1. $f(x,y)=\max(x,y)\in R^2$. 定义在凸集上.

(a).

任取两点 (x_1,y_1), (x_2,y_2), λ.

证明凸函数条件

$$\lambda f(x_1,y_1)+(1-\lambda)f(x_2,y_2)\geq f(\lambda x_1+(1-\lambda)x_2,\ \lambda y_1+(1-\lambda)y_2)$$

即

$$\lambda\max(x_1,y_1)+(1-\lambda)\max(x_2,y_2)\geq\max(\lambda x_1+(1-\lambda)x_2,\ \lambda y_1+(1-\lambda)y_2)$$

满足 $\lambda\max(x_1,y_1)\geq\lambda x_1$ 且 $\lambda\max(x_1,y_1)\geq\lambda y_1$

$(1-\lambda)\max(x_2,y_2)\geq(1-\lambda)x_2$ 且 $(1-\lambda)\max(x_2,y_2)\geq(1-\lambda)y_2$

因此 $\lambda\max(x_1,y_1)+(1-\lambda)\max(x_2,y_2)\geq\max(\lambda x_1+(1-\lambda)x_2,\ \lambda y_1+(1-\lambda)y_2)$

$f(x,y)=\max(x,y)$ 为凸函数

(b). 求 $f(x,y)=\ln(e^x+e^y)$ 的 Hesse 矩阵.

$$\nabla f(x,y)=\left(\frac{e^x}{e^x+e^y},\ \frac{e^y}{e^x+e^y}\right)$$

$$\nabla^2 f(x,y)=\begin{pmatrix}\frac{e^{x+y}}{(e^x+e^y)^2} & \frac{-e^{x+y}}{(e^x+e^y)^2}\\ \frac{-e^{x+y}}{(e^x+e^y)^2} & \frac{e^{x+y}}{(e^x+e^y)^2}\end{pmatrix}=\frac{e^{x+y}}{(e^x+e^y)^2}\begin{pmatrix}1 & -1\\ -1 & 1\end{pmatrix}$$

$|\nabla^2 f(x,y)|=0$, 故 Hesse 矩阵半正定. $f(x,y)=\ln(e^x+e^y)$ 为凸函数

2. 显然二次函数有唯一极大值点

0.618 法:

① $a_0=0$ $b_0=25$ $t_1=15.45$ $t_1'=9.55$

$f(t_1)=-381.39$. $f(t_1')=-66.33$ $f(t_1)<f(t_1')$

② $a_1=0$ $b_1=15.45$ $t_2=9.55$ $t_2'=5.90$

$f(t_2)=-66.33$ $f(t_2')=23.98$ $f(t_2)<f(t_2')$

③ $a_2=0$ $b_2=9.55$ $t_3=5.90$ $t_3'=3.65$

$f(t_3)=23.98$ $f(t_3')=39.87$ $f(t_3)<f(t_3')$

④ $a_3=0$ $b_3=5.90$ $t_4=3.65$ $t_4'=2.25$

$f(t_4)=39.87$ $f(t_4')=34.41$ $f(t_4)>f(t_4')$

⑤ $a_4=2.25$ $b_4=5.90$ $t_5=4.51$ $t_5'=3.65$

$f(t_5)=37.40$ $f(t_5')=39.87$ $f(t_5)<f(t_5')$

⑥ $a_5=2.25$ $b_5=4.51$ $t_6=3.65$ $t_6'=3.11$

$f(t_6)=39.87$ $f(t_6')=39.16$ $f(t_6)>f(t_6')$

⑦ $a_6=3.11$. $b_6=4.51$

此时，已满足区间长度 $2<25\times8\%=2$

故所得极大值点 $t=\frac{a_6+b_6}{2}=3.81$. 极大值 $f(t)=39.75$

斐波那契法：

由于 $F_6=13$ $F_5=8$. 且 $F_n\geq\frac{1}{8\%}=12.5$. 故 $n\geq6$

① $a_0=0$ $b_0=25$ $t_1=15.38$. $t_1'=9.62$.

$f(t_1)=-376.43$ $f(t_1')=-68.84$ $f(t_1)<f(t_1')$

② $a_1=0$ $b_1=15.38$ $t_2=9.62$ $t_2'=5.77$

$f(t_2)=-68.84$ $f(t_2')=25.75$ $f(t_2)<f(t_2')$

③ $a_2=0$ $b_2=9.62$ $t_3=5.77$ $t_3'=3.85$

$f(t_3)=25.75$. $f(t_3')=39.69$ $f(t_3)<f(t_3')$

④ $a_3=0$ $b_3=5.77$ $t_4=3.85$ $t_4'=1.92$.

