

$$\begin{aligned}
 & \min_{x \in \mathbb{R}^d} f_0(x) \\
 \text{s.t.} \quad & f_i(x) \leq 0 \quad i=1 \dots m \\
 & Ax = b
 \end{aligned}$$

$\in \mathbb{R}^{n \times d}$        $\mathbb{R}^n$

Лагранжиан:

$$L(x, \lambda, \nu) = f_0(x) + \sum_{i=1}^n \lambda_i f_i(x) + \nu^T (Ax - b)$$

$\lambda_i \geq 0 \quad (\text{вектор } \lambda)$

Двойственная:

$$g(\lambda, \nu) = \inf_{x \in \mathbb{R}^d} L(x, \lambda, \nu)$$

об. ло:

$$g(\lambda, \nu) \leq f_0(x^*) \quad \forall \lambda \geq 0 \quad \nu \in \mathbb{R}^n$$

$\uparrow$   
решение

### Условие Слейтера

Будем говорить, что для задачи с ограничениями выполняется условие Слейтера, если существует  $x \in \mathbb{R}^d$ , такой что

$$f_i(x) < 0, \quad i = 1, \dots, m \quad \text{и} \quad Ax = b.$$

### Теорема Слейтера

Если в задаче с ограничениями все функции являются выпуклыми и выполняется условие Слейтера, то тогда при построении двойственной задачи выполняется свойство сильной двойственности, а именно

$$\sup_{\lambda \geq 0, \nu \in \mathbb{R}^n} g(\lambda, \nu) = f_0(x^*).$$

## Седловая точка

Точка  $(x^*, \lambda^*, \nu^*) \in \mathbb{R}^d \times \mathbb{R}_+^m \times \mathbb{R}^n$  называется седловой для функции  $L(x, \lambda, \nu)$ , если для любых  $(x, \lambda, \nu) \in \mathbb{R}^d \times \mathbb{R}_+^m \times \mathbb{R}^n$  выполнено

$$L(x, \lambda^*, \nu^*) \geq L(x^*, \lambda^*, \nu^*) \geq L(x^*, \lambda, \nu).$$

## Теорема о седловой точке Куна-Таккера

Для задачи выпуклой оптимизации с выпуклыми ограничениями с выполненным условием Слейтера следующие утверждения эквивалентны:

- для  $x^*$  существует  $\lambda^* \succeq 0$  и  $\nu^* \in \mathbb{R}^n$  такие, что  $(x^*, \lambda^*, \nu^*)$  – седловая точка функции Лагранжа,
- $x^*$  – глобальное решение задачи оптимизации с ограничениями.

Док-во:

$\Rightarrow (x^*, \lambda^* \succeq 0, \nu^*)$  – седловая точка Лагранжа

- $x^*$  – узловая сущн? от граничного

$$f_i : S_i(x^*) \geq 0 \quad (Ax \neq b)$$

$$\Rightarrow \sup_{\lambda \succeq 0, \nu} L(x^*, \lambda, \nu) = +\infty$$

$$L(x^*, \lambda^*, \nu^*) \geq L(x^*, \lambda, \nu) \text{ по опт. седла}$$

$$! \lambda^* \rightarrow \lambda^* = (\lambda_{i-1}^*, 2\lambda_i^*, \lambda_{i+1}^*)$$

$x^*$  – узловая сущн

$$\bullet S_0(x^*) = \sup_{\lambda \succeq 0, \nu} L(x^*, \lambda, \nu)$$

условие Слейтера  
сущн:

$$L(x^*, \lambda^*, \nu^*) \geq L(x^*, \lambda, \nu)$$

$$|| S_0(x^*)$$

сего:

$$L(x, \lambda^*, \gamma^*) \geq L(x^*, \lambda^*, \gamma^*)$$

$$f_0(x) + \sum_{i=0}^n \lambda_i^* f_i(x) + (\gamma^*)^\top (Ax - b) \geq L(x^*, \lambda^*, \gamma^*)$$

