自93 姜水鹏 2019010465 凸优化 2022/11/27

1. (a). $\forall t\in\mathbb{R},\ d\in\mathbb{R}^n,\ g(t)=\nabla f(x+td)$

$$g'(t)=\nabla^2 f(x+td)d$$

$$g(\alpha)-g(0)=\nabla f(x+\alpha d)-\nabla f(x)=\int_0^\alpha g'(t)\,dt$$

即 $\left\|\left(\int_0^\alpha \nabla^2 f(x+td)\,dt\right)d\right\|_2$

$$=\|\nabla f(x+\alpha d)-\nabla f(x)\|_2$$

$$\le \alpha L\|d\|$$

令 $\alpha\to 0^+$ $\|\nabla^2 f(x)d\|\le L\|d\|$

由诱导范数. $\|\nabla^2 f(x)\|=\max_d \frac{\|\nabla^2 f(x)d\|}{\|d\|}\le L.$

在x处 Taylor 展开

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

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

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

$(\forall x,y\in\mathbb{R}^n)$

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

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

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

(c). 由(b).

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

$\because t\le \frac{1}{L}$, 故 $\left(1-\frac{Lt}{2}\right)\cdot t\in\left[0,\frac{1}{2L}\right]$

$$\|\nabla f(x_k)\|_2^2\le 2L\left(f(x_k)-f(x_{k+1})\right)$$

$$\le\frac{2}{t}\left(f(x_k)-f(x_{k+1})\right)$$

(d). 由(c).

$$\sum_{i=0}^k\|\nabla f(x_i)\|_2^2=\|\nabla f(x_0)\|_2^2+\|\nabla f(x_1)\|_2^2+\cdots+\|\nabla f(x_k)\|_2^2$$

$$\le\frac{2}{t}(f(x_0)-f(x_1))+\frac{2}{t}(f(x_1)-f(x_2))+\cdots+\frac{2}{t}(f(x_k)-f(x_{k+1}))$$

$$=\frac{2}{t}\left(f(x_0)-f(x_{k+1})\right)$$

~~迭代停止时 $x_{k+1}=x$~~ $f(x^*)\le f(x_{k+1})$

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

(e). 由(d)

$$\left(\min_{i=0,\dots,k}\|\nabla f(x_i)\|_2\right)^2\cdot(k+1)$$

$$\le\sum_{i=0}^k\|\nabla f(x_i)\|_2^2$$

$$\le\frac{2}{t}\left(f(x_0)-f(x^*)\right)$$

故. $\min_{i=0,\dots,k}\|\nabla f(x_i)\|_2\le\sqrt{\frac{2}{t(k+1)}\left(f(x_0)-f(x^*)\right)}$

2. (a).

$$\mathrm{Prox}_{\alpha f}(v)=\arg\min_x \frac{\alpha}{2}\|Ax-b\|_2^2+\frac{1}{2}\|x-v\|^2$$

梯度条件. $\alpha A^T(Ax-b)+x-v=0$

$$\mathrm{Prox}_{\alpha f}(v)=(\alpha A^TA+I)^{-1}(v+\alpha A^Tb)$$

不动点迭代 $x_{k+1}=\mathrm{Prox}_{\alpha f}(x_k)$

得 $x_{k+1}=\left(A^TA+\frac{1}{\alpha}I\right)^{-1}\left(\frac{x_k}{\alpha}+A^Tb\right)$

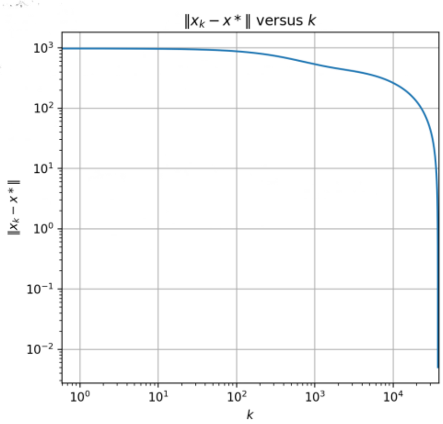
$$=\left(A^TA+\frac{1}{\alpha}I\right)^{-1}\left(\frac{x_k}{\alpha}+A^Tb+A^TAx_k-A^TAx_k\right)$$

