1、

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

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

在 $\hat{X} = (1, 1)$ 处有：

$$\nabla f(\hat{X}) = \left(\frac{3}{2} \ \ \frac{3}{2} \right)^T$$

$$\nabla^2 f(\hat{X}) = \begin{bmatrix} \frac{5}{4} & -\frac{1}{4} \\ -\frac{1}{4} & \frac{5}{4} \end{bmatrix}$$

(1)

牛顿方向 $D = -\nabla^2 f(\hat{X})^{-1} \nabla f(\hat{X})$

$$= -\begin{bmatrix} \frac{5}{6} & \frac{1}{6} \\ \frac{1}{6} & \frac{5}{6} \end{bmatrix} \begin{bmatrix} \frac{3}{2} \\ \frac{3}{2} \end{bmatrix}$$

$$= \left(-\frac{3}{2}, -\frac{3}{2} \right)^T$$

= 两梯度值相等，所以↓

(2)

l_1范数下的最速下降方向为 $(-1, 0)^T$ 或 $(0, -1)^T$

2.

$$\nabla f(\hat{x}) = (4x-4+2y,\ 4y-6+2x)^T$$

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

负梯度法：

1° $X_1 = (1, 1)$, $f(X_1) = -4$

$D_1 = -\nabla f(X_1) = (-2, 0)^T$，记 $X_2 = (1-2t_1, 1)^T$

$\therefore \nabla^T f(\hat{X_2}) D_1 = (2-8t_1, -4t_1) \cdot \begin{pmatrix} -2 \\ 0 \end{pmatrix} = 0$

$\Rightarrow t_1 = \frac{1}{4}$

2° $X_2 = (\frac{1}{2}, 1)^T$, $f(X_2) = -\frac{9}{2}$

$D_2 = -\nabla f(X_2) = (0, 1)^T$，记 $X_3 = (\frac{1}{2}, 1+t_2)^T$

$\therefore \nabla^T f(\hat{X_3}) \cdot D_2 = (2t_2, -1+4t_2) \cdot \begin{pmatrix} 0 \\ 1 \end{pmatrix} = 0$

$\Rightarrow t_2 = \frac{1}{4}$

3° $\underline{X_3 = (\frac{1}{2}, \frac{5}{4})^T}$, $\underline{f(X_3) = -4.625}$

牛顿法：

1° $X_1=(1,1)^T$, $f(X_1)=-4$

$$D=-(\nabla^2 f(X_1))^{-1}\nabla f(X_1)=(-\frac{2}{3},\frac{1}{3})^T$$

记 $X_2=(1-\frac{2}{3}t_1,\ 1+\frac{1}{3}t_1)^T$

则 $f(X_3)=2(1-\frac{2}{3}t_1)^2+2(1+\frac{1}{3}t_1)^2-4(1-\frac{2}{3}t_1)$

$\qquad -6(1+\frac{1}{3}t_1)+2(1-\frac{2}{3}t_1)(1+\frac{1}{3}t_1)$

$$\frac{df(X_3)}{dt_1}=\frac{4}{3}t_1-\frac{4}{3}=0 \Rightarrow t_1=1$$

2° $X_2=(\frac{1}{3},\frac{4}{3})^T$, $f(X_2)=-\frac{14}{3}=-4.67$

此时 $Df(X_2)=(0,0)^T$ ，已达到最优、停止迭代

最优解：$\underline{X^*=(\frac{1}{3},\frac{4}{3})^T}$，此时 $f(X^*)=\underline{-\frac{14}{3}=-4.67}$