Homework 13

1

(a).

It's easy to see that $f(x)$ is monotonically increasing, thus the optimal solution is $x^* = 0$ and the optimal value is f^* =log2.

(b).

The dual function is

$$\phi(\mu) = \inf_{x \in \mathbb{R}} [\log(1 + e^x) - \mu x]$$

Then we can easily see that when $\mu \notin [0, 1]$, $\phi(\mu)$ is unbounded below. When $\mu \in [0, 1]$,we can calculate the explicit expression of $\phi(\mu)$. So the dual function is

$$\phi(\mu) = \begin{cases} (\mu - 1)\log(1 - \mu) - \mu \log \mu, \mu \in [0, 1] \\ -\infty, \text{otherwise} \end{cases}$$

The dual problem is

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

(c).

We only need to consider $\mu \in [0, 1]$, where $\phi'(\mu) = \log(\frac{1-\mu}{\mu})$. We can easily see that ϕ' is monotonically decreasing and $\phi'(\frac{1}{2}) = 0$. So the dual optimal solution is $\mu^* = \frac{1}{2}$ and the dual optimal value is $\phi^* = \log 2$. The strong duality holds.

2

(a).

The Lagrange dual function is

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

We can easily see that when $\mu_1+\mu_2+1 \leq 0$, ϕ is unbounded below. When $\mu_1+\mu_2+1>0$, we can calculate the explicit expression of $\phi(\mu_1,\mu_2)$. So the dual function is

$$\phi(\mu_1,\mu_2)=\begin{cases}1-\frac{(\mu_1-\mu_2)^2+1}{\mu_1+\mu_2+1}, \mu_1+\mu_2+1>0\\ -\infty, \text{otherwise}\end{cases}$$

The dual problem is

$$\begin{aligned}\max_{\mu_1,\mu_2} \quad & \phi(\mu_1,\mu_2)\\ \text{s.t.} \quad & \mu_1,\mu_2 \geq 0\end{aligned}$$

(b).

We only need to consider the case where $\mu_1+\mu_2+1>0$. Its easy to see $\phi<1$ when $\mu_1+\mu_2+1>0$. Also note that ϕ goes to 1 as $\mu_1=\mu_2$ and go to positive infinity together. So the dual optimal value is $\phi^*=1$.
And we already know that the primal optimal value is $f^*=1$, so strong duality holds.

(c).

The Slater's condition does not hold as there is no point in $\text{int}D$ which is strictly feasible. Yet the strong duality still holds, which means that Slater's condition is not necessary for strong duality.

(d).

It is not attained by any dual feasible point because as we have shown in part (b), its optimal value is reached only when μ_1 and μ_2 both go to positive infinity. This is expected because in Problem 2(b) we have shown that the optimal point $\boldsymbol{x}^*$ is not a regular point.

3

(a).

Follow the expression given by this problem, we can derive the explicit expression of $\phi(\mu)$, which is

$$\phi(\mu)=\begin{cases}\mu, \quad \mu\leq 0\\ \mu-\frac{4}{3\sqrt{3}}\mu^{\frac{3}{2}}, \quad \mu>0\end{cases}$$

(b).

When $\mu > 0$, $\phi'(\mu) = 1 - \frac{2}{\sqrt{3}}\mu^{\frac{1}{2}}$. So we can easily see that $\phi(\mu)$ reaches its maximum at point $\mu^* = \frac{3}{4}$. The dual optimal value is $\phi^* = \frac{1}{4}$.

(c).

Firstly we can see that $f(\frac{1}{2}, \frac{1}{2}) = \frac{1}{4}$ where $\boldsymbol{x} = (\frac{1}{2}, \frac{1}{2})$ is feasible. Also note that weak duality always holds, meaning that $f^* \geq \phi^*$. Thus $f^* \geq \frac{1}{4}$. Then we can say that f^* is actually $\frac{1}{4}$ as it is reached by point $\boldsymbol{x} = (\frac{1}{2}, \frac{1}{2})$.

(d).

In this case we can easily see that the dual function

$$\phi(\boldsymbol{\mu}) = \inf_{\boldsymbol{x}\in\mathbb{R}}[x_1^3 + x_2^3 + \mu_1(1 - x_1 - x_2) - \mu_2 x_1 - \mu_3 x_2]$$

is unbounded from below. Thus the dual function is actually

$$\phi(\boldsymbol{\mu}) = -\infty$$

The strong duality does not hold.

4

(a).

