MATH3431 Machine Learning and Neural Networks III Epiphany term 2024

Homework 1

Lecturer: Georgios P. Karagiannis georgios.karagiannis@durham.ac.uk

As formative assessment, submit the solutions to all the Exercises

---

Exercise 1. ($\star$) Let $f:\mathbb{R}^d \to \mathbb{R}$ be a convex and β-smooth function.

(1) Show that for $v, w \in \mathbb{R}^d$

$$f(v) - f(w) \in \left(\langle \nabla f(w), v - w \rangle, \langle \nabla f(w), v - w \rangle + \frac{\beta}{2} \|v - w\|^2 \right)$$

(2) Show that for $v, w \in \mathbb{R}^d$ such that $v = w - \frac{1}{\beta}\nabla f(w)$, it is

$$\frac{1}{2\beta} \|\nabla f(w)\|^2 \leq f(w) - f(v)$$

(3) Additionally assume that $f(x) > 0$ for all $x \in \mathbb{R}^d$. Show that for $w \in \mathbb{R}^d$,

$$\|\nabla f(w)\| \leq \sqrt{2\beta f(w)}$$

Solution.

(1) If $f:\mathbb{R}^d \to \mathbb{R}$ is β-smooth then it is

$$f(v) \leq f(w) + \langle \nabla f(w), v - w \rangle + \frac{\beta}{2} \|v - w\|^2$$

$$f(v) - f(w) \leq \langle \nabla f(w), v - w \rangle + \frac{\beta}{2} \|v - w\|^2$$

If it is convex then it is

$$f(v) \geq f(w) + \langle \nabla f(w), v - w \rangle$$

$$f(v) - f(w) \geq \langle \nabla f(w), v - w \rangle$$

Together these conditions imply upper and lower bounds

$$f(v) - f(w) \in \left(\langle \nabla f(w), v - w \rangle, \langle \nabla f(w), v - w \rangle + \frac{\beta}{2} \|v - w\|^2 \right)$$

(2) For $v, w \in \mathbb{R}^d$ such that $v = w - \frac{1}{\beta}\nabla f(w)$, it is

$$
\begin{aligned}
f(v) \leq f(w) + \langle \nabla f(w), v - w \rangle + \frac{\beta}{2}\|v - w\|_2^2 \quad &\text{(due to smoothness)} \\
\iff f(w) - f(v) \leq f(w) - f(v) & \\
\iff \langle \nabla f(w), v - w \rangle + \frac{\beta}{2}\|v - w\|_2^2 \leq f(w) - f(v) & \\
\iff \left\langle \nabla f(w), \frac{1}{\beta}\nabla f(w) \right\rangle + \frac{\beta}{2}\left\|\frac{1}{\beta}\nabla f(w)\right\|_2^2 \leq f(w) - f(v) & \\
\iff \frac{1}{2\beta}\|\nabla f(w)\|^2 \leq f(w) - f(v) & \\
\|\nabla f(w)\|^2 \leq 2\beta(f(w) - f(v)) &
\end{aligned}
$$

as $f(\cdot) \geq 0$

$$\|\nabla f(w)\|^2 \leq 2\beta f(w)$$

(3) From part 2, this is obvious because $f(x) > 0$ for all $x \in \mathbb{R}^d$, as

$$\|\nabla f(w)\|^2 \leq 2\beta f(w) \Leftrightarrow \|\nabla f(w)\| \leq \sqrt{2\beta f(w)}$$

Exercise 2. ($\star\star$) Let $f : \mathbb{R}^d \to \mathbb{R}$ be a λ-strongly convex function. Assume that w^* is a minimizer of f i.e.

$$w^* = \arg\min_w \{f(w)\}$$

Show that for any $w \in \mathbb{R}^d$ it holds

$$f(w) - f(w^*) \geq \frac{\lambda}{2}\|w - w^*\|^2$$

Hint: Use the definition of λ-strongly convex function, properly rearrange it, and ...

Solution. We use the definition of λ-strongly convex function; i.e. for all w, u, and $\alpha \in (0, 1)$ we have

$$
\begin{aligned}
f(aw + (1-\alpha)u) \leq af(w) + (1-\alpha)f(u) - \frac{\lambda}{2}\alpha(1-\alpha)\|w - u\|^2 \Leftrightarrow \\
\frac{f(aw + (1-\alpha)u) - f(u)}{\alpha} \leq f(w) + f(u) - \frac{\lambda}{2}(1-\alpha)\|w - u\|^2
\end{aligned}
$$

For $u = w^*$ it is

$$\frac{f(aw + (1-\alpha)w^*) - f(w^*)}{\alpha} \leq f(w) + f(w^*) - \frac{\lambda}{2}(1-\alpha)\|w - w^*\|^2$$

When $a \to 0$

$$\frac{\lambda}{2}\alpha(1-\alpha)\|w - w^*\|^2 \to 0$$

I know that w^* is the minimizer of f. So 0 is the minimizer of g with $g(a) = f(aw + (1-\alpha)w^*)$ hencewhen $a \to 0$

$$\frac{f(aw + (1-\alpha)w^*) - f(w^*)}{\alpha} \to \left.\frac{\mathrm{d}}{\mathrm{d}a}g(a)\right|_{a=0}$$

So

$$0 \leq f\left(w\right) + f\left(w^{*}\right) - \frac{\lambda}{2}\left\|w - w^{*}\right\|^{2}$$

which concludes the proof.