NCTU Pattern Recognition, Homework 2

0516003 李智嘉

Part 1, Coding (60%):

Q1: Compute the mean vectors m_i, (i=1,2) of each 2 classes on training data

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

Q2: Compute the within-class scatter matrix S_W on training data

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

Q3: Compute the between-class scatter matrix S_B on training data

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

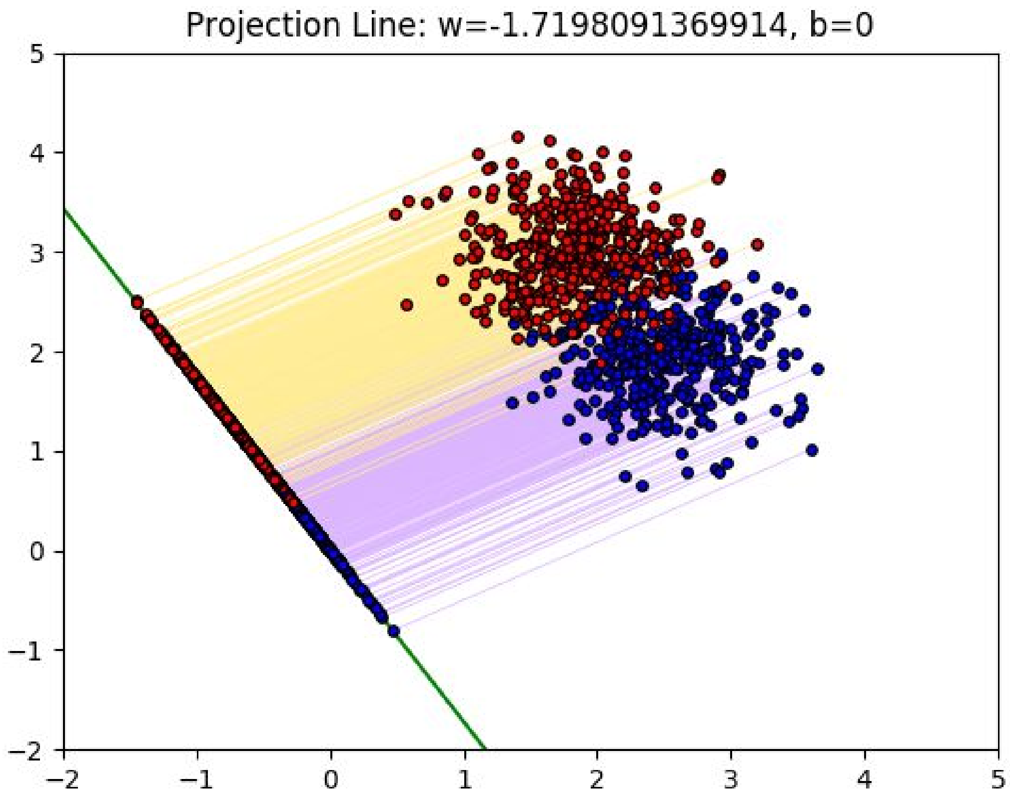
Q4: Compute the Fisher's linear discriminant W on training data

- ```
  Fisher's linear discriminant:
  [[ 0.50266214]
   [-0.86448295]]
  ```

Q5: Project the **testing data** by linear discriminant to get the class prediction by nearest-neighbor rule and calculate your accuracy score on **testing data**

- ```
  Accuracy of test-set 0.912
  ```

Q6: 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.
  ```



Projection Line: w=-1.7198091369914, b=0
5
4
3
2
1
0
−1
−2
−2
−1
0
1
2
3
4
5

Part 2, Questions (40%):

Q1: 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.

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

$$= w^T m_2 - w^T m_1 + \lambda w^T w - \lambda$$

$$\frac{\partial L(\lambda, w)}{\partial w} = m_2 - m_1 + 2\lambda w = 0$$

$$\Rightarrow 2\lambda w = m_1 - m_2$$

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

$$\Rightarrow \underline{w \propto (m_2 - m_1)} \quad \#$$

Q2: Using (eq 1) and (eq 2), derive the result (eq 3) for the posterior class probability in the two-class generative model with Gaussian densities, and verify the results (eq 4) and (eq 5) for the parameters w and w0.

2.

Gaussian density: $P(x|C_i) = \frac{1}{(2\pi)^{\frac{D}{2}}} \frac{1}{|\Sigma|^{\frac{1}{2}}} e^{-\frac{1}{2}(x-\mu_i)^T \Sigma^{-1}(x-\mu_i)}$

$$a = \ln \frac{P(x|C_1)P(C_1)}{P(x|C_2)P(C_2)} = \ln \frac{P(x|C_1)}{P(x|C_2)} + \ln \frac{P(C_1)}{P(C_2)}$$

$$= -\frac{1}{2}(x-\mu_1)^T \Sigma^{-1}(x-\mu_1) + \frac{1}{2}(x-\mu_2)^T \Sigma^{-1}(x-\mu_2) + \ln \frac{P(C_1)}{P(C_2)}$$

$$= -\frac{1}{2}\left(x^T \Sigma^{-1} x - 2x^T \Sigma^{-1}\mu_1 + \mu_1^T \Sigma^{-1}\mu_1\right)$$
$$+ \frac{1}{2}\left(x^T \Sigma^{-1} x - 2x^T \Sigma^{-1}\mu_2 + \mu_2^T \Sigma^{-1}\mu_2\right) + \ln \frac{P(C_1)}{P(C_2)}$$

$$= x^T \Sigma^{-1}(\mu_1 - \mu_2) - \frac{1}{2}\mu_1^T \Sigma^{-1}\mu_1 + \frac{1}{2}\mu_2^T \Sigma^{-1}\mu_2 + \ln \frac{P(C_1)}{P(C_2)}$$

$$= w^T x + w_0$$

where $w = \Sigma^{-1}(\mu_1 - \mu_2)$

$w_0 = -\frac{1}{2}\mu_1^T \Sigma^{-1}\mu_1 + \frac{1}{2}\mu_2^T \Sigma^{-1}\mu_2 + \ln \frac{P(C_1)}{P(C_2)}$ #

$\Rightarrow P(C_1|x) = \sigma(a) = \sigma(w^T x + w_0)$ #