Home assignment 3

Bayan Hendawi bhendawi@edu.hse.ru
Katarina Kučuk kkuchuk@edu.hse.ru

April 26, 2021

Consider a closed and convex subset $K \subset \mathbb{R}^d$ as well as a comex function $f : K \to \mathbb{R}$. We suppose f is $\alpha-$convex, differentiable and $\beta-$smooth for $0 < \alpha < \beta$. For $x_1 \in K$ fixed, we define $\forall t \geqslant 1$:

$$x_{t+1} = \pi_K \left(x_t - \frac{1}{\beta} \nabla f\left(x_t\right) \right)$$

The goal of this home assignment is to prove the following:

Theorem

$\forall t \geqslant 1$

$$f\left(x_{t+1}\right) - f\left(x^*\right) \leqslant \frac{\beta}{2} \left(1 - \frac{\alpha}{\beta}\right)^t \|x_1 - x^*\|_2^2$$

Task 1

Problem: Using a minor modification of the proof of Lemma 4 in Lecture 9, show that $\forall x, y \in K$:

$$f\left(x^+\right) - f(y) \leqslant g_k(x)^\top (x-y) - \frac{1}{2\beta} \|g_K(x)\|_2^2 - \frac{\alpha}{2} \|x-y\|_2^2$$

where $x^+ := \pi_k \left(x - \frac{1}{\beta} \nabla f(x) \right)$ and $g_k(x) := \beta\left(x - x^+\right)$.

Solution: By the angular characterization of the projection onto K, we have that:

$$(x^+ - (x - \frac{1}{\beta} \nabla f(x)))^T \left(x^+ - y\right) \leqslant 0$$

Multiplying both sides by β and rearranging the terms yields, we get:

$$\nabla f(x)^T \left(x^+ - y\right) \leq g_K(x)^T \left(x^+ - y\right)$$

Now:

$$f(x^+) - f(y) = \underbrace{f(x^+) - f(x)}_{I} + \underbrace{f(x) - f(y)}_{II}$$

I :

$$f(x^+) - f(x) \overset{\text{Lemma 1}}{\leqslant} \nabla f(x)^T \left(x^+ - x\right) + \frac{\beta}{2} \|x^+ - x\|_2^2 =$$

$$\overset{g_K(x)=\beta\left(x^+-x\right)}{=} \nabla f(x)^T \left(x^+ - x\right) + \frac{\beta}{2} \frac{1}{\beta^2} \|g_K(x)\|_2^2 =$$

$$= \nabla f(x)^T \left(x^+ - x\right) + \frac{1}{2\beta} \|g_K(x)\|_2^2$$

II :

$$f(x)-f(y) \overset{\alpha-\text{convexity}}{\leqslant} \nabla f(x)^T\left(x-y\right)-\frac{\alpha}{2}\|x-y\|_2^2=$$

$$=\nabla f(x)^T\left(x-x^+\right)+\underbrace{\nabla f(x)^T\left(x^+-y\right)}_{\leq g_K(x)^T\left(x^+-y\right)}-\frac{\alpha}{2}\|x-y\|_2^2=$$

$$\leqslant \nabla f(x)^T\left(x-x^+\right)+g_K(x)^T\left(x^+-y\right)-\frac{\alpha}{2}\|x-y\|_2^2$$

I and II:

$$f(x^+)-f(y) \leq \nabla f(x)^T\left(x^+-x\right)+\frac{1}{2\beta}\|g_K(x)\|_2^2+\nabla f(x)^T\left(x-x^+\right)+g_K(x)^T\left(x^+-y\right)-\frac{\alpha}{2}\|x-y\|_2^2=$$

$$=g_K(x)^T\left(x^+-y\right)+\frac{1}{2\beta}\|g_K(x)\|_2^2-\frac{\alpha}{2}\|x-y\|_2^2=$$

$$=g_K(x)^T\left(x-y\right)+\underbrace{g_K(x)^T\left(x^+-x\right)}_{=-\frac{1}{\beta}\|g_K(x)\|_2^2}+\frac{1}{2\beta}\|g_K(x)\|_2^2-\frac{\alpha}{2}\|x-y\|_2^2=$$

$$=g_K(x)^T\left(x-y\right)-\frac{1}{2\beta}\|g_K(x)\|_2^2-\frac{\alpha}{2}\|x-y\|_2^2$$

Task 2

Problem: Show that $\forall t \geqslant 1$:

$$\|x_{t+1}-x^*\|_2^2=\|x_t-x^*\|_2^2-\frac{2}{\beta}g_K\left(x_t\right)^\top\left(x_t-x^*\right)+\frac{1}{\beta^2}\left\|g_k\left(x_t\right)\right\|_2^2$$

Solution: We defined following as:

$$x_{t+1}=\pi_K\left(x_t-\frac{1}{\beta}\nabla f\left(x_t\right)\right)$$

$$x^*=\underset{x\in K}{\operatorname{argmin}} f(x)$$

Now:

$$\|x_{t+1}-x^*\|_2^2=\|\pi_K(x_t-\frac{1}{\beta}\nabla f(x_t))-x^*\|_2^2=$$

$$\overset{x^*\in K\to\pi(x^*)=x^*}{=} \|x_{t+1}-x^*\|_2^2=\|\pi_K(x_t-\frac{1}{\beta}\nabla f(x_t))-\pi(x^*)\|_2^2 \leqslant$$

$$\overset{\text{proj. is 1-Lip.}}{\leqslant} \|x_t-\frac{1}{\beta}\nabla f(x_t)-x^*\|_2^2=$$

$$=\|x_t-x^*\|_2^2-2\frac{1}{\beta}\nabla f(x_t)^T\left(x_t-x^*\right)+\frac{1}{\beta^2}\|\nabla f(x_t)\|_2^2=$$

$$\overset{\nabla f(x_t)=g_K(x_t)}{=} \|x_t-x^*\|_2^2-\frac{2}{\beta}\,g_K(x_t)^T\left(x_t-x^*\right)+\frac{1}{\beta^2}\left\|g_K(x_t)\right\|_2^2$$

Task 3

Problem: Combining the two above questions, show that, $\forall t \geqslant 1$,

$$\|x_{t+1} - x^*\|_2^2 \leqslant \left(1 - \frac{\alpha}{\beta}\right) \|x_t - x^*\|_2^2 - \frac{2}{\beta} \left(f\left(x_{t+1}\right) - f\left(x^*\right)\right)$$

Solution: We know from Lemma 4 from the lecture that:

$$f(x_{t+1}) - f(x^*) \leqslant g_K(x)^T \left(x_t - x^*\right) - \frac{1}{2\beta} \|g_K(x)\|_2^2 - \frac{\alpha}{2}\|x_t - x^*\|_2^2 =$$

$$= \beta \left(x_t - x_{t+1}\right)^T \left(x_t - x^*\right) - \frac{\beta}{2}\|x_t - x_{t+1}\|_2^2 - \frac{\alpha}{2} \|x_t - x^*\|_2^2$$

Using following:

$$\|a\|^2 = \|a + b\|^2 - 2a^T b + \|b\|^2$$
$$2a^T b - \|b\|^2 = \|a + b\|^2 - \|a\|^2,$$

where $a = x_{t+1} - x^*$, $b = x_t - x_{t+1}$, we have:

$$f(x_{t+1}) - f(x^*) \leqslant \frac{\beta}{2} \left(\|x_t - x^*\|_2^2 - \|x_{t+1} - x^*\|_2^2\right) - \frac{\alpha}{2}\|x_t - x^*\|_2^2$$

$$\frac{\beta}{2}\|x_{t+1} - x^*\|_2^2 \leqslant \frac{\beta}{2}\|x_t - x^*\|_2^2 - \frac{\alpha}{2}\|x_t - x^*\|_2^2 - (f(x_{t+1}) - f(x^*))$$

$$\|x_{t+1} - x^*\|_2^2 \leqslant \|x_t - x^*\|_2^2 - \frac{\alpha}{\beta}\|x_t - x^*\|_2^2 - \frac{2}{\beta} \left(f(x_{t+1}) - f(x^*)\right) =$$

$$= \left(1 - \frac{\alpha}{\beta}\right) \|x_t - x^*\|_2^2 - \frac{2}{\beta} \left(f(x_{t+1}) - f(x^*)\right)$$

Task 4

Problem: Deduce from the previous inequality that the result of the theorem indeed holds.

Solution: From the previous task, we know:

$$f(x_{t+1} - f(x^*)) \leqslant \frac{\beta}{2} \left(1 - \frac{\alpha}{\beta}\right) \|x_t - x^*\|_2^2 - \frac{\beta}{2} \|x_{t+1} - x^*\|_2^2$$

We also know that $f(x_{t+1}) - f(x^*) \geqslant 0$ because $x^* = \underset{x \in K}{\operatorname{argmin}} f(x)$. Now:

$$0 \leqslant f(x_{t+1}) - f(x^*) \leqslant \frac{\beta}{2} \left(1 - \frac{\alpha}{\beta}\right) \|x_t - x^*\|_2^2 \underbrace{-\frac{\beta}{2}\|x_{t+1} - x^*\|_2^2}_{\leqslant 0} \leqslant$$

$$\leqslant \frac{\beta}{2} \left(1 - \frac{\alpha}{\beta}\right) \|x_t - x^*\|_2^2$$

Again, from the previous task, we have:

$$\|x_t - x^*\|_2^2 \leqslant \left(1 - \frac{\alpha}{\beta}\right) \|x_{t-1} - x^*\|_2^2 \underbrace{-\frac{2}{\beta} \left(f(x_t) - f(x^*)\right)}_{\leqslant 0} \leqslant$$

$$\leqslant \left(1 - \frac{\alpha}{\beta}\right) \|x_{t-1} - x^*\|_2^2$$

We repeat the same process and at the end we have:

$$\|x_t - x^*\|_2^2 \leqslant \left(1 - \frac{\alpha}{\beta}\right)^{t-1} \|x_1 - x^*\|_2^2$$

Finally:

$$f(x_{t+1}) - f(x^*) \leqslant \frac{\beta}{2} \left(1 - \frac{\alpha}{\beta}\right) \left(1 - \frac{\alpha}{\beta}\right)^{t-1} \|x_1 - x^*\|_2^2 = \frac{\beta}{2} \left(1 - \frac{\alpha}{\beta}\right)^{t} \|x_1 - x^*\|_2^2$$