Семинар по оптимизации

Материалы подготовлены: Паниным Никитой

По вопросам и неточностям писать:
1) paninnikita83@gmail.com
2) telegram: @nickzooot

Оптимизация : $f(x) \xrightarrow[X]{} \min$

условная оптимизация — безусловная

$x \in X \subset \mathbb{R}^n$ — $X = \mathbb{R}^n$

$\exists$ v - точка локального минимума

- if $v \in \operatorname{int} X$, то необходимо 1) $f'(v) = 0$
  2) $f''(v) \geq 0$

- if $v \in \partial X$, то это не верно (Пример: $f(x) = -x^2$ $X = [1 \leq x \leq 2]$)

Поэтому требуются более глубокие рассуждения на эту тему:

1. (1) $\exists$ $X = \{x \in E^n : \underbrace{g_1(x) \leq 0, \dots, g_m(x) \leq 0}_{\text{ограничение типа нер-в}}, \underbrace{g_{m+1}(x) = 0, \dots, g_s(x) = 0}_{\text{огр. типа равенств}}\}$

$\operatorname{dom} f$ и $\operatorname{dom} g_i = E^n$

(Могут быть и: 1) $m = 0$ (нет огр. нер-в)
2) $m = s$ (нет огр. р-в)
3) $m = s = 0$ (з-ча безусл. опт.-ции)

красным отмечено то, чего не было на семинаре

Для з-чи исследования на экстремум вводим ф-цию Лагранжа:

$$\mathcal{L}(x, \bar{\lambda}) = \lambda_0 f(x) + \sum_{j=1}^{s} \lambda_j g_j(x), \text{ где } x \in E^n \; \bar{\lambda} \in E^{s+1}$$

Th (Каруша - Джона) (необх. условие 1-го порядка)

1) $\exists$ x_* - т. локального минимума ф-ции на X (1)

2) $f(x), g_1(x), \dots, g_m(x)$ - дифф-мы в x_*

3) $g_{m+1}(x), \dots, g_s(x)$ - непрерывно дифф-мы в некоторой окрестности т. x_*

Тогда $\exists$ множители Лагранжа $\overline{\lambda} = (\lambda_0^*, \lambda_1^*, \dots, \lambda_s^*)$:

a) $\overline{\lambda} \neq 0$

b) $\lambda_0^* \geq 0 \quad \lambda_1^* \geq 0, \dots, \lambda_m^* \geq 0$

c) $\mathcal{L}_x(x_*, \overline{\lambda}) = 0$

d) $\lambda_i^* g_i(x_*) = 0, i = 1, \dots, m$ — условие дополняющей нежесткости

Замечание: $\Rightarrow$ Точками локального минимума ф-ции $f(x)$ на X могут быть лишь те точки $x \in E^n$, для которых $\exists$ мн-ли Лагранжа $\overline{\lambda} = (\lambda_0, \dots, \lambda_s)$ такие что пара $(x, \overline{\lambda})$ явл. реш. сист.:

(2)

a) $\overline{\lambda} \neq 0$

b) $\lambda_0 \geq 0, \dots \lambda_m \geq 0$

c) $\mathcal{L}_x(x, \overline{\lambda}) = 0$

d) $\lambda_i g_i(x) = 0 \quad i = 1, \dots m$

+ (вспоминаем про то, чему принадлежит x)

e) $g_i(x) \leq 0, i = \overline{1, m}$

f) $g_i(x) = 0, i = \overline{m+1, s}$

Опр $\exists v$ – фикс. т. локального минимума $f(x)$ на X (см.(1))

Мн-во всех т. $\overline{\lambda}$, для которых $(\overset{x}{\overset{\shortparallel}{v}}, \overline{\lambda})$ явл реш (a)-(f) будем называть множителями Лагранжа (обозн. это мн-во: $\Lambda(v)$)

Замечание: Если $(v, \overline{\lambda})$ реш системы (2) (a-e), то $(v, \alpha\overline{\lambda})$, где $\alpha > 0$ тоже решение этой системы $\Rightarrow \Lambda(v)$ – конус, который будем называть конусом Лагранжа точки v

Опр Ограничение $g_i(x) \leq 0$ наз. активным в т. v, if $g_i(v) = 0$ и пассивным (неактивным) в т. v, if $g_i(v) < 0$.

(Ограничения типа равенств, конечно, активные)

Пр1 $f(x) = \sum_{i=1}^{p} \|x - x_i\|^2$

(стр 65. Васильев) $\bar{X} = \{x \in E^n;\ |x|^2 = \langle x, x \rangle \le 1\}$

Пр2 $f(u) = x$

$X = \{u = (x,y) \in E^2 : g_1(u) = -x \le 0\ g_2(u) = x^2 - y \le 0\ g_3(u) = y - 2x^2 \le 0\}$

Правило множителей Лагранжа в общем случае (необх. усл)

З-ча: $f(x) \to \inf\ x \in X$

(3) $X = \{x \in X_0 \mid \underbrace{g_i(x) \le 0\ i = 1, .., m}_{\text{ограничения типа нер-в}};\ \underbrace{g_i(x) = 0,\ i = m+1, ... s}_{\text{огр. типа равенств}}\}$

a) $X_0 \subset E^n$

b) X_0 — выпуклое мн-во

c) $f(x), g_i(x)$ определены на X_0

(*)

Замечание: Не исключаются следующие случаи:

a) $m = 0$ (нет нер-в)

b) $s = m$ (нет р-в)

c) $m = s = 0$

Ф-ция Лагранжа:

$$\mathcal{L}(x, \bar{\lambda}) = \lambda_0 f(x) + \lambda_1 g_1(x) + \ldots + \lambda_s g_s(x),$$

где $x = (x_1, .., x_n) \in X_0$ $\bar{\lambda} = (\lambda_0, .., \lambda_m) \in E^{s+1}$ $\lambda_1 \ge 0, .., \lambda_m \ge 0$, $\lambda_0 \ge 0$

Th.2. ∃ 1) X взято из условия 3. (X_0-выпуклое!!!)

