1. mean vector of class 1: [2.47107265 1.97913899]
   mean vector of class 2: [1.82380675 3.03051876]

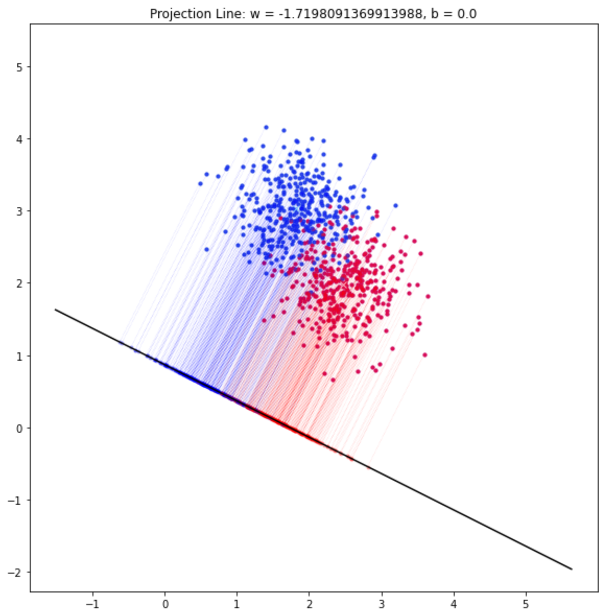
2. Within-class scatter matrix SW:
   [[140.40036447 -5.30881553]
   [-5.30881553 138.14297637]]

3. Between-class scatter matrix SB:
   [[0.41895314 -0.68052227]
   [-0.68052227 1.10539942]]

4. Fisher's linear discriminant:
   [[-0.50266214] [0.86448295]]

5. Accuracy of test-set 0.912

6. 1) best projection line on the training data and show the slope and intercept on the title
   2) colorize the data with each class
   3) project all data points on your projection line.



1.

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

① Taking gradient of L with respect to w, we obtain

$\nabla L = (m_2-m_1) + 2\lambda w$

② Setting $\nabla L=0$ gives

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

from which it follows that $w \propto (m_2-m_1)$

2.

$$eq6 = J(w) = \frac{(m_2-m_1)^2}{s_1^2+s_2^2}$$

① $(m_2-m_1)^2 = (w^T m_2 - w^T m_1)^2$ eq2

$$= w^T(m_2-m_1)(m_2-m_1)^T w$$

$$= w^T S_B w \text{ ...... eq4}$$

② $s_k^2 = \sum_{n\in C_k} (y_n - m_k)^2$ eq3

$$= \sum_{n\in C_k} (w^T x_n - w^T m_k)^2 \text{ ...... eq1}$$

$$= \sum_{n\in C_k} w^T (x_n - m_k)(x_n - m_k)^T w$$

$$= w^T \left[\sum_{n\in C_k} (x_n - m_k)(x_n - m_k)^T\right] w$$

$$\Rightarrow s_1^2 + s_2^2 = w^T \left[\sum_{n\in C_1}(x_n-m_1)(x_n-m_1)^T + \sum_{n\in C_2}(x_n-m_2)(x_n-m_2)^T\right] w$$

$$= w^T S_W w \text{ ...... eq5}$$

According to ①②, $J(w) = \frac{(m_2-m_1)^2}{s_1^2+s_2^2}$

$$= \frac{w^T S_B w}{w^T S_W w}$$

3.

$$E(w) = -\sum_{n=1}^{N} \{t_n \ln y_n + (1-t_n) \ln(1-y_n)\} \text{ ...... eq9}$$

① $\frac{\delta E}{\delta y_n} = -\left[\frac{t_n}{y_n} - \frac{(1-t_n)}{1-y_n}\right]$

$$= \frac{-t_n(1-y_n) + y_n(1-t_n)}{y_n(1-y_n)}$$

$$= \frac{y_n - t_n}{y_n(1-y_n)}$$

② $\frac{\delta y_n}{\delta a_n} = \frac{\delta \sigma(a_n)}{\delta a_n} = \sigma(a_n)[1-\sigma(a_n)]$ eq8

$$= y_n(1-y_n)$$

③ $a_n = w^T \phi_n \Rightarrow \nabla a_n = \phi_n$ where ∇ denotes the gradient with respect to w.

According to ①②③, we have :

$$\nabla E = \sum_{n=1}^{N} \frac{\delta E}{\delta y_n} \frac{\delta y_n}{\delta a_n} \nabla a_n$$

$$= \sum_{n=1}^{N} (y_n - t_n) \phi_n$$