① ∇f is L-Continuous $\not\Rightarrow$ $f(x)$ is L-Continuous

$f(x) = x^2$ is NOT L-Cont

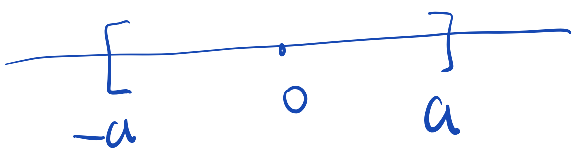
$f'(x) = 2x$ is L-Cont. $\Leftarrow$ $f''(x) = 2$ is bounded

② Even if $\nabla f(x)$ is not L-continuous on $\mathbb{R}$,

Convergence Theorems we proved may still apply

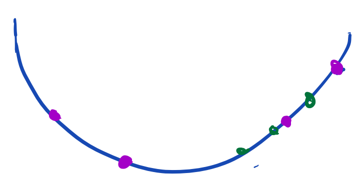
$f(x) = x^4$ is convex

$f'(x) = 4x^3$ is not L-Cont



$x_0 = a$

$x_{k+1} = x_k - \eta \nabla f(x_k)$



③ If ∇f is L-continuous with parameter $L > 0$

$f(x)$ is strongly convex with $\mu > 0$,

then $L \geq \mu$.

$$\langle \nabla f(x) - \nabla f(y), x - y \rangle \geq \mu \|x - y\|^2$$

$$\langle \nabla f(x) - \nabla f(y), x - y \rangle \leq \|\nabla f(x) - \nabla f(y)\| \cdot \|x - y\| \leq L \|x - y\|^2$$

$\frac{L}{\mu}$ is called condition number of $f(x)$

Example: $f(x) = \frac{1}{2} x^T K x - x^T b$

If ∇f is L-continuous with parameter $L > 0$

$f(x)$ is convex, then

$$\langle \nabla f(x) - \nabla f(y), x - y \rangle \geq \frac{1}{L} \|\nabla f(x) - \nabla f(y)\|^2 \quad ①$$

② is ① applied to $\phi(x) = f(x) - \frac{\mu}{2}\|x\|^2$

$$\langle \nabla f(x) - \nabla f(y), x - y \rangle \geq \frac{\mu L}{\mu + L} \|x - y\|^2 + \frac{1}{\mu + L} \|\nabla f(x) - \nabla f(y)\|^2 \quad ②$$

for a strongly convex function?

Case I: If $L = \mu$,

$$\langle \nabla f(x) - \nabla f(y), x-y \rangle \geq \frac{L}{2}\|x-y\|^2 + \frac{1}{2L}\|\nabla f(x) - \nabla f(y)\|^2$$

$$\left.\begin{array}{l}\text{Convexity} \\ L\text{-cont. } \nabla f\end{array}\right\} \Rightarrow \langle \nabla f(x) - \nabla f(y), x-y \rangle \geq \frac{1}{L}\|\nabla f(x) - \nabla f(y)\|^2$$

$$\text{strong convexity} \Rightarrow \langle \nabla f(x) - \nabla f(y), x-y \rangle \geq \mu\|x-y\|^2$$

Case II: If $L > \mu$, define $\phi(x) = f(x) - \frac{\mu}{2}\|x\|^2$.

Then $\begin{cases} 1)\ \phi(x) \text{ is convex} \\ 2)\ \nabla\phi(x) = \nabla f(x) - \mu x \end{cases}$

is L-continuous with $(L-\mu)$.

$$0 \leq \langle \nabla\phi(x) - \nabla\phi(y), x-y \rangle = \langle \nabla f(x) - \nabla f(y), x-y \rangle - \mu\|x-y\|^2 \leq (L-\mu)\|x-y\|^2$$

Lemma $f(x)$ is convex and $\nabla f(x)$ is L-cont. with L

$$\Longleftrightarrow 0 \leq \langle \nabla f(x) - \nabla f(y), x-y \rangle \leq L\|x-y\|^2$$

Proof: "$\Rightarrow$"

$$\text{Convexity} \Longleftrightarrow 0 \leq \langle \nabla f(x) - \nabla f(y), x-y \rangle$$

$$\text{C-S} \Rightarrow \langle \nabla f(x) - \nabla f(y), x-y \rangle \leq \|\nabla f(x) - \nabla f(y)\| \cdot \|x-y\|$$

∇f is L-continuous with parameter $L > 0$

$f(x)$ is strongly convex with $\mu > 0$,

$\Rightarrow \mu\|x-y\|^2 \leq \langle \nabla f(x) - \nabla f(y), x-y \rangle \leq L\|x-y\|^2$

