Седловые задачи

$$\min_{x \in \mathbb{R}^d} f_0(x)$$

$$\text{s.t.} \quad f_i(x) \le 0 \qquad i = 1..m$$

$$Ax = b \qquad A \in \mathbb{R}^{n \times d} \quad b \in \mathbb{R}^n$$

- **Лагранжиан**

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

вектор из λ_i

$$\lambda_i \ge 0 \qquad \nu \in \mathbb{R}^n$$

- **Двойственная функция**

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

В общем случае $\forall \lambda_i \ge 0 \quad \forall \nu \in \mathbb{R}^n$

$$g(\lambda, \nu) \le f_0(x^*)$$

не хорошо :-(

$g(\lambda^*, \nu^*)$ может $< f_0(x^*)$

Условие Слейтера

$$\exists\, x \in \mathbb{R}^d : \quad f_i(x) < 0 \qquad i = 1..m$$

$$Ax = b$$

Теорема (Слейтера)

Если $f_0, f_1 .. f_m$ выпуклые и выполнено условие Слейтера, то тогда

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

Опр. (седловая точка)

Точка $(x^*, \lambda^*, \nu^*) \in \mathbb{R}^d \times \mathbb{R}^m_+ \times \mathbb{R}^n$ – седловая точка функц $L(x, \lambda, \nu)$, если $\forall (x, \lambda, \nu) \in \mathbb{R}^d \times \mathbb{R}^m_+ \times \mathbb{R}^n$

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

Теорема (Куна-Таккера)

Если $f_0, f_1 \ldots f_m$ выпуклые и выполнено усл Слейтера, тогда

$$x^* \text{ – глоб решение}$$

$$\Longleftrightarrow$$

для x^* сущ $\lambda_i^* \geq 0, \nu^* \in \mathbb{R}^n$: (x^*, λ^*, ν^*) – седловая т. Лагранжиана

Док-во:

$\Leftarrow$ x^* удовл. огран.? от противного

пусть $\exists i : f_i(x^*) > 0$ (аналогично $Ax^* \neq b$)

$$\sup_{\lambda_j \geq 0, \nu \in \mathbb{R}^n} L(x^*, \lambda, \nu) = +\infty \quad (\lambda_i \to +\infty)$$

Опред. с. т.: $L(x^*, \lambda^*, \nu^*) \geq L(x^*, \lambda, \nu) \quad \forall \lambda_i \geq 0, \forall \nu \in \mathbb{R}^n$

$\measuredangle \ \lambda = (\lambda_1^*, \ldots \lambda_{i-1}^*, 2\lambda_i^*, \lambda_{i+1}^*, \ldots \lambda_n^*)$

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

противоречие с опред с.т.

$\boxed{x^* \text{ удовл. ограничениям}}$

$\triangleleft$ $\sup\limits_{\lambda_i \geq 0,\, \nu \in \mathbb{R}^n} L(\underbrace{x^*}_{\text{удовл. огран}}, \lambda, \nu) \overset{?}{=} f_0(x^*)$ (sup достигается $\lambda_i = 0,\ \forall \nu \in \mathbb{R}^n$)

из опред. с.т.: $L(x^*, \lambda^*, \nu^*) \geq L(x^*, \lambda, \nu)$

берем sup: $L(x^*, \lambda^*, \nu^*) = \sup L(x^*, \lambda, \nu) = f_0(x^*)$

из 2го пункта опр. с.т.:

$$L(x, \lambda^*, \nu^*) \geq L(x^*, \lambda^*, \nu^*) = f_0(x^*)$$

$$\| $$

$$f_0(x) + \underbrace{\sum_{i=1}^{m} \lambda_i^* f_i(x)}_{\leq 0} + \underbrace{\nu^{*T}(Ax - b)}_{=0} \geq f_0(x^*)$$

$\triangleleft$ x — удовл. огран

$$f_0(x) + \underbrace{\delta}_{\leq 0} \geq f_0(x^*)$$

$$f_0(x) \geq f_0(x^*) \Rightarrow x^* \text{ — глоб. мин.}$$

$\Leftarrow$ см. презентацию ■

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

1 игрок: набор действий задается $x \in \mathbb{X}$

2 игрок: $\lambda \in \Lambda$

$L(x, \lambda)$ — 1 игрок „заплатит" 2ому

Чего хотят игроки? 1 платить меньше, 2 получить больше

Формально, найти $(x^*, \lambda^*) \in \mathbb{X} \times \Lambda$

$$\underbrace{L(x,\lambda^*) \geq L(x^*,\lambda^*) \geq L(x^*,\lambda)}_{\text{седловая задача}}$$

Зависит ли игра от того, кто первый?

