выпукло. Тогда конус $K_A(x_0)$ также выпуклый, причем

$$K_A(x_0) \;=\; \overline{con\,(A - x_0)}.$$

Задача 4.6. Найти конус касательных направлений к графику функции $y = x \sin \frac{1}{x}$ в точке $(0, 0)$.

Задача 4.7*. Доказать, что для выпуклого множества $\rho(x_0 + \varepsilon h, A)/\varepsilon \to 0$ при $\varepsilon \to 0+$ равномерно по всем направлениям $h \in K_A(x_0)$, $|h| \leq 1$. Это означает, что в малой окрестности точки x_0 выпуклое множество, масштабированное на размер этой окрестности, очень мало отличается от своего касательного конуса (в метрике Хаусдорфа), и это отличие стремится к нулю при стремлении диаметра окрестности к нулю. Таким образом, выпуклое множество в любой своей точке хорошо аппроксимируется своим касательным конусом.

Вопрос: Верно ли это свойство для невыпуклых множеств?

Лекция 5. Двойственные множества: поляра и сопряженный конус

Выпуклые объекты — множества, функции и т.д. — обладают тем важным замечательным свойством, что для каждого из них, определенного в своем пространстве, имеется некоторый двойственный объект, определенный в сопряженном пространстве (т.е. в пространстве линейных функционалов на исходном пространстве); при этом исходный объект, если он обладает некоторыми свойствами замкнутости, однозначно восстанавливается по своему двойственному объекту. Отсюда вытекает, что выпуклый объект можно задавать двумя способами — как описанием его самого, так и путем задания двойственного объекта. Этот факт целиком основан на теореме об отделимости, и является фундаментальным в выпуклом анализе.

В этой лекции мы рассмотрим двойственный объект к выпуклому множеству. Так как исходное пространство у нас есть $\mathbb{R}^n$ и в нем задано скалярное произведение, то сопряженное к нему пространство мы отождествляем с тем же $\mathbb{R}^n$.

Двойственные объекты: поляра множества

Пусть A — непустое множество в $\mathbb{R}^n$.

Определение 5.1. Полярой множества A называется множество

$$A^0 \;=\; \{p \in \mathbb{R}^n \mid \; (p, A) \le 1\}.$$

Из этого определения сразу вытекает

Свойство 1. A^0 есть выпуклое замкнутое множество, содержащее 0, ибо

$$A^0 \;=\; \bigcap_{x \in A} \{p \in \mathbb{R}^n \mid \; (p, x) \le 1\}$$

— есть пересечение замкнутых полупространств, содержащих 0.

Установим еще ряд свойств поляры:

Свойство 2. Если $A \subset B$, то $A^0 \supset B^0$. (Очевидно.)

Свойство 3. $(\overline{A})^0 = A^0$, $(co\; A)^0 = A^0$, $(A \cup \{0\})^0 = A^0$, и следовательно,

$$(\;\overline{co\;(A \cup \{0\})}\;)^0 \;=\; A^0\;. \tag{5.1}$$

Таким образом, к множеству можно добавлять точку 0, овыпуклять и замыкать его, не изменяя при этом его поляры.

Докажем первые три равенства. Во всех них, с учетом свойства 2, надо доказать лишь включение $\supset$. Пусть $p \in A^0$. Возьмем произвольный $x_0 \in \overline{A}$. По определению имеется последовательность $x_k \in A,\; x_k \to x_0$. Так как при этом $(p, x_k) \le 1$, то и $(p, x_0) \le 1$, откуда $p \in (\overline{A})^0$, и тем самым первое равенство доказано.

Возьмем теперь произвольный $y \in co\,A$, т.е. $y = \sum \alpha_i x_i$, где все $x_i \in A$, $\alpha_i \geq 0$, $\sum \alpha_i = 1$. Так как $(p_i, x_i) \leq 1$ для всех i, то $(p, y) = \sum \alpha_i (p_i x_i) \leq \sum \alpha_i = 1$, откуда следует, что $p \in (co\,A)^0$, т.е. второе равенство доказано.

Выполнение неравенства $(p,\ A \cup \{0\}) \leq 1$ очевидно, поэтому третье равенство также доказано. □

Свойство 4. Множество A ограничено $\iff$ $0 \in int\,A^0$.

Доказательство. ($\Longrightarrow$) Пусть $|A| \leq r$. Покажем, что тогда $B_{1/r} \subset A^0$. Действительно, если $|p| \leq 1/r$, то $\forall\, x \in A$ имеем $|(p,x)| \leq |p| \cdot |x| \leq \frac{1}{r}\, r = 1$.

($\Longleftarrow$) Пусть $B_\varepsilon \subset A^0$ при некотором $\varepsilon > 0$. Тогда $\forall\, x \in A$ имеем $(A^0, x) \leq 1$, следовательно $(B_\varepsilon, x) \leq 1$, т.е. $\varepsilon(B_1, x) \leq 1$, т.е. $\varepsilon \cdot \sup_{|p| \leq 1}(p, x) = \varepsilon\,|x| \leq 1$, следовательно $|x| \leq 1/\varepsilon \quad \forall\, x \in A$, т.е. $|A| \leq 1/\varepsilon$. □

Свойство 5. Пусть A выпукло. Тогда $0 \in int\ A \iff A^0$ ограничено (и, следовательно, компактно). Доказать самим по аналогии с предыдущим.

Задача. Пусть последовательность выпуклых компактов A_k сходится по Хаусдорфу к компакту A, для которого $0 \in int\,A$. Доказать, что тогда и последовательность поляр A_k^0 сходится по Хаусдорфу к A^0 .

Примеры поляр

0) Если $A = \{0\}$, то $A^0 = \mathbb{R}^n$, и наоборот, если $A = \mathbb{R}^n$, то $A^0 = \{0\}$.

1) Если $A = \{a \neq 0\}$ — точка, то $A^0 = \{p\ \ |(p, a) \leq 1\}$ — полупространство. Если A есть полупространство $\{x\ \ |(a, x) \leq 1\}$, где $a \in \mathbb{R}^n$, то $A^0 = [0, a]$ — отрезок.

2) Пусть $A = B_1(0) = \{x\ \ |\sum_1^n x_i^2 \leq 1\}$ — единичный шар в норме, соответствующей скалярному произведению. Тогда $A^0 = A = B_1(0)$.

Действительно, если $p \in B_1(0)$, т.е. $|p| \leq 1$, то $\forall\, x \in A$ имеем $|(p,x)| \leq |p| \cdot |x| \leq 1$, т.е. $p \in A^0$. Если же $|p| > 1$, то взяв $x = p/|p| \in A$, получим $(p, x) = (p, p/|p|) = |p| > 1$, следовательно, $p \notin A^0$.

Задача. Доказать, что если $A^0 = A$, то $A = B_1(0)$, т.е. других множеств, совпадающих со своей полярой, нет.

3) Если $A = \{x \mid \sum x_i \leq 1,$ все $x_i \geq 0\} = co\,\{0, e_1, \ldots, e_n\}$ — стандартный симплекс, то $A^0 = \{p \mid$ все $p_i \leq 1\}$ — сдвинутый отрицательный ортант.

4) Пусть $A = \{x \mid \sum |x_i| \leq 1\}$ — единичный шар в норме l_1 . Тогда $A^0 = \{p \mid \max |p_i| \leq 1\}$ — единичный шар в норме l_∞ . Обратно: если A — единичный шар в норме l_∞, то A^0 — единичный шар в норме l_1 .

5) Пусть $A = \{x \mid \sum |x_i|^\alpha \leq 1\}$ — единичный шар в норме l_α, где $\alpha > 1$. Тогда $A^0 = \{p \mid \sum |p_i|^\beta \leq 1\}$ — единичный шар в норме l_β, где $1/\alpha + 1/\beta = 1$. (Отсюда при $\alpha = 2$ получаем свойство 3, а при $\alpha \to 1 - 0$ или $\alpha \to \infty$ свойство 4).

6) Пусть $A = \{x \mid -a \le (l,x) \le b\}$ — полоса (здесь $a,b > 0$, $l \in \mathbb{R}^n$). Тогда $A^0 = \{p = \lambda l \mid -1/a \le \lambda \le 1/b\}$ — отрезок. Обратно: если A — отрезок такого вида, то A^0 — соответствующая полоса.

7) На плоскости $\mathbb{R}^2$ рассмотрим множество $A = \{\, x, y > 0,\ \ y \ge 1/x\}$ (надграфик гиперболы в первом октанте). Какова его поляра?

Этот вопрос ставит в тупик многих студентов. На самом же деле решается он очень просто. По свойству 3 поляра множества A совпадает с полярой множества $\overline{co\,(A \cup \{0\})}$. Легко видеть, что последнее множество есть неотрицательный октант $\mathbb{R}^2_+$, поляра которого есть неположительный октант $\mathbb{R}^2_-$. Итак, $A^0 = \mathbb{R}^2_-$.

Переход ко второй поляре. Применим теперь операцию взятия поляры к множеству A^0, т.е. рассмотрим вторую поляру множества A (иногда ее называют биполярой); для краткости вместо $(A^0)^0$ будем писать A^{00}. Можно было заметить, что почти во всех вышеприведенных примерах A^{00} совпадает с исходным множеством A. Оказывается, это совпадение не случайно.

Теорема 5.1 (о второй поляре). Пусть A — выпуклое замкнутое множество, содержащее точку 0. Тогда $A^{00} = A$. Верно и обратное.

Доказательство. Сразу заметим, что обратное утверждение тривиально в силу свойства 1 поляры любого множества. Поэтому надо доказать лишь прямое утверждение.

Итак, пусть A — выпуклое замкнутое множество и $0 \in A$. Обозначим $A^0 = C$. Так как $(C, A) \le 1$, то по определению $A \subset C^0 = A^{00}$. Покажем, что $A \supset A^{00}$, или, что эквивалентно: если $x \notin A$, то $x \notin C^0$.

Действительно, пусть $x \notin A$. Тогда x строго отделим от A некоторым функционалом p, т.е. $(p,x) > \sup(p,A)$. Поскольку $A \ni 0$, то $\sup(p,A) \ge 0$, и следовательно, $(p,x) > 0$. Умножая p на подходящее положительное число, можно добиться того, что будет $(p,x) > 1 \ge \sup(p,A)$. Тогда $p \in A^0 = C$, а неравенство $(p,x) > 1$ говорит о том, что $x \notin C^0$, ч.т.д. $\square$

Из этой теоремы и свойства 3 вытекает

Следствие 1. Пусть A – замкнутое выпуклое множество, содержащее 0. Тогда оно является полярой некоторого замкнутого выпуклого множества B, содержащего 0: $A = B^0$.

Действительно, положим $B = A^0$. Это замкнутое выпуклое множество, содержащее 0, и по теореме 5.1 $B^0 = A$.

Таким образом, множество A и его поляра A^0 двойственны по отношению к друг другу: каждое из них есть поляра другого. Другими словами, операция взятия поляры инволютивна на классе замкнутых выпуклых множеств, содержащих 0, а сам этот класс инвариантен относительно данной операции (она переводит его в себя).

Следствие 2. Пусть A – любое множество. Тогда

$$A^{00} = \overline{co\,(A \cup \{0\})}. \tag{5.2}$$

Доказательство. Обозначим множество, стоящее в правой части, через B. Это замкнутое выпуклое множество, содержащее 0. По теореме 5.1 $B^{00} = B$. Но по свойству 3 $B^0 = A^0$, т.е. получаем $A^{00} = B$.

Установим еще несколько свойств поляры. Пусть A, B — произвольные множества.

Свойство 6. $(A \cup B)^0 = A^0 \cap B^0$ (доказать!), и следовательно,

$$\overline{(co\,(A \cup B))}^0 \;=\; A^0 \cap B^0. \tag{5.3}$$

Это свойство, в частности, позволяет легко находить поляру к многограннику $M = co\,\{a_1, \ldots, a_m\}$. Согласно (5.3), $M^0 = \bigcap_i \{a_i\}^0 \;=\; \{\, p \mid (p, a_i) \le 1 \;\; \forall\, i\,\}$.

Вопрос о вычислении двойственного объекта к *пересечению* двух (а тогда и нескольких) объектов является нетривиальным, в отличие от вычисления двойственного объекта к объединению (который, как правило, намного проще). Это основной вопрос в технике выпуклого анализа. Ожидаемая формула здесь легко находится из соображений двойственности (точнее, инволютивности операции перехода к двойственному объекту), но ее справедливость устанавливается лишь при некоторых дополнительных предположениях "регулярности".

Свойство 7. $(A \cap B)^0 \;\supset\; \overline{co(A^0 \cup B^0)}$ для любых множеств A, B.

Достаточно доказать включение $(A \cap B)^0 \supset A^0 \cup B^0$ (а далее воспользоваться свойством 3 поляры). Но это включение очевидно вытекает из свойства 2: $A \cap B \subset A$, поэтому $(A \cap B)^0 \supset A^0$, и точно так же $(A \cap B)^0 \supset B^0$. □

Упражнение 1. Привести пример двух выпуклых множеств, для которых в свойстве 7 выполнено строгое включение.

Свойство $7'$. Пусть множества A, B выпуклы, замкнуты и содержат 0. Тогда

$$(A \cap B)^0 = \overline{co\,(A^0 \cup B^0)}\;. \tag{5.4}$$

Действительно, пусть $C = A^0$, $D = B^0$. Тогда по теореме 5.1 $A = C^0$, $B = D^0$, и по свойству 6 и следствию 1

$$(A \cap B)^0 = (C^0 \cap D^0)^0 = ((C \cup D)^0)^0 = \overline{co\,(C \cup D)} = \overline{co\,(A^0 \cup B^0)}\;.$$

Упражнение 2. Пусть $A_1, \ldots, A_m$ – выпуклые замкнутые множества, содержащие 0. Доказать по индукции, что тогда

$$(\bigcap \; A_i)^0 \;=\; \overline{co\;(\bigcup A_i^0)}\;. \tag{5.5}$$

Полученная формула для поляры пересечения, если сравнить ее с формулой для поляры объединения (свойство 6), все еще не является в точности двойственной к последней: присутствует "лишнее" замыкание. Оно исчезает в следующем "регулярном" случае.

Свойство 8. Пусть множества A, B выпуклы, замкнуты, и у обоих 0 есть внутренняя точка (условие регулярности). Тогда

$$(A \cap B)^0 \;=\; co\,(A^0 \cup B^0). \tag{5.6}$$

По свойству $7'$ имеем $(A \cap B)^0 = \overline{co(A^0 \cup B^0)}$. Но по свойству 5 оба множества A^0 и B^0 – компакты, и тогда выпуклая оболочка компакта $A^0 \cup B^0$ тоже есть компакт, поэтому справа стоит замкнутое множество, и еще раз замыкать его уже не требуется. □

Нетрудно показать, что требование замкнутости A, B здесь несущественно. Т.е. справедливо более общее

Свойство $8'$. Пусть множества A, B выпуклы (не обязательно замкнуты), и у обоих 0 есть внутренняя точка. Тогда справедлива формула (5.6).

Действительно, в этом случае по свойству 3 $(A \cap B)^0 = (\overline{A \cap B})^0$, а по свойству 2.5 $\overline{A \cap B} = \overline{A} \cap \overline{B}$, т.е. вместо множеств A, B можно рассматривать их замыкания, к которым применить свойство 8. □

Установим, как меняется поляра при линейном отображении.

Свойство 9 (поляра линейного образа). Пусть задано линейное отображение $G : \mathbb{R}^m \to \mathbb{R}^n$, задано произвольное множество $B \subset \mathbb{R}^m$, и пусть $A = GB$. Тогда

$$A^0 = (G^*)^{-1}(B^0) \;=\; \{p \mid G^*p \in B^0\,\}. \tag{5.7}$$

Здесь $G^* : \mathbb{R}^n \to \mathbb{R}^m$ – сопряженное отображение, а верхний индекс -1 означает взятие полного прообраза (а не обратной матрицы, которой может не существовать!). Напомним, что, строго говоря, G^* есть отображение сопряженных пространств $(\mathbb{R}^n)^* \to (\mathbb{R}^m)^*$. Если базисы в этих пространствах сопряжены к базисам в исходных пространствах $\mathbb{R}^n$ и $\mathbb{R}^m$, то G^* задается транспонированной матрицей к матрице оператора G .

Доказательство. Для любого $p \in \mathbb{R}^n$ верно равенство $(p, A) = (p, GB) = (G^*p, B)$. Отсюда $(p, A) \le 1 \iff (G^*p, B) \le 1$, т.е. $p \in A^0 \iff G^*p \in B^0$, а это и означает, что $A^0 = (G^*)^{-1}(B^0)$. □

Из формулы (5.7) очевидно следует включение $G^*A^0 \subset B^0$, но обратного включения, вообще говоря, нет. Рассмотрим такие примеры на плоскости. Пусть $G : \mathbb{R}^2 \to \mathbb{R}^2$ есть проекция на ось x_1, а множество B есть единичный круг или отрезок, соединяющий точки $(-1, 1)$ и $(1, -1)$. В обоих случаях A есть отрезок $[-1, 1]$ на оси x_1, его поляра A^0 есть вертикальная полоса $|x_1| \le 1$, $G^* = G$, $G^*A^0 = [-1, 1]$ на оси x_1, тогда как B^0 есть единичный круг или косая полоса, ограниченная прямыми, проходящими через точки $(-1, 0)$ и $(1, 0)$ под углом 45^0 (см. рис. 5.XX).

Если $m = n$ и G обратимо, то G^* тоже обратимо, и тогда из (5.7) очевидно вытекает равенство $G^*A^0 = B^0$.

Свойство 10 (поляра линейного прообраза). Пусть, как и прежде, задано линейное отображение $G : \mathbb{R}^m \to \mathbb{R}^n$, но теперь задано выпуклое множество $A \subset \mathbb{R}^n$, а $B = G^{-1}(A)$ есть его полный прообраз. Предположим, что $Im\, G \cap\ int\ A \neq \emptyset$ (условие регулярности). Тогда $B^0 = G^* A^0$.

Доказать самим в качестве упражнения. Рассмотреть сначала случай, когда G действует "на". Обозначим $X = \mathbb{R}^m$, $Y = \mathbb{R}^n$. Тогда $X = L + Z$, где $L = Ker\, G$, и $G : Z \to Y$ действует 1–1 и "на". Поэтому $B = L + B_Z$, где $B_Z = (G^{-1}A) \bigcap Z$, и любой $q \in B^0$ очевидно зануляется на L, а значит имеет вид $q = G^* p$, $p \in Y$.

Обратно: любой $q \in X$ такого вида зануляется на L. Поэтому $(q, B) = (G^* p,\, L + B_Z) = (p,\, GB\, z) = (p, A)$, следовательно, $(q, B) \leq 1$ эквивалентно $(p, A) \leq 1$. Отсюда вытекает, что $q \in B^0 \iff q = G^* p$, где $p \in A^0$. □

Завершим рассмотрение свойств поляры следующим полезным примером.

Поляра эллипсоида. Пусть в $\mathbb{R}^n$ задан эллипсоид с полуосями $\sigma_1, ..., \sigma_n$:

$$A = \left\{ x \;\middle|\; \sum \left(\frac{x_i}{\sigma_i}\right)^2 \leq 1 \right\}.$$

Он, очевидно, есть образ единичного шара B_1 при линейном отображении

$$G = diag\,(\sigma_1, \ldots, \sigma_n) :\ \mathbb{R}^n \to \mathbb{R}^n,$$

$$y = (y_1, ..., y_n) \longmapsto x = (\sigma_1 y_1, ..., \sigma_n y_n).$$

Действительно, $y \in B_1 \iff x \in A$, т.е. $GB_1 = A$.

По свойству 9 имеем $A^0 = G^{-1} B_1$ (мы учитываем, что $G^* = G$ и $B_1^0 = B_1$), т.е. $p \in A^0 \iff Gp \in B_1$, или, что то же самое,

$$A^0 = \{p \mid \sum (\sigma_i p_i)^2 \leq 1\}.$$

Это также эллипсоид, который называется сопряженным к исходному эллипсоиду A. Его полуоси имеют длину $1/\sigma_i$ и направлены вдоль осей исходного эллипсоида.

Двойственные объекты: сопряженный конус

Рассмотрим теперь свойства поляры в случае, когда множество есть конус. Для конусов вместо поляры традиционно используется несколько другой двойственный объект — сопряженный конус.

Итак, пусть K — произвольный конус в $\mathbb{R}^n$.

Определение 5.2. Сопряженным конусом называется множество

$$K^* = \{\, p \in \mathbb{R}^n \mid\ (p, K) \geq 0 \,\}.$$

Отметим следующий простой факт: если $(p, K) \geq c$ при некотором $c \in \mathbb{R}$, то обязательно $c \leq 0$ и $(p, K) \geq 0$. (Доказать!)

Свойства сопряженного конуса

0) K^* — замкнутый выпуклый конус. (Очевидно.)

1) $K^0 = -K^*$.

Действительно, $p \in K^0 \iff (p, K) \leq 1 \iff -(p, K) \geq -1$, а это означает, как было сказано выше, что $-(p, K) \geq 0$, т.е. $-p \in K^*$.

Таким образом, сопряженный конус — это та же поляра, только со знаком минус.

2) $K_1 \subset K_2 \Longrightarrow K_1^* \supset K_2^*$.

3) $(\overline{K})^* = K^*$, $(co\ \ K)^* = K^*$, и следовательно,

$$(\overline{co\ \ K}\,)^* \;=\; K^*, \tag{5.8}$$

т.е. сопряженный конус не меняется, если исходный конус овыпуклить и замкнуть. (Следует из свойства поляры 3, но легко доказать и непосредственно.)

4) Пусть L — линейное подпространство. Тогда $L^* = L^\perp$ (ортогональное дополнение к L). Другими словами, если $(p, L) \geq 0$, то $(p, L) = 0$.

Следует из того, что L вместе с каждым элементом x содержит также и $-x$. $\square$

Напомним, что в линейной алгебре доказывалось, что если

$$L = \{\ x \mid\ (b_j, x) = 0, \qquad j = 1, \ldots, m\},$$

где $b_j \in \mathbb{R}^n$, то

$$L^\perp = \{\ p = \sum \beta_j b_j\ \mid \beta_j \in \mathbb{R}\} \tag{5.9}$$

есть линейная оболочка векторов $b_1, \ \ldots\ , b_m$.

5) Пусть $K = \{\ x \mid (p, x) \geq 0\}$. (Если $p \neq 0$, то это полупространство.) Тогда

$$K^* = \ \mathbb{R}_+ p\ = \{\ q = \alpha p \mid \alpha \geq 0\}$$

— замкнутый луч, порожденный вектором p .

Другими словами, если даны два вектора $p, q \in \mathbb{R}^n$, и из $(p, x) \geq 0$ следует $(q, x) \geq 0$ (в этом случае говорят, что вектор q подчинен вектору p), то q коллинеарен p : $\quad q = \alpha p$ при некотором $\alpha \geq 0$.

Доказательство. Включение $\supset$ очевидно. Докажем обратное включение $\subset$. Достаточно рассмотреть случай $p \neq 0$. Возьмем произвольный элемент $q \in K^*$ и представим его в виде $q = \alpha p + l$, где $l \perp p$. Если $l \neq 0$, то вектор $-l \in K$, и для него $(q, l) = -|l|^2 < 0$, что противоречит условию $q \in K^*$. Поэтому $l = 0$ и тогда $q = \alpha p$. Если $\alpha < 0$, то для вектора $p \in K$ опять получаем противоречие: $(q, p) = \alpha |p|^2 < 0$. Следовательно, $\alpha \geq 0$, ч.т.д. $\square$

Можно было рассуждать и так. По условию имеем импликацию $(p, x) \geq 0 \Longrightarrow (q, x) \geq 0$. В частности, из $(p, x) = 0$ следует $(q, x) \geq 0$ и тогда $(q, x) = 0$ (почему?). Отсюда, как было сказано выше, $q = \alpha p$, и тогда ясно, что $\alpha \geq 0$, иначе указанная импликация нарушается. $\square$

Теорема о второй поляре приобретает здесь следующий вид.

Теорема 5.2 (о втором сопряженном конусе).
Пусть K —замкнутый выпуклый конус. Тогда $K^{**} = K$.

(Вытекает очевидно из теоремы 5.1, так как в силу замкнутости $K \ni 0$, и все условия теоремы 5.1 выполнены. Однако, полезно также доказать и непосредственно. Опять, как и в теореме 5.1, надо будет применить теорему об отделимости.)

Следствие 1. Пусть K — выпуклый замкнутый конус. Тогда он является сопряженным к некоторому выпуклому замкнутому конусу C : $K = C^*$.

(Действительно, надо взять $C = K^*$, тогда по теореме 5.2 $C^* = K^{**} = K$.)

Таким образом, операция перехода к сопряженному конусу инволютивна на классе выпуклых замкнутых конусов.

Следствие 2. Для произвольного конуса K справедливо равенство:

$$K^{**} \;=\; \overline{co\, K}\;. \tag{5.10}$$

Также вытекает из следствия 1 к теореме 5.1, ибо $\overline{co\, K \cup \{0\}} = \overline{co\, K}$.
А можно, как и в случае поляры, доказать и непосредственно. Положим $Q = \overline{co\, K}$. Это выпуклый замкнутый конус, причем $Q^* = K^*$ (свойство 3). Тогда $K^{**} = Q^{**}$, а по теореме 5.2 $Q^{**} = Q = \overline{co\, K}$.

Для выпуклых множеств A и B мы рассматривали множества $A \bigcap B$ и $co\,(A \bigcup B)$, и находили формулы для их поляр. Если A и B — выпуклые конусы, то, как мы знаем, $co\,(A \bigcup B) = A + B$, поэтому для конусов естественны операции пересечения и суммы. Посмотрим, что будет происходить при этом с сопряженными конусами.

Пусть K_1 и K_2 — произвольные конусы (не обязательно выпуклые). Тогда:

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

7) $(K_1 \cap K_2)^* \supset \overline{K_1^* + K_2^*}$. (Очевидно.)

Как и в случае поляры, равенство в последнем свойстве имеет место при некотором дополнительном предположении.

7′) Пусть K_1, K_2 — выпуклые замкнутые конусы. Тогда

$$(K_1 \cap K_2)^* = \overline{K_1^* + K_2^*}\;. \tag{5.11}$$

Действительно, пусть $C_1 = K_1^*$, $C_2 = K_2^*$. Тогда по теореме 5.2 $K_1 = C_1^*$, $K_2 = C_2^*$, и пользуясь свойством 6 и следствием 1 из теоремы 5.2, получаем

$$(K_1 \cap K_2)^* = (C_1^* \cap C_2^*)^* = (C_1 + C_2)^{**} = \overline{C_1 + C_2} = \overline{K_1^* + K_2^*}\;.$$

Упражнение 3. Доказать по индукции, что для любого конечного числа выпуклых замкнутых конусов

$$(K_1 \cap \ldots \cap K_m)^* \;=\; \overline{K_1^* + \ldots + K_m^*}\,. \tag{5.12}$$

Убрать замыкание в правой части (5.11) (а значит, и в (5.12)), пользуясь свойством 8 для поляры пересечения, непосредственно нельзя, так как там присутствует требование, чтобы точка 0 была внутренней для обоих множеств, а для конусов это требование приводит к тривиальному случаю (тогда конус есть все пространство).

Задача 5.1. Показать, что действительно убрать замыкание в (5.11), вообще говоря, нельзя. Для этого достаточно привести пример двух выпуклых замкнутых конусов $C_1,\ C_2,$ сумма которых $C_1 + C_2$ есть незамкнутый конус.

(*Указание:* в пространстве $\mathbb{R}^3$ рассмотреть конус $C_1,$ образованный кругом $(z-1)^2 + x^2 \le 1,\quad y = 1,$ и точкой 0. В качестве C_2 взять ось y. Показать, что конус $C_1 + C_2$ не замкнут.)

Напомним, что если два непустых выпуклых конуса не пересекаются, то они отделяются гиперплоскостью, проходящей через 0. Докажем обобщение этого свойства на случай нескольких конусов. Нам понадобится следующий простой факт.

Лемма 5.1. Пусть A — произвольное множество, и линейный функционал $p \ne 0$ таков, что $(p, A) \ge c.$ Тогда $\forall\, x \in \operatorname{int} A$ выполнено $(p,x) > c.$ Другими словами, ненулевой линейный функционал не может достигать минимума (и точно так же максимума) во внутренней точке множества.

Доказательство. Допустим, что $\exists\, x_0 \in \operatorname{int} A,$ для которого $(p, x_0) = c.$ Возьмем любой вектор $v \in \mathbb{R}^n$ такой, что $(p, v) < 0$ (такой всегда есть, ибо $p \ne 0$). Тогда при малом $\varepsilon > 0$ точка $x_0 + \varepsilon v \in A,$ и на ней $(p,\ x_0 + \varepsilon v) = c + \varepsilon (p,v) < c,$ противоречие. $\square$

Теорема 5.3 (Дубовицкого–Милютина о непересечении конусов).

Пусть $K_1, \ldots,\ K_m,\ H$ – непустые выпуклые конусы, из которых первые m открыты. Тогда пересечение всех этих конусов пусто $\iff$ существуют функционалы $p_i \in K_i^*,\ \ i = 1, \ldots, m,\ \ q \in H^*,$ не все равные нулю (т.е. набор этих функционалов нетривиален), сумма которых равна нулю:

$$p_1 + \ \ldots\ + p_m\, +\, q\ =\ 0. \qquad (5.13)$$

Равенство (5.13) называется уравнением Эйлера–Лагранжа. Это кажущееся странным название объясняется тем, что все необходимые условия локального минимума первого порядка в самых различных задачах на экстремум, в том числе и уравнение Эйлера–Лагранжа в задачах вариационного исчисления, и даже Принцип максимума Понтрягина (!) могут быть получены из этого, на первый взгляд примитивного, равенства. Теорема Дубовицкого–Милютина является удобным и эффективным инструментом в различных вопросах теории экстремума, в особенности при выводе необходимых условий оптимальности; она является одним из ключевых пунктов т.н. схемы (или метода) Дубовицкого–Милютина. Подробнее об этом см. [55].

Доказательство. Импликация $(\Longleftarrow)$ доказывается просто. От противного: допустим, что выполнено уравнение Эйлера, но $\exists\, x \in \bigcap K_i\, \bigcap H.$ Если все $p_i = 0,$ то из (5.13) следует, что и $q = 0,$ т.е. набор функционалов тривиален, противоречие.

Поэтому хотя бы один $p_s \neq 0$. Но тогда из неравенства $(p_s, K_s) \geq 0$, учитывая, что конус K_s открыт, по лемме 5.1 получаем $(p_s, x) > 0$, и тогда $\sum_1^m (p_i, x_i) + (q, x) > 0$ (все слагаемые неотрицательны, и по крайней мере одно положительно), что противоречит уравнению Эйлера.

