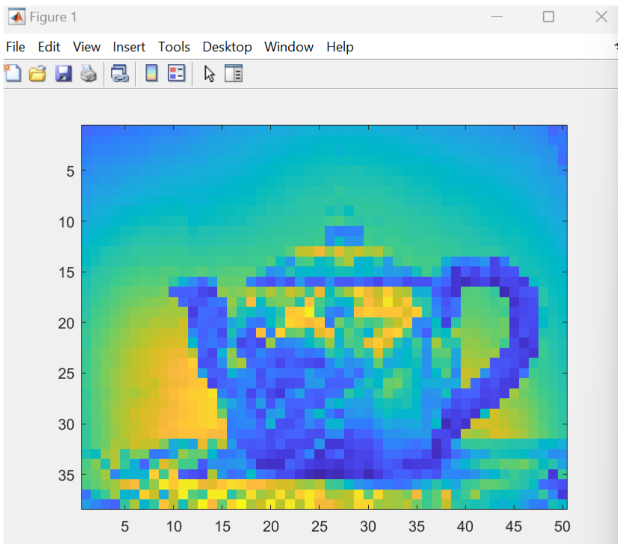
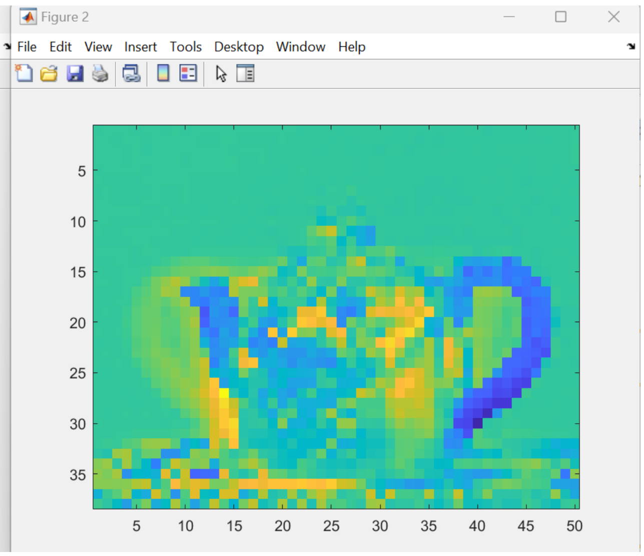
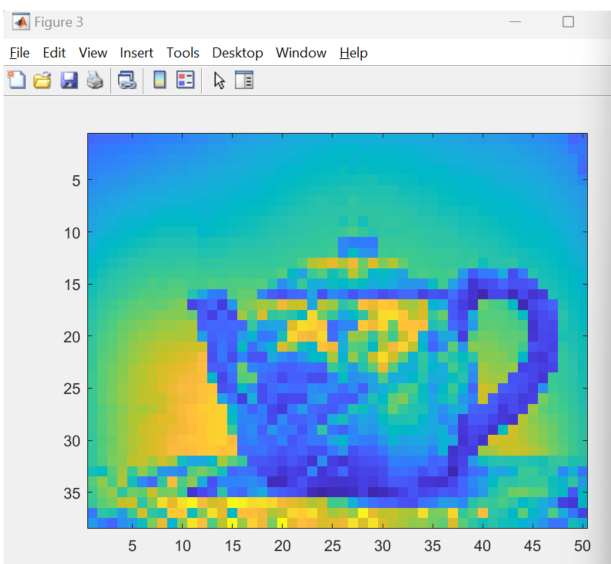
# HW4

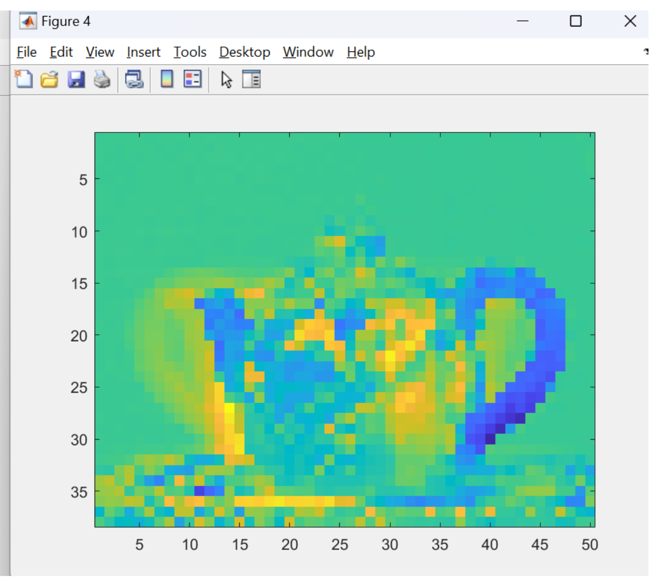
## Q1

After observing the following 10 sets of pictures (the original one on the left, the generated one on the right), it can be concluded that PCA reduced the dimensions of the original data by keeping the most relevant and effective features.