$\geq 0$        $\leq 0$        $= 0$        $= f_0(x^*)$

$x$  *гол. оптимум*

$$f_0(x) - f_0(x^*) \geq 0 \quad \forall x \text{ гол. оптимум.}$$

$$f_0(x^*) - f_0(x^*) \geq 0 \quad \forall x \text{ гол. оптимум.} \quad \Rightarrow x^* \text{ -}\begin{cases} \text{гол. оптимум.} \\ \text{решение} \end{cases}$$

$\Leftarrow x^*$  - решение

условия  $\lambda^*, \gamma^*$  - решения глобальной задачи

по оптимизации:

$$f_0(x^*) = g(\lambda^*, \gamma^*) = \inf_x L(x, \lambda^*, \gamma^*)$$

множ.

$$f_0(x^*) \leq L(x^*, \lambda^*, \gamma^*)$$

$$= f_0(x^*) + \sum_{i=0}^n \lambda_i^* f_i(x^*) + (\gamma^*)^\top (Ax^* - b)$$

$\leq 0$       *решение*       $= 0$

$$L(x^*, \lambda^*, \gamma^*) = f_0(x^*) = \inf_x L(x, \lambda^*, \gamma^*)$$

*небольшое замечание*

$$L(x^*, \lambda^*, \gamma^*) \leq L(x, \lambda^*, \gamma^*)$$

$$L(x^*, \lambda^*, \nu^*) = f_0(x^*) + \underbrace{\sum \lambda_i f_i(x^*)}_{\geq 0} + \underbrace{\nu^T(Ax^* - b)}_{= 0}$$

$L(x^*, \lambda, \nu)$   
бюджетные ограничения

$$L(x^*, \lambda^*, \nu^*) \geq L(x^*, \lambda, \nu) \quad \blacksquare$$

$$L(x, \lambda) : \bar{X} \times \Lambda \rightarrow \mathbb{R}$$

Проблема оптимизации логарифм:

2 условия:

- 1 условие: логарифм  $x \in \bar{X}$
- 2 условие: логарифм  $\lambda \in \Lambda$

$L(x, \lambda)$ :

- задача о логарифме
- некоторый из всех максимумов логарифма

$$L(x, \lambda)$$

Чтобы: некоторый максимум логарифма  
бюджетные ограничения

Узлы:  $x^*, \lambda^*$

$$L(x, \lambda^*) \geq L(x^*, \lambda^*) \geq L(x^*, \lambda)$$

↑  
известные  
составлены X  
представляют ему  
максимум

↑  
аналогично

Важен ли порядок?

1) сначала 1 нер. , потом 2 нер.

$$\inf_x \sup_\lambda L(x, \lambda)$$

2) наоборот

$$\sup_\lambda \inf_x L(x, \lambda)$$

$$\inf_{\hat{x}} L(\hat{x}, \lambda) \leq L(x, \lambda)$$

$$\sup_\lambda \inf_{\hat{x}} L(\hat{x}, \lambda) \leq \sup_\lambda L(x, \lambda)$$

$$\cancel{\inf_x \sup_\lambda \inf_{\hat{x}} L(\hat{x}, \lambda)} \leq \inf_x \sup_\lambda L(x, \lambda)$$

а когда равенство?

### Теорема о седловой точке

Множество седловых точек функции  $L: \mathcal{X} \times \Lambda \rightarrow \mathbb{R}$  непустое тогда и только тогда, когда обе задачи  $\sup_\lambda \inf_x L(x, \lambda)$  и  $\inf_x \sup_\lambda L(x, \lambda)$  имеют решение и эти решения совпадают.

$$\bullet \text{ седло} = \inf \sup = \sup \inf$$

min max = max min

$$\text{а когда } \inf \sup = \sup \inf$$

максим.

миним.

