Homework 1

Zihan Zhao

1001103708

1

(a)

First determine $E[Z]$

$$E[Z] = E[(X - Y)^2]$$
$$= E[(X^2 - 2XY + Y^2]$$
$$= E[X^2] - E[2XY] + E[Y^2]$$
$$= E[X^2]2E[X]E[Y] + E[Y^2]$$

Since X and Y both are variables from uniform distribution in $[0, 1]$, $E[x] = E[Y]$ and $E[X^2] = E[Y^2]$. Therefore,

$$continue:$$
$$E[Z] = 2E[X^2] + 2E[X]^2$$
$$= 2\int_0^1 x^2 f(x)dx - 2(\int_0^1 xf(x)dx)^2$$

Note that $f(x)$ is the pdf of X. So $f(x) = \frac{1}{1-0} = 1$. Therefore,

$$continue:$$
$$E[Z] = 2\int_0^1 x^2 dx - 2(\int_0^1 xdx)^2$$
$$= 2*\frac{1}{3} - 2*(\frac{1}{2})^2$$
$$= \frac{1}{6} = 0.16667$$

Second determine $Var[Z]$

$$Var[Z] = E[Z^2] - E[Z]^2$$
$$= E[(X - Y)^4] - (\frac{1}{6})^2$$
$$= E[X^4 - 4X^3Y + 6X^2Y^2 - 4XY^3 + Y^4] - (\frac{1}{6})^2$$
$$= E[X^4] - 4E[X^3]E[Y] + 6E[X^2]E[Y^2] - 4E[X]E[Y^3] + E[Y^4] - (\frac{1}{6})^2$$

Since X and Y both are variables from uniform distribution in $[0, 1]$, $E[x] = E[Y]$, $E[X^2] = E[Y^2]$, $E[X^3] = E[Y^3]$, and $E[X^4] = E[Y^4]$ Therefore,

$$Var[Z] = 2E[X^4] - 8E[X^3]E[X] + 6E[X^2]^2 - (\frac{1}{6})^2$$

$$= 2\int_0^1 x^4 f(x)dx - 8(\int_0^1 x^3 f(x)dx)(\int_0^1 xf(x)dx) + 6(\int_0^1 x^2 f(x)dx)^2 - (\frac{1}{6})^2$$

$$= 2 * \frac{1}{5} - 8 * \frac{1}{4} * \frac{1}{2} + 6(\frac{1}{3})^2 - (\frac{1}{6})^2 = \frac{7}{180} = 0.38889$$

(b)

First determine $E[R]$. Since $X_1, .., Y_d$ are i.i.d, $(X_1 - Y_1)^2, ..., (X_d - Y_d)^2$ are independent. Therefore,
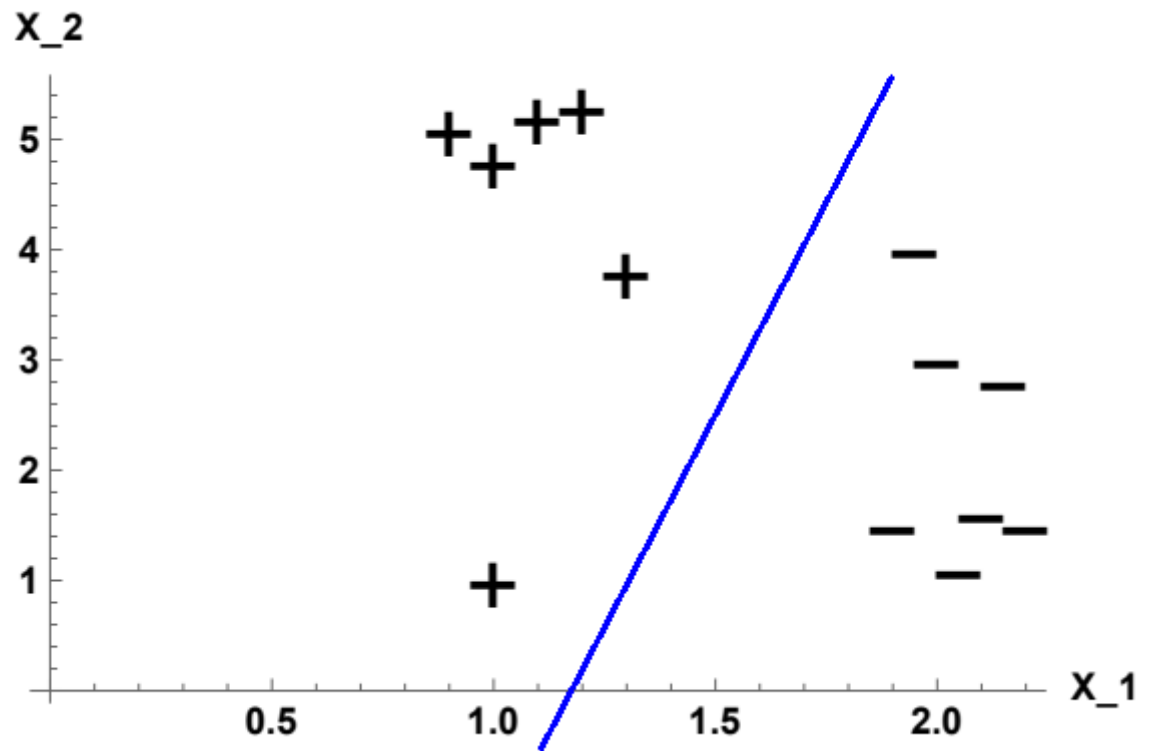
$$E[R] = E[\sum_{i=1}^d Z_i]$$

