HOMEWORK 10

QUESTION 1

(a.)

$$
\begin{aligned}
\min_{\boldsymbol{x}} \quad & F(\boldsymbol{x}) = \boldsymbol{c}^T\boldsymbol{x} + \tfrac{1}{t}B(\boldsymbol{x}) \\
\text{s.t.} \quad & \boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}
\end{aligned} \tag{1}
$$

Where $B(\boldsymbol{x}) = -\sum_{j=1}^{n} log(x_j)$, x_j is the j^{th} component of $\boldsymbol{x}$.

(b.)

The gradient is:

$$
\begin{aligned}
\nabla F(\boldsymbol{x}) &= \boldsymbol{c} + \frac{1}{t}\nabla B(\boldsymbol{x}) \\
&= \boldsymbol{c} + \sum_{j=1}^{n} \frac{1}{t} \cdot (-\frac{1}{x_j}) \cdot (0, 0, 0, \cdots, 1, \cdots, 0, 0, 0)^T
\end{aligned} \tag{2}
$$

The Hessian matrix is:

$$
\nabla^2 F(\boldsymbol{x}) = \sum_{j=1}^{n} \frac{1}{x_j^2} \nabla g_j(\boldsymbol{x}) \nabla g_j(\boldsymbol{x})^T \tag{3}
$$

Where $g_j(\boldsymbol{x}) = -x_j, \nabla g_j(\boldsymbol{x}) = (0, 0, 0, \cdots, -1, \cdots, 0, 0, 0)^T$

(d.)

By adding the slack

Variables, we could convert the original LP into the standard form:

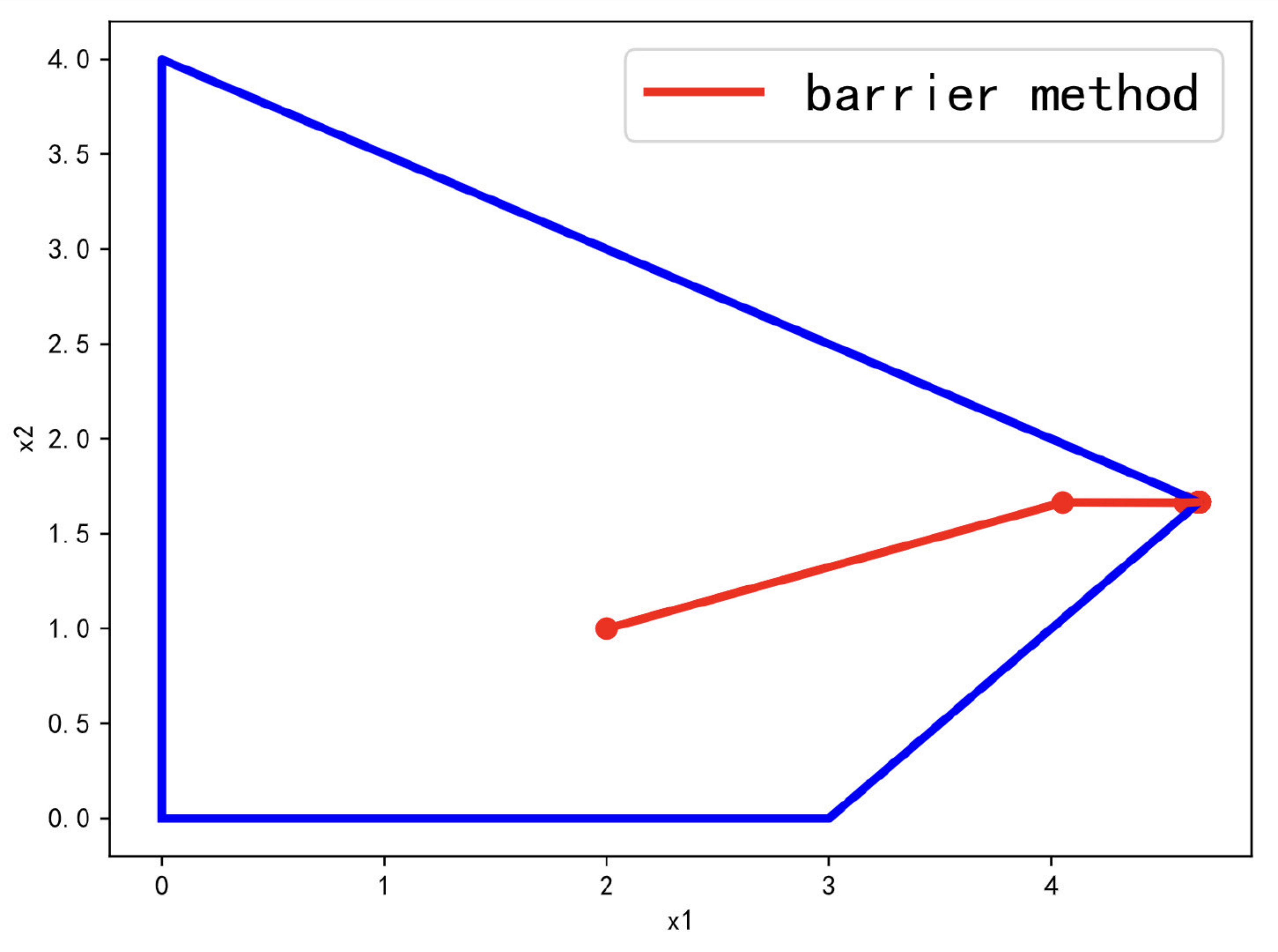
$$
\begin{aligned}
\min_{\boldsymbol{x} \in \mathbb{R}^n} \quad & \boldsymbol{c}^T\boldsymbol{x} \\
\text{s.t.} \quad & \boldsymbol{A}\boldsymbol{x} = \boldsymbol{b} \\
& \boldsymbol{x} \geq \boldsymbol{0}
\end{aligned} \tag{4}
$$