## Теорема Сиона-Какутани

Пусть  $\mathcal{X}, \Lambda$  выпуклые компактные множества, пусть также  $L : \mathcal{X} \times \Lambda \rightarrow \mathbb{R}$  непрерывна, выпукла по  $x$  (для любого фиксированного  $\lambda$ ) и вогнута по  $\lambda$  (для любого фиксированного  $x$ ). Тогда  $L$  имеет седловые точки на  $\mathcal{X} \times \Lambda$ .

## Теорема Сиона-Какутани

Пусть  $\mathcal{X}, \Lambda$  выпуклые множества, и  $\mathcal{X}$  или  $\Lambda$  дополнительно компактно, пусть также  $L : \mathcal{X} \times \Lambda \rightarrow \mathbb{R}$  непрерывна, выпукла по  $x$  (для любого фиксированного  $\lambda$ ) и вогнута по  $\lambda$  (для любого фиксированного  $x$ ). Тогда (гарантий существования тут нет)

$$\sup_{\lambda \in \Lambda} \inf_{x \in \mathcal{X}} L(x, \lambda) = \inf_{x \in \mathcal{X}} \sup_{\lambda \in \Lambda} L(x, \lambda)$$

Вывод: влече седло решени

$$\min_{x \in \mathcal{X}} \max_{\lambda \in \Lambda} L(x, \lambda)$$

для  $\min_{x \in \mathcal{X}} \max_{\lambda \in \Lambda} L(x, \lambda, \lambda)$

как решать?

граф. спуск  $\Rightarrow$  граф. метод - подъем

$$x^{k+1} = x^k - \nabla_x L(x^k, \lambda^k)$$

$$\lambda^{k+1} = \lambda^k + \nabla_\lambda L(x^k, \lambda^k)$$

подъем по  $\max$

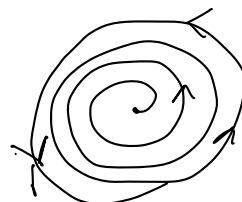
Гипотеза:  $\min_{x \in \mathbb{R}^d} \max_{\lambda \in \mathbb{R}} L(x, \lambda)$

$$x^* = 1 \quad \lambda^* = 1 \quad \text{решение: } x^* = 0 \quad \lambda^* = 0$$

$$\nabla_x L(x^*, \lambda^*) = x^* = 1 \quad \nabla_\lambda L(x^*, \lambda^*) = \lambda^* = 1$$

$$\text{глобально: } \begin{pmatrix} -1 \\ 1 \end{pmatrix}$$

$$\begin{pmatrix} -1 & 1 \\ 1 & 1 \end{pmatrix} \xrightarrow{\frac{\pi i}{2}} \begin{pmatrix} 0 & 0 \end{pmatrix}$$



расходимся

## Алгоритм 2 Экстраградиентный метод

Вход: размер шага  $\gamma > 0$ , стартовая точка  $x^0 \in \mathbb{R}^d$ , количество итераций  $K$

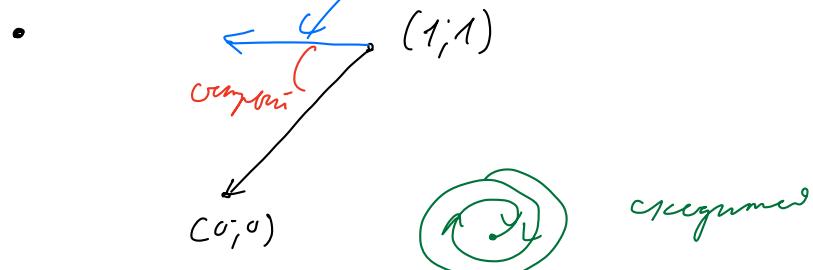
```

1: for  $k = 0, 1, \dots, K-1$  do
2:    $x^{k+1/2} = x^k - \gamma \nabla_x L(x^k, \lambda^k)$  ] неградиентный шаг
3:    $\lambda^{k+1/2} = \lambda^k + \gamma \nabla_\lambda L(x^k, \lambda^k)$  ] шаг из Баг метода
4:    $x^{k+1} = x^{k+1/2} - \gamma \nabla_x L(x^{k+1/2}, \lambda^{k+1/2})$ 
5:    $\lambda^{k+1} = \lambda^{k+1/2} + \gamma \nabla_\lambda L(x^{k+1/2}, \lambda^{k+1/2})$ 
6: end for

```

Выход:  $\frac{1}{K} \sum_{k=0}^{K-1} x^{k+1/2}, \frac{1}{K} \sum_{k=0}^{K-1} \lambda^{k+1/2}$

Итогущий: выполнение энтузиазма.