1) 1 игрок первый

$$\inf_{x\in \underline{X}}\left[\sup_{\lambda\in\underline{\Lambda}} L(x,\lambda)\right]$$

2) 2 игрок первый

$$\sup_{\lambda\in\underline{\Lambda}}\left[\inf_{x\in \underline{X}} L(x,\lambda)\right]$$

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

$$\inf_{\tilde{x}} L(\tilde{x},\lambda) \leq L(x,\lambda) \quad \forall x\in \underline{X}$$

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

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

<u>Теорема</u> (Сиона - Какутани)

Мн-во седл. точек $L(x,\lambda)$ непусто $\Longleftrightarrow$ $\sup\inf L(x,\lambda)$ и $\inf\sup L(x,\lambda)$ имеют решение и эти решения совпадают

Теорема (Сион-Какутани)

$\underline{X}, \underline{\Lambda}$ – выпуклые компакты

$L(x,\lambda)$ – выпукло-вогнутая функция $\underline{X} \times \underline{\Lambda}$

тогда L имеет седловые точки на $\underline{X} \times \underline{\Lambda}$

поиск седловой точки (если существует)

$$\Longleftrightarrow$$

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

$$\Longleftrightarrow$$

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

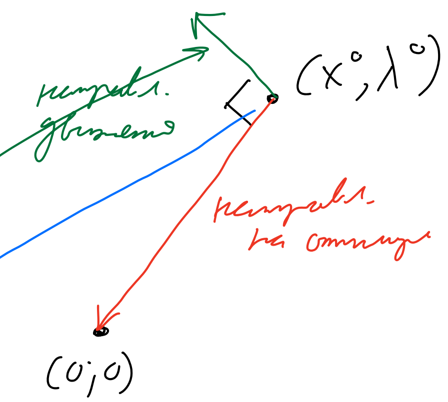
$$\min_{x\in\mathbb{R}^d} f(x) \qquad\qquad \min_{x\in\mathbb{R}^d} \max_{\lambda\in\mathbb{R}^n} L(x,\lambda)$$

$$x^{k+1} = x^k - \gamma \nabla f(x^k) \quad \Longrightarrow \quad \begin{aligned} x^{k+1} &= x^k - \gamma \nabla_x L(x^k,\lambda^k) \\ \lambda^{k+1} &= \lambda^k + \gamma \nabla_\lambda L(x^k,\lambda^k) \end{aligned}$$

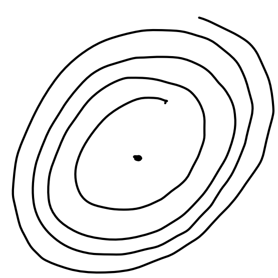
Пример $\min_{x\in\mathbb{R}} \max_{\lambda\in\mathbb{R}} x\lambda$ $\qquad x^* = 0$, $\lambda^* = 0$

старт: $(x^0, \lambda^0) \neq (0;0)$

направл: $\gamma\begin{pmatrix}-\nabla_x \\ \nabla_\lambda\end{pmatrix} = \gamma\begin{pmatrix}-\lambda^0 \\ x^0\end{pmatrix}$



$$\left\langle \gamma\begin{pmatrix}-\lambda^0 \\ x^0\end{pmatrix}; \begin{pmatrix}x^0 \\ \lambda^0\end{pmatrix}\right\rangle = 0 \qquad 90^\circ$$



• Метод Экстраградиента (Г. Корпелевич)

$$\boxed{\begin{aligned} x^{k+1/2} &= x^k - \gamma \nabla_x L(x^k, \lambda^k) \\ \lambda^{k+1/2} &= \lambda^k + \gamma \nabla_\lambda L(x^k, \lambda^k) \\ x^{k+1} &= x^k - \gamma \nabla_x L(x^{k+1/2}, \lambda^{k+1/2}) \\ \lambda^{k+1} &= \lambda^k - \gamma \nabla_\lambda L(x^{k+1/2}, \lambda^{k+1/2}) \end{aligned}}$$

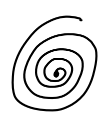
• неявная схема:

$$x^{k+1} = x^k - \gamma \nabla f(x^{k+1}) \quad \leftarrow \text{неявная схема}$$

$$x^{k+1} = x^k - \gamma \nabla f(x^k) - \gamma \nabla r(x^{k+1}) \quad \leftarrow \text{прокс. метод}$$

EG — имитация неявной схемы

Угол в примере выше между EG-направлением и направлением на решение — острый



⊕ простота, как у GD

⊕ лучше скорость, чем у GD для седловых задач

⊕ оптимален для седл. задач

⊖ два раза вычисление градиента

Решение ⊖ :

- **Оптимистичный Экстраградиент:**

$$\boxed{\begin{aligned}
x^{k+1/2} &= x^k - \gamma \nabla_x L(x^{k-1/2}, \lambda^{k-1/2}) \\
\lambda^{k+1/2} &= \lambda^k + \gamma \nabla_\lambda L(x^{k-1/2}, \lambda^{k-1/2}) \\
x^{k+1} &= x^k - \gamma \nabla_x L(x^{k+1/2}, \lambda^{k+1/2}) \\
\lambda^{k+1} &= \lambda^k - \gamma \nabla_\lambda L(x^{k+1/2}, \lambda^{k+1/2})
\end{aligned}}$$

используем старые градиенты