Where $\boldsymbol{c} = (-3, -1, 0, 0)^T, \boldsymbol{x} = (x_1, x_2, s_1, s_2)^T, \boldsymbol{b} = (8, 3)^T$

$$
\boldsymbol{A} = \begin{pmatrix} 1 & 2 & 1 & 0 \\ 1 & -1 & 0 & 1 \end{pmatrix} \tag{5}
$$

Output

```
iteration 0: [2 1 4 2]
iteration 1: [4.05162269 1.66475112 0.61887506 0.61312843]
iteration 2: [4.60203413 1.66223521 0.07349545 0.06020108]
iteration 3: [4.66016983 1.66617239 0.00748539 0.00600256]
iteration 4: [4.66601116e+00 1.66661630e+00 7.56240149e-04 6.05137244e-04]
iteration 5: [4.66656565e+00 1.66665890e+00 1.16557159e-04 9.32491734e-05]
iteration 6: [4.66660301e+00 1.66666177e+00 7.34522347e-05 5.87631561e-05]
iteration 7: [4.66660602e+00 1.66666200e+00 6.99796075e-05 5.59849280e-05]
iteration 8: [4.66660632e+00 1.66666202e+00 6.96355818e-05 5.57096953e-05]
iteration 9: [4.66660634e+00 1.66666203e+00 6.96032699e-05 5.56838446e-05]
[-3 -1  0  0]
[4.66660634e+00 1.66666203e+00 6.96032699e-05 5.56838446e-05]
optimal value: -15.666481055899224
```



QUESTION 2

(a.)