∃ 2) $x_* \in X$ — точка локального минимума в з-че (*)

] 3) $f(x), g_1(x), \ldots, g_m(x)$ - дифф-мы в т. x_*

$\|x_* - x\| < \varepsilon$

4) $g_{m+1}(x), \ldots g_s(x)$ - непрерывно дифф-мы в некоторой окр-ти $O(x_*, \varepsilon) \cap X_0$ точки x_*.

множители Лагранжа (удовл a, b, c)

Тогда $\exists \lambda = (\lambda_0, \lambda_1, \ldots, \lambda_s)$:

a) $\bar\lambda = (\lambda_0, \lambda_1, \ldots \lambda_s) \neq 0$ $\lambda_0 \geq 0, \lambda_1 \geq 0, \ldots, \lambda_m \geq 0$

Вариационное неравенство

b) $\langle \mathcal{L}_x(x_*, \bar\lambda), x - x_* \rangle \geq 0$ $\forall x \in \underline{X_0}$ (не X !!!)

c) $\lambda_i g_i(x_*) = 0$ - условие дополняющей нежесткости

Замечание:

① Точками лок. мин. могут быть лишь те $x = v$, для которых $\exists$ множители Лагранжа $\bar\lambda = (\lambda_0, \lambda_1, \ldots, \lambda_s)$: $(v, \bar\lambda) \in E^{n+s+1}$ явл. реш. системы:

$(**)$

a) $\langle \lambda_0 f(v) + \lambda_1 g_1'(v) + \ldots + \lambda_s g_s'(v), x - v \rangle \geq 0$ $\forall x \in X_0$

b) $\lambda_i g_i(v) = 0$ $i = 1, \ldots, m$

c) $v \in X_0$ $g_1(v) \leq 0, \ldots, g_m(v) \leq 0$ $g_{m+1}(v) = 0, \ldots, g_s(v) = 0$

d) $\bar\lambda \neq 0$ $\lambda_0 \geq 0, \ldots, \lambda_m \geq 0$

② Если x_* - внутренняя точка X_0, то вариационное

нер-во (Th 2, b))

равносильно условию

$$\mathcal{L}_x'(x_*, \bar\lambda^*) = \lambda_0^* f'(x_*) + \sum_{j=1}^{m+s} \lambda_j^* g_j'(x_*) = 0$$

(то же открытое)

(т.е если $X_0 = E^n$ или открытое, то автоматически выполнено)

Опр $\exists\ v$ — фикс. т. локального минимума $f(x)$ на X (см.(1))

Мн-во всех т. $\bar{\lambda}$, для которых $(\overset{x}{\overset{\|}{v}}, \bar{\lambda})$ явл реш (a)-(d)

будем называть множителями Лагранжа (обозн. это мн-во: $\Lambda(v)$)

Замечание: Если $(v, \bar{\lambda})$ реш система (2) (a-e), то $(v, \alpha\bar{\lambda})$, где

$\alpha > 0$ тоже решение этой системы $\Rightarrow$ $\Lambda(v)$ - конус, который

будем называть конусом Лагранжа точки v $\Rightarrow$

$\Rightarrow$ можем если $\lambda_0 > 0$ делить на $\alpha = \lambda_0$ и

рассматривать $\bar{\lambda} = (1, \lambda_1, \lambda_2 \ldots \lambda_n)$

(3) Как обычно решать? (не во всех задачах работает)

3.1) проверить, что з-ча записана в каноническом виде

3.2) Выписать ф-цию Лагранжа

3.3) Использовать теорему, перебирая.

сначала $\lambda_0 = 0$, потом $\lambda_0 = 1$

$\vec{\lambda} \neq 0$

$\parallel$

$(\lambda_0, \lambda_1, \dots, \lambda_s)$

$\Downarrow$

находим стационарные точки

Достаточное условие оптимальности

Th 3

1) $\exists$ в з-че $(*)$ X_0 - выпукло

2) $f(x), g_1(x), \dots g_s(x)$ – непр. дифф-мы в некоторой окрестности точки $x_* \in X$

3) $f(x), g_1(x), \dots g_s(x)$ – дважды дифф-мы в x_*

4) $(x_*\ \bar{\lambda}^*)$ - решение системы $(**)$

5) $< \mathcal{L}''_{xx}(x_*, \bar{\lambda}^*) h, h > > 0$

для $\forall h \in E^n$, $h \neq 0$:

5.1 $\lambda_0^* < f'(x_*), h > \le 0$

5.2 $< g_j'(x_*), h > \le 0$ $\forall j \in I(x_*)$

5.3 $< g_j'(x_*), h > = 0$ $j = m+1, \dots s$

(if $X_0 = E^n$ и $m=0$, то эти усл. не нужны)

мн-во индексов активных ограничений (вводили выше)

Тогда в x_* реализуется строгий локальный минимум

(if $\lambda_0 \le 0$ то строгий локальный максимум)

! Замечание: Во всех сформулированных выше теоремах для максимума надо $\lambda_0 \geq 0$ заменить на $\lambda_0 \leq 0$ и все!!!

