



Convex Optimization
Week 12

清华大学
(2/5/22)

2/5/22
14:27
2021/2/5

$$1. \quad f(x_{k+1}) = f(x_k - t \nabla f(x_k))$$

1)

$$\begin{aligned} &= f(x_k) - t \nabla f(x_k)^T \nabla f(x_k) + \frac{t^2}{2} \nabla^2 f(x_k) \nabla f(x_k)^T \nabla f(x_k) + o(t^2) \\ &= f(x_k) - t \nabla f(x_k)^T \nabla f(x_k) + \frac{t^2}{2} \|\nabla f(x_k)\|_2^2 + o(t^2) \\ &\quad \nabla^2 f(x_k) \leq L I \\ &\quad 0 \leq t \leq \frac{2}{L} \Rightarrow \frac{t^2}{2} \|\nabla f(x_k)\|_2^2 \geq 0 \end{aligned}$$

$$\Rightarrow f(x_{k+1}) \leq f(x_k) - (1 - \frac{Lt}{2}) \frac{t}{2} \|\nabla f(x_k)\|_2^2$$

$$2) \quad \frac{1}{2} \|\nabla f(x_k)\|_2^2 \leq$$

$$\begin{aligned} &(1 - \frac{Lt}{2}) \frac{t}{2} \|\nabla f(x_k)\|_2^2 \geq f(x_k) - f(x_{k+1}) \\ &0 \leq t \leq \frac{2}{L} \Rightarrow 1 - \frac{Lt}{2} \geq 0, t > 0, \\ &\frac{1}{2} \|\nabla f(x_k)\|_2^2 \geq \end{aligned}$$

$$(1 - \frac{Lt}{2}) \frac{t}{2} \|\nabla f(x_k)\|_2^2 \leq f(x_k) - f(x_{k+1})$$

$$\Rightarrow 0 < t \leq \frac{2}{L} \Rightarrow (1 - \frac{Lt}{2}) \frac{t}{2} \|\nabla f(x_k)\|_2^2 \leq \frac{2}{-Lt^2 + 2t} (f(x_k) - f(x_{k+1}))$$

$$\Rightarrow t \leq \frac{2}{L} \Rightarrow -Lt^2 + 2t \geq t$$

$$2t - Lt^2 \geq t$$

$$\frac{2}{-Lt^2 + 2t} \leq \frac{2}{t}$$

$$\|\nabla f(x_k)\|_2^2 \leq \frac{2}{t} (f(x_k) - f(x_{k+1}))$$

3) 由(2)可知(2)里#号,有

$$\sum_{i=0}^k \|\nabla f(x_i)\|_2^2 \leq \frac{2}{t} (f(x_0) - f(x_{k+1}))$$

$$+ f(x_0) - f(x_{k+1})$$

$$\sum_{i=0}^k \|\nabla f(x_i)\|_2^2 \leq \frac{2}{t} (f(x_0) - f(x_{k+1}))$$

$$\therefore f(x_{k+1}) \geq f(x^*)$$

$$\therefore \sum_{i=0}^k \|\nabla f(x_i)\|_2^2 \leq \frac{2}{t} (f(x_0) - f(x^*))$$

$$4) \quad \lim_{i \rightarrow \infty} \|\nabla f(x_i)\|_2^2 \leq \frac{2}{t(k+1)} (f(x_0) - f(x^*))$$

$$\therefore \lim_{i \rightarrow \infty} \|\nabla f(x_i)\|_2 \leq \sqrt{\frac{2}{t(k+1)} (f(x_0) - f(x^*))}$$

$$2. (1) p_{\alpha}(x) = \arg \min_x \frac{\alpha}{2} (Ax-b)^T (Ax-b) + \frac{1}{2} (x-x_0)^T (x-x_0)$$

$$= \arg \min_x \frac{\alpha}{2} x^T A^T A x - \alpha (A^T b)^T x + \frac{1}{2} x^T x - \frac{1}{2} x_0^T x + \frac{1}{2} x_0^T x_0 = G(x)$$