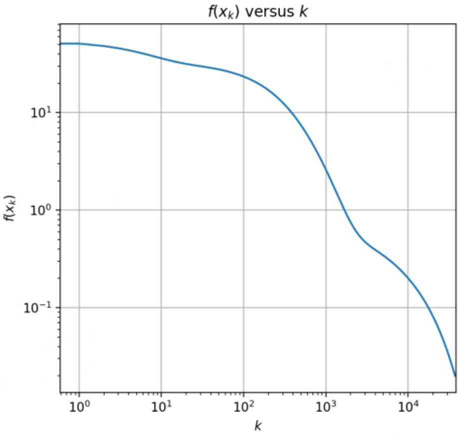
$$=x_k+\left(A^TA+\frac{1}{\alpha}I\right)^{-1}A^T(b-Ax_k)$$

(b). 选取 $\alpha=16.0$

初值 $x_0=[0, 0, ..., 0]^T$

图像如下





$$3.(a).\ \partial h(x) = \begin{cases} a_i^T(Ax-b)+1 & \text{if. } x_i>0 \\ [a_i^T(Ax-b)-1,\ a_i^T(Ax-b)+1] & \text{if. } x_i=0 \\ a_i^T(Ax-b)-1 & \text{if. } x_i<0 \end{cases}$$

其中 a_i 组成A的列. $A=[a_1,\cdots,a_n]$

$x_{k+1}=x_k-\alpha_k g_k$. 设 $x_{k,i}=0$ 时 取 $g_{k,i}=a_i^T(Ax_k-b)\in\partial h(x_k)$.

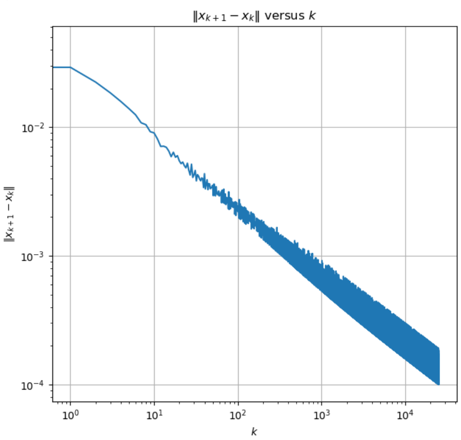
(b). $f(y)-f(x)-\nabla f(x)^T(y-x)=\frac{1}{2}(Ay-b)^T(Ay-b)-\frac{1}{2}(Ax-b)^T(Ax-b)-[A^T(Ax-b)]^T(y-x)$