<u>Задача 1</u> (исследуем на экстремум) [λ_0 - может равняться нулю]

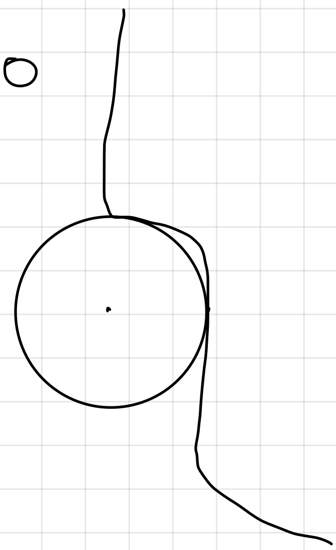
$$\begin{cases} f(u) = x \\ u \in X = \{u = (x, y) \in E^2 : g_1(u) = x^2 + y^2 - 1 = 0 \quad g_2(u) = x^3 + y^3 - 1 = 0\} \end{cases}$$

(1) $\mathcal{L}_u'(u, \bar{\lambda}) = (\lambda_0 + 2\lambda_1 x + 3\lambda_2 x^2; \quad 2\lambda_1 y + 3\lambda_2 y^2)$

$u \in E^2 \quad \bar{\lambda} \in R^3$

$$\mathcal{L}_{uu}(u, \bar{\lambda}) = \begin{pmatrix} 2\lambda_1 + 6\lambda_2 x & 0 \\ 0 & 2\lambda_1 + 6\lambda_2 y \end{pmatrix}$$

$$\begin{cases} \lambda_0 + 2\lambda_1 x + 3\lambda_2 x^2 = 0 \\ 2\lambda_1 y + 3\lambda_2 y^2 = 0 \\ x^2 + y^2 = 1 \\ x^3 + y^3 = 1 \\ \lambda \neq 0 \end{cases}$$



$u_1 = (0,1) \qquad u_2 = (1,0)$

$u_1 = (0,1)$

$$\begin{cases} \lambda_0 = 0 \\ 2\lambda_1 + 3\lambda_2 = 0 \end{cases}$$

среди них есть подходящие $\overline{\lambda}$, чтобы теорема выполнялась.

$\exists\ \lambda_0 = 0 \quad \lambda_1 = 3 \quad \lambda_2 = -2$

Тогда для $h \in E^n$ $h \neq 0$:

$\langle g_1'(u_1), h \rangle = 2h_2 = 0$

$\langle g_2'(u_1), h \rangle = 3h_2 = 0$

Получаем $\langle \mathcal{L}''_{uu}(u, \overline{\lambda}^*)h, h \rangle =$

$$= \left\langle \begin{pmatrix} 6 & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} h_1 \\ h_2 \end{pmatrix}, \begin{pmatrix} h_1 \\ h_2 \end{pmatrix} \right\rangle =$$

$= 6h_1^2 > 0$

$\Downarrow$

$f'(u_1) = 0$

локального непредставления и ...

$\Downarrow$

u_1 – устр. точка

$(x+y)(x^2+y^2-xy)=1$

$(x^2+y^2)=1$

$(\overbrace{x+y}^{b})(1-xy)=1$

$(x+y)^2 - 2\overbrace{xy}^{a} = 1$

$b(1-a)=1$

$b^2-2a=1$

$-ab+b=1$

$b^2-2a=1$

$a = \frac{b^2-1}{2}$

$\frac{1-b^2}{2}b+b=1$

$-\frac{b^3}{2}+\frac{b}{2}+b=1$

$-b^3+3b-2=0$

$-b^3+b+2b-2=0$

$b(1-b^2)+2(b-1)=0$

$-b(b^2-1)+2(b-1)=0$

$(-b(b+1)+2)(b-1)=0$

Задача 2

$$f(u) = x^2 + y^2 \to extr$$

$$X = \{(x,y) \in X_0,\ g(u) = x + y - 2 \le 0\}$$

где $X_0 = \{u = (x,y) \in E^2 : x \geq 1\ y \geq 0\}$

$$\mathcal{L}(u, \bar{\lambda}) = \lambda_0 (x^2 + y^2) + \lambda(x + y - 2) \qquad \bar{\lambda} \neq 0 \quad \lambda \geq 0$$

$$(x,y) \in X_0$$

$$\mathcal{L}'_u(u, \bar{\lambda}) = (2\lambda_0 x + \lambda,\ 2\lambda_0 y + \lambda)$$

$$\mathcal{L}''_u(u, \bar{\lambda}) = \begin{pmatrix} 2\lambda_0 & 0 \\ 0 & 2\lambda_0 \end{pmatrix} = 2\lambda_0 I$$