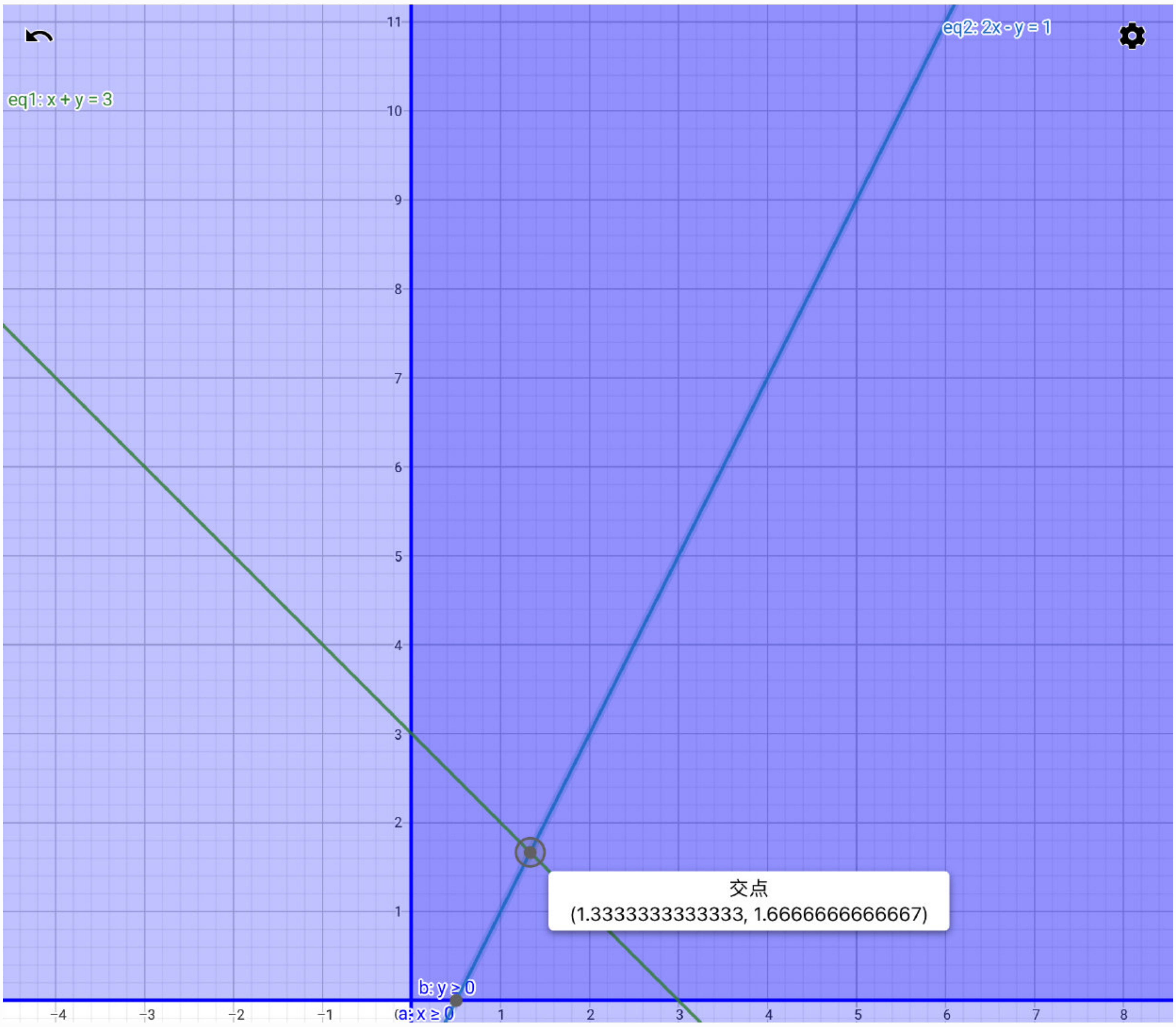
$$\begin{aligned} \max_{\boldsymbol{\mu} \in \mathbb{R}^4} \quad & -8\mu_1 - 3\mu_2 \\ \text{s.t.} \quad & \mu_1 + \mu_2 - \mu_3 = 3 \\ & 2\mu_1 - \mu_2 - \mu_4 = 1 \\ & \mu_1 \geq 0 \\ & \mu_2 \geq 0 \\ & \mu_3 \geq 0 \\ & \mu_4 \geq 0 \end{aligned} \tag{6}$$

