

Содержание

1	Лекция 1. Введение в математическую оптимизацию	3
1.1	Содержание курса	3
1.2	Обозначения	5
1.3	Дополнение: обозначения для матриц и векторов	8
2	Лекция 2. Метод Ньютона. Two-slope тест	9
2.1	Введение	9
2.2	Необходимый критерий	10
2.3	Достаточный критерий второго порядка	10
2.4	Идея метода Ньютона	11
2.5	Минимизация с использованием Two-Slope-Test	12
2.6	Анализ метода Ньютона	14
2.7	Выпуклые множества	14
2.8	Оболочки	15
3	Лекция 3. Выпуклые множества	16
3.1	Полупространства. Гиперплоскости	16
3.2	Полиэдры (многогранники)	17
3.3	Полиэдры: примеры	18
3.4	Суммы Минковского и масштабирование	18
3.5	Теоремы отделимости для выпуклых множеств	18
3.6	Топологическое замыкание	21
4	Лекция 4. Конусы	22
4.1	Определение	22
4.2	Свойства конусов	22
4.3	Важные примеры конусов	23
4.4	Теорема отделимости для выпуклых конусов	24
4.5	Конические оболочки	24
4.6	Выпуклые конические оболочки	25
4.7	Конечнопорожденные конусы	25
4.8	Теорема Каратеодори	25
5	Лекция 5. Полярные конусы	27
5.1	Определение	27
5.2	Свойства полярных конусов	27
5.3	Поляр пересечений	28

6	Лекция 6. Полиэдральный конус	32
6.1	Полиэдральный конус	32
6.2	Свойства полиэдрального конуса	32
6.3	Теоремы и леммы про полиэдральные конусы	32
6.4	Усиление Леммы 2.44	33
7	Лекция 7. Теорема о поляре полиэдрического конуса	37
7.1	Теорема о поляре полиэдрического конуса	37
7.2	Лемма Фаркаша	38
7.3	Теорема о матрице	39
7.4	Глава 3. Оптимальные условия для выпуклой задачи оп- тимизации	40
8	Лекция 8. Проблема выпуклой оптимизации	40
8.1	Постановка задачи	40
8.2	Радиальные и нормальные конусы	41
8.3	Радиальные и нормальные конусы для полиэдра	42
8.4	Оптимизация дифференцируемой функции на выпуклом множестве.	42
8.5	Некоторые нормальные конусы	44
8.6	Условия Каруша-Куна-Таккера. Постановка задачи	45
8.7	Нормальный конус для регулярной тройки	45
9	Лекция 9. Дополнительное условие выпуклости и диффе- ренцируемости, ККТ	46
9.1	Нормальный конус для регулярной тройки	46
9.2	Дополнительное условие выпуклости и дифференцируемости	47
9.3	Условия Каруша - Кунна - Такера (дифференцируемость, выпуклость)	49
9.4	Условия Каруша - Кунна - Такера для задачи линейной оптимизации (1 вариант)	49
9.5	ККТ для задачи линейной оптимизации (2 вариант)	50

1 Лекция 1. Введение в математическую оптимизацию

1.1 Содержание курса

1. Оптимизация без дополнительных условий (ограничений)
2. Выпуклые множества и конусы
3. Условия оптимальности для выпуклых задач
4. Двойственность и коническая оптимизация
5. Полиномиальные методы выпуклой оптимизации
6. Геометрия линейной оптимизации
7. Симплекс-метод
8. Целочисленная и комбинаторная оптимизация

Глава 1. Оптимизация без дополнительных условий (ограничений)

Структура лекции

1. Функции многих переменных
 - (a) Отображение
 - (b) Глобальный и локальный экстремум
 - (c) Выпуклые и вогнутые функции
2. Дифференцируемые функции
 - (a) Производные первого порядка
 - (b) Производные второго порядка
 - (c) Квадратичные функции. Определение матрицы
3. Критерий оптимальности для дифференцируемых функций
 - (a) Необходимые критерии
 - (b) Достаточные критерии
4. Метод Ньютона

1.2 Обозначения

Задача оптимизации: $\min \{f(x)|x \in X\}$ или $\max \{f(x)|x \in X\}$.

(Не обязательно может быть принято.)

Написание: $\min f(x) \max f(x)$, где $x \in X$.

$X \subseteq \mathbb{R}^n (n \geq 1)$: Область допустимых значений (область определения).

Целевая функция: $f : X \rightarrow \mathbb{R} \cup \{-\infty, \infty\}$

Оптимальное значение: $\inf \{f(x)|x \in X\}$ или $\sup \{f(x)|x \in X\}$ (из $\mathbb{R} \cup \{-\infty, \infty\}$)

Оптимальное решение (оптимальная величина):

$x^* \in X$, $f(x) = \text{Опт.Величина}$, т.е. $f(x^*) \leq f(x)$ или $f(x^*) \geq f(x)$ для всех $x \in X$

Недопустимая задача оптимизации: $X = \emptyset$.

Неограниченная задача оптимизации: Оптимальное значение $= +\infty$ или $= -\infty$.

(В частности, неограниченная задача допустима).

Определение 1.1 График функции – это такая область, что $\{(x, f(x))|x \in X\} \subset \mathbb{R}^{n+1}$.

Определение 1.2 Поверхность уровня (линии уровня для $n=2$) функции f для величины (значения) a – это такая область, что $\{x \in X|f(x) = a\} \subset \mathbb{R}^n$

Определение 1.3 Функция f принимает в точке x^* свой глобальный минимум на X , если условие $f(x^*) \leq f(x)$ выполняется для всех $x \in X$.

- Минимум часто называют глобальным минимумом.
- Оптимальное решение для $\min \{f(x)|x \in X\}$ – это в точности те точки, в которых функция принимает свое минимальное значение на области X .
- Функция может принимать минимальное значение в бесконечном числе точек (например, постоянная функция – константа или $f(x_1, x_2) = \cos(x_1) \sin(x_2)$)
- Функция может также и вовсе не принимать свое минимальное значение (например, постоянная линейная функция)

Определение 1.4 Функция принимает в $x^* \in X$ свое (локальное) минимальное значение, если $\exists \varepsilon > 0$, что $f(x^*) \leq f(x)$, $\forall x \in X$ что $\|x - x^*\| < \varepsilon$.

- Глобальный минимум является также и локальным минимумом.
- Не всякий локальный минимум является глобальным минимумом.

- Функция может принимать бесконечно много локальных минимумов, либо не принимать их вообще.

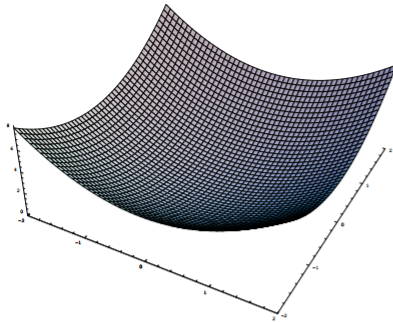
Определение 1.5 Функция называется (строго) выпуклой, если для всех $x, y \in \mathbb{R}^n$ (при $x \neq y$) и $0 < \lambda < 1$ выполняются условия:

$$f(\lambda x + (1 - \lambda)y) \leq (<) \lambda f(x) + (1 - \lambda)f(y)$$

Определение 1.6 Функция называется (строго) вогнутой, если для всех $x, y \in \mathbb{R}^n$ (при $x \neq y$) и $0 < \lambda < 1$ выполняются условия:

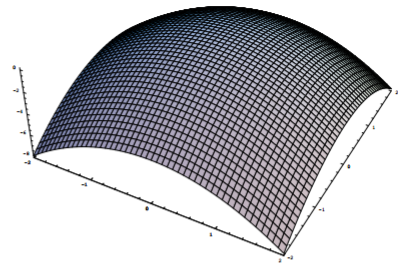
$$f(\lambda x + (1 - \lambda)y) \geq (>) \lambda f(x) + (1 - \lambda)f(y)$$

Наблюдение 1.7 Функция f является выпуклой(вогнутой) тогда и только тогда, когда ее график для всех точек $x, y \in \mathbb{R}^n$ лежит под(над) (или возможно на) отрезке, соединяющем две точки $(x, f(x))$ и $(y, f(y))$.



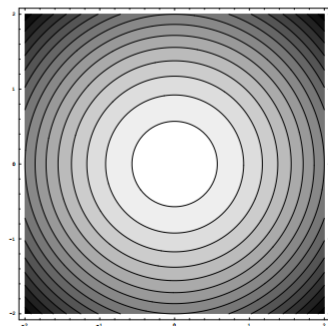
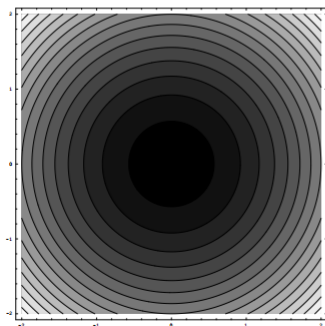
$$f(x_1, x_2) = x_1^2 + x_2^2$$

streng konvex



$$f(x_1, x_2) = -x_1^2 - x_2^2$$

streng konkav



Предложение 1.8 Если выпуклая(вогнутая) функция принимает в точке значение локального минимума(максимума), то она также принимает

значение глобального минимума(максимума) в этой точке.

Замечание 1.9 Если у выпуклой(вогнутой) функции есть глобальный минимум(максимум), то существует единственная (единственно определенная) точка, в которой она его принимает.

Предложение 1.10 Функция одновременно и выпуклая и вогнутая тогда и только тогда, когда она аффинная.

Определение 1.11 Функция называется аффинной, если она представляется в виде суммы линейной и постоянной функций.

$a \in \mathbb{R}^n, \beta \in \mathbb{R}$:

$$f(x) = \langle a, x \rangle + \beta$$

для $\forall x \in X$.

Определение 1.12 Скалярным произведением двух векторов $v, w \in \mathbb{R}^n$ называется

$$\langle v, w \rangle = \sum_{i=1}^n v_i w_i$$

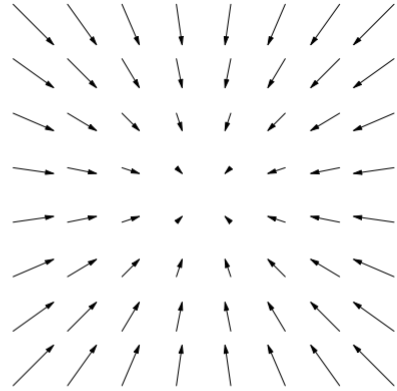
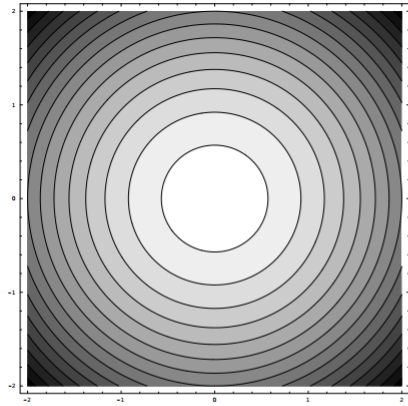
Определение 1.13 Если функция дифференцируема в некоторой точке $x^* \in X$, то градиентом этой функции в точке x^* называется

$$\text{grad}_{x^*} f := \left(\frac{\partial f}{\partial x_1}(x^*), \frac{\partial f}{\partial x_2}(x^*), \dots, \frac{\partial f}{\partial x_n}(x^*) \right)$$

- Если $v \in \mathbb{R}^n, \|v\| = 1$, то $\langle \text{grad}_{x^*} f, v \rangle \in \mathbb{R}$ — это увеличение (рост) f в точке x^* в направлении v .
- $(\|v\| := \sqrt{\sum_{i=1}^n v_i^2})$ Евклидова норма
- Аффинная функция $\varphi(x) := \langle \text{grad}_{x^*} f, x - x^* \rangle + f(x^*)$ локально хорошо аппроксимирует (приближает) f в точке x^* :

$$\lim_{x \rightarrow x^*} \frac{f(x) - \varphi(x)}{\|x - x^*\|} = 0$$

Замечание 1.14 Если f дифференцируема в точке $x^* \in \text{int}(X)$, $\alpha = \text{grad}_{x^*} f$ ортогонален линии уровня функции f с параметром α и показывает направление наибольшего роста функции f в точке x^* .



$$f(x_1, x_2) = -x_1^2 - x_2^2$$

Определение 1.15 Если Функция f дважды дифференцируема в точке $x^* \in \text{int}(X)$, то матрицей Гессе f в точке x^* называется матрица:

$$\text{hess}_{x^*} f := \begin{pmatrix} \frac{\partial^2 f}{\partial x_1 \partial x_1}(x^*) & \dots & \frac{\partial^2 f}{\partial x_1 \partial x_n}(x^*) \\ \dots & \dots & \dots \\ \frac{\partial^2 f}{\partial x_n \partial x_1}(x^*) & \dots & \frac{\partial^2 f}{\partial x_n \partial x_n}(x^*) \end{pmatrix}$$

- Если f дважды дифференцируема, то матрица $\text{hess}_{x^*} f$ симметричная.
- Квадратичная функция $q(x) := \frac{1}{2}(x-x^*)^T (\text{hess}_{x^*} f) (x-x^*) + \langle \text{grad}_{x^*} f, x - x^* \rangle + f(x^*)$ аппроксимирует функцию f локально хорошо в точке x^* :

$$\lim_{x \rightarrow x^*} \frac{f(x) - q(x)}{\|x - x^*\|^2} = 0$$

1.3 Дополнение: обозначения для матриц и векторов

- $[p] := \{1, 2, \dots, p\}$
- $\mathbb{R}^{[m] \times [n]}$: Область (пространство) действительных матриц $[m] \times [n]$.
- Особые матрицы:
 - $\mathbb{O}_p \in \mathbb{R}^p, \mathbb{O}_{[m] \times [n]}$: нуль-вектор, нуль-матрица.
 - $\mathbb{1}_p \in \mathbb{R}^p, \mathbb{1}_{[m] \times [n]}$: единичный вектор, единичная матрица.