Докажем $(\Longrightarrow)$. Пусть $\bigcap K_i \bigcap H = \emptyset$. Обозначим $\mathbb{R}^n = X$, и рассмотрим пространство $Z = X^m$ с элементами $z = (x_1, \ldots, x_m)$, где каждый x_i есть элемент пространства $\mathbb{R}^n$. В этом пространстве рассмотрим выпуклые конусы $\Omega = K_1 \times \ldots \times K_m$ (он очевидно открыт) и

$$D = \{\ z = (x_1, \ldots, x_m)\ |\ x_1 = \ldots = x_m \in H\ \} \quad - \text{ "диагональ" конуса } H^m.$$

Мы утверждаем, что $\Omega \cap D = \emptyset$. Действительно, в противном случае найдется z, у которого все компоненты $x_i \in K_i$ (ибо $z \in \Omega$), и с другой стороны, все x_i равны между собой: $x_i = x \ \ \forall\, i$, и этот $x \in H$ (ибо $z \in D$). Но тогда вектор x принадлежит всем конусам K_i и H, т.е. их пересечение непусто, противоречие.

Так как Ω и D не пересекаются, то по теореме 3.3 их можно разделить: существует $l \in Z$, $\ l = (p_1, \ldots, p_m) \neq 0$ такой, что $(l, \Omega) \geq 0$ и $(l, D) \leq 0$.

Для любого $z = (x_1, \ldots, x_m) \in Z$ по определению имеем $(l, z) = \sum_1^m (p_i, x_i)$. Покажем, что из $(l, \Omega) \geq 0$ вытекает, что $\forall\, i \ \ (p_i, K_i) \geq 0$. Действительно, возьмем произвольные $x_i \in K_i$, $i = 1, \ldots, m$. Тогда $\forall\, \varepsilon > 0$ точка $z(\varepsilon) = (x_1, \varepsilon x_2, \ldots, \varepsilon x_m) \in \Omega$, поэтому $(l, z(\varepsilon)) \geq 0$, и при $\varepsilon \to 0+$ получим $\lim\,(l, z(\varepsilon)) = (p_1, x_1) \geq 0$. Точно так же $(p_i, x_i) \geq 0 \ \ \forall\, i$, поэтому все $p_i \in K_i^*$.

Положим теперь $q = -\sum p_i$. Тогда равенство (5.13) автоматически выполнено. Неравенство $(l, D) \leq 0$ означает, что $\forall\, x \in H \quad \sum (p_i, x) \leq 0$, т.е. $(q, x) \geq 0$, и следовательно, $q \in H^*$.

Итак, существование требуемых p_i , q установлено, ч.т.д. $\square$

В случае, когда $m = 1$, т.е. имеются лишь два конуса K_1 и H, теорема Дубовицкого–Милютина фактически является переформулировкой теоремы об отделимости для конусов; в общем же случае она представляет собой некую (очень удачную!) адаптацию теоремы об отделимости для целей теории оптимизации.

Из теоремы 5.3 в частности вытекает удобная формула для сопряженного к пересечению конусов, упрощающая полученную ранее формулу (5.11).

Теорема 5.4. Пусть K_1, K_2 открытые выпуклые конусы, имеющие непустое пересечение: $K_1 \cap K_2 \neq \emptyset$. Тогда

$$(K_1 \cap K_2)^* = K_1^* + K_2^*. \tag{5.14}$$

(По сравнению с (5.11) здесь отсутствует замыкание в правой части.)

Доказательство. Включение $\supset$ тривиально. Докажем $\subset$.

Пусть вектор $p \in (K_1 \cap K_2)^*$. Нам надо представить его в виде суммы $p = p_1 + p_2$, где $p_1 \in K_1^*$, $p_2 \in K_2^*$. Считаем $p \neq 0$ (иначе имеется тривиальное представление: $0 = 0 + 0$).

Введем еще один конус: $K_3 = \{\ x \mid (p, x) < 0\}$ — открытое полупространство. Очевидно, $K_1 \cap K_2 \cap K_3 = \emptyset$. (Действительно, иначе $\exists x \in K_1 \cap K_2$ такой, что $x \in K_3$, т.е. $(p, x) < 0$, а это противоречит выбору p.)

Тогда по теореме Дубовицкого–Милютина найдутся $p_i \in K_i^*$, $i = 1, 2, 3$, не все равные 0, такие что $p_1 + p_2 + p_3 = 0$. Но по свойству 5 $p_3 = -\alpha p$, где $\alpha \geq 0$, т.е. $p_1 + p_2 = \alpha p$. Если $\alpha = 0$, то имеем $p_1 + p_2 = 0$, и хотя бы один из этих функционалов ненулевой. Но тогда, опять по теореме Дубовицкого–Милютина (!), получаем, что $K_1 \cap K_2 = \emptyset$, что противоречит условию теоремы. Следовательно, $\alpha > 0$, и тогда $p = \frac{1}{\alpha}(p_1 + p_2) \in K_1^* + K_2^*$, ч.т.д. $\square$

Обратим внимание, что в этом доказательстве мы дважды пользовались теоремой Дубовицкого–Милютина: один раз в одну сторону, а второй раз в другую.

Упражнения.

4) Пусть K_1, K_2 — выпуклые конусы, внутренности которых имеют общие точки: $(int\ K_1) \cap (int\ K_2) \neq \emptyset$. Показать, что тогда верна та же формула (5.14).

Показать, что здесь недостаточно требовать, чтобы $K_1 \cap K_2 \neq \emptyset$. (Рассмотреть контрпример из задачи 5.1.)

5) Доказать формулу (5.14), опираясь на формулу (5.6) для поляры пересечения, с помощью перехода от выпуклого множества A к выпуклому конусу $\overline{con}\,(A, 1)$ (см. лекцию 1).

6) Показать, что конус касательных направлений (см. лекцию 4) и конус внешних нормалей (см. лекцию 3) к выпуклому множеству A в произвольной точке $x \in A$ являются взаимно двойственными: $N_A(x) = -K_A^*(x)$.

7) Пусть выпуклые множества A, C имеют общие внутренние точки.
Тогда $\forall\, x \in A \cap C$ касательный конус к пересечению этих множеств есть пересечение касательных конусов:

$$K_{A\cap C}(x) \ = \ K_A(x) \bigcap K_C(x),$$

а конус внешних нормалей к пересечению есть сумма конусов внешних нормалей к каждому из множеств:

$$N_{A\cap C}(x) \ = \ N_A(x) \ + \ N_C(x).$$

Лекция 6. Полиэдральные множества и двойственные к ним

Здесь мы рассмотрим вопрос о том, какой вид имеет поляра или сопряженный конус в случае, когда исходное множество (соответственно конус) полиэдральное, т.е. задается конечным числом линейных неравенств и равенств. В первую очередь нас будет интересовать случай полиэдрального (конечногранного) конуса.

Вместе с понятием конечногранного конуса имеется также понятие конечнопорожденного конуса.

Определение 6.1. Конус C называется *конечнопорожденным,* если

$$C = \sum_{i=1}^{m} \mathbb{R}_+ a_i \,,$$

где $a_1, \ldots, a_m$ – некоторые векторы из $\mathbb{R}^n$; они называются *образующими* конуса C. (Точнее, образующими называются те из них, которые нельзя в этом равенстве отбросить.)

Определение 6.2. Векторы $a_1, ..., a_m \in \mathbb{R}^n$ называются *позитивно зависимыми,* если $\exists\, \lambda_i \ \geq\ 0$ такие, что $\sum \lambda_i \ >\ 0$ и $\sum \lambda_i a_i \ = 0$.

Очевидно, позитивная зависимость векторов эквивалентна тому, что их выпуклая оболочка содержит 0 (ибо условие $\sum \lambda_i > 0$ можно заменить на $\sum \lambda_i = 1$). Например, если хотя бы один $a_i = 0,$ то весь набор заведомо позитивно зависим.

Если указанных λ_i не существует, то данные векторы называются *позитивно независимыми.* Или, другими словами: векторы позитивно независимы, если их выпуклая оболочка не содержит 0.

Понятие позитивной зависимости (независимости) хотя и похоже на понятие линейной зависимости (независимости), имеет существенное отличие от последнего.
А именно, тогда как в $\mathbb{R}^n$ число линейно независимых векторов не может превосходить $n,$ число позитивно независимых векторов может быть любым. Например, любой ненулевой вектор $a,$ повторенный любое число раз: $a, a, \ldots, a,$ образует позитивно независимую систему.

Более интересный пример: пусть $a_1, \ldots, a_m \in \mathbb{R}^2$ – любые векторы, у которых вторая координата положительна. Покажите, что такой набор позитивно независим. Этот пример есть частный случай следующего общего утверждения.

Лемма 6.1.
Векторы a_i позитивно независимы $\iff$ $\exists\, p \in \mathbb{R}^n : \ (p, a_i) > 0 \ \ \forall\, i$.

Доказательство. $(\Longleftarrow)$ Пусть такой p существует. Если при этом векторы a_i позитивно зависимы, то из равенства $\sum \lambda_i a_i \ = 0$ получаем $\sum \lambda_i (p_i, a_i) \ = 0.$ Но так как все $\lambda_i \ \geq\ 0$ и некоторые из них > 0, а все $(p_i, a_i) > 0$, то слева в последнем равенстве стоит положительное число. Противоречие.

($\Longrightarrow$) Так как $\{\, a_i \,\}$ позитивно независимы, то их выпуклая оболочка A не содержит точку 0. Множество A выпукло и замкнуто (как выпуклая оболочка компакта), поэтому существует функционал p, строго отделяющий A от точки 0 : $\min\,(p,A) > (p,0) = 0$. Отсюда $(p,a_i) > 0 \quad \forall\, i$. $\square$

Следствие. Если векторы $a_1,...,a_m$ позитивно независимы, то порожденный ими конус $\sum \mathbb{R}_+\, a_i$ замкнут.

Доказательство. Если a_i позитивно независимы, то по лемме 6.1 открытые конусы $K_i = \{\, p \,|\, (p,a_i) > 0 \,\}$ имеют непустое пересечение, а тогда по теореме 5.4 и свойству 5 сопряженного конуса имеем

$$\Big(\bigcap K_i\Big)^* = \sum K_i^* = \sum \mathbb{R}_+\, a_i \ .$$

Но конус слева замкнут (как сопряженный к некоторому), а тогда замкнут и конус, стоящий справа. $\square$

Как мы знаем, алгебраическая сумма двух замкнутых множеств не обязательно замкнута. Однако, справедливо следующее утверждение.

Лемма 6.2. Пусть $\mathbb{R}^n$ разбито в прямую сумму: $\mathbb{R}^n = L \oplus M$, и пусть C_L, C_M – замкнутые множества в подпространствах L, M соответственно. Тогда их сумма $C = C_L + C_M$ также замкнута.

Доказательство. Пусть дана последовательность $x_k = y_k + z_k \ \in C$, где $y_k \in C_L$, $z_k \in C_M$, такая что $x_k \to x_0$. Нам надо показать, что $x_0 \in C$.

Так как $x_0 = y_0 + z_0$, где $y_0 \in L$ и $z_0 \in M$ есть проекции x_0 на L и M соответственно, а y_k и z_k есть проекции x_k на L и M, то из того, что $x_k \to x_0$, вытекает, что $y_k \to y_0$ и $z_k \to z_0$ (в силу непрерывности проекции). Из замкнутости C_L, C_M вытекает, что $y_0 \in C_L$, $z_0 \in C_M$, а тогда по определению $x_0 = y_0 + z_0 \in C$. $\square$

Задача. Пусть выпуклый конус K содержит подпространство L, и M есть прямое дополнение к L в $\mathbb{R}^n$. Тогда в подпространстве M имеется выпуклый конус K_M такой, что $K = K_M \oplus L$. (В качестве K_M надо взять проекцию конуса K на подпространство M вдоль L.) Конус K_M есть, так сказать, фактор-конус конуса K по содержащемуся в нем подпространству L .

Лемма 6.3. Пусть $a_1,...,a_m$ — произвольный конечный набор векторов. Тогда конус $C = \sum \mathbb{R}_+\, a_i$ замкнут.

Доказательство. Индукция по m. Для $m = 1$ утверждение выполнено тривиально. Пусть оно выполняется для любых векторов, число которых меньше некоторого m. Рассмотрим произвольный набор из m векторов.

Если a_i позитивно независимы, то конус C замкнут по следствию из леммы 6.1. Пусть теперь a_i позитивно зависимы, т.е. $\exists\, \lambda_i \ge 0$, такие что $\sum \lambda_i > 0$ и

$\sum \lambda_i a_i = 0$. Без нарушения общности считаем $\lambda_m = 1$. (Почему?) Тогда

$$-a_m = \lambda_1 a_1 + \ldots + \lambda_{m-1} a_{m-1}\,,$$

и этот вектор принадлежит C, ибо справа стоит позитивная комбинация элементов выпуклого конуса C. Таким образом, $\pm\, a_m \in C$, и следовательно, прямая $L = \mathbb{R}\, a_m \subset C$. Тогда (по свойству выпуклого конуса) вместе с каждым $x \in C$ будет $x + L \subset C$.

Пусть $\mathbb{R}^n = L \oplus M$, где $M -$ любое прямое дополнение к L. Тогда $a_i = l_i + a_i^{\perp}$, где $l_i \in L$, $a_i^{\perp} \in M$. Рассмотрим конус $C_M = \sum_{i=1}^{m-1} \mathbb{R}_+ \, a_i^{\perp}$, содержащийся в подпространстве M. По предположению индукции он замкнут. Ясно, что наш $C = L + C_M$ (показать!), и тогда по лемме 6.2 он также замкнут. □

Следствие. Пусть $a_1, ..., a_m$ — произвольный конечный набор векторов, $L -$ произвольное подпространство. Тогда конус $C = L + \sum \mathbb{R}_+ \, a_i$ замкнут.

Доказательство. Пусть $\{\, b_i \,\}$ есть любой базис в L. Тогда $L = \sum \mathbb{R}_+ \,(\pm b_i)$, следовательно, конус

$$C = \sum \mathbb{R}_+ a_i \, + \, \sum \mathbb{R}_+ \,(\pm b_i)$$

замкнут по лемме 6.3. (А можно было перейти к $a_i^{\perp}$ и воспользоваться леммой 6.2.) □

Упражнение. Верно ли, что если $K -$ произвольный выпуклый замкнутый конус, а $L -$ подпространство, то конус $K + L$ замкнут? Привести контрпример.

Теорема 6.1. (Лемма Фаркаша). Пусть задан конечногранный конус

$$K = \{\, x \in \mathbb{R}^n \mid \quad (a_i, x) \geq 0, \;\; i = 1, ..., m; \qquad (b_j, x) = 0, \;\; j = 1, ..., s \},$$

где $a_i,\, b_j$ - произвольные векторы из $\mathbb{R}^n$. Тогда

$$K^* = \sum \mathbb{R}_+ \, a_i \, + \sum \mathbb{R}\, b_j \, = \{\, p = \sum \alpha_i a_i + \beta_j b_j \mid \alpha_i \geq 0 \,\}.$$

Доказательство. Введем конусы $K_i = \{\, x \,|\, (a_i, x) \geq 0\}$ и подпространство $M = \{\, x \,|\, (b_j, x) = 0 \;\forall\, j \,\}$. Тогда $K = \bigcap K_i \bigcap M$, следовательно, по формуле (5.12)

$$K^* = \overline{\sum K_i^* \, + \, M^{\perp}} \; = \; \overline{\sum \mathbb{R}_+ a_i \, + \sum \mathbb{R} b_j}\,,$$

а по следствию из леммы 6.3 справа под чертой стоит замкнутый конус, поэтому замыкание брать уже не надо. □

Лемма Фаркаша утверждает, что если вектор p "подчинен" векторам a_i, b_j в том смысле, что из условий $(a_i, x) \geq 0 \;\forall\, i$ и $(b_j, x) = 0 \;\forall\, j$ следует $(p, x) \geq 0$, то этот вектор есть просто позитивно-линейная комбинация векторов $a_i,\, b_j$, т.е. $p = \sum \alpha_i a_i \, + \sum \beta_j b_j$, где все $\alpha_i \geq 0$, а β_j произвольны.

Обратное утверждение (если p имеет такой вид, то он подчинен векторам a_i, b_j) выполняется тривиально.

Упражнение. Доказать лемму Фаркаша с помощью теоремы Дубовицкого–Милютина. (Указание: рассмотреть сначала случай, когда ограничений равенства в определении конуса K нет, а его с помощью индукции свести к случаю, когда вектора a_i позитивно независимы.)

Перейдем теперь от многогранных конусов к многогранным множествам.

Полупространство. Пусть для начала множество Π задано одним неравенством $(l,x) \le c$, где $l \ne 0$, $c > 0$.

Тогда любой вектор $p \in \Pi^0$ имеет вид $p = \lambda l$, $\lambda \ge 0$ (почему?), при этом из $(l,x) \le c$ должно следовать $(p,x) = \lambda(l,x) \le 1$, поэтому $\lambda \le 1/c$. Итак,

$$\Pi^0 = \left\{ p = \lambda l \mid \ 0 \le \lambda \le \frac{1}{c} \right\} \ = \ \frac{1}{c}\,[0,\, l]\ .$$

Если $c = 0$, то Π есть конус и, как мы знаем, $\Pi^0 = \{\, p = \lambda l \mid \lambda \ge 0\,\} \ = \ \mathbb{R}_+ l$.

Многогранное множество, внутренность которого содержит 0. Пусть теперь множество A задано конечным числом неравенств $(l_k, x) \le c_k$, $\quad k = 1, \ldots, r$, где для начала все $c_k > 0$. Если ввести полупространства

$$\Pi_k = \{\, x \mid (l_k, x) \le c_k \,\}, \qquad k = 1, \ldots, r, \tag{6.1}$$

то в каждом из них 0 есть внутренняя точка, $A = \bigcap_k \Pi_k$, и по свойству 5.8

$$A^0 = \ co\, \left(\bigcup_k \Pi_k^0 \right) \ = \ co\, \left\{ 0,\, \frac{1}{c_1} l_1\, , \ldots ,\, \frac{1}{c_m} l_m \right\}$$

— это многогранник.

Общий случай многогранного множества, содержащего 0. Пусть M задано конечным числом неравенств (6.1), где все $c_k > 0$, а также ограничениями

$$(a_i, x) \le 0, \qquad i = 1, \ldots, m, \tag{6.2}$$

$$(b_j, x) = 0, \qquad j = 1, \ldots, s, \tag{6.3}$$

где $a_i,\, b_j \in \mathbb{R}^n$.

Тогда $M = A \cap K$, где A задано неравенствами (6.1), а K есть конус, заданный ограничениями (6.2), (6.3). По лемме Фаркаша

$$K^0 = (-K)^* = \ \sum \mathbb{R}_+ a_i \ + \sum \mathbb{R} b_j\ ,$$

это конечнопорожденный конус. По свойству поляры 7' (лекция 5) $M^0 = \overline{co\,(A^0 \cup K^0}$, и так как A^0 – выпуклый компакт, содержащий 0, а K^0 – замкнутый выпуклый конус, то по лемме 4.6

$$M^0 = A^0 + K^0 = \ co\, \left\{ 0,\, \frac{1}{c_1} l_1\, , \ldots ,\, \frac{1}{c_m} l_m \right\} \ + \sum \mathbb{R}_+ a_i \ + \sum \mathbb{R} b_j\ ,$$

т.е. M^0 состоит из всех векторов вида

$$p = \sum_{k=1}^{r} \frac{\alpha_k}{c_k} l_k + \sum_{i=1}^{m} \lambda_i a_i + \sum_{j=1}^{s} \mu_j b_j \,,$$

где $\alpha_k \geq 0$, $\sum \alpha_k \leq 1$, $\lambda_i \geq 0$, $\mu_j \in \mathbb{R}$.

Двойственное описание полиэдральных множеств

Напомним, что *полиэдром* (полиэдральным, или многогранным множеством) называется множество решений системы неравенств

$$(a_i, x) \leq c_i\,, \qquad i = 1, \ldots, k\,;$$

(сюда можно добавить и равенства, что, впрочем, не меняет общности), т.е. множество, являющееся пересечением конечного числа замкнутых полупространств.

В частности, полиэдральный (многогранный или конечногранный) конус — это множество решений системы неравенств $(a_i, x) \leq 0$, $i = 1, \ldots, k$.

С другой стороны, имеются также следующие множества (мы их уже встречали в предыдущих лекциях):

Выпуклый многогранник — это выпуклая оболочка любого конечного множества точек: $co\,\{x_1, \ldots, x_m\}$;

Конечнопорожденный конус $\sum \mathbb{R}_+ a_i$ (минимальный замкнутый конус, порожденный векторами $a_1, \ldots, a_m$).

Таким образом, первые два класса множеств задаются в виде пересечения конечного числа полупространств, а вторые — в виде выпуклой (или конической) оболочки конечного числа векторов.

Следующая замечательная теорема устанавливает, что фактически это одни и те же классы множеств.

Теорема 6.2. (Минковского-Вейля).

1) Множество является ограниченным полиэдром $\Longleftrightarrow$ оно есть выпуклый многогранник.

2) Множество является многогранным конусом $\Longleftrightarrow$ оно является конечнопорожденным конусом.

3) Множество полиэдрально $\Longleftrightarrow$ оно есть сумма выпуклого многогранника и конечнопорожденного конуса.

4) Поляра к полиэдральному множеству есть полиэдральное множество.

Геометрически справедливость этих утверждений не вызывает сомнений, однако доказательства первых трех из них отнюдь не тривиальны (утверждение 4 легко следует из трех первых). Мы их здесь не приводим, отсылая заинтересованного читателя к книге [R, §19]. Эта теорема имеет такие следствия:

1) Пересечение конечного числа многогранников есть многогранник.

2) Сумма конечного числа полиэдров есть полиэдр (в частности она замкнута).

3) Образ и полный прообраз полиэдра при аффинном отображении также являются полиэдрами.

4) Если два полиэдра не пересекаются, то их можно строго отделить друг от друга.

Доказательство этих следствий оставляем читателю в качестве упражнений.

Понятие особого ограничения

Конечногранные множества естественным образом возникают в качестве линейных ограничений (равенства и неравенства) в различных задачах оптимизации. Среди ограничений неравенства могут быть так называемые *особые* ограничения. Цель данного раздела — познакомиться с этим понятием.

Рассмотрим следующую систему равенств и строгих неравенств:

$$\left.\begin{aligned} (\, l_i, x\,) < 0, \qquad & i = 1, \ldots, m\,, \\ (\, h_j, x\,) = 0, \qquad & j = 1, \ldots, s\,, \end{aligned}\right\} \tag{6.4}$$

где l_i , h_j – некоторые векторы из $\mathbb{R}^n$.

Лемма 6.4 (частный случай теоремы 5.3). Система (6.4) не имеет решений $\Longleftrightarrow$ существуют $\alpha_i \ge 0$, $i = 1, \ldots, m$; β_j, $j = 1, \ldots, s$, такие что $\sum \alpha_i > 0$ и

$$\sum \alpha_i l_i \; + \; \sum \beta_j h_j \; = 0\,. \tag{6.5}$$

Равенство (6.5) (как и равенство (5.13)) называется уравнением Эйлера–Лагранжа, а коэффициенты α_i, β_j называются множителями Лагранжа.

Доказательство. ($\Longleftarrow$) Пусть указанные α_i, β_j существуют, но, тем не менее, система имеет решение x. Тогда, суммируя эту систему с данными коэффициентами α_i, β_j , получим

$$\sum \alpha_i (l_i, x) + \sum \beta_j (h_j, x) = \Big(\sum \alpha_i l_i + \sum \beta_j h_j\Big)\; x = 0. \tag{6.6}$$

Но $\sum \alpha_i (l_i, x) < 0$, ибо все $\alpha_i \ge 0$ и хотя бы одно $\alpha_i > 0$, а $\sum \beta_j (h_j, x) = 0$, поэтому левая часть равенства (6.6) отрицательна, а справа стоит 0. Противоречие.

($\Longrightarrow$) Если какой-то вектор $l_{i_0} = 0$, то возьмем $\alpha_{i_0} = 1$, и положим остальные коэффициенты $\alpha_i = \beta_j = 0$. Тогда (6.5) очевидно выполнено. Поэтому далее считаем, что все векторы $l_i \ne 0$.

Рассмотрим непустые открытые конусы $K_i = \{\, x \mid (l_i, x) < 0 \,\}$ и подпространство $N = \{\, x \mid (h_j, x) = 0 \;\; \forall\, j \,\}$. По условию $K_1 \bigcap \ldots \bigcap K_m \bigcap N = \emptyset$, поэтому $\exists\, p_i \in K_i^*$ и

$q \perp N$, не все равные 0, такие что $\sum p_i + q = 0$. Но каждый $p_i = -\alpha_i l_i$, где $\alpha_i \ge 0$, а $q = -\sum \beta_j h_j$. Если все $\alpha_i = 0$, то все $p_i = 0$ а тогда и $q = 0$, противоречие. Следовательно, $\sum \alpha_i > 0$ и выполнено (6.5). □

Рассмотрим подробнее случай, когда система (6.4) не имеет решений. Согласно лемме 6.4, это эквивалентно тому что $\exists \alpha_i \ge 0,\ \beta_j$ такие, что $\sum \alpha_i > 0$ и выполнено уравнение Эйлера–Лагранжа (6.5).

Определение 6.3. Будем говорить, что индекс i_0 *особый* (или i_0 -ое ограничение неравенства *особое*), если для любого набора множителей $(\alpha_i,\ \beta_j)$ с указанными свойствами выполнено равенство $\alpha_{i_0} = 0$.

Таким образом, особый индекс это такой, который ни в какой комбинации не участвует в уравнении Эйлера–Лагранжа.

Если же существует набор (α_i, β_j) с указанными свойствами, в котором $\alpha_{i_0} > 0$, то индекс i_0 называется *неособым.*

Вместе с системой (6.4), в которой все неравенства строгие, рассмотрим также соответствующую систему с нестрогими неравенствами:

$$\left.\begin{aligned} &(\, l_i, x\,) \le 0, \qquad i = 1, ..., m\,, \\ &(\, h_j, x\,) = 0, \qquad j = 1, ..., s\,, \end{aligned}\right\} \tag{6.7}$$

Равенства обеих систем можно кратко записывать в матричном виде: $Hx = 0$, где матрица H образована строками h_j.

Основное свойство особого ограничения (или индекса) дается следующей леммой.

Лемма 6.5. Индекс i_0 — особый $\Longleftrightarrow$ существует решение $\widehat{x}$ системы (6.7) такое, что $(l_{i_0}, \widehat{x}) < 0$.

Доказательство. $(\Longleftarrow)$ Если указанный $\widehat{x}$ существует, но i_0 не особый, то для соответствующего набора α_i, β_j (где $\alpha_{i_0} > 0$, а остальные $\alpha_i \ge 0$), умножая уравнение (6.5) на данный $\widehat{x}$, получим противоречие: слева будет отрицательная величина, а справа 0.

$(\Longrightarrow)$ От противного. Допустим, что требуемого $\widehat{x}$ нет. Это означает, что множества (выпуклые конусы)

$$A \,=\, \{\, x \,|\, (l_{i_0}, x) < 0 \,\} \quad \text{и}$$

$$C \,=\, \{\, x \,|\, (l_i, x) \le 0,\ i \ne i_0, \quad Hx = 0 \,\}$$

не пересекаются. Тогда их можно отделить: $\exists\, p \in A^*$, $q \in C^*$, не равные оба нулю и такие, что $p + q = 0$. Как мы уже знаем, $p = -\alpha_{i_0} l_{i_0}$, где $\alpha_{i_0} \ge 0$. Ясно, что $\alpha_{i_0} > 0$ (иначе $p = 0$, а тогда и $q = 0$, что не годится). Поэтому можем считать $\alpha_{i_0} = 1$. Тогда $q = -p = l_{i_0} \in C^*$, а по лемме Фаркаша

$$q = \,-\sum_{i \ne i_0} \alpha_i l_i \,-\, \sum \beta_j h_j\,,$$

где все $\alpha_i \geq 0$. При этом $p + q = 0$ означает, что

$$l_{i_0} + \sum_{i \neq i_0} \alpha_i l_i + \sum \beta_j h_j \ = \ 0,$$

а это и есть уравнение Эйлера–Лагранжа, в котором $\alpha_{i_0} = 1$. Следовательно индекс i_0 не особый. $\square$

Итак, особое ограничение (или индекс) можно определить двумя способами: либо по определению 6.3 (с использованием двойственных переменных, т.е. множителей Лагранжа), либо по лемме 6.5 (с использованием свойств системы (6.7) в исходном пространстве).

Пусть I_0 есть множество всех особых индексов системы (6.7). Для каждого $i \in I_0$ имеется свой вектор $\widehat{x}_i$, указанный в лемме 6.5. Возьмем их сумму: $\widehat{x} \ = \sum_{i \in I_0} \widehat{x}_i$. Тогда $\widehat{x}$ по-прежнему удовлетворяет системе (6.7), и при этом для всех $i \in I_0$ одновременно выполнено $(l_i, \widehat{x}) < 0$, т.е. вектор $\widehat{x}$ направлен строго внутрь всех особых ограничений неравенства.

Замечание. Рассмотренное здесь понятие особого ограничения в точности соответствует известному понятию *особого режима* в задачах оптимального управления. (Более правильно было бы называть эти режимы особыми по отношению к некоторому ограничению неравенства на управление.) В этих бесконечномерных ситуациях также полезно иметь в виду указанную двойственную природу особого ограничения, в частности, существование направления, идущего внутрь особого ограничения.

Лекция 7. Выпуклые функции

Имеются два определения выпуклой функции. Одно из них, традиционное, состоит в следующем. Пусть D есть непустое выпуклое множество в $\mathbb{R}^n$.

Определение 1. Функция $f : D \to \mathbb{R}$ называется выпуклой, если она удовлетворяет следующему свойству: $\forall\, x_1,\ x_2 \in D, \quad \forall\, \alpha_1,\ \alpha_2 \geq 0 : \ \ \alpha_1 + \alpha_2 = 1,$ выполнено *неравенство Йенсена*:

$$f(\alpha_1 x_1 + \alpha_2 x_2) \ \leq\ \alpha_1 f(x_1) + \alpha_2 f(x_2). \tag{7.1}$$

Другими словами, график функции лежит не выше любой своей хорды.

Однако часто бывает удобно считать, что f задана на всем пространстве $\mathbb{R}^n$, но может иметь как конечные значения, так и $\pm\infty$:

$$f:\ \mathbb{R}^n\ \to\ \overline{\mathbb{R}} = \mathbb{R} \cup \{\pm\infty\}\,. \tag{7.2}$$

В этом случае принимается следующее (более современное) определение выпуклой функции.

Определение 2. Функция f выпукла, если ее надграфик

$$epi f = \{(x, z) \in \mathbb{R}^{n+1} \mid z \geq f(x)\}$$

есть выпуклое множество.

Обозначение $epi f$ происходит от слова *epigraph* — надграфик. Мы часто будем обозначать надграфик через E_f , или даже просто E, когда ясно, о какой функции идет речь.

Вместе с надграфиком для функции (7.2) определяется также множество

$$D = dom f = \ \{x \in \mathbb{R}^n \mid \ f(x) < +\infty\},$$

которое называется ее *эффективным множеством* (или эффективной областью) (от слова *domain*).

