Chapter 2 Homework

2015011313 徐鉴劲 计54

Problem 1

(a)

Error probability $P(error|x) = max[P(\omega_1|x), P(\omega_2|x)$, in which $P(\omega_1|x) = \frac{P(x|\omega_1)P((\omega_1)}{p(x)}$, and $P(\omega_2|x)$ is similar. Minimize error probability gives the following decision rule:

Select ω_1 if $P(x|\omega_1) > P(x|\omega_2)$. Select ω_2 otherwise.

(b)

Suppose $R(\omega|x)$ is the risk of selecting ω when x is observed. This error risk matrix gives the following risk expression:

$R(\omega_1|x) = P(\omega_2|x)$, $R(\omega_2|x) = 0.5P(\omega_1|x)$. To minimize the risk, the following dicision rule can be reached:

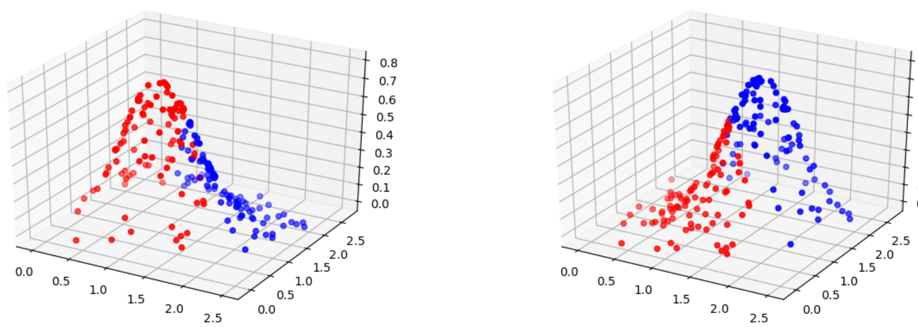
If $\frac{P(\omega_1|x)}{P(\omega_2|x)} < \frac{1}{2}$, select ω_2. Otherwise, select ω_1.

Experiment

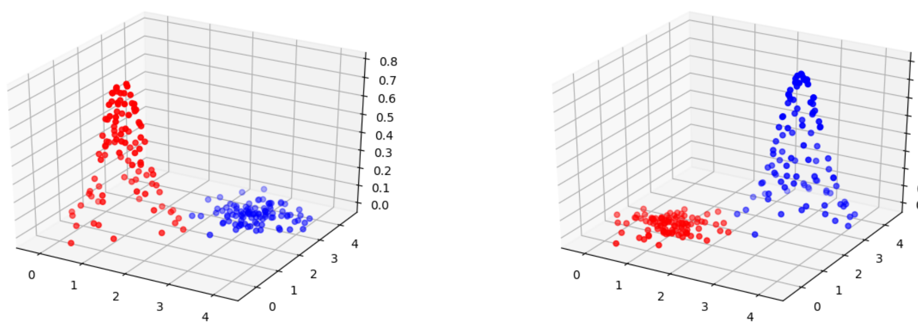
Run the experiment:

`python hw2.py`

Set μ to be 1.5, the accuracy is about 70% ~ 80%, which varies greatly. To be specific, the accuracy of ω_1 and ω_2 are close to each other. In addition, the figure below shows the feature probability $P(x|\omega_1)$. Red and blue represent ω_1, ω_2 respectively.



Set μ to be 3, the accuracy is 100%. The feature probability is also shown below.



Problem 1

(a)

As the covariance matrix Σ can be divided into 2 blocks,
$P(x_1, x_2, x_3) = P(x_1)P(x_2, x_3) = \mathcal{N}(x_1; 1, 1)\mathcal{N}(x_1, x_2; \begin{bmatrix} 2 \\ 2 \end{bmatrix}, \begin{bmatrix} 5 & 2 \\ 2 & 5 \end{bmatrix})$.

The formulae for two dimensional normal distribution is $\frac{1}{2\pi\sqrt{|\Sigma|}} e^{-\frac{1}{2}(x-\mu)^T\Sigma^{-1}(x-\mu)}$. $|\Sigma| = 21, \Sigma^{-1} = \frac{1}{21}\begin{bmatrix} 5 & -2 \\ -2 & 5 \end{bmatrix}$.

$$P(x_0|\omega) = \frac{1}{\sqrt{2\pi}} e^{-\frac{(x-1)^2}{2}} = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{8}} \sim 0.35206532676.$$

$$P(x_1, x_2|\omega) = \frac{1}{42\pi} e^{-\frac{1}{42}[2,1]\begin{bmatrix} 5 & -2 \\ -2 & 5 \end{bmatrix}\begin{bmatrix} 2 \\ 1 \end{bmatrix}} = \frac{1}{42\pi} e^{-\frac{17}{42}} = 5.056092087 \times 10^{-3}$$

So $P(\mathbf{x}_0) = 1.78 \times 10^{-3}$.