- $\mathbb{I}_{[m] \times [n]}$: матрица инцидентности
- Единичный вектор (полагаю замечание относительно контекста имеет отношение к нему)
- $e_i = (0, \dots, 0, 1, 0, \dots, 0)$: i -ый единичный вектор
- Для $A \in \mathbb{R}^{[m] \times [n]}$, $i \in [m]$, $j \in [n]$:
 - $A_{(i,j)} = A_{ij} = a_{ij}$: значение матрицы на i строке, j столбце.
 - $A_{(i,*)} \in \mathbb{R}^n$: i -ая строка матрицы A
 - $A_{(*,j)} \in \mathbb{R}^m$: j -ый столбец матрицы A
- В терминах произведения матриц: отождествление \mathbb{R}^p с $\mathbb{R}^{[p] \times [1]}$ (вектор-столбцы)

2 Лекция 2. Метод Ньютона. Two-slope тест

2.1 Введение

Опр.1.16 Функция $q : \mathbb{R}^n \rightarrow \mathbb{R}$ называется **квадратичной**, если матрица $Q \in \mathbb{R}^{[n] \times [n]}$ симметричная и при $c \in \mathbb{R}^n$, $\gamma \in \mathbb{R}$ выполняется для всех $x \in \mathbb{R}^n$:

$$q(x) = x^T Q x + \langle c, x \rangle + \gamma$$

Опр.1.17 Матрица $Q \in \mathbb{R}^{[n] \times [n]}$ **положительно определена, положительно полуопределена, отрицательно определена, отрицательно полуопределена**, когда для всех $x \in \mathbb{R}^n \setminus \{0\}$ соответственно:

$$x^T Q x > 0,$$

$$x^T Q x \geq 0,$$

$$x^T Q x < 0,$$

$$x^T Q x \leq 0.$$

Матрица не определена, когда она ни положительно, ни отрицательно полуопределена.

Теорема 1.18 Симметрическая матрица $Q \in \mathbb{R}^{[n] \times [n]}$ положительно определена, положительно полуопределена, отрицательно определена, отрицательно полуопределена, когда все собственные значения Q положительны, не отрицательны, отрицательны, не положительны. Матрица не определена, когда она содержит и положительные, и отрицательные собственные значения.

Теорема 1.19 Пусть функция $q(x) : \mathbb{R}^n \rightarrow \mathbb{R}$ — квадратичная и $q(x) = x^T Q x + \langle c, x \rangle + \gamma$, $q(x)$ (строго) выпукла вниз/вверх, когда для всех $x \in \mathbb{R}^n$ матрица Q положительно/отрицательно полуопределена (определена).

Теорема 1.20 Дважды непрерывно дифференцируемая функция $f : \mathbb{R}^n \rightarrow \mathbb{R}$ выпукла вниз/вверх, когда для всех $x \in \mathbb{R}^n$ матрица Гессе положительно/отрицательно полуопределена. Если при этом матрица Гессе всюду положительно/отрицательно определена, то f строго выпукла вниз/вверх.

Замечание 1.21

Для квадратичной функции $q(x) = x^T Q x + \langle c, x \rangle + \gamma$ характер выпуклости полностью определяется частью с квадратами, так как для любого x^* :

1. $\text{grad}_{x^*} q = 2Qx^* + c$
2. $\text{hess}_{x^*} q = 2Q$

2.2 Необходимый критерий

Теорема 1.22 (необходимый критерий) Если функция $f : \mathbb{R}^n \supseteq X \rightarrow \mathbb{R}$ принимает локальный экстремум во внутренней точке $x^* \in \text{int}(X)$, в которой f дифференцируема, то $\text{grad}_{x^*} f = \mathbb{O}_n$.

Опр.1.23 Для каждой дифференцируемой функции $f : \mathbb{R}^n \supseteq X \rightarrow \mathbb{R}$ (X открыто) точка $x^* \in X$, для которой $\text{grad}_{x^*} f = \mathbb{O}_n$, называется **стационарной точкой**.

Следствие 1.24 Стационарные точки — единственные кандидаты на локальный экстремум дифференцируемой функции (для открытых множеств).

2.3 Достаточный критерий второго порядка

Теорема 1.25 (достаточный критерий второго порядка) Пусть $f : \mathbb{R}^n \supseteq X \rightarrow \mathbb{R}$ (X открыто) дважды непрерывно дифференцируема на X и пусть $x^* \in X$ — стационарная точка f . Тогда:

1. Если $\text{hess}_{x^*} f$ положительно определена, то f принимает в x^* локальный минимум;

2. Если $\text{hess}_{x^*} f$ отрицательно определена, то f принимает в x^* локальный максимум;
3. Если $\text{hess}_{x^*} f$ не определена, то в x^* нет локального экстремума.

2.4 Идея метода Ньютона

- Пусть $f : \mathbb{R}^n \rightarrow \mathbb{R}$ дважды непрерывно дифференцируема на \mathbb{R}^n и:
 - $\text{hess}_x f$ положительно определена для всех $x \in \mathbb{R}^n$
 - f имеет глобальный минимум, который также достигается в единственной стационарной точке $x^{min} \in \mathbb{R}^n$
- Стратегия: генерируем последовательность $x^{(0)}, x^{(1)}, \dots \in \mathbb{R}^n$, причем $\lim_{k \rightarrow \infty} x^{(k)} = x^{min}$
- Для всех k :
 - выбираем необходимое направление $d^{(k)} \in \mathbb{R}^n$
 - определяем $x^{(k+1)}$ как (приблизительное) оптимальное решение одномерной проблемы минимума $\min\{f(x^{(k)} + td^{(k)}) \mid t \in \mathbb{R}\}$.

Определение направления $d^{(k)}$

✓ Квадратичная функция

$$q^{(k)}(x) := \frac{1}{2}(x - x^{(k)})^T (\text{hess}_{x^{(k)}} f) (x - x^{(k)}) + \langle \text{grad}_{x^{(k)}} f, x - x^{(k)} \rangle + f(x^{(k)}) =$$

$$= x^T Q^{(k)} x + \langle c^{(k)}, x \rangle + \gamma^{(k)}$$

где

$$Q^{(k)} = \frac{1}{2} \text{hess}_{x^{(k)}} f$$

$$c^{(k)} = \text{grad}_{x^{(k)}} f - (\text{hess}_{x^{(k)}} f) x^{(k)}$$

$$\gamma^{(k)} = f(x^{(k)}) + \frac{1}{2} x^{(k)T} (\text{hess}_{x^{(k)}} f) x^{(k)}$$

хорошо приближает функцию f в окрестности $x^{(k)}$.

✓ функция $q^{(k)}$ строго выпукла с единственной точкой минимума $\tilde{x}/$

$$\checkmark d^{(k)} = \tilde{x} - x^{(k)}$$

$$\checkmark \mathbb{O}_n = \text{grad}_{\tilde{x}} q = 2Q^{(k)} \tilde{x} + c^{(k)}$$

✓ $Q^{(k)}$ положительно определена и обратима, то после подстановки получаем:

$$\tilde{x} = x^{(k)} - (hess_{x^{(k)}} f)^{-1} grad_{x^{(k)}} f \text{ значит } d^{(k)} = -(hess_{x^{(k)}} f)^{-1} grad_{x^{(k)}} f$$

2.5 Минимизация с использованием Two-Slope-Test

✓ Пусть $x^{(k)}$ не стационарная точка f .

✓ Пусть $\phi_k : \mathbb{R}^n \rightarrow \mathbb{R}$ такая, что $\phi_k(t) = f(x^{(k)} + td^{(k)})$.

✓ $d^{(k)}$ - направление спуска, значит $\phi'_k(0) < 0$.

✓ Идеальный случай: $x^{(k+1)} = x^{(k)} + \tau_k^* d^{(k)}$ для аргумента минимизации τ_k^* задачи $\min(\phi_k(t) | t \in \mathbb{R})$

✓ Вместо этого: выберем исходные параметры $1 > \alpha_1 > \frac{1}{2} > \alpha_0 > 0$ а затем найдем такой $x^{(k+1)} = x^{(k)} + \tau_k$ для $\tau_k > 0$ такого, что условия Two-Slope выполняются

$$\phi_k(0) + \alpha_1 \phi'_k(0) \tau \leq \phi_k(\tau) \leq \phi_k(0) + \alpha_2 \phi'_k(0) \tau$$

✓ Определяем при помощи проверки такую $u = 1, 2, 4, 8, \dots$, что $\phi_k(0) + \alpha_1 \phi'_k(0) u \leq \phi_k(u)$

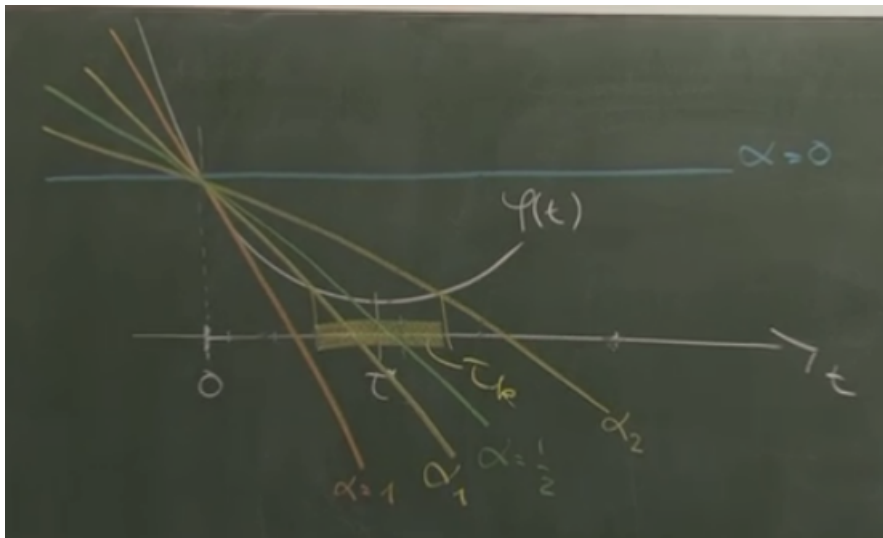
✓ В случае $\phi_k(u) \leq \phi_k(0) + \alpha_2 \phi'_k(0) u : \tau_k = u$

✓ В противном случае $l = 0$ и делаем тоже самое для $\phi_k(l) \leq \phi_k(0) + \alpha_2 \phi'_k(0) l$

✓ Проверяем оба неравенства из условий Two-Slope для $\tau = \frac{u-l}{2}$. Если выполняются оба, то $\tau_k = \tau$

Если не выполняется первое, то $l = \tau$ и снова проверяем условия для пересчитанного τ

Если не выполняется второе, то $u = \tau$ и снова проверяем условия для пересчитанного τ



Замечание Метод находит нужное τ за конечное число шагов

2.6 Анализ метода Ньютона

Предложение 1.26 Пусть $f : \mathbb{R}^n \rightarrow \mathbb{R}$ дважды непрерывно дифференцируемая на \mathbb{R}^n со всюду положительно определенной матрице Гессе и $x^{(0)} \in \mathbb{R}^n$ такое, что множество $X^0 := \{x \in \mathbb{R}^n | f(x) \leq f(x^{(0)})\}$ ограничено, то сходится последовательность, генерируемая методом Ньютона и тестом Two-Slope к точке, в которой f достигает своего минимума.

Замечание Для всех достаточно больших k выполняется условие:

$$\|x^{(k+1)} - x^{min}\| \leq \frac{L}{2m} \|x^{(k)} - x^{min}\|^2$$

где L - константа Липшица для функции $x \rightarrow hess_x f$ и m - наименьшее собственное значение $hess_x f$ на X^0 . В частности, при хорошем выборе начальной точки получается квадратичная сходимость.

Лекция 2: выпуклые множества и конусы

2.7 Выпуклые множества

Опр.2.1 Множество $X \subseteq \mathbb{R}^n$ называется **выпуклым** в случае, если для любых $x, y \in X$: $\lambda x + (1 - \lambda)y \in X$ для всех $0 \leq \lambda \leq 1$.

Множество выпуклое, если оно с любыми двумя своими точками содержит отрезок их соединяющий.



konvex



nicht konvex

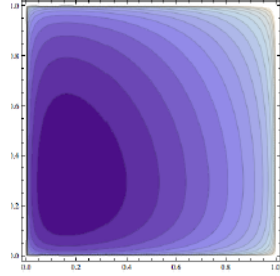
Выпуклые функции над выпуклыми множествами

Опр.2.2 Пусть $X \subseteq \mathbb{R}^n$ выпукло. Функция $f : X \rightarrow \mathbb{R}$ называется выпуклой вниз(вверх), если для всех $x, y \in X$, $0 \leq \lambda \leq 1$:

$$f(\lambda x + (1 - \lambda)y) \leq (\geq) \lambda f(x) + (1 - \lambda)f(y)$$

Замечание 2.3 Если выпуклая вниз/вверх функция на выпуклом множестве принимает в точке локальный минимум/максимум, то она там также принимает глобальный минимум/максимум.