$\forall u$ - подозр на экстремум:

$$\langle \mathcal{L}'_u(u, \bar{\lambda}), v - u \rangle \qquad \forall v \in X_0$$

$$\parallel$$

$$\begin{cases} (2\lambda_0 x + \lambda)(v_1 - x) + (2\lambda_0 y + \lambda)(v_2 - y) \geq 0 \\ \lambda(x + y - 2) = 0 \quad \bar{\lambda} = (\lambda_0, \lambda) \neq 0 \quad \lambda \geq 0 \end{cases}$$

Замечание: $\exists\ x^*, y^*$ - экстремум $\Rightarrow$ $v = (x^*, v_2)$. Тогда $(2\lambda_0 y^* + \lambda)(v_2 - y^*) \geq 0$ $\forall (x^*, v_2) \in X_0$ чтобы это выполнялось для $\forall v_2 \geq 0$ (узуен) надо чтобы $2\lambda_0 y^* + \lambda = 0$.

I сл	II сл
$u \in \operatorname{int} X_0$	$u \in \partial X_0$

Тогда

$$\left.\begin{matrix} x > 1 \\ y > 0 \end{matrix}\right. \quad \begin{matrix} (2\lambda_0 x + \lambda, 2\lambda_0 y + \lambda) = (0,0) \\ \lambda(x + y - 2) = 0 \quad (\lambda_0, \lambda) \neq 0 \quad \lambda \geq 0 \end{matrix} \qquad (1)$$

(2) $x = 1$, $y \neq 0$ $\quad \mathcal{L}'_y(u, \bar{\lambda}) = 2\lambda_0 y + \lambda = 0 \quad \lambda(x + y - 2) = 0.\ (\lambda_0, \lambda) \neq 0,\ \lambda \geq 0$

(3) $y = 0$, $x \neq 1$ $\quad \mathcal{L}'_x(u, \bar{\lambda}) = 2\lambda_0 x + \lambda = 0 \quad \lambda(x + y - 2) = 0\ (\lambda_0, \lambda) \neq 0,\ \lambda \geq 0$

(4) $x=1$, $y=0$. $\lambda(x+y-2)=0 \Rightarrow \lambda=0$ $\Downarrow$

$2\lambda_0 - v_1 = 0$ $v_1 \geq 0$

$\lambda_0 > 0$

1) $\lambda_0 = 0$ - очевидно для всех случаев противоречие.

2) $\lambda_0 \neq 0$

- Для (1) нет решений

- Для (2) $x=1$ if $\lambda=0 \Rightarrow y=0 \Rightarrow$ не подходит.

  $x=1$ if $\lambda \neq 0 \Rightarrow y=1 \Rightarrow \lambda = -2\lambda_0 > 0$ $\Downarrow$ $\lambda_0 < 0$ $\left(\begin{matrix}\lambda \geq 0 \\ \lambda \neq 0\end{matrix}\right)$

- Для (3) $y=0$ if $\lambda=0 \Rightarrow x=0$ - не подходит.

  $y=0$ if $\lambda \neq 0$ $x=2 \Rightarrow \lambda = -4\lambda_0 > 0$ $\Downarrow$ $\lambda_0 < 0$

- Для (4) $x=1$, $y=0 \Rightarrow \lambda=0$ $2\lambda_0 - v_1 \geq 0$, $\lambda_0 > 0$

- для $u_1 = (1,0)$ $\lambda=0$ $\lambda_0=1$. Тогда $\langle \mathcal{L}''_{uu}(u_1,\bar{\lambda})h,h\rangle = 2\|h\|^2 > 0$

  $\Rightarrow (1,0)$ - строгий лок мин.

- для $u_2 = (2,0)$, $u_3 = (1,1)$ $\langle \mathcal{L}''_{uu}(u_i,\bar{\lambda})h,h\rangle = -2\|h\|^2 < 0$ для $\forall h \neq 0$

  $\Rightarrow$ теорема не применима

Воспользуемся теор. Вейерштр. для компактного мн-ва.

$$f(x_1)=1 < f(x_2)=2 < f(x_3)=4.$$

$\Downarrow$

x_1 – точка глобального минимума

x_3 – точка глобального максимума

Что с x_2 – ?
$x_2 = (1,1)$

$$\ast \left(1, 1-\frac{1}{k}\right) \text{ и } \left(1+\frac{1}{k}, 1-\frac{1}{k}\right) \quad k=1,2\ldots$$

$$f\left(1, 1-\frac{1}{k}\right)=1+\left(1-\frac{1}{k}\right)^2 < 2 < \left(1+\frac{1}{k}\right)^2+\left(1-\frac{1}{k}\right)^2=2+\frac{2}{k^2}$$

$\Downarrow$

x_2 – не является т. экстремума

Условия Каруша-Куна-Таккера

Поговорим про регулярность:

Опр. З-чу $(\ast)$ называют **регулярной**, если $\lambda_0^* > 0$, тогда ф-ция Лагранжа называется <u>нормальной</u> и имеет вид:

$$\mathcal{L}(x,\lambda)=f(x)+\sum_{i=1}^{s}\lambda_i g_i(x), \text{ где}$$

$x \in X_0$, а $\lambda=(\lambda_1,\lambda_2,\ldots,\lambda_s) \in \Lambda_0 = \{\lambda=(\lambda_1,\ldots,\lambda_s)\in E_s : \lambda_1 \geq 0,\ldots,\lambda_m \geq 0\}$.

