Опр. (проксимальный оператор)

$r: \mathbb{R}^d \to \mathbb{R} \cup \{+\infty\}$. Проксимальный оператор, порожденный функц. r есть

$$\mathrm{prox}_r(x) = \operatorname*{argmin}_{\tilde{x} \in \mathbb{R}^d} \left\{ r(\tilde{x}) + \frac{1}{2}\|x - \tilde{x}\|_2^2 \right\}$$

Сущ. решения:

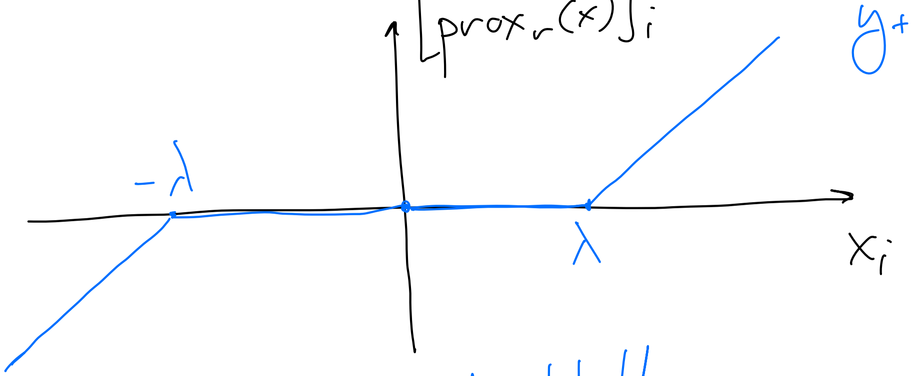
- $\exists\, \hat{x} \in \mathbb{R}^d \hookrightarrow r(\hat{x}) < +\infty$ (конечна), тогда если r дополнительно выпукла, то такой оператор однозначн. определен.

Пример

- $r(x) = \lambda \|x\|_1$, где $\lambda > 0$

$$[\mathrm{prox}_r(x)]_i = \left[|x_i| - \lambda\right]_+ \mathrm{sign}(x_i)$$

$$y_+ = \begin{cases} y, & y \geq 0 \\ 0, & y < 0 \end{cases}$$



trash hold
(feature selection)

• $r(x) = \frac{\lambda}{2}\|x\|_2^2$, где $\lambda > 0$

$$\text{prox}_r(x) = \frac{x}{1+\lambda}$$

• $r(x) = \mathbb{I}_{\mathcal{X}}(x)$, где $\mathcal{X}$ – выпуклое замкн. мн-во

$$\mathbb{I}_{\mathcal{X}}(x) = \begin{cases} 0 & x \in \mathcal{X} \\ +\infty & x \notin \mathcal{X} \end{cases}$$

$$\text{prox}_r(x) = \arg\min_{\tilde{x} \in \mathbb{R}^d} \left\{ \mathbb{I}_{\mathcal{X}}(\tilde{x}) + \frac{1}{2}\|x - \tilde{x}\|_2^2 \right\}$$

$$= \begin{cases} \arg\min\limits_{\tilde{x} \in \mathbb{R}^d} \left\{ \frac{1}{2}\|x - \tilde{x}\|_2^2 \right\} & \tilde{x} \in \mathcal{X} \\ \text{нет решений} & \end{cases}$$

$$= \text{proj}_{\mathcal{X}}(x)$$

Свойство 1 (техническое)

$r: \mathbb{R}^d \to \mathbb{R} \cup \{+\infty\}$ выпуклая функция,

prox_r определен

Тогда $\forall\, x, y \in \mathbb{R}^d$ следующие св-ва эквивалентны

1) $\text{prox}_r(x) = y$
2) $x - y \in \partial r(y)$
3) $\langle x - y, z - y \rangle \le r(z) - r(y) \quad \forall z \in \mathbb{R}^d$

Док-во:

1) $y = \arg\min\limits_{\tilde{x} \in \mathbb{R}^d} \left\{ r(\tilde{x}) + \frac{1}{2}\|x - \tilde{x}\|_2^2 \right\}$

Условие оптимальности

$$0 \in \partial\left(r(\tilde{x}) + \tfrac{1}{2}\|x-\tilde{x}\|_2^2\right)\Big|_{\tilde{x}=y}$$

$$= \partial r(y) + y - x$$

$$y - x \in \partial r(y)$$ доказано $1 \Leftrightarrow 2$

Определение субград. для выпуклых функций

$$\langle \underset{\in \partial r(y)}{g}; z-y\rangle \le r(z)-r(y) \quad \forall z \in \mathbb{R}^d$$

$y - x \in \partial r(y)$

$$\langle y-x; z-y\rangle \le r(z)-r(y)$$ доказано $2 \Leftrightarrow 3$

Свойство 2:

$r: \mathbb{R}^d \to \mathbb{R} \cup \{+\infty\}$ выпуклая и prox_r определен, тогда $\forall x, y \in \mathbb{R}^d \rightarrow \|\mathrm{prox}_r(x) - \mathrm{prox}_r(y)\|_2 \le \|x-y\|_2$