Наблюдение 2.4 Пусть функция $f : \mathbb{R}^n \rightarrow \mathbb{R}$ выпуклая, то для каждого $\alpha \in \mathbb{R}$ множество $\{x \in \mathbb{R}^n : f(x) \leq \alpha\}$ выпукло.

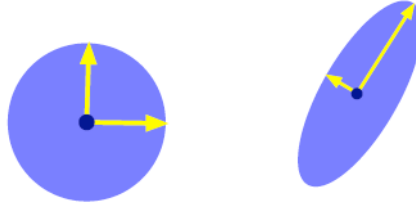


Эллипсоиды

Замечание 2.5 Для положительно определенной симметричной матрицы $Q \in \mathbb{R}^{n \times n}$ и $z \in \mathbb{R}^n$ определяется выпуклый и компактный эллипсоид с центром в z :

$$Ell(z, Q) := \{x \in \mathbb{R}^n | (x - z)^T Q^{-1} (x - z) \leq 1\}$$

Самым простым эллипсоидом является шар. Это эллипсоид с единичной матрицей. Полуоси эллипсоида равны корням из соответствующих собственных значений матрицы.



Наблюдение 2.6 Пусть I — множество индексов (любой мощности) и $X_i \subseteq \mathbb{R}^n$ — выпуклое множество ($i \in I$). Тогда $\bigcap_{i \in I} X_i$ также выпукло.

2.8 Оболочки

Выпуклые оболочки

Опр.2.7 Для $X \subseteq \mathbb{R}^n$ будем называть

$$\text{conv} X := \bigcap \{X^* \subseteq \mathbb{R}^n \mid X \subseteq X^*, X^* \text{ выпукло}\}$$

выпуклой оболочкой X .

Линейная оболочка:

$$\text{lin}X := \bigcap \{L \subseteq \mathbb{R}^n \mid X \subseteq L, L \text{ линейное подпространство}\}$$

Аффинная оболочка:

$$\text{aff}X := \bigcap \{A \subseteq \mathbb{R}^n \mid X \subseteq A, A \text{ аффинное подпространство}\}$$

Для $x^{(1)}, \dots, x^{(r)} \in \mathbb{R}^n$ и $\lambda_1, \dots, \lambda_r \in \mathbb{R}$

$$\sum_{i=1}^r \lambda_i x^{(i)}$$

линейная комбинация $x^{(1)}, \dots, x^{(r)}$.

- В случае $\sum_{i=1}^r \lambda_i = 1$: **аффинная комбинация**
- В случае $\lambda_1, \dots, \lambda_r \geq 0$: **коническая комбинация**
- Конические аффинные комбинации: **выпуклые комбинации**

Замечание 2.8 Выпуклая / линейная / аффинная оболочка $X \subseteq \mathbb{R}^n$ — это множество всех выпуклых / линейных / аффинных комбинаций (конечного числа) точек из X .

3 Лекция 3. Выпуклые множества

3.1 Полупространства. Гиперплоскости

Определение 2.9

Для $a \in \mathbb{R}^n \setminus \{\mathbf{0}_n\}$ и $\beta \in \mathbb{R}$ множества

$$H^{\leq}(a, \beta) := \{x \in \mathbb{R}^n : \langle a, x \rangle \leq \beta\}$$

и

$$H^=(a, \beta) := \{x \in \mathbb{R}^n : \langle a, x \rangle = \beta\}$$

называют определёнными (a, β) (аффинным) полупространством и (аффинной) гиперплоскостью соответственно. (В случае $\beta = 0$: линейными).

Замечание 2.10

- Полупространства являются выпуклыми и замкнутыми.
- Гиперплоскости являются выпуклыми.
- Аффинные подпространства являются выпуклыми.
- Пересечение любого количества полупространств является выпуклым.

3.2 Полиэдры (многогранники)

Определение 2.11

Подмножество $P \subseteq \mathbb{R}^n$ называют (выпуклым) полиэдром, если P является пересечением конечного количества аффинных полупространств.

- $P = \emptyset$ и $P = \mathbb{R}^n$ (пересечение по пустому индексному множеству) являются полиэдрами.
- Аффинные подпространства являются полиэдрами.

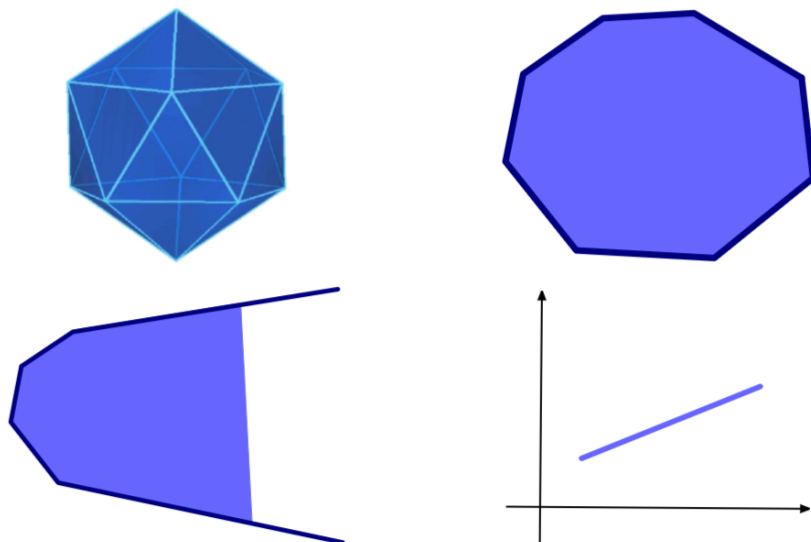
Замечание 2.12

- Полиэдры являются выпуклыми и (топологически) замкнутыми.
- Множество допустимых решений (допустимое множество)

$$P^{\leq}(A, b) := \{x \in \mathbb{R}^n : Ax \leq b\}$$

задачи линейной оптимизации является полиэдром.

3.3 Полиэдры: примеры



3.4 Суммы Минковского и масштабирование

Определение 2.13

Для множеств $X_1, \dots, X_q \subseteq \mathbb{R}^n$ выражение

$$\sum_{i=1}^q X_i = X_1 + \dots + X_q := \left\{ \sum_{i=1}^q x^{(i)} : x^{(i)} \in X_i, \forall i \in [q] \right\}$$

называют суммой Минковского множеств X_1, \dots, X_q .

Замечание 2.14

Суммы Минковского и масштабирование выпуклых множеств являются выпуклыми множествами.

$X \subseteq \mathbb{R}^n, \alpha \in \mathbb{R} : \alpha X := \{\alpha x \mid x \in X\}$ — масштабирование множества X .

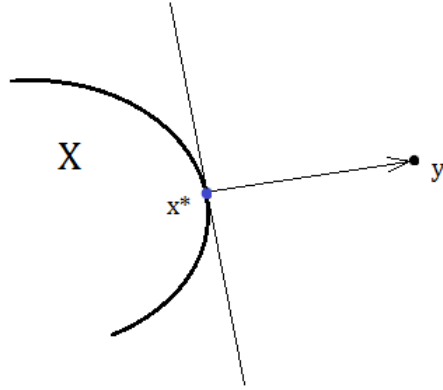
3.5 Теоремы отделимости для выпуклых множеств

Теорема 2.15 (Теорема о разделяющей гиперплоскости)

Пусть $X \subseteq \mathbb{R}^n$ является выпуклым и замкнутым и $y \in \mathbb{R}^n \setminus X$, тогда существуют такие $a \in \mathbb{R}^n \setminus \{0_n\}$ и $\varepsilon > 0$, что $\langle a, x \rangle \leq \langle a, y \rangle - \varepsilon$ для $\forall x \in X$.

◀ Сделаем допущение, что $X \neq \emptyset$. \emptyset замкнуто и выпукло, но для него теорема очевидно выполняется, т. к. $\forall x \in \emptyset$.

Итак, $X \neq \emptyset$. Выберем $a \in \mathbb{R}^n \setminus \{\mathbb{O}_n\}$, $\varepsilon > 0$



Будем строить гиперплоскость. Покажем, что существует минимум расстояния от y до X (обозначим его x^*). Для этого важно свойство замкнутости. Если бы X было ограниченным, то X было бы компактом. Пересечём X , если оно не ограничено, с шаром большего радиуса — получим компакт.

Пусть $x^* \in X$ такой, что $\|y - x^*\| = \min \{\|y - x\| : x \in X\}$. Причём, x^* существует, т. к., если выбрать $\bar{x} \in X \neq \emptyset$,

$$\bar{X} := \{x \in X : \|y - x\| \leq \|y - \bar{x}\|\}$$

$$\inf \{\|y - x\| : x \in X\} = \inf \{\|y - x\| : x \in \bar{X}\} \quad (*)$$

функция $x \mapsto \|y - x\|$ — непрерывна, \bar{X} — компакт, и эта функция принимает значения инфимумов (*).

Определим: $a = y - x^*$ и $\varepsilon = \|a\|^2$.

Покажем, что выбор a и ε корректен.

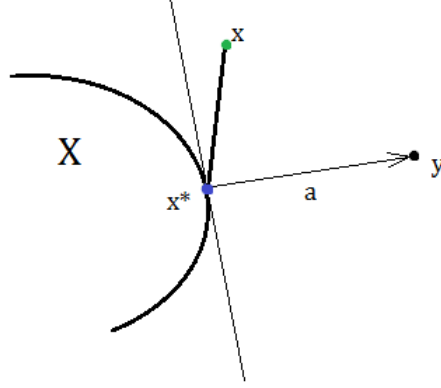
Предположим, что $\exists x \in X \mid \langle a, x \rangle > \langle a, y \rangle - \varepsilon$ (**), причём $\langle a, y \rangle - \varepsilon = \langle a, y \rangle - \langle a, a \rangle = \langle a, y - a \rangle = \langle a, x^* \rangle$.

Покажем, что этого не может быть.

Противоречие:

По соединяющему x^* и x отрезку точки должны иметь меньшее расстояние до y .

Посмотрим, что происходит с расстоянием от y до точек прямой при дви-



жении в направлении x .

Определим функцию $\phi : \mathbb{R} \rightarrow \mathbb{R}$, где $\phi(t) := \|y - (x^* + t(x - x^*))\|^2$, причём $x^* + t(x - x^*) = (1 - t)x^* + tx$.

Достаточно показать, что $\phi'(0) < 0$ потому, что если производная < 0 , то $\exists 0 < \tilde{t} \leq 1$ такое, что $\phi(0) > \phi(\tilde{t}) = \|y - \tilde{x}\|^2$, где $\tilde{x} = x^* + \tilde{t}(x - x^*) \in X$, т. к. X — выпуклое, а $\phi(0) = \|y - x^*\|^2$. А это противоречие с минимальностью x^* .

$$\phi'(t) = \frac{d}{dx} \langle y - (x^* + t(x - x^*)), y - (x^* + t(x - x^*)) \rangle = \frac{d}{dx} (\langle y - x^* \rangle^2 - 2\langle y - x^*, x - x^* \rangle t + \langle x - x^* \rangle^2 t^2) = -2\langle y - x^*, x - x^* \rangle + 2\langle x - x^* \rangle^2 t$$

$$\phi'(0) = -2\langle y - x^*, x - x^* \rangle = 2\langle a, x - x^* \rangle = 2(\langle a, x^* \rangle - \langle a, x \rangle) < 0,$$

т. к. из (***) вытекает, что $\langle a, x^* \rangle < \langle a, x \rangle$. ►

Лемма 2.16а

Пусть $X \subseteq \mathbb{R}^n$ — выпуклое, $y \in \mathbb{R}^n \setminus X$. Тогда существует $a \in \mathbb{R}^n \setminus \{\mathbb{O}\}$ что $\langle a, x \rangle \leq \langle a, y \rangle$, $\forall x \in X$

◀

- Достаточно построить последовательность $y^{(k)} \in \mathbb{R}^n \setminus cl(X)$, чтобы $\lim_{k \rightarrow \infty} y^{(k)} = y$.
 - По теореме 2.15 (т.к. $cl(X)$ замкнуто и выпукло) существует $a^{(k)} \in \mathbb{R}^n \setminus \{\mathbb{O}\}$, что $\langle a^{(k)}, x \rangle \leq \langle a^{(k)}, y^{(k)} \rangle$, $\forall k \in \mathbb{N}$, $\forall x \in cl(X)$
 - С помощью масштабирования сделаем $\|a^{(k)}\| = 1$, $\forall k \in \mathbb{N}$, по Теореме Больцано-Вейерштрасса существует $K \subseteq \mathbb{N}$ что

$\lim a^{(k)} = a \in \mathbb{R}^n$, при $k \rightarrow \infty$, $k \in K$

– Т.к. $\|a^{(k)}\| = 1$ и $\|\cdot\|$ - непрерывное отображение, то $\|a\| = 1$ и $\Rightarrow a \neq \mathbb{O}$

– Имеем $\forall x \in X$:

$$* \langle a^{(k)}, x \rangle \leq \langle a^{(k)}, y^{(k)} \rangle, \forall k \in \mathbb{N}$$

$$* \text{Т.к. скалярное произведение непрерывно} \Rightarrow \langle a, x \rangle \leq \langle a, y \rangle$$

- Существование последовательности $y^{(k)} \in \mathbb{R}^n \setminus cl(X)$, что $\lim_{k \rightarrow \infty} y^{(k)} = y$ следует из того, что для $k \in \{1, 2, \dots\}$ не все точки

$$v^{(0)} := y - \frac{1}{k} \mathbb{1}, v^{(1)} := y + \frac{1}{k} e_1, v^{(n)} := y + \frac{1}{k} e_n$$

лежат в $cl(X)$.