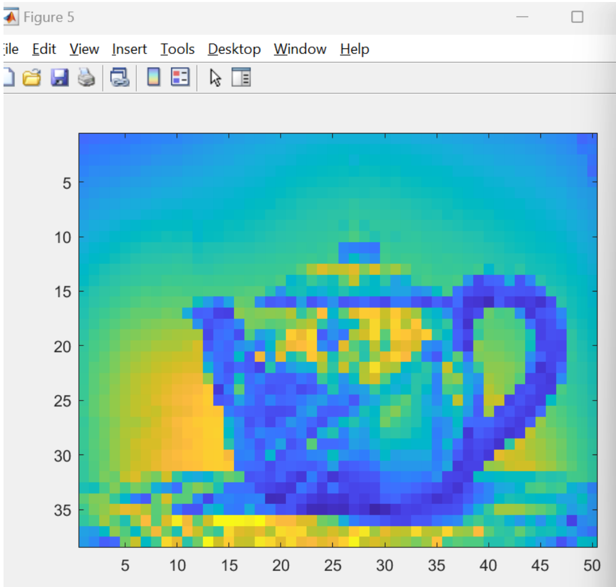


Figure 5
ile Edit View Insert Tools Desktop Window Help
5 10 15 20 25 30 35
5 10 15 20 25 30 35 40 45 50

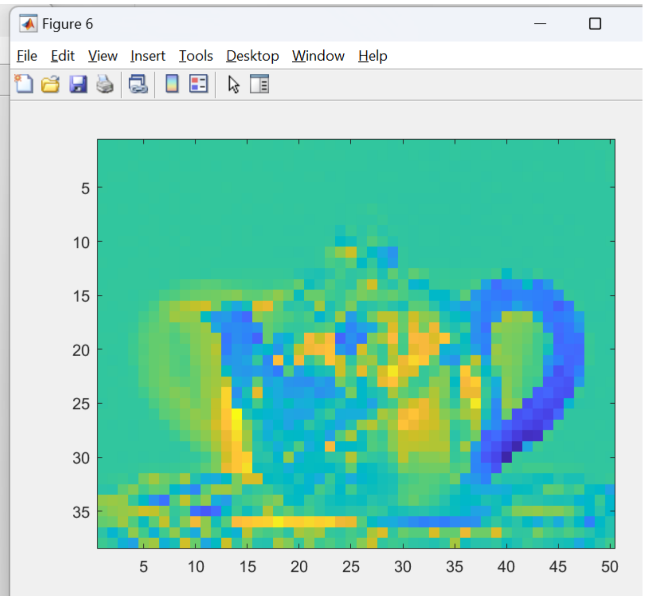


Figure 6
File Edit View Insert Tools Desktop Window Help
5 10 15 20 25 30 35
5 10 15 20 25 30 35 40 45 50

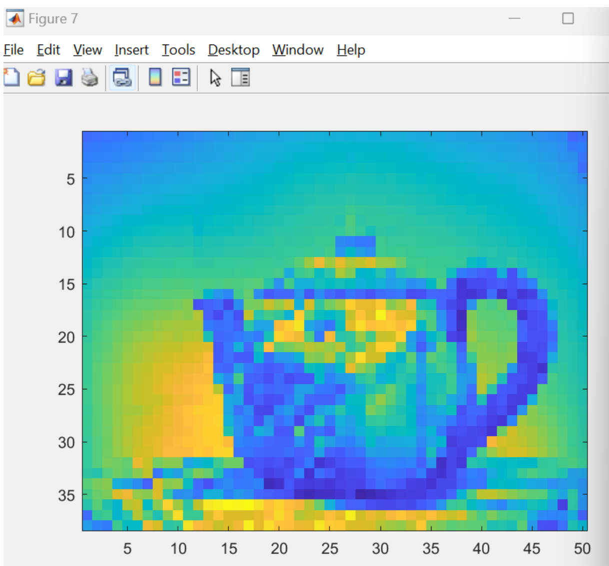


Figure 7
File Edit View Insert Tools Desktop Window Help
5 10 15 20 25 30 35
5 10 15 20 25 30 35 40 45 50

