Practice exam Q4 - v1

What if we run AGD-G until the norm of the gradient halves and then restart from that point?

To get $\|\nabla f(x_A^{out})\|_2 \leq \underbrace{\frac{1}{2}\|\nabla f(x_0)\|_2}_{=\varepsilon}$:

we need $O\left(\sqrt{\frac{L(f(x_0)-f(x^*))}{(\frac{1}{2})\|\nabla f(x_0)\|_2}}\right)$

$$= O\left(\sqrt{\frac{2L(f(x_0)-f(x^*))}{\|\nabla f(x_0)\|_2}}\right)$$

iterations.

After j calls to AGD-G, we will have

$$\|\nabla f(x_{A_j}^{out})\|_2 \leq \frac{1}{2^j}\|\nabla f(x_0)\|_2 \leq \varepsilon$$

$$2^j \geq \frac{\|\nabla f(x_0)\|_2}{\varepsilon}$$

$$j \geq \log_2\left(\frac{\|\nabla f(x_0)\|_2}{\varepsilon}\right)$$

The total number of iterations would be

$$\underbrace{O\left(\sqrt{\frac{2L(f(x_0)-f(x^*))}{\|\nabla f(x_0)\|_2}}\right)}_{\#\text{ of iterations per call}} \cdot \underbrace{\log_2\left(\frac{\|\nabla f(x_0)\|_2}{\varepsilon}\right)}_{\#\text{ of calls}}$$

from the inequality in Q3:

$$\sqrt{\frac{2L(f(x_0)-f(x^*))}{\|\nabla f(x_0)\|_2}} \leq \sqrt{\frac{2L(\frac{1}{2m})\|\nabla f(x_0)\|_2^2}{\|\nabla f(x_0)\|_2}}$$

$$= \sqrt{\frac{L}{m}\|\nabla f(x_0)\|_2}$$