Задачи с ограничениями

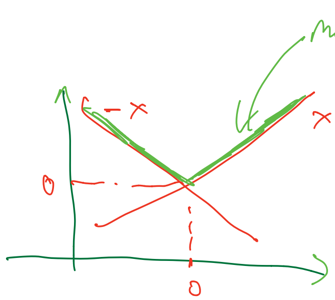
$$f(x) \to \min$$
$$Ax = b \quad \lambda$$
$$\vec{g}(x) \le 0 \quad \mu$$
$$x \in Q$$

задача выпуклой оптимизации

выпуклые

$x \in Q$ ← мн-во простой структуры

$$\begin{pmatrix} g_1(x) \\ g_2(x) \\ \vdots \\ g_m(x) \end{pmatrix} \le 0 \Leftrightarrow \begin{cases} g_1(x) \le 0 \\ \dots \\ g_m(x) \le 0 \end{cases}$$



$$\underbrace{g(x) = \max_{k=\overline{1,m}} g_k(x)}_{\text{негладкая ф-ция}} \le 0$$

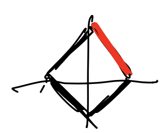
$\{x : g(x) \le 0\} = \hat{Q} \to$ не мн-во простой структуры

$g(x) = \|x\|_2^2 \to$ простая ф-ция (потому что просто проекц. на шар)

Q — шары в разных нормах или их подмножества (простые)

Пример. $S_n(1) \subseteq B_1^n(1)$

$$= \{x \in \mathbb{R}^n_+ : \sum_{i=1}^n x_i = 1\}$$



Принцип множителей Лагранжа

$$f(x) + \underbrace{\max_{\lambda} \langle \lambda, Ax - b \rangle}_{\begin{cases} 0, & Ax = b \\ +\infty, & Ax \ne b \end{cases}} + \underbrace{\max_{\mu \ge 0} \langle \mu, g(x) \rangle}_{\begin{cases} 0, & g(x) \le 0 \\ +\infty, & g(x) > 0 \end{cases}} \to \min_{x \in Q}$$

$$\min_{x\in Q}\left\{\max_{\lambda}\ \max_{\mu\ge 0}\left\{\underbrace{f(x)+\langle\lambda,Ax-b\rangle+\langle\mu,g(x)\rangle}_{L(x;\lambda,\mu)}\right\}\right\}$$

Т. Сиона-Какутани
(фон Неймана)
min max = max min

$\vee$ – вып. по x (min-пер.)

$\cap$ – вып. по μ,λ (max-пер.)

$x\in Q$ (Q – компакт)

$$\max_{\lambda;\,\mu\ge 0}\ \min_{x\in Q} L(x;\lambda,\mu)$$

$x(\lambda,\mu)$ – реш. задачи

$$\varphi(\lambda,\mu) = L(x(\lambda,\mu);\lambda,\mu)$$ – двойственная ф-ция

Двойственная задача

$$\varphi(\lambda,\mu)\to\max_{\lambda,\,\mu\ge 0} \qquad (D)$$

Алгоритм. Решаем (D)

находим λ^N,μ^N

$\downarrow\ x(\lambda^N,\mu^N)$

находим $x^N = x(\lambda^N,\mu^N)$?

$x_* = x(\lambda_*,\mu_*)$ – верно

Проблема $\lambda^N,\mu^N \xrightarrow[N\to\infty]{} (\lambda_*,\mu_*)$ – Не обязательно!!!

$$\varphi(\lambda,\mu) = \min_{x\in Q}\{f(x)+\langle\lambda,Ax-b\rangle+\langle\mu,g(x)\rangle\} \to x(\lambda,\mu)$$ реш. задачи

$$\nabla_\lambda\varphi(\lambda,\mu) = Ax(\lambda,\mu)-b\ ;\quad \nabla_\mu\varphi(\lambda,\mu) = g(x(\lambda,\mu))$$

ф-ла Демьянова-Данскина

Пояснение

Если $\varphi(\lambda) = \min\limits_{x} \mathcal{L}(x,\lambda)$, то $\nabla\varphi(\lambda) = \mathcal{L}_\lambda(x(\lambda),\lambda)$

○

$$\frac{\partial \mathcal{L}(x,\lambda)}{\partial x}\Big|_{x=x(\lambda)} \equiv 0 \quad (*)$$

$$L_x(x(\lambda),\lambda) \equiv 0$$

$$\nabla\varphi(\lambda) = \nabla\mathcal{L}(x(\lambda),\lambda) = \underbrace{\frac{\partial x}{\partial\lambda}\nabla_x \mathcal{L}(x,\lambda)\Big|_{x=x(\lambda)}}_{0} + \nabla_\lambda \mathcal{L}(x,\lambda)\Big|_{x=x(\lambda)} =$$

