27. 11. 17 МОМО XII

Ускоренний оптимальний метод Нестерова

$f(x) \to \min\limits_{x \in \mathbb{R}^n}$, $f \in C^1$

FOM: $x_k = x_0 - \sum\limits_{i=0}^{k-1} a_i \nabla f(x_i)$

Утв $\forall\ 0 \le k \le \frac{n-1}{2}$ $\exists f \in C_L^{1,1}$ и вып. (μ сильно вып.):

$\forall$ FOM $f(x_k) - f(x_{opt}) \geq \dfrac{3L \| x_0 - x_{opt}\|^2}{32(k+1)^2}$

$\left(f(x_k) - f(x_{opt}) \geq \dfrac{\mu}{2} \left(\dfrac{\sqrt{\mu/L} - 1}{\sqrt{\mu/L} + 1} \right)^{2k} \| x_0 - x_{opt} \|^2 \right)$

Идея док-ва:

$f(x) = const \left(\frac{1}{2} \langle Ax, x \rangle - \langle x, e_1 \rangle \right)$

$e_1 = (1, 0 \ldots 0)^T$, $A = \begin{bmatrix} \ddots & & 0 \\ & \ddots & \\ 0 & & \ddots \end{bmatrix}$

$\nabla f(x) = const (Ax - e_1)$, $x_0 = 0$

Ф-ия	Град. спуск	Ускор. град. спуск
$\in C_L^{1,1}$ и μ сильно вып.	$O\left(\left(1 - \frac{\mu}{L}\right)^k\right)$	$O\left(\left(1 - \sqrt{\frac{\mu}{L}}\right)^k\right)$
$\in C^{1,1}$ и вып.	$O(1/k)$	$O(1/k^2)$

$F(x) = f(x) + h(x) \to \min\limits_x$

$f \in C_L^{1,1}$ и вып., $h \in C, \notin C^1$, вып. простая

$F(x) \le m_k(x) = f(x_k) + \nabla f(x_k)^T (x - x_k) + \frac{L}{2} \| x - x_k \|^2 + h(x) \to \min\limits_x$

$\operatorname{prox}_h(x) = \operatorname*{argmin}_u \left(\frac{1}{2}\|u-x\|^2 + h(u)\right)$

Prox - Grad : $x_{k+1} = \operatorname*{argmin}_x m_k(x) =$

градиентное $\quad = \operatorname{prox}_{\frac{1}{L}h}\left(x_k - \frac{1}{L}\nabla f(x_k)\right) =$

отображение: $\quad = x_k - \frac{1}{L} G_{1/L}(x_k)$

$G_{1/L}(x_k) = L(x_k - x_{k+1}) = L\left(x_k - \operatorname{prox}_{\frac{1}{L}h}\left(x_k - \frac{1}{L}\nabla f(x_k)\right)\right)$

критерий останова: $\|G_{1/L}(x_k)\|^2 \le \varepsilon$

оценочная посл-ть: $\psi_k(x) = \frac{1}{2}\|x - x_0\|^2 +$

$$+ \sum_{i=1}^{k} a_i \left(f(x_i) + \langle \nabla f(x_i), x - x_i\rangle + h(x)\right)$$

$A_k = \sum_{i=1}^{k} a_i$, $A_0 = 0$

$\psi_k^* = \min_x \psi_k(x)$, $v_k = \operatorname*{argmin}_x \psi_k(x) = \operatorname{prox}_{A_k h}\left(x_0 - \sum_{i=1}^{k} a_i \nabla f(x_i)\right)$

$\{x_k\}_{k=0}^{\infty}$, $\{v_k\}_{k=0}^{\infty}$, $\{y_k\}_{k=0}^{\infty}$

$y_k = \dfrac{A_k x_k + a_{k+1} v_k}{A_k + a_{k+1}}$, $x_{k+1} = \operatorname{prox}_{1/L\, h}\left(y_k - \frac{1}{L}\nabla f(y_k)\right)$

$\exists\, (x_k, \psi_k(x))$:

① $A_k F(x_k) \le \psi_k^*$

② $\psi_k(x) \le \frac{1}{2}\|x - x_0\|^2 + A_k F(x) \quad \forall x$, тогда:

$$F(x_k) - F(x_{opt}) \overset{①}{\le} \frac{\psi_k^*}{A_k} - F(x_{opt}) \le \frac{\psi_k(x_{opt})}{A_k} - F(x_{opt}) \overset{②}{\le}$$

$$\overset{②}{\le} \frac{1}{2A_k}\|x_{opt} - x_0\|^2 + \cancel{F(x_{opt})} - \cancel{F(x_{opt})} = \frac{1}{2A_k}\|x_0 - x_{opt}\|^2$$

②

$$\psi_k(x) = \frac{1}{2}\|x - x_0\|^2 + \sum_{i=1}^{k} a_i \big(\underbrace{f(x_i) + \langle \nabla f(x_i), x - x_i\rangle}_{\le f(x)} + h(x)\big) \le$$

$$\le \frac{1}{2}\|x - x_0\|^2 + A_k F(x) \quad \forall a_i$$

① по индукции

$k = 0$: $\psi_0^* = 0$; $A_0 = 0$

Хотим: $\psi_{k+1}^* \geq A_{k+1} F(x_{k+1})$

$$\psi_{k+1}^* = \min_x \psi_{k+1}(x) = \min_x \big(\psi_k(x) + a_{k+1}\big(f(x_{k+1}) +$$

$$+ \langle \nabla f(x_{k+1}), x - x_{k+1}\rangle + h(x)\big) \geq$$

// $\psi_k(x) \geq \psi_k^* + \frac{1}{2}\|x - v_k\|^2$; $\psi_k^* \geq A_k F(x_k)$
от сильной выпуклости, $\mu \geq 1$; индукционный переход

$$\geq \min_x \big(A_k F(x_k) + \frac{1}{2}\|x - v_k\|^2 + a_{k+1}\big(f(x_{k+1}) +$$

$$+ \langle \nabla f(x_{k+1}), x - x_{k+1}\rangle + h(x)\big)\big) \geq$$

$$h(x_{k+1}) + \langle \partial h(x_{k+1}), x - x_{k+1}\rangle$$

// $F(x_k) \geq F(x_{k+1}) + \langle \partial F(x_{k+1}), x_k - x_{k+1}\rangle$

$$\geq \min_x \big(A_{k+1} F(x_{k+1}) + \langle \partial F(x_{k+1}), A_k x_k - A_k x_{k+1}\rangle +$$

$$+ \frac{1}{2}\|x - v_k\|^2 + a_{k+1}\langle \partial F(x_{k+1}), x - x_{k+1}\rangle\big) =$$

// $A_k x_k = A_{k+1} y_k - a_{k+1} v_k$

$$= \min_x \big(A_{k+1} F(x_{k+1}) + A_{k+1}\langle \partial F(x_{k+1}), y_k - x_{k+1}\rangle +$$

$$+ a_{k+1}\langle \partial F(x_{k+1}), x - v_k\rangle + \frac{1}{2}\|x - v_k\|^2\big) =$$

$$= \left\{ \operatorname*{argmin}_x = v_k - a_{k+1}\partial F(x_{k+1}) \right\} = A_{k+1} F(x_{k+1}) +$$

$$+ A_{k+1}\langle \partial F(x_{k+1}), y_k - x_{k+1}\rangle - \frac{a_{k+1}^2}{2}\|\partial F(x_{k+1})\|^2 =$$

$$= A_{k+1}\Big(F(x_{k+1}) + \langle \partial F(x_{k+1}), y_k - x_{k+1}\rangle -$$

$$- \frac{a_{k+1}^2}{2A_{k+1}}\|\partial F(x_{k+1})\|^2\Big) \geq A_{k+1} F(x_{k+1})$$

$$\Updownarrow$$

$$\langle \partial F(x_{k+1}), y_k - x_{k+1}\rangle \geq \frac{a_{k+1}^2}{2A_{k+1}}\|\partial F(x_{k+1})\|^2$$

Утв. $f \in C_L^{1,1}$, $x = prox_{\frac{1}{L}h}(y - \frac{1}{L}\nabla f(y)) \Rightarrow$

$\exists\ F'(x) \in \partial F(x) : \langle F'(x), y - x\rangle \geq \frac{1}{2L}\|F'(x)\|^2$

