Proof.

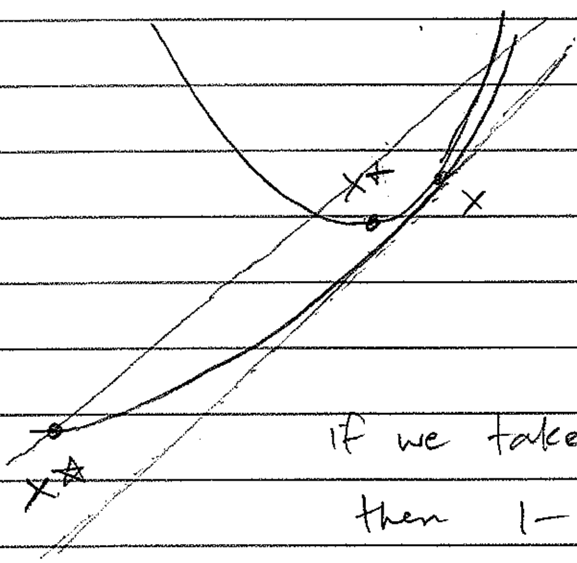
∇f Lipschitz w/ constant L

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

$\iff$

$$f(y) \leq f(x) + \nabla f(x)^T (y-x) + \frac{L}{2}\|y-x\|_2^2$$

for all y, x.



$$y = x^+ = x - t\nabla f(x)$$

$$f(x^+) \leq f(x) - \nabla f(x)^T t \nabla f(x) + \frac{L}{2} t^2 \|\nabla f(x)\|_2^2$$

$$= f(x) - \left(1 - \frac{Lt}{2}\right) \cdot t \, \|\nabla f(x)\|_2^2$$

If we take $t \leq \frac{1}{L}$

then $1 - \frac{Lt}{2} \geq \frac{1}{2}$, so $\underline{f(x^+) \leq f(x) - \frac{1}{2}t\|\nabla f(x)\|_2^2}$

$$f(x^\star) \geq f(x) + \nabla f(x)^T (x^\star - x)$$

$$f(x) \leq f(x^\star) + \nabla f(x)^T (x - x^\star) \quad \leftarrow \text{plug in}$$

$$f(x^+) \leq f(x^\star) + \nabla f(x)^T (x - x^\star) - \frac{t}{2}\|\nabla f(x)\|_2^2$$

$$f(x^+) - f(x^\star) \leq \frac{1}{2t}\left(\|x - x^\star\|_2^2 - \|\underbrace{x - t\nabla f(x)}_{x^+} - x^\star\|_2^2 \right)$$

$$\leq \frac{1}{2t}\left(\|x - x^\star\|_2^2 - \|x^+ - x^\star\|_2^2 \right)$$

sum this up over $i = 1 \ldots k$

$$\sum_{i=1}^{k} f(x^{(i)}) - f(x^\star) \leq \frac{1}{2t}\|x^{(0)} - x^\star\|_2^2 - \cancel{\frac{1}{2t}\|x^{(k)} - x^\star\|_2^2}$$

$$\leq \frac{1}{2t}\|x^{(0)} - x^\star\|_2^2$$

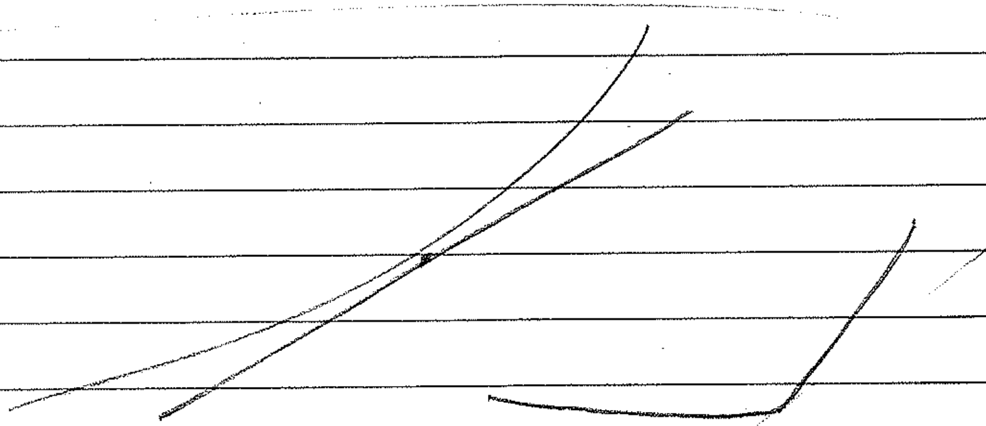
$$f(x^{(k)}) - f(x^\star) \le \frac{1}{k}\sum_{i=1}^{k} f(x^{(i)}) - f(x^\star)$$