Ясно, что D есть проекция множества E на подпространство x, т.е. D состоит из тех точек x, в которые проектируется хотя бы одна точка из E (проверить!); при этом

$$E = \ \{\, (x, z) \ \mid x \in D,\ \ z \geq f(x)\,\}\,.$$

Очевидно, что D непусто $\iff$ E непусто. Если E выпукло, то, как мы знаем, D также выпукло. (Обратное, конечно, неверно.)

Определение 3. Функция f вида (2) называется собственной, если всюду $f(x) > -\infty$ и хотя бы в одной точке $f(x) < +\infty$.

Другими словами, функция собственная, если ее надграфик непуст и не содержит "вертикальных" прямых (т.е. прямых, параллельных базисному вектору e_{n+1}).

Как правило мы будем иметь дело с собственными функциями, а несобственные функции будут получаться как "нежелательный" результат операций с собственными.

Примеры.

1. $f(x) \equiv +\infty$ есть несобственная выпуклая функция; у нее $E = \emptyset$, $D = \emptyset$.

2. $f(x) \equiv -\infty$ есть несобственная выпуклая функция; $E = \mathbb{R}^{n+1}$, $D = \mathbb{R}^n$.

3. $f(x) \equiv c \in \mathbb{R}$ — есть собственная выпуклая функция; $E = \{(x, z) \mid z \geq c\}$ — полупространство, $D = \mathbb{R}^n$.

4. Аффинная функция $f(x) = (a, x) + b$, где $a \in \mathbb{R}^n$, $b \in R$. Это собственная выпуклая функция, ее надграфик $E = \{(x, z) \mid (a, x) + b \leq z\}$ есть замкнутое полупространство, $D = \mathbb{R}^n$.

Нетрудно заметить, что и обратно, любое замкнутое полупространство в $\mathbb{R}^{n+1}$, не являющееся "вертикальным" (т.е. не инвариантное относительно сдвигов вдоль вертикальной прямой $\mathbb{R}\, e_{n+1}$), есть надграфик некоторой аффинной функции.

Как связаны два приведенных выше определения выпуклой функции?

Теорема 7.1. *Пусть f — собственная выпуклая функция в смысле определения 2. Тогда существует непустое выпуклое множество D, вне которого $f = +\infty$, а на D функция f конечна и является выпуклой в смысле определения 1.*

Обратно: пусть f выпукла на непустом выпуклом множестве D в смысле определения 1. Тогда, считая $f(x) = +\infty$ вне D, получаем собственную выпуклую функцию в смысле определения 2.

Доказательство. (2) $\Rightarrow$ (1). Пусть f — собственная выпуклая функция в смысле определения 2. Тогда множество $E = epi f$ и его проекция $D = dom f$ непусты и выпуклы. Осталось проверить неравенство Йенсена. Пусть даны точки x_1, $x_2 \in D$ и числа α_1, $\alpha_2 \geq 0$, $\alpha_1 + \alpha_2 = 1$. Тогда по определению

$$(x_1, z_1 = f(x_1)) \in E, \qquad (x_2, z_2 = f(x_1)) \in E,$$

и в силу выпуклости E

$$\alpha_1(x_1, z_1) + \alpha_2(x_2, z_2) \in E,$$

т.е.

$$(\alpha_1 x_1 + \alpha_2 x_2, \quad \alpha_1 z_1 + \alpha_2 z_2) \in E,$$

а это означает по определению, что

$$f(\alpha_1 x_1 + \alpha_2 x_2) \leq \alpha_1 z_1 + \alpha_2 z_2 = \alpha_1 f(x_1) + \alpha_2 f(x_2).$$

(1) $\Rightarrow$ (2). Пусть $f : D \to \mathbb{R}$ выпукла в смысле определения 1, и мы положили $f(x) = +\infty$ вне D. Тогда всюду $f(x) > -\infty$, и на непустом множестве D по условию $f(x) \in \mathbb{R}$, следовательно, f собственная. Надо лишь проверить выпуклость ее надграфика $E = \{(x, z) \mid x \in D, \quad f(x) \leq z\}$.

Возьмем любые точки (x_1, z_1), $(x_2, x_2) \in E$ и числа $\alpha_1, \alpha_2 \geq 0$, $\alpha_1 + \alpha_2 = 1$. По определению $f(x_1) \leq z_1$ и $f(x_2) \leq z_2$, а в силу неравенства Йенсена имеем

$$f(\alpha_1 x_1 + \alpha_2 x_2) \ \leq \ \alpha_1 f(x_1) + \alpha_2 f(x_2) \ \leq \ \alpha_1 z_1 + \alpha_2 z_2,$$

откуда $(\alpha_1 x_1 + \alpha_2 x_2,\ \alpha_1 z_1 + \alpha_2 x_z) \in E$, т.е. $\alpha_1(x_1, z_1) + \alpha_2(x_2, z_2) \in E$, и следовательно, E выпукло. Теорема доказана. $\square$

Таким образом, если ограничиться рассмотрением собственных функций, оба определения выпуклой функции по сути дела дают одно и то же.

Отметим несколько простых свойств выпуклых функций.

Свойство 1. Пусть функция $f : D \to \mathbb{R}$ выпукла. Тогда $\forall\, x_1, \ldots, x_k \in D$, $\forall\, \alpha_1, \ldots, \alpha_k \geq 0$, $\sum \alpha_i = 1$, справедливо "многоточечное" неравенство Йенсена

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

Доказательство проводится индукцией по k; мы оставляем его читателю в качестве несложного упражнения.

С выпуклыми функциями можно производить ряд операций, не выходя из класса выпуклых функций; важнейшими из них являются сумма функций $f(x) + g(x)$ и максимум функций $max\,(f(x), g(x))$.

Свойство 2. Если f и g собственны и выпуклы, то $f + g$ и $max\,(f, g)$ тоже выпуклы. Обе эти новые функции имеют $dom = D_f \cap D_g$.

Доказательство. Выпуклость $f + g$ удобно проверять с помощью неравенства Йенсена, а выпуклость $max(f, g)$ через выпуклость надграфика, заметив предварительно, что

$$E_{max(f,g)} \ = \ E_f \bigcap E_g \ .$$

Эту элементарную проверку мы оставляем читателю. $\square$

Обратим внимание, что если f и g были собственными, но их эффективные множества D_f и D_g не пересекались (или, что то же самое, их надграфики E_f и E_g не пересекались), то функции $f + g$ и $max\,(f, g)$ уже несобственные: они всюду равны $+\infty$.

Из свойства 2 вытекает простое, но важное

Следствие. Если $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ – выпуклая, а $g = (a, x) + b : \mathbb{R}^n \to \mathbb{R}$ – аффинная, то $f + g$ выпуклая, и ее эффективная область есть D_f .

Другими словами, к выпуклой функции можно прибавлять (а значит, и вычитать из нее) любую аффинную, сохраняя выпуклость и не меняя эффективного множества.

Операцию суммы можно рассматривать для любого конечного числа функций. Более того, если $f_1, ..., f_k$ – собственные выпуклые функции, и $\alpha_1, \ldots, \alpha_k$ – произвольные положительные числа, то функция $F(x) = \sum \alpha_i f(x_i)$ также выпукла, и ее эффективное множество

$$D_F \; = \; \bigcap_{i=1}^{k} D_{f_i} \; .$$

Если D_F непусто, то полученная функция F собственная; в противном случае $F \equiv +\infty$ несобственная. Проверка этих утверждений не представляет трудностей.

Несколько интересней обстоит дело с операцией максимума. Отметим, что в классическом анализе, который посвящен главным образом изучению гладких (т.е. дифференцируемых, и даже непрерывно дифференцируемых) функций, эта операция отсутствует (точнее, не рассматривается), ибо она сразу же выводит за рамки класса гладких функций (максимум двух гладких функций, как правило, уже не есть гладкая функция). В выпуклом же анализе гладкость изучаемых функций не является обязательным требованием (выпуклая функция может не быть гладкой), и поэтому операция максимума вполне "законна".

Более того, операция максимума может быть применена не только к конечному, но и к любому бесконечному числу выпуклых функций. А именно, пусть $\{f_\mu(x), \mu \in M\}$ есть семейство выпуклых функций, где индекс μ пробегает некоторое множество M любой мощности (и любой природы). Тогда функция

$$F(x) \; = \; \sup_{\mu \in M} f_\mu(x)$$

также будет выпуклой; у нее, очевидно,

$$E_F = \bigcap_{\mu} E_{f_\mu} \, , \qquad D_F = \bigcap_{\mu} D_{f_\mu} \; .$$

Таким образом, максимуму (точнее, супремуму) функций соответствует пересечение надграфиков. Этот факт сразу же наталкивает на следующее соображение.

Пусть дана выпуклая функция, у которой надграфик E замкнут. Тогда E, как выпуклое замкнутое множество, есть пересечение замкнутых полупространств. Замкнутые же полупространства в $\mathbb{R}^{n+1}$ (все, кроме "вертикальных"), как мы уже знаем, есть надграфики некоторых аффинных функций. Таким образом, если отвлечься от возможного присутствия этих вертикальных полупространств, мы приходим к выводу, что выпуклая функция, имеющая замкнутый надграфик, должна быть супремумом некоторого семейства аффинных функций.

Этот факт называется теоремой Минковского и ниже будет строго доказан. А пока мы рассмотрим несколько примеров выпуклых функций, порождаемых выпуклыми множествами.

5. **Индикаторная функция множества.** Пусть A — непустое выпуклое множество. Тогда

$$\delta_A(x) = \begin{cases} 0, & \text{если } \; x \in A, \\ +\infty, & \text{если } \; x \notin A, \end{cases}$$

есть собственная выпуклая функция; она называется *индикаторной функцией* множества A. Ясно, что $dom\, \delta_A = A$.

6. **Сужение выпуклой функции.** Пусть $f : \mathbb{R}^n \to \mathbb{R}^\bullet = \mathbb{R} \cup \{+\infty\}$ есть собственная выпуклая функция, A — выпуклое множество в $\mathbb{R}^n$. Тогда функция

$$(f\, |_A)(x) = \begin{cases} f(x), & \text{если } x \in A, \\ +\infty, & \text{если } x \notin A, \end{cases}$$

— сужение функции f на множество A — также выпукла. Она очевидно равна $f(x) + \delta_A(x)$; при этом $dom\,(f|_A) = (dom f) \cap A$. Если $dom\, A = \emptyset$, полученная функция $f|_A$ несобственная (она $\equiv +\infty$).

7. **Функция Минковского.** Пусть A — выпуклое множество, содержащее 0. Функция

$$\mu_A(x) = \inf \{r > 0 \mid x \in rA\}$$

называется *функцией Минковского* множества A. (Иногда ее называют также *калибровочной* функцией.) Ясно, что всегда $\mu_A(x) \geq 0$ и $\mu_A(0) = 0$, поэтому μ_A — собственная функция.

Лемма 7.1. *Функция* μ_A *выпукла.*

Доказательство. Пусть даны произвольные точки $x, y \in \mathbb{R}^n$ и числа $\alpha, \beta \geq 0$, $\alpha + \beta = 1$. Покажем, что $\mu(\alpha x + \beta y) \leq \alpha\mu(x) + \beta\mu(y)$.

Считаем, что $\mu(x)$ и $\mu(y)$ конечны, т.е. меньше $+\infty$, иначе требуемое неравенство выполнено тривиально.

Возьмем любые числа $r_x > \mu(x)$ и $r_y > \mu(y)$. Тогда $x \in r_x A$, $y \in r_y A$. Следовательно,

$$\alpha x + \beta y \in \alpha r_x A + \beta r_y A = (\alpha r_x + \beta r_y)\, A$$

(последнее равенство — из-за выпуклости A), откуда по определению

$$\mu(\alpha x + \beta y) \leq \alpha r_x + \beta r_y.$$

Так как это неравенство справедливо $\forall\, r_x > \mu(x)$, $\forall\, r_y > \mu(y)$, то оно справедливо и для их предельных значений:

$$\mu(\alpha x + \beta y) \leq \alpha\mu(x) + \beta\mu(y), \qquad \text{ч. т. д.}$$

Пример: если A есть единичный шар в некоторой норме $\|\cdot\|$, то $\mu_A(x) = \|x\|$.

8. **Опорная функция множества.** Пусть A — произвольное непустое множество в $\mathbb{R}^n$. Функция

$$\varphi(x) = \sup_{a \in A}\, (a, x)$$

называется опорной функцией множества A. Ясно, что она собственная, выпуклая (как супремум линейных), и имеет $f(0) = 0$.

Нетрудно видеть, что при замене A на $co\, A$ и на $\overline{A}$ опорная функция не меняется. (Доказать!) Таким образом, при изучении опорных функций всегда можно считать, что множество A выпукло и замкнуто.

9. Еще один способ построения выпуклой функции – взять выпуклое множество в $\mathbb{R}^{n+1}$ и рассмотреть функцию, график которой есть его нижняя огибающая.

Лемма 7.2. Пусть Q – произвольное выпуклое множество в $\mathbb{R}^{n+1}$. Тогда функция $f : \mathbb{R}^n \to \overline{\mathbb{R}}$, определенная по формуле

$$f(x) = \inf \{ \mu \mid (x, \mu) \in Q \},$$

выпукла. Она собственная $\iff$ Q непусто и не содержит лучей, направленных "вертикально вниз", т.е. лучей вида $(x, \mu) - \mathbb{R}_+ e_{n+1}$.

Доказательство. Надграфик f есть, очевидно, $Q + \mathbb{R}_+ e_{n+1}$, а это выпуклое множество (как сумма двух выпуклых), поэтому f выпукла. Далее, отсутствие направленных вниз лучей эквивалентно тому, что всюду $f(x) > -\infty$, а непустота Q эквивалентна тому, что $f(x) < +\infty$ для некоторого $x \in \mathbb{R}^n$. $\square$

10. Следующие примеры показывают, что эффективное множество у выпуклой функции (точнее, ее в некотором естественном смысле *максимальное* эффективное множество) может быть как замкнутым, так и незамкнутым.

Функции

$$f(x) = \begin{cases} -\sqrt{1 - x^2}, & \text{если } |x| \le 1 \\ +\infty, & \text{иначе} \end{cases}$$

и

$$f(x) = \begin{cases} \frac{1}{\sin x}, & \text{если } 0 < x < \pi \\ +\infty, & \text{иначе} \end{cases}$$

имеют $D = [-1, 1]$ и $(0, \pi)$ соответственно.

При работе с выпуклыми функциями часто используется следующее

Свойство 3. Если f — выпуклая функция, то $\forall\, c$ лебегово множество ее подуровня

$$L_c(f) = \{ x \in \mathbb{R}^n \mid f(x) \le c \} \quad \text{выпукло.}$$

Доказательство. Для случая собственной функции рассуждения следующие. Ясно, что $L_c(f) \subset D_f$, а на D_f справедливо неравенство Йенсена, поэтому, если точки $x_1, x_2 \in L_c(f)$, то для их выпуклой комбинации имеем

$$f(\alpha_1 x_1 + \alpha_2 x_2) \le \alpha_1 f(x_1) + \alpha_2 f(x_2) \le c,$$

откуда $\alpha_1 x_1 + \alpha_2 x_2 \in L_c(f)$, и поэтому $L_c(f)$ выпукло.

Для несобственной функции продумать доказательство самостоятельно. $\square$

Вопрос: верно ли обратное утверждение? Ответ — нет. Нетрудно заметить, что у любой монотонной функции $f : \mathbb{R} \to \mathbb{R}$ множества подуровней выпуклы, но не любая монотонная функция выпукла.

Полезно также иметь в виду следующее

Свойство 4. Пусть функция $f : \mathbb{R}^n \to \mathbb{R}$ – выпукла, и $x,\, h \in \mathbb{R}^n$ таковы, что $f(x+h) - f(x) = b$. Тогда $\forall\, t \geq 1$ выполнено неравенство $f(x+th) - f(x) \geq b\,t$. Т.е. график выпуклой функции лежит выше (точнее, не ниже) прямой, полученной из любой хорды, вне отрезка этой хорды.

Доказательство. Считаем $x = 0, \quad f(x) = 0$. Тогда $f(h) = b$. Так как $h = \frac{1}{t}(t\,h) + (1 - \frac{1}{t})\,0$, то $f(h) \leq \frac{1}{t}\, f(t\,h)$, откуда $f(t\,h) \geq t\, f(h) = t\,b$. □

Сублинейные функции

В классе всех выпуклых функций имеется важный подкласс, элементы которого играют роль, аналогичную роли конусов в классе всех выпуклых множеств. Это функции, обладающие, кроме выпуклости, также свойством положительной однородности.

Напомним, что функция $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ называется *положительно однородной* (первой степени), если $\forall\, x \in \mathbb{R}^n,\ \forall\, \alpha > 0 \qquad f(\alpha x) = \alpha f(x)$.

Определение 4. Функция f называется *сублинейной* (иногда говорят – линейно-выпуклой), если она выпукла и положительно однородна.

Лемма 7.3. Функция f сублинейна $\Longleftrightarrow$ ее надграфик E_f – конус.

Доказательство. ($\Longrightarrow$) Пусть $(x, z) \in E_f$, т.е. $f(x) \leq z$. Тогда $\forall\, \alpha > 0$ $f(\alpha x) = \alpha f(x) \leq \alpha\, z$, поэтому $(\alpha x,\, \alpha z) \in E_f$.

($\Longleftarrow$) Пусть $x \in \mathbb{R}^n$, $\alpha > 0$. Надо показать, что $f(\alpha x) = \alpha f(x)$. Так как $(x,\, z = f(x)) \in E_f$, то $(\alpha x,\, \alpha z) \in E_f$, т.е. $f(\alpha x) \leq \alpha\, z$. Если $f(\alpha x) < \alpha\, z$, т.е. $f(\alpha x) = \alpha\, z'$, $z' < z$, то $(\alpha x,\, \alpha z') \in E_f$, и так как E_f – конус, то $(x, z') \in E_f$, а тогда $f(x) \leq z' < z$, противоречие. Поэтому $f(\alpha x) = \alpha f(x)$. □

Упражнение. Покажите, что собственная функция f сублинейна $\Longleftrightarrow$ она положительно однородна и субаддитивна. (Последнее свойство означает, что $\forall\, x, y \in \mathbb{R}^n \quad f(x+y) \leq f(x) + f(y)$.) Сравните с аналогичным утверждением для выпуклого конуса.

Пусть f – сублинейная и собственная функция. Каково ее значение в нуле?
Из собственности следует, что $f(0) > -\infty$. Нетрудно показать, что если $f(0) < +\infty$, т.е. $f(0) \in \mathbb{R}$, то $f(0) = 0$. Действительно, для $x = 0$ и любого $\alpha > 0$ в силу однородности имеем $f(0) = \alpha\, f(0)$, откуда $f(0) = 0$.

Итак, для собственной сублинейной функции значение $f(0)$ либо 0, либо $+\infty$.

Пример, когда $f(0) = +\infty$: пусть на плоскости $\mathbb{R}^2$ надграфик f есть угол, образованный биссектрисой первого квадранта и осью ординат, не включая самой оси. Тогда $f(x) = x$ при $x > 0$, и $f(x) = +\infty$ при $x \leq 0$. (Вместо биссектрисы можно, конечно, взять просто ось абсцисс.)

Пример несобственной сублинейной функции $f : \mathbb{R} \to \overline{\mathbb{R}}$. Положим $f(x) = -\infty$ при $x \le 0$ и $f(x) = +\infty$ при $x > 0$. (Каков ее надграфик?) Мы таких функций будем по возможности избегать.

Свойства сублинейных функций. Сумма и положительная комбинация конечного числа сублинейных функций сублинейна. Верхняя грань любого числа сублинейных функций также сублинейна.

Доказательства этих свойств элементарно; оставляем их читателю.

Примеры сублинейных функций — линейная функция (но не аффинная!), максимум конечного числа линейных функций, любая норма, опорная функция любого множества, функция Минковского выпуклого множества. Выпуклость двух последних функций была установлена выше, а их положительная однородность очевидна (проверить!).

К сублинейным функциям мы еще вернемся, а сейчас зададимся вопросом, как можно узнать, выпукла данная функция или нет.

Критерии выпуклости функций

Установим сначала следующее важное свойство.

Лемма 7.4. Функция $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ выпукла тогда и только тогда, когда ее сужение на любую прямую есть выпуклая функция одного переменного.

Доказательство. Импликация $\Rightarrow$ тривиальна, поэтому требуется доказать лишь обратную импликацию. Если f собственная, то достаточно брать прямые, пересекающие множество $D = dom\, f$. Пусть $\forall\, x_0 \in D,\ \forall\, h \in R^n$ функция $\varphi(t) = f(x_0{+}th)$ выпукла. Из неравенства Йенсена вытекает, что тогда пересечение D с прямой $x_0 + Rh$ выпукло (почему?), и на нем выполняется это неравенство. Отсюда следует, что и само D выпукло, и на нем выполнено неравенство Йенсена, и поэтому f выпукла по определению 1.

В общем случае (если f несобственна) данное свойство вытекает из того факта (а на самом деле ему эквивалентно), что E_f выпукло тогда и только тогда, когда его пересечение с любой двумерной вертикальной плоскостью (т.е. плоскостью, параллельной базисному вектору e_{n+1}) выпукло. (Проверить!) $\square$

Таким образом, выпуклость функции, как и выпуклость множества — это "одномерное" свойство (в отличие, скажем, от непрерывности или дифференцируемости функции).

Напомним известные критерии выпуклости функций одного переменного.

Пусть $f : (a, b) \to \mathbb{R}$ — дифференцируемая функция.

Теорема 7.2. *Следующие три свойства эквивалентны:*
1) f *выпукла на* (a, b),

2) $\forall\, x_0, x \in (a,b)$ *функция Вейерштрасса неотрицательна:*

$$\mathcal{E}(x_0,x) = f(x) - f(x_0) - f'(x_0)(x-x_0) \geq 0$$

(т.е. график f *лежит не ниже своей касательной, проведенной в любой его точке),*
3) $f'(x)$ *монотонно неубывает на* (a,b).

Если f *дважды дифференцируема, то перечисленные свойства эквивалентны также следующему:*
4) $\forall\, x \in (a,b) \quad f''(x) \geq 0.$

Доказательство этой теоремы можно найти в любом учебнике математического анализа. (Полезно также доказать ее самостоятельно.)

Посмотрим теперь, какие критерии выпуклости получаются отсюда для функций нескольких переменных. Пусть D — непустое открытое выпуклое множество в $\mathbb{R}^n$, и на нем задана дифференцируемая функция $f: D \to \mathbb{R}$.

Теорема 7.3. *Следующие три свойства эквивалентны:*
1) f *выпукла на* D,
2) $\forall\, x_0, x \in D$ *функция Вейерштрасса неотрицательна:*

$$\mathcal{E}(x_0,x) = f(x) - f(x_0) - (f'(x_0),\, x-x_0) \geq 0, \tag{7.5}$$

3) $\forall\, x_0, x_1 \in D$

$$(f'(x_1) - f'(x_0),\, x_1 - x_0) \geq 0 \tag{7.6}$$

(в этом случае говорят, что отображение $x \mapsto f'(x)$ *монотонно).*

Если f *дважды дифференцируема на* D, *то перечисленные свойства эквивалентны также следующему:*
4) $\forall\, x \in D$ *матрица вторых производных неотрицательно определена:*

$$f''(x) \geq 0.$$

(Здесь и далее неравенство $S \geq 0$ для симметрической матрицы S означает, что $\forall\, h \in \mathbb{R}^n \quad (Sh,h) \geq 0$).

Доказательство. согласно свойству 4, достаточно рассмотреть функцию $\varphi(t) = f(x_0 + t\bar{x})$, где $x_0 \in D$, $\bar{x} \in \mathbb{R}^n$. Для нее

$$\varphi'(t) = f'(x_0 + t\bar{x})\,\bar{x}, \qquad \varphi''(t) = (f''(x_0+t\bar{x})\,\bar{x},\, \bar{x}),$$

функция Вейерштрасса в точности совпадает с (7.5) (при $\bar{x} = x - x_0$, $t = 1$), неравенство (7.6) превращается в свойство монотонности $\varphi'(t)$, а неравенство $\varphi''(0) = (f''(x_0)\bar{x}, \bar{x}) \geq 0 \;\; \forall\, \bar{x} \in \mathbb{R}^n$ в точности означает, что матрица $f''(x_0)$ неотрицательно определена. $\square$

Напомним, что согласно критерию Сильвестра симметрическая матрица $S \geq 0$ тогда и только тогда, когда все ее главные миноры (не только угловые, а любые миноры с одинаковыми номерами строк и столбцов) неотрицательны.

Мы пишем $S > 0$, если $\forall\, h \neq 0 \quad (Sh, h) > 0$. Из компактности единичной сферы в $\mathbb{R}^n$ вытекает, что тогда на единичной сфере $(Sh, h) \geq \alpha > 0$, и следовательно, $\forall\, h \in \mathbb{R}^n$ выполнена оценка $(Sh, h) \geq \alpha |h|^2$. Такая матрица S называется положительно определенной. Критерий Сильвестра утверждает, что $S > 0$ тогда и только тогда, когда все ее угловые (а тогда и все главные) миноры положительны.

Примеры.

11. Квадратичная функция

$$f(x) = (Sx, x) + (p, x) + r \ : \ \mathbb{R}^n \to \mathbb{R}$$

выпукла тогда и только тогда, когда $S \geq 0$.

Действительно, в любой точке x вторая производная $f''(x) = S$; далее надо применить теорему 7.3.

12. Функция $f(x) = |Ax - b|^2$, где $A -$ $n \times n-$ матрица, $b \in \mathbb{R}^n$, всегда выпукла. Действительно, $f(x) = (Ax - b, Ax - b) = (A^* Ax, x) - 2(Ax, b) + |b|^2$,
а матрица A^*A всегда неотрицательно определена.

В завершении этой лекции рассмотрим вопрос о суперпозиции выпуклых функций.

Лемма 7.5. Пусть $D \subset \mathbb{R}^n$ — выпуклое множество, $f : D \to \mathbb{R}$ — выпуклая функция, $f(D) \subset (a, b)$, и дана функция $g : (a, b) \to \mathbb{R}$, которая выпукла и монотонно неубывает (случаи $a = -\infty$ или $b = +\infty$ не исключаются). Тогда их суперпозиция $\Phi(x) = g(f(x))$ выпукла на D.

Доказательство. Функция f удовлетворяет на D неравенству Йенсена:

$$f(\alpha_1 x_1 + \alpha_2 x_2) \leq \alpha_1 f(x_1) + \alpha_2 f(x_2),$$

откуда в силу монотонности g имеем:

$$g(f(\alpha_1 x_1 + \alpha_2 x_2)) \leq g(\alpha_1 f(x_1) + \alpha_2 f(x_2)),$$

а последняя величина в силу выпуклости g

$$\leq \alpha_1 g(f(x_1)) + \alpha_2 g(f(x_2)).$$

Таким образом, мы получили $\Phi(\alpha_1 x_1 + \alpha_2 x_2) \leq \alpha_1 \Phi(x_1) + \alpha_2 \Phi(x_2)$,
т.е. Φ удовлетворяет неравенству Йенсена на D. $\square$

Следствие. Пусть $f : D \to \mathbb{R}$ выпукла.
Если $f > 0$ на D, то $\forall\, p \geq 1$ функции $f^p(x)$, $e^{f(x)}$ также выпуклы на D.
Если $f < 0$ на D, то выпуклы функции $-\frac{1}{f(x)}$, $e^{f(x)}$, $-\ln(-f(x))$.

Вогнутые функции. Вместе с выпуклыми иногда рассматриваются также *вогнутые функции* — это те функции $f(x)$, у которых $-f(x)$ есть выпуклая функция. Другими словами, это те функции, которые имеют выпуклый *подграфик*, или для которых неравенство Йенсена выполнено с противоположным знаком. Самостоятельного значения вогнутые функции не имеют.

Выпуклые функции иногда называют функциями, выпуклыми вниз, а вогнутые — выпуклыми вверх.

Если f дважды дифференцируема, то ее вогнутость эквивалентна тому, что $\forall\, x$ $f''(x) \leq 0$. Для использования критерия Сильвестра в этом случае надо иметь в виду, что знак любого главного минора зависит от четности его размера, поэтому удобнее проверять не вогнутость f, а выпуклость $-f$.

Задачи. 1) Пусть $\pm f : \mathbb{R}^n \to \mathbb{R}$ – выпуклые функции. Тогда f аффинна. Другими словами, если функция одновременно и выпукла, и вогнута, то она аффинна.

2) Сформулируйте и докажите аналог леммы 7.5 для вогнутых функций.

Лекция 8. Непрерывность выпуклых функций

Свойство выпуклости функции оказывается настолько сильным, что оно автоматически обеспечивает ряд других важных свойств функции, в частности ее непрерывность. Доказательство этого факта удобно разбить на ряд этапов.

Пусть $D \subset \mathbb{R}^n$ – выпуклое множество, $f : D \to \mathbb{R}$ – выпуклая функция, и дана точка $x_0 \in D$.

Лемма 8.1. Пусть $f(x) - f(x_0) \le C$ на некотором шаре $B_r(x_0) \subset D$, $r > 0$. Тогда и $|f(x) - f(x_0)| \le C$ на $B_r(x_0)$.

Доказательство. Без нарушения общности считаем $x_0 = 0$, $f(x_0) = 0$. Тогда $\forall\, x \in B_r(0)$ будет $-x \in B_r(0)$, поэтому $f(x) \le C$ и $f(-x) \le C$.

Так как $0 = \frac{x+(-x)}{2}$, то в силу выпуклости, $0 = f(0) \le \frac{1}{2}(f(x) + f(-x))$,

следовательно, $-f(x) \le f(-x) \le C$. Отсюда $|f(x)| \le C$ на $B_r(0)$, ч.т.д. $\square$

Полученную оценку можно существенно усилить, опять опираясь на выпуклость функции.

Лемма 8.2. Пусть $f(x) - f(x_0) \le C$ на шаре $B_r(x_0) \subset D$. Тогда $\forall\, \varepsilon \in [0, 1]$, $\forall\, x \in B_{\varepsilon r}(x_0)$ выполнена оценка

$$|f(x) - f(x_0)| \le \varepsilon C . \tag{8.1}$$

Доказательство. Считая опять $x_0 = 0$, $f(x_0) = 0$, имеем $\forall\, z \in D :\ \ \varepsilon z = (1-\varepsilon)\cdot 0 + \varepsilon \cdot z$, поэтому $f(\varepsilon z) \le (1-\varepsilon) f(0) + \varepsilon f(z) = \varepsilon f(z)$. Если z пробегает $B_r(0)$, то $x = \varepsilon z$ пробегает $B_{\varepsilon r}(0)$, поэтому для $x \in B_{\varepsilon r}(0)$ получаем $f(x) \le \varepsilon f(z) \le \varepsilon C$. Отсюда по лемме 8.1 получаем $|f(x)| \le \varepsilon C$ на $B_{\varepsilon r}(0)$. $\square$

Перепишем полученную оценку (8.1) в более удобном виде. Для любого $x \in B_r(x_0)$ положим $\varepsilon = |x - x_0|/r$. Тогда очевидно $x \in B_{\varepsilon r}(x_0)$, ибо $|x - x_0| \le \varepsilon r$. Поэтому оценка (8.1) принимает вид :

$$|f(x) - f(x_0)| \le \frac{|x - x_0|}{r}\, C ,$$

т.е.

$$|f(x) - f(x_0)| \le \frac{C}{r}\, |x - x_0| . \tag{8.2}$$

Таким образом, при выполнении условия леммы 8.2 (или леммы 8.1), на $B_r(x_0)$ выполнена оценка (8.2). $\square$

Отсюда вытекает

Лемма 8.3. Пусть выпуклая функция $f : D \to \mathbb{R}$ ограничена сверху в некоторой окрестности точки $x_0 \in \operatorname{int} D$. Тогда она непрерывна в x_0 .

