Завдання 2

Микола Коломієць

6 грудня 2023 р.

Зміст

1 задача

Умова

Доведіть, що для гладкої задачі опуклого програмування $f \to \min\limits_C$ має місце:

$$f(x) = \min_C f \Leftrightarrow x \in C \quad \text{та} \quad (\nabla f(x), y - x) \geq 0, \quad \forall y \in C$$

Для опуклої функції виконується нерівність (минула дз)

$$f(y) \geq f(x) + (\nabla f(x), y - x),$$

Якщо $\forall y \in C$ маємо $(\nabla f(x), y - x) \geq 0$ то

$$f(y) \geq f(x) + (\nabla f(x), y - x) \geq f(x) \quad \forall y \in C,$$

тобто $f(x) = \min\limits_C f$.

Тепер навпаки(зправа наліво), запишемо тепер наближення першого порядку для f в x :

$$f(x) \leq f(y) = f(x) + (\nabla f(x), y - x) + o(\|y - x\|), \quad \forall y \in C.$$

Оскільки допустима множина C опукла, то разом із точками x та y до неї входять довільні їхні опуклі комбінації, тобто точки вигляду $(1-\lambda)x + \lambda y$, де $\lambda \in (0, 1)$. Підставляемо ї у останно нерівність:

$$0 \leq \lambda(\nabla f(x), y - x) + o(\lambda\|y - x\|), \quad \forall y \in C$$

Спрямовуючи $\lambda \to 0$ бачимо, що знак правої частини визначається перпим доданком (властивості o-малого), а тому маємо нерівність

$$0 \leq (\nabla f(x), y - x), \quad \forall y \in C.$$

2 задача

Умова

Нехай $g : E \to \mathbb{R} \cup \{+\infty\}$ - власна замкнена опукла функція. Довести, що

$$z = prox_g(x) \quad \Leftrightarrow \quad g(y) - g(z) \geq (x - z, y - z) \quad \forall y \in E$$

Нехай

$$z = \text{prox}_g(x) = \text{argmin}_{y \in E} \left\{ g(y) + \frac{1}{2}\|y - x\|^2 \right\}.$$

Тоді для $t \in (0, 1)$

$$g(z + t(y - z)) + \frac{1}{2}\|z + t(y - z) - x\|^2 \geq g(z) + \frac{1}{2}\|z - x\|^2.$$

Опуклість g дає

$$(1 - t)g(z) + \text{tg}(y) + \frac{1}{2}\|z + t(y - z) - x\|^2 \geq g(z) + \frac{1}{2}\|z - x\|^2.$$

Далі

$$g(y) - g(z) \geq \frac{t(x - z, y - z) + \frac{t^2}{2}\|y - z\|^2}{t}.$$

Переходимо до границі при $t \to 0$.

Тепер навпаки. Нехай

$$g(y) - g(z) \geq (x - z, y - z) \quad \forall y \in E.$$

Оскільки

$$2(x - z, y - z) = \|z - x\|^2 - \|y - x\|^2 + \|y - z\|^2,$$

то

$$g(y) + \frac{1}{2}\|y - x\|^2 \geq g(z) + \frac{1}{2}\|z - x\|^2 + \frac{1}{2}\|y - z\|^2 \quad y \in E.$$

Тому

$$z = \text{prox}_g(x) = \text{argmin}_{y \in E} \left\{ g(y) + \frac{1}{2}\|y - x\|^2 \right\}.$$

3 задача

Умова

Нехай $g : E \to \mathbb{R} \cup \{+\infty\}$ - власна замкнена та опукла функція. Довести, що

$$x = prox_g(x) \quad \Leftrightarrow \quad x \in argmin_{y \in E} g(y)$$

Нехай$x = \mathrm{prox}_g(x)$. З минулої задачі маємо

$$g(y) - g(x) \geq (x - x, y - x) = 0 \quad \forall y \in E$$

$$g(y) \geq g(x) \quad \forall y \in E \to x \in argmin_{y \in E} g(y)$$

В обернену сторону. Нехай $x \in argmin_{y \in E} g(y)$ тоді

$$\mathrm{prox}_g(x) = \mathrm{argmin}_{y \in E}\{g(y) + \frac{1}{2}\|y - x\|^2\} = x$$

адже в x функція g досягає свого мінімуму і $\|y - x\|$ не може бути меншою ніж у самій точці x.