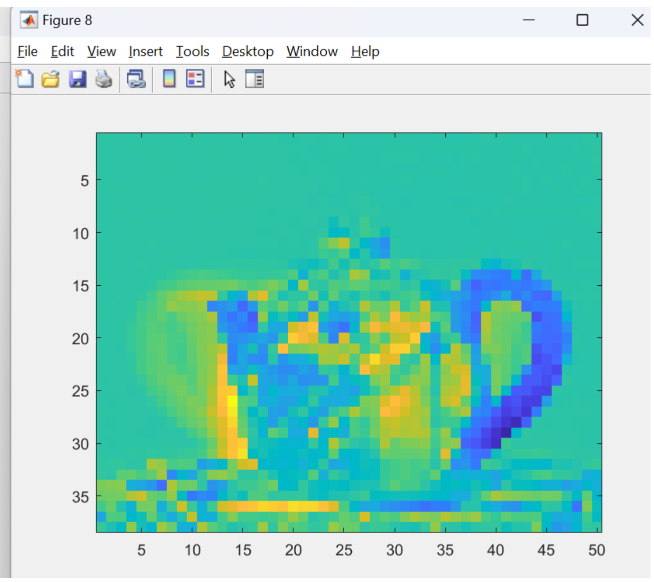


Figure 8
File Edit View Insert Tools Desktop Window Help
5 10 15 20 25 30 35
5 10 15 20 25 30 35 40 45 50

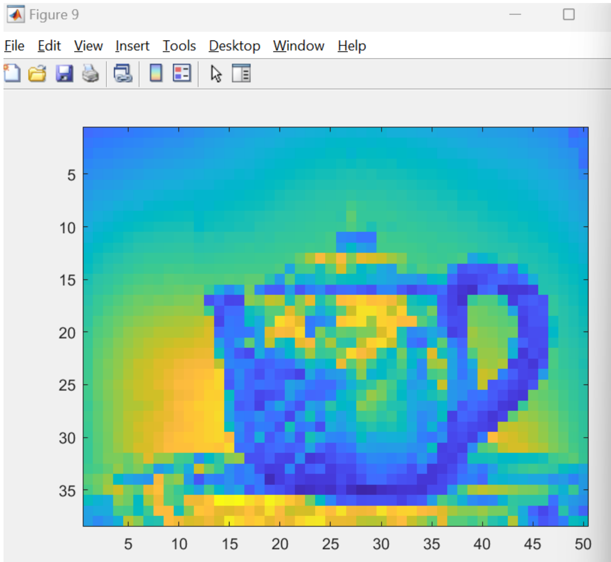


Figure 9
File Edit View Insert Tools Desktop Window Help
5
10
15
20
25
30
35
5 10 15 20 25 30 35 40 45 50

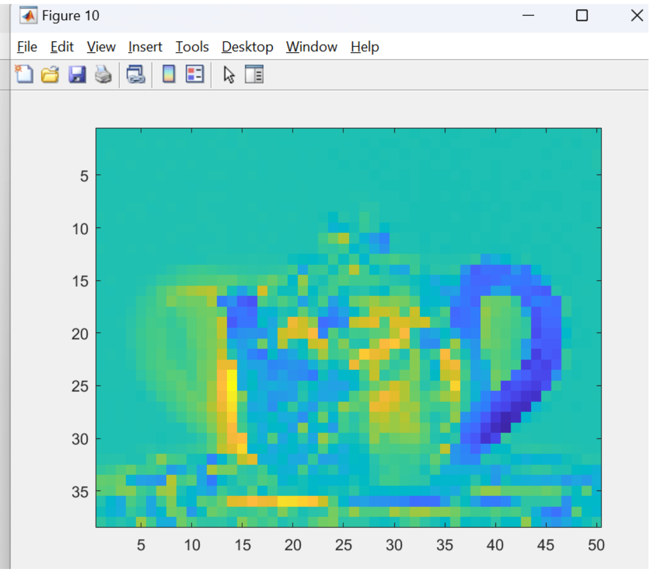


Figure 10
File Edit View Insert Tools Desktop Window Help
5
10
15
20
25
30
35
5 10 15 20 25 30 35 40 45 50