из $(*) \Rightarrow L_x(x(\lambda),\lambda) = 0$

$$= \mathcal{L}_\lambda(x(\lambda),\lambda)$$

●

Двойственная задача

$$\varphi(\lambda,\mu) \to \max_{\lambda;\,\mu\geq 0}$$

$$\downarrow$$

$$\hat\varphi = -\varphi \mapsto \hat\varphi(\lambda,\mu) \to \min_{\lambda;\,\mu\geq 0}$$

$$\hat\varphi(\mu) \to \min_{\mu\geq 0}$$

x_* – реш. исходной задачи $\Rightarrow g(x_*) \leq 0$

$$\hat\varphi(\mu) = -\min_{x\in Q}\{f(x) + \langle\mu, g(x)\rangle\} = \max_{x\in Q}\{-f(x) - \langle\mu, g(x)\rangle\}$$

$\mu \geq 0$

$$-f(x(\mu)) - \langle\mu, g(x(\mu))\rangle \underset{\text{из опр. } x(\mu)}{\geq} -f(x_*) - \langle\underset{\geq 0}{\mu}, \underset{\leq 0}{g(x_*)}\rangle \geq -f(x_*)$$

$$f(x(\mu)) - f(x_*) \leq -\langle\mu, \underbrace{g(x(\mu))}_{= -\nabla\hat\varphi(\mu)}\rangle = \langle\mu, \nabla\hat\varphi(\mu)\rangle$$

$f(x(\mu)) - f(x_*) \leq \langle\mu, \nabla\hat\varphi(\mu)\rangle$ → отвечает за качество восст. x по μ

$$f(x(\lambda,\mu)) - f(x_*) \leq \langle\lambda, \nabla\hat\varphi(\lambda)\rangle + \langle\mu, \nabla\hat\varphi(\mu)\rangle$$

$$\langle \mu^N, \nabla \tilde{\varphi}(\mu^N)\rangle \xrightarrow[N\to\infty]{} 0 \Rightarrow f(x(\mu^N)) - f(x_*) \xrightarrow[N\to\infty]{} 0$$

Это все верно, если μ^N - огранич. (для чего нужно μ_* - огранич)

$$\tilde{\varphi}(0) \geq \tilde{\varphi}(\mu_*) = \max_{x\in Q}\left\{-f(x) - \sum_{i=1}^m \mu_*^i g_i(x)\right\} \geq$$

$\exists \bar{x}\in Q:$ $\max\limits_{i=\overline{1,m}} g_i(\bar{x}) \leq -\gamma$ условие Слейтера

$$\geq -f(\bar{x}) - \sum_{i=1}^m \mu_*^i g_i(\bar{x}) \geq$$

$$\geq -f(\bar{x}) + \gamma\|\mu_*\|_1$$

$$\|\mu_*\|_1 \leq \frac{\tilde{\varphi}(0) + f(\bar{x})}{\gamma} \leq \frac{f(\bar{x}) - \min_{x\in Q} f(x)}{\gamma}$$

сепарабельный ф-л

<u>Пример</u>. $f(x) = \sum\limits_{i=1}^n f_i(x_i) \to \min\limits_{\substack{Ax=b\\ x\in\mathbb{R}^n}}$

$$\mathcal{L}(x;\lambda) := \sum_{i=1}^n f_i(x_i) + \langle\lambda, Ax-b\rangle \to \min_{x\in\mathbb{R}^n}$$

$$\sum_{i=1}^n \left\{f_i(x_i) + [A^T\lambda]_i x_i - b_i\right\} \to \min_{x\in\mathbb{R}^n}$$

$$\min_{x_i\in\mathbb{R}} f_i(x_i) + [A^T\lambda]_i x_i - b_i$$

$x_i(\lambda)$ сложность $O(\ln\frac{1}{\varepsilon})$

$O(n\ln\frac{1}{\varepsilon}) \to x(\lambda)$

без учета умнож. $\underbrace{A^T \text{ на } \lambda}_{nnz(A)}$ и вычисл $f_i'(x_i)$ $\Rightarrow O(nnz(A) + n\ln\frac{1}{\varepsilon})$

$x(\lambda)$ – считается не точно $\Longrightarrow$ $\nabla\varphi(\lambda)$ – считается неточно

$\tilde{\varepsilon} = \varepsilon / N_{\text{двойств}}(\varepsilon)$, $\begin{cases} f(x^N) - f(x_*) \le \varepsilon \\ \underbrace{\|Ax^N - b\|_2}_{\nabla\varphi(\lambda)} \le \varepsilon / R \end{cases}$

$N_{\text{двойств}}$ ↘ число итераций метода для $\tilde{\varphi}(\lambda) \to \min$

$R = \|\lambda_*\|_2$

$x^N \to x(\lambda^N)$

$f(x(\lambda^N)) - f(x_*) \le \underbrace{\langle \lambda^N, \nabla\tilde{\varphi}(\lambda^N)\rangle}_{\varepsilon}$