Что же такое условие ККТ? (здесь считаем, что з-ча задана в смысле условий т№2)

По сути условие ККТ это правило множителей Лагранжа при условии, что з-ча регулярная.

Зам 1. Все замечания для правил. мн. Лагранжа (ПМЛ) здесь справедливы

Опр Мн-во $X \subset \mathbb{R}^n$ наз. полиэдром, если оно представляет собой:

$$X = \{x \in \mathbb{R}^n \mid \underset{\in \mathbb{R}^{m\times n}}{A}x \le b\} = \{x \in \mathbb{R}^n \mid \langle a_i, x\rangle \le b_i,\ i=1,\dots,m\}$$

Замечание: 1) Может быть равенства $(\langle a_i, x\rangle = b_i)$

Примеры: 1) $X = E^n$ 2) $X = E^n_+$ 3) $X = \{x=(x_1,\dots,x_n) : x_i \ge 0\ \forall i \in I \subset \{1,2,\dots,n\}\}$

4) $X = \{x = (x_1,\dots,x_n) : \alpha_i \le x^i \le \beta_i,$ $\alpha_i \le \beta_i$ причем некоторые $\alpha_i = -\infty$ или $\beta_j = +\infty$.

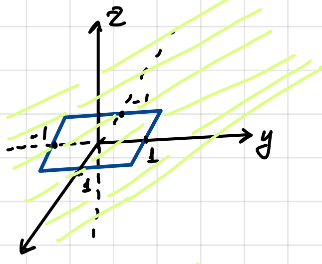
(affine hull)

Опр Аффинной оболочкой мн-ва $C \subseteq \mathbb{R}^n$ наз.

$\forall k \quad \theta_i \in \mathbb{R}$

$$\mathrm{aff}\, C = \{\theta_1 x_1 + \dots + \theta_k x_k \mid \forall x_1,\dots,x_k \in C,\ \theta_1+\dots+\theta_k = 1\}$$

Замечание:



aff-hull

1. По инд-ции доказывается, что

$$\{\theta_1 x_1 + \dots + \theta_k x_k \mid \forall x_1,\dots,x_k \in C,\ \theta_1+\dots+\theta_k = 1\} =$$

$$= \{\theta x_1 + (1-\theta) x_2 \mid \forall x_1, x_2 \in C \quad \theta \in \mathbb{R}\}$$

2. affine hull это наименьший affine set содержащий C : if S - aff.set : $C \subseteq S \Rightarrow$
$\Rightarrow \operatorname{aff} C \subseteq S$.

Опр. Аффинное множество (affine set) $C \subseteq \mathbb{R}^n$ if
$\forall x_1, x_2 \in C$ и $\theta \in \mathbb{R}$ $\theta x_1 + (1-\theta)x_2 \in C$.

Опр Relative interior (relint, ri) $C \subseteq \mathbb{R}^n$

$$\operatorname{relint}(C) = \{x \in C \mid B(x,r) \cap \operatorname{aff} C \subseteq C \text{ для какого-то } r > 0\}$$

$B(x,r) = \{y \mid \|y - x\| \leq r\}$ $\|\cdot\|$ - $\forall$ нормы т.к. они эквивалентны.

$\Downarrow$

Одинаковый relint

Замечание: ① $\operatorname{relint}(E^n) = E^n$ ② if $\operatorname{int} C \neq \emptyset$, то $\operatorname{int} C = \operatorname{ri} C$
$\Rightarrow$ ③ у открытого мн-ва $\operatorname{int} C = C \neq \emptyset = C = \operatorname{ri} C$.

Условие регулярности:

1) $\exists X_0$ - выпукло в $(*)$
2) $f, g_1, \ldots, g_m$ - дифф-мы в $x_* \in \bar{X}$
3) $g_1, \ldots, g_m$ - выпуклы на $\bar{X}_0$
4) $g_{m+1}, \ldots, g_s$ - линейны. (лин. ф-ция - это $\langle a, x \rangle + b$)

Если выполнено хотя бы одно из условий в исследуемой т. x_*, то $\exists\ \bar{\lambda} = (\lambda_0, \ldots \lambda_s)$ для нее : удовл усл ККТ ($\lambda_0 = 1$)

Усл.: a) $m' = s$ (нет огр. р-в) и $\exists\ \bar{x} \in X_0 : g_i(\bar{x}) < 0$
$\forall i = 1, \ldots m$

Условие Слейтера (SC)

Условие линейности (LCQ) → b) X_0 - полиэдр, $g_1, \dots, g_m$ - линейные

Модифицированное условие Слейтера
c) X_0 - полиэдр, $g_{\ell+1}, \dots, g_m$ $(0 \le \ell \le m)$ – линейные

и $\exists$ т. $\bar{x} \in X$: $g_i(\bar{x}) < 0$ для $\forall i = 1, \dots, \ell$

d) $g_{\ell+1}, \dots, g_m$ $(0 \le \ell \le m)$ - линейны и $\exists \bar{x} \in \underline{ri\, X_0} \cap X$: $g_i(\bar{x}) < 0$ при всех $i = 1, \dots \ell$)

Замечание: 1) рав-в и нерав-в может не быть
2) if $X_0 = \mathbb{R}^n$, то c и d совпадают

Опр Выпуклой з-чей наз (*), в которой f, g_i $i = 1, \dots m$ - выпуклы и X_0 - выпуклое мн-во, а g_j $j = m+1 \dots s$ - линейные.