$$\nabla G(x) = (A^T A + I) x - \alpha A^T b - x_0$$

$$\nabla^2 G(x) = A^T A + I \quad \because \alpha \gg 1$$

$$\therefore \nabla^2 G(x) \approx I$$

$$\therefore x = (\alpha A^T A + I)^{-1} (\alpha A^T b + x_0)$$

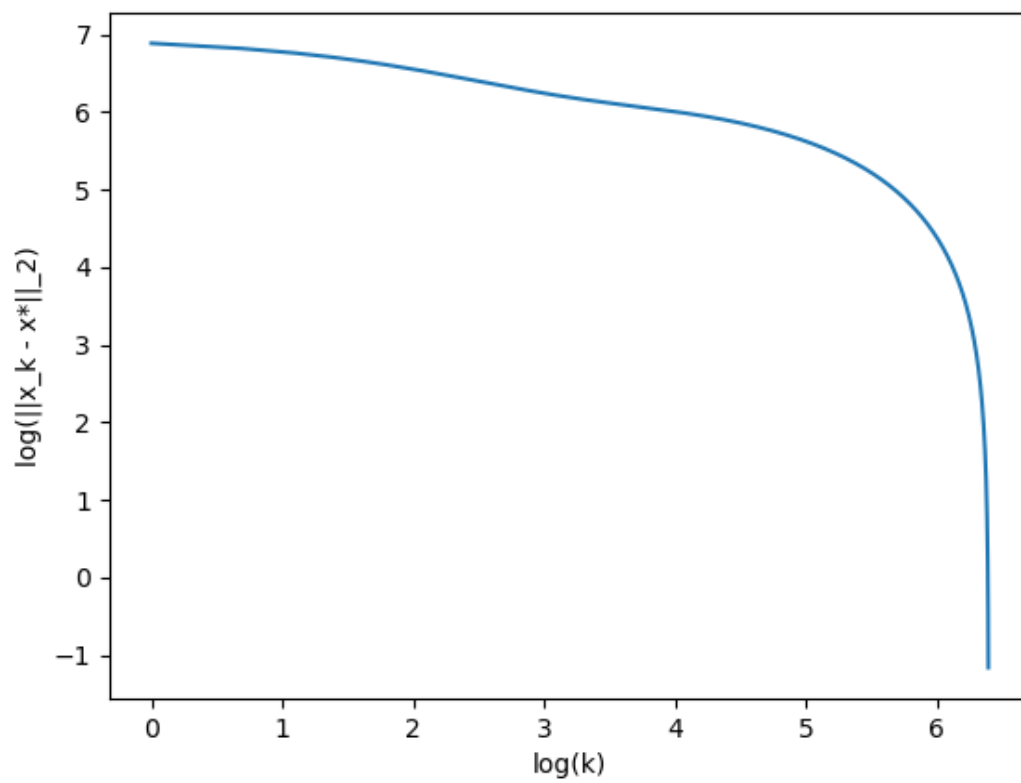
$$x_{k+1} = \frac{\alpha A^T A + I}{(\alpha A^T A + I)}^{-1} (\alpha A^T b + x_k)$$

