

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

$\swarrow \in \mathbb{R}^{n \times d} \quad \nwarrow \mathbb{R}^n$

Лагранжиан:

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

\uparrow
 $\lambda_i \geq 0 \quad (\text{вероятно})$

Двойственная:

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

св-во:

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

\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 эквиваленты:

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

Доказ.

$\Rightarrow (x^*, \lambda^* \geq 0, \nu^*)$ – седловая точка Лагранжиана

- x^* – глоб. оптим.? от предполож. $\exists i: S_i(x^*) > 0 \quad (Ax^* \neq b)$

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

$$L(x^*, \lambda^*, \nu^*) \geq L(x^*, \lambda, \nu) \text{ из отрез. сегм.}$$

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

x^* – глоб. оптимальным

$$\cdot \quad J_0(x^*) = \sup_{\lambda \geq 0, \nu} L(x^*, \lambda, \nu)$$

условие Слейтера
сегм.:

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

$$\parallel \\ J_0(x^*)$$

целю:

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

$$f_0(x) + \sum_{\substack{\geq 0 \\ \nearrow \leq 0}} \lambda_i^* f_i(x) + (J^*)^T (Ax - b) \geq L(x^*, \lambda^*, J^*) = f_0(x^*)$$

$\nearrow = 0$

x глоб. оптимальным

$$f_0(x) - \underbrace{\sum_{\geq 0}}_{\geq 0} \geq f_0(x^*) \quad \forall x \text{ глоб. оптимальным}$$

x^* - глоб. оптимальный

$$f_0(x) \geq f_0(x^*) \quad \forall x \text{ глоб. оптимальным} \quad \Rightarrow x^* - \text{решение}$$

$\Leftarrow x^*$ - решение

иногда λ^*, J^* - решение глобальной задачи

по Силверману:

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

мы знаем

$$f_0(x^*) \leq L(x^*, \lambda^*, J^*) = f_0(x^*) + \sum_{\substack{\geq 0 \\ \nearrow \leq 0}} \lambda_i^* f_i(x^*) + (J^*)^T (Ax^* - b) = f_0(x^*) + \underbrace{\sum_{\geq 0}}_{=0} \lambda_i^* f_i(x^*) + \underbrace{(J^*)^T (Ax^* - b)}_{=0}$$

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

вероятно минимальное значение

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

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

$L(x^*, \lambda, \nu)$
всегда такое же

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

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

Примеры свойств:

2 свойства: 1 свойс. достигается $x \in \bar{X}$
2 свойс. достигается $\lambda \in \Lambda$

$L(x, \lambda)$: значение на границе
первого свойства максимум достигается
 $L(x, \lambda)$

цель: первое значение
второе значение

Ищем цель: x^*, λ^*

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

↑
значение
свойства x
при этом
мы

↑
анализируем

Важны ли порядок?

1) сначала 1 порядок, потом 2 порядок

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

2) наоборот

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

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

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

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

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.

- $\text{седло} = \inf \sup = \sup \inf$
 $\min \max = \max \min$

а когда $\inf \sup = \sup \inf$

миним.

максим.

Теорема Сиона-Какутани

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

Теорема Сиона-Какутани

Пусть \mathcal{X}, Λ выпуклые множества, и \mathcal{X} или Λ дополнительно компактно, пусть также $L : \mathcal{X} \times \Lambda \rightarrow \mathbb{R}$ непрерывна, выпукла по x (для любого фиксированного λ) и вогнута по λ (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: в нем себе решаем

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

Лагранж $\min_x \max_{\lambda \in [0,1]} L(x, \lambda, \lambda)$

Как решить?

граду. спуск \Rightarrow граду. спуск - поочередно

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

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

поочередно по max

x^{k+1}

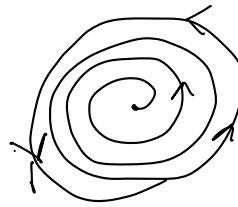
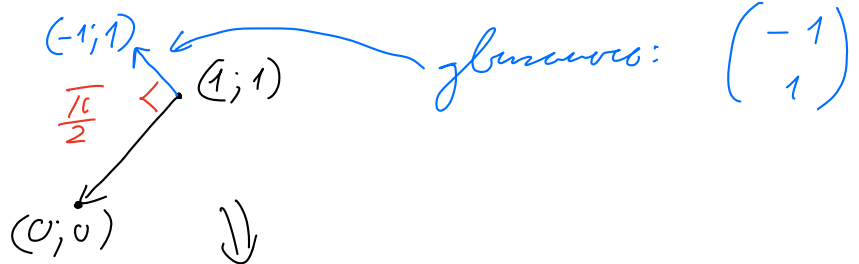
Пример: $\min_{x \in \mathbb{R}^2} \max_{\lambda \in \mathbb{R}} x \lambda$

$$x^0 = 1 \quad \lambda^0 = 1$$

решение: $x^* = 0 \quad \lambda^* = 0$

$$\nabla_x L(x^0, \lambda^0) = \lambda^0 = 1$$

$$\nabla_\lambda L(x^0, \lambda^0) = x^0 = 1$$



расхождение

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: $x^{k+1/2} = x^k - \gamma \nabla_x L(x^k, \lambda^k)$

3: $\lambda^{k+1/2} = \lambda^k + \gamma \nabla_\lambda L(x^k, \lambda^k)$

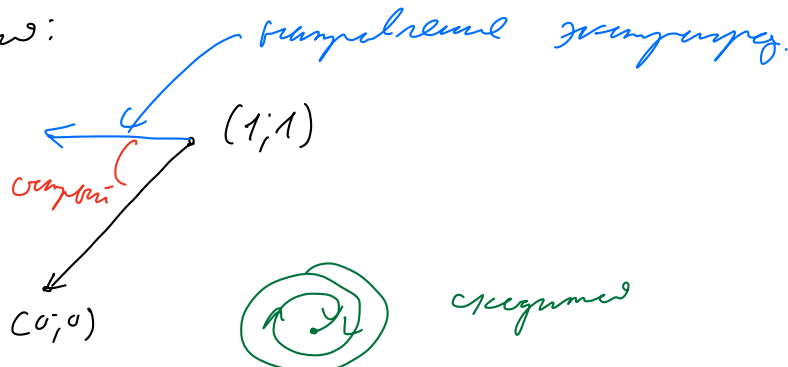
4: $x^{k+1} = x^{k+1/2} - \gamma \nabla_x L(x^{k+1/2}, \lambda^{k+1/2})$

5: $\lambda^{k+1} = \lambda^{k+1/2} + \gamma \nabla_\lambda L(x^{k+1/2}, \lambda^{k+1/2})$

6: **end for**

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

Примечание:



• величина проекционных методов (разновидности)

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

итерация:

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

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) λ не fixed

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

а есть λ есть $u_\lambda \rightarrow \lambda^* \quad u_x \rightarrow x^*$

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

пример поиска минимума?

$$\min \max (x-1)(\lambda+1)$$

решение: $x=1 \quad \lambda=-1$

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

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

2) Итоговое:

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

Задача:

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

↑ *выпуска* ↑ *выпуска*

Примеры:

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

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

- $\min f(x) + L(Ax)$

↑ *разр* ↑ *вес* ↑ *генери*

если L само сопряжен.

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

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

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 точно \Rightarrow субс - оптимальн
 $x^k \rightarrow 2x^{k+1} - x^k \leftarrow$ экстраградиент