- Предположим, что $v^{(0)}, v^{(1)}, \dots, v^{(n)} \in cl(X)$, тогда $y = \sum_{i=1}^n \frac{1}{n+1} v^{(i)} \in int(cl(X)) \subseteq X$. Противоречие, т.к. $y \notin X$. ►

Терема 2.16

Пусть $X, Y \subseteq \mathbb{R}^n$ — выпуклые множества и $X \cap Y = \emptyset$. Тогда существует такой $a \in \mathbb{R}^n \setminus \{\mathbb{O}_n\}$, что $\langle a, x \rangle \leq \langle a, y \rangle$ для $\forall x \in X, y \in Y$.

◀

- Множества $X, Y \in \mathbb{R}^n$ -выпуклы и $X \cap Y = \emptyset$
- Множество $X + ((-1)Y) := X - Y$ - выпукло (т.к. сумма Минковского выпуклых множеств является выпуклым множеством) и $\mathbb{O} \notin X - Y$
- Из Леммы 2.16а $\Rightarrow \exists a \in \mathbb{R}^n \setminus \{\mathbb{O}\} : \langle a, x - y \rangle \leq \langle a, \mathbb{O} \rangle, \forall x \in X, \forall y \in Y$.
- Т.к. $\langle a, \mathbb{O} \rangle = 0$, то $\langle a, x - y \rangle \leq \langle a, \mathbb{O} \rangle \iff \langle a, x \rangle \leq \langle a, y \rangle$. ►

3.6 Топологическое замыкание

Замечание 2.17

Для любого выпуклого множества $X \subseteq \mathbb{R}^n$ его топологическое замыкание $cl(X)$ также является выпуклым.

Следствие 2.18

Топологическое замыкание любого выпуклого множества есть пересечение всех содержащих это множество полупространств.

Следствие 2.18

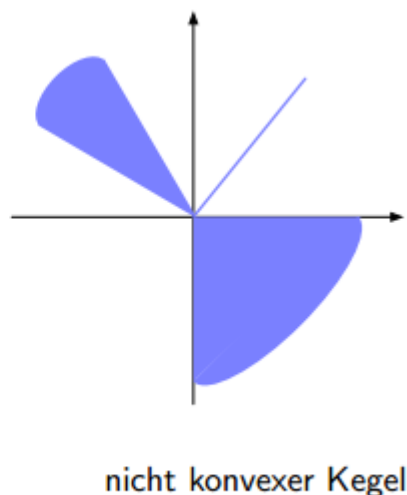
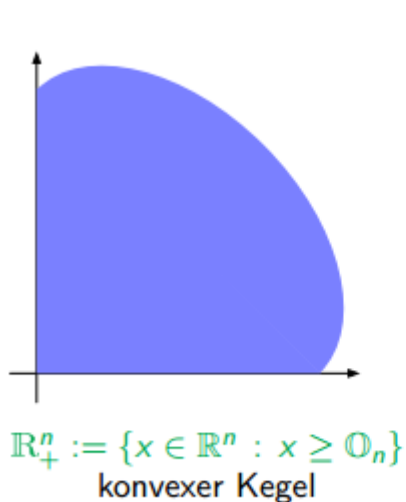
Пересечения (сколь угодно большого количества) полупространств являются замкнутыми выпуклыми множествами.

Пересечения конечного количества полупространств являются полиэдрами.

4 Лекция 4. Конусы

4.1 Определение

Определение 2.20 Подмножество $K \subseteq \mathbb{R}^n$ является **конусом**, когда $K \neq \emptyset$ и $\forall x \in K$ и $\alpha \geq 0$ выполняется $\alpha x \in K$.



4.2 Свойства конусов

Замечание 2.21

Пусть I является множеством индексов (произвольной мощности) и $K_i \subseteq \mathbb{R}^n$ - конусы ($i \in I$), тогда множество пересечений $\bigcap_{i \in I} K_i$ является конусом.

Замечание 2.22

Непустое множество $\emptyset \neq K \subseteq \mathbb{R}^n$ является выпуклым конусом тогда и

только тогда, когда K содержит все конические комбинации своих элементов.

◀ Требуется показать, что $\forall x, y \in K, \forall \lambda, \mu > 0 : \lambda x + \mu y \in K$

- В связи с выпуклостью K имеем: $\frac{\lambda}{\lambda + \mu}x + \frac{\mu}{\lambda + \mu}y \in K$
- $z := \frac{\lambda}{\lambda + \mu}x + \frac{\mu}{\lambda + \mu}y$
- Т.к. K - конус $\implies \lambda x + \mu y = (\lambda + \mu)z \in K$. ▶

4.3 Важные примеры конусов

- **Неотрицательный октант**

$$\mathbb{R}_+^n := \{x \in \mathbb{R}^n | x \geq \mathbb{O}^n\}$$

- **Конус симметрических неотрицательно определенных матриц**

$$\mathbb{S}_+^k := \{A \in \mathbb{S}^k | x^T A x \geq 0, \forall x \in \mathbb{R}^k\}$$

◀ $A, B \in \mathbb{S}_+^k, \lambda, \mu \geq 0$: матрица A положительно полуопределена, если

$$\forall x \in \mathbb{R}^k : x^T A x \geq 0$$

Коническая комбинация симметрических матриц является симметрической матрицей:

$$\implies x^T (\lambda A + \mu B) x = x^T \lambda A x + x^T \mu B x \geq 0$$

$\implies \mathbb{S}_+^k$ - выпуклый конус. ▶

- \mathbb{S}_+^k и \mathbb{R}_+^n выпуклые и замкнутые множества.

$$\blacktriangleleft \mathbb{S}_+^k = \{A \in \mathbb{S}^k : x^T A x \geq 0, \forall x \in \mathbb{R}^k\} = \bigcap_{x \in \mathbb{R}^k} \{A \in \mathbb{S}^k : x^T A x \geq 0\}$$

Определим отображение $\varphi_x : \mathbb{S}^k \longrightarrow \mathbb{R}$ как $\varphi_x(A) := x^T A x$ Имеем:

$$\{A \in \mathbb{S}^k : x^T A x \geq 0\} = \varphi_x^{-1}([0, \infty[)$$

Множество $[0, \infty[$ - замкнуто, прообраз замкнутого - замкнут, пересечение любого количества замкнутых множеств - замкнуто.

$\implies \mathbb{S}_+^k$ - замкнуто. ▶

4.4 Теорема отделимости для выпуклых конусов

Теорема 2.23

Пусть $K \subseteq \mathbb{R}^n$ есть замкнутый выпуклый конус и $y \in \mathbb{R}^n$ K - точка, находящаяся вне K , тогда есть $a \in \mathbb{R}^n$ что

$$\langle a, x \rangle \leq 0 \text{ для любого } x \in K \text{ и } \langle a, y \rangle = 1$$

◀

По Теореме 2.15

$$\implies \exists a \in \mathbb{R}^n \setminus \{\mathbb{O}_n\} : \langle a, x \rangle < \langle a, y \rangle \forall x \in K \quad (1)$$

Т.к. $\mathbb{O}_n \in K$, то $\langle a, y \rangle > 0$

$$\text{При масштабировании: } \left(a \longleftarrow \frac{1}{\langle a, y \rangle} a \right) \text{ получим } \langle a, y \rangle = 1 \quad (2)$$

Пусть существует $x \in K$ что $\langle a, x \rangle > 0$. По определению конуса $\forall \lambda > 0$ $\lambda x \in K$ и $\langle a, \lambda x \rangle = \lambda \langle a, x \rangle$.

Получили противоречие с условиями (1) и (2), т.к. можно выбрать $\lambda > 0$ при котором $\lambda \langle a, x \rangle > 1$. ▶

4.5 Конические оболочки

Определение 2.24

Пусть $X \subseteq \mathbb{R}^n$. Конической оболочкой множества X называют множество

$$\text{cone}X := \cap \{K \subseteq \mathbb{R}^n | X \subseteq K, K - \text{конус}\}$$

Замечание 2.25

$$\text{cone}X = \{\alpha x | x \in X, \alpha \geq 0\} \cup \{\mathbb{O}^n\}$$

Замечание 2.26

- Для любого $X \subseteq \mathbb{R}^n$ коническая оболочка $\text{cone}X$ является конусом.
- Для любого выпуклого множества X коническая оболочка $\text{cone}X$ является выпуклым конусом.

4.6 Выпуклые конические оболочки

Определение 2.27

Для $X \subseteq \mathbb{R}^n$ множество

$$\text{scope}X := \cap \{K \subseteq \mathbb{R}^n | X \subseteq K, K - \text{выпуклый конус}\}$$

является выпуклой конической оболочкой множества X .

Замечание 2.28

Для любого $X \subseteq \mathbb{R}^n$ выпуклая коническая оболочка $\text{scope}X$...

- является выпуклым конусом.
- есть множество всех конических комбинаций элементов из X .

4.7 Конечнопорожденные конусы

Определение 2.29

Конус называют **конечнопорожденным** для конечного множества $X \subseteq \mathbb{R}^n$, если он является выпуклой конической оболочкой данного множества

$$\text{scope}X = \left\{ \sum_{x \in X} \lambda_x x \mid \lambda_x \geq 0, \forall x \in X \right\}$$

Если $X \subseteq \mathbb{R}^n$ линейно-независимо, то $\text{scope}X$ называют **симплициальным** конусом.

Замечание 2.30

Конечнопорожденный конус является выпуклым.

4.8 Теорема Каратеодори

Теорема 2.31

Пусть $X \subseteq \mathbb{R}^n$ и $x \in \text{scope}X$, тогда есть линейно-независимое подмножество $X_1 \subseteq X$, что $x \in \text{scope}X_1$ (В частности $|X_1| \leq n$).

◀ Выберем $X_1 \subseteq X$ как минимальное по мощности подмножество X , что коэффициенты $\lambda_y \geq 0$ ($y \in X_1$) и $x = \sum_{y \in X_1} \lambda_y y$.

В частности $\lambda_y > 0 \forall y \in X_1$.

Предположим, что X_1 линейно-зависимо. Следовательно существует набор $\mu_y \in \mathbb{R}^n$ ($y \in X_1$), где не все μ_y равны нулю, и $\sum_{y \in X_1} \mu_y y = 0$

Смасштабируем μ_y так, чтобы

$$\max \left\{ \frac{\mu_y}{\lambda_y} : y \in X_1 \right\} = 1$$

Рассмотрим $x = \sum_{y \in X_1} \lambda_y y - \sum_{y \in X_1} \mu_y y = \sum_{y \in X_1} (\lambda_y - \mu_y) y$

В сумме $\sum_{y \in X_1} (\lambda_y - \mu_y) y$ присутствует на один вектор меньше, чем в изначальной. Получили противоречие с выбором минимальности X_1 . ►

Теорема 2.32

Конечнопорожденный конус является выпуклым и замкнутым.

◀

- Пусть $X \subseteq \mathbb{R}^n$, $|X| < \infty$, $K := \text{ccone}(X)$.
- K , очевидно, выпукло.
- По теореме 2.31 $\implies \text{ccone}(X) = \bigcup_{\hat{X} \subseteq X} \text{ccone}(\hat{X})$, где $\hat{X} \subseteq X$, \hat{X} -линейно-независимы.
- Т.к. конечное объединение замкнутых множеств - замкнуто, достаточно показать, что для каждого линейно-независимого подмножества $\hat{X} \subseteq \mathbb{R}^n$ симплицальный конус $\text{ccone}(\hat{X})$ замкнут.
- Дополним \hat{X} до базиса из \mathbb{R}^n как $\hat{X} \sqcup Y$ и определим линейное отображение $\varphi: \mathbb{R}^n \longrightarrow \mathbb{R}^{\hat{X}}$ как $\varphi(x) := e_x$, $\forall x \in \hat{X}$ и $\varphi(y) := 0$, $\forall y \in Y$
- Прообраз $\varphi^{-1}(\mathbb{R}_+^{\hat{X}})$ множества $\mathbb{R}_+^{\hat{X}}$ замкнуто, т.к. φ непрерывно и $\mathbb{R}_+^{\hat{X}}$ замкнуто.
- $\varphi^{-1}(\mathbb{R}_+^{\hat{X}}) = \left\{ \sum_{x \in \hat{X}} \lambda_x x + \sum_{y \in Y} \mu_y y : \lambda_x \geq 0, \forall x \in \hat{X}, \mu_y \in \mathbb{R}, \forall y \in Y \right\} = \text{ccone}(\hat{X}) + \text{lin}(Y)$
- Т.к. $\text{lin}(\hat{X}) \cap \text{lin}(Y) = \{0\}$: $\varphi^{-1}(\mathbb{R}_+^{\hat{X}}) \cap \text{lin}(\hat{X}) = \text{ccone}(\hat{X})$. Множества $\varphi^{-1}(\mathbb{R}_+^{\hat{X}})$ - замкнуто и $\text{lin}(\hat{X})$ - замкнуто.
- $\implies \text{ccone}(\hat{X})$ - замкнуто. ►

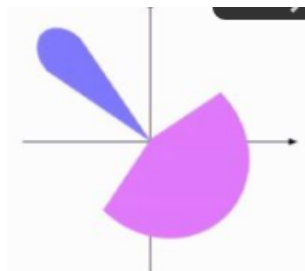
5 Лекция 5. Полярные конусы

5.1 Определение

Определение 2.33

Для конуса $K \subseteq \mathbb{R}^n$ полярным конусом называется множество

$$K^\circ := \{y \in \mathbb{R}^n : \langle y, x \rangle \leq 0, \forall x \in K\}$$



5.2 Свойства полярных конусов

- **Замечание 2.34** Если $K_1 \subseteq K_2$, то $K_2^\circ \subseteq K_1^\circ$
- **Замечание 2.35** Для конуса $K \subseteq \mathbb{R}^n$ $cl(K)$ является конусом, что выполняется $K^\circ = (cl(K))^\circ$
- **Замечание 2.36** Полярный конус является выпуклым замкнутым конусом.

