Lipschitz continuity

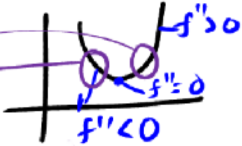
- Intuitively, a Lipschitz continuous function is limited in how fast it changes: there exists a definite real number such that, for every pair of points on the graph of the gradient, the absolute value of the slope of the line connecting them is not greater than this real number
    - This bound is called the function's Lipschitz constant, $L > 0$
- Thus, $\nabla f(x)$ is Lipschitz continuous if
  $$\underbrace{\left\|\nabla f(x) - \nabla f(y)\right\|_2 \leq L}_{\text{Rate of change of gradient is upper bounded}} \|x - y\|_2$$

Interpretation of Lipschitz continuity

- Consider $\nabla f(x) \in \mathbf{R}$, and $\nabla f(x) = \frac{df}{dx} = f'(x)$
- $|f'(x) - f'(y)| \leq L|x - y|$
  $\implies \frac{f'(x) - f'(y)}{|x-y|} \leq L$
  $\implies \left|\frac{f'(x+h) - f'(x)}{h}\right|$ (putting $y = x + h$)
- Taking limit $h \to 0$, we get
  $|f''(x)| \leq L$
- f'' represents curvature

(one can also show that if $|f''(x)| \leq L$ then f' is Lipschitz continuous

$|f''(x)|$ might be high in these regions



f' is decreasing

For a Lipschitz continuous $\nabla f \colon \mathbf{R}^n \to \mathbf{R}^n$, we can show that for any vector v,

- $v^\top \nabla^2 f(x) v \leq v^\top L v$
  $\implies v^\top (\nabla^2 f(x) - LI) v \leq 0$
- That is, $\nabla^2 f(x) - LI$ is negative semi-definite
- This can be written as:

$$\nabla^2 f(x) \preceq LI$$

Example: $f(x) = \frac{x^3}{3}$

- $f(x) = \frac{x^3}{3} \implies f'(x) = x^2$
- **Claim:** $f'(x)$ is locally Lipschitz continuous but not globally

  → in closed bounded interval it is Lipschitz cts
- Consider $x \in \mathbf{R}$
- $\sup_{y \in (x-1,x+1)} |f''(y)| = \sup_{y \in (x-1,x+1)} |2y| \leq 2|x| + 1$
- Applying mean value theorem:
  $\exists\, (y, z) \in (x-1, x+1)^2, \lambda$
  $f''(\lambda) = \frac{f'(y) - f'(z)}{y - z}$

- $|f'(y) - f'(z)| = |f''(\lambda)(y - z)|$
  $\leq \big|2|x| + 1\big|\, |y - x|,\ \forall (y, z) \in (x - 1, x + 1)^2$
- Thus, $L = \big|2|x| + 1\big|$
- Therefore, f' is Lipschitz continuous in $(x - 1, x + 1)$
- But as $x \to \infty$, $L \to \infty$
- This implies that f' may not be Lipschitz continuous everywhere
- Consider $y \neq 0$, and
  $\frac{f'(y) - f'(0)}{|y - 0|} = |y|$
- $|y| \to \infty$ as $y \to \infty$
- Thus, f' is proved to not be Lipschitz continuous globally

Another example

- Consider
$$f(x) = \begin{cases} x^2 sin\left(\frac{1}{x^2}\right) & \text{if } x \neq 0 \\ 0 & \text{if } x = 0 \end{cases}$$
- We can verify that this function is continuous and differentiable everywhere
i.e. $f''(0) = 0$ from left and right
- However, we can show that $f'(x)$ is *not* Lipschitz continuous

Lipschitz continuity: another example

- **Consider:** $f'(x) = |x|$
- Since $|f'(x) - f'(y)| = \big||x| - |y|\big| \leq |x - y|$,
  f' is Lipschitz continuous with $L = 1$
- However, it is not differentiable everywhere (not at 0)
- In fact, if f is continuously differentiable everywhere, it is also Lipschitz continuous
- $|x|$
  For functions over a closed and bounded subset of the real line:
  f is continuous $\supseteq$ f is differentiable (almost everywhere) $\supseteq$ f is Lipschitz continuous $\supseteq$ f' is continuous $\supseteq$ f' is differentiable

Considering gradients in Lipschitz continuity

- If ∇f is Lipschitz continuous, then

$$\left\|\nabla f(x) - \nabla f(y)\right\| \leq L\|x - y\|$$

- **Taylor's theorem** states that if f and its first n derivatives $f', f'', \ldots, f^{(n)}$ are continuous in the closed interval $[a, b]$, and differentiable in (a, b), then there exists a number $c \in (a, b)$ such that

$$f(b) = f(a) + f'(a)(b-a) + \frac{1}{2!}f''(a)(b-a)^2 + \ldots + \frac{1}{n!}f^{(n)}(a)(b-a)^n + \frac{1}{(n+1)!}f^{(n+1)}(c)(b-a)^{n+1}$$

- We will invoke Taylor's theorem up to the second degree:

$$f(y) = f(x) + f'(x)(y-x) + \frac{1}{2}f''(c)(y-x)^2$$

  where $c \in (x, y)$ and $x, y \in \mathbf{R}$
- Let us generalize to $f\colon \mathbf{R}^n \to \mathbf{R}$:

$$f(y) = f(x) + \nabla^\top f(x)(y-x) + \frac{1}{2}(y-x)^\top \nabla^2 f(c)(y-x)$$

  where $c = x + \Gamma(y-x)$, $\Gamma \in (0,1)$, and $x, y \in \mathbf{R}^n$
- If ∇f is Lipschitz continuous,

$$f(y) \le f(x) + \nabla^\top f(x)(y-x) + \frac{L}{2}\|y-x\|^2$$

- Convexity:

$$f(y) \geq f(x) + \nabla^\top f(x)(y - x)$$

- Strict convexity:

$$f(y) > f(x) + \nabla^\top f(x)(y - x)$$

- Strong convexity:

$$f(y) \geq f(x) + \nabla^\top f(x)(y - x) + \frac{m}{2}\|y - x\|^2$$

  - Strong convexity implies strict convexity
  - $\frac{m}{2}\|y - x\|^2$ can be 0 only when $y = x$