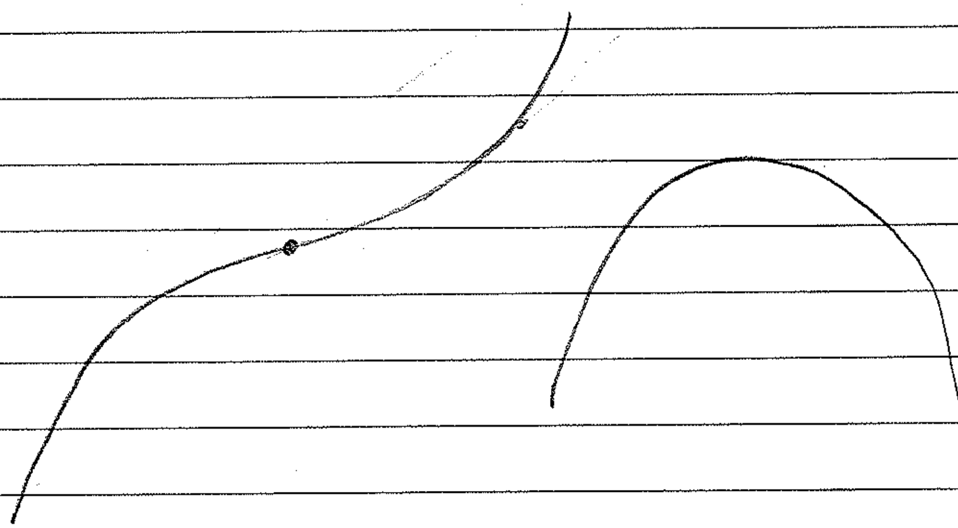
$$\le \frac{1}{2tk}\|x^{(0)} - x^\star\|_2^2 \qquad \square$$

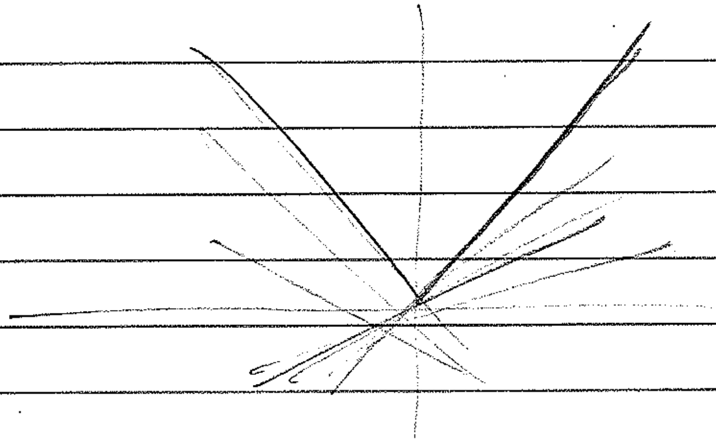
ε st $f(x^{(k)}) - f(x^\star) \le \varepsilon$