$$= \sum_{i=1}^d E[Z_i]$$

$$= dE[Z] = \frac{d}{6}$$

Second determine $Var[R]$. Since $Z_1, .., Z_d$ are independent, $Cov(Z_i, Z_j) = 0$ where i and j are arbitrary picks from 1 to d. Therefore,

$$Var[R] = Var[\sum_{i=1}^d Z_i]$$

$$= \sum_{i=1}^d Var[Z_i] + 2\sum_{i<j}^d Cov[Z_i, Z_j]$$

$$= dVar[Z] = \frac{7d}{180}$$
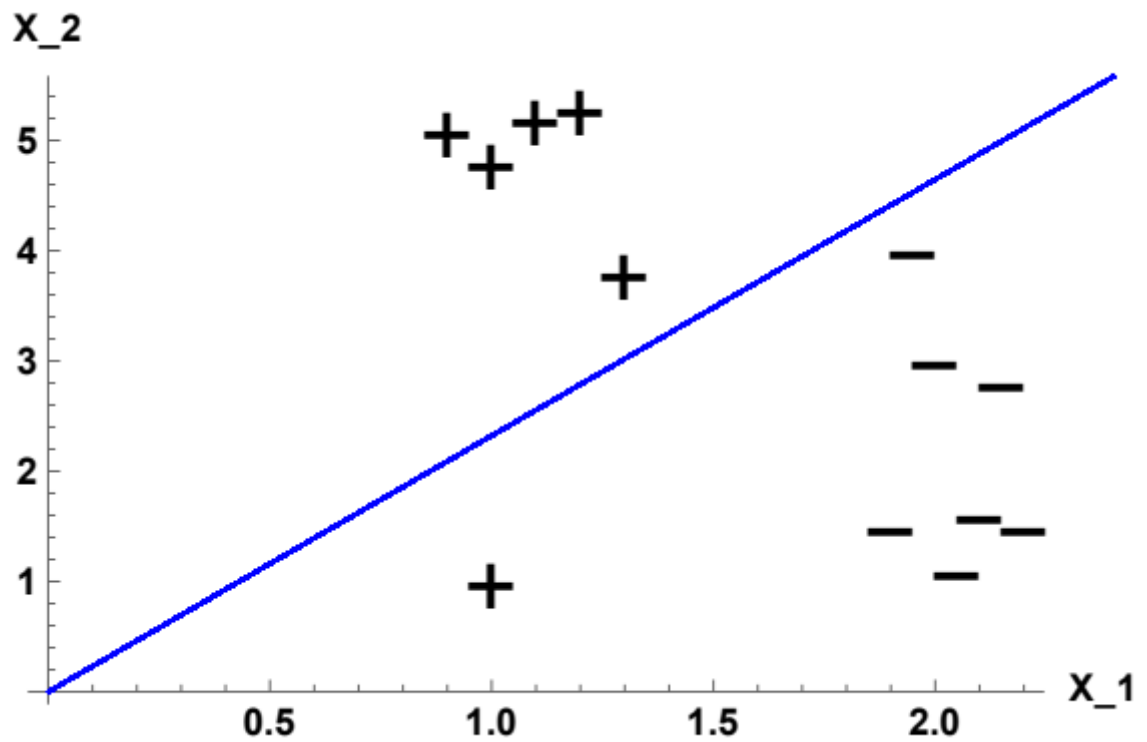
2

(a)

(b)

No, it is not unique.

(c)

0. The left side of the boundary marks "+" and the right side marks "-". All "+" points are in left side and all "-" points are in right side.