4 задача

Умова

Нехай $g : E \to \mathbb{R} \cup +\infty$ — власна замкнена та опукла функція. Доведіть, що

$$(prox_g(x) - prox_g(y), x - y) \geq \|prox_g(x) - prox_g(y)\|_2^2 \quad \forall x, y \in E$$

З задачі 2 випливає, що при $z_1 = prox_g(y), z_2 = prox_g(x)$

$$g(p) - g(prox_g(y)) \geq (y - prox_g(y), p - prox_g(y)) = (y - prox_g(y), p) - (y - prox_g(y), prox_g(y))$$

$$g(p) - g(prox_g(x)) \geq (x - prox_g(x), p - prox_g(x)) = (x - prox_g(x), p) - (x - prox_g(x), prox_g(x))$$

Виберемо $p = prox_g(x)$ і $p = prox_g(y)$ відповідно. Отримаємо

$$g(prox_g(x)) - g(prox_g(y)) \geq (y - prox_g(y), prox_g(x)) - (y - prox_g(y), prox_g(y))$$

$$g(prox_g(y)) - g(prox_g(x)) \geq (x - prox_g(x), prox_g(y)) - (x - prox_g(x), prox_g(x))$$

Додамо нерівності

$$(y - prox_g(y), prox_g(x)) - (y - prox_g(y), prox_g(y)) + (x - prox_g(x), prox_g(y)) - (x - prox_g(x), prox_g(x)) \leq 0$$

$$(y, prox_g(x)) - (prox_g(y), prox_g(x)) - (y, prox_g(y)) +$$

$$+ (x, prox_g(y)) - (prox_g(x), prox_g(y)) - (x, prox_g(x)) + \|prox_g(x) - prox_g(y)\|^2 \leq 0$$

$$(prox_g(y) - prox_g(x), x - y) + \|prox_g(x) - prox_g(y)\|^2 \leq 0$$

Доведено

5 задача

Умова

Нехай $f : E \to R$ опукла та L-гладка відносно норми $\|\|_2$ функція, $g : E \to \mathbb{R} \cup +\infty$ власна замкнена та опукла функція. Доведіть, що

$$x = prox_{\lambda g}(x - \lambda \nabla f(x)) \quad \Leftrightarrow \quad x \in argmin_{y \in E}\{f(y) + g(y)\}$$

Зліва направо впипливає з задачі 3 і з градієнтного методу. Доведемо зправа наліво. Нехай $x \in argmin_{y \in E}\{f(y) + g(y)\}$. Тоді знову ж таки з задачі 3 і з того, що градієнтний метод збігається, маємо доведення. Не знаю, як по іншому записати це доведення, сподіваюсь пройде.

6 задача

Умова

Нехай $f : E \to \mathbb{R}$ μ-сильно опукла L-гладка функція. Розглянемо

$$\begin{cases} \ddot{x} + 2\sqrt{\mu}\dot{x} = -\nabla f(x) \\ x(0) = x_0, \quad \dot{x}(0) = 0, \quad \mu > 0 \end{cases}$$

Доведіть, що

$$f(x(t)) - f_* \leq 2e^{-\sqrt{\mu}t}(f(x_0) - f_*) \quad \forall t > 0$$

$$\ddot{x} + 2\sqrt{\mu}\dot{x} = 0$$

$$x(t) = e^{-2\sqrt{\mu}t} + C_1$$

З минулої задачі маємо, що

$$x^* \in argmin_{y\in E}\{f(y) + x(y)\} \quad \Leftrightarrow \quad x^* = prox_{\lambda x}(x^* + \lambda\nabla f(x))$$

$$x^* = prox_{\lambda x}(x^* + \lambda\nabla f(x^*)) = argmin_{y\in E}\{\lambda x(y) + \frac{1}{2}\|x^* + \lambda\nabla f(x^*) - y\|^2\}$$

$f_* = f_{min}$ $x^*(0) = x_0, \quad \dot{x}^*(0) = 0$

$$\frac{1}{2\sqrt{\mu}}\dot{x} + x = f(x)$$

Не знаю, як далі, заплутався.

7 задача

Умова

Доведіть, що для методу

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

послідовність

$$V_k = (2k+1)L\left(f(x_k) - f(x^*)\right) + k(k+2)\left\|\nabla f(x_k)\right\|_2^2 + L^2\left\|x^k - x^*\right\|_2^2$$

незростаюча.