$f(t_4)=39.69$ $f(t_4')=31.41$ $f(t_4)>f(t_4')$

⑤ $a_4=1.92$ $b_4=5.77$ $t_5=3.88$ $t_5'=3.85$

其中，由于 $F_1=1$. $F_2=2$ 故 t_5 的计算公式为 $t_5=a_4+(\frac{1}{2}+\varepsilon)(b_4-a_4)$

取 $\varepsilon=0.01$

$f(t_5)=39.64$. $f(t_5')=39.69$ $f(t_5)<f(t_5')$

⑥ $a_5=1.92$ $b_5=3.88$

此时，已满足区间长度 $2<25\times8\%=2$

故所得极大值点 $t_5'=3.85$. $f(t_5')=39.69$

3. $\nabla f=(\frac{\partial f}{\partial x},\frac{\partial f}{\partial y})=(\frac{e^x}{e^x+e^y}+x,\ \frac{e^x}{e^x+e^y}+y)$

$$\nabla^2 f=\begin{pmatrix}\frac{\partial^2 f}{\partial x^2} & \frac{\partial^2 f}{\partial x\partial y}\\ \frac{\partial^2 f}{\partial y\partial x} & \frac{\partial^2 f}{\partial y^2}\end{pmatrix}=\begin{pmatrix}\frac{e^{x+y}}{(e^x+e^y)^2}+1, & \frac{-e^{x+y}}{(e^x+e^y)^2}\\ \frac{-e^{x+y}}{(e^x+e^y)^2} & \frac{e^{x+y}}{(e^x+e^y)^2}+1\end{pmatrix}$$

代入，得

$\nabla f(1,1)=(\frac{3}{2},\frac{3}{2})$

$$\nabla^2 f(1,1)=\begin{pmatrix}\frac{5}{4} & -\frac{1}{4}\\ -\frac{1}{4} & \frac{5}{4}\end{pmatrix}$$

牛顿方向 $D=-(\nabla^2 f(1,1))^{-1}\nabla f(1,1)=-\begin{pmatrix}\frac{3}{2}\\ \frac{3}{2}\end{pmatrix}$

l_1 范数下的最速下降方向

$\nabla f(1,1)=(\frac{3}{2},\frac{3}{2})$. $\left|\frac{\partial f(x)}{\partial x_i}\right|=\|\nabla f(x)\|_\infty$

考虑 l_1 范数，有 $\hat{D}=(-\lambda,\lambda-1)$，$\lambda\in[0,1]$

4. 负梯度法：

$\nabla f=(4x-4+2y,\ 4y-6+2x)$

第一次迭代：$\nabla f(1,1)=\begin{pmatrix}2\\0\end{pmatrix}$ 故 $D=-\nabla f(1,1)=\begin{pmatrix}-2\\0\end{pmatrix}$

令 $\nabla^T f(\hat{x}+t\hat{D})\nabla f(\hat{x})=0$

即 $(4(1-2t)-4+2,\ 0)\begin{pmatrix}2\\0\end{pmatrix}=0$ 得 $t=\frac{1}{4}$

更新后的点为 $(\frac{1}{2},1)$

第二次迭代 $\nabla f(\frac{1}{2},1)=\begin{pmatrix}0\\-1\end{pmatrix}$ 故 $D=-\nabla f(\frac{1}{2},1)=\begin{pmatrix}0\\1\end{pmatrix}$

令 $\nabla^T f(\hat{x}+t\hat{D})\nabla f(\hat{x})=0$

即 $4(1+t)-6+1=0$ 得 $t=\frac{1}{4}$.

更新后的点为 $(\frac{1}{2},\frac{5}{4})$

$f(\frac{1}{2},\frac{5}{4})=-\frac{37}{8}$

牛顿法：

$\nabla f=(4x-4+2y,\ 4y-6+2x)$

$\nabla^2 f=\begin{pmatrix}4&2\\2&4\end{pmatrix}$ $(\nabla^2 f)^{-1}=\begin{pmatrix}\frac{1}{3}&-\frac{1}{6}\\-\frac{1}{6}&\frac{1}{3}\end{pmatrix}$

第一次迭代：$\nabla f(1,1)=\begin{pmatrix}2\\0\end{pmatrix}$

$D=-(\nabla^2 f)^{-1}\nabla f(1,1)=\begin{pmatrix}-\frac{2}{3}\\\frac{1}{3}\end{pmatrix}$

令 $\nabla f(\hat{x}+tD)\cdot D=0$. 即 $4(1-\frac{2}{3}t)-4+2(1+\frac{1}{3})t=0$. 解得 $t=1$

更新后的点为 $\left(\frac{1}{3}, \frac{4}{3}\right)$

第二次迭代，$\nabla f\left(\frac{1}{3}, \frac{4}{3}\right) = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$. $D = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$.

已经为最优解. $t=0$.

故最终更新后的点为 $\left(\frac{1}{3}, \frac{4}{3}\right)$. $f\left(\frac{1}{3}, \frac{4}{3}\right) = -\frac{14}{3}$