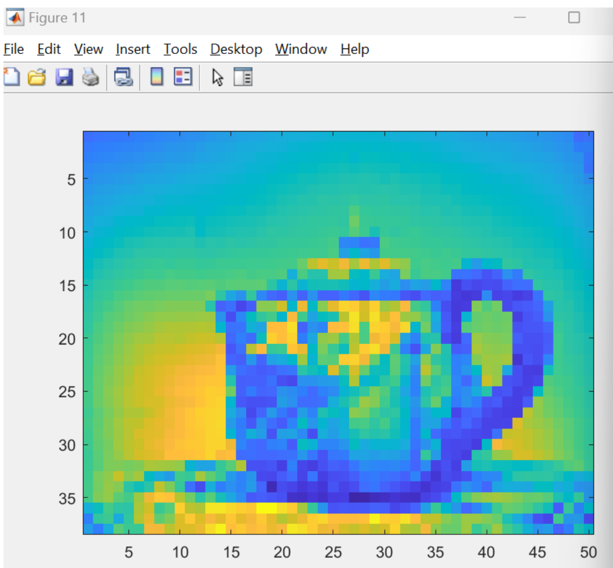
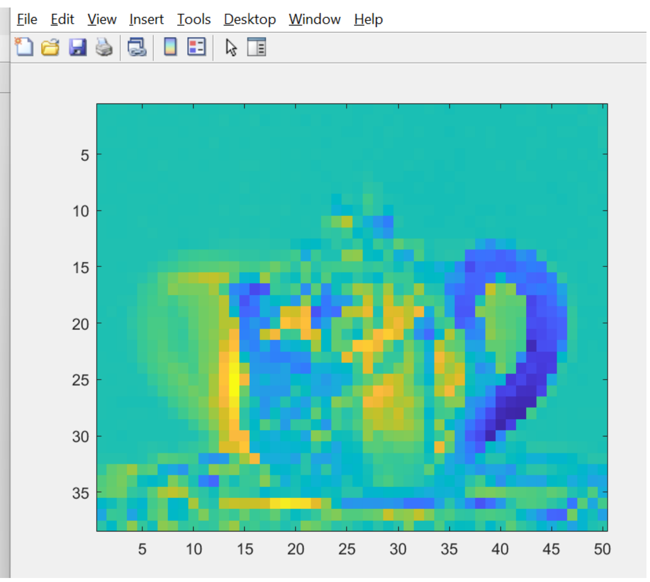


Figure 11
File Edit View Insert Tools Desktop Window Help
5
10
15
20
25
30
35
5 10 15 20 25 30 35 40 45 50


File Edit View Insert Tools Desktop Window Help
5
10
15
20
25
30
35
5 10 15 20 25 30 35 40 45 50

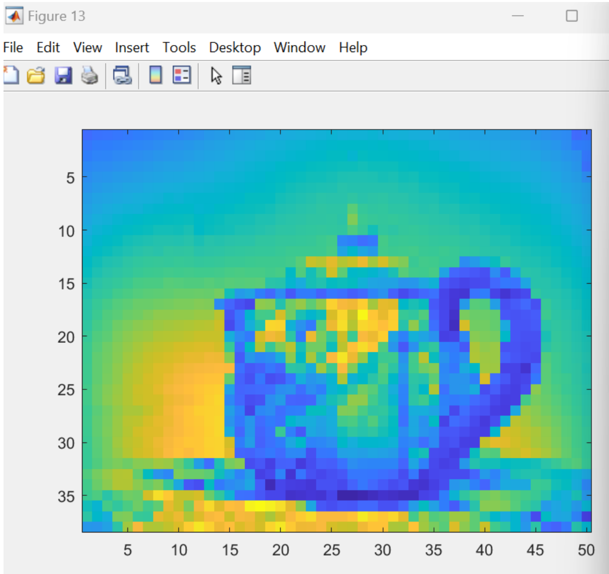


Figure 13
File Edit View Insert Tools Desktop Window Help
5 10 15 20 25 30 35
5 10 15 20 25 30 35 40 45 50