$\square$ $G_{1/L}(y) = L(y-x)$; $G_{1/L}(y) - \nabla f(y) \in \partial h(x)$

$\|\nabla f(y) - \nabla f(x)\|^2 \leq L^2\|y-x\|^2$

$\partial F(x) \ni F'(x) = G_{1/L}(y) - \nabla f(y) + \nabla f(x)$

$\|L(y-x) - F'(x)\|^2 \leq L^2\|y-x\|^2$

$L^2\|y-x\|^2 - 2L\langle y-x, F'(x)\rangle + \|F'(x)\|^2 \leq L^2\|y-x\|^2$

$\langle F'(x), y-x\rangle \geq \frac{1}{2L}\|F'(x)\|^2$ ■

$\frac{a_{k+1}^2}{2A_{k+1}} = \frac{1}{2L} \Leftrightarrow \frac{a_{k+1}^2}{A_k + a_{k+1}} = \frac{1}{L}$, отсюда выражаем a_{k+1}

Схема проксимального ускоренного метода Нестерова

$x_0 = 0$, $A_0 = 0$, L, $v_0 = x_0$

для $k = 0, 1, 2, \dots$

Найти a : $\frac{a^2}{A_k + a} = \frac{1}{L}$ $\underline{(*)}$

$y = \frac{A_k x_k + a \cdot v_k}{A_k + a}$; $A_{k+1} = A_k + a$

$x_{k+1} = prox_{\frac{1}{L}h}(y_k - \frac{1}{L}\nabla f(y_k))$

если $f(x_{k+1}) > f(y_k) + \langle\nabla f(y_k), x_{k+1} - y_k\rangle +$

$+ \frac{L}{2}\|x_{k+1} - y_k\|^2$, то $L \leftarrow L \cdot 2$ то $(*)$

если $\|L(y_k - x_{k+1})\|^2 \leq \varepsilon$, то выход

$v_{k+1} = prox_{A_{k+1}h}(x_0 - \sum_{i=1}^{k+1} a_i \nabla f(x_i))$

$L \leftarrow L/2$

$$F(x_k) - F(x_{opt}) \le \frac{1}{2A_k} \| x_0 - x_{opt} \|^2 \le$$

$$\frac{a_{k+1}^2}{a_{k+1} + A_k} = \frac{1}{L}, \quad \frac{(A_{k+1} - A_k)^2}{A_{k+1}} = \frac{1}{L},$$

$$A_{k+1} = L(A_{k+1} - A_k)^2 = L(\sqrt{A_{k+1}} - \sqrt{A_k})^2(\sqrt{A_{k+1}} + \sqrt{A_k})^2 \le$$

$$\le \{ A_{k+1} \ge A_k \} \le L(\sqrt{A_{k+1}} - \sqrt{A_k})^2 4 A_{k+1}$$

$$\sqrt{A_{k+1}} \ge \sqrt{A_k} + \frac{1}{2\sqrt{L}} \ge \sqrt{A_{k-1}} + \frac{2}{2\sqrt{L}} \ge \ldots \ge \underset{\overset{\|}{0}}{\sqrt{A_0}} + \frac{k+1}{2\sqrt{L}} = \frac{k+1}{2\sqrt{L}}$$

$$A_{k+1} \ge \frac{(k+1)^2}{4L}$$

$\le \frac{2L}{(k~~+1~~)^2} \| x_0 - x_{opt} \|^2$ сходимость оптимальная с точностью для ~~б~~ ближней константы

$F(\bar{x}_k) = \min\limits_{0 \le i \le k} F(x_i)$ немонотонное уменьшение функционала

семинар

(2) $\partial \delta_{\mathbb{R}^n_+} = ?$

стандартные классы выпуклых задач

коническое и полуопределённое программирование

$\boxed{1}$ линейное программирование (LP)

$\min\limits_{x \in \mathbb{R}^n} \underline{\langle c, x \rangle}$ $\leftrightarrow$ $\langle a_i, x \rangle \le b_i$, $i = 1..m$

s. t. $\boxed{Ax \le b}$, $Gx = h$ $\leftrightarrow$

$\boxed{2}$ кв. прогр. (QP)