Доказательство. По условию при некотором $r > 0$ функция f ограничена сверху на $B_r(x_0)$, поэтому мы находимся в условиях леммы 8.2, согласно которой выполнена оценка (8.2), а из нее очевидно вытекает что $f(x) \to f(x_0)$ при $x \to x_0$. $\square$

Следствие. Если собственная выпуклая функция $f : \mathbb{R}^n \to \mathbb{R}^\bullet$ ограниченна сверху ($f(x) \le \text{const}$) на некотором открытом множестве $\mathcal{U}$, то f непрерывна в каждой его точке.

Доказательство. Так как всюду $f > -\infty$, а на $\mathcal{U}$ к тому же $f(x) \le \text{const}$, то $f(\mathcal{U}) \subset \mathbb{R}$. А так как открытое множество по определению есть окрестность любой своей точки, то по лемме 8.3 f непрерывна в каждой точке $\mathcal{U}$. $\square$

Теперь мы готовы доказать объявленный выше факт.

Теорема 8.1. Пусть $f : D \to \mathbb{R}$ – выпуклая функция, определенная на выпуклом множестве $D \subset \mathbb{R}^n$. Тогда она непрерывна на $\operatorname{int} D$.

Доказательство. Возьмем произвольную точку $x_0 \in \operatorname{int} D$. В свете леммы 8.3 нам достаточно показать, что найдется окрестность точки x_0, на которой f ограничена сверху. Будем искать эту окрестность как внутренность некоторого многогранника, например симплекса (или, скажем, куба).

Пусть M есть произвольный многогранник в $\mathbb{R}^n$, у которого 0 – внутренняя точка. (Ясно, что такой многогранник найдется — достаточно взять любой многогранник с непустой внутренностью и переместить начало координат в любую его внутреннюю точку.) Тогда $\forall\, \varepsilon > 0$ множество $A_\varepsilon = x_0 + \varepsilon M$ также есть многогранник, и x_0 – его внутренняя точка. Так как $x_0 \in \operatorname{int} D$, то при достаточно малых ε будет $A_\varepsilon \subset D$, поэтому на всем A_ε функция f конечна.

Пусть $a_1, \dots, a_N$ есть множество вершин A_ε, т. е. $A_\varepsilon = \operatorname{co}\{a_1, \dots, a_n\}$. Положим $C = \max\limits_{1 \le i \le N} f(a_i)$. Тогда $C < +\infty$. А так как каждая точка $x \in A_\varepsilon$ представима в виде выпуклой комбинации вершин:

$$x = \sum_{i=1}^{N} \gamma_i a_i, \qquad (\gamma_i \ge 0, \;\; \sum \gamma_i = 1),$$

то в силу выпуклости f имеем $f(x) \le \sum \gamma_i f(a_i) \le C$.

Итак, $\forall\, x \in A_\varepsilon$ справедлива оценка $f(x) \le C$. В частности, эта оценка выполнена на открытом множестве $int\, A_\varepsilon$, содержащем точку x_0. Тем самым мы нашли окрестность точки x_0, на которой f ограничена сверху. Теорема доказана. $\square$

Полученный результат можно сформулировать в следующем виде: *Если $f : \mathbb{R}^n \to \mathbb{R}^\bullet$ – собственная выпуклая функция, $D = dom\, f$, то f непрерывна на $int\, D$.*

В граничных точках множества D непрерывности f, вообще говоря, нет. Пример (рис. **??**): $D = [0, +\infty)$, $f(x) = 0$ при $x > 0$, $f(0) = 1$.

Замечание. Лемма 8.3 и ее следствие справедливы в любом нормированном пространстве, так как в их доказательстве мы нигде не пользовались конечномерностью пространства. Однако теорема 8.1 в бесконечномерном нормированном пространстве становится неверной. (Построить контрпример!) Для ее справедливости надо потребовать, чтобы f была ограниченной сверху на некотором шаре, содержащемся в $int\, D$.

Из теоремы 8.1 вытекает следующее свойство сублинейных функций.

Лемма 8.4. Пусть $f : \mathbb{R}^n \to \mathbb{R}^\bullet$ есть сублинейная функция. Тогда $dom\, f = \mathbb{R}^n \iff f$ ограничена в некоторой окрестности нуля $\iff$ она ограничена на единичном шаре $\iff \exists\, c : \;\; \forall\, x \;\; |f(x)| \le c|x| \iff \exists\, c : \;\; f$ липшицева на $\mathbb{R}^n$ с константой c. Во всех этих свойствах ограниченность эквивалентна ограниченности сверху.

Сублинейные функции, обладающие любым из указанных свойств (и значит, всеми остальными), называются ограниченными (по аналогии с ограниченными линейными функционалами в нормированных пространствах).

Доказательство. Основная цепочка импликаций здесь такова: если $dom\, f = \mathbb{R}^n$, т.е. f конечна на всем $\mathbb{R}^n$, то по теореме 8.1 f непрерывна на $\mathbb{R}^n$, тогда она ограничена в некоторой окрестности нуля, а следовательно, ограничена на единичном шаре некоторой константой c. Отсюда в силу однородности $\forall\, x \;\; |f(x)| \le c|x|$, а тогда в силу субаддитивности $\forall\, x, h \;\; f(x+h) \le f(x) + f(h)$, следовательно $f(x+h) - f(x) \le c|h|$, и точно так же $f(x) \le f(x+h) + f(-h)$, откуда $f(x) - f(x+h) \le c|h|$, поэтому $|f(x+h) - f(x)| \le c|h|$, а это и есть липшицевость с константой c. Обратная цепочка импликаций очевидна. $\square$

Обратим еще раз внимание на оценку (8.2) — она наталкивает на мысль, что в окрестности каждой точки из $\operatorname{int} D$ выпуклая функция f не только непрерывна, но более того, липшицева. Это действительно так, и мы сейчас строго установим это свойство.

Нам потребуется следующий простой, но важный факт, справедливый для произвольного метрического пространства X (не только для $X = \mathbb{R}^n$).

Лемма 8.5. Пусть Q — компакт в метрическом пространстве X, содержащийся в открытом множестве $\mathcal{U}$. Тогда $\exists\, \varepsilon > 0$ такое, что его $\varepsilon-$ *раздутие*

$$B_\varepsilon(Q) = \bigcup_{x \in Q} B_\varepsilon(x) \quad \text{также содержится в } \mathcal{U} .$$

(Для нормированного пространства X очевидно $B_\varepsilon(Q) = Q + B_\varepsilon(0)$.)

Доказательство. Допустим, что такого $\varepsilon > 0$ нет. Это означает, что $\forall\, k$ при $\varepsilon = 1/k$ существует точка $x_k \in Q$, для которой шар $B_\varepsilon(x_k)$ содержит точки из дополнения к $\mathcal{U}$, т.е. $\exists y_k \in \mathcal{U}' = X \setminus \mathcal{U}$ такой, что $\rho(y_k, x_k) \le 1/k$. Так как Q – компакт, то некоторая подпоследовательность $x_{k_m} \to x_0 \in Q$. При этом

$$\rho(y_{k_m}, x_0) \le \ \rho(y_{k_m}, x_{k_m}) + \rho(x_{k_m}, x_0) \ \le \ \frac{1}{k_m} + \rho(x_{k_m}, x_0) \ \to 0,$$

т.е. $y_{k_m} \to x_0$. Поскольку $\mathcal{U}'$ замкнуто (как дополнение к открытому), то $x_0 \in \mathcal{U}'$, т.е. $x_0 \notin \mathcal{U}$. А это противоречит тому, что по условию $x_0 \in Q \subset \mathcal{U}$. □

Следующее простое свойство справедливо для любого *нормированного* пространства, и в частности для пространства $\mathbb{R}^n$.

Лемма 8.6. Пусть Q — выпуклое множество, функция $f : Q \to \mathbb{R}$ такова, что существуют числа $r > 0$, L такие, что $\forall\, x, y \in Q$, $\|x - y\| < r$ выполнена оценка

$$|f(x) - f(y)| \le L\,\|x - y\|\,. \tag{8.3}$$

Тогда эта оценка выполнена $\forall\, x, y \in Q$.

Доказательство. Возьмем произвольные точки $x, y \in Q$. Соединим их отрезком и разобьем его конечным числом точек x_i, $i = 1, \dots, s+1$, которые "возрастают" от $x_1 = x$ до $x_{s+1} = y$. (Т. е. для любой тройки индексов $i < j < k$ точка x_j принадлежит отрезку $[x_i, x_k]$, см. рисунок (**??**). Тогда очевидно:

$$y - x = \sum_{i=1}^{s}(x_{i+1} - x_i), \qquad f(y) - f(x) = \sum_{i=1}^{s}(f(x_{i+1}) - f(x_i)),$$

следовательно,

$$f(y) - f(x) \le \sum_{i=1}^{s} |f(x_{i+1}) - f(x_i)| \le \sum_{i=1}^{s} L\,\|x_{i+1} - x_i\| = L\,\|x_{s+1} - x_1\| = L\,\|x - y\|\,.$$

Здесь последнее неравенство написано в силу имеющейся оценки (8.3) для "близких" точек, а предпоследнее равенство – в силу того, что все точки x_i расположены на одной прямой, причем в монотонном порядке. □

Замечание. Эта лемма остается верной и в следующем усиленном варианте. Пусть L таково, что $\forall\, x \in Q$ $\exists\, r = r(x) > 0$ такое, что f липшицева с константой L на $Q \cap B_r(x)$ (т.е. f *локально липшицева на* Q с единой константой L). Тогда f липшицева на всем Q с той же константой L. (Доказать в качестве упражнения.)

Теорема 8.2. Пусть $f : D \to \mathbb{R}$ — выпуклая функция на выпуклом множестве D. Тогда она липшицева на любом компакте, содержащемся внутри D. Т. е. если компакт $Q \subset \operatorname{int} D$, то $\exists\, L = L(Q)$ такое, что $\forall\, x, y \in Q$

$$|f(x) - f(y)| \le L\,|x - y|\,.$$

Более того (уточнение): если $r > 0$ таково, что $\widetilde{Q} = Q + B_r \subset \operatorname{int} D$, и на $\widetilde{Q}$ выполнено неравенство $M_1 \le f(x) \le M_2$, то можно взять

$$L(Q) = \frac{M_2 - M_1}{r} .$$

Доказательство. Без нарушения общности считаем, что компакт Q выпуклый (иначе можно перейти к его выпуклой оболочке $\operatorname{co} Q$, которая также содержится внутри D). По лемме 8.5 существует такое $r > 0$, что $\widetilde{Q} = Q + B_r \subset \operatorname{int} D$. Множество $\widetilde{Q}$ есть, очевидно, компакт (как алгебраическая сумма двух компактов). По теореме 8.1 f непрерывна на $\widetilde{Q}$, поэтому f ограничена на нем сверху и снизу, т. е. $\exists\, M_1, M_2$ такие, что $M_1 \le f(x) \le M_2$ на $\tilde{Q}$. Так как $\forall\, x_0 \in Q$ шар $B_r(x_0) \subset Q + B_r = \tilde{Q}$, то на этом шаре $f(x) - f(x_0) \le C = M_2 - M_1$.

Тогда по лемме 8.2 $\forall\, x \in B_r(x_0)$ имеем

$$|f(x) - f(x_0)| \le \frac{C}{r}\, |x - x_0| . \tag{8.4}$$

В частности, эта оценка выполнена $\forall\, x \in Q \cap B_r(x_0)$, т. е. $\forall\, x_0, x \in Q$ таких, что $|x - x_0| \le r$. При этом константа $L = C/r$ не зависит от точек x_0, x. Отсюда по лемме 8.6 неравенство (8.4) выполнено $\forall\, x_0, x \in Q$, ч.т.д. $\square$

Замечания. 1. Условие "Q — компакт, содержащийся в $\operatorname{int} D$" — существенно. Если оно нарушается, то, как показывают следующие примеры, теорема 8.2 перестает быть верной (см. рис. **??**):

а) Q – не замкнуто: $f(x) = \frac{1}{x}$ на полупрямой $Q = D = \{x > 0\}$,

б) Q не ограничено: $f(x) = x^2$ на прямой $Q = D = \mathbb{R}$,

в) Q – компакт, но содержится в D, а не в $\operatorname{int} D$:

$$f(x) = -\sqrt{1 - x^2} \quad \text{на отрезке} \quad Q = D = [-1, 1].$$

Здесь при $x \to \pm 1$, $f'(x) \to \pm\infty$, поэтому липшицевости на всем отрезке нет.

2. Тем не менее, освободиться от требования компактности Q в теореме 8.2 все-таки можно, если предположить, что $Q + B_r \subset D$ при некотором $r > 0$, и что f ограничена на $Q + B_r$ и сверху и снизу некоторыми константами M_2 и M_1. (Напомним, что мы эти факты доказывали, а не постулировали). В такой форме теорема 8.2 справедлива и в бесконечномерном нормированном пространстве. (Продумать доказательство самостоятельно.)

Несколько позже, в лекции 12, будет рассмотрен вопрос о дифференцируемости выпуклых функций. А сейчас рассмотрим следующий вопрос.

Глобальная липшицевость выпуклых функций. Пусть $f : \mathbb{R}^n \to \mathbb{R}$ – выпуклая функция, всюду принимающая конечные значения (следовательно, она автоматически собственная). Как мы уже знаем, в этом случае она непрерывна на всем пространстве $\mathbb{R}^n$ и липшицева на любом шаре (а значит, и на любом ограниченном множестве). Интересно выяснить вопрос — когда f липшицева на всем пространстве?

Для сублинейных функций этот вопрос решается просто: согласно лемме 8.4, сублинейная функция, конечная на всем пространстве, является липшицевой на всем пространстве. Её константа Липшица равна максимальному значению модуля функции на единичном шаре.

Для произвольной выпуклой функции $f : \mathbb{R}^n \to \mathbb{R}$ рассмотрим множество $E = epi\, f$. По определению оно выпукло. Кроме того, так как оно задано неравенством $f(x) \le z$, а f непрерывна, то E замкнуто. Так как оно неограничено, то его рецессивный конус $\operatorname{Rec} E$ (который всегда замкнутый и выпуклый) содержит ненулевые элементы. Оказывается, для липшицевости f надо, чтобы этих элементов было "достаточно много".

А именно, пусть $\pi : \mathbb{R}^{n+1} \to \mathbb{R}^n$ есть проекция $(x, z) \mapsto x$. Тогда справедлива

Теорема 8.3. Функция f липшицева на $\mathbb{R}^n \iff \pi(\operatorname{Rec} E_f) = \mathbb{R}^n$.

Для доказательства нам потребуется следующее свойство выпуклых функций, представляющее независимый интерес.

Лемма 8.7. Пусть выпуклая функция $g : \mathbb{R}^n \to \mathbb{R}$ липшицева с константой L, а выпуклая функция $f : \mathbb{R}^n \to \mathbb{R}$ есть ее миноранта: $f(x) \le g(x) \quad \forall x \in \mathbb{R}^n$.
Тогда f тоже липшицева с константой L.

Доказательство. От противного: допустим $\exists\, x_0,\, x_1$ такие что $|f(x_1) - f(x_0)| > L\,|x_1 - x_0|$. Без нарушения общности считаем $f(x_1) - f(x_0) > 0$. Положим $h = x_1 - x_0$. Тогда $|h| > 0$,

$$f(x_0 + h) - f(x_0) = (L + \delta)\,|h|, \qquad \delta > 0,$$

и по свойству 4 выпуклых функций (см. лекцию 7) на луче $x_0 + th$ при $t \ge 1$ будем иметь

$$f(x_0 + th) - f(x_0) \;\ge\; (L + \delta)\,|h|\,t.$$

Пусть $f(x_0) = g(x_0) + C$. Тогда, с одной стороны, $g(x_0 + th) - g(x_0) \le L\,|h|\,t$
(в силу липшицевости), а с другой

$$g(x_0 + th) - g(x_0) \;\ge\; f(x_0 + th) - f(x_0) + C \;\ge\; (L + \delta)\,|h|\,t + C,$$

следовательно, при всех $t \ge 1$ имеем

$$L\,|h|\,t \;\ge\; (L + \delta)\,|h|\,t + C\,.$$

Ясно, что при достаточно больших t получим здесь противоречие. $\square$

Доказательство теоремы 8.3. Без нарушения общности считаем $f(0) = 0$. Тогда точка $(0, 0) \in E$, поэтому $\operatorname{Rec} E_f \subset E_f$.

($\Longrightarrow$) Пусть f липшицева с некоторой константой L. Рассмотрим функцию $g(x) = L\,|x|$. По условию $\forall\, x$ имеем $f(x) = f(x) - f(0) \le L\,|x| = g(x)$, поэтому $E_f \subset E_g$. Так как g принимает конечные значения на всем пространстве, то $\pi\,(E_g) = D_g = \mathbb{R}^n$. Так как g сублинейна и непрерывна, то E_g есть выпуклый замкнутый конус, и следовательно, $\operatorname{Rec} E_g = E_g$. Отсюда и из включения $\operatorname{Rec} E_f \supset \operatorname{Rec} E_g$ вытекает $\pi\,(\operatorname{Rec} E_f) \supset \pi\,(\operatorname{Rec} E_g) = \pi\,(E_g) = \mathbb{R}^n$, ч.т.д. $\square$

($\Longleftarrow$) Здесь мы воспользуемся некоторыми понятиями и фактами из следующей лекции 9. Теорема 8.3 в следующей лекции не используется, поэтому такие действия корректны.

Пусть $\pi\,(\operatorname{Rec} E_f) = \mathbb{R}^n$. Так как множество E_f "выдерживает поднятие", то $\operatorname{Rec} E_f$ содержит вектор e_{n+1} и, будучи конусом, выдерживает добавление луча $\mathbb{R}_+\, e_{n+1}$, т.е. тоже выдерживает поднятие. Кроме того, все его "вертикальные сечения" замкнуты, ибо оно само замкнуто. Тогда по лемме 9.3 $\operatorname{Rec} E_f$ является надграфиком некоторой собственной функции Φ, которая является сублинейной, ибо её надграфик есть выпуклый конус. Из условия $\pi\,(\operatorname{Rec} E_f) = \mathbb{R}^n$, т.е. $\pi\,(E_\Phi) = \mathbb{R}^n$ следует, что на всем пространстве $\Phi < +\infty$, а так как $E_\Phi = \operatorname{Rec} E_f \subset E_f$, то всюду $\Phi \ge f$. Так как Φ есть сублинейная функция, конечная на всем пространстве, то по лемме 8.4 она липшицева с некоторой константой L, а тогда из неравенства $f \le \Phi$ согласно лемме 8.7 следует, что и f липшицева на всем пространстве с той же константой. $\square$

Задача. Указанная здесь функция Φ может быть явно вычислена по формуле

$$\Phi(h) \;=\; \lim_{t\to\infty} \frac{f(th)}{t}\,, \qquad \forall\, h \in \mathbb{R}^n,$$

и тогда глобальная константа Липшица функции f есть

$$L = \max_{|h|=1} |\Phi(h)| \;=\; \max_{|h|=1} \lim_{t\to\infty} \frac{f(th)}{t}\ .$$

Лекция 9. Представление выпуклой функции в виде максимума аффинных

Пусть $f : \mathbb{R}^n \longrightarrow \overline{\mathbb{R}}$ – произвольная функция.

Определение 1. Функция f называется *полунепрерывной снизу в точке* x_0, если

$$f(x_0) \;\le\; \varliminf_{x \mapsto x_0} f(x), \tag{9.1}$$

и просто *полунепрерывной снизу*, если неравенство (9.1) верно для всех $x_0 \in \mathbb{R}^n$.

Из определения нижнего предела вытекает, что свойство полунепрерывности в точке x_0 эквивалентно следующему свойству:

$$\text{если } \; x_k \to x_0 \;\; \text{ и } \;\; f(x_k) \to \mu_0\,, \qquad \text{то } \;\; f(x_0) \le \mu_0\,. \tag{9.2}$$

Кроме того, оно эквивалентно также следующему (формально более сильному) свойству:

$$\text{если } \;\; x_k \to x_0, \;\;\; f(x_k) \le \mu_k, \;\;\; \mu_k \to \mu_0, \qquad \text{то } \;\; f(x_0) \le \mu_0\,. \tag{9.3}$$

Докажем последнюю эквивалентность.

($\Leftarrow$) Положив в (9.3) все $\mu_k = f(x_k)$, получим свойство (9.2), которое эквивалентно (9.1).

($\Rightarrow$) Если верно (9.1), то (9.3) вытекает из следующей цепочки соотношений:

$$f(x_0) \;\le\; \varliminf_{k} f(x_k) \;\le\; \varliminf_{k} \mu_k \;= \lim_{k} \mu_k \;= \mu_0\,.$$

Теорема 9.1. Пусть $f : \mathbb{R}^n \longrightarrow \overline{\mathbb{R}}$ — произвольная функция. Тогда следующие три свойства эквивалентны:

а) f полунепрерывна снизу на $\mathbb{R}^n$;

б) множество $E = \operatorname{epi} f$ замкнуто в $\mathbb{R}^{n+1}$;

в) $\forall\, \mu \in \mathbb{R}$ множество подуровня $L_\mu = \{x \mid f(x) \le \mu\}$ замкнуто в $\mathbb{R}^n$.

Доказательство. Замкнутость E означает, что если $(x_k, \mu_k) \in E$, т. е. $f(x_k) \le \mu_k$, и $(x_k, \mu_k) \to (x_0, \mu_0)$, то $(x_0, \mu_0) \in E$, т. е. $f(x_0) \le \mu_0$. Но это в точности свойство (9.3), которое эквивалентно свойству (а). Таким образом, (а) $\Longleftrightarrow$ (б).

Покажем, что (а) $\Longrightarrow$ (в). Свойство (в) означает, что если $x_k \to x_0$ и $f(x_k) \le \mu_0 \;\; \forall\, k$, то $f(x_0) \le \mu_0$. Но это очевидно следует из (9.3), если положить в нем все $\mu_k = \mu_0$.

Наконец, покажем, что (в) $\Longrightarrow$ (а). Пусть $x_k \to x_0$, $f(x_k) \to \mu_0$, и нам нужно доказать, что $f(x_0) \le \mu_0$. Для любого $\mu > \mu_0$ при достаточно больших k имеем $f(x_k) < \mu$, следовательно $x_k \in L_\mu$. В силу замкнутости L_μ тогда и $x_0 \in L_\mu$, т. е. $f(x_0) \le \mu$. Поскольку это верно $\forall\, \mu > \mu_0$, то $f(x_0) \le \mu_0$.

Итак, (б) $\Longleftrightarrow$ (а) $\Longleftrightarrow$ (в). $\square$

Отметим в частности, что если f непрерывна, то ее надграфик E_f обязательно замкнут.

Далее, любую функцию f можно восстановить по ее надграфику:

$$f(x) = \inf \{ \mu \mid (x,\mu) \in E \} . \tag{9.4}$$

(Напомним, что $\inf \emptyset = +\infty$.)

В связи с этой формулой возникает естественный вопрос. Пусть $\mathcal{E}$ — некоторое множество в $\mathbb{R}^{n+1}$. Когда оно есть надграфик некоторой функции?

Определение 2. Будем говорить, что $\mathcal{E}$ *выдерживает поднятие*, если $\forall (x,\mu) \in \mathcal{E}$, $\forall \mu' > \mu$ точка $(x,\mu') \in \mathcal{E}$.

Т.е. луч, идущий от любой точки множества $\mathcal{E}$ в направлении $n+1$ -го базисного вектора e_{n+1} ("вертикально вверх"), целиком содержится в $\mathcal{E}$. (Это означает, что направление e_{n+1} является рецессивным для $\mathcal{E}$).

Будем говорить, что $\mathcal{E}$ имеет замкнутые вертикальные сечения, если $\forall x \in \mathbb{R}^n$ множество $\mathcal{E} \cap (x + \mathbb{R} \cdot e_{n+1})$, равное $\{ \mu \mid (x,\mu) \in \mathcal{E} \}$, замкнуто.

Лемма 9.1. Множество $\mathcal{E} \subset \mathbb{R}^{n+1}$ является надграфиком некоторой функции $\varphi : \mathbb{R}^n \to \overline{\mathbb{R}}$ $\iff$ оно выдерживает поднятие и все его вертикальные сечения замкнуты. При этом

$$\varphi(x) = \inf \{ \mu \mid (x,\mu) \in \mathcal{E} \} . \tag{9.5}$$

Доказательство. Если $\mathcal{E} = \operatorname{epi} \varphi$ для некоторой функции φ, то указанные свойства и соотношение (9.5) были установлены выше.

Пусть теперь $\mathcal{E}$ обладает указанными свойствами. Построим функцию φ по формуле (9.5) и покажем, что для нее $\operatorname{epi} \varphi = \mathcal{E}$.

По определению $(x,\mu) \in \operatorname{epi} \varphi \iff \varphi(x) \le \mu$, т. е.

$$\varphi(x) = \inf \left\{ \mu' \mid \left(x,\mu'\right) \in \mathcal{E} \right\} \le \mu . \tag{9.6}$$

Первое равенство здесь — согласно (9.4). Мы утверждаем, что последнее неравенство эквивалентно тому, что $(x,\mu) \in \mathcal{E}$. Действительно, если $(x,\mu) \in \mathcal{E}$, то указанное неравенство очевидно. Обратно, пусть выполнено (9.6). Тогда множество $\left\{ \mu' \mid (x,\mu') \in \mathcal{E} \right\}$ непусто, и его $\inf \le \mu$. Если в этом множестве имеется $\mu' \le \mu$, то и само μ ему принадлежит, поскольку $\mathcal{E}$ выдерживает поднятие, и тогда $(x,\mu) \in \mathcal{E}$. Если же в этом множестве все $\mu' > \mu$, то в силу (9.6) $\inf \mu' = \mu$, и тогда в силу замкнутости данного множества опять μ ему принадлежит. Лемма доказана. □

Пусть $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ – произвольная функция. Как уже говорилось, справедлива формула (9.4), где $E = \operatorname{epi} f$. Рассмотрим множество $\overline{E}$ – замыкание множества E. Ясно, что оно по-прежнему выдерживает поднятие (проверьте!) и что его вертикальные сечения замкнуты. Определим функцию

$$\overline{f}(x) = \inf \left\{ \mu \mid (x,\mu) \in \overline{E} \right\} , \tag{9.7}$$

которая называется замыканием функции f.

Согласно лемме 9.1, $\operatorname{epi}(\overline{f}) = \overline{E}$, и так как $\overline{E} \supset E$, то $\overline{f}(x) \le f(x)$.

Определение 3. Функция f называется *замкнутой*, если она совпадает со своим замыканием: $\overline{f}(x) = f(x) \quad \forall\, x \in \mathbb{R}^n$.

Лемма 9.2. Функция f замкнута $\iff$ ее надграфик E замкнут.

Доказательство. $(\Leftarrow)$ Следует из (9.4) и (9.7).
$(\Rightarrow)$ Если $\overline{f} = f$, то надграфики этих функций совпадают: $\overline{E} = E$, а это и означает, что E замкнуто. $\square$

Из этой леммы и теоремы 9.1 вытекает

Следствие. Собственная функция f замкнута $\iff$ она полунепрерывна снизу.

Если f была выпуклой, т.е. ее надграфик E был выпуклым множеством, то и множество $\overline{E}$ выпукло, и поэтому функция $\overline{f}$ также выпукла.

Примеры замыкания выпуклых функций

1)

$$f(x) = \begin{cases} 0, & \text{если } x > 0 \\ +\infty, & \text{если } x \le 0, \end{cases}$$

$$\overline{f}(x) = \begin{cases} 0, & \text{если } x \ge 0 \\ +\infty, & \text{если } x < 0. \end{cases}$$

Т. е. было $f(0) = +\infty$, стало $\overline{f}(0) = 0$.

2) Пусть D – произвольный круг в $\mathbb{R}^2$,

$$f(x) = \begin{cases} 0, & x \in \operatorname{int} D \\ +\infty, & x \notin D \\ \text{любое число} \in [0, +\infty], & x \in \partial D. \end{cases}$$

(Проверить, что это выпуклая функция.) Тогда

$$\overline{f}(x) = \left. \begin{cases} 0, & x \in D \\ +\infty, & x \notin D \end{cases} \right| = \delta_D(x)\,.$$

3)

$$f(x_1, x_2) = \begin{cases} 0, & x_1 > 0 \\ +\infty, & x_1 < 0 \\ \varphi(x_2), & x_1 = 0, \end{cases}$$

где $\varphi \ge 0$ — любая выпуклая функция одной переменной. Тогда

$$\overline{f}(x_1, x_2) = \begin{cases} 0, & x_1 \ge 0 \\ +\infty, & x_1 < 0. \end{cases}$$

Задачи. 1) Пусть f – выпуклая функция, и $\exists\, x_0$, в которой $f(x_0) = -\infty$. Тогда $\overline{f}$ принимает не более двух значений: $\pm\infty$.

2) Будет ли замкнутой сумма двух выпуклых замкнутых функций? (Привести контрпример).

Отметим следующее простое

Свойство. Пусть $f(x) = \sup\limits_{\alpha} f_\alpha(x)$, где $\{\alpha\}$ есть некоторое множество индексов, а каждая f_α – выпуклая замкнутая функция. Тогда f также выпукла и замкнута.

Это вытекает из того, что $\operatorname{epi} f = \bigcap\limits_{\alpha} \operatorname{epi} f_\alpha$.

Следующий факт уже нетривиален.

Лемма 9.3. Пусть f – собственная выпуклая функция. Тогда существует аффинная функция $l(x) = (p, x) + b$, такая что $\forall\, x \quad l(x) \le f(x)$.

Таким образом, каждая собственная выпуклая функция обладает *аффинной минорантой.*

Доказательство. Пусть $D = \operatorname{dom} f$. Достаточно рассмотреть случай, когда $\operatorname{int} D \ne \emptyset$. (Почему?). Возьмем произвольную точку $x_0 \in \operatorname{int} D$ и число $\mu_0 < f(x_0)$. Тогда $(x_0, \mu_0) \notin E = epi f$.

По теореме 3.2 об отделимости $\exists$ вектор $(p, \alpha) \ne 0$ такой, что $\forall\, (x, \mu) \in E$

$$px + \alpha\mu \;\le\; px_0 + \alpha\mu_0\,. \tag{9.8}$$

Так как множество E выдерживает поднятие (т.е. неограниченное увеличение компоненты μ), то очевидно $\alpha \le 0$. Если $\alpha = 0$, то $p \ne 0$ и $\forall\, x \in D$ имеем $px \le px_0$. Но этого не может быть, поскольку $x_0 \in \operatorname{int} D$. Таким образом $\alpha < 0$, и поэтому можно считать $\alpha = -1$.