Here we have $f(\boldsymbol{w}^*, b^*) = f^*$ and $\phi(\boldsymbol{\mu}^*) = \phi^*$. Also, note that here all inequality constraints are affine. So here feasibility means Slater's condition holds for the primal problem. Then by Slater's Theorem strong duality holds, i.e. $f^* = \phi^*$. Under these conditions, KKT conditions hold.
So by complementary slackness for any i with $\mu_i^* > 0$, $g_i(\boldsymbol{w}^*, b^*)$ must be zero. Which means $y_i(\boldsymbol{x}_i^T\boldsymbol{w}^* + b^*) = 1$. Note that $y_i^2 = 1$ for any i, so $b^* = y_i - \boldsymbol{x}_i^T\boldsymbol{w}^*$.

(b).

The output and the figure are given below.
primal optimal:

w = [-1.09090908 1.45454545]
b = [-0.09090911]

dual optimal:
mu = [1.65289255e+00 0.00000000e+00 0.00000000e+00 0.00000000e+00 0.00000000e+00 0.00000000e+00 0.00000000e+00 1.65289254e+00 0.00000000e+00 0.00000000e+00 7.11813924e-09 0.00000000e+00 0.00000000e+00]

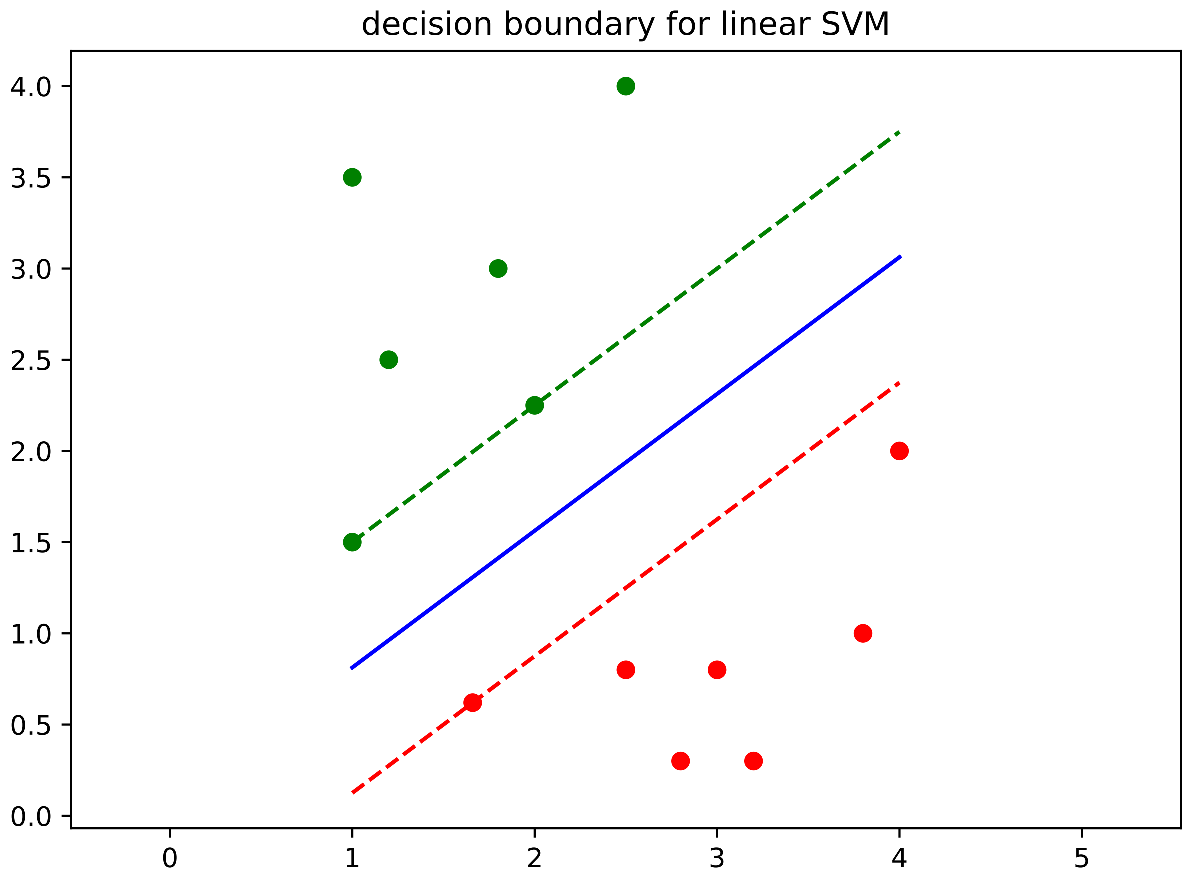

Figure 1: SVM result