- **Замечание 2.37** Для $X \subseteq \mathbb{R}^n$ выполняется

$$(ccone X)^\circ = \{y \in \mathbb{R}^n \mid \langle x, y \rangle \leq 0, \forall x \in X\}$$

Теорема 2.38 Для любого замкнутого выпуклого конуса K выполняется $K^{\circ\circ} = K$



- Очевидно, что $K \subseteq K^{\circ\circ}$.
- Допустим, что $y \in K^{\circ\circ} \setminus K$
- Т.к. K замкнутый и выпуклый конус:

По Теореме 2.23:

$$\implies \exists a \in \mathbb{R}^n \setminus \{0^n\} : \langle a, x \rangle \leq 0, \forall x \in K \quad (1)$$

$$\langle a, y \rangle = 1 > 0 \quad (2)$$

- Из (1) $\implies a \in K^\circ$
- Из (2) и того, что $y \in K^{\circ\circ} \implies a \notin K^\circ$

Получили противоречие. ►

5.3 Поляр пересечений

Теорема 2.39

Если $K_1, \dots, K_q \subseteq \mathbb{R}^n$ - выпуклые конусы и

$$K_1 \cap (\cap_{i=2}^q \text{int}(K_i)) \neq \emptyset, \quad (1)$$

тогда

$$(\cap_{i=1}^q K_i)^\circ = \sum_{i=1}^q K_i^\circ \quad (2)$$

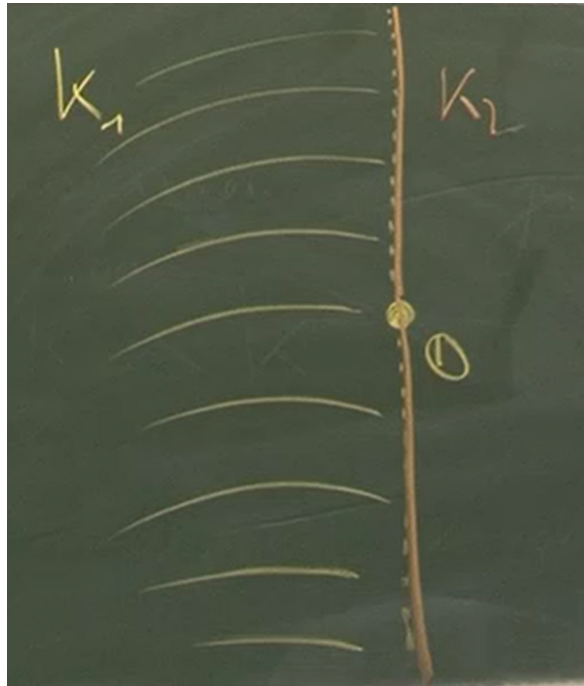
где $\text{int}(X)$ - множество внутренних точек из $X \subseteq \mathbb{R}^n$.

Мотивация теоремы 2.39



- $K_k^\circ \subseteq (\cap_{i=1}^q K_i)^\circ, \forall k \in [q]$
- Т.к. $(\cap_{i=1}^q K_i)^\circ$ - выпуклый конус $\implies \sum_{k=1}^q K_k^\circ \subseteq (\cap_{i=1}^q K_i)^\circ$
- $(\cap_{i=1}^q K_i)$ - очень небольшой конус, который содержит только все общие элементы всех конусов \implies у этого конуса огромный поляр, который будет содержать всю $\sum_{k=1}^q K_k^\circ$.

Частный случай Теоремы 2.39



- Первый конус задается как $K_1 = \{(0, 0)\} \cup \{(x, y) : x < 0\}$
- Второй конус: $K_2 = \{(0, y) : y \in \mathbb{R}\}$
- Рассмотрим пересечение этих конусов:

$$K_1 \cap K_2 = \emptyset$$

- Поляр пересечения равен всему пространству:

$$(K_1 \cap K_2)^\circ = \mathbb{R}^2$$

- Поляр к первому конусу: $K_1^\circ = \{(x, 0) : x \geq 0\}$
- Ко второму: $K_2^\circ = \{(x, 0) : x \in \mathbb{R}\}$
- Сумма Минковского поляр: $K_1^\circ + K_2^\circ = \{(x, 0) : x \in \mathbb{R}\}$
- Как видно, сумма Минковского полностью содержится в поляре пересечения.
- Однако, теорема не выполняется в следствие того, что нарушены условия: пересечение проходит только по нулевому вектору.

Переходим к доказательству теоремы 2.39: ◀

" \supseteq ": Очевидно, смотри мотивацию к теореме.

" \subseteq ": Необходимо показать, что поляр содержится в сумме Минковского:

- Это, очевидно, выполняется для нулевого вектора:

$$0 \in \sum_{i=1}^q K_i^\circ$$

- Достаточно для каждого $y \in (\bigcap_{i=1}^q K_i)^\circ \setminus \{0\}$ показать, что

$$y \in \sum_{i=1}^q K_i^\circ$$

- Определим множество $K_0 := \{x \in \mathbb{R}^n : \langle y, x \rangle > 0\}$. Данное множество выпукло и замкнуто.
- Зададим 2 выпуклых множества:

$$\mathcal{K} := K_0 * K_1 * \dots * K_q,$$

Где $\mathcal{K} \subseteq \mathbb{R}^{(q+1)n}$.

$$\mathcal{D} := \{(x, x, x, \dots, x) \in \mathbb{R}^{(q+1)n} : x \in \mathbb{R}^n\} \subseteq \mathbb{R}^{(q+1)n}.$$

Эти множества являются выпуклыми (\mathcal{K} является декартовым произведением выпуклых, а \mathcal{D} – линейной оболочкой, что тоже является выпуклым множеством), при этом $\mathcal{K} \cap \mathcal{D} = \emptyset$, т. к.:

- Во множестве K_0 лежат только те вектора x , с которыми $\langle x, y \rangle > 0$, а в $K_i, i \in \{1, \dots, q\}$ лежат только такие векторы, которые дают неположительное скалярное произведение с y . \implies В \mathcal{K} не содержится элементов вида такого, как в \mathcal{D} .

- Из теоремы 2.16:

$$\exists a \in \mathbb{R}^{(q+1)n} : \langle a, z \rangle \leq \langle a, \tilde{z} \rangle, \text{ для любого } z \in \mathcal{K} \text{ и } \tilde{z} \in \mathcal{D}.$$

где $a = (y^{(0)}, y^{(1)}, \dots, y^{(q)}) \neq \mathbb{O}$.

Таким образом (отсюда надо проверять дальше):

$$\forall X^{(0)} \in K_0, X^{(1)} \in K_1, \dots, X^{(q)} \in K_q, \forall x \in \mathbb{R}^n$$

$$\sum_{i=0}^q \langle y^i, x^i \rangle \leq \sum_{i=0}^q \langle y^i, x \rangle = \langle \sum_{i=0}^q y^i, x \rangle \quad (*)$$

Не все $y^i = 0$, это гарантируется условием теоремы.

- Из (*) следует, что:

$$\exists M \in \mathbb{R} : \forall i \in \{0, \dots, q\}, \forall x^{(i)} \in K_i : \langle y^i, x^i \rangle \leq M \text{ отсюда следует, что } \langle y^i, x^i \rangle \leq 0, \forall x^{(i)} \in K_i, \forall i \in \{0, \dots, q\}$$

- $y^{(i)} \in K_i^\circ, \forall i \in \{0, 1, \dots, q\}$

Допустим, что $y^{(0)} = \alpha y$, где $\alpha \leq 0$. Т.к. $y^{(0)} = \alpha y + u$ и $u \perp y$, т.е. $\langle y, u \rangle = 0$

Возьмём $\langle y^{(0)}, u \rangle = \|u\|^2 \succ 0$, тогда \exists окрестность U точки u , где $\langle y^{(0)}, x \succ 0, \forall x \in U$. Получим противоречие, т. к. $U \cap K_0 \neq \emptyset$. Тогда $y^{(0)} = \alpha y, y \in K_0 \rightarrow 0 \geq \langle y^{(0)}, y \rangle = \alpha \|y\|^2, y \neq 0$

- Из (*) следует, что:

$$\exists L \in \mathbb{R} : \forall x^{(i)} \in R^L : \langle \sum_{i=0}^q y^i, x \rangle \geq L, \sum y^{(i)} = 0(**)$$

- Из (**) следует, что:

$$-\alpha y = -y^{(0)} = \sum_{i=1}^q y^{(i)} \in \sum_{i=1}^q K_i$$

- $\exists x \in K_1 \cap (\bigcap_{i=2}^q \text{int}(K_i))$ Предположим, что $\alpha = 0$, то $y^{(0)} = 0$, тогда в (**):

$$\sum_{i=1}^q y^{(i)} = 0 (***)$$

Откуда следует, что:

$$\sum_{i=1}^q \langle y^i, x \rangle = \langle \sum_{i=1}^q y^i, x \rangle = 0$$

Откуда в свою очередь следует, что:

$$\forall i \in \{1, \dots, q\}, \langle y^i, x \rangle = 0(***)$$

Сумма неположительных равна нулю.

- Т. к. $x \in \text{int} K_i \forall i \in \{2, \dots, q\} \exists \varepsilon \succ 0$ и $x + \varepsilon y^{(i)} \in K_i$

$$\forall i \in \{2, \dots, q\} : 0 \geq \langle y^i, x + \varepsilon y^{(i)} \rangle = \langle y^i, x \rangle + \varepsilon \|y\|^2$$

Из этого следует, что все $y^{(i)} = 0$

Получаем противоречие.



6 Лекция 6. Полиэдральный конус

6.1 Полиэдральный конус

Определение 2.40

Полиэдральный конус - это конус, который является полиэдром.

6.2 Свойства полиэдрального конуса

Замечание 2.41

Множество $K \subseteq \mathbb{R}^{m \times n}$ — полиэдральный конус тогда и только тогда, когда $\exists A \in \mathbb{R}^{m \times n}$, удовлетворяющее условию

$$K = P^{\leq}(A, \mathbb{O}_n) = \{x | Ax \leq \mathbb{O}_n\}$$

Замечание 2.42

Полиэдральный конус выпуклый и замкнутый.

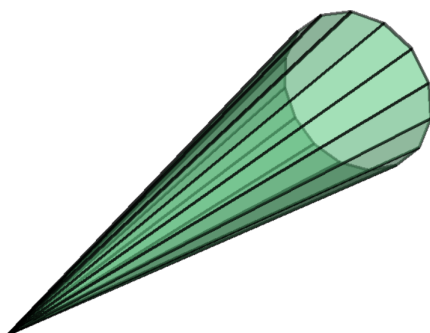
Замечание 2.43

Поляры конечно порожденных конусов являются полиэдральными конусами.

6.3 Теоремы и леммы про полиэдральные конусы

Лемма 2.44

Каждый полиэдральный конус является конечно порожденным.
(Смотри картинку ниже, черные линии - вектора для выпуклой конической оболочки)



Теорема 2.45

Конус полиэдральный тогда и только тогда, когда он конечно порожденный.



- Покажем, что каждый конечно порожденный конус K является полиэдральным.
- Пусть K конечно порожденный конус
 - $\implies K^\circ$ полиэдральный (следует из замечания 2.43).
 - $\implies K^\circ$ конечно порожденный (следует из леммы 2.44).
 - $\implies K^{\circ\circ}$ полиэдральный конус (следует из замечания 2.43).
- Покажем, что $K = K^{\circ\circ}$.
- K выпуклый, замкнутый $\implies K = K^{\circ\circ}$ (следует из теоремы 2.32, 2.38).



6.4 Усиление Леммы 2.44

Определение

$\forall M \in \mathbb{R}^{m \times n}$ следующие множества:

- $\triangleright \delta(M) = \{\det M_{I \times J} \mid I \subseteq [m], J \subseteq [n], |I| = |J|\}$. Судя по описанию, это – множество миноров в матрице M .
- $\triangleright \Delta(M) = \{\frac{p}{q} \mid p, q \in \delta(M) \cup (-\delta(M)), q \neq 0\}$ – множество дробей, что числитель и знаменатель лежат в указанных множествах миноров, причем $q \neq 0$.

Лемма 2.44*

$\forall A \in \mathbb{R}^{m \times n}, \exists X \subseteq \Delta(A)^n, |X| < \infty$ такое, что $\{x \mid Ax < 0\} = \text{ccone}(X)$.

$(\Delta(A)^n$ по логике означает n -кратное декартово произведение множества $\Delta(A)$).