$\langle \nabla\phi(x) - \nabla\phi(x), x-y \rangle$

$\begin{cases} 1) & \phi(x) \text{ is convex} \\ 2) & \nabla\phi(x) = \nabla f(x) - \mu x \\ & \text{is L-continuous with } (L-\mu). \end{cases}$

$\Rightarrow \langle \nabla\phi(x) - \nabla\phi(y), x-y \rangle \geq \frac{1}{L-\mu} \|\nabla\phi(x) - \nabla\phi(y)\|^2$

$\Rightarrow \langle \nabla f(x) - \nabla f(y), x-y \rangle - \mu\|x-y\|^2$

$$\geq \frac{1}{L-\mu}\|\nabla f(x)-\nabla f(y)-\mu(x-y)\|^2$$

$$=$$

$$\frac{1}{L-\mu}\|\nabla f(x)-\nabla f(y)\|^2+\frac{\mu^2}{L-\mu}\|x-y\|^2$$

$$-\frac{2\mu}{L-\mu}\langle \nabla f(x)-\nabla f(y), x-y\rangle$$

$$\Rightarrow \frac{L+\mu}{L-\mu}\langle \nabla f(x)-\nabla f(y), x-y\rangle \geq \frac{L\mu-\mu^2+\mu^2}{L-\mu}\|x-y\|^2$$

$$+\frac{1}{L-\mu}\|\nabla f(x)-\nabla f(y)\|^2$$

$$\Rightarrow$$

$$\langle \nabla f(x)-\nabla f(y), x-y\rangle \geq \frac{\mu L}{\mu+L}\|x-y\|^2+\frac{1}{\mu+L}\|\nabla f(x)-\nabla f(y)\|^2$$

$f(x_k)-f(x_*) = O(\frac{1}{k})$

<u>Theorem</u>

Assume ∇f is L-continuous. $0<\eta\leq\frac{2}{L}$

Assume $f(x)\geq f(x_*)$, $\forall x\in\mathbb{R}^n$

Assume $f(x)$ is strongly convex with μ

Then for $x_{k+1}=x_k-\eta\nabla f(x_k)$

with any $\eta\in\left(0,\frac{2}{L+\mu}\right]$:

$$\|x_k-x^*\|^2\leq\left(1-\frac{2\eta\mu L}{\mu+L}\right)^k\|x_0-x^*\|^2$$

$$r_{k+1}^2 = \| x_{k+1} - \eta \nabla f(x_k) \|^2$$

$$= \| x_k - x^* - \eta \nabla f(x_k) \|^2$$

$$= r_k^2 + 2\langle -\eta \nabla f(x_k), x_k - x^* \rangle + \eta^2 \| \nabla f(x_k) \|^2$$

$$= r_k^2 - 2\eta \langle \nabla f(x_k) - \nabla f(x_*), x_k - x^* \rangle$$

$$+ \eta^2 \| \nabla f(x_k) \|^2$$

$$\langle \nabla f(x) - \nabla f(y), x - y \rangle \geqslant \frac{\mu L}{\mu + L} \| x - y \|^2 + \frac{1}{\mu + L} \| \nabla f(x) - \nabla f(y) \|^2$$

$$\langle \nabla f(x_k) - \nabla f(x_*), x_k - x_* \rangle \geqslant \frac{\mu L}{\mu + L} \| x_k - x_* \|^2 + \frac{1}{\mu + L} \| \nabla f(x_k) \|^2$$

$$\leqslant r_k^2 - 2\eta \frac{\mu L}{\mu + L} \| x_k - x_* \|^2 - 2\eta \frac{1}{\mu + L} \| \nabla f(x_k) \|^2$$

$$+ \eta^2 \| \nabla f(x_k) \|^2$$

$$= \left(1 - 2\eta \frac{\mu L}{\mu + L}\right) r_k^2 + \eta \left(\eta - \frac{2}{\mu + L}\right) \| \nabla f(x_k) \|^2$$

So $\eta \in \left(0, \frac{2}{\mu + L}\right] \Rightarrow$

$$r_{k+1}^2 \leqslant \left(1 - 2\eta \frac{\mu L}{\mu + L}\right) r_k^2$$

If $\eta = \frac{2}{L+\mu}$, $1-2\eta\frac{\mu L}{\mu+L} = 1-\frac{2\mu L}{(\mu+L)^2}$

$$= \left(\frac{\mu-L}{\mu+L}\right)^2$$

$$= \left[\frac{L/\mu-1}{L/\mu+1}\right]^2$$