Pattern Recognition, Homework 2

- 309553018 林孟學

1. Part 1

(1) (5%) Compute the mean vectors m_i (i=1, 2) of each 2 classes on **training data**

```
mean vector of class 1:
[[ 1.3559426 ]
 [-1.34746216]]
mean vector of class 2:
[[-1.29735587]
 [ 1.29096203]]
```

(2) (5%) Compute the within-class scatter matrix S_W on **training data**

```
Within-class scatter matrix SW:
[[ 388.64001349 -228.92177708]
 [-228.92177708  665.56910433]]
```

(3) (5%) Compute the between-class scatter matrix S_B on **training data**

```
Between-class scatter matrix SB:
[[ 7.03999279 -7.00052687]
 [-7.00052687  6.9612822 ]]
```

(4) (5%) Compute the Fisher's linear discriminant W on **training data**

```
 Fisher's linear discriminant:
[[-0.94096648]
 [ 0.33849976]]
```

(5) (20%) Project the **testing data** by Fisher's linear discriminant to get the class prediction by nearest-neighbor rule and calculate your accuracy score on **testing data** (you should get accuracy over 0.9)

```
y_pred = predict(x_train, y_train, w, x_test)


def accuracy_score(y_test, y_pred):
    n = len(y_test)
    return (n-sum(abs(y_test-y_pred))) / n

acc = accuracy_score(y_test, y_pred)
print(f"Accuracy of test-set {acc}")
```

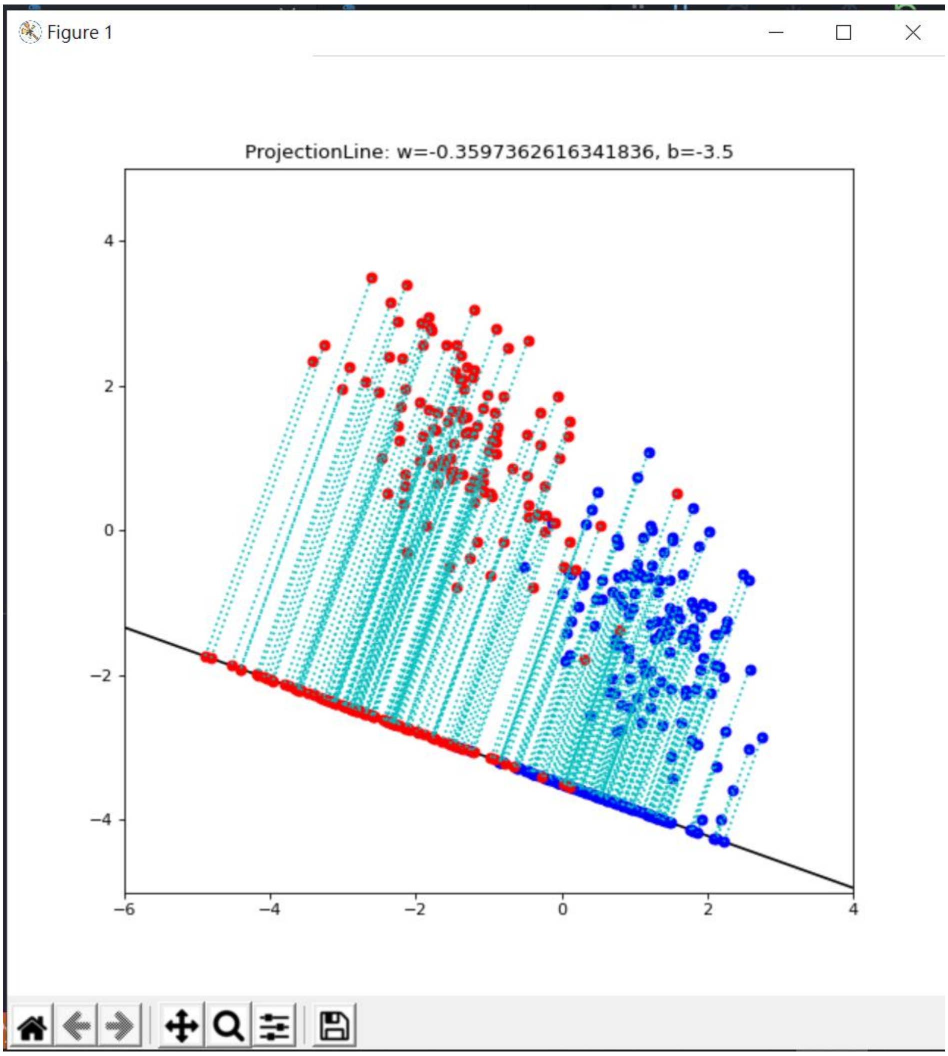
```
Accuracy of test-set 0.916
```

(6) (20%) Plot the

1) **best projection line** on the **training data** and show the slope and intercept on the title (you can choose any value of **intercept** for better visualization)