$c^k = \varepsilon$, $k = O(\log(1/\varepsilon))$ strong convexity

vs $O(1/\varepsilon)$ w/o







$$\nabla f(x) = \frac{x}{\|x\|_2}$$

$f(y) = \|y\|_2 \geq 0 + g^T y$ for all y

g any vector s.t. $\|g\|_2 \leq 1$

$$\|x\|_1 = \sum_{j=1}^{n} |x_j|$$

$g_i = \text{sign}(x_i)$ if $x_i \neq 0$
$\in [-1, 1]$ if $x \neq 0.$

$\partial f(x)$

$g_1, g_2 \in \partial f(x)$

$\alpha g_1 + (1-\alpha) g_2 = g$

$\alpha \cdot f(y) \geq (f(x) + g_1^T (y-x)) \, \alpha$

$(1-\alpha) f(y) \geq (f(x) + g_2^T (y-x))(1-\alpha)$

$f(y) \geq f(x) + g^T (y-x)$

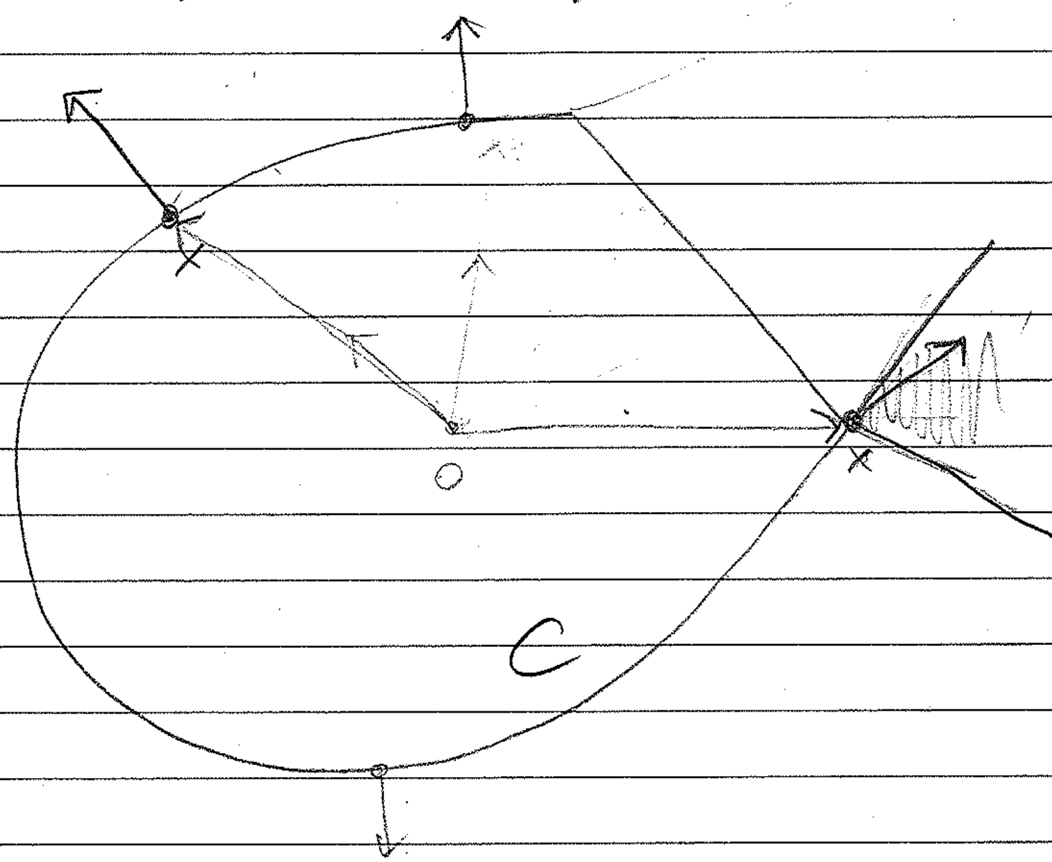
$C \subseteq \mathbb{R}^n$ convex

$$I_C(x) = \begin{cases} 0 & x \in C \\ \infty & x \notin C \end{cases}$$

if $x \in C$

$$\partial I_C(x) = N_C(x) = \{ g : g^T x \geq g^T y \text{ if } y \in C \}$$

$$I_C(y) \geq I_C(x) + g^T(y-x)$$



$$0 \in \partial f(x^*) \iff x^* \text{ minimizes } f$$

$$\min_{x \in C} f(x) = \min_{x} f(x) + I_C(x)$$

$$0 \in \partial f(x^*) + \partial I_C(x^*) \Leftarrow$$