(b.)

$$\begin{aligned} \max_{\boldsymbol{\mu} \in \mathbb{R}^2} \quad & -8\mu_1 - 3\mu_2 \\ \text{s.t.} \quad & -\mu_1 - \mu_2 \leq -3 \\ & -2\mu_1 + \mu_2 \leq -1 \\ & \mu_1 \geq 0 \\ & \mu_2 \geq 0 \end{aligned} \tag{7}$$

(c.)

optimal value in 1(d) is $v_1 = -15.66648106$.



From the graph we know the optimal point for (7) is $\boldsymbol{\mu} = (4/3, 5/3)^T$

The corresponding optimal value v_2 is $-8 \cdot 4 \div 3 - 5 = -15.666666 \approx v_1$

(d.)

```
iteration 0: [4. 1. 2. 6.]
iteration 1: [1.61583503 1.63097966 0.24681469 0.60069039]
iteration 2: [1.36062986 1.66109969 0.02172955 0.06016003]
iteration 3: [1.33604959 1.66609656 0.00214615 0.00600263]
iteration 4: [1.33360862e+00 1.66660873e+00 2.17352772e-04 6.08520299e-04]
iteration 5: [1.33338218e+00 1.66665638e+00 3.85604457e-05 1.07967126e-04]
iteration 6: [1.33336727e+00 1.66665952e+00 2.67957938e-05 7.50271977e-05]
iteration 7: [1.33336603e+00 1.66665978e+00 2.58097263e-05 7.22662827e-05]
iteration 8: [1.33336590e+00 1.66665981e+00 2.57137843e-05 7.19976522e-05]
iteration 9: [1.33336589e+00 1.66665981e+00 2.57014474e-05 7.19631098e-05]
dual optimal value: -15.666906545270626
```

Dual optimal solution $\boldsymbol{\mu} = (\mu_1, \mu_2, \mu_3, \mu_4)^T$=
$[1.33336589e+001.66665981e+002.57014474e-057.19631098e-05]$

Dual optimal value = -15.66690655

QUESTION 3

(a.)

Since $f(x) = log(2 + e^x)$ is monotonically increasing, thus the optimal solution is $x = 0$ and the optimal value $f^* = log(3)$

(b.)

The Lagrangian is

$$\mathcal{L}(x,\mu) = log(2+e^x) - \mu x \tag{8}$$

Considering that $\frac{\partial\mathcal{L}}{\partial x} = 1-\mu-\frac{2}{e^x+2} \in [-\mu, 1-\mu]$

Case 1: $\mu < 0$

$\because \frac{\partial\mathcal{L}}{\partial x} \ge 0$

$\therefore \phi(\mu) = -\infty$

Case 2: $\mu = 0$

$\because \frac{\partial\mathcal{L}}{\partial x} \ge 0$

$\therefore \phi(\mu) = log2$

Case 3: $0 < \mu < 1$

Let $\frac{\partial\mathcal{L}}{\partial x} = 0$, we have $x = log(\frac{2\mu}{1-\mu})$

Thus

$$\phi(\mu) = log(\frac{2}{1-\mu}) - \mu log(\frac{2\mu}{1-\mu}) \tag{9}$$

Case 4: $\mu = 1$

$\because \frac{\partial\mathcal{L}}{\partial x} \le 0 $

$\therefore \phi(\mu) = \lim_{x\to+\infty}\mathcal{L}(x,\mu) = 0$

Case 5: $\mu > 1$

$\because \frac{\partial\mathcal{L}}{\partial x} \le 0 $

$\therefore \phi(\mu) = \lim_{x\to+\infty}\mathcal{L}(x,\mu) = -\infty$

To sum up, the dual function is

$$\phi(\mu) = \begin{cases} -\infty & \mu < 0 \\ log2 & \mu = 0 \\ log(\frac{2}{1-\mu}) - \mu log(\frac{2\mu}{1-\mu}) & 0 < \mu < 1 \\ 0 & \mu = 1 \\ -\infty & \mu > 1 \end{cases} \tag{10}$$

