Gradient Descent: Questions to Ponder

- Suppose $\eta = 0.1$ works well for $f(x)$, what about $g(x) = f(10x)$?
  - Do we want bigger steps or smaller steps?
  - How does the magnitude of the gradient compare between $g(x)$ and $f(x)$?
  - How does the Lipschitz constant compare between $g(x)$ and $f(x)$?

  Send your answer to: sysucsers@163.com titled "ml&dm-hw1-name-sid" by 23:59, March 17th, 2021.

---

1. $x' = x - \eta \nabla f(x)$

$g(\frac{x}{10}) = f(x') = f(x - \eta \nabla f(x)) = g(\frac{x}{10} - \frac{\eta}{10} \nabla f(x)) = g(\frac{x}{10} - \frac{\eta}{100} \nabla g(\frac{x}{10}))$

∴ 是原本 $\frac{1}{100}$ 倍

2. $g(x) = f(10x)$

$\nabla g(x) = 10 \nabla f(10x)$

∴ 是原本 10 倍

3. $\|\nabla f(x) - \nabla f(y)\| \leq L \|x - y\|$

$\Longleftrightarrow \|\frac{1}{10} \nabla g(\frac{x}{10}) - \frac{1}{10} \nabla g(\frac{y}{10})\| \leq 10 L \|\frac{x}{10} - \frac{y}{10}\|$

$\Longleftrightarrow \|\nabla g(\frac{x}{10}) - \nabla g(\frac{y}{10})\| \leq 100 L \|\frac{x}{10} - \frac{y}{10}\|$

∴ L 是原本 100 倍