$\min\limits_{x \in \mathbb{R}^n} \frac{1}{2} \langle Qx, x \rangle + \langle p, x \rangle$ $\quad Q \in S^n_+$

s. t. $Ax \le b$, $Gx = h$

③ Кв. прогр. с кв. огр. (QCQP)

$$\min \frac{1}{2} \langle Q_0 x, x \rangle + \langle p_0, x \rangle$$

$$\text{s.t.} \quad \frac{1}{2} \langle Q_i x, x \rangle + \langle p_i, x \rangle \le \hat{c}_i \quad \forall i = 1 \dots m$$

$$Gx = h \quad , \quad Q_0, Q_1 \dots Q_m \in S^n_+$$

④ Конические прогр. 2-го порядка (SOCP)

second order conic programming

$$\min_{x \in \mathbb{R}^n} \langle c, x \rangle$$

$$\text{s.t.} \quad \| A_i x + b_i \|_2 \le \langle c_i, x \rangle + d_i$$

$$i = 1 \dots m$$

$$Gx = h$$

$x \in K \Rightarrow t \cdot x \in K \quad \forall t \ge 0$, K - конус

Конус Лоренца

$$(A_i x + b_i, \langle c_i, x \rangle + d_i) \in K_2 = \{ (y, t) \in \mathbb{R}^m \times \mathbb{R} : \| y \|_2 \le t \}$$

коническое программирование

$$\min_{x \in \mathbb{R}^n} \langle c, x \rangle$$

s.t. $Gx = h$, $x \in K$, конус

хорошие конусы:

① $K = \mathbb{R}^n_+$, LP

② $K = K_2$, Лоренца

③ $K = S^n_+$

⑤ Полуопределённое программирование (SDP)

$$\min_{x \in \mathbb{R}^n} \langle c, x \rangle$$

матричное $\le$

$$\text{s.t.} \quad x_1 A_1^{(i)} + \dots + x_n A_n^{(i)} \le A_0$$

$$i = 1 \dots m \quad , \quad Gx = h \quad , \quad A_j^{(i)} \in S^n$$

$$\min_{X \in S_n^+} \langle C, X \rangle$$

эквивалентно

s.t. $GX = H$, $\cancel{X \in}$

$$LP \subset QP \subset QCQP \subset SOCP \subset SDP$$

Примеры:

① $\min\limits_{x \in \mathbb{R}^n} \|Ax - b\|_2 \iff \min\limits_{x \in \mathbb{R}^n} \|Ax - b\|_2^2$

$x \geq 0$ $\qquad Q = A^T A \qquad$ (QP)

② $\min\limits_{x \in \mathbb{R}^n} \|Ax - b\|_1 \quad , \; x \geq 0$

$\|Ax - b\|_1 = \langle 1_m, |Ax - b| \rangle = \langle 1_m, t \rangle$

$t_i \geq (Ax - b)_i \;, \quad t_i \leq -(Ax - b)_i$

$\min\limits_{t \in \mathbb{R}^m} \langle 1_m, t \rangle \qquad \| \|_2 , \| \|_\infty \Rightarrow$ (LP)

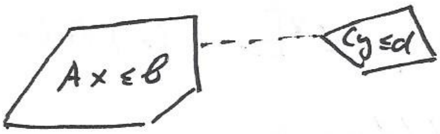
$\left[\begin{array}{l} t_i \geq (Ax - b)_i \\ t_i \leq -(Ax - b)_i \end{array}\right.$ $\quad \| \|_2 : t_i \geq |(Ax - b)_i|$ $\quad$ epi f $\quad t \in \mathbb{R}^m_+, \min \langle 1, t \rangle$

$\| \|_\infty : t_i \geq |(Ax - b)_i| \; \forall i \quad t \in \mathbb{R}_+ , \min t$

③ $\min\limits_{x, y \in \mathbb{R}^n} \|x - y\|_2 \qquad \min\limits_{z \in \mathbb{R}^n} Tz$

s.t. $Ax \leq b , \; Cy \leq d \sim \qquad T = [I, -I]$

$z = (x, y)$



④

$\frac{1}{2} \langle Q_i x, x \rangle + \langle p_i, x \rangle \leq r_i \;, \; \exists Q_i \in \cancel{S^n_+} \; S^n$