2) **colorize the data** with each class

3) project all data points on your projection line. Your result should look like the below image (This image is for reference, not the answer)



2. Part 2

(1) Show that maximization of the class separation criterion given by $L(\lambda, w) = w^T (m2 - m1) + \lambda(w^T w - 1)$ with respect to w, using a Lagrange multiplier to enforce the constraint $w^T w = 1$, leads to the result that $w \propto (m2 - m1)$.

(1)

$$L(\lambda, w) = w^T(m_2 - m_1) + \lambda(w^T w - 1)$$

$L(w) = w^T(m_2 - m_1)$ with constraint $g(w) = w^T w = 1$

let $\frac{\partial L}{\partial w} = 0$ and $\frac{\partial L}{\partial \lambda} = 0$

$$\frac{\partial L}{\partial w} = (m_2 - m_1) + 2\lambda w = 0 \quad \text{——— ①}$$

$$\frac{\partial L}{\partial \lambda} = w^T w - 1 = 0 \quad \text{——— ②}$$

$$w = -\frac{1}{2\lambda}(m_2 - m_1)$$

$$w^T w = 1$$

$$\left(-\frac{1}{2\lambda}(m_2 - m_1)\right)^T\left(-\frac{1}{2\lambda}(m_2 - m_1)\right) = 1$$

$$\frac{1}{4\lambda^2}(m_2 - m_1)^T(m_2 - m_1) = 1$$

$$(m_2 - m_1)^T(m_2 - m_1) = 4\lambda^2$$

$$\lambda = \pm\frac{1}{2}\sqrt{(m_2 - m_1)^T(m_2 - m_1)}$$ // constant

① ———

$$w = \frac{-1}{2\lambda}(m_2 - m_1)$$

$$w = \frac{-1}{\pm\sqrt{(m_2 - m_1)^T(m_2 - m_1)}}(m_2 - m_1)$$

可計算之常數

$$\Rightarrow w \propto (m_2 - m_1)$$

(2) (20%) Show that the logistic sigmoid function satisfies the property

$\sigma(-a) = 1 - \sigma(a)$ and its inverse is given by $\sigma^{-1}(y) = \ln\left(\frac{y}{1-y}\right)$.

(2)

$$\sigma(x) = \frac{1}{1+e^{-x}}$$

① prove $\sigma(-a) = 1 - \sigma(a)$

$$\sigma(-a) = \frac{1}{1+e^{a}} = \frac{1+e^{a}-e^{a}}{1+e^{a}} = 1 - \frac{e^{a}}{1+e^{a}}$$

$$= 1 - \frac{e^{a}}{1+e^{a}} \times \frac{e^{-a}}{e^{-a}}$$

$$= 1 - \frac{e^{a+(-a)}}{e^{-a}+e^{a+(-a)}} = 1 - \frac{1}{e^{-a}+1} = 1 - \sigma(a) \#$$

② prove $\sigma^{-1}(y) = \ln\left(\frac{y}{1-y}\right)$

let $y = \sigma(a)$,

$a = \sigma^{-1}(y)$

prove that $\sigma^{-1}(y) = \ln\left(\frac{y}{1-y}\right) = \ln\left(\frac{\sigma(a)}{1-\sigma(a)}\right) = a$

$$\ln\left(\frac{\sigma(a)}{1-\sigma(a)}\right) = \ln\left(\frac{1-\sigma(-a)}{\sigma(-a)}\right) = \ln\left(\frac{1-\frac{1}{1+e^{a}}}{\frac{1}{1+e^{a}}}\right) = \ln\left(\left(1-\frac{1}{1+e^{a}}\right)(1+e^{a})\right)$$

$$= \ln\left(1 + e^{a} \underbrace{- \frac{1}{1+e^{a}} - \frac{e^{a}}{1+e^{a}}}_{=-1}\right)$$

$$= \ln(1+e^{a}-1) = \ln(e^{a}) = a \#$$