Th (Теорема Куна-Таккера в дифференциальной форме)

Пусть исх. задача является $\underline{\text{выпуклой}}$. Тогда если для набора переменных $X \times \mathbb{R}^m_+ \times \underset{(m+1)}{\mathbb{R}} \times \dots \times \underset{(s)}{\mathbb{R}}$ выполнены все усл. ККТ, то x_* – точка глобального минимума.

Задачи на ККТ.

(1.7)

$$\begin{cases} (x-4)^2 + (y-4)^2 \to \min \\ x+y \le 4 \\ x+3y \le 9. \end{cases}$$

Задача выпуклая

(+)

Все линейно

$\Downarrow$

ККТ - необх и дост.

$$\mathcal{L} = (x-4)^2 + (y-4)^2 + \lambda_1 (x+y-4) + \lambda_2 (x+3y-9)$$

Усл. ККТ

$$\begin{cases} 2(x-4) + \lambda_1 + \lambda_2 = 0 \\ 2(y-4) + \lambda_1 + 3\lambda_2 = 0. \\ x+y \le 4 \\ x+3y \le 9 \\ \lambda_1 \geqslant 0 \\ \lambda_2 \geqslant 0 \\ \lambda_1 (x+y-4) = 0 \\ \lambda_2 (x+3y-9) = 0 \end{cases}$$

Последовательно перебираем λ_1 и λ_2.

(1) $\exists\ \lambda_1 > 0$ и $\lambda_2 > 0$

тогда $\begin{cases} x+y=4 \\ x+3y=9 \end{cases} \Rightarrow \begin{matrix} x = \frac{3}{2} \\ y = \frac{5}{2} \end{matrix} \Rightarrow$

Подставляем в первые два:

$$\begin{cases} (3-8) + \lambda_1 + \lambda_2 = 0 \\ (5-8) + \lambda_1 + 3\lambda_2 = 0 \end{cases}$$

$\lambda_2 = -1$ - противореч.

(2) Пусть $\lambda_1 = 0$ $\lambda_2 = 0$. Тогда $x = 4$, $y = 4$ - противоречие.

(3) $\lambda_1 > 0$ $\lambda_2 = 0$

$$\begin{cases} x + y - 4 = 0 \\ 2(x-4) + \lambda_1 = 0 \\ 2(y-4) + \lambda_1 = 0 \end{cases}$$

$$\begin{cases} -2y + \lambda_1 = 0 \\ 2y - 8 + \lambda_1 = 0 \end{cases}$$

$4y = 8$

$y = 2$

$x = 2$ $\Rightarrow$ решение подходит.

$\lambda_1 = 4$

$\lambda_2 = 0$

(4)
$$\begin{cases} x + 3y - 9 = 0 \\ 2(x-4) + \lambda_1 = 0 \\ 2(y-4) + \lambda_1 = 0 \end{cases}$$

$$\begin{cases} 2(5-3y) + \lambda_1 = 0 \\ 2(y-4) + \lambda_1 = 0 \end{cases}$$

$x = 9 - 3y = 9 - \frac{27}{4}$

$10 - 6y - 2y + 8 = 0$

$-8y = -18$

$y = \frac{9}{4}$ $x = \frac{36 - 27}{4} = \frac{9}{4}$ -

- противоречие -

(2) $\begin{cases} \min \frac{1}{2}\|x\|_2^2 \\ Ax - b = 0 \end{cases}$ – выпуклая + линейные огр.

$L = \frac{1}{2}\|x\|_2^2 + \lambda^T(Ax-b)$ $\left\{ \lambda^T Ax = \overbrace{(x, A^T\lambda)}^{*} \right\}$ $d^* = A^T\lambda$

$\nabla_x L = x + A^T\lambda = 0$

$\begin{cases} x = -A^T\lambda \\ Ax - b = 0 \end{cases} \Rightarrow -AA^T\lambda = b \quad (\lambda = -(AA^T)^{-1}b)$

$\Downarrow$

$x^* = A^T(AA^T)^{-1}b$

(3) Энтропия $E \log p(\xi)$ – энтропия.

Будем макс энтр $\Rightarrow$

$\begin{cases} \min \sum_1 x_i \log x_i \\ \sum_1 x_i = 1 \\ x_i \geq 0 \end{cases}$

$\mathcal{L} = \sum_{i=1}^n x_i \log x_i + \lambda\left(\sum_{i=1}^n x_i - 1\right) - \sum_{i=1}^n \mu_i x_i$

$[\nabla_x \mathcal{L}]_k = \log x_k + \frac{x_k}{x_k} + \lambda - \mu_k = 0$

$-\mu_i x_i = 0 \quad \mu_i \geq 0$

$\sum_1 x_i = 1 \quad x_i \geq 0$

$x_k = \exp(\mu_k - \lambda - 1) > 0$

$\Downarrow$

$\mu_i = 0$

$$\Downarrow$$

$$x_k = \exp(-\lambda - 1)$$

т.к. $\sum x_k = 1$, то

$$1 = n \exp(-\lambda - 1)$$

$$\Downarrow$$

$$\exp(-\lambda - 1) = \frac{1}{n}$$

$$\Downarrow$$

$$x_k = \frac{1}{n}.$$

Замечание: если p огр. типа нер-в, то в общем случае надо перебрать 2^p вариантов