$\|Ax(\lambda^N) - b\|_2 \le \|\nabla\tilde{\varphi}(\lambda^N)\|_2 \sim \varepsilon / R$

$\|\lambda^N\|_2 \sim \|\lambda_*\|_2 \sim R$

$\lambda^0 = 0$

$$\tilde{\varphi}(\lambda^N) - \tilde{\varphi}_* \le \frac{4L_\varphi \|\lambda^N\|_2^2}{N^2}$$

f_i – μ-сильно вып. в 2-нор. $L_\varphi \sim \frac{1}{\mu}$

$\rotatebox{90}{$\le$}$

$\frac{1}{2L_\varphi}\|\nabla\tilde{\varphi}(\lambda^N)\|_2^2$ $\Longrightarrow$ $\|\nabla\tilde{\varphi}(\lambda^N)\|_2 \sim \frac{1}{N}$ → плохо!!!

(D) $\tilde{\varphi}(\lambda) \to \min\limits_\lambda$

(D_ε) $\underbrace{\tilde{\varphi}(\lambda) + \frac{\varepsilon}{R^2}\|\lambda\|_2^2}_{\tilde{\varphi}_\varepsilon} \to \min\limits_\lambda$ $\longrightarrow$ $\varepsilon/2$-реш. (D_ε) будет ε-реш. (D)

Далее, см. упр. 4.4. пособия МОиМО

для ускор. метода

$$\frac{1}{2L_{\tilde{\varphi}_\varepsilon}}\|\nabla\tilde{\varphi}_\varepsilon(\lambda^N)\|_2^2 \le \tilde{\varphi}_\varepsilon(\lambda^N) - \tilde{\varphi}_{\varepsilon,*} \lesssim L_{\tilde{\varphi}_\varepsilon} R^2 \exp\left(-\sqrt{\frac{\varepsilon}{L_{\tilde{\varphi}_\varepsilon} R^2}}\, N\right)$$

$L_{\tilde{\varphi}_\varepsilon} \gtrsim L_\varphi$

$$\|\nabla\tilde{\varphi}_\varepsilon(\lambda^N)\|_2 \le L_{\tilde{\varphi}_\varepsilon} R \exp\left(-\frac{1}{2}\sqrt{\frac{\varepsilon}{L_{\tilde{\varphi}_\varepsilon} R^2}}\, N\right)$$

Общий факт: с точн. до мн.-множ. $\Delta f^N \simeq \Delta \varphi^N$

ЭЛП: $\quad f_i(x_i) = x_i \ln x_i$

$\exists$ явные ф-лы

Метод штрафных ф-ций

$$f(x) \to \min_{\substack{Ax=b \\ x\in Q}} \quad \Big|\ \lambda_*$$

вып. задача

$\|\lambda_*\|_2 \le R_\lambda$

$$f(x^N) - f(x_*) \le \varepsilon, \qquad \frac{R_\lambda^2}{\varepsilon}\|Ax-b\|_2^2 \le \varepsilon$$

$$f(x) \to \min_{x\in Q} \qquad \frac{R_\lambda^2}{\varepsilon}\|Ax-b\|_2^2 \to \min_{x\in Q}$$

$$F_\varepsilon(x) = f(x) + \frac{R_\lambda^2}{\varepsilon}\|Ax-b\|_2^2 \to \min_{x\in Q} \qquad (*)$$

$$f(x) \to \min_{\substack{\|Ax-b\|_2^2 \le \varepsilon^2 \\ x\in Q}} \quad \Big|\ \tilde{\lambda}$$

$$f(x) + \underbrace{\tilde{\lambda}}_{\frac{R_\lambda^2}{\varepsilon}}\left(\|Ax-b\|_2^2 - \cancel{\varepsilon^2}\right) \to \min_{x\in Q}$$

$$F_\varepsilon(x^N) - \min_{x\in Q} F_\varepsilon(x) \le \varepsilon$$

$$\Downarrow$$

$$f(x^N) - \min_{\substack{Ax=b \\ x\in Q}} f(x) \le \varepsilon$$

$$\|Ax^N - b\|_2 \le 2\varepsilon / R_\lambda$$

Не обяз.
Вып. оптим.

$$f(x) \to \min$$
$$h(x) = 0 \quad | \ \lambda$$

интеграл. ф-ции

$$\longrightarrow \quad f(x) + K\|h(x)\|_2^2 \to \min \quad \longrightarrow \quad x(K)$$

$$K\,h(x(K)) \xrightarrow[K\to\infty]{} \lambda$$

h не обяз. скаляр!

$$h(x(K)) \simeq \frac{\lambda}{K}$$

$$f(x^K) - f(x_*) = O\left(\frac{\|\lambda\|_2^2}{K}\right)$$