(d)

As λ approaches ∞, $J(w)$ heavily depends on $\lambda \omega_0^2$. In order to minimize $\lambda \omega_0^2$, ω_0 should be 0. So the boundary should pass through (0,0). The sketch is shown below.
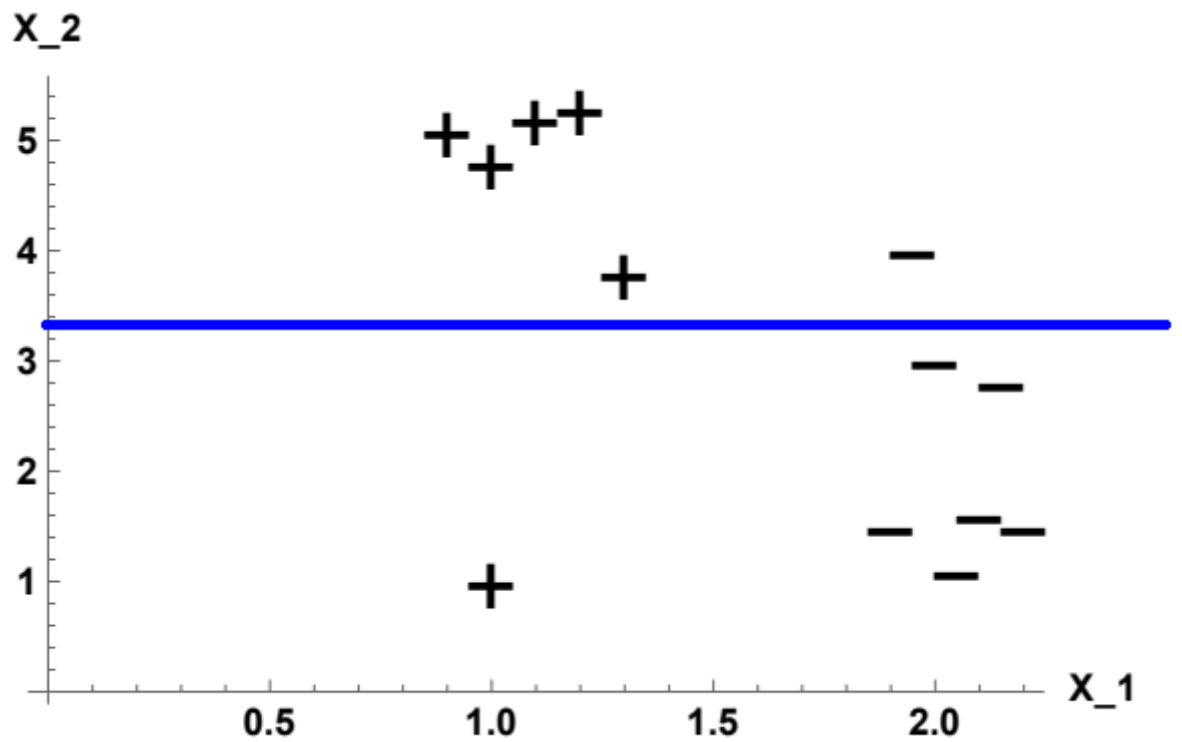
(e)

1. The upper side of the boundary marks "+" and the down side marks "-". All "+" points are in left side except one point is at down side. All "-" points are in right side. So the error is 1.

(f)

Similar with (d), in order to minimize $\lambda \omega_1^2$, ω_1 should be 0. So the boundary line is $\omega_0 + \omega_2 x_2 = 0$. It should be a horizontal line. The sketch is shown below.
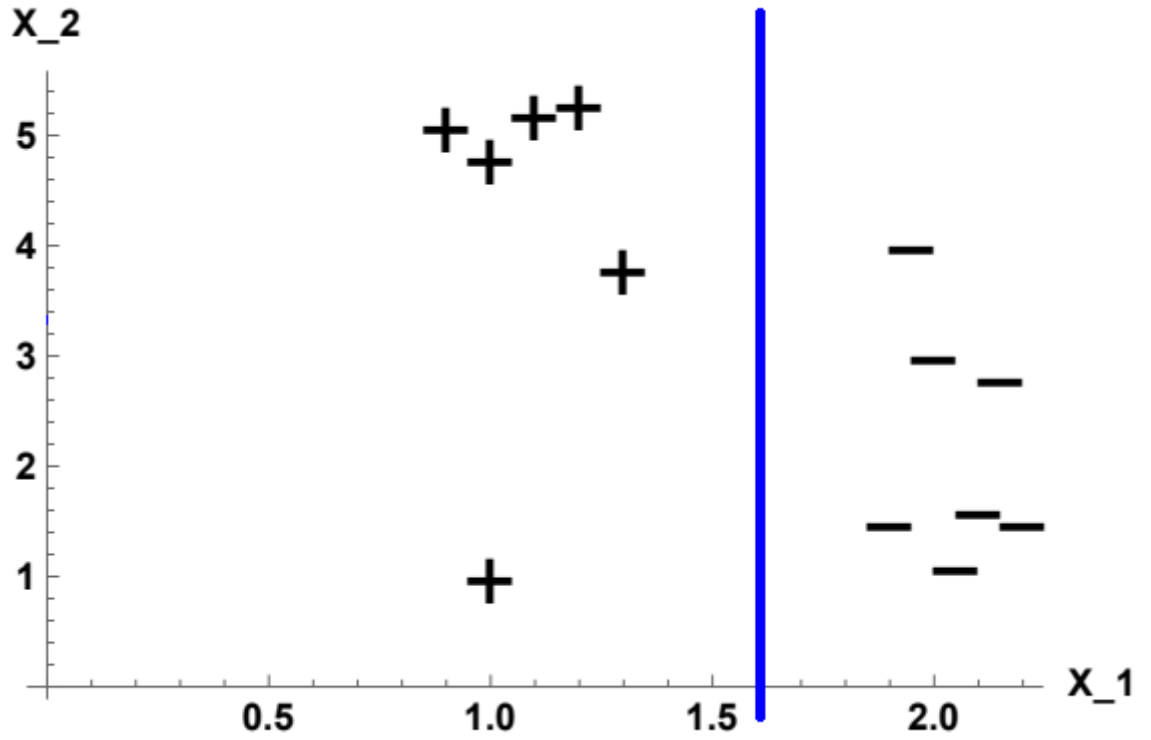
(g)