The dual problem is

$$\begin{aligned} \max_{\mu} \quad & \phi(\mu) \\ \text{s.t.} \quad & \mu \ge 0 \end{aligned} \tag{11}$$

(c.)

When $0 < \mu < 1$, let $\phi'(\mu) = -log(\frac{2\mu}{1-\mu}) = 0 \longrightarrow \mu = \frac{1}{3}$. $\phi(\mu)$ is increasing in $(0, \frac{1}{3})$ and decreasing in $(\frac{1}{3}, 1)$

After comparing with other cases, we conclude that the dual optimal value $\phi^* = \phi(\frac{1}{3})$ = $log(3) = f^*$, **thus the strong duality holds.**

QUESTION 4

(a.)

The Lagrangian is

$$
\begin{aligned}
\mathcal{L}(\boldsymbol{x}, \boldsymbol{\lambda}, \boldsymbol{\mu}) &= f(\boldsymbol{x}) + \sum_{i=1}^{k} \lambda_i h_i(\boldsymbol{x}) + \sum_{j=1}^{m} \mu_j g_j(\boldsymbol{x}) \\
&= f(\boldsymbol{x}) + \sum_{j=1}^{2} \mu_j g_j(\boldsymbol{x})
\end{aligned} \tag{12}
$$

Where

$$
\begin{aligned}
f(\boldsymbol{x}) &= x_1^2 + x_2^2 \\
g_1(\boldsymbol{x}) &= (x_1 - 2)^2 + (x_2 - 1)^2 - 1 \\
g_2(\boldsymbol{x}) &= (x_1 - 2)^2 + (x_2 + 1)^2 - 1
\end{aligned} \tag{13}
$$

Let

$$
\nabla_{\boldsymbol{x}} \mathcal{L}(\boldsymbol{x}, \boldsymbol{\mu}) = 0 \tag{14}
$$

We get

$$
\begin{cases} 2x_1 + 2\mu_1(x_1 - 2) + 2\mu_2(x_1 - 2) = 0 \\ \\ 2x_2 + 2\mu_1(x_2 - 1) + 2\mu_2(x_2 + 1) = 0 \end{cases} \longrightarrow \begin{cases} \hat{x_1} = \frac{2\mu_1 + 2\mu_2}{1 + \mu_1 + \mu_2} \\ \\ \hat{x_2} = \frac{\mu_1 - \mu_2}{1 + \mu_1 + \mu_2} \end{cases} \tag{15}
$$

The dual function is

$$
\begin{aligned}
\phi(\boldsymbol{\mu}) &= \mathcal{L}(\hat{\boldsymbol{x}}, \boldsymbol{\mu}) \\
&= \frac{1}{(1 + \mu_1 + \mu_2)^2} (-\mu_1^3 - \mu_2^3 + \mu_1^2 \mu_2 + \mu_1 \mu_2^2 + +3\mu_1^2 + \mu_2^2 + 10\mu_1\mu_2 + 4\mu_1 + 4\mu_2)
\end{aligned} \tag{16}
$$

The dual problem is

$$
\begin{aligned}
\max_{\boldsymbol{\mu}} \quad & \phi(\boldsymbol{\mu}) = \frac{1}{(1 + \mu_1 + \mu_2)^2} (-\mu_1^3 - \mu_2^3 + \mu_1^2 \mu_2 + \mu_1 \mu_2^2 + +3\mu_1^2 + 3\mu_2^2 + 10\mu_1\mu_2 + 4\mu_1 + 4\mu_2) \\
\text{s.t.} \quad & \boldsymbol{\mu} \geq 0
\end{aligned} \tag{17}
$$

Considering the symmetry, WOLG we set $\mu_1 = \mu_2 = \mu'$ in the formula above, then we derive $\gamma(\mu') = \frac{8\mu'}{1+2\mu'}$, which takes 4 as its upper bound. Thus $\phi^* = 4$.