(b)

To transform the matrix into identity matrix, first we do eigen value decomposition:

$B\begin{bmatrix} 5 & 2 \\ 2 & 5 \end{bmatrix}B^T = \begin{bmatrix} 3 & 0 \\ 0 & 7 \end{bmatrix}$, in which $B = \begin{bmatrix} \frac{\sqrt{2}}{2} & -\frac{\sqrt{2}}{2} \\ \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \end{bmatrix}$.

Let the original random variables to be $X = \begin{bmatrix} X_1 \\ X_2 \\ X_3 \end{bmatrix}$, and $\tilde{B} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \frac{\sqrt{2}}{2} & -\frac{\sqrt{2}}{2} \\ 0 & \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \end{bmatrix}$ the tranformation would be

$\tilde{X} = diag(1, \frac{1}{\sqrt{3}}, \frac{1}{\sqrt{7}})\tilde{B}(X - \mu)$

(c)

Apply the transformation to $\mathbf{x}_0$, the result is $\tilde{\mathbf{x}}_0 = \begin{bmatrix} \frac{1}{2} \\ \frac{1}{\sqrt{6}} \\ \frac{3}{\sqrt{14}} \end{bmatrix}$

(d)

Mahalanobis distance is $B_M(x) = \sqrt{(x-\mu)^T\Sigma^{-1}(x-\mu)}$.

For original distribution, $d_1 = \sqrt{\frac{1}{21}[\frac{1}{2}, 2, 1]\begin{bmatrix} 21 & 0 & 0 \\ 0 & 5 & -2 \\ 0 & -2 & 5 \end{bmatrix}\begin{bmatrix} \frac{1}{2} \\ 2 \\ 1 \end{bmatrix}} = \frac{1}{2}\sqrt{\frac{89}{21}}$.

For transformed distribution, $d_2 = \sqrt{|||[\frac{1}{2}, \frac{1}{\sqrt{6}}, \frac{3}{\sqrt{14}}]|||^2} = \sqrt{\frac{1}{4} + \frac{1}{6} + \frac{9}{14}} = \frac{1}{2}\sqrt{\frac{89}{21}}$.

$d_1 = d_2$.

(e)

Original probability density is $P(\mathbf{x}_0) = Ce^{-\frac{1}{2}(x_0-\mu)^T\Sigma^{-1}(x_0-\mu)}$.

The tranformed probability density is

$P(\tilde{\mathbf{x}}_0) = Ce^{-\frac{1}{2}(\tilde{x}_0 - T^t\mu)^T(T^t\Sigma T)^{-1}(\tilde{x}_0 - T^t\mu)}$

in which $\tilde{\mathbf{x}}_0 = T^t\mathbf{x}_0$.

Thus we have

$P(\tilde{\mathbf{x}}_0) = Ce^{-\frac{1}{2}(x_0-\mu)^t T(T^t\Sigma T)^{-1}T^t(x_0-\mu)}$

As T is a linear tranformation, it is not singular, we have

$(T^t\Sigma T)^{-1} = T^{-1}\Sigma^{-1}(T^{-1})^t$

so all the T can be canceled out:

$P(\tilde{\mathbf{x}}_0)$ = $Ce^{-\frac{1}{2}(x_0-\mu)^t T T^{-1} \Sigma^{-1} (T^{-1})^t T^t (x_0-\mu)}$ = $Ce^{-\frac{1}{2}(x_0-\mu)^T \Sigma^{-1}(x_0-\mu)} = P(\mathbf{x}_0)$.

(f)

Let the gaussian random vector to be X, and the original parameter to be μ and Σ. Apply the whitening transformation $\tilde{X} = \Phi\Lambda^{-\frac{1}{2}}X$, then $\mathcal{N}(\mu, \Sigma)$ is tranformed into $\mathcal{N}(\Phi\Lambda^{-\frac{1}{2}}\mu, \Phi\Lambda^{-\frac{1}{2}}\Sigma(\Phi\Lambda^{-\frac{1}{2}})^T)$.

As $\Phi\Sigma\Phi^T = \Lambda$, so $\Phi\Lambda^{-\frac{1}{2}}\Sigma(\Lambda^{-\frac{1}{2}})^T\Phi^T = \Lambda^{-\frac{1}{2}}\Phi\Sigma\Phi^T(\Lambda^{-\frac{1}{2}})^T = I$.

Problem 2

(a)

$P(x_0, x_1, x_2, x_3|\omega_1, \omega_3, \omega_3, \omega_2) = P(0.6|\omega_1)P(0.1|\omega_3)P(0.9|\omega_3)P(1.1|\omega_2) = \frac{1}{4\pi^2}e^{-\frac{0.4^2}{2}}e^{-\frac{0.9^2}{2}}e^{-\frac{0.1^2}{2}}e^{-\frac{0.6^2}{2}} = \frac{1}{4\pi^2}e^{-\frac{0.16+0.81+0.01+0.36}{2}} = 0$

(b)

$P(0.6, 0.1, 0.9, 1.1|\omega_1, \omega_2, \omega_2, \omega_3) = \frac{1}{4\pi^2}e^{-\frac{0.36+0.16+0.16+0.01}{2}} = 0.018$

(c)

The sequence is $\omega_2, \omega_1, \omega_3, \omega_3$.