Тогда из (9.8) при $\mu = f(x)$ получаем, что $\forall\, x \in D$

$$px - f(x) \;\le\; px_0 - \mu_0\,,$$

т.е. $px + b \le f(x)$, где $b = \mu_0 - px_0$. $\square$

Лемма 9.4. Пусть $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ – произвольная функция, и l есть ее аффинная миноранта: $l \le f$. Тогда l есть миноранта и для $\overline{f}$: $l \le \overline{f}$.

Доказательство. Пусть $E = \operatorname{epi} f$, а $\Pi = \operatorname{epi} l$. Так как $l \le f$, то $\Pi \supset E$. Но, так как l непрерывна, Π есть замкнутое множество (полупространство), поэтому $\Pi \supset \overline{E} = \operatorname{epi} \overline{f}$. А это эквивалентно тому, что $l \le \overline{f}$. Лемма доказана. $\square$

Обратное утверждение: «если $l \le \overline{f}$, то $l \le f$ » тривиально вытекает из неравенства $\overline{f} \le f$. Таким образом, любая функция f и ее замыкание $\overline{f}$ имеют одно и то же семейство аффинных минорант.

Лемма 9.5. Пусть f – собственная выпуклая функция. Тогда $\overline{f}$ – собственная выпуклая замкнутая функция.

Доказательство. Выпуклость и замкнутость $\overline{f}$ содержатся в ее определении, поэтому надо доказать лишь собственность. Так как f – собственная, то $\exists\, x_0$, в котором $f(x_0) < +\infty$, а тогда и $\overline{f}(x_0) \le f(x_0) < +\infty$. Осталось показать, что всюду $\overline{f}(x) > -\infty$. По лемме 9.3 f имеет аффинную миноранту l, а по лемме 9.4 она является минорантой и для $\overline{f}$. Тогда $\forall\, x$ получаем $\overline{f}(x) \ge l(x) > -\infty$, ч.т.д. $\square$

Теперь мы готовы доказать обнаруженный еще в лекции 7 факт о представлении выпуклой функции в виде супремума аффинных.

Теорема 9.2 (Г. Минковский). Пусть f есть собственная выпуклая замкнутая функция, и $\mathcal{L} = \{l\}$ есть множество всех ее аффинных минорант. Тогда $\forall\, x$

$$f(x) = \sup_{l \in \mathcal{L}} l(x)\ .$$

Таким образом, всякая собственная выпуклая замкнутая функция есть верхняя грань своих аффинных минорант.

Доказательство. По условию $E = \operatorname{epi} f$ есть непустое выпуклое замкнутое множество в $\mathbb{R}^{n+1}$, и значит по следствию 1 из теоремы 3.1 оно есть пересечение всех замкнутых полупространств, его содержащих. Любое замкнутое полупространство в $\mathbb{R}^{n+1}$ с координатами (x, μ), $x \in \mathbb{R}^n$, $\mu \in \mathbb{R}$, имеет вид:

$$\{(x,\mu) \ \mid\ (p,x) + \alpha\mu + b \le 0\,\}, \tag{9.9}$$

где $p \in \mathbb{R}^n$, $\alpha \in \mathbb{R}$, $(p,\alpha) \ne (0,0)$. Если такое полупространство содержит E, то (опять благодаря возможности неограниченно увеличивать μ) очевидно $\alpha \le 0$. Если $\alpha < 0$, то можно считать $\alpha = -1$, и тогда полупространство (9.9) имеет вид

$$\{(x,\mu) \ \mid\ (p,x) + b \le \mu\,\}\,. \tag{9.10}$$

Это есть надграфик аффинной функции $l(x) = (p,x) + b$. Назовем такое полупространство *наклонным.* Так как это полупространство содержит E, то $l \le f$, т.е. l — аффинная миноранта. Обратно, надграфик любой аффинной миноранты является наклонным полупространством вида (9.10) и содержит E. Согласно лемме 9.3, хотя бы одно такое полупространство для нашей f имеется; оно соответствует некоторой аффинной миноранте $l_0(x)$.

Если же $\alpha = 0$, то вектор $p \ne 0$, и полупространство (9.9) имеет вид

$$\{(x,\mu) \ \mid\ l(x) = (p,x) + b \le 0\,\}. \tag{9.11}$$

Это полупространство цилиндрическое относительно μ, поэтому оно не является надграфиком никакой собственной функции. Назовем такое полупространство *вертикальным.* Неравенство $l(x) \le 0$ выполнено $\forall\,(x,\mu) \in E$, т. е. $\forall\, x \in D = \operatorname{dom} f$.

Итак, множество E есть пересечение всех наклонных полупространств, его содержащих, и всех вертикальных полупространств, его содержащих. Покажем, что вертикальные полупространства можно не включать в это пересечение, т.е. что E есть пересечение только всех наклонных полупространств, его содержащих. Если это верно, то тогда $E = \operatorname{epi} f$ есть пересечение надграфиков всевозможных аффинных минорант функции f :

$$E_f = \bigcap_{l \in \mathcal{L}} E_l \ .$$

Но, как мы знаем, такое равенство эквивалентно тому, что функция f есть супремум этих аффинных функций:

$$f(x) = \sup_{l \in \mathcal{L}} l(x),$$

и тем самым теорема будет доказана. Таким образом, осталось показать, что каждое вертикальное полупространство можно безболезненно исключить из участников пересечения.

Возьмем любое вертикальное полупространство

$$V = \{ (x, \mu) \ \mid \ l_1(x) = (p_1, x) + b_1 \le 0 \},$$

содержащее E. Нам достаточно показать, что если некоторая точка $(x_0, \mu_0) \notin V$, то она не принадлежит и некоторому наклонному полупространству. (Почему?)
Т.е. достаточно показать, что если $l_1(x_0) > 0$, то найдется аффинная функция $l(x)$ такая, что $l \le f$ и $(x_0, \mu_0) \notin E_l$, т. е. $l(x_0) > \mu_0$.

Как уже отмечалось, имеется аффинная $l_0 \le f$. Тогда $\forall\, x \in D = \operatorname{dom} f$ выполнены неравенства $l_1(x) \le 0$, $l_0(x) \le f(x)$. Следовательно, $\forall\, \lambda \ge 0$

$$l_0(x) + \lambda l_1(x) \le f(x) \qquad \forall\, x \in D.$$

Но для $x \notin D$ это и подавно верно, ибо слева стоит некоторая конечная величина, а справа $+\infty$. Таким образом, $\forall\, x \in \mathbb{R}^n$

$$l(x) = l_0(x) + \lambda l_1(x) \le f(x),$$

т.е. l есть аффинная миноранта для f. Кроме того, так как $l_1(x_0) > 0$, то при достаточно большом $\lambda > 0$ получим $l(x_0) = l_0(x_0) + \lambda l_1(x_0) > \mu_0$, что и требовалось. Теорема доказана. $\square$

Следствие. Пусть $f -$ собственная выпуклая (но, вообще говоря, не замкнутая) функция, и $\mathcal{L} -$ семейство ее аффинных минорант. Тогда

$$\sup_{l \in \mathcal{L}} l(x) \ = \ \overline{f}(x). \tag{9.12}$$

Доказательство. По лемме 9.5 функция $\overline{f} -$ собственная выпуклая и замкнутая. Пусть $\overline{\mathcal{L}}$ есть множество всех ее аффинных минорант. Согласно только что доказанной теореме,

$$\sup_{l \in \overline{\mathcal{L}}} l(x) \ = \ \overline{f}(x). \tag{9.13}$$

Но по лемме 9.4 $\overline{\mathcal{L}} = \mathcal{L}$, и тогда (9.13) превращается в (9.12), ч.т.д. $\square$

Установим также некоторое уточнение теоремы 9.2, состоящие в том что супремум можно брать не по всем аффинным минорантам, а лишь по "максимальным", которые определяются следующим образом.

Пусть $f : \mathbb{R}^n \to \mathbb{R}^\bullet$ есть произвольная собственная функция.

Определение 4. Аффинная функция $l(x) = (p, x) - b$ называется *опорной к* f, если

а) она есть миноранта: $\forall x \quad l(x) \le f(x)$,

б) $\forall \varepsilon > 0$ функция $l(x) + \varepsilon$ уже не является минорантой: $\exists x$, для которого $l(x) + \varepsilon > f(x)$.

Нетрудно сообразить, что для опорной функции $b = \sup\limits_x \{(p, x) - f(x)\}$. Обратим внимание, что даже для выпуклой f опорная функция не обязательно совпадает с f в некоторой точке; она может быть всюду строго меньше f. (Например, $l(x) \equiv 0$ опорна к $f(x) = e^{-x}$, $x \in \mathbb{R}$.)

Лемма 9.6. Пусть $l \le f$. Тогда $\exists c \ge 0$ такое, что $l + c$ опорна к f .

Доказательство. Положим $c = \sup\{c' \mid l(x) + c' \le f(x) \ \ \forall x\}$. Так как множество, стоящее в скобках, содержит $c' = 0$, то оно непусто и его $\sup \ge 0$. Так как f собственная, то для некоторого x_0 значение $f(x_0)$ конечно, поэтому все $c' \le \mathrm{const}$, и следовательно, $c < +\infty$. Наконец, переходя в неравенстве $l(x) + c' \le f(x)$ к пределу при $c' \to c$, получим $l(x) + c \le f(x)$, т.е. $l(x) + c$ по-прежнему есть миноранта. Максимальность c ясна из его определения, поэтому $l(x) + c$ опорна к $f(x)$. $\square$

Таким образом, к любой аффинной миноранте можно добавить неотрицательную константу (как бы "поднять" ее), так что она превратится в опорную.

Из определения 4 следует, что если две функции f_1 и f_2 имеют одно и то же множество аффинных минорант, то и множество опорных у них одно и то же. (Покажите!) Из этого соображения и леммы 9.4 вытекает

Лемма 9.7. Аффинная функция l опорна к $f \iff l$ опорна к $\overline{f}$.

Вернемся к выпуклым функциям. Следующая теорема уточняет теорему 9.2.

Теорема 9.3 (Г. Минковский). Пусть f — собственная выпуклая замкнутая функция, и $\mathcal{L}_0$ есть множество всех ее опорных. Тогда $\forall x$

$$f(x) \ = \ \sup_{l \in \mathcal{L}_0} l(x). \tag{9.14}$$

Таким образом, всякая собственная выпуклая замкнутая есть супремум своих опорных.

Доказательство. По теореме 9.2

$$f(x) = \sup_{l \in \mathcal{L}} l(x). \tag{9.15}$$

По лемме 9.6 для каждой $l \in \mathcal{L}$ имеется $c_l \ge 0$, такое, что $l + c_l \in \mathcal{L}_0$. Тогда, с одной стороны,

$$\sup_{l \in \mathcal{L}} l(x) \ \le \ \sup_{l \in \mathcal{L}} \{l(x) + c_l\},$$

а с другой —

$$\sup_{l\in\mathcal{L}} \{l(x)+c_l\} \ \le\ \sup_{l\in\mathcal{L}_0} l(x),$$

поскольку каждая функция из левой части принадлежит $\mathcal{L}_0$.

Наконец, неравенство $\sup\limits_{\mathcal{L}_0} l(x) \le \sup\limits_{\mathcal{L}} l(x)$ очевидно в силу включения $\mathcal{L}_0 \subset \mathcal{L}$. Из полученных неравенств следует, что $\sup\limits_{\mathcal{L}} l(x) = \sup\limits_{\mathcal{L}_0} l(x)$, а тогда из (9.15) получаем (9.14), ч.т.д. $\square$

Следствие. Пусть f — собственная выпуклая функция, и $\mathcal{L}_0 = \mathcal{L}_0(f)$ есть множество всех ее опорных. Тогда $\forall\, x$

$$\sup_{l\in\mathcal{L}_0} l(x) \ =\ \overline{f}(x)\,.$$

Для доказательства надо применить теорему 9.3 к функции $\overline{f}$ и учесть, что по лемме 9.7 ее множество опорных $\mathcal{L}_0(\overline{f}) = \mathcal{L}_0(f)$.

Упражнения. 1) Найти все опорные к следующим функциям:

$$f = |x|,\ \ |x|^2 \ \text{ в } \ \mathbb{R}^n, \qquad f = e^x,\ \ \sqrt{1+x^2} \ \text{ в } \ \mathbb{R}^1.$$

2) Доказать, что если l опорна к f, то $\forall\,\alpha > 0$, $\forall\, c$ аффинная функция $\alpha l + c$ опорна к $\alpha f + c$.

Лекция 10. Субдифференциал выпуклой функции в точке

Как мы уже видели на примерах, выпуклая функция не обязательно дифференцируема в каждой точке. Тем не менее, она обладает некоторыми свойствами, близкими к дифференцируемости. Перейдем к их рассмотрению.

Пусть f – собственная выпуклая функция, $D = \operatorname{dom} f$.

Определение 10.1. Аффинная функция $l(x) = (p, x) + b$ называется опорной к f в точке $x_0 \in D$, если

$$\forall\, x \in \mathbb{R}^n \quad l(x) \le f(x), \ \text{ и } \ l(x_0) = f(x_0). \tag{10.1}$$

(Ясно, что первое неравенство достаточно проверять лишь для $x \in D$, поскольку $f(x) = +\infty$ вне D, и тогда это неравенство выполнено очевидным образом.)

Определение 10.2. Вектор $p \in \mathbb{R}^n$ называется субградиентом выпуклой функции f в точке $x_0 \in D$, если $\forall\, x \in \mathbb{R}^n$

$$(p,\, x - x_0) \ \le\ f(x) - f(x_0). \tag{10.2}$$

Другими словами, если существует число b, с которым аффинная функция $l = (p, x) + b$ является опорной к f в точке x_0 , поскольку (10.2) эквивалентно тому, что

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

При этом требуемое число $b = -px_0 + f(x_0)$.

Множество всех субградиентов в точке x_0 называется субдифференциалом функции f в точке x_0 и обозначается $\partial f(x_0)$.

Установим некоторые свойства субдифференциала. Легко видеть, что всегда $\partial f(x_0)$ есть выпуклое замкнутое множество (может быть, пустое). Однако можно утверждать и больше.

Лемма 10.1. $\forall\, x_0 \in \operatorname{int} D$ множество $\partial f(x_0)$ есть непустой выпуклый компакт.

Доказательство. Как уже было сказано, выпуклость и замкнутость очевидна. Докажем ограниченность. Пусть $\varepsilon > 0$ таково, что шар $B_\varepsilon(x_0) \subset int\, D$. По теореме 8.2 функция f липшицева на этом шаре с некоторой константой L, поэтому $\forall\, p \in \partial f(x_0)$, $\forall\, \bar{x} \in B_\varepsilon(0)$ имеем

$$(p, \bar{x}) \ \le\ f(x_0 + \bar{x}) - f(x_0) \le L|\bar{x}|,$$

и точно так же $(p, -\bar{x}) \le L|\bar{x}|$, поэтому $|(p, \bar{x})| \le L|\bar{x}|$. Отсюда $|p\,| \le L$. Таким образом, все векторы из $\partial f(x_0)$ ограничены по модулю числом L .

Докажем, что $\partial f(x_0)$ непусто. Рассмотрим в пространстве $\mathbb{R}^{n+1}$ множество $E = \operatorname{epi} f$. Точка $(x_0,\, \mu_0 = f(x_0))$ является в нем граничной (почему?), поэтому существует ненулевой вектор $(p,\alpha) \in \mathbb{R}^{n+1}$, опорный в этой точке к множеству E : $\forall\, (x,\mu) \in E$

$$px + \alpha\mu \;\le\; px_0 + \alpha\mu_0\,. \tag{10.3}$$

Ясно, что здесь $\alpha \le 0$ (иначе при $\mu \to \infty$ получим противоречие). Если $\alpha = 0$, то $p \ne 0$, и тогда $\forall\, x \in D$ имеем $px \le px_0$, а это противоречит условию $x_0 \in \operatorname{int} D$.

Итак, $\alpha < 0$, и тогда можно считать $\alpha = -1$. Тогда из (10.3) при $\mu = f(x)$ получаем: $\forall\, x \in D$

$$px - f(x) \;\le\; px_0 - f(x_0),$$

т.е. $p(x - x_0) \le f(x) - f(x_0)$, а это и означает, что $p \in \partial f(x_0)$, и следовательно, $\partial f(x_0)$ непусто. $\square$

Посмотрим, что такое субдифференциал функции f в точке x_0 в случае, когда она имеет в этой точке обычный дифференциал (то есть градиент).

Лемма 10.2. Пусть $f : D \to \mathbb{R}$ выпукла и в точке $x_0 \in \operatorname{int} D$ дифференцируема. Тогда ее субдифференциал в этой точке состоит из единственного вектора $p = f'(x_0)$:

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

Доказательство. Так как нам надо доказать совпадение двух множеств, покажем, что каждое из них содержится в другом.

($\subset$) Пусть $p \in \partial f(x_0)$. Тогда $\forall\, \overline{x} \in \mathbb{R}^n$ при $\varepsilon \to 0+$ имеем

$$\varepsilon(p, \overline{x}) \le f(x_0 + \varepsilon\overline{x}) - f(x_0) = \varepsilon f'(x_0)\,\overline{x} + o(\varepsilon),$$

откуда $p\,\overline{x} \le f'(x_0)\,\overline{x} + o(1)$, и в пределе $p\,\overline{x} \le f'(x_0)\,\overline{x}$. Это неравенство остается верным при замене $\overline{x} \mapsto -\overline{x}$, поэтому на самом деле имеет место равенство $p\,\overline{x} = f'(x_0)\,\overline{x} \quad \forall\, \overline{x} \in \mathbb{R}^n$. А оно означает, что $p = f'(x_0)$.

($\supset$) Так как f выпукла, то по критерию Вейерштрасса $\forall\, x$

$$\mathcal{E}(x, x_0) = f(x) - f(x_0) - f'(x_0)(x - x_0) \ge 0,$$

а это и означает выполнение (10.2) для вектора $p = f'(x_0)$, поэтому $f'(x_0) \in \partial f(x_0)$. Лемма доказана. $\square$

Итак, для дифференцируемой выпуклой функции субдифференциал сводится к обычному дифференциалу (градиенту): имеется всего лишь один субградиент, который совпадает с градиентом. В частности, для аффинной функции $l(x) = (a, x) + b$ будет $\partial l(x) = \{a\}$ при всех $x \in \mathbb{R}^n$.

Справедливо также и обратное утверждение к лемме 10.2.

Теорема 10.1. Пусть $f : D \to \mathbb{R}$ выпукла, и в некоторой точке $x_0 \in \operatorname{int} D$ ее субдифференциал состоит из единственного вектора: $\partial f(x_0) = \{p_0\}$. Тогда f дифференцируема в точке x_0 и $f'(x_0) = p_0$.

Мы докажем эту теорему несколько позже, в лекции 12.

Таким образом, выпуклая функция дифференцируема во внутренней точке своей области определения тогда и только тогда, когда ее субдифференциал в этой точке состоит из единственного элемента.

Отметим также следующий факт. Как мы знаем, выпуклая функция не обязательно дифференцируема в произвольной точке $x \in \operatorname{int} D$. Спрашивается, много ли таких точек, где нет производной $f'(x)$, и обязательно ли существуют точки, в которых производная есть? Ответ на этот вопрос дает следующая замечательная

Теорема 10.2 (Андерсон, Кли). Пусть функция $f : D \to \mathbb{R}$ выпукла. Тогда почти всюду на $\operatorname{int} D$ она имеет производную. Т.е. множество тех $x \in \operatorname{int} D$, в которых $f'(x)$ не существует, имеет лебегову меру 0.

Доказательство довольно сложно, поэтому здесь мы его не приводим, отсылая интересующихся к книге [R].

Следующее свойство вытекает непосредственно из определения 10.2. (Проверьте!).

Лемма 10.3. Пусть f – собственная выпуклая функция. Тогда при добавлении к ней любой аффинной функции $l(x) = ax + b$ во всех точках $x \in \mathbb{R}^n$ справедливо равенство

$$\partial(f + l)(x) \;=\; \partial f(x) + a,$$

т.е. субдифференциал сдвигается на вектор a и не зависит от константы b.
(Если одна из частей этого равенства пуста, то пуста и другая.)

Лемма 10.4. Пусть $f,\ g$ – собственные выпуклые функции, $\mathcal{U}$ – открытое множество, и на нем $f \equiv g$. Тогда $\forall\, x \in \mathcal{U} \qquad \partial f(x) = \partial g(x)$.

Доказательство. Возьмем любую точку $x_0 \in \mathcal{U}$. Достаточно показать, что если $p \in \partial f(x_0)$, то $p \in \partial g(x_0)$. Без нарушения общности считаем $f(x_0) = g(x_0) = 0$, а с учетом леммы 10.3 можно также считать $p = 0$. Тогда имеем

$$f(x) \;\ge\; 0 \qquad \forall\, x \in \mathbb{R}^n. \tag{10.4}$$

Покажем, что это выполнено и для g. Допустим, это не так: $\exists\, x_1$, для которого $g(x_1) < 0$. Тогда рассмотрим точку $x_\varepsilon = \varepsilon x_1 + (1 - \varepsilon) x_0$, $0 < \varepsilon < 1$. Для нее в силу выпуклости $g(x_\varepsilon) \;\le\; \varepsilon g(x_1) < 0$. Но при малом $\varepsilon > 0$ имеем $x_\varepsilon \in \mathcal{U}$, поэтому $f(x_\varepsilon) = g(x_\varepsilon) < 0$, что противоречит неравенству (10.4). $\square$

Из доказанной леммы вытекает, что $\partial f(x)$ зависит только от значений функции f в произвольно малой окрестности данной точки x. Это свойство "локальности" субдифференциала аналогично свойству локальности обычной производной для дифференцируемых функций.

Задача. Пусть f, g — собственные выпуклые функции, $\mathcal{U}$ — открытое выпуклое множество, и на нем $\partial f(x) = \partial g(x)$. Доказать, что тогда на этом множестве $f(x) = g(x) + \mathrm{const}$. Существенна ли здесь выпуклость $\mathcal{U}$?

Рассмотрим два примера нахождения субдифференциала.

Субдифференциал индикаторной функции. Пусть A — выпуклое множество, и δ_A — его индикаторная функция. Вычислим её субдифференциал в произвольной точке $x_0 \in A$. По определению

$$p \in \partial\delta_a(x_0) \iff \forall\, x \quad (p, x - x_0) \;\le\; \delta_A(x) - \delta_A(x_0),$$

и, как легко заметить, последнее неравенство достаточно проверять только для всех $x \in A$. (Почему?) Таким образом, получаем, что

$$\partial\delta_a(x_0) \;=\; \{\, p \;\mid\; (p, x - x_0) \;\le\; 0 \quad \forall\, x \in A \,\},$$

а это по определению есть конус внешних нормалей к множеству A в точке x_0. Итак,

$$\partial\delta_a(x_0) \;=\; N_A(x_0)\,.$$

Напомним, что, в свою очередь, $N_A(x_0) \;=\; -K_A^*(x_0)$, где $K_A(x_0)$ есть касательный конус к множеству A в точке x_0. Если $x_0 \in int\, A$, то очевидно $K_A(x_0) = \mathbb{R}^n$ и $N_A(x_0) = \{0\}$. Если же $x_0 \in \partial\, A$, то, как мы знаем, $N_A(x_0)$ содержит ненулевые элементы.

Субдифференциал функции Минковского. Пусть $\mu_A(x)$ есть функция Минковского выпуклого множества A, содержащего внутри себя ноль. Тогда $p \in \partial\mu_A(0)$ означает $px \le \mu_A(x)\;\; \forall\, x$, т.е. из $\mu_A(x) = 1$ должно следовать $px \le 1$. Это эквивалентно требованию $(p, A) \le 1$ (покажите!), т.е. $p \in A^0$. Итак, $\partial\mu_A(0) = A^0$.

Случай $x_0 \ne 0$ будет рассмотрен позже.

Следующая теорема позволяет находить субдифференциал суммы выпуклых функций.

Теорема 10.3 (Моро–Рокафеллар). Пусть $F \;=\; f_1 + f_2$, где f_1, f_2 – собственные выпуклые функции, и $\exists\, \hat{x} \in (\mathrm{int}\, D_1) \bigcap D_2$. Тогда $\forall\, x \in D_1 \bigcap D_2$

$$\partial F(x) \;=\; \partial f_1(x) + \partial f_2(x).$$

(Существование указанной точки $\hat{x}$ играет здесь роль условия регулярности.)

Доказательство. ($\supset$, тривиальная часть) Пусть $q = p_1 + p_2$, где $p_1 \in \partial f_1(x)$, $p \in \partial f_2(x)$. Тогда $\forall\, x'$ имеем

$$(q, x'-x) = (p_1, x'-x) + (p_2, x'-x) \le f_1(x') - f_1(x) + f_2(x') - f_2(x) = F(x') - F(x),$$

следовательно, $q \in \partial F(x)$.

($\subset$) Без нарушения общности считаем $x = 0$, $f_1(0) = f_2(0) = 0$. (Почему?) Пусть $q \in \partial F(0)$. Надо показать, что имеется представление $q = p_1 + p_2$, где $p_1 \in \partial f_1(0)$, $p_2 \in \partial f_2(0)$. Считаем $q = 0$ (почему?), и таким образом,

$$F(x) \ge 0 \qquad \forall\, x \in \mathbb{R}^n. \tag{10.5}$$

Надо показать, что $0 = p_1 + p_2$, то есть что $\exists\, p \in \partial f_1(0)$ такой, что $-p \in \partial f_2(0)$.

Рассмотрим в пространстве $\mathbb{R}^{n+1}$ два множества:

$$C_1 = \ \{(x,\mu) \,|\ \ \mu > f_1(x), \ \ x \in \operatorname{int} D_1\}\,,$$

$$C_2 = \ \{(x,\mu) \,|\ \ \mu \le -f_2(x), \ \ x \in D_2\,\}\,.$$

Ясно, что $C_1 = \operatorname{int} E_1$, а C_2 есть зеркальное отражение множества $E_2 = \operatorname{epi} f_2$ относительно "горизонтальной" плоскости x. Множества C_1, C_2 оба выпуклы, непусты, и C_1 открыто. Мы утверждаем, что они не пересекаются. (Действительно, если существует точка $(x,\mu) \in C_1 \cap C_2$, то по определению $f_1(x) < \mu$, $f_2(x) \le -\mu$, и тогда $F(x) = f_1(x) + f_2(x) < 0$, что противоречит (10.5).)

Тогда по теореме отделимости существует ненулевой вектор $(p,\alpha) \in \mathbb{R}^{n+1}$ такой, что

$$\sup_{C_1} (p\,x - \alpha\mu) \ \le \ \inf_{C_2} (p\,x - \alpha\mu). \tag{10.6}$$

Ясно, что здесь $\alpha \ge 0$. (Если $\alpha < 0$, то, увеличивая μ в левой части, получим нарушение этого неравенства.)

Если $\alpha = 0$, то $p \ne 0$, и тогда (10.6) означает, что

$$\sup_{\operatorname{int} D_1} \ p\,x \ \le \ \inf_{D_2} p\,x\,.$$

Возьмем точку $\hat{x} \in (\operatorname{int} D_1) \bigcap D_2$, которая существует по условию теоремы. Тогда $\inf\limits_{D_2} p\,x \le p\,\hat{x}$, и отсюда

$$\sup_{\operatorname{int} D_1} \ p\,x \ \le \ p\,\hat{x},$$

т.е. линейный функционал $p \ne 0$ достигает своего максимума на открытом множестве $\operatorname{int} D_1$ в некоторой точке этого множества, а такого не может быть. Противоречие.

Следовательно, $\alpha > 0$, и поэтому считаем $\alpha = 1$. Тогда (10.6) означает, что $\forall\,(x_1,\mu_1) \in C_1$, $\forall\,(x_2,\mu_2) \in C_2$

$$p\,x_1 - \mu_1 \ \le \ p\,x_2 - \mu_2.$$

В силу непрерывности левой части этого неравенства относительно (x_1, μ_1), оно справедливо и $\forall\,(x_1, \mu_1) \in \overline{C}_1$, а поскольку $C_1 = \operatorname{int} E_1$, то $\overline{C}_1 \supset E_1$, поэтому $\forall\, x_1 \in D_1$, $\forall\, x_2 \in D_2$, полагая $\mu_1 = f_1(x_1)$, $\mu_2 = -f_2(x_2)$, имеем

$$p\,x_1 - f_1(x_1) \;\le\; p\,x_2 + f_2(x_2).$$

Полагая здесь $x_2 = 0$, получаем $\forall\, x_1 \in D_1$ неравенство $px_1 \le f_1(x_1)$, т.е. $p \in \partial f_1(0)$, а полагая $x_1 = 0$, получаем $\forall\, x_2 \in D_2$ неравенство $-px_2 \le f_2(x_2)$, т.е. $-p \in \partial f_2(0)$, ч. т. д. $\square$

Вопрос о нахождении субдифференциала максимума выпуклых функций будет рассмотрен в следующих лекциях.

Лекция 11. Субдифференциал сублинейной функции

Напомним, что собственная функция $\varphi : \mathbb{R}^n \to \mathbb{R}^\bullet$ является сублинейной $\Longleftrightarrow$ она положительно однородна и субаддитивна (см. лекцию 7). Далее мы будем рассматривать только замкнутые сублинейные функции. Нетрудно показать, что тогда $0 \in dom\,\varphi$ и $\varphi(0) = 0$. (Как мы знаем, для собственной сублинейной функции имеется альтернатива: $\varphi(0) = 0$ или $\varphi(0) = +\infty$. Так как $epi\,\varphi$ есть непустой замкнутый конус, то $(0,0) \in\ epi\,\varphi$, поэтому $\varphi(0) \le 0$, а тогда остается лишь вариант $\varphi(0) = 0$.)

Для произвольной выпуклой функции было введено понятие опорной аффинной функции (лекция 9). Посмотрим, что оно означает для сублинейных функций.

Лемма 11.1. Аффинная функция $l(x) = px + b$ опорна к сублинейной $\varphi \Longleftrightarrow$ $b = 0$ и $px \le \varphi(x)$ для любого x .

Доказательство. ($\Longleftarrow$) Если $b = 0$ и $px \le \varphi(x)$ для любого x, то функция $l(x) = px$ есть миноранта к φ. При любом $b > 0$ функция $l(x) = px + b$ уже не будет минорантой т.к. $l(0) = b > \varphi(0) = 0$. Отсюда по определению следует, что $l(x) = px$ есть опорная к φ.