$\lambda g(x) = 0$ $\quad \lambda = 0$ и $g(x) = 0$

(4) $\|x - x_0\|_2^2 \to \min.$ (Проекция на евклидов шар)

$X = \{x \in \mathbb{R}^n \mid \|x\| \leq 1\}$ $\quad \exists\, x_0 \notin X$

$\|x\|^2 - 1 \leq 0$

Выпуклая ф-ия + выполнено усл. Слейтера.

$$\mathcal{L} = \|x - x_0\|_2^2 + \lambda(\|x\|^2 - 1)$$

$$\begin{cases} \nabla \mathcal{L}_x = 2(x - x_0) + 2\lambda x = 0 \\ \lambda \geq 0 \\ \|x\|^2 \leq 1 \\ \lambda(\|x\|^2 - 1) = 0 \end{cases}$$

if $\lambda = 0 \Rightarrow x = x_0 = \|x\|^2 = \|x_0\|^2 \leq 1.$

if $\lambda > 0 \Rightarrow \|x\|^2 = 1$

$\nabla L_x \quad (\cdot, x)$

$2\|x\|^2 - 2\langle x, x_0\rangle + 2\lambda\|x\|^2 = 0$

$$\Downarrow \quad \lambda = \langle x, x_0\rangle - 1$$

$$(x - x_0) + (\langle x, x_0\rangle - 1)x = 0$$

$$x - x_0 + (x^T x_0)x - \cancel{x} = 0$$

$$\Downarrow$$

$$\underbrace{(x^T x_0)}_{\text{скаляр}} x = x_0.$$

$$\Downarrow$$

$$x = \lambda' x_0 \quad |\cdot \|\cdot\|$$

$$\underbrace{\|x\|}_{=1} = \lambda' \|x_0\|$$

$$\lambda' = \frac{1}{\|x_0\|} \Rightarrow x = \frac{x_0}{\|x_0\|} \text{ if } \|x_0\| > 1$$

Теория двойственности.

Будем рассматривать регулярные з-чи $(\lambda_1, \dots \lambda_s)$

$$\text{(*)} \quad \begin{cases} f(x) \to \min\limits_{x \in X} \\ X = \{x \in \mathbb{R}^n : g_1(x) \le 0, \dots, g_m(x) \le 0; \\ \qquad g_{m+1}(x) = 0, \dots, g_s(x) = 0 \} \end{cases} \qquad \Lambda_0 = \{ \lambda : \lambda_1 \ge 0, \dots \lambda_m \ge 0, \lambda_{m+1} \in \mathbb{R}, \dots \lambda_s \in \mathbb{R} \}$$

$\text{(*)}\ \mathcal{L}(x, \lambda) = f(x) + \sum\limits_{i=1}^{s} g_i(x)\lambda_i$ – нормальная ф-ция Лагранжа

Давайте предложим метод: как для $\forall$ условной з-чи оптимизации $\min\limits_{x \in X} f(x)$ записать новую задачу (двойственную), $\max\limits_{y \in \Omega} g(y)$; что прямая задача оценивается двойственной снизу: $g(y) \le f(x) \ \forall \ y \in \Omega, x \in X$

Поскольку нер-во справедливо для всех допустимых $x, y \Rightarrow$

$$\Rightarrow \qquad \max_{y \in \Omega} g(y) \le \min_{x \in X} f(x)$$

(+) Двойственной з-чи.

(a) Построение оценки снизу на решение прямой з-чи.

Исходную может быть сложно решить. Но если есть двойственная к ней, то мы можем взять $\forall y \in \Omega$ и подставить в $g(y)$ – получим некоторую оценку $g(y) \le \min\limits_{x \in X} f(x)$

(максимизация естественное желание)

(b) Проверка на допустимость з-чи и огр-ничений:

Из нер-ва $\max\limits_{y \in \Omega} g(y) \le \min\limits_{x \in X} f(x) \Rightarrow$ if $\min\limits_{x \in X} f(x) = -\infty \Leftrightarrow$

$\Leftarrow\!\!> \Omega = \varnothing$

(c) Двойственную з-чу бывает решить легче, чем прямую.

При этом if выполнена сильная двойственность:
$g(y^*) = f(x^*)$, то ничего не теряем.

(d) Получение оценки сверху на невязку по ф-ции:

$\exists$ $f^* = f(x^*)$. Тогда из того что мы требуем от двойственной ф-ции $g(y)$ $\Rightarrow$

$$f^* \geq g(y)$$

$$-f^* \leq -g(y) \quad \Big| + f(x)$$

$$f(x) - f^* \leq \underbrace{f(x) - g(y)}_{\text{зазор}} \quad \forall y \in \Omega, \ \forall x \in X$$

Опишем возможный способ построения двойственной з-чи, когда $X = \{x \in E^n, g_1(x) \leq 0, \ldots, g_m(x) \leq 0, g_{m+1}(x) = 0, \ldots, g_s(x) = 0\}$