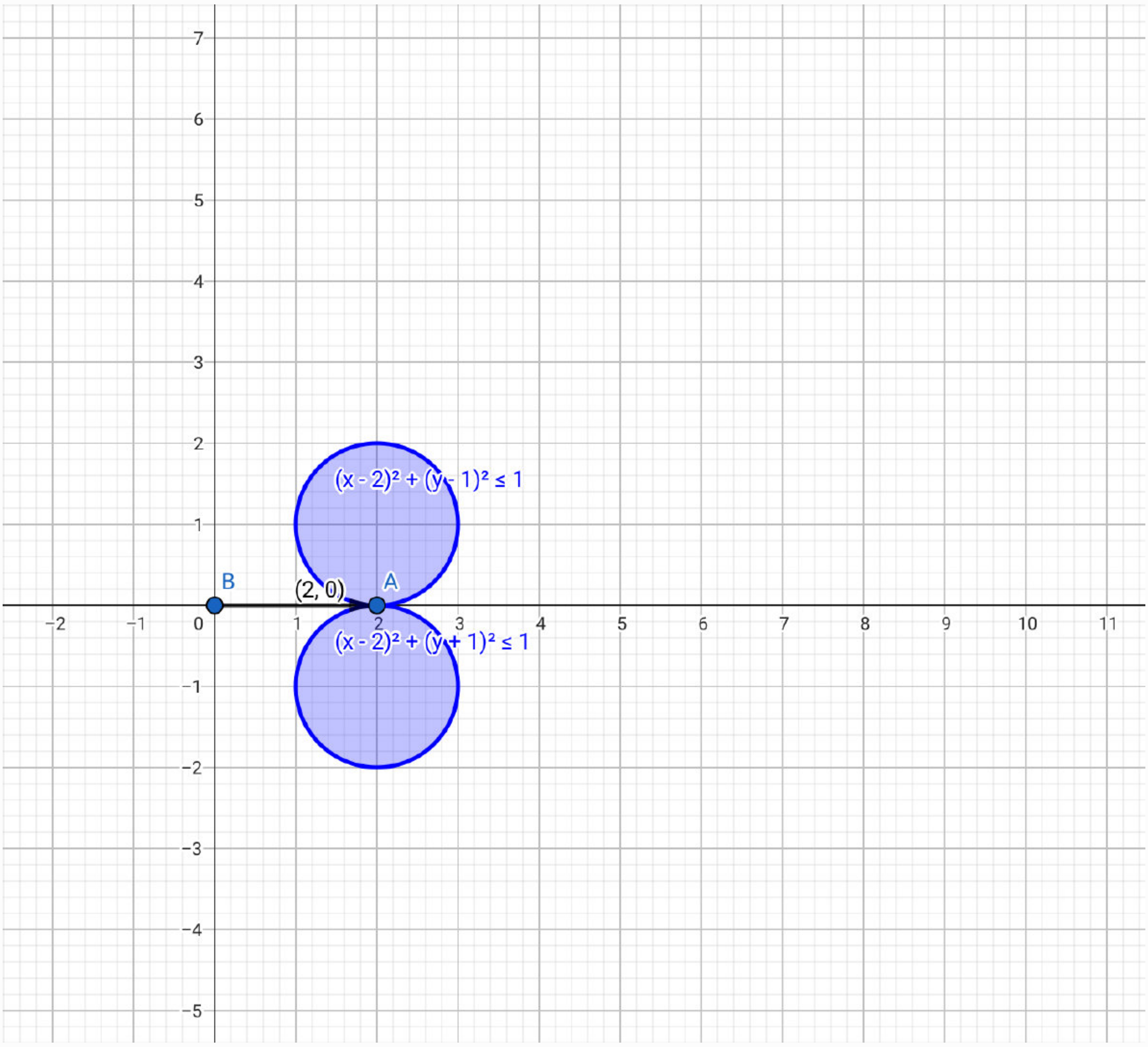
(b.)