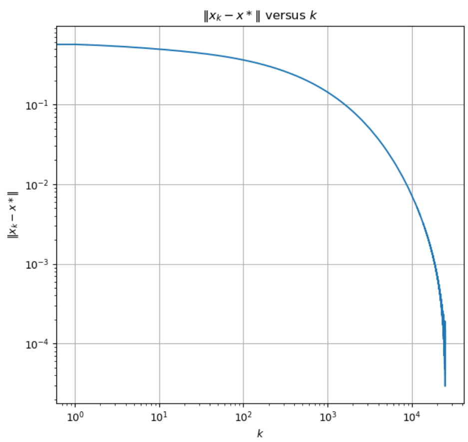
$=\frac{1}{2}(y-x)^TA^TA(y-x)$

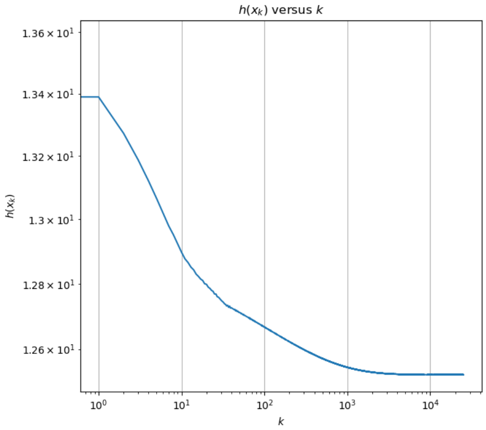
$$m\le\frac{\frac{1}{2}(y-x)^TA^TA(y-x)}{\frac{1}{2}(y-x)^T(y-x)}=\frac{\|A(y-x)\|_2^2}{\|y-x\|_2^2}=\left(\frac{\|Az\|_2}{\|z\|_2}\right)^2$$

即A最大奇异值的平方=0.756945502784552

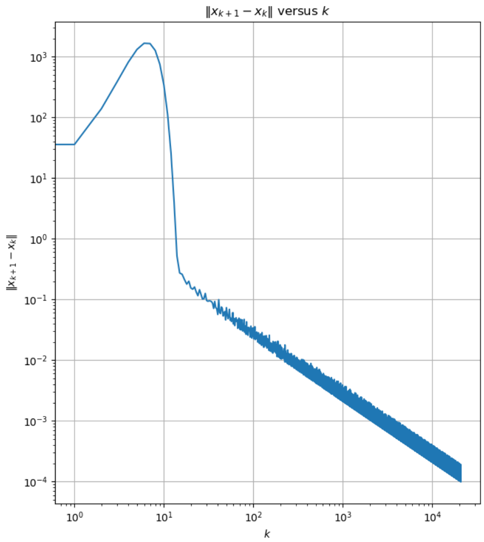
(a) 图像

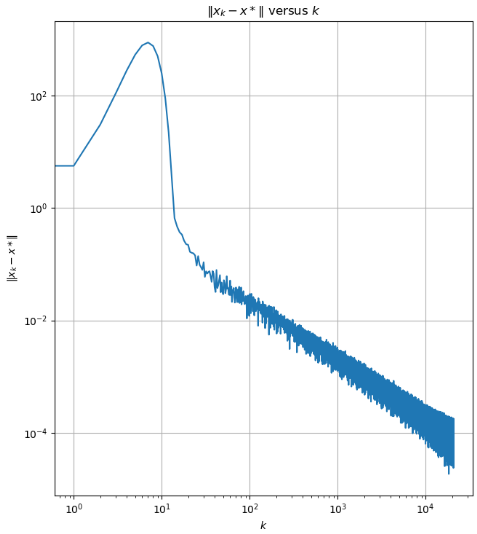


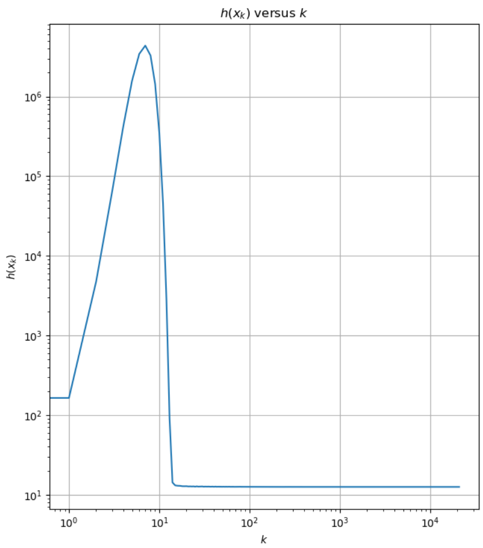



$h(x_k)$ versus k
$h(x_k)$
k

(b) 图像


$\|x_{k+1} - x_k\|$ versus k
$\|x_{k+1} - x_k\|$
k


$\|x_k - x^*\|$ versus k
$\|x_k - x^*\|$
k


$h(x_k)$ versus k
$h(x_k)$
k