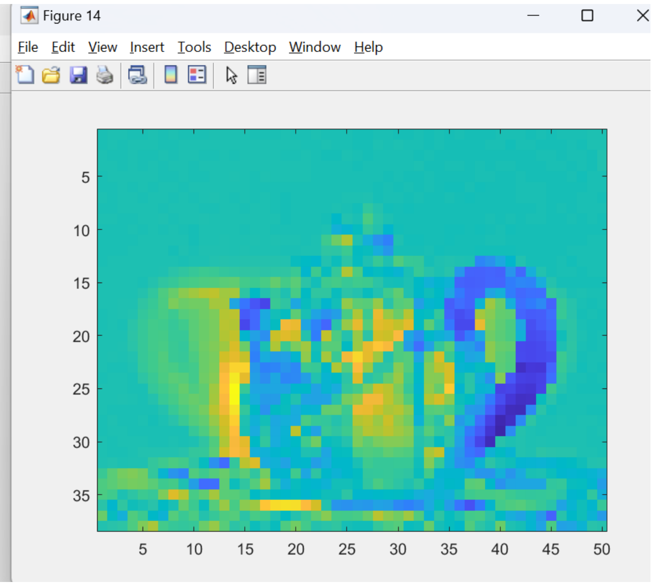


Figure 14
File Edit View Insert Tools Desktop Window Help
5 10 15 20 25 30 35
5 10 15 20 25 30 35 40 45 50

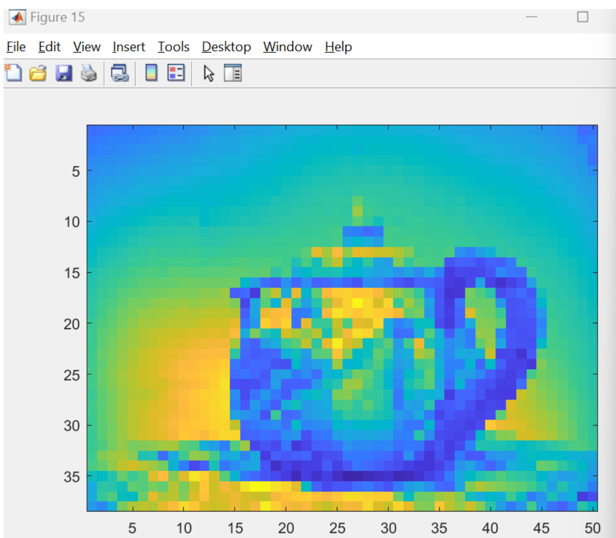


Figure 15
File Edit View Insert Tools Desktop Window Help
5 10 15 20 25 30 35
5 10 15 20 25 30 35 40 45 50

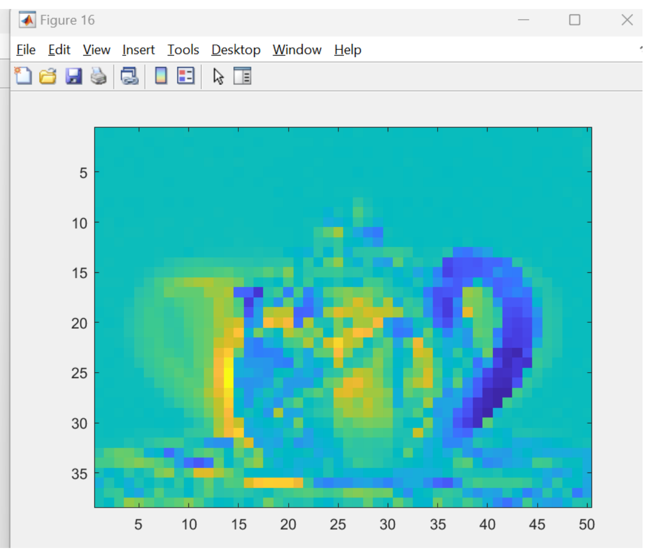


Figure 16
File Edit View Insert Tools Desktop Window Help
5 10 15 20 25 30 35
5 10 15 20 25 30 35 40 45 50

