$$f(y) = f(x) + \nabla f(x)^T (y-x) + \frac{1}{2}(y-x)^T \underbrace{\nabla^2 f(\theta(x-y)+y)}_{\frac{1}{t} I \quad \text{for some } \theta \in [0,1]}(y-x)$$

$$f(y) \approx \underbrace{f(x) + \nabla f(x)^T (y-x)}_{\text{linear approximation}} + \underbrace{\frac{1}{2t}\|y-x\|_2^2}_{\text{proximity term}} = g(y)$$

$\min g(y)$ $\quad \nabla g(y) = 0 \iff \nabla f(x) + \frac{1}{t}(y-x) = 0$

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

$$\| \quad x^+$$

$f(x) = x^2$ $\quad x_0 = 2$ $\quad t = 2$

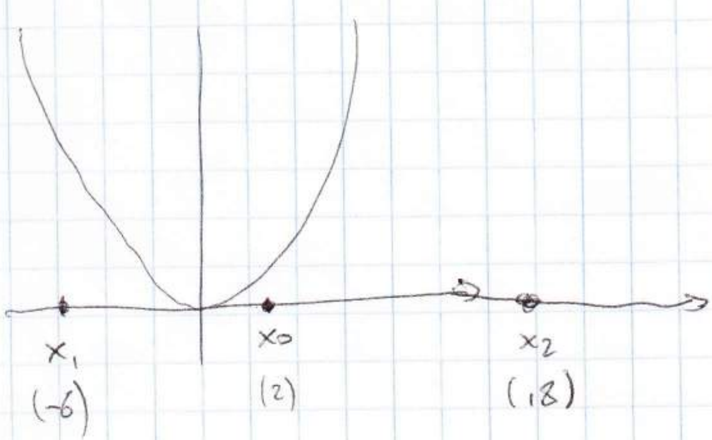
$\nabla f(x) = 2x$

$x_1 = x_0 - t\nabla f(x_0) = 2 - 2 \cdot 4 = 2 - 8 = -6$

$x^* = 0$

$x_2 = x_1 - t\nabla f(x_1) = -6 - 2\cdot(-12) = -6 + 24$

$= 18$



Try with $t = 1$!

Try with $t < 1$!

$\boxed{t \leq \frac{1}{L}}$

$$f(\cancel{x^{(k)} - x^*}) \quad f(x^{(k)}) - f(x^*) \leq \underbrace{\frac{\|x^{(0)} - x^*\|_2^2}{2tk}}_{\varepsilon} = O\left(\frac{1}{k}\right)$$

$$\varepsilon = \frac{\|x^{(0)} - x^*\|_2^2}{2tk} = \frac{C}{k} \iff k = \frac{C}{\varepsilon} = O\left(\frac{1}{\varepsilon}\right)$$

∇f being Lipschitz continuous with constant $L > 0$

$$\iff$$

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

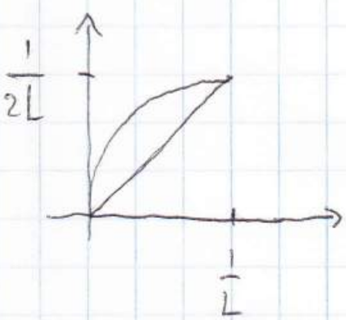
$$y = x^+ = x - t \nabla f(x) \iff \nabla f(x) = \frac{x - x^+}{t}$$

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

$$= f(x) - t \|\nabla f(x)\|_2^2 + \frac{Lt^2}{2} \|\nabla f(x)\|_2^2$$

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

Note that $\left(1 - \frac{Lt}{2}\right) t \geq \frac{t}{2}$ for $t \in \left(0, \frac{1}{L}\right]$



Backtracking $f(x^+) \leq f(x) - \alpha t \|\nabla f(x)\|_2^2$

$$(*) \leq f(x) - \frac{t}{2} \|\nabla f(x)\|_2^2 = f(x) - \frac{t}{2} \left\| \frac{x - x^+}{t} \right\|_2^2$$

$$= f(x) - \frac{1}{2t} \|x^+ - x\|_2^2 \qquad (**)$$

By the convexity of f

$$f(x^*) \geq f(x) + \nabla f(x)^T (x^* - x) \iff$$
$$f(x) \leq f(x^*) + \nabla f(x)^T (x - x^*)$$

$$\circledast\circledast \leq f(x^*) + \nabla f(x)^T (x - x^*) - \frac{1}{2t} \| x^+ - x \|_2^2$$

$$= f(x^*) + \left(\frac{x - x^+}{t} \right)^T (x - x^*) - \frac{1}{2t} \| x^+ - x \|_2^2$$

$$= f(x^*) + \frac{1}{t} (x - x^+)^T (x - x^*) - \frac{1}{2t} \| x^+ - x \|_2^2$$

$$= f(x^*) + \frac{1}{2t} \left[2 (x - x^+)^T (x - x^*) - \| x^+ - x \|_2^2 \right]$$

$$\| x - x^* \|_2^2 - \| x - x^* \|^2$$

$$= f(x^*) + \frac{1}{2t} \left[\| x - x^* \|_2^2 - \| x^+ - x^* \|_2^2 \right]$$

Now I have $f(x^+) - f(x^*) \leq \frac{1}{2t} \left[\| x - x^* \|_2^2 - \| x^+ - x^* \|_2^2 \right]$

$$\sum_{i=1}^{k} \left(f(x^{(i)}) - f(x^*) \right) = \frac{1}{2t} \sum_{i=1}^{k} \left(\| x^{(i-1)} - x^* \|_2^2 - \| x^{(i)} - x^* \|_2^2 \right)$$
$$= \frac{1}{2t} \left(\| x^{(0)} - x^* \|_2^2 - \| x^{(k)} - x^* \|_2^2 \right)$$

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

$$\frac{1}{k} \sum_{i=1}^{k} \left(f\left(x^{(i)}\right) - f\left(x^*\right) \right) \leq \frac{1}{2tk} \|x^{(0)} - x^*\|_2^2$$

$$f\left(x^{(k)}\right) - f\left(x^*\right) \leq \frac{1}{k} \sum_{i=1}^{k} \left(f\left(x^{(i)}\right) - f\left(x^*\right) \right) \leq \frac{1}{2tk} \|x^{(0)} - x^*\|_2^2$$