6.4 $f(x) = \max_{1\le i\le k} x_i + \frac{1}{2}\|x\|^2$

(a)

记 $r = \max_{1\le i\le k} x_i$

若 $r \geq 0$, $f(x) \geq r + \frac{1}{2}r^2 \geq 0$

若 $r < 0$ $f(x) \geq r + \frac{k}{2}r^2 \geq -\frac{1}{2k}$, 当且仅当 $r = -\frac{1}{k}$, $x_1 = \cdots = x_k = -\frac{1}{k}$ 取等

故最优解 $x^* = -\frac{1}{k}$ $f^* = -\frac{1}{2k}$

(b) 设 $\|x^1\| \le \frac{1}{\sqrt{k}}$, $\|x^2\| \le \frac{1}{\sqrt{k}}$

$|f(x^1) - f(x^2)| = \left| \max_{1\le i\le k} x_i^1 - \max_{1\le i\le k} x_i^2 + \frac{1}{2}\|x^1\|^2 - \frac{1}{2}\|x^2\|^2 \right|$

$\le \frac{1}{2}\left|\|x^1\|^2 - \|x^2\|^2\right| + \left|\max_{1\le i\le k} x_i^1 - \max_{1\le i\le k} x_i^2\right|$ (因为 $\left\|\nabla_x \frac{1}{2}\|x\|^2\right\| = \|x\| \le \frac{1}{\sqrt{k}}$)

$\le \frac{1}{\sqrt{k}}\|x^1 - x^2\| + \left|\max_{1\le i\le k} x_i^1 - \max_{1\le i\le k} x_i^2\right|$

$\le \left(\frac{1}{\sqrt{k}} + 1\right)\|x^1 - x^2\|$ (记 $x^1 = (a_1 \cdots a_n)^T$ $x^2 = (b_1 \cdots b_n)^T$, $i_0 = \arg\max_i a_i$, $j_0 = \arg\max_j b_j$. $\left|\max_{1\le i\le k} x_i^1 - \max_{1\le j\le k} x_i^2\right| = |a_{i_0} - b_{j_0}|$ 若 $a_{i_0} \geq b_{j_0}$, $|a_{i_0} - b_{j_0}| = a_{i_0} - b_{j_0}$ $\le a_{i_0} - b_{i_0} \le \|x^1 - x^2\|_2$. 若 $a_{i_0} < b_{j_0}$, $|a_{i_0} - b_{j_0}| = b_{j_0} - a_{i_0} \le b_{j_0} - a_{j_0} \le \|x^1 - x^2\|_2$.)

#

(c) $x^0 = 0$, $x^{k+1} = x^k - \alpha_k g^k$, $g^k = x^k + e_j$. j 为任 $x_j^k = \max_{1\le i\le k} x_i^k$ 成立的最小整数. 则 $\forall y, x \in \mathbb{R}^n$, $f(y) - f(x) = \max_j y_j - \max_i x_i$

$:= \max_j y_j - x_{i_0} \geq y_{i_0} - x_{i_0} = g^{kT}(y - x)$ 故 $g^k \in \partial f(x^k)$

我们断言 当 $k < K$ 时, $\hat{f}^k - f^* \geq \frac{GR}{2(1+\sqrt{K})}$

证明: 这等价于 $\min_{0\le i\le k} f(x^i) \geq -\frac{1}{2k} + \frac{1}{2k}$ $i=0$ 显然成立.

$\Longleftrightarrow f(x^i) \geq 0$, $0 \le i \le k$

由数学归纳法可得

$x^k = (-(1-\alpha_{k-1})\cdots(1-\alpha_1)\alpha_0, -(1-\alpha_{k-1})\cdots\alpha_1, \cdots -\alpha_{k-1}, 0, \cdots 0)^T$

$\forall 1 \le k < K$ 我们让 $0 < \alpha_i < 1$, $\forall i$

有 $\min_{1\le i\le k} x_i^k = 0$ $f(x^k) = \frac{1}{2}\|x^k\|^2 \geq 0$ 证毕!