Замечания к $\Delta(M)$

- $M \in \mathbb{Q}^{m \times n} \implies \Delta(M) \subseteq \mathbb{Q}$.
- Определим понятие **Длины кода**:

- $a \in \mathbb{Q}, a = \frac{p}{q}, p \in \mathbb{Z}, q \in \mathbb{N}/\{0\}, p, q$ взаимно простые: $\langle a \rangle := 1 + \lceil \log_2(|p| + 1) \rceil + \lceil \log_2(q + 1) \rceil$. $\langle \rangle$ -длина кода.
- Количество бит для хранения вектора $v \in \mathbb{Q}^n : \langle v \rangle := n + \sum_{i=1}^n \langle v_i \rangle$.
- Количество бит для хранения матрицы $M \in \mathbb{Q}^{m \times n} : \langle M \rangle := mn + \sum_{i=1}^n \sum_{j=1}^m \langle M_{ij} \rangle$
- Максимальная длина кода в матрице $\max\{\langle z \rangle : z \in \Delta(M)\}$ ограничена полиномом от длины кода матрицы $M \forall M \in \mathbb{Q}^{m \times n}$

Для доказательства Леммы 2.44*:

В лемме 2.44* важно показать, что $X \subseteq \Delta(A)^n$.

Индукция по $p = 0, 1, \dots$:

Для всех $B \in \mathbb{R}^{p \times n}, C \in \mathbb{R}^{q \times n} (p+q \geq 1, n \geq 1), A = \begin{pmatrix} B \\ C \end{pmatrix} \in \mathbb{R}^{(p+q) \times n}, \exists X \subseteq \Delta(A)^n, |X| < \infty$, такое что $K := \{x \in \mathbb{R}^n | Bx \leq 0, Cx = 0\} = \text{ccone}X$.

Лемма 2.44a

Пусть $B \in \mathbb{R}^{p \times n}, C \in \mathbb{R}^{q \times n} (p+q \geq 1, n \geq 1), A = \begin{pmatrix} B \\ C \end{pmatrix} \in \mathbb{R}^{(p+q) \times n}$,

$K := \{x \in \mathbb{R}^n | Bx \leq 0, Cx = 0\}$.

- В случае, если $\dim(\text{Ker}(A)) \geq \dim(\text{Ker}(C)) - 1 : \exists X \subseteq \Delta(A)^n, |X| < \infty$, такое что $K = \text{ccone}X$.
- В противном случае $\exists z \in \text{Ker}(C)$, такое что $z \notin K, -z \notin K$.

Illustration 1 des Beweises von Lemma 2.44a

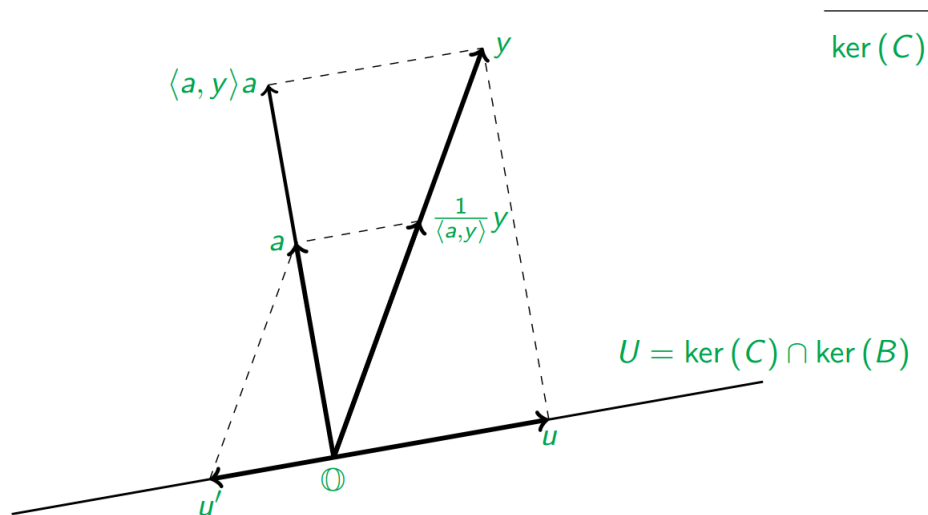
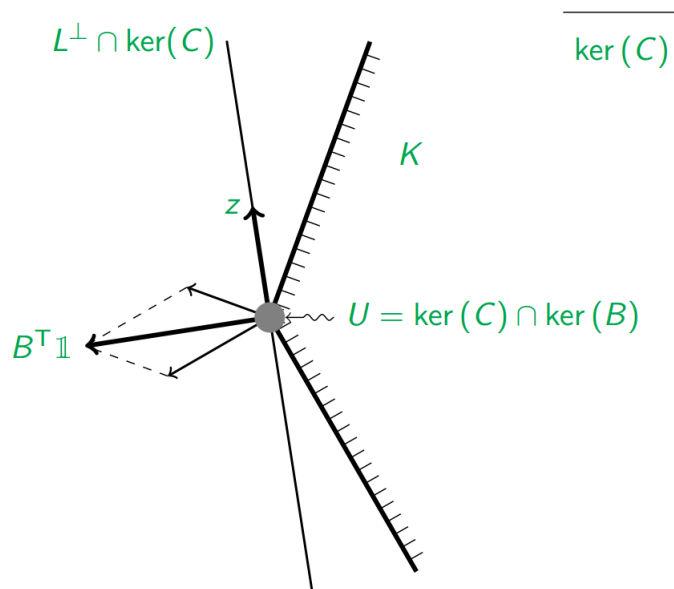


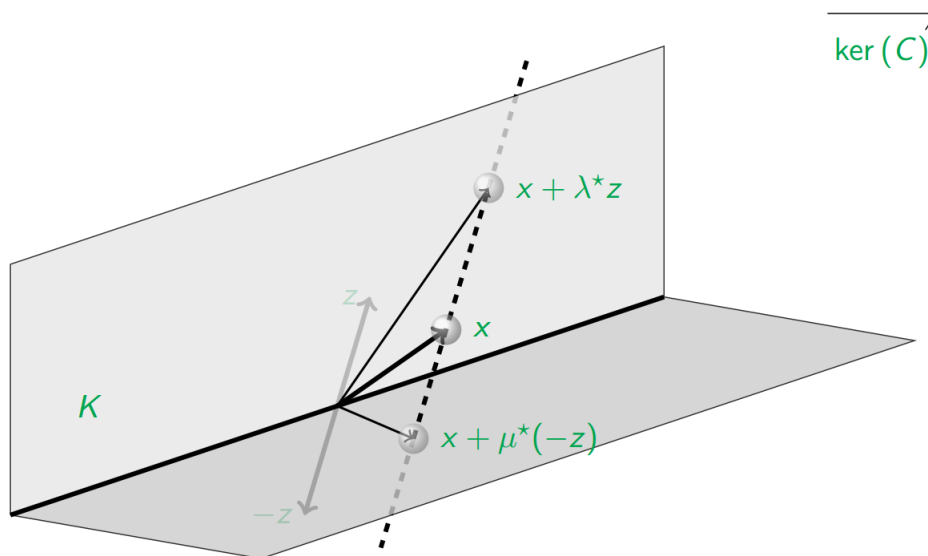
Illustration 2 des Beweises von Lemma 2.44a



Докажем лемму 2.44* по индукции

- $p = 0$: Доказательство вытекает из Леммы 2.44а
- $p \geq 1$: из Леммы 2.44а следует, что можно принять $z \in Ker(C)$, такое что $z, -z \notin K$.
 - Для всех $i \in [p] : B^{(i)} \in \mathbb{R}^{(p-1) \times n}, K_i := \{x \in \mathbb{R}^n : B^{(i)}x \leq 0, \langle B_{i,*}, x \rangle = 0, Cx = 0\}$.
 - По индукционному предположению $\exists X_i \subseteq \Delta(A)^n, |X_i| < \infty$, такое что $K_i = ccone X_i$.
Достаточно показать, что $K \subseteq ccone X, X := \cup_{i=1}^p X_i$ ("⊇" очевидно).
 - Пусть $x \in K$.
 - $I := \{i \in [p] : \langle B_{i,*}, z \rangle > 0\} \neq \emptyset$ (т.к. $Bz \leq 0, Cz = 0, z \notin K$)
 - $\forall i \in I : \lambda_i := \max\{\lambda \geq 0 : \langle B_{i,*}, x + \lambda z \rangle \leq 0\} \geq 0$ ($Bx \leq 0$, т.к. $x \in K$).
 - Выберем $i^* \in I, \lambda^* := \lambda_{i^*} = \min\{\lambda_i, i \in I\} \geq 0$.
Тогда $B(x + \lambda^* z) \leq 0$ (т.к. $Bx \leq 0$), $\langle B_{i^*,*}, x + \lambda^* z \rangle = 0, C(x + \lambda^* z) = 0$ ($Cx = 0, Cz = 0$).
 - $\implies x + \lambda^* z \in K_{i^*} = ccone X_{i^*} \subseteq ccone X, \lambda^* \geq 0$
 - Аналогично: для $-z \in Ker C \setminus K, \mu^* \geq 0$ можно построить $x + \mu^*(-z) \in ccone X$
 $\implies x - \mu^* z \in ccone X, \mu^* \geq 0$
 - Т.к. $ccone X$ выпуклый, то
 $x = \frac{\mu^*}{\lambda^* + \mu^*}(x + \lambda^* z) + \frac{\lambda^*}{\lambda^* + \mu^*}(x - \mu^* z) \in ccone X$, т.к. является выпуклой комбинацией двух точек, ему принадлежащих.

Illustration des Induktionsschritts (Beweis Lemma 2.44^{*})



7 Лекция 7. Теорема о поляре полиэдрического конуса

7.1 Теорема о поляре полиэдрического конуса

Теорема 2.46 Для поляра полиэдрического конуса $K = P^{\leq}(A, \mathbb{O}_m)$ (где $A \in \mathbb{R}^{m \times n}$) выполнено

$$K^{\circ} = \text{ccone}\{A_{1,*}, \dots, A_{m,*}\}$$

В частности: поляр полиэдрического конуса конечно порожденный.

Доказательство

Предположим, что это не так. Пусть $\tilde{K} = \text{ccone}\{A_{1,*}, \dots, A_{m,*}\}$.

Из замечания 2.37 $\tilde{K}^{\circ} = P^{\leq}(A, \mathbb{O}_m)$

Т.к. конус \tilde{K} замкнут, то из теорем 2.32, 2.38 $\tilde{K}^{\circ\circ} = \tilde{K}$

$\tilde{K}^{\circ\circ} = P^{\leq}(A, \mathbb{O}_m)^{\circ} = K^{\circ} \square$

Следствие 2.47 Пусть $K_1, \dots, K_q \subseteq \mathbb{R}^n$ полиэдрические конусы. Тогда

$$(\bigcap_{i=1}^q K_i)^{\circ} = \sum_{i=1}^q K_i^{\circ}$$

Поляр пересечения, из предыдущей теоремы, есть выпуклая коническая оболочка, натянутая на строки.

А это есть ни что иное как сумма Минковского поляр.

7.2 Лемма Фаркаша

Лемма 2.48 (Фаркаш) Пусть $A \in \mathbb{R}^{m \times n}$ и $b \in \mathbb{R}^m$. Тогда если $P^{\leq}(A, b) = \emptyset$, то существует $\lambda \in \mathbb{R}_+^m$ и

$$\lambda^T A = \mathbb{O}_n^T \text{ и } \langle \lambda, b \rangle = -1$$

Доказательство

В обратную сторону все должно быть понятно:

Система неразрешима, если существует

$$\lambda \in \mathbb{R}_+^m \text{ и } \lambda^T A = \mathbb{O}_n^T \text{ и } \langle \lambda, b \rangle = -1$$

Т.к. имеем $(\lambda^T A) x \leq \lambda^T b$ и получаем $-1 \leq 0$

Пусть $Ax \leq b$ неразрешима. Определим

$$\bar{A} = \begin{bmatrix} A & -b \\ \mathbb{O}_n^T & -1 \end{bmatrix} \in \mathbb{R}^{(m+1) \times (n+1)}$$

Из того, что $Ax \leq b$ неразрешима следует, что $P^{\leq}(\bar{A}, \mathbb{O}_{m+1}) \subseteq H^{\leq}(e_{n+1}, 0)$.

Предположим, что это не так. Тогда $\bar{A}x \leq \mathbb{O}_{m+1}$ и $x_{n+1} > 0$.

$$Ax_{[n]} \leq x_{n+1}b;$$

$$A \frac{1}{x_{n+1}} x_{[n]} \leq b.$$

Т.е мы нашли решение системы \implies противоречие.

Следовательно, $e_{n+1} \in (P^{\leq}(\bar{A}, \mathbb{O}_{m+1}))^\circ$. Значит, по теореме 2.46, т.к. поляр есть выпуклая коническая оболочка, натянутая на строки матрицы, существует коническая комбинация для разложения e_{n+1}

$$\exists \lambda \in \mathbb{R}_+^m, \lambda_0 \in \mathbb{R}_+ : \lambda^T A = \mathbb{O}_n, \lambda^T(-b) + \lambda_0(-1) = 1$$

$\lambda^T A = \mathbb{O}_n$ — первых n нулевых координат e_{n+1}

$$\lambda^T b = -\lambda_0 - 1 < 0$$

Умножая на $\frac{1}{|-\lambda_0 - 1|}$, получаем требуемое. \square

7.3 Теорема о матрице

Теорема 2.49 Для любой матрицы $A \in \mathbb{R}^{m \times n}$ и $b \in \mathbb{R}^m$ выполняется:

или

$$P^{\leq}(A, \mathbb{O}_m) = \{x \in \mathbb{R}^n | Ax \leq b\} \neq \emptyset$$

или

$$\{y \in \mathbb{R}^m | A^T y = \mathbb{O}_n, \langle b, y \rangle = -1, y \geq \mathbb{O}_m\} \neq \emptyset$$

(но не оба).