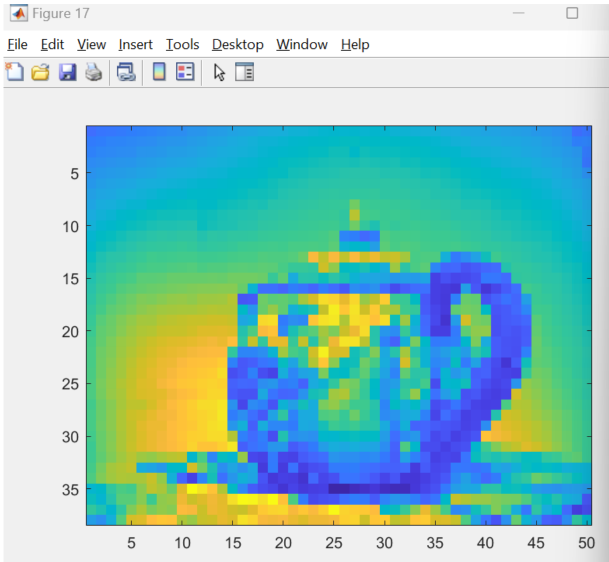


Figure 17
File Edit View Insert Tools Desktop Window Help
5 10 15 20 25 30 35
5 10 15 20 25 30 35 40 45 50

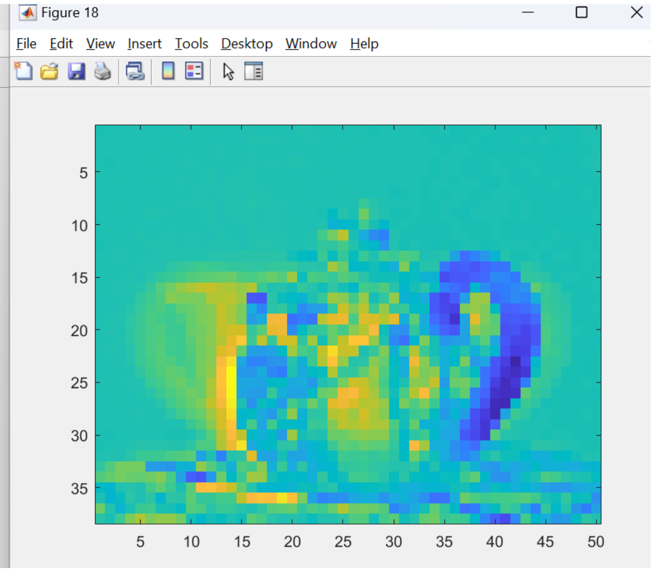


Figure 18
File Edit View Insert Tools Desktop Window Help
5 10 15 20 25 30 35
5 10 15 20 25 30 35 40 45 50

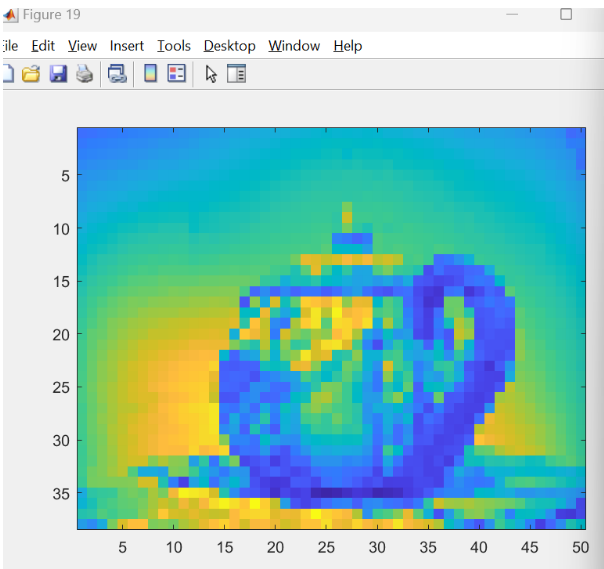


Figure 19
ile Edit View Insert Tools Desktop Window Help
5 10 15 20 25 30 35
5 10 15 20 25 30 35 40 45 50

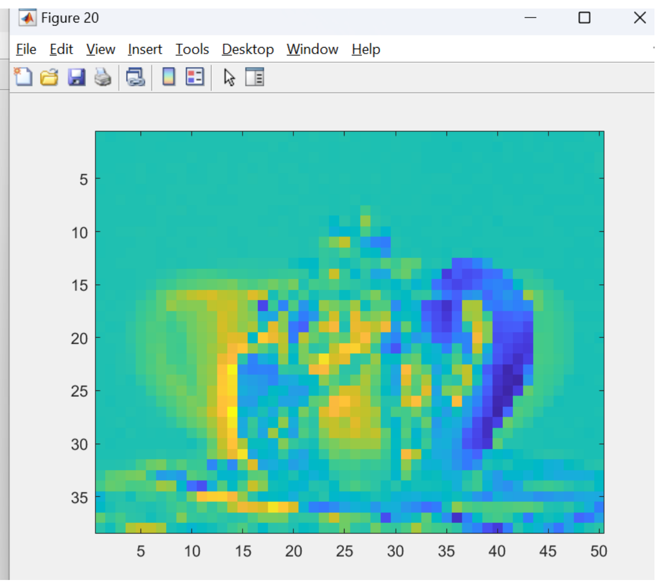


Figure 20
File Edit View Insert Tools Desktop Window Help
5 10 15 20 25 30 35
5 10 15 20 25 30 35 40 45 50