0 1 2 3

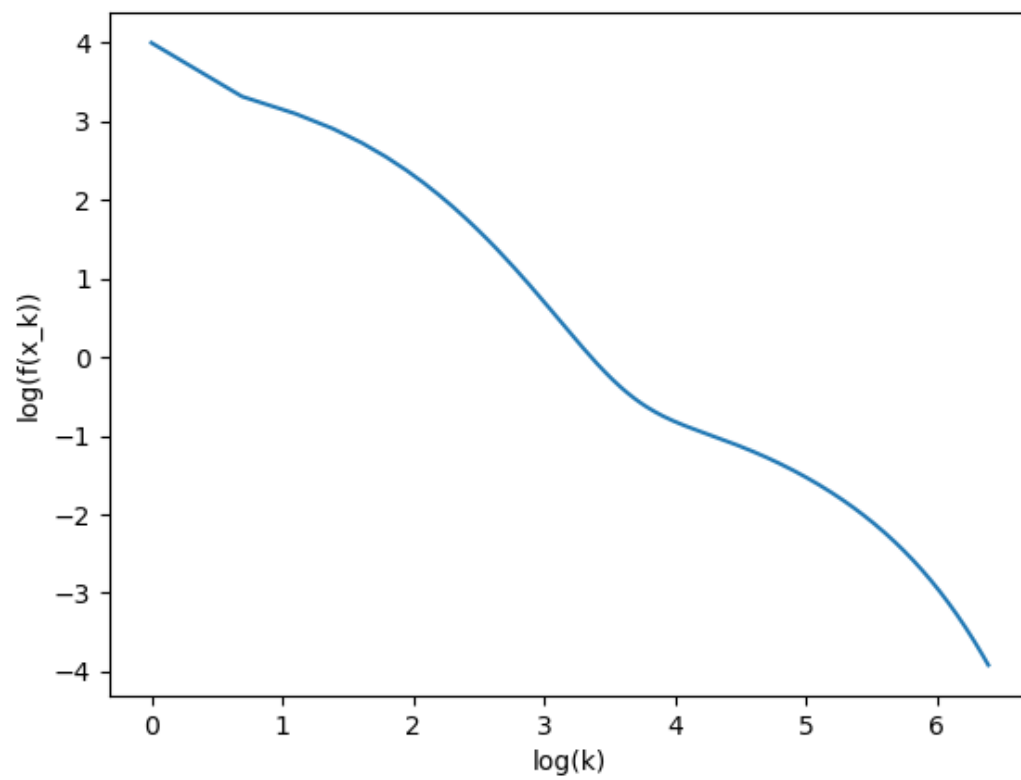
(2)

采用 $\alpha=1000$ 时, 收敛到 $f(x)=0.01991554109994043$, 此时 x 结果见根目录下result/Q2/x.csv文件

$\log(k)$ 和 $\log(\|x_k - x^*\|^2)$ 的图像如下:



$\log(k)$ 和 $\log(f(x_k))$ 的图像如下：



$\min_x \frac{1}{2} \|Ax - b\|_2^2 + \|x\|_1$ 的次梯度 $g(x)$ 为

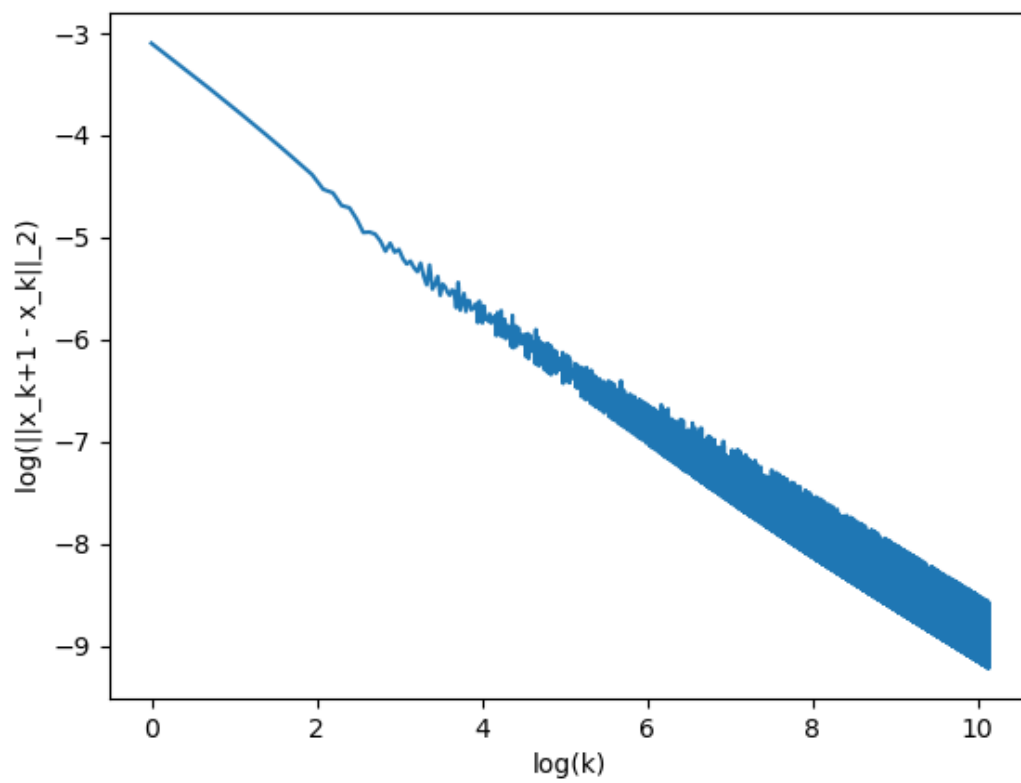
$$A^T Ax - A^T b + \text{sign}(x), \text{ 其中 } \text{sign}(x)_i = \begin{cases} 1 & x_i > 0 \\ 0 & x_i = 0 \\ -1 & x_i < 0 \end{cases}$$