• Вспомним проекционные методы  
(где есть схема)

$$x^{k+1} = x^k - \gamma \nabla f(x^{k+1})$$

стриктический:

$$x^{k+1} = x^k - \gamma \triangleright f(x^{k+1/2})$$

### Теорема о сходимости экстраградиентного метода

Пусть дана непрерывно дифференцируемая по обеим группам переменным выпуклая-вогнутая  $L$ -гладкая функция  $L : \mathbb{R}^d \times \mathbb{R}^n \rightarrow \mathbb{R}$ , тогда для экстраградиентного метода справедлива следующая оценка сходимости для любого  $u \in \mathbb{R}^d \times \mathbb{R}^n$  и для любого  $\gamma \leq \frac{1}{L}$ :

$$\left( L\left(\frac{1}{K} \sum_{k=0}^{K-1} x^{k+1/2}, u_\lambda\right) - L\left(u_x, \frac{1}{K} \sum_{k=0}^{K-1} \lambda^{k+1/2}\right) \right) \leq \frac{\|z^0 - u\|_2^2}{2\gamma K}$$

1)  $\lambda$  не бико

$$L(\hat{x}^k) - L(u_x) \leq \dots$$

$$\text{а если } \lambda \text{ есть } u_\lambda \rightarrow \lambda^* \quad u_x \rightarrow x^*$$

$$L(x^k, \lambda^*) - L(x^*, \lambda^k) = 0$$

пример наименьшего?

$$\min_{x^*} \max_{\lambda} (x-1)(\lambda+1)$$

решение:  $x=1 \quad \lambda=-1 \quad (x^*-1)(\lambda+1)=0$

$$(x-1)(\lambda^*+1)=0$$

2) Числодук:

$$\max_{\lambda} L(x^k, \lambda) - \min_x L(x, \lambda^k)$$

Задача:

$$\min_x \max_{\lambda} f(x) - \lambda^T A x - g(\lambda)$$

↑  
внешняя  
↑  
внутренняя

Примеч:

- $\min f(x)$   
s.t.  $Ax = b$

$$\min_x \max_{\lambda} L(x, \lambda) = f(x) - \lambda^T A x$$

- $\min f(x) + L(Ax)$   
репл.      ищ.  
гамма

если  $L$  выпуклая.

$$L(Ax) = L^*(Ax) = \max_{\lambda} \{ -L^*(\lambda) + \lambda^T A x \}$$

$$\min_x \max_{\lambda} f(x) - L^*(\lambda) + \lambda^T A x$$

Метод:

---

#### Алгоритм 4 Прямо-двойственный алгоритм

---

**Вход:** стартовая точка  $x^0 \in \mathbb{R}^d$ , количество итераций  $K$

- 1: **for**  $k = 0, 1, \dots, K-1$  **do**
- 2:      $x^{k+1} = x^k - \eta (\nabla f(x^k) - A^T \lambda^k)$
- 3:      $\lambda^{k+1} = \lambda^k - \eta (\nabla g(\lambda^k) + A(2x^{k+1} - x^k))$
- 4: **end for**

**Выход:**  $\frac{1}{K} \sum_{k=1}^K x^k, \frac{1}{K} \sum_{k=1}^K \lambda^k$

если  $x^k$  присвоено  $\Rightarrow$  сух - погреш.

$x^k \rightarrow 2x^{k+1} - x^k$  ← экстраполация