Док-во: $u = \mathrm{prox}_r(x)$, $v = \mathrm{prox}_r(y)$

Из пред. свойства:

$$\langle x-u; z_1-u\rangle \le r(z_1)-r(u) \quad \forall z_1 \in \mathbb{R}^d$$

$$\langle y-v; z_2-v\rangle \le r(z_2)-r(v) \quad \forall z \in \mathbb{R}^d$$

$z_1 = v \qquad z_2 = u$

$$\left.\begin{aligned} &\langle x-u, v-u\rangle \le r(v)-r(u) \\ &\langle y-v; u-v\rangle \le r(u)-r(v) \end{aligned}\right| +$$

$$\langle x-u, v-u\rangle + \langle v-y; v-u\rangle \le 0$$

$$\langle v-u+x-y; v-u\rangle \le 0$$

$$\underbrace{\langle v-u; v-u\rangle}_{\|v-u\|_2^2} \le \underbrace{\langle x-y; u-v\rangle}_{\text{КБШ}}$$

$$\|v-u\|_2^{\cancel{2}} \le \|x-y\|_2 \cdot \cancel{\|u-v\|_2}$$

$$\|\mathrm{prox}_r(x)-\mathrm{prox}_r(y)\|_2 \le \|x-y\|_2 \quad \blacksquare$$

$$\min_{x\in\mathbb{R}^d} \left[f(x)+r(x)\right] \longleftarrow \text{композитная задача}$$

- f – выпуклая, L-гладкая
- r – выпуклая, проксимально дружественная (prox_r считается аналит.)

Проксимальный градиентный спуск

$$\boxed{x^{k+1} = \mathrm{prox}_{\gamma r}\left(x^k-\gamma\nabla f(x^k)\right)}$$

Задача

- $\mathrm{prox}_{\gamma r}\left(x^k-\gamma\nabla f(x^k)\right) = \arg\min\limits_{x\in\mathbb{R}^d}\left\{\gamma r(x) + \frac{1}{2}\|x-x^k+\gamma\nabla f(x^k)\|_2^2\right\}$

$$= \arg\min_{x\in\mathbb{R}^d} \left\{ \gamma r(x) + \frac{1}{2}\|x-x^k\|_2^2 + \gamma\langle \nabla f(x^k); x - x^k\rangle + \cancel{\frac{\gamma^2}{2}\|\nabla f(x^k)\|_2^2} \right\}$$

$$+ f(x^k)\cdot\gamma$$

$$= \arg\min_{x\in\mathbb{R}^d} \left\{ \gamma\cdot\Big(\underbrace{f(x^k) + \langle \nabla f(x^k); x-x^k\rangle}_{\text{линейная аппрокс } f} + \underbrace{r(x)}_{\text{точная инф. } r}\Big) + \frac{1}{2}\|x-x^k\|_2^2 \right\}$$

- r дифференцируема

$$\arg\min_{x\in\mathbb{R}^d} \underbrace{\left\{ \gamma r(x) + \frac{1}{2}\|x - x^k + \gamma\nabla f(x^k)\|_2^2 \right\}}_{\nabla = 0}$$

$$\gamma \nabla r(x) + x - x^k + \gamma \nabla f(x^k) = 0$$

$$x^{k+1}$$

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

неявная схема

Свойство (другое новое)

$$\forall \gamma > 0 \qquad x^* = \mathrm{prox}_{\gamma r}(x^* - \gamma \nabla f(x^*))$$

решение $\min f(x) + r(x)$

Док-во: Условие оптим.

$$0 \in \nabla f(x^*) + \partial r(x^*) \quad \Big|\cdot\gamma \qquad \pm x^*$$

$$x^* - \gamma\nabla f(x^*) - x^* \in \gamma\,\partial r(x^*)$$

Свойство нерастяж. оператора (лекц. 1+2)

$$x^* = \operatorname{prox}_{\gamma r}(x^* - \gamma \nabla f(x^*))$$

Док-во сходимости:

$$\|x^{k+1} - x^*\|_2^2 = \|\operatorname{prox}_{\gamma r}(x^k - \gamma \nabla f(x^k)) - x^*\|_2^2$$

св-во о фикс. точке

$$= \|\operatorname{prox}_{\gamma r}(x^k - \gamma \nabla f(x^k)) - \operatorname{prox}_{\gamma r}(x^* - \gamma \nabla f(x^*))\|_2^2$$

$$\|\operatorname{prox}(x) - \operatorname{prox}(y)\|_2 \le \|x-y\|_2$$

$$\le \|x^k - x^* - \gamma \nabla f(x^k) + \gamma \nabla f(x^*)\|_2^2$$

Дальше, как и в град. спуске с проекцией

⊕ сходится, как град. спуск для гладкой задачи

⊕ понятная физика

⊖ для „простых" функций r