具体推导如下:

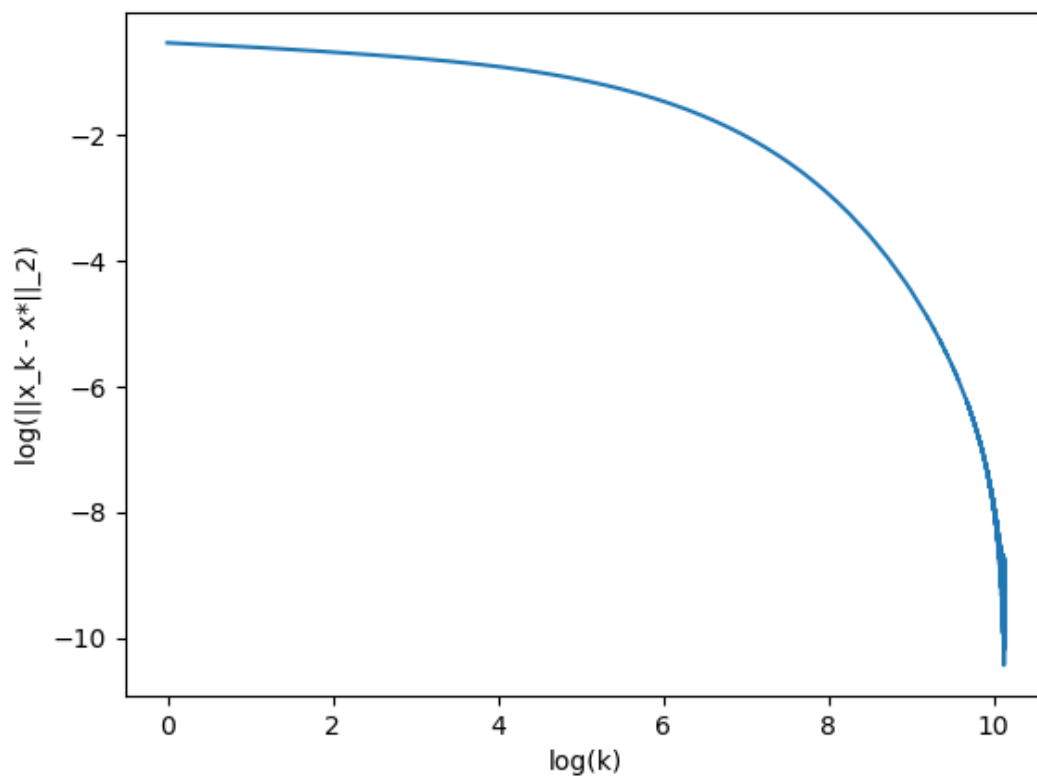
3.1 $h(x) = \|Ax - b\|_2^2 + \|x\|_1$,
 $\Rightarrow \|Ax - b\|_2^2 = f(x)$ 在 \mathbb{R}^n 上可微,
 $\nabla f(x) = A^T Ax - A^T b$,
 \therefore 只需求 $\|x\|_1$ 在 \mathbb{R}^n 上的次梯度 $\varphi(x)$,
 $g(x) = \nabla f(x) + \varphi(x)$
 任取 y, x , $\|y\|_1 - \|x\|_1 \geq (y-x)^T \varphi(x)$
 $\therefore \|y\|_1 = \sum_{i=1}^n |y_i|, \|x\|_1 = \sum_{i=1}^n |x_i|$
 $\therefore \|y\|_1 - \|x\|_1 = \sum_{i=1}^n |y_i| - |x_i|$
 $(y-x)^T \varphi(x) = \sum_{i=1}^n \varphi(x_i) y_i - \varphi(x_i) x_i$
 任取 y_i, x_i ,
 有 $|y_i| - |x_i| \geq \varphi(x_i) y_i - \varphi(x_i) x_i$, 当 $x_i > 0$ 时, $\varphi(x_i) = 1$, 有
 即 $|y_i| \geq \varphi(x_i) y_i$, $|y_i| \geq \varphi(x_i) y_i$,
 $\varphi(x_i) x_i \geq |x_i|$ $-|x_i| = -x_i$, 左对右
 当 $x_i < 0$ 时, $\varphi(x_i) = -1$,
 有 $|y_i| \geq \varphi(x_i) y_i$,
 $-|x_i| = x_i$, 左对右
 当 $x_i = 0$ 时, $\varphi(x_i) = 0$
 有 $0 = \varphi(x_i) y_i \geq 0$
 即 $\varphi(x_i) = \begin{cases} 1 & x_i > 0 \\ 0 & x_i = 0 \\ -1 & x_i < 0 \end{cases}$ 即可