($\Longrightarrow$) Пусть $l(x) = px + b$ опорна к φ. Для любого $z \in \mathbb{R}^n$ и любого $\lambda > 0$ при $x = \lambda z$ имеем

$$(p, \lambda z) + b \le \lambda \varphi(z),$$

следовательно

$$(p, z) + \frac{b}{\lambda} \le \varphi(z),$$

откуда при $\lambda \to +\infty$ в пределе получаем $(p,z) \le \varphi(z)$, т.е. функция (p,z) есть миноранта для $\varphi(z)$. Далее, так как $l(x) = (p,x) + b \le \varphi(x)$, то при $x = 0$ имеем $b \le 0$. Если $b < 0$, то функция $l(x) + |b| = (p,x)$, как только что установлено, является минорантой к φ, а это противоречит опорности l. Следовательно, $b = 0$.
$\square$

Таким образом, все аффинные опорные к сублинейной функции на самом деле просто линейные: $l(x) = (p,x)$; необходимость добавления константы $b = 0$ отпадает. Поэтому можно дать следующее

Определение 11.1. Линейный функционал (вектор) $p \in \mathbb{R}^n$ называется опорным к сублинейной функции φ, если для любого $x \in \mathbb{R}^n$ $\quad (p,x) \le \varphi(x)$.

Множество всех опорных векторов обозначается $\partial\varphi$. Итак,

$$\partial\varphi = \ \{p \mid \ (p,x) \le \varphi(x) \ \text{ для любого } \ x \in \mathbb{R}^n \} \ .$$

Так как у любой собственной выпуклой функции φ имеется хотя бы одна аффинная опорная (см. лекцию 9), то у любой собственной сублинейной функции множество $\partial\varphi$ непусто. Более того, оно всегда выпукло и замкнуто, ибо является пересечением семейства замкнутых полупространств:

$$\partial\varphi = \bigcap_{x\in\mathbb{R}^n} \{p \mid (p,x) \le \varphi(x)\}.$$

Далее, если функция φ замкнута (а здесь мы рассматриваем только такие), то по теореме Минковского она восстанавливается по своим опорным: для любого $x \in \mathbb{R}^n$

$$\varphi(x) = \sup_{p\in\partial\varphi} (p,x). \tag{11.1}$$

Таким образом, любая замкнутая сублинейная функция имеет вид

$$\varphi(x) = \sup_{p\in A} (p,x),$$

где A некоторое выпуклое замкнутое множество. Покажем, что для данной функции φ такое множество A единственно.

Лемма 11.2. Пусть выпуклое замкнутое множество A таково, что $\forall\, x \in \mathbb{R}^n$

$$\varphi(x) = \sup_{a\in A} (a,x).$$

Тогда $\partial\varphi = A$.

Доказательство. Из условия следует, что для любого $a \in A$ справедливо неравенство $(a,x) \le \varphi(x)$ $\forall\, x$, поэтому $a \in \partial\varphi$. Таким образом, $A \subset \partial\varphi$.

Для доказательства обратного включения покажем, что если $\hat{a} \notin A$, то $\hat{a} \notin \partial\varphi$. Действительно, если $\hat{a} \notin A$, то по теореме о строгой отделимости точки от выпуклого замкнутого множества найдется вектор $\hat{x} \in \mathbb{R}^n$ такой, что

$$(\hat{a},\hat{x}) > \sup_{a\in A} (a,\hat{x}) = \varphi(\hat{x}),$$

а из этого вытекает, что $\hat{a} \notin \partial\varphi$. $\square$

Лемма 11.3. Сублинейная замкнутая функция $\varphi : \mathbb{R}^n \to \mathbb{R}^\bullet$ положительна вне нуля тогда и только тогда, когда $0 \in int\,\partial\varphi$.

Доказательство. $(\Longrightarrow)$ Пусть $\varphi(x) > 0$ при $x \ne 0$. Так как φ замкнута, то она полунепрерывна снизу (см. лекцию 9), а так как $\varphi > 0$ на единичной сфере, то на ней $\varphi(x) \ge r > 0$. Тогда для любого $x \in \mathbb{R}^n$ имеется оценка $\varphi(x) \ge r|x|$. (Почему?) Из этой оценки следует, что шар $B_r(0)$ содержится в $\partial\varphi$, ибо если $|p| \le r$, то для любого x

$$(p,x) \le |p|\cdot|x| \le r|x| \le \varphi(x).$$

($\Longleftarrow$) Пусть при некотором $r > 0$ шар $B_r(0) \subset \partial\varphi$. Тогда для любого x

$$\varphi(x) = \sup_{p \in \partial\varphi} (p, x) \;\geq\; \sup_{|p| \leq r} (p, x) \;=\; r|x|,$$

и следовательно, $\varphi(x) > 0$ при $x \neq 0$. □

Наибольший интерес представляют сублинейные функции, определенные на всем пространстве. Как мы уже знаем (см. лекцию 8), это те сублинейные функции, которые ограничены на единичном шаре – они называются ограниченными сублинейными функциями.

Лемма 11.4. Сублинейная функция φ является ограниченной тогда и только тогда, когда ее субдифференциал $\partial\varphi$ есть ограниченное множество (и значит, он является выпуклым компактом).

Доказательство вытекает из того факта, что ограниченность сублинейной функции по определению означает, что при некотором C выполняется оценка

$$|\varphi(x)| \leq C|x| \qquad \forall\, x \in \mathbb{R}^n,$$

а она эквивалентна тому, что $|\partial\varphi| \leq C$. Детали оставляем читателю. □

Обратимся теперь к понятию субдифференциала в точке и посмотрим, каким будет $\partial\varphi(x_0)$ для сублинейной функции φ. Если $x_0 = 0$, то по определению

$$p \in \partial\varphi(0) \iff px \leq \varphi(x) \iff p \in \partial\varphi.$$

Таким образом $\partial\varphi(0) = \partial\varphi$. Буква ∂ употребляется здесь в двух разных смыслах: как субдифференциал сублинейной φ "вообще", и как субдифференциал выпуклой функции φ в данной точке. Это не совсем корректно, но так уж принято. Из контекста всегда ясно, какой субдифференциал имеется в виду.

Каким будет $\partial\varphi(x_0)$ для произвольной точки $x_0 \in \mathbb{R}^n$?

Теорема 11.1.

$$\partial\varphi(x_0) = \{p \in \partial\varphi \mid p\, x_0 = \varphi(x_0)\}$$

– это те элементы из $\partial\varphi$, значения которых в данной точке x_0 совпадают с $\varphi(x_0)$.

Доказательство. Случай $x_0 = 0$ рассмотрен выше, поэтому считаем $x_0 \neq 0$.

($\supset$) Пусть $p \in \partial\varphi$ и $p\, x_0 = \varphi(x_0)$. Тогда для любого x имеем $p\, x \leq \varphi(x)$. Вычитая отсюда предыдущее равенство, получим

$$p(x - x_0) \leq \varphi(x) - \varphi(x_0) \quad \forall\, x \in \mathbb{R}^n, \tag{11.2}$$

а это и означает, что $p \in \partial\varphi(x_0)$.

($\subset$) Пусть $p \in \partial\varphi(x_0)$, то есть выполнено (11.2). Возьмем произвольный $\bar{x} \in \mathbb{R}^n$ и положим $x = x_0 + \bar{x}$. Тогда из (11.2) получаем $(p, \bar{x}) \leq \varphi(x_0 + \bar{x}) - \varphi(x_0)$.

Но в силу сублинейности φ имеем $\varphi(x_0+\bar{x}) \le \varphi(x_0)+\varphi(\bar{x})$, поэтому $(p,\bar{x}) \le \varphi(\bar{x})$. Так как $\bar{x} \in \mathbb{R}^n$ произвольно, то это означает, что $p \in \partial\varphi$.

Положим теперь в (11.2) $x = \alpha\, x_0$, где $\alpha > 0$. Тогда

$$(\alpha-1)(p,x_0) \le (\alpha-1)\varphi(x_0).$$

При $\alpha > 1$ отсюда получаем $(p,x_0) \le \varphi(x_0)$, а при $\alpha < 1$ получаем $(p,x_0) \ge \varphi(x_0)$, и значит $(p,x_0) = \varphi(x_0)$. $\square$

В качестве следствия равенства $\partial\varphi(0) = \partial\varphi$ нетрудно получить формулу для субдифференциала суммы двух ограниченных сублинейных функций:

$$\partial(\varphi_1+\varphi_2) \;=\; \partial\varphi_1 + \partial\varphi_2\,.$$

Действительно, пользуясь указанным равенством и теоремой Моро–Рокафеллара, получаем

$$\partial(\varphi_1+\varphi_2) \;=\; \partial(\varphi_1+\varphi_2)(0) \;=\; \partial\varphi_1(0)+\partial\varphi_2(0) \;=\; \partial\varphi_1+\partial\varphi_2\,.$$

Примеры.

1) $\varphi(x) = |x|$, $x \in \mathbb{R}$. Легко видеть, что здесь

$$(\,p\,x \le |x| \;\text{ для любого } x\,) \iff |p| \le 1. \tag{11.3}$$

Таким образом, $\partial\varphi = [-1,1]$. Тогда для произвольного $x \in \mathbb{R}^n$ получаем

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

Это многозначное отображение удобно обозначить как $Sign\, x$. При $x \ne 0$ оно совпадает с общепринятой однозначной функцией $sign\, x$, но при $x = 0$ они отличаются: $sign\, 0 = 0$, тогда как $Sign\ 0 = [-1,1]$. Отображение $Sign\, x$ часто встречается в задачах оптимального управления при нахождении управления из условия максимума функции Понтрягина (и отнюдь не всегда его можно заменить однозначной функцией $sign\, x$!).

2) $\varphi(x) = |x| = \sqrt{x_1^2+...+x_n^2}\,$, $x \in \mathbb{R}^n$. Нетрудно сообразить, что здесь, как и раньше, справедливо утверждение (11.3), поэтому $\partial\,\varphi = B_1(0) -$ единичный шар. А поскольку вне нуля φ дифференцируема, то получаем формулу

$$\partial\varphi(x) = \begin{cases} x/|x|\,, & x \ne 0,\\ B_1(0), & x = 0. \end{cases}$$

Это многозначное отображение также удобно обозначить как $Sign\,x$ ("векторный сигнум" векторного аргумента). Для данной точки x оно дает единичный вектор в направлении x. При $x = 0$ направление не определено; в качестве него принимается весь единичный шар.

3) $\varphi(x) = \max\{x_1, x_2\}, \quad x \in \mathbb{R}^2$. Покажите, что здесь

$$\partial\varphi(x) = \begin{cases} (1,0), & \text{если } x_1 > x_2, \\ (0,1), & \text{если } x_1 < x_2, \\ \text{отрезок } \{(p_1,p_2)\,|\,p_1 \geq 0,\ p_2 \geq 0,\ p_1 + p_2 = 1\}, \\ \qquad\qquad \text{если } x_1 = x_2. \end{cases}$$

Следующий пример заслуживает особого расмотрения.

4) Пусть $\varphi : \mathbb{R}^n \to \mathbb{R}$ есть ограниченая сублинейная функция, положительная вне нуля. Из лемм 11.3 и 11.4 следует, что тогда множество $A = \partial\varphi$ есть выпуклый компакт и $0 \in int\,A$. При этом

$$\varphi(x) = \max_{y \in A}\,(y, x) \qquad \forall\, x \in \mathbb{R}^n. \tag{11.4}$$

Положим $B = \{x\,|\,\varphi(x) \leq 1\}$. Это выпуклое замкнутое множество. Из ограниченности φ следует, что $0 \in int\,B$, а из положительности φ вне нуля — что B ограничено (покажите!), следовательно B есть выпуклый компакт. Определим сублинейную функцию

$$\psi(y) = \max_{x \in B}\,(y, x), \qquad y \in \mathbb{R}^n. \tag{11.5}$$

Оказывается, множества A, B и функции φ, ψ являются двойственными друг к другу в следующем смысле.

Теорема 11.2.

1) $A^0 = B, \quad B^0 = A$,

$$A = \{y\,|\,\psi(y) \leq 1\}, \qquad B = \partial\psi. \tag{11.6}$$

2) Функции Минковского множеств A и B имеют вид:

$$\begin{cases} \mu_A(y) = \psi(y), \\ \mu_B(x) = \varphi(x). \end{cases} \tag{11.7}$$

3) Справедливо обобщенное неравенство Коши-Буняковского: $\forall\, x,\, y \in \mathbb{R}^n$

$$(x, y) \leq \varphi(x)\,\psi(y). \tag{11.8}$$

4) Равенство $(x, y) = \varphi(x)\,\psi(y)$ эквивалентно тому, что $\tilde{x} = x/\varphi(x) \in \partial\psi(y)$, а также тому, что $\tilde{y} = y/\psi(y) \in \partial\varphi(x)$.

Доказательство. 1) $y \in B^0 \iff$ (из $\varphi(x) \leq 1$ следует $(y, x) \leq 1$) $\iff$ $\forall\, x \quad (y, x) \leq \varphi(x)$ (покажите!) $\iff$ $y \in \partial\varphi = A$. Таким образом, $B^0 = A$, а тогда по теореме о второй поляре $A^0 = B$. Равенство $B = \partial\psi$ следует из (11.5).

Далее, $y \in A = B^0$ означает, что $\max(y, B) \leq 1$, то есть $\psi(y) \leq 1$. Пункт 1 полностью доказан.

2) Второе равенство в (11.7) очевидно следует из определения множества B, а тогда первое равенство по аналогии следует из (11.6).

3) Пусть сначала x, y таковы, что $\varphi(x) \leq 1$, $\psi(y) \leq 1$. Тогда $x \in B$, $y \in A$, и поскольку $A^0 = B$, имеем $(x, y) \leq 1$. Отсюда и из положительной однородности функций φ, ψ вытекает, что для произвольных $x, y \in \mathbb{R}^n$ справедливо (11.8) (Покажите!).

4) Пусть $\tilde{y} = y/\psi(y) \in \partial\varphi(x)$. Тогда по теореме 11.1 $(\tilde{y}, x) = \varphi(x)$, т.е. $(y, x) = \psi(y)\,\varphi(x)$, ч.т.д.

Обратно. Пусть для некоторых x, y выполнено равенство $(y, x) = \psi(y)\,\varphi(x)$. Тогда для $\tilde{y} = y/\psi(y)$ имеем

$$(\tilde{y}, x) = \varphi(x). \tag{11.9}$$

Для любого x' в силу (11.8) $(y, x') \leq \psi(y)\,\varphi(x')$, поэтому $(\tilde{y}, x') \leq \varphi(x')$, следовательно, $\tilde{y} \in \partial\varphi$. Отсюда с учетом (11.9) получаем $\tilde{y} \in \partial\varphi(x)$. □

Описанные здесь сублинейные функции φ и ψ, соответствующие взаимно полярным множествам A и B, называются взаимно двойственными (иногда говорят — взаимно сопряженными). В частности, если функция φ является симметричной (то есть $\varphi(-x) = \varphi(x)$ для любого x), то $\varphi(x)$ есть некоторая норма в $\mathbb{R}^n$, а $\psi(y)$ — сопряженная норма; множества B и A суть единичные шары в этих нормах.

Например, если $\varphi(x) = \|x\|_1 = \sum |x_i|$, то $\psi(y) = \|y\|_\infty = \max |y_i|$, и наоборот. Если $\varphi(x) = \|x\|_p = \sqrt[p]{|x_1|^p + ... + |x_n|^p}$, где $1 < p < \infty$, то $\psi(y) = \|y\|_q = \sqrt[q]{|y_1|^q + ... + |y_n|^q}$, где $q \in (1, \infty)$ определяется из равенства $\frac{1}{p} + \frac{1}{q} = 1$. Такие показатели p и q также называются взаимно сопряженными, а неравенство (11.8) в этом случае называется неравенством Гёльдера; при $p = q = 2$ оно превращается в неравенство Коши-Буняковского.

Следующая важная теорема позволяет находить субдифференциал максимума нескольких сублинейных функций.

Теорема 11.3 (Дубовицкий–Милютин). Пусть $\varphi_1, \varphi_2 : \mathbb{R}^n \to \mathbb{R}$ – ограниченные сублинейные функции, и $\Phi(x) = \max\{\varphi_1(x), \varphi_2(x)\}$. (Ясно, что это тоже ограниченная сублинейная функция.) Тогда

$$\partial\,\Phi = \, co\,(\partial\varphi_1 \bigcup \partial\varphi_2).$$

Доказательство. Из условия следует, что множества $A_1 = \partial\varphi_1$ и $A_2 = \partial\varphi_2$ — выпуклые компакты. Тогда по теореме Каратеодори множество $A = co(A_1 \bigcup A_2)$ также есть выпуклый компакт. Для любого $x \in \mathbb{R}^n$ имеем

$$\max_{p \in A}(p, x) = \max_{\alpha \in [0,1],\, p_1 \in A_1,\, p_2 \in A_2} (\alpha\, p_1 x + (1-\alpha) p_2\, x) =$$

$$= \max_{p_1 \in A_1,\, p_2 \in A_2} \max_{\alpha \in [0,1]} (\alpha\, p_1\, x + (1-\alpha) p_2\, x) = \max_{p_1 \in A_1,\, p_2 \in A_2} \max\, [(p_1, x), (p_2, x)] =$$

$$= \max\, [\max_{p_1 \in A_1} (p_1, x), \max_{p_2 \in A_2} (p_2, x)] \ = \ \max\, [\varphi_1(x), \varphi_2(x)] \ = \Phi(x).$$

Отсюда по лемме 11.2 делаем заключение, что $A = \partial \Phi$, ч.т.д. $\square$

Следствие. Пусть $\varphi_1, ..., \varphi_m$ — ограниченные сублинейные функции, и

$$\Phi(x) = \max\, \{\varphi_1(x), ..., \varphi_m(x)\}.$$

Тогда

$$\partial\, \Phi = \ co \left(\partial \varphi_1 \bigcup \ldots \bigcup \partial \varphi_m \right), \tag{11.10}$$

то есть $p \in \partial \Phi$ означает, что $p = \sum_{i=1}^m \alpha_i\, p_i$, где все $p_i \in \partial \varphi_i$, все $\alpha_i \geq 0$ и $\sum_{i=1}^m \alpha_i = 1$.

Доказательство. Индукция по m. Для $m = 2$ доказано выше. Пусть теперь $m > 2$ и формула (11.10) верна для любых $m' < m$. Рассмотрим случай m сублинейных функций. Положим

$$g(x) = \max_{1 \leq i \leq m-1} \varphi_i(x).$$

По предположению индукции

$$\partial g = \ co \bigcup_{1 \leq i \leq m-1} \partial \varphi_i.$$

При этом

$$\Phi(x) = \ \max\, \{g(x),\, \varphi_m(x)\},$$

и по теореме 11.2

$$\partial\, \Phi = \ co \left(\partial g \bigcup \partial \varphi_m \right) \ = \ co \left((co \bigcup_{i \leq m-1} \partial \varphi_i) \bigcup \partial \varphi_m \right) = \ co \left(\bigcup_{i \leq m} \partial \varphi_i \right).$$

$\square$

Чуть позднее мы докажем аналогичную теорему для произвольных выпуклых функций. А сейчас приведем весьма полезное обобщение теоремы 11.3 на случай бесконечного числа сублинейных функций.

Пусть S есть некоторое компактное топологическое пространство, и для каждого $s \in S$ задана ограниченная сублинейная функция $\varphi_s : \mathbb{R}^n \to \mathbb{R}$. Таким образом, имеется семейство сублинейных функций $\varphi_s(x)$, параметризованное элементами s из компакта S. Будем предполагать, что при любом фиксированном x функция $\varphi_s(x)$ полунепрерывно сверху зависит от s. Положим для $x \in \mathbb{R}^n$

$$\Phi(x) = \ \max_{s \in S} \varphi_s(x).$$

Ясно, что максимум здесь достигается (в силу обобщенной теоремы Вейерштрасса). Поэтому для любого $x \in \mathbb{R}^n$ имеем $\Phi(x) < +\infty$, то есть построенная сублинейная функция Φ ограничена, и мы хотим найти ее субдифференциал. В этой ситуации справедлив полный аналог формулы (11.10).

Теорема 11.4.

$$\partial \Phi \; = \; co \left(\bigcup_{s \in S} \partial \varphi_s \right).$$

Доказательство. Так как все φ_s — ограниченные сублинейные функции, то по лемме 11.4 все множества $A_s = \partial \varphi_s$ суть выпуклые компакты. Так как сублинейная функция Φ тоже ограничена, то $\partial\Phi$ — тоже выпуклый компакт. Пусть

$$M = \bigcup_{s \in S} A_s \; .$$

Так как все $\varphi_s \le \Phi$, то все $A_s \subset \partial\Phi$, а тогда и $M \subset \partial\Phi$. Покажем, что M замкнуто. Пусть $p_k \in M$ и $p_k \to p_0$. Тогда для любого k имеем $p_k \in A_{s_k}$ при некотором $s_k \in S$. Переходя к подпоследовательности и учитывая, что S — компакт, считаем, что $s_k \to s_0 \in S$. При этом для любого $x \in \mathbb{R}^n$ имеем по определению:

$$(p_k,\, x) \le \varphi_{s_k}(x), \qquad k = 1, 2, \dots \; .$$

Переходя здесь к пределу и учитывая полунепрерывность сверху функций φ_s относительно s, получаем $(p_0, x) \le \varphi_{s_0}(x)$, что в силу произвольности $x \in \mathbb{R}^n$ означает, что $p_0 \in \partial\varphi_{s_0}$, и следовательно, $p_0 \in M$.

Таким образом, M есть замкнутое подмножество компакта $\partial\Phi$, и поэтому само является компактом. Тогда по теореме Каратеодори множество $A = co\, M$ есть выпуклый компакт. При этом для любого $x \in \mathbb{R}^n$ имеем

$$\Phi(x) = \max_{s\in S} \max_{p \in A_s} (p, x) = \max_{p \in M} (p, x) = \max_{p \in A} (p, x),$$

откуда по лемме 11.2 получаем, что $A = \partial\Phi$. □

Следствие. Из теоремы 11.4 и теоремы Каратеодори вытекает, что любой элемент $p \in \partial\Phi$ представим в виде

$$p = \sum_{i=1}^{m} \alpha_i \, p_i \; ,$$

где $m \le n+1$, все $\alpha_i \ge 0$, $\sum_1^m \alpha_i = 1$, а каждый $p_i \in \partial\varphi_{s_i}$ для некоторого $s_i \in S$. (Это утверждение называется теоремой об очистке.) Верно и обратное: любой вектор p указанного вида принадлежит $\partial\Phi$ (это очевидно).

Сопряженный к конусу, заданному сублинейным неравенством. В качестве приложения понятия субдифференциала рассмотрим следующий вопрос. Пусть φ есть ограниченная сублинейная функция. Рассмотрим множество

$$C = \{x \mid \varphi(x) \le 0\}.$$

Это, очевидно, замкнутый выпуклый конус. (Проверьте!) В теории экстремума часто возникает необходимость нахождения сопряженного к такому конусу.

Теорема 11.5. Предположим, что существует $\hat{x}$, для которого $\varphi(\hat{x}) < 0$. Тогда

$$C^* = \{p = -\alpha\, l \mid \ \alpha \ge 0, \ \ l \in \partial\varphi\},$$

то есть

$$C^* = -\mathbb{R}_+\, \partial\varphi. \tag{11.11}$$

Требование существования указанного здесь $\hat{x}$ называется условием Слейтера. Оно обеспечивает невырожденность сублинейного ограничения $\varphi(x) \le 0$.

Доказательство. $(\supset)$ Если $p = -\alpha l$, где $\alpha \ge 0$, $l \in \partial\varphi$, то для любого $x \in C$ имеем $(l,x) \le \varphi(x) \le 0$, тогда $(p,x) = -\alpha(l,x) \ge 0$, и потому $p \in C^*$.

$(\subset)$ Пусть $p \in C^*$, $p \ne 0$. Надо показать, что $p = -\alpha l$, где $\alpha \ge 0$, $l \in \partial\varphi$. (Для $p = 0$ имеется очевидное представление $p = -0 \cdot l$, где в качестве l можно взять любой элемент из $\partial\varphi$.) По определению

$$\varphi(x) \le 0 \quad \Longrightarrow \quad (p,x) \ge 0. \tag{11.12}$$

В пространстве $\mathbb{R}^{n+1}$ рассмотрим множества

$$\Omega_1 \ = \ \{(x,z) \mid \varphi(x) < z\},$$

$$\Omega_2 = \{(x,z) \mid (p,x) < 0, \ \ z = 0\}.$$

Это выпуклые конусы, оба по условию непустые (из них Ω_1 открыт). Мы утверждаем, что $\Omega_1 \bigcap \Omega_2 = \emptyset$. Действительно, иначе найдется точка $(x,0) \in \Omega_1 \bigcap \Omega_2$, то есть $\varphi(x) < 0$ и $(p,x) < 0$, что противоречит (11.12). По теореме отделимости существует функционал $(l,\beta) \in \mathbb{R}^{n+1}$, который меньше нуля на Ω_1 и больше или равен нуля на Ω_2, то есть

$$\varphi(x) < z \ \ \Rightarrow \ (l,x) + \beta z < 0, \tag{11.13}$$

$$(p,x) < 0 \ \ \Rightarrow \ (l,x) \ge 0. \tag{11.14}$$

Из (11.13) при $x = 0$, $z = 1$ следует, что $\beta < 0$, поэтому можно положить $\beta = -1$. Тогда (11.13) означает:

$$\varphi(x) < z \quad \Rightarrow \quad (l,x) < z.$$

Отсюда следует, что $\forall\, x$ будет $(l,x) \le \varphi(x)$ (покажите!), то есть $l \in \partial\varphi$. Из (11.14) следует, что $l = -\alpha p$ при некотором $\alpha \ge 0$. Если $\alpha = 0$, то $l = 0$, тогда $0 \in \partial\varphi$, то есть всюду $\varphi(x) \geq 0$, а это противоречит существованию $\hat{x}$, для которого $\varphi(\hat{x}) < 0$. Поэтому $\alpha > 0$, и тогда $p = -\frac{1}{\alpha}\, l$, ч.т.д. $\square$

Отметим также следующее полезное свойство.

Лемма 11.5. При выполнении условий теоремы 11.5 внутренность конуса C непуста, причем

$$x \in int\, C \iff \varphi(x) < 0.$$

Доказательство. $(\Longleftarrow)$ Следует из непрерывности φ.

$(\Longrightarrow)$ Пусть $x \in int\, C$. Тогда при достаточно малом $\varepsilon > 0$ точка $x - \varepsilon\hat{x} \in C$, поэтому $\varphi(x - \varepsilon\hat{x}) \leq 0$, и мы имеем $x = (x - \varepsilon\hat{x}) + \varepsilon\hat{x}$, откуда

$$\varphi(x) \leq \varphi(x - \varepsilon\hat{x}) + \varphi(\varepsilon\hat{x}) \leq \varepsilon\varphi(\hat{x}) < 0, \qquad \text{ч.т.д.}$$

$\square$

Лекция 12. Производная выпуклой функции по направлению

Пусть дана произвольная функция $f : \mathbb{R}^n \to \mathbb{R}^\bullet$, и пусть точка x_0 такова, что $f(x_0) < +\infty$.

Определение 12.1. Производной функции f в точке x_0 по направлению $\bar{x} \in \mathbb{R}^n$ называется величина

$$f'(x_0, \bar{x}) \;=\; \lim_{t\to 0+} \frac{f(x_0 + t\bar{x}) - f(x_0)}{t}\,,$$

если этот предел существует (конечный или бесконечный).

Покажем, что для выпуклых функций этот предел всегда существует. Как следует из определения 1, для этого достаточно рассмотреть случай выпуклой функции одной переменной $(g(t) = f(x_0 + t\bar{x}))$. Установим сначала следующее свойство.

Лемма 12.1. Пусть $f : [a, b] \to \mathbb{R}$ есть выпуклая функция одной переменной, и пусть даны точки $t_0 < t < t_1$ на отрезке $[a, b]$. Тогда

$$\frac{f(t) - f(t_0)}{t - t_0} \;\le\; \frac{f(t_1) - f(t_0)}{t_1 - t_0} \;\le\; \frac{f(t_1) - f(t)}{t_1 - t}\,. \tag{12.1}$$

Другими словами, отношение $\frac{f(t)-f(t_0)}{t-t_0}$ не возрастает при $t \searrow t_0$ (и поэтому существует его предел при $t \to t_0+$).

Доказательство. Без нарушения общности считаем $t_0 = 0$, $t_1 = 1$, $t \in (0, 1)$. Тогда $t \;=\; (1 - t)\cdot 0 + t\cdot 1$, поэтому

$$f(t) \;\le\; (1 - t)f(0) + tf(1)\,, \tag{12.2}$$

отсюда $f_t - f_0 \;\le\; t(f_1 - f_0)$, а это и есть первое неравенство в (12.1).

Перепишем теперь (12.2) так:

$$-(1 - t)f_0 \;\le\; tf_1 - f_t$$

и добавим к обеим частям $f_1(1 - t)$. Получим

$$(f_1 - f_0)(1 - t) \;\le\; f_1 - f_t\;,$$

а это и есть второе неравенство в (12.1). □

Для функции одной переменной имеются лишь два неколлинеарных направления: $+1$ и -1; при этом, очевидно,

$$f'(x_0, +1) = f'_{\text{пр}}(x_0), \qquad f'(x_0, -1) = -f'_{\text{лев}}(x_0).$$

Оба этих предела существуют, так как монотонно убывающая функция $f(t_0+\varepsilon)/\varepsilon$ всегда имеет предел, конечный или бесконечный. Если $x_0 = b$ (крайняя правая точка $dom f$), то для всех $x > b$ очевидно $f(x) = +\infty$, и поэтому $f'_{\text{пр}}(x_0) = +\infty$. Если же $x_0 = a$ (крайняя левая точка $dom f$), то может случиться, что $f'_{\text{пр}}(x_0) = -\infty$. Аналогичные свойства справедливы и для $f'_{\text{лев}}(x_0)$.

Пример. $D = [-1, 1]$, $f(x) = -\sqrt{1-x}$. Тогда

$$f'_{\text{лев}}(-1) = -\infty, \qquad f'_{\text{пр}}(-1) = -\infty, \qquad f'_{\text{лев}}(1) = +\infty, \qquad f'_{\text{пр}}(1) = +\infty.$$

Для выпуклой функции многих переменных $f : \mathbb{R}^n \to \mathbb{R}^\bullet$ из леммы 12.1 следует, что она имеет производную по любому направлению $\bar{x} \in \mathbb{R}^n$, так как монотонная функция $(f(x_0+\varepsilon\bar{x}) - f(x_0))/\varepsilon$ всегда имеет предел. При этом для любого $\varepsilon > 0$

$$f'(x_0, \bar{x}) \;\le\; \frac{f(x_0+\varepsilon\bar{x}) - f(x_0)}{\varepsilon}\,. \tag{12.3}$$

Отметим, что здесь $\varepsilon > 0$ действительно любое, так как если $x_0+\varepsilon\bar{x} \notin dom f$, то $f(x_0+\varepsilon\bar{x}) = +\infty$, и неравенство (12.3) заведомо выполнено.

Из определения (12.1) видно, что для любого $\alpha > 0$ $f'(x_0, \alpha\bar{x}) = \alpha f'(x_0, \bar{x})$, то есть функция $\varphi(\bar{x}) = f'(x_0, \bar{x})$ положительно однородна (первой степени). Оказывается, если f выпукла, то φ также будет выпуклой.