2. The upper side of the boundary marks "+" and the down side marks "-". All "+" points are in left side except one point is at down side. All "-" points are in right side except one point is at upper side. So the error is 2.

(h)

Similar with (d), in order to minimize $\lambda\omega_2^2$, ω_2 should be 0. So the boundary line is $\omega_0 + \omega_1 x_1 = 0$. It should be a vertical line. The sketch is shown below.

(i)

0. The left side of the boundary marks "+" and the right side marks "-". All "+" points are in left side and all "-" points are in right side.

3

(a)

Since $p(x)$ is possibility mass, $0 \leq p(x) \leq 1$. Therefore,

$$p(x) \leq 1$$
$$\frac{1}{p(x)} \leq 1$$
$$\log(\frac{1}{p(x)}) \geq 1 \geq 0$$

since $p(x) \geq 0$,

$$p(x) \log(\frac{1}{p(x)}) \geq 0$$
$$H(X) = \sum_x p(x) \log(\frac{1}{p(x)}) \geq 0$$

6

(b)

In order to prove $KL(p||q) \geq 0$, it is equivalent to prove $\sum_x^X p(x) \log(\frac{p(x)}{q(x)}) \geq 0$. Let's define new variable Y which is $y = f(x) = \frac{q(x)}{p(x)}$. Then let's define a new function $\phi(y) = -\log y$, which is convex function. According to Jensen's Inequality,

$$E[\phi(Y)] \geq \phi(E[Y])$$

$$\sum_x^X p(x)\phi(f(x)) \geq \phi(E[Y])$$

$$-\sum_x^X p(x) \log(\frac{q(x)}{p(x)}) \geq \phi(E[Y])$$

$$\sum_x^X p(x) \log(\frac{p(x)}{q(x)}) \geq \phi(E[Y])$$

$$KL(p||q) \geq \phi(E[Y])$$
$$KL(p||q) \geq \phi(E_x[f(x)])$$
$$KL(p||q) \geq \phi(\sum_x^X p(x)\frac{q(x)}{p(x)})$$
$$KL(p||q) \geq \phi(\sum_x^X q(x))$$
$$KL(p||q) \geq \phi(1)$$
$$KL(p||q) \geq -\log(1)$$
$$KL(p||q) \geq 0$$

(c)

In order to show that

$$I[Y;X] = KL(p(x,y)||p(x)p(y))$$

On left side,

$$I[Y;X] = H(Y) - H(Y|X)$$
$$= \sum_y p(y) \log(\frac{1}{p(y)}) - (-\sum_y \sum_x p(x,y) \log(p(y|x)))$$
$$= \sum_y p(y) \log(\frac{1}{p(y)}) + \sum_y \sum_x p(x,y) \log(p(y|x))$$

On right side,

$$KL(p(x,y)||p(x)p(y)) = \sum_y \sum_x p(x,y) \log(\frac{p(x,y)}{p(x)p(y)})$$

note that $p(y|x) = \frac{p(x,y)}{p(x)}$

$$KL(p(x,y)||p(x)p(y)) = \sum_y \sum_x p(x,y) \log(\frac{p(y|x)}{p(y)})$$

$$= \sum_y \sum_x (p(x,y)(\log(p(y|x)) + \log(\frac{1}{p(y)})))$$

$$= \sum_y \sum_x (p(x,y)(\log(p(y|x))) + \sum_y \sum_x (p(x,y) \log(\frac{1}{p(y)})))$$

$$= \sum_y \sum_x (p(x,y)(\log(p(y|x))) + \sum_y (p(y) \log(\frac{1}{p(y)})))$$

$$= I[Y;X]$$

Therefore right side is equal to left side.
Done.