According to (a.), $\phi^* = 4$

And by solving the original problem graphically, we know that $f^* = 4$

$\because f^* = 2^2 + 0^2 = \phi^*$

$\therefore$ **<u>strong duality holds.</u>**



(c.)

According to the graph above, we can see that there exists no point $\boldsymbol{x} \in int\ dom f \cap (\bigcap_{j=1}^{2} \operatorname{dom} g_j)$

s.t. $g_j(\boldsymbol{x}) < 0$ for $j = 1, 2$

$\therefore$ Slater's condition doesn't hold.

However, from (b.) we know that for this problem strong duality holds, which proves that **Slater's condition is sufficient but not necessary** for strong duality.

(d.)

The dual optimal value couldn't be obtained by any feasible point.

This shows us that KKT condition doesn't hold at the primal optimal solution.

Because we know that if KKT condition holds at $\boldsymbol{x}^*$ with Lagrange multipliers $\boldsymbol{\mu}^*$, then $\boldsymbol{\mu}^*$ should be the dual optimal solution, i.e. $\phi(\mu^*) = \phi^*$.

However, we have already figured out that the dual optimal value couldn't be obtained by any feasible point, which means $\phi(\mu^*) = \phi^*$ fails.

QUESTION 5

(a.)

$$\begin{aligned}\phi(\mu) &= \inf_{\boldsymbol{x}\in\{\boldsymbol{x}\in\mathbb{R}^2:\boldsymbol{x}\geq 0\}} x_1^5 + x_2^5 + \mu(1 - x_1 - x_2), \mu \geq 0 \\ &= 2(\frac{\mu}{5})^{\frac{5}{4}} + \mu(1 - 2(\frac{\mu}{5})^{\frac{1}{4}})\end{aligned} \tag{18}$$

(b.)

$$\phi'(\mu) = 1 - 2(\frac{\mu}{5})^{\frac{1}{4}} = 0 \quad \longrightarrow \mu^* = \frac{5}{16} \tag{19}$$

Thus $\phi^* = \phi(\mu^*) = \frac{1}{16}$

(c.)

Let $\boldsymbol{x_1} = (x_1, 0)^T, \boldsymbol{x_2} = (x_2, 0)^T$, $x_1 \geq 0, x_2 \geq 0$

Since $f(\boldsymbol{x})$ is convex on its domain, we have

$$\frac{f(\boldsymbol{x_1}) + f(\boldsymbol{x_2})}{2} \geq f(\frac{\boldsymbol{x_1} + \boldsymbol{x_2}}{2}) \tag{20}$$

which induces

$$x_1^5 + x_2^5 \geq \frac{(x_1 + x_2)^5}{16} \geq \frac{1}{16} \tag{21}$$

The equality could be attained when $x_1 = x_2 = \frac{1}{2}$

$\therefore$ the primal optimal value is $f^* = \frac{1}{16}$

(d.)

The dual function is

$$\phi(\boldsymbol{\mu}) = \inf_{\boldsymbol{x}\in\mathbb{R}^2}[x_1^5 + x_2^5 + \mu_1(1 - x_1 - x_2) - \mu_2 x_1 - \mu_3 x_2] \tag{22}$$

We make $x_1 \to -\infty, x_2 \to -\infty$, then $\phi(\boldsymbol{\mu}) \to -\infty$

Thus, the strong duality doesn't hold for P2.