解得最优的 $h(x) = 12.522447423125461$, x 存储在根目录下的 `result/Q3/x1.csv` 中

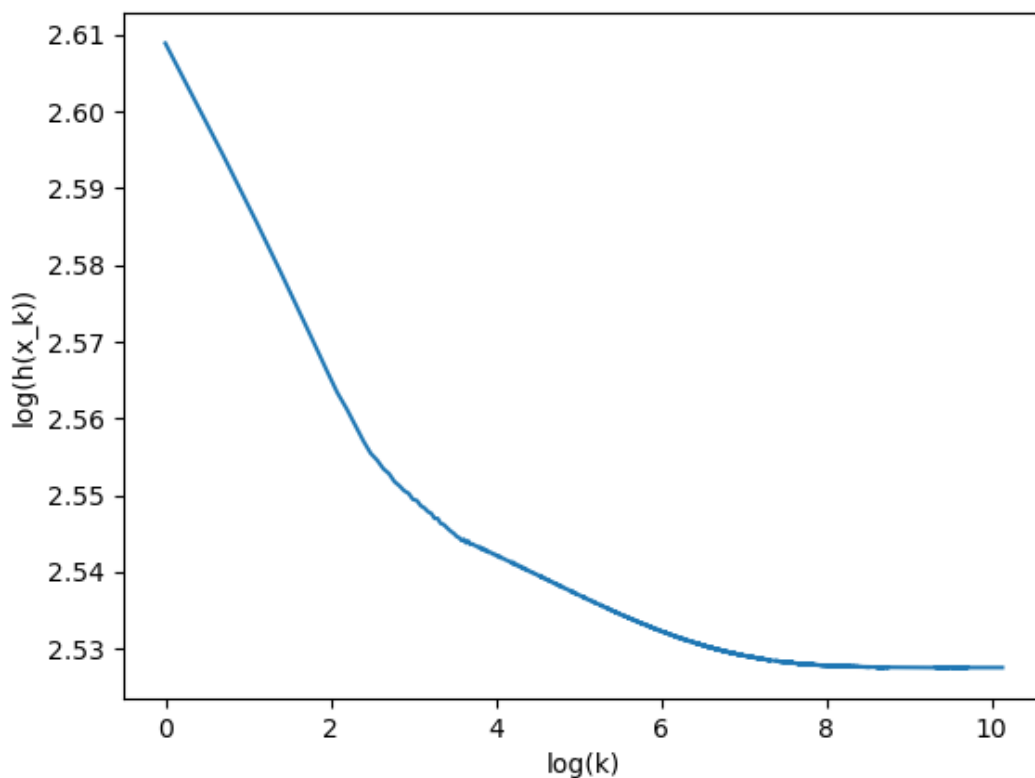
$\log(k)$ 和 $\log(\|x_{k+1} - x_k\|_2)$ 的图像如下:



$\log(k)$ 和 $\log(\|x_k - x^*\|_2)$ 的图像如下:



$\log(x)$ 和 $\log(h(x))$ 的图像如下:



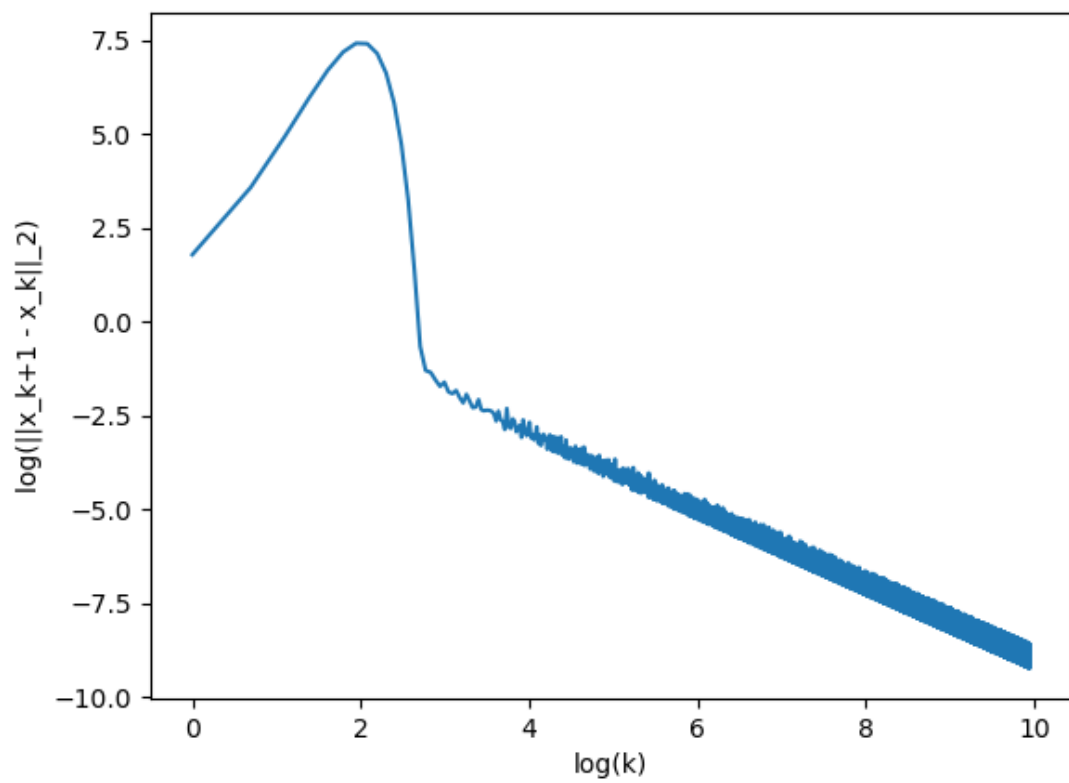
(2)

$$\begin{aligned} \text{对 } f \text{ 做泰勒展开, 有 } f(y) &= f(x) + \nabla f(x)^T(y-x) + \frac{1}{2}(y-x)^T \nabla^2 f(x)(y-x) \\ &\geq f(x) + \nabla f(x)^T(y-x) + \frac{1}{2}m(y-x)^T(y-x), \text{ 其中 } \nabla^2 f(x) = A^T A \end{aligned}$$