Опр **Двойственной ф-цией** к $f(x)$ наз. Ф-цию $g(\lambda) : \in$, $\lambda = (\lambda_1, \lambda_2, \ldots, \lambda_s)$

$$g(\lambda) = \inf_{x \in \mathbb{R}^n} \mathcal{L}(x, \lambda), \quad \lambda = (\lambda_1, \lambda_2, \ldots, \lambda_s)$$

Утв 1 Для $\forall$ $x \in X$, $\lambda \in \Lambda_0$ верно:

{т.к $\lambda_i \geq 0$ $g_i \leq 0$} $f(x) \geq g(\lambda)$

$$\forall x \in X \quad f(x) \geq f(x) + \sum_{i=1}^{m} \lambda_i g_i(x) = \mathcal{L}(x, \lambda) \geq \inf_{x' \in \mathbb{R}^n} L(x', \lambda) = g(\lambda)$$

Утв2 $g(\lambda)$ – вогнута

$$-g(\lambda) = -\inf_{x\in\mathbb{R}^n} \mathcal{L}(x,\lambda) = \sup_{x\in\mathbb{R}}(-\mathcal{L}(x,\lambda))$$

1) $-\mathcal{L}(x,\lambda)$ – линейна отн λ => выпуклая

2) $\sup(\cdot)$ – выпуклая ф-ция

[$f(x) = \sup_{i\in I} f_i(x)$ $x\in X$ ← выпуклое мн-во; I ← семейство выпуклых ф-ций

Док-во

$x_\alpha = x\alpha + (1-\alpha)v$ $\forall x,v$

$\sup |$ $f_i(x_\alpha) \le \alpha f_i(x) + (1-\alpha) f_i(v)$ ■

$\Downarrow$

$-g(\lambda)$ – выпуклая => $g(\lambda)$ – вогнутая.

Двойственной з-чей будем наз-ть з-чу вида:

$$\begin{cases} g(\lambda) \to \max_{\lambda[1:m]\in\mathbb{R}^m_+,\ \lambda[m+1:p]\in\mathbb{R}^{p-m+1}} \\ \lambda_1 \ge 0, \dots \lambda_m \ge 0 \end{cases}$$

Получим выпуклую задачу оптимизации (при том, что исходная может быть не выпуклой)

Из определения дв. з-чи => $\underbrace{g(\lambda^*) \le f(x^*)}_{\text{слабая двойственность}}$

if $g(\lambda^*) = f(x^*)$, то это сильная двойственность

Условие сильной двойственности:

1) Выпуклая з-ча + условие Слейтера

Weak Slater 2) Выпуклая з-ча + $g_i(\bar{x}) < 0$ для тех ф-ций типа неравенств, которые не вида $\langle a, x\rangle + b$ (не линейные)

Замечание:

Если выполнена сильная двойственность, то решение прямой з-чи x_* и двойственной (λ^*, μ^*) тоже связаны.

$$x_* \in \operatorname{Arg}\min_{x \in \mathbb{R}^n} \mathcal{L}(x, \lambda^*, \mu^*)$$

Задачи на двойственность:
Построить двойственную ф-цию.

(1) $\begin{cases} \|x\|^2 \to \min\limits_x \\ Ax = b \end{cases}$

Договоримся что $(\underbrace{\lambda_1, \dots \lambda_m}_{\mu}, \underbrace{\lambda_{m+1}, \dots \lambda_s}_{\nu})$

$$\mathcal{L}(x, \nu) = \|x\|^2 + \nu^T (Ax - b)$$

$$\nabla \mathcal{L}_x = 2x + A^T \nu = 0$$

$$x = -\frac{1}{2} A^T \nu$$

$$\Downarrow$$

т.к. $g(\lambda) = \inf\limits_x \mathcal{L} = \frac{1}{4} \nu^T A A^T \nu + -\frac{1}{2} \nu^T A A^T \nu - \nu^T b =$

$$= -\frac{1}{2} \nu^T A A^T \nu - \nu^T b.$$

Получаем двойств. задачу

$$-\frac{1}{2} \nu^T A A^T \nu - \nu^T b \to \max_\nu$$

Дифф-руем

$$\varphi'(\nu) = -AA^T\nu - b$$

$\Downarrow$ т.к. з-ча безусл.

$$AA^T \nu = -b$$

т.к. усл. Слейтера выполн

то $\exists \lambda^*$: $AA^T\lambda^* = -b$

и находим решение x $x = -\frac{1}{2}A^T\lambda^*$

(2) $\begin{cases} c^T x \to \min \\ Ax = b \\ Ex \le d \end{cases}$ $\Rightarrow$ усл. Слейтера выполнено

$$\mathcal{L} = c^T x + \lambda^T (Ex - d) + \mu^T (Ax - b)$$

$$\nabla_x \mathcal{L} = c + E^T \lambda + A^T \mu = 0.$$

$$q(\lambda, \mu) = \begin{cases} -\lambda^T d - \mu^T b & \nabla \mathcal{L} = 0 \\ -\infty & \nabla \mathcal{L} \neq 0. \end{cases}$$

Двойств з-ча $-\infty$ исключается.

$$\begin{cases} \lambda^T d + \mu^T b \to \min\limits_{\lambda, \mu} \\ \lambda \ge 0 \\ c + E^T \lambda + A^T \mu = 0 \end{cases}$$

(Здесь x^* не можем найти((()