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

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

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

- 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 $\forall\ \lambda_i\ge 0 \quad \forall \nu\in\mathbb{R}^n$

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

не хорошо

Условие Слейтера

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

$$Ax=b$$

Теорема Слейтера

Если для задачи с выпуклыми $f_0, f_1 \dots f_m$ выполняется 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 ф. $L(x,\lambda,\nu)$, 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 для задачи оптимизации с выпуклыми $f_0, f_1 .. f_m$ выполнено условие Слейтера, то тогда

x^* — глоб. решение задачи

$$\Longleftrightarrow$$

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

Док-во:

($\Leftarrow$) x^* удовл. ограничен.? от противного

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

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

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

$$\lambda=(\lambda_1^*,\dots\lambda_{i-1}^*,\,2\lambda_i^*,\,\lambda_{i+1}^*\dots\lambda_m^*)$$

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

противоречие, т.к. (x^*,λ^*,ν^*) — с.т.

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

$\lhd$ $\sup\limits_{\lambda_i \geq 0, \nu \in \mathbb{R}^n} L(x^*, \lambda, \nu) \overset{?}{=} f_0(x^*)$ (sup достигается на любых $\lambda_i \geq 0$ $\nu \in \mathbb{R}^n$)

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

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

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

$$f_0(x) + \sum \lambda_i^* f_i(x) + (\nu^*)^T(Ax-b)$$

$$\geq f_0(x^*) + \sum \lambda_i^* f_i(x^*) + (\nu^*)^T(Ax^*-b)$$

$$= f_0(x^*)$$

$$\forall x \in \mathbb{R}^d \to \quad f_0(x) + \underbrace{\sum \lambda_i^* f_i(x)}_{\leq 0} + \underbrace{(\nu^*)^T(Ax-b)}_{=0} \geq f_0(x^*)$$

x - удовл. огран.:

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

$$\Downarrow$$

$f_0(x) \geq f_0(x^*)$ для x удовл. огран.

$\Rightarrow$ Что дает условие Слейтера:

$$f_0(x^*) = g(\lambda^*, \nu^*) = \inf_{x \in \mathbb{R}^d} L(x, \lambda^*, \nu^*)$$

решение двойств. задачи

Тогда

$$f_0(x^*) = \boxed{\inf L(x, \lambda^*, \nu^*) \leq L(x^*, \lambda^*, \nu^*)}$$

одна часть из опред. седловой задачи

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

$$f_0(x^*) \le L(x^*,\lambda^*,\nu^*) = f_0(x^*) + \le 0$$

$$\boxed{= f_0(x^*)}$$

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

$$\|$$

$$L(x^*,\lambda^*,\nu^*) + \le 0 = L(x^*,\lambda,\nu)$$

$$\boxed{L(x^*,\lambda^*,\nu^*) \ge L(x^*,\lambda,\nu)} \qquad \forall \lambda_i \ge 0,\ \forall \nu \in \mathbb{R}^n$$

Второй игрок стр. с.м. ∎

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

1 игрок: выбор действия, зависит от $x \in X$

2 игрок: $\lambda \in \Lambda$

$L(x,\lambda)$ – прибыль игрока 2 при выборах действий (убыль для 1 игрока) x от 1 игрока и λ от 2 игрока

Чего хотят игроки? 1 платить меньше, 2 получить больше

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

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

Зависит ли игра от того, кто первый?

1) 1 игрок выбирает первым:

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

2) 2 игрок выбирает первым:

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

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

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

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

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

<u>Теорема</u> (Сиона-Какутани)

X, Λ — выпуклые компакты,

$L(x,\lambda)$ — выпукло-вогнутая на $X\times\Lambda$

тогда L имеет седловые точки на $X\times\Lambda$

<u>Теорема</u>

Мн-во седл. точек L непусто $\Leftrightarrow$ $\sup\limits_{\lambda}\inf\limits_{x} L(x,\lambda)$ и $\inf\limits_{x}\sup\limits_{\lambda} L(x,\lambda)$ имеют решения и они совпадают.

поиск седла (если существует)

$$\Leftrightarrow$$

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

$$\Leftrightarrow$$

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

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

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

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

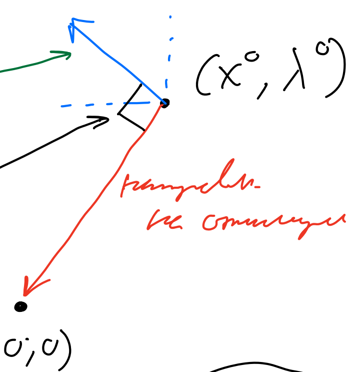
старт.: (x^0, λ^0)

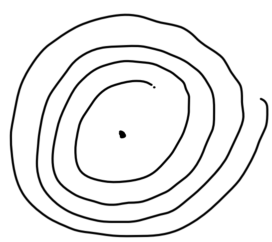
направление: $\gamma\begin{pmatrix} -\lambda^0 \\ x^0 \end{pmatrix}$

направл. вдоль.

90°

$$\langle \gamma\begin{pmatrix} -\lambda^0 \\ x^0 \end{pmatrix}; \begin{pmatrix} x^0 \\ \lambda^0 \end{pmatrix}\rangle = 0$$





- 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 / прокс оператор

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

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

EG – имитация неявной схемы

- Неточный шаг вперед в точке градиента

Угол в примере всегда будет острый

⊕ такой же простой, как и GD

⊕ лучше гарантии, чем у GD

⊕ оптимален для седл. задач

⊖ два раза вычисление градиента

Решение минуса:

- 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