Цей результат дає оцінку

$$\left\|\nabla f(x_k)\right\|_2^2 \leq \frac{L}{k(k+2)}\left(L\left\|x^0 - x^*\right\|_2^2 + f(x_0) - f(x^*)\right).$$

$$V_k = (2k+1)L\left(f(x_k) - f(x^*)\right) + k(k+2)\left\|\nabla f(x_k)\right\|_2^2 + L^2\left\|x^k - x^*\right\|_2^2$$

$$V_{k+1} = (2k+3)L\left(f(x_{k+1}) - f(x^*)\right) + (k+1)(k+3)\left\|\nabla f(x_{k+1})\right\|_2^2 + L^2\left\|x^{k+1} - x^*\right\|_2^2$$

$$V_{k+1} - V_k = (2k+1)L(f(x_{k+1}) - f(x_k)) + 2Lf(x_{k+1}) - 2Lf(x^*) +$$

$$+k(k+2)(\|\nabla f(x_{k+1})\|^2 - \|\nabla f(x_k)\|^2) + (2k+3)\nabla f(x_{k+1}) +$$

$$+L^2(\|x^{k+1} - x^*\|^2 - \|x^k - x^*\|^2)$$

Проксимальний оператор нерозтягуючий. Використаємо це.

$$\begin{aligned}
&\|x_{k+1} - x^*\|^2 = \\
&= \left\|\mathrm{prox}_{\lambda g}\left(x_k - \lambda\nabla f(x_k)\right) - \mathrm{prox}_{\lambda g}\left(x^* - \lambda\nabla f(x^*)\right)\right\|^2 \leq \\
&\quad \leq \|x_k - \lambda\nabla f(x_k) - x^* + \lambda\nabla f(x^*)\|^2 = \|x_k - x^*\|^2 - \\
&\quad - 2\lambda\left(\nabla f(x_k) - \nabla f(x^*), x_k - x^*\right) + \lambda^2\|\nabla f(x_k) - \nabla f(x^*)\|^2.
\end{aligned}$$

Тоді

$$V_{k+1} - V_k \leq (2k+1)L(f(x_{k+1}) - f(x_k)) + 2Lf(x_{k+1}) - 2Lf(x^*) +$$

$$k(k+2)(\|\nabla f(x_{k+1})\|^2 - \|\nabla f(x_k)\|^2) + (2k+3)\nabla f(x_{k+1}) -$$

$$-2L(\nabla f(x_k) - \nabla f(x^*), x_k - x^*) + \|\nabla f(x_k) - \nabla(x^*)\|^* \leq$$

Ну тут я старався, не знаю як тут далі треба, але сподіваюсь що хоч хід думок був правильний.

8 задача

Умова

8. Доведіть, що для градієнтного методу

$$x_{k+1} = x_k - \lambda \nabla f\left(x_k\right)$$

з $\lambda \in \left(0, \frac{2}{L}\right)$ має місце оцінка

$$f\left(x_k\right) - f\left(x^*\right) = O\left(\frac{1}{k}\right).$$

$$f(x_k) - f(x^*) = f(x_k) - f^* + \lambda(\nabla f(x_k), x_k) + O(\|\nabla f(x_k)\|^2)$$

З минулого завдання $\nabla f(x_k) = O(\frac{1}{k^2})$ отже данна сумма має як мінімум порядок $O(\frac{1}{k})$.

9 задача

Умова

9. Нехай $x-$ розв'язок задачі Коші

$$\begin{cases} \frac{d}{dt}x = -A^*Ax + A^*b \\ x(0) = x_0 \end{cases}$$

де $b, x_0 \in \mathbb{R}^n$, A - матриця $n \times n$. Що можна сказати про $\lim_{t\to+\infty} x(t)$? Відповідь обгрунтуйте.

Занадто абстрактне питання. Думаю що залежно від матриці і вектора можна отримати різні відповіді. Наприклад якщо $A = 0$ $b = 0$ відповідь буде $x(t) = 0$ якщо візьмемо $A = -E$ і $b! = 0$ то $x(t)$ буде розбігатися на нескінченності.

10 задача

Умова

Нехай x - розв'язок задачі Коші

$$\begin{cases} \frac{d}{dt}x = -Ax + b, \\ x(0) = x_0 \end{cases}$$

де $b, x_0 \in \mathbb{R}^n, A$ - симетрична додатньо визначена матриця $n \times n$. Що можна сказати про $\lim_{t \to +\infty} x(t)$? Відповідь обгрунтуйте.

Те саме, що і в минулом завданні тільки тепер $A \neq 0$. Тоді $x(t)$ буде розбігатися на нескінченності, якщо взяти її одиничною, а вектор b взяти оберненого знака до x_0 або довільного, якщо той нуль. І буде прямувати до x_0, якщо вектор b взяти його рівним x_0. Тож однозначної відповіді немає.