2. (a)

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

$f(x^+) \geq f(x) + \nabla f(x)^T (x^+ - x)$ ①

$x^+ - x = -t\nabla f(x)$ ②

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

$\therefore f(x^+) \leq f(x) + \nabla f(x)^T (x^+-x) + \frac{L}{2}\|x^+-x\|_2^2$ ③

②代入③ $f(x^+) \leq f(x) + \nabla f(x)^T (x^+-x) + \frac{L}{2}t^2\|\nabla f(x)\|_2^2$

$t \leq \frac{1}{L}$ $\therefore \frac{L}{2}t^2 \leq \frac{t}{2}$

$\therefore$ 得 $f(x^+) \leq f(x) + \underbrace{\nabla f(x)^T (x^+-x)}_{-t\|\nabla f(x)\|_2^2} + \frac{t}{2}\|\nabla f(x)\|_2^2$ ④ 得 $f(x^+) \leq f(x) - \frac{t}{2}\|\nabla f(x)\|_2^2$ ⑤ (题干)

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

$f(x) \leq f^* - \nabla f(x)^T (x^* - x)$ ⑥

⑤+⑥: $f(x^+) \leq f^* + \nabla f(x)^T (x - x^*) - \frac{t}{2}\|\nabla f(x)\|_2^2$ 得证

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

$\frac{1}{t^2}\|x^+-x\|_2^2$

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

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

$\therefore$ ①② 结合

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

(c) 无需证

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

凸 f 满足 $f(\frac{1}{k}\sum_{i=1}^{k}x^{(i)}) \leq \frac{1}{k}\sum_{i=1}^{k}f(x^{(i)})$

$\bar{x}^{(k)} = \frac{1}{k}\sum_{i=1}^{k}x^{(i)}$

$f(\bar{x}^{(k)}) - f^* \leq \frac{1}{k}\sum(f(x^{(i)})-f^*)$

∵函数值 $f(x^{(i)})$ 是单调下降的

$f(x^{(1)}) \geq f(x^{(2)}) \geq \cdots \geq f(x^{(k)})$

∴ $f(x^{(k)}) - f^* \leq f(\bar{x}^{(k)}) - f^*$

∴由①得 $f(\bar{x}^{(k)}) - f^* \leq \frac{1}{2tk}\|x^{(0)}-x^*\|_2^2$

⇓

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

3.

(a) 若 h 是适当的闭凸函数，假设 h 至少在定义域一点处存在次梯度

$m(u)=h(u)+\frac{1}{2}\|u-x\|_2^2$

$h(u)$ 凸函数. 且至少在一点处存在次梯度. ∴ $h(u)$ 有全局[illegible]

$h(u) \geq h(v)+\theta^T(u-v)$ $\theta \in \partial h(v)$

$m(u) = h(u)+\frac{1}{2}\|u-x\|^2 \geq h(v)+\theta^T(u-v)+\frac{1}{2}\|u-x\|^2$

∴ $m(u)$ 有二次下界，当 $\|u\| \to +\infty$, $m(u) \to +\infty$ ∴ $m(u)$ 有强制性

∴有最小值 (Homework 2)

唯一性: $m(u)$ 强凸，$m(u)$ 最小值唯一

(由于 f 为强凸且存在最小值，∴ f 最小值唯一)

(b)

$$\min_z \left\{ \frac{1}{2t}\|x-z\|_2^2 + \frac{1}{2}z^TAz - b^Tz \right\}$$

$$x^{(k+1)} = \arg\min_z \left\{ \frac{1}{2t}\|x^{(k)}-z\|_2^2 + \frac{1}{2}z^TAz - b^Tz \right\}$$

$$\frac{\partial}{\partial z} = \frac{1}{t}(z - x^{(k)}) + Az - b = 0$$

得 $z = \left(\frac{1}{t}I + A\right)^{-1}\left(\frac{1}{t}x^{(k)} + b\right)$

$\therefore x^{(k+1)} = \left(\frac{1}{t}I + A\right)^{-1}\left(\frac{1}{t}x^{(k)} + b\right)$

$\varepsilon = \frac{1}{t}$ 得

$$x^{(k+1)} = (A + \varepsilon I)^{-1}(b - Ax^{(k)} + Ax^{(k)})$$
$$= x^{(k)} + (A + \varepsilon I)^{-1}(b - Ax^{(k)})$$

证毕