Теорема 12.1. Пусть $f : D \to \mathbb{R}$ выпуклая функция, и $x_0 \in int D$. Тогда для любого $\bar{x} \in \mathbb{R}^n$ значение производной по направлению $f'(x_0, \bar{x})$ конечно. Более того, существует такое число L, что

$$|f'(x_0, \bar{x})| \;\le\; L\,|\bar{x}| \qquad \forall\, \bar{x} \in \mathbb{R}^n.$$

Кроме того, $f'(x_0, \bar{x})$ выпукла по $\bar{x}$. Таким образом, $\varphi(\bar{x}) = f'(x_0, \bar{x})$ есть ограниченная сублинейная функция.

Доказательство. Пусть $r > 0$ таково, что шар $B_r(x_0) \subset int\, D$. По теореме 8 f липшицева на этом шаре с некоторой константой L. Тогда для любого $\bar{x}$ при достаточно малых $\varepsilon > 0$ точка $x_0+\varepsilon\bar{x} \in B_r(x_0)$, поэтому

$$|f'(x_0, \bar{x})| = \lim_{\varepsilon\to 0+} \left| \frac{f(x_0+\varepsilon\bar{x}) - f(x_0)}{\varepsilon} \right| \;\le\; \frac{L\,|\varepsilon\bar{x}|}{\varepsilon} \;=\; L\,|\bar{x}|,$$

а это и означает ограниченность φ.

Докажем выпуклость φ. Пусть $\bar{x}_1, \bar{x}_2 \in \mathbb{R}^n$, $\alpha_1, \alpha_2 \ge 0$, $\alpha_1+\alpha_2 = 1$. Тогда

$$f'(x_0,\, \alpha_1\bar{x}_1+\alpha_2\bar{x}_2) = \lim_{\varepsilon\to 0+} \frac{f(x_0+\varepsilon(\alpha_1\bar{x}_1+\alpha_2\bar{x}_2)) - f(x_0)}{\varepsilon} =$$

$$= \lim_{\varepsilon\to 0+} \frac{f(\alpha_1(x_0+\varepsilon\bar{x}_1)+\alpha_2(x_0+\varepsilon\bar{x}_2))-f(x_0)}{\varepsilon} \le$$

$$\le \lim_{\varepsilon\to 0+} \frac{\alpha_1 f(x_0+\varepsilon\bar{x}_1)+\alpha_2 f(x_0+\varepsilon\bar{x}_2)-\alpha_1 f(x_0)-\alpha_2 f(x_0)}{\varepsilon} =$$

$$= \alpha_1 \lim_{\varepsilon\to 0+} \frac{f(x_0+\varepsilon\bar{x}_1)-f(x_0)}{\varepsilon} \;+\; \alpha_2 \lim_{\varepsilon\to 0+} \frac{f(x_0+\varepsilon\bar{x}_2)-f(x_0)}{\varepsilon} =$$

$$= \;\; \alpha_1 f'(x_0,\bar{x}_1) \;+\; \alpha_2 f'(x_0,\bar{x}_2).$$

□

Установим теперь связь между субдифференциалом выпуклой функции f и ее производной по направлению $\varphi(\bar{x}) = f'(x_0,\bar{x})$.

Теорема 12.2. Для любого $x_0 \in int\, D$ справедливо равенство

$$\partial f(x_0) \;=\; \partial f'(x_0,\cdot),$$

то есть $\partial f(x_0) = \partial\varphi$, где $\varphi(\bar{x}) = f'(x_0,\bar{x})$, $\bar{x}\in\mathbb{R}^n$.

Доказательство. ($\subset$) Пусть $p \in \partial f(x_0)$. Тогда $\forall\,\bar{x}\in\mathbb{R}^n$, $\forall\,\varepsilon>0$ по определению имеем $\varepsilon(p,\bar{x}) \le f(x_0+\varepsilon\bar{x}) - f(x_0)$, откуда

$$(p,\bar{x}) \;\le\; \lim_{\varepsilon\to 0+} \frac{f(x_0+\varepsilon\bar{x})-f(x_0)}{\varepsilon} \;=\; f'(x_0,\bar{x}),$$

то есть $p \in \partial f'(x_0,\cdot)$.

($\supset$) Обратно. Пусть $p \in \partial f'(x_0,\cdot)$, то есть $\forall\,\bar{x}$ выполнено $(p,\bar{x}) \le f'(x_0,\bar{x})$. В силу (12.3) при $\varepsilon = 1$ получаем $f'(x_0,\bar{x}) \le f(x_0+\bar{x}) - f(x_0)$, и тогда

$$(p,\bar{x}) \;\le\; f(x_0 \;+\; \bar{x}) \;-\; f(x_0) \qquad \forall\,\bar{x}\in\mathbb{R}^n,$$

а это и означает, что $p \in \partial f(x_0)$. □

Так как сублинейная функция $\varphi(\bar{x}) = f'(x_0,\bar{x})$ ограничена, то она непрерывна на $\mathbb{R}^n$, и следовательно замкнута, а тогда по теореме Минковского

$$\varphi(\bar{x}) = \max_{p\in\partial\varphi}\, (p,\bar{x}), \qquad \forall\,\bar{x}\in\mathbb{R}^n,$$

а отсюда и из теоремы 12.2 вытекает очевидное

Следствие. Для любого $x_0 \in int\, D$

$$f'(x_0,\bar{x}) \;=\; \max_{p\in\partial f(x_0)} (p,\bar{x}) \qquad \forall\,\bar{x}\in\mathbb{R}^n. \tag{12.4}$$

Таким образом, для вычисления производной выпуклой функции по направлению нет нужды находить предел (12.1) — достаточно знать её субдифференциал, который определяется без использования каких-либо пределов.

Из формулы (12.4) вытекает

Лемма 12.2. Пусть выпуклая функция $f : D \to \mathbb{R}$ в некоторой точке $x_0 \in int\, D$ имеет единственный субградиент p. Тогда $\forall\, \bar{x} \in \mathbb{R}^n \quad f'(x_0, \bar{x}) = (p, \bar{x})$. (В этом случае говорят, что p есть производная Гато функции f в точке x_0).

Следующая лемма показывает, что для выпуклых функций в конечномерном пространстве понятия производной Гато и обычной производной (т.е. производной Фреше) совпадают.

Лемма 12.3. Пусть $D \subset \mathbb{R}^n$, и выпуклая функция $f : D \to \mathbb{R}$ в некоторой точке $x_0 \in int\, D$ имеет производную Гато: $f'_\Gamma(x_0) = p$. Тогда p есть ее производная Фреше, т.е. обычная производная.

Доказательство. Считаем $x_0 = 0, \quad p = 0$. Тогда $\forall\, x \quad f(x) \geq 0$, и

$$f(\varepsilon x) = o(\varepsilon) \qquad \text{при} \quad \varepsilon \to 0+ \ . \tag{12.5}$$

Нам надо показать, что $f(x) = o(|x|)$ при $|x| \to 0$ (т.е. что "о малое" в формуле (12.5) является равномерным по всем направлениям). Так как всюду $f \geq 0$, то достаточно показать, что $f(x) \leq o(|x|)$ при $|x| \to 0$.

Возьмем любой многогранник A, у которого 0 есть внутренняя точка. Пусть $z_1, \ldots, z_N$ есть его вершины. Согласно (12.5), $\forall\, i = 1, \ldots, N$ при $\varepsilon \to 0$ имеем

$$f(\varepsilon z_i) = \varepsilon\, g_i(\varepsilon), \qquad \text{где} \quad g_i(\varepsilon) \to 0.$$

Положим $G(\varepsilon) = \max\limits_{1 \leq i \leq N} g_i(\varepsilon)$. Тогда $\forall\, i \quad f(\varepsilon z_i) \leq \varepsilon\, G(\varepsilon)$, и $G(\varepsilon) \to 0$.

Пусть $\mu(x)$ есть функция Минковского множества A. Заметим, что если $|x| \to 0$, то и $\mu(x) \to 0$, и наоборот (докажите!).

Тогда любой $x \in \mathbb{R}^n$ представим в виде $x = \mu(x)\, y$, где $y \in A$, т.е.

$$x = \mu(x) \sum_{i=1}^{N} \alpha_i(x)\, z_i\ , \qquad \text{где} \quad \alpha_i(x) \geq 0, \quad \sum_{i=1}^{N} \alpha_i(x) = 1,$$

и в силу выпуклости f получаем

$$f(x) = f\left(\sum \alpha_i(x)\, (\mu(x)\, z_i)\right) \leq \sum \alpha_i(x)\, \mu(x)\, G(\mu(x)) =$$

$$= \mu(x)\, G(\mu(x)) = o(\mu(x)) = o(|x|), \qquad \text{ч.т.д.}$$

□

Из этих двух лемм вытекает сформулированная выше теорема 10.1.

Производная по направлению и субдифференциал максимума.

Как хорошо известно, максимум дифференцируемых функций может не быть дифференцируемым в данной точке. Тем не менее, в любой точке он имеет производную по любому направлению. Покажем это. Сначала установим одно простое свойство максимума нескольких функций (не обязательно выпуклых).

Лемма 12.4. Пусть открытом множестве $U \in \mathbb{R}^n$ заданы функции $f_i : U \to \mathbb{R}$, $i = 1, ..., m$, которые непрерывны в некоторой точке $x_0 \in U$ и имеют производные по некоторому направлению $\bar{x} \in \mathbb{R}^n$. Тогда функция

$$F(x) = \max_{1 \le i \le m} f_i(x), \qquad x \in U,$$

также имеет в точке x_0 производную по данному направлению $\bar{x}$, которая выражается формулой

$$F'(x_0, \bar{x}) \;=\; \max_{i \in I(x_0)} \; f_i'(x_0, \bar{x}),$$

где $I(x_0) = \{\, i \mid f_i(x_0) = F(x_0)\,\}$ — т.н. множество активных индексов.

Доказательство. Так как все функции f_i непрерывны в x_0, и для любого $j \notin I(x_0)$ имеем $f_j(x_0) < F(x_0)$, то в некоторой окрестности $\mathcal{O}_j(x_0)$ будем иметь $f_j(x) < F(x)$. Пусть $\mathcal{V}$ есть пересечение $\mathcal{O}_j(x_0)$ по всем неактивным индексам. Тогда для любого $x \in \mathcal{V}$

$$F(x) = \max_{i \in I(x_0)} \; f_i(x).$$

Возьмем произвольный $\bar{x} \in \mathbb{R}^n$. Так как при достаточно малых $\varepsilon > 0$ точка $x_0 + \varepsilon\bar{x} \in \mathcal{V}$, то по определению

$$F'(x_0, \bar{x}) = \lim_{\varepsilon \to 0+} \frac{F(x_0 + \varepsilon\bar{x}) - F(x_0)}{\varepsilon} = \lim_{\varepsilon \to 0+} \; \max_{i \in I(x_0)} \; \frac{f_i(x_0 + \varepsilon\bar{x}) - f_i(x_0)}{\varepsilon} =$$

$$= \max_{i \in I(x_0)} \; \lim_{\varepsilon \to 0+} \frac{f_i(x_0 + \varepsilon\bar{x}) - f_i(x_0)}{\varepsilon} \;=\; \max_{i \in I(x_0)} f_i'(x_0, \bar{x}).$$

$\square$

(Почему перестановка предела и максимума здесь законна?)

Теперь мы готовы доказать обобщение формулы (11.10) для произольных выпуклых функций.

Теорема 12.3 (Дубовицкого-Милютина о субдифференциале максимума выпуклых функций). Пусть $D \subset \mathbb{R}^n$ – выпуклое множество, $f_i : D \to \mathbb{R}$, $i = 1, ..., m$, – выпуклые функции, $F(x) = \max_i f_i(x)$. Пусть $x_0 \in int\, D$, $I(x_0) = \{\, i \mid f_i(x_0) = F(x_0)\}$ – множество активных индексов. Тогда

$$\partial\, F(x_0) \;=\; co \left(\bigcup_{i \in I(x_0)} \partial f_i(x_0) \right). \tag{12.6}$$

Доказательство. Введем сублинейные функции

$$\varphi_i(\bar{x}) = f_i'(x_0, \bar{x}), \qquad i \in I(x_0), \qquad \bar{x} \in \mathbb{R}^n .$$

Положим $\Phi(\bar{x}) = \max\limits_{i \in I(x_0)} \varphi_i(\bar{x})$, это тоже сублинейная функция.

По лемме 12.2 $F'(x_0, \bar{x}) = \Phi(\bar{x}) \quad \forall \bar{x}$, а по теореме 11.3

$$\partial \Phi \;=\; co \left(\bigcup_{i \in I(x_0)} \partial \varphi_i \right) . \tag{12.7}$$

Наконец, по теореме (12.2) $\partial F(x_0) = \partial \Phi$, а $\partial f_i(x_0) = \partial \varphi_i \quad \forall i$, поэтому (12.7) превращается в (12.6). □

Конус внешних нормалей к множеству подуровня выпуклой функции

Пусть, как и раньше, $f : D \to \mathbb{R}$ – выпуклая функция, $x_0 \in int\, D$. Рассмотрим множество

$$A = \{\, x \in D \mid f(x) \le f(x_0)\}$$

— т.н. нижнее лебегово множество, или множество подуровня данной функции. Ясно, что оно выпукло. Множества такого типа являются одним из основных объектов в задачах на экстремум. (Например, они могут являться ограничениями задачи.)

Требуется найти конус касательных направлений $K_A(x_0)$ и сопряженный к нему, или, что эквивалентно, конус внешних нормалей $N_A(x_0)$.

Будем предполагать, что существует $x_1 \in D$, для которого $f(x_1) < f(x_0)$, т.е. что x_0 не есть точка минимума функции f на данном множестве D (условие Слейтера). При этих условиях справедлива

Теорема 12.4.

$$K_A(x_0) = \; \{\, \bar{x} \in \mathbb{R}^n \mid f'(x_0, \bar{x}) \le 0 \,\} . \tag{12.8}$$

Доказательству предпошлем следующую

Лемму 12.5. Точка $x_1 \in D$, для которой $f(x_1) < f(x_0)$ существует тогда и только тогда, когда существует $\bar{x} \in \mathbb{R}^n$, для которого $f'(x_0, \bar{x}) < 0$.

Доказательство. ($\Rightarrow$) Полагая $\bar{x} = x_1 - x_0$ и применяя неравенство (12.3) с $\varepsilon = 1$, имеем $f'(x_0, \bar{x}) \le f(x_0 + \bar{x}) - f(x_0) < 0$.

($\Leftarrow$) При $\varepsilon \to 0+$ точка $x_0 + \varepsilon\bar{x} \in D$, и для нее при достаточно малых $\varepsilon > 0$ имеем $f(x_0 + \varepsilon\bar{x}) = f(x_0) + \varepsilon f'(x_0, \bar{x}) + o(\varepsilon) < f(x_0)$. $\square$

С помощью этой леммы дадим

Доказательство теоремы 12.4. ($\subset$) Пусть $\bar{x} \in K_A(x_0)$. По определению это означает, что $x_0+\varepsilon\bar{x}+r(\varepsilon) \in A$ при $\varepsilon \to 0+$, где $r(\varepsilon) = o(\varepsilon)$ − некоторая "поправка". Тогда $f(x_0+\varepsilon\bar{x}+r(\varepsilon)) \le f(x_0)$, а с учетом липшицевости f в некоторой окрестности точки x_0 имеем $f(x_0 + \varepsilon\bar{x}) \le f(x_0) + o(\varepsilon)$, и поэтому

$$f'(x_0, \bar{x}) = \lim_{\varepsilon \to)+} \frac{f(x_0 + \varepsilon\bar{x}) - f(x_0)}{\varepsilon} \le 0.$$

($\supset$) Обозначим множество в правой части (12.8) через Ω. Ясно, что оно замкнуто. По лемме 11.5, учитывая лемму 12.3, имеем

$$int\,\Omega \;=\; \{\, \bar{x} \mid f'(x_0, \bar{x}) < 0\},$$

и это множество непусто. Кроме того, мы знаем, что

$$\overline{int\,\Omega} \;=\; \overline{\Omega} \;=\; \Omega.$$

Если $\bar{x} \in int\,\Omega$, то есть $f'(x_0, \bar{x}) = -c < 0$, то при достаточно малых $\varepsilon > 0$ будет

$$f(x_0 + \varepsilon\bar{x}) - f(x_0) = \varepsilon f'(x_0, \bar{x}) + o(\varepsilon) \le -\frac{c}{2}\,\varepsilon,$$

поэтому $x_0 + \varepsilon\bar{x} \in A$, и следовательно, $\bar{x} \in K_A(x_0)$. Таким образом, $int\,\Omega \subset K_A(x_0)$. Но тогда, в силу замкнутости любого касательного конуса, и $\Omega \subset K_A(x_0)$. $\square$

Итак, мы доказали, что $K_A(x_0)$ задается сублинейным неравенством $\varphi(\bar{x}) = f'(x_0, \bar{x}) \le 0$. При этом, согласно лемме 12.3, выполнено условие Слейтера: существует $\bar{x}$, для которого $\varphi(\bar{x}) < 0$. Отсюда по теореме 11.5 имеем

$$N_A(x_0) = -K_A^*(x_0) \;=\; \mathbb{R}_+ \cdot \partial\varphi \;=\; \mathbb{R}_+ \cdot \partial f(x_0),$$

то есть в итоге получаем

Следствие из теоремы 12.4.

$$N_A(x_0) = \;\mathbb{R}_+ \cdot \partial f(x_0).$$

Эта формула, как и (11.11), часто используется при исследовании задач оптимизации.

Лекция 13. Преобразование Лежандра-Юнга-Фенхеля

В этой лекции мы определим двойственный объект к выпуклой функции и установим основные свойства операции перехода к двойственной функции.

Пусть $f : \mathbb{R}^n \to \overline{\mathbb{R}} = \mathbb{R} \cup \{\pm\infty\}$ есть произвольная функция.

Определение 13.1. *Сопряженной (двойственной)* функцией к функции f называется функция $f^* : \mathbb{R}^n \to \overline{\mathbb{R}}$, определяемая равенством

$$f^*(p) = \sup_{x \in \mathbb{R}^n} \{(p, x) - f(x)\}. \tag{13.1}$$

Как мы видим в этой формуле x играет роль параметра, а аргументом функции является p. (Строго говоря, p есть элемент сопряженного пространства $(\mathbb{R}^n)^*$, но мы отождествляем его с исходным пространством $\mathbb{R}^n$.)

Операция перехода от функции f к ее сопряженной f^* по формуле 13.1 называется *преобразованием Лежандра-Юнга-Фенхеля*.

При каждом фиксированном x функция $(p, x) - f(x)$ аффинна по p, поэтому f^*, как супремум аффинных, всегда выпукла и замкнута. Отметим также, что из формулы 13.1 сразу вытекает *неравенство Юнга:*

$$\forall\, p, x \qquad f^*(p) + f(x) \geq (p, x). \tag{13.2}$$

Свойство 1. Если функция f собственная и выпуклая, то f^* — собственная выпуклая замкнутая.

Доказательство. В доказательстве нуждается только собственность f^*.
Так как f собственная, то $\exists\, x_0$, для которого $f(x_0) \in \mathbb{R}$. Тогда $\forall\, p$ имеем $f^*(p) \geq (p, x_0) - f(x_0) > -\infty$. Первое условие собственности выполнено. Далее, по лемме 9.3 f имеет аффиную миноранту, то есть имеются $\hat{p}$, $\hat{b}$ такие, что $l(x) = (\hat{p}, x) - \hat{b} \leq f(x) \quad \forall\, x \in \mathbb{R}^n$. Тогда

$$f^*(\hat{p}) = \sup_{x \in \mathbb{R}^n} [(\hat{p}, x) - f(x)] \leq \sup_{x \in \mathbb{R}^n} [(\hat{p}, x) - l(x)] = \hat{b} < +\infty,$$

то есть выполнено и второе условие собственности f^*. □

Свойство 2. Если $f(x) \leq g(x) \quad \forall\, x$, то $f^*(p) \geq g^*(p) \quad \forall\, p$.

Доказательство. Действительно,

$$g^*(p) = \sup_x [\, p x - g(x)\,] \leq \sup_x [\, p x - f(x)\,] = f^*(p). \qquad \square$$

Свойство 3. $(\overline{f})^* \equiv f^*, \qquad (co\, f)^* \equiv f^*.$

Доказательство. Пусть $E = epi f$. Тогда $\forall p \in \mathbb{R}^n$ имеем

$$f^*(p) = \sup_x \sup_{\mu \geq f(x)} (px - \mu) = \sup_{(x,\mu)\in E} (px - \mu) = \sup_{(x,\mu)\in \overline{E}} (px - \mu) =$$

$$= \sup_x \sup_{\mu:\,(x,\mu)\in \overline{E}} (px - \mu) = \sup_x [px - \overline{f}(x)] = (\overline{f})^*(p).$$

Здесь мы воспользовались тем, что

$$\sup_{\mu:\,(x,\mu)\in\overline{E}} (-\mu) = -\inf_{\mu:\,(x,u)\in\overline{E}} (\mu) = -\overline{f}(x).$$

Итак, первое равенство установлено. Второе устанавливается аналогично. □

К функции f^* также можно применить операцию взятия сопряженной, получив $f^{**} -$ вторую сопряженную к f.

Свойство 4. $f^{**} \leq f.$

Доказательство. Зафиксируем любой $x \in \mathbb{R}^n$. Согласно неравенству Юнга, $\forall p$ имеем $f(x) \geq p\,x - f^*(p)$, и тогда

$$f(x) \geq \sup_p [p\,x - f^*(p)] = f^{**}(x).$$

Последнее равенство — по определению сопряженной функции к f^*. □

Примеры.

1) Пусть $l(x) = (a,x) - b$ (аффинная функция). Тогда

$$l^*(x) = \sup_x [((p-a),x) + b] = \begin{cases} b, & \text{если } p = a \\ +\infty, & \text{если } p \neq a \end{cases}$$

Далее,

$$l^{**}(x) = \sup_p [(p,x) - l^*(p)] = (a,x) - b = l(x).$$

(При $p \neq a$ функция в скобках равна $-\infty$, поэтому такие p при взятии супремума можно не учитывать.)

Таким образом для аффинной функции l непосредственным вычислением получаем $l^{**}(x) = l(x) \quad \forall x$. Запомним этот факт.

2) $f(x) = \frac{1}{p}|x|^p, \qquad 1 < p < \infty \qquad (x \in \mathbb{R}).$
(здесь $p -$ не аргумент сопряженной функции, а традиционный параметр.)

Тогда, пользуясь *неравенством Гельдера*, нетрудно показать, что

$$f^*(y) = \frac{1}{q}|y|^q, \qquad \text{где } q \in (1, \infty) -$$

— т.н. *сопряженный показатель,* определяемый равенством $1/p + 1/q = 1$.

3) $f : \mathbb{R}^n \to \mathbb{R}, \qquad f(x) = \frac{1}{2}(x, x)$.

Так как

$$\max_x \left[(y, x) - \frac{1}{2}(x, x)\right]$$

достигается при $x_* = y$, то $f^*(y) = \frac{1}{2}(y, y)$, то есть $f^* = f$.

Возникает вопрос: есть ли еще функции, совпадающие со своими сопряженными? Ответ — нет.

Утверждение. Пусть $\varphi^* \equiv \varphi$. Тогда $\varphi(x) = \frac{1}{2}(x, x)$.

Доказательство. Согласно неравенству Юнга, $\forall x$ имеем $2\varphi(x) = \varphi(x) + \varphi^*(x) \geq (x, x) = 2f(x)$, то есть $\varphi(x) \geq f(x)$, и тогда по свойству 2 $\varphi^* \leq f^*$. Снимая "звездочки" (в силу равенств $\varphi^* = \varphi$, $f^* = f$), получаем $\varphi \leq f$, и следовательно $\varphi = f$. $\square$

4) Пусть L есть подпространство в $\mathbb{R}^n$, а $\delta_L(x) -$ его индикаторная функция. Тогда $\delta_L^*(p) = \delta_{L^\perp}(p)$ есть индикаторная функция ортогонального дополнения к L; при этом $\delta_L^{**}(x) = \delta_L(x)$, то есть опять вторая сопряженная функция совпадает с исходной.

5) Пусть $f(x) = e^x$. Тогда

$$f^*(p) = \begin{cases} p \ln p - p, & \text{если } p > 0, \\ 0, & \text{если } p = 0, \\ +\infty, & \text{если } p < 0. \end{cases}$$

Здесь мы видим, что даже для функции, принимающей всегда конечные значения, сопряженная функция может принимать значение $+\infty$. Поэтому, чтобы не выходить из заданного класса функций, принято с самого начала рассматривать функции со значениями в расширенной прямой $\mathbb{R} \cup \{+\infty\}$.

Если теперь еще раз перейти к сопряженной функции, то нетрудно убедиться, что $f^{**} = f$.

6) Пусть

$$f(x) = \begin{cases} -\sqrt{a^2 - x^2}, & \text{если } |x| \leq a, \\ +\infty, & \text{если } |x| > a. \end{cases}$$

Тогда $f^*(p) = a\sqrt{1 + p^2}, \qquad p \in \mathbb{R}$, и опять $f^{**} = f$.

7) Пусть $\varphi -$ произвольная замкнутая сублинейная функция, то есть

$$\varphi(x) = \sup_{a \in A}(a, x) \qquad \forall\, x \in \mathbb{R}^n,$$

где $A -$ непустое выпуклое замкнутое множество. Тогда

$$\varphi^*(x) = \delta_A(p) = \begin{cases} 0\ , & \text{если } p \in A, \\ +\infty, & \text{если } p \notin A \end{cases}$$

— индикаторная функция множества A.

Доказательство. Если $p \in A$, то по определению $(p, a) \le \varphi(x)\ \forall\, x,$ поэтому

$$\varphi^*(p) = \sup_x (p\,x - \varphi(x))$$

достигается при $x = 0$, то есть $\varphi^*(p) = 0$.

Если же $p \notin A$, то по определению $\partial\varphi$ существует $\hat{x}$ такой, что $p\,\hat{x} > \varphi(\hat{x})$, поэтому для $x_\lambda = \lambda\,\hat{x}$ при $\lambda \to \infty$ получаем

$$\varphi^*(p) = \sup_x\ [p\,x - \varphi(x)] \ge \sup_{\lambda > 0}\ [p\,x_\lambda - \varphi(x_\lambda)] = \sup_{\lambda > 0} \lambda\,[p\,\hat{x} - \varphi(\hat{x})] = +\infty\,.$$

□

Возьмем еще раз сопряженную функцию: покажем, что $\delta_A^* = \varphi(x)$.

Действительно, по определению

$$\delta_A^* = \sup_p\ [(p, x) - \delta_A(p)]\,.$$

Если $p \notin A$, то выражение в скобках $= -\infty$, и оно не влияет на супремум, поэтому достаточно рассматривать $p \in A$, и следовательно

$$\delta_A^* = \sup_{p \in A}(p, x) = \varphi(x)\,.$$

Таким образом, если $A -$ выпуклое замкнутое множество, и $\varphi_A(x)$ есть его опорная функция, то

$$\varphi_A^*(p) = \delta_A(p)\,, \qquad\qquad \delta_A^*(p) = \varphi_A(x)\,.$$

Во всех приведенных примерах, как мы видели, вторая сопряженная функция совпадает с исходной. Случаен ли этот факт?

Теорема 13.1. (Фенхель–Моро). Пусть $f : \mathbb{R}^n \to \mathbb{R}^\bullet$ есть собственная выпуклая замкнутая функция. Тогда $f^{**} = f$.

Доказательство. Как мы уже знаем, для аффинной функции l всегда выполнено равенство $l^{**} = l$. Тогда, если $l \le f$, то по свойствам 2 и 4

$$l = l^{**} \le f^{**} \le f\,.$$

Отсюда
$$l(x) \le f^{**}(x) \le f(x) \qquad \forall\, x \in \mathbb{R}^n\,.$$
Пусть $\mathcal{L} = \{\, l\}$ есть множество всех аффинных минорант функции f. Тогда по теореме Минковского $\forall\, x \in \mathbb{R}^n$
$$f(x) = \sup_{l \in \mathcal{L}} l(x) \le f^{**}(x) \le f(x)\,.$$
В начале и конце этой цепочки неравенств стоит одно и то же число $f(x)$, поэтому все неравенства обращаются в равенства, и следовательно, $f^{**}(x) = f(x)$. $\square$

Следствие 1.

а) Если функция f собственная и выпуклая, то $f^{**} = \overline{f}$.
б) Если функция f собственная, то $f^{**} = \overline{co\ f}$.

Доказательство. По свойству 3 и теореме 13.1

а) $(f^*)^* = (\overline{f}^*)^* = \overline{f}$, ибо $\overline{f}$ собственная выпуклая и замкнутая;

б) $(f^*)^* = ((co\ f)^*)^* = (co\ f)^{**} = \overline{co\ f}$. $\square$

Из теоремы 13.1 вытекает следующий интересный факт. Обозначим через $\mathfrak{M}$ множество всех собственных выпуклых замкнутых функций на $\mathbb{R}^n$, а через $\mathfrak{M}_0 -$ подмножество $\mathfrak{M}$, состоящее из неотрицательных функций, равных нулю в нуле.

Согласно свойству 1, операцию перехода к сопряженной функции $(f \mapsto f^*)$ можно рассматривать как отображение
$$\mathfrak{M} \xrightarrow{\ *\ } \mathfrak{M}\,, \tag{13.8}$$
а из теоремы 13.1 следует, что это отображение взаимно-однозначно и "на". Действительно, $\forall\, g \in \mathfrak{M}$ найдется такая $f \in \mathfrak{M}$, что $f^* = g$ (достаточно положить $f = g^*$, тогда $f^* = g^{**} = g$). Если $f_1 \ne f_2$, то $f_1^* \ne f_2^*$, ибо, если $f_1^* = f_2^*$, то $f_1 = f_1^{**} = f_2^{**} = f_2$, противоречие.

Подмножество $\mathfrak{M}_0$ является инвариантным при отображении (13.8), то есть
$$\mathfrak{M}_0 \xrightarrow{\ *\ } \mathfrak{M}_0\,. \tag{13.9}$$
Действительно, если $f \in \mathfrak{M}_0$, т.е. $f(x) \ge 0 \ \ \forall\, x$, и $f(0) = 0$, то $\forall\, p$
$$f^*(p\,) = \sup_x\,(p\,x - f(x) \ge (p\,x - f(x))\Big|_{x\,=\,0} = 0\,,$$
$$f^*(0\,) = \sup_x\,(-f(x)) = -f(0) = 0\,,$$
поэтому $f^* \in \mathfrak{M}_0$.

Отображение (13.9) также является взаимно-однозначным и "на". Первое установлено выше, а второе, как и раньше, вытекает из теоремы 13.1: $\forall g \in \mathfrak{M}_0$ надо взять функцию $f = g^*$, которая согласно (13.9) также принадлежит $\mathfrak{M}_0$; тогда $f^* = g$.

