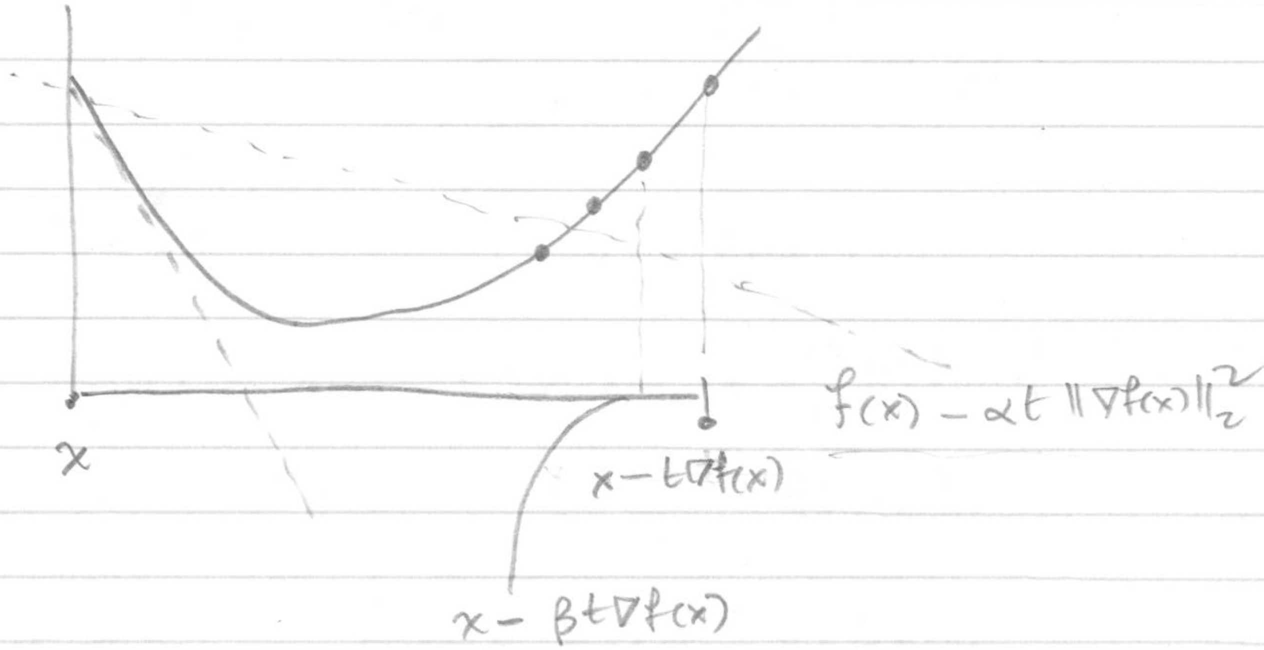
$$\hat{f}(y) = f(x) + \nabla f(x)^T (y-x) + \frac{1}{2t} \|y-x\|_2^2$$

$$\nabla \hat{f}(y) = \nabla f(x) + \frac{1}{t}(y-x) = 0$$

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



$$k = \frac{\|x^{(0)} - x^*\|_2^2}{2t} \cdot \frac{1}{\varepsilon}$$

- ∇f Lipschitz, constant L

$$\Rightarrow \quad f(y) \le f(x) + \nabla f(x)^T (y-x) + \frac{L}{2}\|y-x\|_2^2 \quad \text{all } x, y$$

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

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

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

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

(if $t \le \frac{1}{L}$, $f(x^+) \le f(x) - \frac{t}{2}\|\nabla f(x)\|_2^2$

take $0 < t \le \frac{1}{L}$.

$$f^* = f(x^*) \ge f(x) + \nabla f(x)^T (x^* - x)$$

by convexity

$$f(x) \le f(x^*) + \nabla f(x)^T (x - x^*)$$

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

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

check by plugging in $x^+ = x - t\nabla f(x)$

$$\|x^+ - x^*\|_2^2 = \|x - t\nabla f(x) - x^*\|_2^2$$
$$= \|x - x^*\|_2^2 - 2t\nabla f(x)^T(x - x^*) + t^2\|\nabla f(x)\|_2^2$$

arrived at

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

sum over $i = 1, \dots k$

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

since $f(x^{(i)})$ are monotone decreasing

$$f(x^{(k)}) - f^* \le \frac{1}{2tk}\|x^{(0)} - x^*\|_2^2 \checkmark$$