7.4 Глава 3. Оптимальные условия для выпуклой задачи оптимизации

Глава 3

Оптимальные условия для выпуклой задачи оптимизации

- **Выпуклая задача оптимизации**

$$\min\{f(x) \mid x \in X\}$$

- **Частный случай** Максимизировать/минимизировать линейную функцию на выпуклом множестве

$$\max\{f(x)\} = -\min\{-f(x)\}$$

1. Задача линейной оптимизации

- f аффинная, g аффинная, $X_0 = \mathbb{R}_+^r \times \mathbb{R}_-^s \times \mathbb{R}^m$

2. Полуопределенная задача оптимизации

Как и в линейном программировании, но в X_0 множитель \mathbb{S}_+^l .

$$\mathbb{S}_+^l = \{M \in \mathbb{S}^l \mid M \text{ положительно полуопределена}\}$$

$$\mathbb{S}^l = \{M \in \mathbb{R}^{l \times l} \mid M \text{ симметрическая}\}$$

Стандартный вид:

$$\min/\max \langle C, X \rangle$$

$$\langle A^{(i)}, X \rangle = b_i, \forall i \in [p]$$

$$X \in \mathbb{S}_+^k, C \in \mathbb{S}^k, A^{(i)} \in \mathbb{S}^k, b_i \in \mathbb{R}, i \in [p]$$

Каждую полуопределенную задачу (также линейную задачу) можно привести к полуопределенной задаче в стандартном виде.

- Скалярное произведение определяется

$$\langle M, M' \rangle = \sum_{i,j} M_{i,j} M'_{i,j}$$

- $\mathbb{S}^k \subseteq \mathbb{R}^{k \times k}$ является подпространством размерности $\frac{k(k+1)}{2}$

8 Лекция 8. Проблема выпуклой оптимизации

8.1 Постановка задачи

- Задача выпуклой оптимизации: $\min\{f(x) : x \in X\}$.

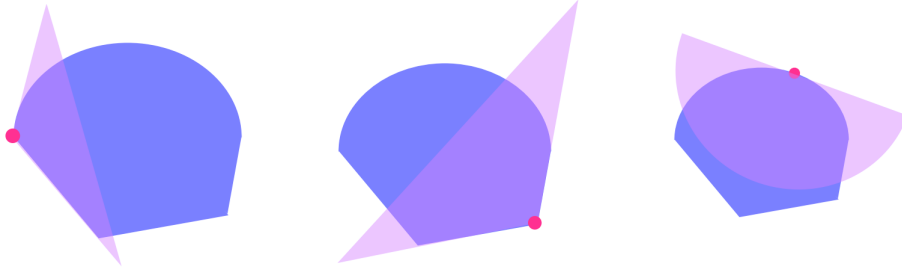
- Целевая функция $f : \mathbb{R}^n \rightarrow \mathbb{R}$ дифференцируемая и выпуклая.
- Множество допустимых решений:
 $X = \{x \in X^{(0)} : g^{(i)}(x) \leq 0, i \in [m], h^{(i)}(x) = 0, i \in [p]\}$
- $X^{(0)} \subseteq \mathbb{R}^n$ замкнутое выпуклое множество (простой структуры).
- $g^{(1)}, \dots, g^{(m)} : \mathbb{R}^n \rightarrow \mathbb{R}$ дифференцируемые и выпуклые функции.
- $h^{(1)}, \dots, h^{(p)} : \mathbb{R}^n \rightarrow \mathbb{R}$ аффинные функции.
- Таким образом, X замкнутое выпуклое множество.

8.2 Радиальные и нормальные конусы

Определение 3.1

Пусть $X \subseteq \mathbb{R}^n$ и $x^* \in X$. Тогда радиальный конус (конус направлений) множества X в точке x^* – это множество:

$$K_{x^*}(X) := \text{cone}(X - \{x^*\}) = \{\lambda(x - x^*) | x \in X, \lambda \geq 0\}$$

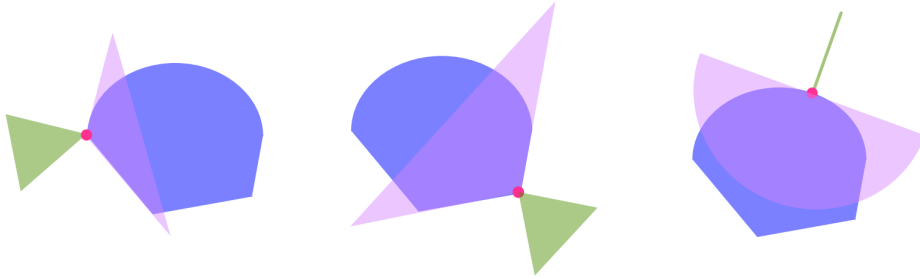


Наблюдение 3.2

Если x^* внутренняя точка из X , тогда $K_{x^*}(X) = \mathbb{R}^n$

Определение 3.3

Пусть $X \subseteq \mathbb{R}^n$ и $x^* \in X$. Тогда нормальный конус множества X в точке x^* – это множество: $N_{x^*}(X) := K_{x^*}(X)^\circ$



Наблюдение 3.4

Если x^* внутренняя точка из X , тогда $N_{x^*}(X) = \{\mathbb{O}_n\}$

8.3 Радиальные и нормальные конусы для полиэдра

Пусть $A \in \mathbb{R}^{m \times n}$, $b \in \mathbb{R}^m$ и $x^* \in \mathbb{R}^n$, причём $Ax^* \leq b$:

$$Eq_{Ax \leq b}(x^*) := \{i \in [m] \mid \langle A_{i,*}, x^* \rangle = b_i \subseteq [m]\}$$

Радиальный конус для полиэдра $Ax \leq b$ в точке x^* – это все направления y , такие, что:

$$\begin{aligned} \exists \lambda > 0 : A(x^* + \lambda y) &\leq b \\ Ax^* + \lambda Ay &\leq b \end{aligned}$$

Строки $\langle A_{i,*}, x^* \rangle \leq b_i$ задают любое направление y , а строки $\langle A_{i,*}, x^* \rangle = b_i$ дают только те направления y , для которых $\langle A_{i,*}, y \rangle \leq 0$.

Замечание 3.5

Пусть $x^* \in P^{\leq}(A, b) = \{x \in \mathbb{R}^n \mid Ax \leq b\}$ ($A \in \mathbb{R}^{m \times n}$, $b \in \mathbb{R}^m$). Тогда справедливо:

$$K_{x^*}(P^{\leq}(A, b)) = \{y \in \mathbb{R}^n \mid \langle A_{i,*}, y \rangle \leq 0 \mid i \in Eq_{Ax \leq b}(x^*)\}$$

$$N_{x^*}(P^{\leq}(A, b)) = \text{ccone}\{A_{i,*} \mid i \in Eq_{Ax \leq b}(x^*)\}$$

В частности, радиальный и нормальный конусы полиэдра – полиэдрические конусы.

8.4 Оптимизация дифференцируемой функции на выпуклом множестве.

Теорема 3.6

Пусть $f : X \rightarrow \mathbb{R}$ дифференцируемая функция, заданная на выпуклом множестве $X \subseteq \mathbb{R}^n$.

1. Если f принимает в точке $x^* \in X$ локальный минимум на X , то $-\text{grad}_{x^*}(f) \in N_{x^*}(X)$
2. Если f выпуклая и $-\text{grad}_{x^*}(f) \in N_{x^*}(X)$, то $f(x^*) = \min\{f(x) \mid x \in X\}$

◀ Пусть:

1. f принимает в x^* локальный минимум на X .

- Предположим $-\text{grad}_{x^*}(f) \notin N_{x^*}(X)$.
- Это значит: $-\text{grad}_{x^*}(f) \notin K_{x^*}(X)^\circ$
- $\implies \exists \lambda \geq 0, x \in X : \langle -\text{grad}_{x^*}(f), \lambda(x - x^*) \rangle > 0$,
- $\lambda(x - x^*) \in K_{x^*}(X)$
- т.к. $(\lambda \neq 0) \implies \langle \text{grad}_{x^*}(f), x - x^* \rangle < 0$.
- Рассмотрим функцию $\phi : t \rightarrow f(x^* + t(x - x^*))$, тогда $\langle \text{grad}_{x^*}(f), x - x^* \rangle = \phi'(0)$
- $\phi'(0) < 0 \implies \exists$ достаточно малое $0 < \epsilon \leq 1 : f(x^* + \epsilon(x - x^*)) = \phi(\epsilon) < \phi(0) = f(x^*)$, $x^* + \epsilon(x - x^*) \in X$
- X выпуклое множество, $x, x^* \in X$ и $x^* + \epsilon(x - x^*)$ принадлежит окрестности x^* .
- Получили противоречие с тем, что в x^* достигается локальный минимум.
- Следовательно, $-\text{grad}_{x^*}(f) \in N_{x^*}(X)$.

2. Пусть f выпуклая функция и $-\text{grad}_{x^*}(f) \in N_{x^*}(X)$.

- Пусть $x \in X$ произвольная точка.
- Для всех $t \in]0, 1]$ справедливо:

$$f(x^* + t(x - x^*)) = f((1-t)x^* + tx) \leq (1-t)f(x^*) + tf(x) = f(x^*) + t(f(x) - f(x^*))$$

$$\implies \frac{f(x^* + t(x - x^*)) - f(x^*)}{t} \leq f(x) - f(x^*)$$

- $\frac{f(x^* + t(x - x^*)) - f(x^*)}{t} \xrightarrow{t \rightarrow 0} \phi'(0) = \langle \text{grad}_{x^*}(f), x - x^* \rangle$
- $\langle \text{grad}_{x^*}(f), x - x^* \rangle \leq f(x) - f(x^*)$

$$\langle \text{grad}_{x^*}(f), x - x^* \rangle \geq 0 \text{ [т.к. } x - x^* \in K_{x^*}(X)] \implies f(x^*) \leq f(x). \blacktriangleright$$

Следствие 3.7

Дифференцируемая выпуклая функция f принимает свой (глобальный) минимум в точке x^* на множестве X тогда и только тогда, когда $-\text{grad}_{x^*}(f) \in N_{x^*}(X)$.

Замечания к теореме 3.6 (следствию 3.7)

1. Если $x^* \in \text{int}(X)$, тогда $N_{x^*}(X) = \{\mathbb{O}\}$ и тогда в условии теоремы $\text{grad}_{x^*}(f) = \mathbb{O}$.
2. Теорема 3.6 / следствие 3.7 не говорят о существовании оптимального решения.
Пример: $\min x$, $x \in \mathbb{R}$ или $\min y$, при условии $y \geq 1/x$, $x > 0$.

8.5 Некоторые нормальные конусы

Лемма 3.8

Если $K \subseteq \mathbb{R}^n$ и $x^* \in K$, тогда $N_{x^*}(K) = \{y \in K^\circ \mid \langle x^*, y \rangle = 0\}$.

Замечание 3.9

1. $x^* \in \mathbb{R}_+^n$ (1-й ортант), тогда $N_{x^*}(\mathbb{R}_+^n) = \{y \in \mathbb{R}_-^n \mid y_i = 0 \text{ } i \in [n], x_i^* > 0\}$
2. $X^* \in \mathbb{S}_+^k$ (положительно определённая матрица), тогда $N_{X^*}(\mathbb{S}_+^k) = \{Y \in \mathbb{S}_-^k \mid \langle X^*, Y \rangle = 0\}$ (скалярное произведение Фробениуса)

Замечание 3.10

$K_i \subseteq \mathbb{R}^{n_i}$ ($i \in [r]$) и $(x^{(1)}, \dots, x^{(r)}) \in K_1 \times \dots \times K_r$, тогда $N_{(x^{(1)}, \dots, x^{(r)})}(K_1 \times \dots \times K_r) = N_{x^{(1)}}(K_1) \times \dots \times N_{x^{(r)}}(K_r)$

8.6 Условия Каруша-Куна-Таккера. Постановка задачи

$$(X_0, (g_{(i)})_{i \in [m]}, (h_{(i)})_{i \in [p]})$$

- $X_0 \subseteq \mathbb{R}^n$ выпуклое.
- $g_i : \mathbb{R}^n \rightarrow \mathbb{R}$ выпуклые и дифференцируемые ($i \in [m]$)
- $h_i : \mathbb{R}^n \rightarrow \mathbb{R}$ аффинные ($i \in [p]$)

Множество допустимых решений:

$$X = \{x \in X_0 \mid g^{(i)}(x) \leq 0, i \in [m], h^{(i)}(x) = 0, i \in [p]\}$$

Иная запись

$$X = X_0 \cap \bigcap_{i \in [m]} G_i \cap \bigcap_{i \in [p]} H_i$$

где $G_i := g_i^{-1}(\mathbb{R}_-)$ и $H_i := h_i^{-1}(0) \subseteq \mathbb{R}^n$.