**Q2**

$X$: draw Apple $\quad E_1$: Box 1 $\quad E_2$: Box 2

$$P(E_1|X) = \frac{P(X|E_1)\,P(E_1)}{P(X)} = \frac{P(X|E_1)P(E_1)}{P(X|E_1)P(E_1) + P(X|E_2)P(E_2)}$$

$$= \frac{\frac{8}{12} \cdot \frac{1}{2}}{\frac{8}{12} \cdot \frac{1}{2} + \frac{10}{12} \cdot \frac{1}{2}} = \frac{4}{9}$$

**Q3**

$$P(y|\theta) = \alpha^{y}(1-\alpha)^{1-y}$$

$$\sum_{i=1}^{N} \log P(y_i|\alpha) = \sum_{i=1}^{N} \log \alpha^{y_i}(1-\alpha)^{1-y_i}$$

$$\Rightarrow \quad \frac{\partial}{\partial\alpha}\sum_{i=1}^{N} \log \alpha^{y_i}(1-\alpha)^{1-y_i} = 0$$

$$\frac{\partial}{\partial\alpha}\sum_{i=1}^{N} y_i \log\alpha + (1-y_i)\log(1-\alpha) = 0$$

$$\frac{\partial}{\partial\alpha}\sum_{i\in \text{Class } 1} \log\alpha + \sum_{i\in\text{Class } 0} \log(1-\alpha) = 0$$

$$\sum_{i\in\text{Class }1} \frac{1}{\alpha} + \sum_{\dots} \frac{-1}{1-\alpha} = 0$$

$$N_1 \frac{1}{\alpha} - N_0 \frac{1}{1-\alpha} = 0$$

$$N_1(1-\alpha) - N_0\alpha = 0$$

$$N_1 - \alpha(N_1+N_0) = 0$$

$$\boxed{\alpha = \frac{N_1}{N_1+N_0}}$$

$$\sum_{i=1}^{N} \log P(\vec{x}_i|\vec{\mu},\Sigma) = \sum_{i=1}^{N} \log \frac{1}{(2\pi)^{D/2}\sqrt{|\Sigma|}} \exp\left(-\frac{1}{2}(\vec{x}_i-\vec{\mu})^T \Sigma^{-1}(\vec{x}_i-\vec{\mu})\right)$$

$\rightarrow$ Max over $\mu$

$$\frac{\partial}{\partial\mu}\left(\sum_{i=1}^{N} \log \frac{1}{(2\pi)^{D/2}\sqrt{|\Sigma|}} \exp(-\frac{1}{2}(\vec{x}_i-\vec{\mu})\Sigma^{-1}(\vec{x}_i-\vec{\mu})\right) = 0$$

$$\sum_{i=1}^{N} (\vec{x}_i - \vec{\mu})\Sigma^{-1} = \vec{0}$$

$$\boxed{\vec{\mu} = \frac{1}{N}\sum_{i=1}^{N} \vec{x}_i}$$

$$l = \sum_{i=1}^{N} -\frac{D}{2}\log 2\pi - \frac{1}{2}\log|\Sigma| - \frac{1}{2}(\vec{x}_i - \vec{\mu})^T \Sigma^{-1} (\vec{x}_i - \vec{\mu})$$

$$= -\frac{ND}{2}\log 2\pi + \frac{N}{2}\log|\Sigma^{-1}| - \frac{1}{2}\sum_{i=1}^{N} tr\left[(\vec{x}_i - \vec{\mu})^T \Sigma^{-1} (\vec{x}_i - \vec{\mu})\right]$$

$$= -\frac{ND}{2}\log 2\pi + \frac{N}{2}\log A - \frac{1}{2}\sum_{i=1}^{N} tr\left[(\vec{x}_i - \vec{\mu})(\vec{x}_i - \vec{\mu})^T A\right]$$