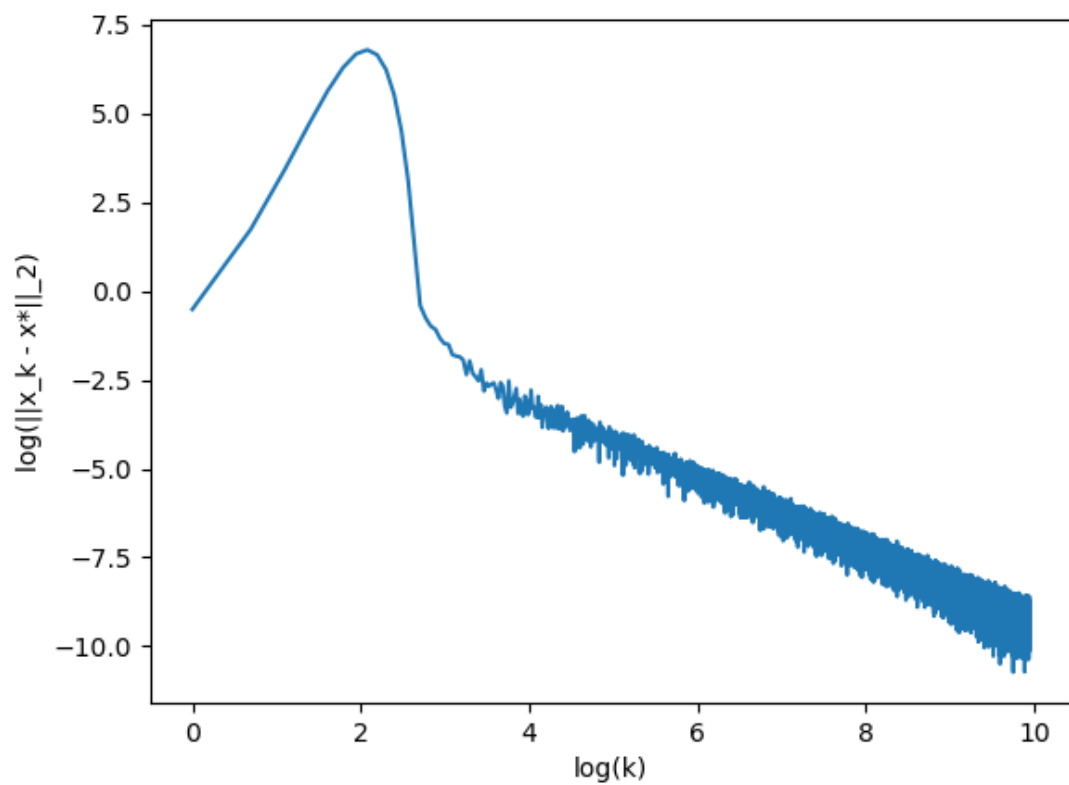
因此有对于任意 $x, y \in \text{dom}(f)$, 有 $(y-x)^T(A^T A - mI)(y-x) \geq 0$, 也就是 $A^T A - mI$ 半正定
 m 最大值为 $A^T A$ 最小特征值

解得 $m=0.7569455027845506$, 最优的 $h(x)=12.52240996025351$, x 存储在根目录下的
 result/Q3/x2.csv 中

$\log(k)$ 和 $\log(\|x_{k+1} - x_k\|_2)$ 的图像如下:



$\log(k)$ 和 $\log(\|x_k - x^*\|_2)$ 的图像如下:



$\log(k)$ 和 $\log(h(x_k))$ 的图像如下:

