

If f has M -Lipschitz Hessian, then

$$\left| f(y) - f(x) - \langle \nabla f(x), y - x \rangle - \frac{1}{2} \langle \nabla^2 f(x)(y - x), (y - x) \rangle \right| \leq \frac{M}{6} \|y - x\|^3$$

Andersen Ang

Mathématique et recherche opérationnelle
UMONS, Belgium

manshun.ang@umons.ac.be Homepage: angms.science

First draft: December 16, 2018

Last update : July 28, 2019

The statement

For a twice differentiable function $f : \mathbb{R}^n \rightarrow \mathbb{R}$, it has M -Lipschitz continuous Hessian if

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

where $M > 0$. Here x, y are vectors in \mathbb{R}^n and $\|\cdot\|$ is L_2 norm.

If f is twice differentiable and has M -Lipschitz continuous Hessian, then

$$\left| f(y) - f(x) - \langle \nabla f(x), y - x \rangle - \frac{1}{2} \langle \nabla^2 f(x)(y - x), (y - x) \rangle \right| \leq \frac{M}{6} \|y - x\|^3$$

This document : prove this inequality.

This inequality comes from Lemma 1.2.4 of the book “Introductory Lectures on Convex Optimization - A basic course” by Nesterov. This document is to show the second part of the lemma.

We first show that we have

$$\|\nabla^2 f(x + \alpha(x - y)) - \nabla^2 f(x)\| \leq M\alpha\|x - y\|. \quad (1)$$

Proof of (1) : as f has M -Lipschitz Hessian, then

$$\|\nabla^2 f(a) - \nabla^2 f(b)\| \leq M\|a - b\|.$$

Put $a = x + \alpha(x - y)$, $b = x$ with $\alpha \in [0, 1]$,

$$\begin{aligned} \|\nabla^2 f(x + \alpha(x - y)) - \nabla^2 f(x)\| &\leq M\|\alpha(x - y)\| \\ &\leq M|\alpha| \cdot \|x - y\| \\ [\alpha \geq 0] &= M\alpha\|x - y\|. \end{aligned}$$

As f is twice differentiable, we have

$$\begin{aligned} f(y) &= f(x) + \langle \nabla f(x), y - x \rangle \\ &\quad + \int_{\tau=0}^{\tau=1} \int_{\alpha=0}^{\alpha=\tau} \langle \nabla^2 f(x + \alpha(y - x))(y - x), y - x \rangle d\alpha d\tau \end{aligned}$$

Standard mathematics trick

$$\begin{aligned} f(y) &= f(x) + \langle \nabla f(x), y - x \rangle \\ &\quad + \int_0^1 \int_0^\tau \langle (\nabla^2 f(x + \alpha(y - x)) + \nabla^2 f(x) - \nabla^2 f(x))(y - x), y - x \rangle d\alpha d\tau \\ &= f(x) + \langle \nabla f(x), y - x \rangle + \int_0^1 \int_0^\tau \langle \nabla^2 f(x)(y - x), y - x \rangle d\alpha d\tau \\ &\quad + \int_0^1 \int_0^\tau \langle (\nabla^2 f(x + \alpha(y - x)) - \nabla^2 f(x))(y - x), y - x \rangle d\alpha d\tau \end{aligned}$$

continue from last page

$$\begin{aligned} f(y) &= f(x) + \langle \nabla f(x), y - x \rangle + \underbrace{\left\langle \nabla^2 f(x)(y - x), y - x \right\rangle \int_0^1 \int_0^\tau d\alpha d\tau}_{\frac{1}{2}} \\ &\quad + \int_0^1 \int_0^\tau \left\langle \left(\nabla^2 f(x + \alpha(y - x)) - \nabla^2 f(x) \right) (y - x), y - x \right\rangle d\alpha d\tau \\ &= f(x) + \langle \nabla f(x), y - x \rangle + \frac{1}{2} \left\langle \nabla^2 f(x)(y - x), y - x \right\rangle \\ &\quad + \int_0^1 \int_0^\tau \left\langle \left(\nabla^2 f(x + \alpha(y - x)) - \nabla^2 f(x) \right) (y - x), y - x \right\rangle d\alpha d\tau \end{aligned}$$

Then

$$\begin{aligned} & f(y) - f(x) - \langle \nabla f(x), y - x \rangle + \frac{1}{2} \left\langle \nabla^2 f(x)(y - x), y - x \right\rangle \\ &= \int_0^1 \int_0^\tau \left\langle \left(\nabla^2 f(x + \alpha(y - x)) - \nabla^2 f(x) \right) (y - x), y - x \right\rangle d\alpha d\tau \end{aligned}$$

continue from last page, take absolute value

$$\begin{aligned} & \left| f(y) - f(x) - \langle \nabla f(x), y - x \rangle + \frac{1}{2} \langle \nabla^2 f(x)(y - x), y - x \rangle \right| \\ &= \left| \int_0^1 \int_0^\tau \langle (\nabla^2 f(x + \alpha(y - x)) - \nabla^2 f(x))(y - x), y - x \rangle d\alpha d\tau \right| \end{aligned}$$

i.e., to prove the inequality, we show the absolute value of the integral is upper bounded by $\frac{M}{6} \|y - x\|^3$.

By inequality of absolute value on integral, we have

$$\begin{aligned} & \left| \int_0^1 \int_0^\tau \langle (\nabla^2 f(x + \alpha(y - x)) - \nabla^2 f(x))(y - x), y - x \rangle d\alpha d\tau \right| \\ & \leq \int_0^1 \int_0^\tau \left| \langle (\nabla^2 f(x + \alpha(y - x)) - \nabla^2 f(x))(y - x), y - x \rangle \right| d\alpha d\tau \\ & \leq \int_0^1 \int_0^\tau \|\nabla^2 f(x + \alpha(y - x)) - \nabla^2 f(x)\| d\alpha d\tau \cdot \|y - x\|^2 \end{aligned}$$

We now use (1) to bound $\|\nabla^2 f(x + \alpha(y - x)) - \nabla^2 f(x)\|$ inside the integral

$$\begin{aligned} & \int_0^1 \int_0^\tau \|\nabla^2 f(x + \alpha(y - x)) - \nabla^2 f(x)\| d\alpha d\tau \cdot \|y - x\|^2 \\ & \stackrel{(1)}{\leq} \int_0^1 \int_0^\tau M\alpha \|x - y\| d\alpha d\tau \cdot \|y - x\|^2 \\ & = \underbrace{\int_0^1 \int_0^\tau \alpha d\alpha d\tau}_{\frac{1}{6}} \cdot M \|y - x\|^3 = \frac{M}{6} \|y - x\|^3 \quad \square \end{aligned}$$

End of document