$$\frac{\partial l}{\partial A} = -0 + \frac{N}{2}(A^{-1})^T - \frac{1}{2}\sum_{i=1}^{N}\left[(\vec{x}_i - \vec{\mu})(\vec{x}_i - \vec{\mu})^T\right]^T$$

$$= \frac{N}{2}\Sigma - \frac{1}{2}\sum_{i=1}^{N}(\vec{x}_i - \vec{\mu})(\vec{x}_i - \vec{\mu})^T \Rightarrow 0$$

$$\Sigma = \frac{1}{N}\sum_{i=1}^{N}(\vec{x}_i - \vec{\mu})(\vec{x}_i - \vec{\mu})^T$$

$Y = \arg\max P(\hat{Y}|X) \quad \hat{Y} \in \{0,1\}$

$\Rightarrow Y = \underset{\hat{Y} \in \{0,1\}}{\arg\max}\, P(X|\hat{Y}) P(\hat{Y})$

$$P(X|\hat{Y}) = \frac{1}{(2\pi)^{d/2} |\Sigma_Y|^{\frac{1}{2}}} \exp\left(-\frac{1}{2}(X-\mu_Y)^T \Sigma_Y^{-1} (X-\mu_Y)\right)$$

$$\Rightarrow Y = \underset{\hat{Y} \in \{0,1\}}{\arg\max} \left(-\frac{1}{2}(X-\mu_Y)^T \Sigma_Y^{-1} (X-\mu_Y) + \log P(Y) - \frac{d}{2}\log(2\pi) - \frac{1}{2}\log(\Sigma_Y)\right)$$

assume $\Sigma_0 = \Sigma_1 = \Sigma$

~~$\Rightarrow Y = \underset{\hat{Y} \in \{0,1\}}{\arg\max}$~~

$$\Rightarrow Y = \underset{\hat{Y} \in \{0,1\}}{\arg\min} \left((X-\mu_Y)^T \Sigma^{-1} (X-\mu_Y) - 2\log P(Y)\right)$$

$$Y = \underset{\hat{Y} \in \{0,1\}}{\arg\min} \left(X^T \Sigma^{-1} X - 2X^T \Sigma^{-1} \mu_Y + \mu_Y^T \Sigma^{-1} \mu_Y - 2\log P(Y)\right)$$

$$\Rightarrow Y = \arg\max \left(X^T \Sigma^{-1} \mu_Y - \frac{1}{2}\mu_Y^T \Sigma^{-1} \mu_Y + \log P(Y)\right)$$

Define $W_0 = \Sigma^{-1}\mu_0 - \frac{1}{2}\mu_0^T \Sigma^{-1} \mu_0 + \log P(0)$

$W_1 = \Sigma^{-1}\mu_1 - \frac{1}{2}\mu_1^T \Sigma^{-1} \mu_1 + \log P(1)$

$\Rightarrow Y = \arg\max X^T W_Y$

$(W_0 - W_1)^T X = 0$ decision boundary, it is linear.

$$\log P(x \mid Y=0) = -\frac{d}{2}\log(2\pi) - \frac{1}{2}\log|\Sigma_y| - \frac{1}{2}(x-\mu_y)^T \Sigma_y (x-\mu_y)$$

$$\Rightarrow -\frac{d}{2}\log(2\pi) - \frac{1}{2}\log\Sigma_0 - \frac{1}{2}(x-\mu_0)^T\Sigma_0^{-1}(x-\mu_0) + \log P(Y=0) =$$
$$-\frac{d}{2}\log(2\pi) - \frac{1}{2}\log\Sigma_1 - \frac{1}{2}(x-\mu_1)^T\Sigma_1^{-1}(x-\mu_1) + \log P(Y=1)$$

$$\Rightarrow \frac{1}{2}\log(\Sigma_0/\Sigma_1) - \frac{1}{2}(x^T\Sigma_0^{-1}x - x^T\Sigma_1^{-1}x - 2x^T\Sigma_0^{-1}\mu_0 + 2x^T\Sigma_1^{-1}\mu_1 + \mu_0^T\Sigma_0^{-1}\mu_0$$
$$-\mu_1^T\Sigma_1^{-1}\mu_1) + \log(P(Y=0)/P(Y=1)) = 0$$

This is a quadratic equation in $x$, therefore the decision boundary is quadratic.