6.5 $\min f(x) = \frac{1}{2}\|Ax-b\|_2^2 + \mu\|x\|_2$, $A \in R^{m\times n}$

(a) $A^TA = I$ 该问题凸，最优值存在

由复合优化一阶必要条件有

$$\nabla_x \frac{1}{2}\|Ax-b\|_2^2 = -\mu\, \partial\|x\|_2 \quad 当\ x = x^*$$

$$x^* - A^Tb = -\mu\, \partial\|x\|_2$$

当 $x^* \neq 0$ 这意味着

$$x^* - A^Tb = -\mu \frac{x^*}{\|x^*\|_2}$$

$$\Rightarrow x^* = \frac{A^Tb}{\|A^Tb\|_2}(\|A^Tb\|_2 - \mu)$$ （设 $\|x^*\|_2 = t$，有 $t(1+\frac{\mu}{t}) = \|A^Tb\|_2$ 即得 $t = \|A^Tb\|_2 - \mu$，代回即得 x^*）

$$= A^Tb\left(1 - \frac{\mu}{\|A^Tb\|_2}\right)$$

$$f(0) = \frac{1}{2}b^Tb$$

$$f(x^*) = \mu\left|\|A^Tb\|_2 - \mu\right| + \frac{1}{2}\left\|AA^Tb\left(1-\frac{\mu}{\|A^Tb\|_2}\right) - b\right\|_2^2$$

$$= \mu\left|\|A^Tb\|_2 - \mu\right| + \frac{1}{2}(A^Tb)^TA^Tb\left(1-\frac{\mu}{\|A^Tb\|_2}\right)^2 - (A^Tb)^TA^Tb\left(1-\frac{\mu}{\|A^Tb\|_2}\right) + \frac{1}{2}b^Tb$$

$$= \mu\left|\|A^Tb\|_2 - \mu\right| - \frac{1}{2}(A^Tb)^TA^Tb\left(1-\frac{\mu}{\|A^Tb\|_2}\right)\left(1+\frac{\mu}{\|A^Tb\|_2}\right) + \frac{1}{2}b^Tb$$

当 $\|A^Tb\|_2 < \mu$ 时：　$f(x^*) > f(0)$

当 $\|A^Tb\|_2 = \mu$ 时　$f(x^*) = f(0)$

当 $\|A^Tb\|_2 > \mu$ 时　$f(x^*) - f(0) = \mu\|A^Tb\|_2 - \frac{1}{2}\|A^Tb\|_2^2 - \frac{1}{2}\mu^2 < 0$

故最优解为 $\begin{cases} x^* & \|A^Tb\|_2 \geq \mu \\ 0 & \|A^Tb\|_2 < \mu \end{cases}$

(b) $x \neq 0$ 时,

$$\nabla_x\left(\frac{1}{2}\|Ax-b\|_2^2+\mu\|x\|_2\right) = A^TAx - A^Tb + \mu\frac{x}{\|x\|_2}$$

Algorithm: hyperparameter x_0, α, N, ε_1, ε_2

Compute $f(0) = c$

for $k=1$ to N, do:

$\quad x_k = x_{k-1} - \alpha\left(A^TAx_{k-1} - A^Tb + \mu\frac{x_{k-1}}{\|x_{k-1}\|_2}\right)$

$\quad$ if $\|\nabla f(x_k)\|_2 < \varepsilon_1$ or $|f(x_k) - f(x_{k-1})| < \varepsilon_2$:

$\qquad$ if $f(x_k) < c$ return x_k else return 0

$\qquad$ end if

end for

find $i_0 = \min_{1\le i\le N} f(x_i)$

if $f(i_0) < c$ return x_{i_0} else return 0

这个算法收敛(由6.2节理论),可以得出最优解 x^*.