8.7 Нормальный конус для регулярной тройки

Лемма 3.11

Если $(X_0, (g_{(i)})_{i \in [m]}, (h_{(i)})_{i \in [p]})$ регулярная, тогда:

$$N_{x*}(X) = N_{x*}(X_0) + \sum_{i \in [m]} N_{x*}(G_i) + \sum_{i \in [p]} N_{x*}(H_i).$$

$(X_0, (g_{(i)})_{i \in [m]}, (h_{(i)})_{i \in [p]})$ регулярная если:

1. Множество X_0 – полиэдр и функции g_1, \dots, g_m аффинные или
2. Множество $X \cap \text{int}(X_0)$ не пусто и функции g_1, \dots, g_m аффинные или
3. Существует точка $x^{(s)} \in X$ ($x^{(s)} \in \text{int}(X_0), p \neq 0$) такая, что $g_i(x^{(s)}) < 0$ для всех $i \in [m]$ (условие Слейтера)

◀ (Лемма 3.11)

$$N_{x*}(x) = (K_{x*}(x))^\circ = (K_{x*}(X_0) \cap \bigcap_{i=1}^m K_{x*}(G_i) \cap \bigcap_{i=1}^p K_{x*}(H_i))^\circ$$

1. • Все радиальные конусы полиэдрические, из следствия 2.47:
 $\Rightarrow (K_{x*}(X_0) \cap \bigcap_{i=1}^m K_{x*}(G_i) \cap \bigcap_{i=1}^p K_{x*}(H_i))^\circ = N_{x*}(X)$

- $N_{x*}(X) = N_{x*}(X_0) + \sum_{i \in [m]} N_{x*}(G_i) + \sum_{i \in [p]} N_{x*}(H_i)$
- 2.
- $K_1 := \bigcap_{i=1}^m K_{x*}(G_i) \cap \bigcap_{i=1}^p K_{x*}(H_i)$
 - $K_2 := K_{x*}(X_0)$ т.к. $X \cap \text{int}(X_0)$ не пусто, то $K_1 \cap \text{int}(K_2) \neq \emptyset$
 - Из теоремы 2.39 $\Rightarrow (K_1 \cap K_2)^\circ = K_1^\circ + K_2^\circ$
 - По следствию 2.47 $\Rightarrow (K_{x*}(X_0) \cap \bigcap_{i=1}^m K_{x*}(G_i) \cap \bigcap_{i=1}^p K_{x*}(H_i))^\circ = N_{x*}(X)$
 - $N_{x*}(X) = N_{x*}(X_0) + \sum_{i \in [m]} N_{x*}(G_i) + \sum_{i \in [p]} N_{x*}(H_i)$
- 3.
- $p = 0$:
 - $K_1 := K_{x*}(X_0)$, $K_{i+1} := K_{x*}(G_i)$ $i \in [m]$
 - $\Rightarrow_{x(s)} K_1 \cap \bigcap_{i=2}^{m+1} \text{int}(K_i) \neq \emptyset$
 - $\Rightarrow_{2.39} (K_{x*}(X_0) \cap \bigcap_{i=1}^m K_{x*}(G_i))^\circ = N_{x*}(X_0) + \sum_{i \in [m]} N_{x*}(G_i)$
 - $p > 0$:
 - $K_1 := K_{x*}(H_i)$, $K_2 := K_{x*}(X_0)$, $K_{i+2} := K_{x*}(G_i)$ $i \in [m]$
 - $\Rightarrow_{x(s)} K_1 \cap \bigcap_{i=2}^{m+2} \text{int}(K_i) \neq \emptyset$
 - По теореме 2.39, следствию 2.47:

$$(K_{x*}(X_0) \cap \bigcap_{i=1}^m K_{x*}(G_i) \cap \bigcap_{i=1}^p K_{x*}(H_i))^\circ = N_{x*}(X)$$
 - $N_{x*}(X) = N_{x*}(X_0) + \sum_{i \in [m]} N_{x*}(G_i) + \sum_{i \in [p]} N_{x*}(H_i)$



9 Лекция 9. Дополнительное условие выпуклости и дифференцируемости, ККТ

9.1 Нормальный конус для регулярной тройки

Лемма 3.11

Если $(X_0, (g_{(i)})_{i \in [m]}, (h_{(i)})_{i \in [p]})$ регулярная, тогда:

$$N_{x*}(X) = N_{x*}(X_0) + \sum_{i \in [m]} N_{x*}(G_i) + \sum_{i \in [p]} N_{x*}(H_i).$$

$(X_0, (g_{(i)})_{i \in [m]}, (h_{(i)})_{i \in [p]})$ регулярная если:

1. Множество X_0 – полиэдр и функции g_1, \dots, g_m аффинные
или
2. Множество $X \cap \text{int}(X_0)$ не пусто и функции g_1, \dots, g_m аффинные
или

3. Существует точка $x^{(s)} \in X$ ($x^{(s)} \in \text{int}(X_0), p \neq 0$) такая, что $g_i(x^{(s)}) < 0$ для всех $i \in [m]$ (условие Слейтера)

9.2 Дополнительное условие выпуклости и дифференцируемости

Теорема 1.16

Если $g : \mathbb{R}^n \rightarrow \mathbb{R}$ дифференцируемая, выпуклая и существует точка $x^{(s)} \in \mathbb{R}^n$ такая, что $g(x^{(s)}) < 0$, то для всех $x^* \in \mathbb{R}^n$, для которых $g(x^*) \leq 0$:

$$N_{x^*}(g^{-1}(\mathbb{R}_-)) =$$

$$\begin{cases} \text{cone}\{\text{grad}_{x^*}g\}, g(x^*) = 0 \\ \{\mathbb{O}_n\}, g(x^*) < 0 \end{cases}$$

Доказательство

1. Сначала определим множество $G := g^{-1}(\mathbb{R}_-) = \{x \in \mathbb{R}^n : g(x) \leq 0\}$.

- Предположим, что точка $x^{(s)} \in \mathbb{R}^n$, такая что $g(x^{(s)}) < 0$, существует.
- Если $g(x^*) < 0$, тогда $x^{(s)} \in \text{int}(G)$ (Так как, если непрерывная функция меньше нуля в конкретной точке, то она меньше нуля и в некоторой окрестности этой точки.)
И, также, $N_{x^*}(G) = \{\mathbb{O}\}$ Пусть теперь $g(x^*) = 0$. Нужно показать, что $N_{x^*}(G) = \text{cone}\{\text{grad}_{x^*}g\}$

2. " \supseteq "

- Покажем, что $N_{x^*}(G) \supseteq \text{cone}\{\text{grad}_{x^*}g\}$ т.е. надо показать, что градиент содержится в нормальном конусе, так как $\text{cone}\{\text{grad}_{x^*}g\}$ содержит градиент
Будем доказывать от противного. Положим, что

$$\text{cone}\{\text{grad}_{x^*}g\} \notin N_{x^*}(G) = (\text{cone}(G - \{x^*\}))^\circ$$

- Соответственно, существует $x \in G$ такой, что $\langle \text{grad}_{x^*}g, x - x^* \rangle > 0$ (Так как в поларе лежат все векторы, которые более

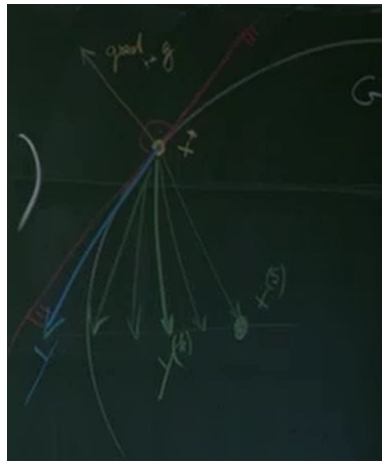
чем ортогональны с исходным. А само это утверждение вытекает из предположения, что градиент не лежит в поляре.)

Значит, существуют точки на сколь угодно большом расстоянии от x^* такие, что их можно выразить следующим образом $y(t) = x^* + t(x - x^*)$ с $g(y(t)) > g(x^*)$ ($t \geq 0$)

- Далее выберем такое t , что $0 \leq t \leq 1 \Rightarrow g(y(t)) > g(x^*) = 0$
С другой стороны $y(t) \in \text{conv}\{x^*, x\} \supseteq G$. Но на множестве G $g(x) \leq 0$. Получили противоречие.

3. " \subseteq "

- Достаточно показать, что для всех $y \in \mathbb{R}^n$ таких, что $\langle \text{grad}_{x^*} g, y \rangle \leq 0$: $y \in \text{cl}(K_{x^*}(G))$
(Так как $(\text{cone}\{\text{grad}_{x^*} g\})^\circ \subseteq \text{cl}(K_{x^*}(G))$) Следовательно $N_{x^*}(G) = (K_{x^*}(G))^\circ = (\text{cl}(K_{x^*}(G)))^\circ \subseteq ((\text{cone}(\text{grad}_{x^*}(g)))^\circ)^\circ = \text{conegrad}_{x^*}(g)$



- ПОЯСНЕНИЯ. Любой вектор, который более чем ортогонален градиенту, должен лежать в допустимом направлении или, по крайней мере, в замыкании допустимых направлений, исходящих из x^* . Если $\langle \text{grad}_{x^*}, y \rangle = 0$ - это недопустимое направление, но y лежит в топологическом замыкании допустимых направлений. Мы строим последовательность векторов, которые гарантированно лежат в радиальном конусе и сходятся к y . И тут нам нужна точка $x^{(s)}$ из условий теоремы, которая точно лежит в G .
- Пусть $y \in \mathbb{R}^n$ такой, что $\langle \text{grad}_{x^*}, y \rangle \leq 0$
- Определим последовательность $y^{(k)} := \frac{1}{k}(x^{(s)} - x^*) + \frac{k-1}{k}y$. Заметим, что все $y^{(k)}$ лежат в радиальном конусе.

- Так как $\lim_{k \rightarrow \infty} y^k = y$ достаточно показать, что для всех $ky^{(k)} \in K_{x^*}(G)$.
- $\langle grad_{x^*} g, y^{(k)} \rangle = \frac{1}{k} \langle grad_{x^*} g, x^{(s)} - x^* \rangle + \frac{k-1}{k} \langle grad_{x^*} g, y \rangle$
Первое слагаемое строго меньше нуля ($g(x^*) = 0, g(x^{(s)}) < 0$), g — выпуклая), а второе меньше либо равно.
- Получаем $\langle grad_{x^*} g, y^{(k)} \rangle < 0$, т.е. функция g убывает в направлении $y^{(k)}$, т.е. на любом расстоянии от x^* много меньших значений функции g .
Следовательно, существует $t > 0$, с $g(x^* + ty^{(k)}) < 0$ и $y^{(k)} \in cone\{x - x^*\} \subseteq cone\{cone(G - \{x^*\})\} = K_{x^*}(G)$.

□

9.3 Условия Каруша - Кунна - Такера (дифференцируемость, выпуклость)

Теорема 3.13

Точка $x^* \in X$ является оптимальным решением задачи

$$\min \{f(x) | x \in X\}$$

когда \exists множители $\lambda_1, \dots, \lambda_m \in \mathbb{R}_+$ и $\mu_1, \dots, \mu_p \in \mathbb{R}$ такое, что

$$grad_{x^*} f + \sum_{i=1}^m \lambda_i grad_{x^*} g_i + \sum_{i=1}^p \mu_i grad_{x^*} h_i \in -N_{x^*}(x_0) \quad (1)$$

и

$$\lambda_i = 0 \text{ для всех } i \in [m], \text{ что } g_i(x^*) < 0 \quad (2)$$

9.4 Условия Каруша - Кунна - Такера для задачи линейной оптимизации (1 вариант)

Теорема 3.14 (Условия дополняющей нежесткости 1)

Пусть $A \in \mathbb{R}^{m \times n}$, $b \in \mathbb{R}^m$ и $c \in \mathbb{R}^n$. Точка $x^* \in \mathbb{R}_+^n$ с $Ax^* = b$ является оптимальным решением задачи

$$\min \{ \langle c, x \rangle | Ax = b, x \in \mathbb{R}_+^n \},$$

когда существует $\mu \in \mathbb{R}^m$ с $\mu^T \leq c^T$ и

$$\mu^T A_{*,j} = c_j, j \in [n], x_+^* > 0.$$

Теорема 3.14 следует из теоремы 3.13 с $f(x) = \langle c, x \rangle, x_0; x_0 \in \mathbb{R}_+, m = 0, h_i(x) = \langle A_{i,*}, x \rangle - b_i (i \in [p])$ - начальные условия. \Rightarrow регулярно (тип 1), X_0 - полиэдр. Проверка выполнения условий теоремы 3.13: (2) \emptyset , т.к. $m=0$
(1) $c + \sum_{i=1}^p \mu A_{i,*} \in \{-x : x \in \mathbb{R}_-, \langle x, x^* \rangle = 0\}$ (Лемма 3.8) $\Leftrightarrow c - \bar{\mu}^T A \in \{x \in \mathbb{R}_+, \langle x, x_* \rangle = 0\} \Leftrightarrow c - \bar{\mu}^T A \geq 0$ и $c_j - \bar{\mu}^T A_{*,j} = 0 \forall j \in [n], x^*_j > 0$.

9.5 ККТ для задачи линейной оптимизации (2 вариант)

Теорема 3.15

Пусть $A \in \mathbb{R}^{m \times n}, b \in \mathbb{R}^n, c \in \mathbb{R}^n$, точка $x^* \in \mathbb{R}^n, Ax^* \leq b$ является решением задачи:

$$\max \{ \langle c, x \rangle \mid Ax \leq b, x \in \mathbb{R}^n \},$$

тогда \exists вектор $\lambda \in \mathbb{R}_+^m, \lambda^T A = c^T$ и

$$\lambda_i = 0 \quad i \in [m], \langle A_{i,*}, x^* \rangle < b_i$$