Преобразование Лежандра

Пусть $f \in C^2(\mathbb{R}^n)$, то есть это дважды гладкая функция на всем пространстве, и пусть $\forall x$ матрица $f''(x)$ положительно определена: $f''(x) > 0$. Тогда $\forall p$ либо

$$\sup_x (px - f(x)) = +\infty,$$

либо этот супремум достигается в некоторой точке x_p, определяемой из равенства

$$p - f'(x_p) = 0. \tag{13.4}$$

Так как всюду $\det f''(x) \neq 0$, то из этого равенства x_p однозначно выражается через p, и тогда

$$f^*(p) = (p, x_p) - f(x_p). \tag{13.5}$$

Это и есть классическое преобразование Лежандра.

Преобразование Лежандра в классической механике

Пусть дана система из N материальных точек с координатами $x_1, \ldots, x_N \in \mathbb{R}^3$ и массами $m_1, \ldots, m_N$. При движении системы координаты точек зависят от времени t, а их скорости равны $\dot{x}_1, \ldots, \dot{x}_N$, соответственно. Кинетическая энергия системы есть $T(\dot{x}_1, \ldots, \dot{x}_N) = \sum_{k=1}^{N} m_k \dot{x}_k^2$, а потенциальная энергия пусть выражается некоторой функцией координат $U(x_1, \ldots, x_N)$. Такая система называется *консервативной.* Для краткости введем обозначения: $\mathbf{x} = (x_1, \ldots, x_N)$, $\dot{\mathbf{x}} = (\dot{x}_1, \ldots, \dot{x}_N)$. Согласно *принципу наименьшего* (а точнее, *стационарного*) *действия,* движение системы происходит по такой траектории, для которой функционал

$$I = \int_{t_0}^{t_1} (T(\dot{x}_1, \ldots, \dot{x}_N) - U(x_1, \ldots, x_N))\, dt,$$

называемый *действием,* принимает стационарное значение, то есть траектория движения должна удовлетворять необходимому условию в задаче $I(\mathbf{x}) \to \min$ – уравнению Эйлера. Подинтегральная функция

$$L(x, \dot{x}) = T(\dot{x}) - U(x) \tag{13.6}$$

называется *лагранжианом.* Уравнение Эйлера, как известно из классического вариационного исчисления, имеет вид

$$(L_{\dot{x}})^{\bullet} = L_x \,,$$

то есть в нашем случае

$$m_k \, \ddot{x_k} = -U'_{x_k} \,, \qquad k = 1, \ldots, N \,, \tag{13.7}$$

— это в точности есть *второй закон Ньютона* (минус градиент U по x_k есть сила, действующая на k -ую точку). Собственно говоря, само выражение для лагранжиана $L = T - U$ как раз и было выбрано так, чтобы уравнение Эйлера для него превратилось в закон Ньютона (13.7).

Будем теперь смотреть на $L(x, \dot{x})$ как на функцию от второго аргумента $\dot{x}$, считая первый аргумент x просто параметром. Тогда L есть выпуклая функция от $\dot{x}$, а ее преобразование Лежандра по $\dot{x}$ есть

$$L^*(x, p) = \sup_{\dot{x}} \left[\sum_{k=1}^{N} (p_k, \dot{x_k}) - \frac{1}{2} \sum_{k=1}^{N} m_k \dot{x_k}^2 + U(x) \right].$$

Супремум здесь очевидно достигается при

$$p_k = m_k \, \dot{x_k} \,, \qquad k = 1, \ldots, N \,; \tag{13.7}$$

величина p_k называется *импульсом* k-ой точки.

Тогда

$$L^*(x, p) = \sum_k \frac{1}{2} \frac{p_k^2}{m_k} + U(x) \,.$$

Первое слагаемое, стоящее справа, есть не что как кинетическая энергия, выраженная через импульсы, а вся сумма есть *полная энергия системы* (кинетическая + потенциальная). Полученная функция обозначается $\mathcal{H}(x, p)$ и называется функцией Гамильтона или *гамильтонианом.* Таким образом, преобразование Лежандра по $\dot{x}$ от лагранжиана $L(x, \dot{x})$ есть гамильтониан $\mathcal{H}(x, p)$.

Преобразование Лежандра в оптимальном управлении

Рассмотрим задачу вариационного исчисления

$$J = \int_{t_0}^{t_1} L(t, x, u) dt \to \min \,,$$

$$\dot{x} = u, \qquad K(x_0, x_1) = 0 \,.$$

Преобразование Лежандра по u от функции $L(t, x, u)$, если t, x в ней считать параметрами, задается формулой

$$L^*(t, x, p) = \max_u \, [(p, u) - L(t, x, u)] \,. \tag{13.9}$$

Стоящая в скобках функция $H(t,x,p,u) = (p,u) - L(t,x,u)$ называется функцией Понтрягина, а (13.9) есть функция Гамильтона $\mathcal{H}(t,x,p)$. Обе эти функции, как и лагранжиан задачи $L(t,x,u)$, играют важную роль в вариационном исчислении и оптимальном управлении.

Итак, гамильтониан $\mathcal{H}(t,x,p)$ есть преобразование Лежандра по u от лагранжиана; при этом

$$\mathcal{H}(t,x,p) = \max_u H(t,x,p,u)\,,$$

т.е. гамильтониан есть максимум функции Понтрягина по управлению.

Лекция 14. Выпуклые экстремальные задачи.

Определение 14.1. *Выпуклой экстремальной задачей* называется задача поиска минимума выпуклой функции f на выпуклом множестве A:

$$f(x) \to \min, \qquad x \in A. \tag{14.1}$$

Здесь A — непустое выпуклое множество в некотором векторном пространстве X (у нас в $\mathbb{R}^n$), оно задает ограничение задачи и называется *множеством допустимых точек*, а f — собственная выпуклая функция на X. (Без нарушения общности можно было бы считать, что $f < +\infty$ на всем A, иначе вместо A можно рассмотреть множество $A_0 = A \cap dom f$.)

Обратим внимание, что у выпуклой функции ищется именно минимум, а не максимум: задача на поиск максимума выпуклой функции является, так сказать, неправильной — в такой задаче выпуклость функции мало чем помогает. В задаче же на минимум выпуклость функции в полной мере проявляет свои преимущества, что позволяет довести теорию таких задач до полного логического завершения уже на уровне условий первого порядка.

Установим сразу следующий важный факт, справедливый для выпуклых задач (и только для них!).

Лемма 14.1. Пусть $x_0 \in A$ есть точка локального минимума в задаче (14.1). Тогда в ней достигается и глобальный минимум в задаче (14.1).

Доказательство. Наличие локального минимума в точке x_0 по определению означает, что существует её окрестность $U(x_0)$, такая что

$$\forall\, x \in U(x_0) \cap A \qquad f(x) \geq f(x_0). \tag{14.2}$$

Допустим, что при этом глобального минимума в точке x_0 нет, т.е. что существует точка $x_1 \in A$, в которой $f(x_1) < f(x_0)$. Соединим обе эти точки отрезком и рассмотрим на нем точки $x_\varepsilon = (1-\varepsilon)x_0 + \varepsilon x_1, \quad 0 < \varepsilon \leq 1$.

При достаточно малых $\varepsilon > 0$ будем, очевидно, иметь $x_\varepsilon \in U(x_0) \cap A$, и при этом в силу выпуклости f

$$f(x_\varepsilon) \;\leq\; (1-\varepsilon)f(x_0) + \varepsilon f(x_1) \;<\; (1-\varepsilon)f(x_0) + \varepsilon f(x_0) \;=\; f(x_0), \tag{14.3}$$

что противоречит условию (14.2). (Строгое неравенство здесь за счет того, что $\varepsilon > 0$ и $f(x_1) < f(x_0)$.) $\square$

Отметим, что для задачи на максимум f этого факта мы бы уже не получили: неравенство Йенсена в (14.3) справедливо только в эту сторону, поэтому никакого противоречия не было бы. Задачи на максимум естественно рассматривать для вогнутых функций, которые суть выпуклые функции со знаком минус, и поэтому поиск их максимума соответствует поиску минимума выпуклых функций.

Итак, в выпуклых задачах все локальные минимумы являются и глобальными, "чисто локальных" минимумов нет. Поэтому можно говорить просто о минимумах.

Перейдем теперь к вопросу о необходимых условиях минимума. Рассмотрим сначала простейший случай, когда допустимое множество A есть все пространство $\mathbb{R}^n$, т.е. ограничения на x нет:

$$f(x) \to \min, \qquad x \in \mathbb{R}^n. \tag{14.4}$$

В этой задаче наличие минимума в некоторой точке x_0 по определению означает, что $f(x_0) \le f(x) \quad \forall x \in \mathbb{R}^n$, а это эквивалентно тому, что

$$0 \in \partial f(x_0). \tag{14.5}$$

Соотношение (14.5) есть необходимое и достаточное (!) условие минимума в точке x_0. Таким образом, для выпуклой функции необходимое условие минимума является также и достаточным. (Для невыпуклой функции этого свойства, конечно же, нет.)

Вернемся к задаче (14.1). Она очевидным образом эквивалентна следующей задаче без ограничения:

$$\tilde{f}(x) = f(x) + \delta_A(x) \to \min, \qquad x \in \mathbb{R}^n. \tag{14.6}$$

Согласно предыдущему, необходимое и достаточное условие минимума в точке x_0 в этой задаче состоит в том, что

$$0 \in \partial \tilde{f}(x_0). \tag{14.7}$$

Найдем вид $\partial \tilde{f}(x_0)$ в терминах данной функции f и множества A.

Согласно теореме Рокафеллара–Моро, если $A \cap int\, D_f$ непусто, то $\partial \tilde{f}(x_0) = \partial f(x_0) + \partial \delta_A(x_0)$ для любой точки $x_0 \in A \cap D_f$, и условие (14.7) означает, что $\partial f(x_0) + N_A(x_0) \ni 0$, ибо $\partial \delta_A(x_0)$ есть конус внешних нормалей к множеству A в точке x_0. Отсюда вытекает следующая

Теорема 14.1. Пусть $A \cap int\, D_f$ непусто (т.е. f непрерывна хотя бы в одной точке множества A). Тогда $x_0 \in A \cap D_f$ есть точка минимума в задаче (14.1) $\Longleftrightarrow$ $\exists\, p \in \partial f(x_0)$ такой, что $-p \in N_A(x_0)$.

В случае, когда множество A задано неравенством $g(x) \le g(x_0)$, где $g -$ тоже собственная выпуклая функция, непрерывная в точке x_0 (т.е. $x_0 \in int\, D_g$) и не имеющая минимума в точке x_0, теорема (14.1) дает $N_A(x_0) = \mathbb{R}_+ \partial g(x_0)$, и тогда условие минимума в задаче (14.1) приобретает вид:

$$\partial f(x_0) + \mathbb{R}_+ \partial g(x_0) \ni 0,$$

т.е. $\exists\, p \in \partial f(x_0), \quad \lambda \ge 0, \quad q \in \partial g(x_0)$ такие, что $p + \lambda q = 0$.

Если, кроме того, f и g дифференцируемы в точке x_0, то мы получаем

$$f'(x_0) + \lambda g'(x_0) = 0$$

— обычное правило множителей Лагранжа.

Рассмотрим теперь более сложную задачу:

$$\begin{cases} f_0(x) \to \min, \\ f_i(x) \le 0, \qquad i = 1, ..., m, \\ x \in A, \end{cases} \tag{14.8}$$

где все f_i, $i = 0, 1, ..., m$ – собственные выпуклые функции, A – по-прежнему непустое выпуклое множество. (Здесь множество допустимых точек состоит из тех $x \in A$, которые удовлетворяют указанным неравенствам.)

Для вывода условий минимума в этой задаче можно рассуждать следующим образом. Пусть все функции f_i непрерывны в x_0. Не нарушая общности, будем считать $f_0(x_0) = 0$.

Лемма 14.2. Пусть x_0 есть точка минимума в задаче (14.8). Тогда она доставляет минимум и в задаче

$$F(x) = \max_{i \in I} f_i(x) \to \min, \qquad x \in A, \tag{14.9}$$

где $I = \{i \ge 0 \mid f_i(x_0) = 0\}$ есть множество активных индексов. (Отметим, что всегда $i = 0 \in I$.)

Доказательство. Действительно, если x_0 не есть точка минимума задачи (14.9), то $\exists\, x_1 \in A$, для которого $F(x_1) < F(x_0) = 0$. Тогда $\forall\, i \in I \quad f_i(x_1) < 0$. Соединим x_1 и x_0 отрезком и рассмотрим на нем точки $x_\varepsilon = (1-\varepsilon)x_0 + \varepsilon x_1$ при $0 < \varepsilon \le 1$. При любом таком ε для любого активного индекса получим, по аналогии с (14.3), $f_i(x_\varepsilon) < 0$, т.е. точка $x_\varepsilon \in A$ удовлетворяет ограничениям неравенства при всех $i \in I$, $i \ge 1$, и кроме того $f_0(x_0) < 0$. Чтобы удовлетворить ограничениям неравенства при $i \notin I$ учтем, что $f_i(x_\varepsilon) \to f_i(x_0) < 0$ при $\varepsilon \to 0+$. Поэтому для достаточно малого $\varepsilon > 0$ эти ограничения также будут выполнены. Тогда точка x_ε является допустимой в задаче (14.8), и в ней $f_0(x_\varepsilon) < f_0(x_0)$, что противоречит минимуму в точке x_0. Лемма доказана. $\square$

К задаче (14.9) мы можем применить полученное выше условие минимума:

$$\exists\, p \in \partial F(x_0): \qquad -p \in N_A(x_0), \tag{14.10}$$

а по теореме Дубовицкого–Милютина

$$p = \sum_{i \in I} \alpha_i\, p_i,$$

где

$$p_i \in \partial f_i(x_0), \qquad \alpha_i \geq 0, \qquad \sum_{i \in I} \alpha_i = 1.$$

Введем функцию

$$L(x) = \sum_{i \in I} \alpha_i f_i(x),$$

которая называется функцией Лагранжа задачи (14.8).

Так как $\partial L(x_0) = \sum_{i \in I} \alpha_i \, \partial f_i(x_0)$, то $p \in \partial L(x_0)$, и из (14.10) следует, что

$$\partial L(x_0) + N_A(x_0) \ni 0, \tag{14.11}$$

а согласно теореме 14.1 это означает, что функция L достигает минимума на множестве A в точке x_0.

Итак, из наличия минимума в точке x_0 в задаче (14.8) вытекает, что найдутся такие множители Лагранжа α_i, что соответствующая им функция Лагранжа удовлетворяет условиям (14.11). Таким образом, существование α_i, для которых выполнено (14.11), есть *необходимое условие минимума* в задаче (14.8).

Однако это условие пока не является достаточным, так как переход от минимума в задаче (14.8) к минимуму в задаче (14.9) справедлив лишь в одну сторону (!).

Можно показать (и ниже мы это покажем), что условие (14.11) становится достаточным, если ограничения задачи (14.8) являются в некотором смысле *регулярными*, и тогда аналогия с условиями для задачи (14.2) будет восстановлена.

Некоторый дефект предложенного способа рассуждений состоит в том, что для его проведения требуется непрерывность всех функций f_i в точке x_0.

Сейчас мы рассмотрим другой способ получения условия (14.11), свободный от указанного предположения и, кроме того, справедливый для произвольных векторных пространств, а не только для конечномерных.

Пусть X есть произвольное векторное пространство любой размерности (и, вообще говоря, без какой-либо топологии), $A -$ непустое выпуклое множество в нем.

Рассмотрим задачу вида (14.8), в которой все $f_i : A \to \mathbb{R}, \quad i = 0, 1, ..., m$ есть выпуклые функции на A, т.е. функции, удовлетворяющие на множестве A неравенству Йенсена.

Следующая теорема дает необходимые и достаточные условия минимума в этой задаче.

Теорема 14.2. (Кун–Таккер).

1) Если x_0 – точка минимума, то найдутся множители Лагранжа $\alpha_0, \alpha_1, ..., \alpha_m \geq 0$ такие, что

$$\sum_{i=0}^{m} \alpha_i > 0 \qquad \text{(условие нетривиальности)},$$

$$\alpha_i \, f_i(x_0) = 0, \qquad i = 1, ..., m \qquad \text{(условия дополняющей нежесткости)},$$

и соответствующая им функция Лагранжа $L(x)=\Sigma_{i=0}^m \alpha_i\, f_i(x)$ достигает минимума по множеству A в точке x_0.

2) Если для некоторой допустимой точки x_0 найдутся множители α_i с указанными свойствами, и при этом $\alpha_0>0$, то $x_0 -$ точка минимума.

3) Если $\exists\, \hat{x}\in A$, для которой $f_i(\hat{x})<0$ при всех $i=1,...,m$ (условие Слейтера регулярности ограничений), и для некоторой допустимой точки x_0 имеются множители Лагранжа с указанными в пункте 1 свойствами, то обязательно $\alpha_0>0$ (и, следовательно, $x_0 -$ точка минимума).

Доказательство. 1) В доказательстве первого утверждения теоремы применяется прием "работы в образе", который состоит в том, что теорема отделимости (которая, конечно же, должна где-то использоваться в доказательстве) применяется не в пространстве X (в котором ее, вообще говоря, и невозможно применить, т.к здесь нет топологии), а в пространстве образов конечномерного отображения

$$f:\; X\to \mathbb{R}^{m+1}, \qquad x\longmapsto (f_0(x_0),\ldots,f_m(x)).$$

А именно, в пространстве $\mathbb{R}^{m+1}$ рассмотрим множество C, состоящее из всех тех векторов $\mu=(\mu_0,...,\mu_m)$, для которых $\exists\, x\in A$ такой, что $f_i(x)<\mu_i$ при всех $i=0,1,...,m$. (Это множество есть проекция "векторного надграфика" отображения f на пространство $\mathbb{R}^{m+1}$.) Ясно, что C непусто.

а) Покажем, что C выпукло. Действительно, если $\mu',\mu''\in C$, т.е. имеются $x',x''\in A$, для которых $f_i(x')<\mu_i'$, $\;f_i(x'')<\mu_i''\;\; \forall\, i\geq 0$, то для любых $\alpha>0$, $\beta>0$, $\;\alpha+\beta=1$ точка $\alpha x'+\beta x''\in A$, и для нее $\forall\, i$

$$f_i(\alpha x'+\beta x'')\;\leq\;\alpha f_i(x')+\beta f_i(x'')\;<\;\alpha\mu_i'+\beta\mu_i''\;,$$

и следовательно, вектор $\alpha\mu'+\beta\mu''\in C$.

б) Покажем, что $0\notin C$. Действительно, если бы $0\in C$, то нашелся бы $x\in A$ такой, что $f_i(x)<0\;\;\forall\, i\geq 1$, т.е. точка x была бы допустимой в задаче (14.8), и $f_0(x)<0=f_0(x_0)$, что противоречит минимуму в точке x_0 .

Воспользуемся конечномерной теоремой об отделимости: существует ненулевой вектор $\alpha=(\alpha_0,...,\alpha_m)$, который отделяет точку 0 от выпуклого множества C :

$$\forall\,\mu\in C \qquad (\alpha,\mu)=\sum_{i=0}^m \alpha_i\mu_i\geq 0. \tag{14.12}$$

Установим теперь ряд свойств вектора α.

в) Все $\alpha_i\geq 0$, ибо у точек множества C можно неограниченно увеличивать любую координату μ_i и при этом неравенство (14.12) должно сохраняться.

Так как $\alpha\neq 0$, то будем далее считать, что

$$\sum_{i=0}^m \alpha_i\;=\;1. \tag{14.13}$$

г) Покажем, что выполнены условия дополняющей нежесткости:

$$\alpha_i\, f_i(x_0) = 0 \qquad \forall\, i = 1, ..., m.$$

Достаточно рассмотреть произвольный индекс i_0, для которого $f_{i_0}(x_0) < 0$.

Возьмем в пространстве $\mathbb{R}^{m+1}$ точку $\mu = (0, ..., 0, f_{i_0}(x_0), 0, ..., 0) + (\delta, \delta, ..., \delta)$, где $\delta > 0$. Так как все $f_i(x_0) \le 0$, то $f_i(x_0) < \mu_i$, поэтому $\mu \in C$. Для нее с учетом (14.13) имеем $(\alpha, \mu) = \alpha_{i_0} f_{i_0}(x_0) + \delta \ge 0$. В пределе при $\delta \to 0+$ получаем $\alpha_{i_0} f_{i_0}(x_0) \ge 0$. Так как $f_{i_0}(x_0) < 0$, то не может быть $\alpha_{i_0} > 0$, поэтому $\alpha_{i_0} = 0$, что и требовалось доказать.

д) Покажем, что функция Лагранжа $L = \sum \alpha_i f_i$ имеет в точке x_0 минимум по множеству A. Действительно, $\forall\, x \in A$, $\forall\, \delta > 0$ вектор μ с координатами $\mu_i = f_i(x) + \delta$ принадлежит C, поэтому $(\alpha, \mu) = \sum \alpha_i f_i(x) + \delta = L(x) + \delta \ge 0$, а тогда и $L(x) \ge 0$. Но в силу условия дополняющей нежесткости $L(x_0) = 0$, поэтому $L(x) \ge L(x_0)$. Утверждение 1 теоремы доказано.

2) Так как $\alpha_0 > 0$, то можно считать $\alpha_0 = 1$, и тогда $L = f_0 + \sum_{i=1}^m f_i$.
По условию, $\forall\, x \in A \quad L(x) \ge L(x_0)$, т.е.

$$f_0(x) + \sum \alpha_i f_i(x) \;\ge\; f_0(x_0) + \sum \alpha_i f_i(x_0) = f(x_0). \tag{14.14}$$

Конечное равенство справедливо из-за того, что последняя сумма равна 0 в силу условий дополняющей нежесткости.

Если точка x удовлетворяет еще и ограничениям неравенства $f_i(x) \le 0$, то $f_0(x) \ge f_0(x) + \sum \alpha_i f_i(x)$, поэтому для любой допустимой точки x, учитывая (14.14), получаем $f_0(x) \ge f_0(x_0)$, а это и означает, что $x_0 -$ точка минимума.

3) Допустим, что указанная точка $\hat{x}$ и множество α_i существуют, но при этом $\alpha_0 = 0$. Тогда $\sum_{i=1}^m \alpha_i > 0$, и следовательно,

$$L(\hat{x}) \;=\; \sum_{i=1}^{m} \alpha_i\, f_i(\hat{x}) < 0,$$

в то время как $L(x_0) = 0$ в силу условий дополняющей нежесткости, что противоречит минимуму L в точке x_0. Теорема доказана. □

Теорема Куна–Таккера при выполнении условия Слейтера представляет собой реализацию т.н. *принципа снятия ограничений*, состоящего в том, что вместо поиска минимума в задаче (14.8) с ограничениями $f_i \le 0$ можно искать минимум в задаче

$$L(x) = \; f_0(x) \;+\; \sum_{i=1}^{m} \alpha_i\, f_i(x) \to \min, \qquad x \in A, \tag{14.15}$$

где эти ограничения уже отсутствуют: задающие их функции прибавлены к целевой функции с некоторыми неотрицательными коэффициентами.

Этот принцип был высказан Лагранжем в конце XVIII века (для задач с ограничениями равенства, возникающих из классической механики) и в сформулированном здесь виде справедлив только для выпуклых задач. (Для невыпуклых задач надо искать не минимум, а стационарные точки задачи (14.15), т.е. точки, удовлетворяющие лишь необходимому условию минимума.)

Теорема Куна-Таккера справедлива и для общего класса выпуклых задач, включающего также и ограничения равенства:

$$\begin{cases} f_0(x) \to \min, \\ f_i(x) \le 0, \qquad i = 1, ..., m, \\ g_i(x) = 0, \qquad j = 1, ..., s, \\ x \in A, \end{cases} \tag{14.16}$$

Функции g_i здесь должны быть аффинными, иначе ограничения равенства задают, вообще говоря, невыпуклое множество. Таким образом, $g_j(x) = (a_j, x) + c_j$, где $a_j \in \mathbb{R}^n$, $c_j \in \mathbb{R}$. Без нарушения общности можно считать, что векторы $a_1, ..., a_s$ линейно независимы. Остальные функции и множество A те же, что и в теореме 14.2 (т.е. выпуклы). Для такой задачи справедлива модификация теоремы 14.2, состоящая в том, что функция Лагранжа теперь имеет вид

$$L(x) = \sum_{i=0}^{m} \alpha_i f_i(x) + \sum_{1}^{s} \beta_j g_j(x),$$

где $\beta_j \in \mathbb{R} -$ множители Лагранжа, соответствующие ограничениям равенства. Обратим внимание, что знак множителей β_j может быть любым, тогда как множители при ограничениях неравенства всегда имеют определенный знак: $\alpha_i \ge 0$.

Условия нетривиальности и дополняющей нежесткости остаются прежними, а условие Слейтера становится, естественно, таким: $\exists\, \hat{x} \in A$, для которых $f_i(\hat{x}) < 0$ при всех $i = 1, ..., m$, и $g_j(\hat{x}) = 0$ при всех $j = 1, ..., s$.

Доказательство новой теоремы проводится по прежней схеме. (Проведите его!).

Экономическая интерпретация.

Выпуклые задачи широко применяются в моделях математической экономики. Более того, это основной класс задач, который там используется. Правда, вместо выпуклых функций там, как правило, используются вогнутые функции (и рассматривается задача их максимизации).

Рассмотрим следующую задачу:

$$\begin{cases} f_0(x) \to \max, \\ f_i(x) \ge c_i, \qquad i = 1, ..., m, \\ x \in A, \end{cases} \tag{14.17}$$

где $A \subset \mathbb{R}^n$ – непустое выпуклое множество, а $f_i : \mathbb{R}^n \to \mathbb{R}, \quad i = 0, 1, ..., m$ – вогнутые функции. (Как правило, они предполагаются возрастающими относительно любой координаты вектора x.)

Точки $x \in A$ будем интерпретировать как технологические способы (параметры) производства некоторого предприятия; множество A при этом называется технологическим множеством. (Одна из фундаментальных гипотез математической экономики состоит в том, что технологические способы можно смешивать в любых пропорциях, и поэтому технологическое множество должно быть выпуклым.)

При данном способе производства $x \in A$ величина $f_i(x)$ означает объем выпуска i-го продукта, измеренного в некоторых единицах. Всего выпускается m продуктов.

Величина c_i (параметр задачи) означает тот объем i-го продукта, который данное предприятие обязано поставить некоторому заказчику (скажем, государству). Величина $f_0(x)$ – это доход предприятия при данном технологическом способе x.

Задача предприятия состоит в максимизации своего дохода при заданном ограничении на способы производства $(x \in A)$ и выполнении всех обязательных поставок ($f_i(x) \geq c_i$).

Будем считать, что условие Слейтера выполнено. Тогда достижение максимума в некоторой точке x_0 эквивалентно тому, что существуют множители $\alpha_0 = 1$, $\alpha_1 \geq 0$, ..., $\alpha_m \geq 0$, удовлетворяющие условиям дополняющей нежесткости, такие что x_0 есть решение задачи

$$L(x) = f_0(x) + \sum_{i=1}^{m} \alpha_i \left(f_i(x) - c_i\right) \to \max, \qquad x \in A. \tag{14.18}$$

Эту новую задачу можно интерпретировать следующим образом. Предприятие не обязано производить заданные объемы продуктов c_i. Но если оно произведет $i-$ го продукта меньше c_i, то его прибыль уменьшится на величину $\alpha_i(c_i - f_i(x))$, а если произведет больше c_i, то его прибыль увеличится на $\alpha_i(f_i(x) - c_i)$. Тогда коэффициент α_i играет роль рыночной цены $i-$ го продукта: при недостаточном выпуске $f_i(x) < c_i$ предприятие, чтобы выполнить план поставок, как бы докупает недостающую часть i-го товара на рынке по цене α_i, за счет чего и происходит уменьшение прибыли; при избыточном же выпуске $f_i(x) > c_i$ предприятие может продать излишек на рынке по цене α_i, увеличив тем самым свою прибыль.

Пусть x_0 есть решение задачи (14.17). Теорема Куна–Таккера утверждает, что существуют такие цены α_i, при которых предприятию можно поставить задачу (14.18) без жесткого указания на объемы произведенной продукции, при решении которой ему, тем не менее, будет выгодно выбрать ту же технологию x_0 и, таким образом, произвести требуемые объемы продукции.

Эта интерпретация множителей Лагранжа α_i как неких воображаемых, идеальных цен была предложена советским математиком Л.В. Канторовичем в конце 1930-х годов при изучении задач линейного программирования (когда все функции линейны, а множество A полиэдрально); они были названы им *объективно обусловленными оценками (О.О.О.)*. В зарубежной литературе коэффициенты α_i называются *теневыми ценами* (shadow prices).

Другой способ интерпретации множителей α_i как цен состоит в следующем. Будем рассматривать не одну задачу (14.17), а семейство задач с векторным параметром $c = (c_1, ..., c_m)$, меняющимся в некотором открытом множестве $\mathcal{O} \in \mathbb{R}^m$ (например, в окрестности некоторой точки $\hat{c}$). Допустим, что $\forall\, c \in \mathcal{O}$ задача имеет единственное решение $x(c)$ и единственный набор множителей Лагранжа $\alpha(c) = (\alpha_1(c), ..., \alpha_m(c))$ с множителем при целевой функции $\alpha_0(c) \equiv 1$. Более того, предположим, что $x(c)$ и $\alpha(c)$ гладким образом зависят от $c \in \mathcal{O}$. (Эти предположения довольно сильные, но в некоторых типичных случаях они выполняются.)

Обозначим через $V(c) = f_0(x(c))$ максимальное значение функционала в задаче (14.17). Тогда в силу условий дополняющей нежесткости имеем

$$V(c) = \; f_0(x(c)) + \sum_{i=1}^{m} \alpha_i(c)\, (f_i(x(c)) - c_i) \; = \; L(x(c), \alpha(c), c).$$

Найдем производную функции V :

$$V'(c) = \; L'_x\, x'(c) \; + \; L'_\alpha\, \alpha'(c) \; + \; L'_c\, .$$

Но $L'_x(x(c)) = 0$ в силу стационарности функции Лагранжа, и, очевидно, $L'_c = -\alpha$.

Рассмотрим средний член

$$L'_\alpha\, \alpha'(c) = \; \sum_{i=1}^{m} L'_{\alpha_i}\, \alpha'_i(c) \; = \; \sum_{i=1}^{m} (f_i(x(c)) - c_i)\, \alpha'_i(c)\; . \tag{14.19}$$

Если индекс i активный, т.е. $f_i(x(c)) - c_i = 0$, то i -го слагаемого здесь нет; если же индекс i неактивный, то он остается неактивным в целой окрестности данной точки c, поэтому в этой окрестности $\alpha_i \equiv 0$ и, следовательно, $\alpha'_i(c) = 0$, поэтому опять данное слагаемое пропадает.

Таким образом, вся сумма (14.19) равна 0, и тогда

$$V'(c) \; = \; -\,\alpha(c).$$

Мы опять пришли к тому, что $\alpha_i(c)$ есть цена i -го ресурса: при изменении планового задания на Δc прибыль предприятия изменяется в первом порядке на величину $\Delta V(c) = -\alpha(c)\, \Delta c$.

Оглавление