$x_i \in \{-1, 1\} \iff x_i^2 = 1 \iff \begin{cases} x_i^2 \leq 1 & \text{вып.} \\ -x_i^2 \leq -1 & \text{вогн.} \end{cases}$ $\quad$ NP

⑤

$\min\limits_t p(t)$

$\min\limits_t at^3 + bt^2 + ct + d$ $\qquad x_3 \geq x_1 \cdot x_2$

$x_1 = t, \; x_2 = t, \; x_3 = x_1 \cdot x_2 \qquad -x_3 \geq -x_1 \cdot x_2$

$x_1 \cdot x_2 \qquad t^3$ $\qquad$ невыпуклое ограничение

минимизация произвольных многочленов NP

⑥ $\min\limits_{x \in \mathbb{R}^n} \underbrace{\|Ax - b\|_2}_{\ell_2} + \underbrace{\|x\|_2}_{\ell_2}$ конические програм.

$$\min_{x, t_1, t_2} \; t_1 + t_2$$

$$\text{s.t.} \quad \|Ax - b\|_2 \le t_1, \quad \|x\|_2 \le t_2$$

неверно будет ставить кв. $\|Ax - b\|_2^2 \le t_1^2, \; \|x\|_2^2 \le t_2^2$

$$x^2 - t^2 \le 0 \qquad \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \notin S^n_+$$

⑦ $\min\limits_{x \in \mathbb{R}^n} \|A_0 - x_1 A_1 - \ldots - x_n A_n\|_F$, ор. кв.п., полуопр.

$$A_0, A_1, \ldots, A_n \in S^n$$

$$\min_{x \in \mathbb{R}^n} \lambda_{max} (\underbrace{A_0 - x_1 A_1 - \ldots - x_n A_n}_{A(x)})$$

$$\min_{x, t} t, \quad \text{s.t.} \quad \lambda_{max}(A(x)) \ge t \iff A(x) \ge tI$$

$$\min_{x, t} t, \quad \text{s.t.} \quad t \ge \sigma_{max}(A(x))$$

$$t^2 \ge \lambda_{max}(A(x)A(x)^T), \quad t^2 I \ge A(x)A(x)^T$$

Лемма о дополнении Шура

$$M = \begin{bmatrix} A & B \\ B^T & C \end{bmatrix} \qquad A \in S^n, \; C \in S^m, \; B \in \mathbb{R}^{n \times m}$$

Пусть $A > 0$. Тогда $M \ge 0 \iff \underbrace{C - B^T A^{-1} B}_{\text{доп-е Шура}} \ge 0$

□ $z = (x, y) \in \mathbb{R}^n \times \mathbb{R}^m$

$$0 \le \langle Mz, z \rangle = \langle Ax, x \rangle + 2\langle x, By \rangle + \langle Cy, y \rangle \to \min_x$$

$$Ax^* + By = 0, \quad x^* = -A^{-1}By, \quad \langle (C - B^T A^{-1} B) y, y \rangle \ge 0 \; ■$$

тогда переход от нелинейных нер-в к линейным

$$\begin{bmatrix} tI & A(x)^T \\ A(x) & tI \end{bmatrix} \ge 0 \iff tI - \frac{A(x)A(x)^T}{t} \ge 0$$

$$t^2 I \ge A(x)A(x)^T$$

① LP

$Ax \le b \iff \mathrm{Diag}\{b - Ax\} \succeq 0$ (SDP)

② QP

$\frac{1}{2} \langle Qx, x \rangle + \langle p, x \rangle \le z$

$\frac{1}{2} x^T Q x \le z - p^T x$

$Q = L^T L \Rightarrow \frac{1}{2} x^T L^T L x \le z - p^T x$

лемма Шура: $\Updownarrow$

$$\begin{bmatrix} 2I & Lx \\ (Lx)^T & (z - \langle p, x \rangle) I \end{bmatrix} \succeq 0 \qquad \begin{bmatrix} tI & x \\ x^T & t \end{bmatrix} \succeq 0$$

$\Updownarrow$

③ SOCP

$\|x\|_2 \le t \iff \|x\|_2^2 \le t^2 \iff t^2 - x^T x \ge 0$