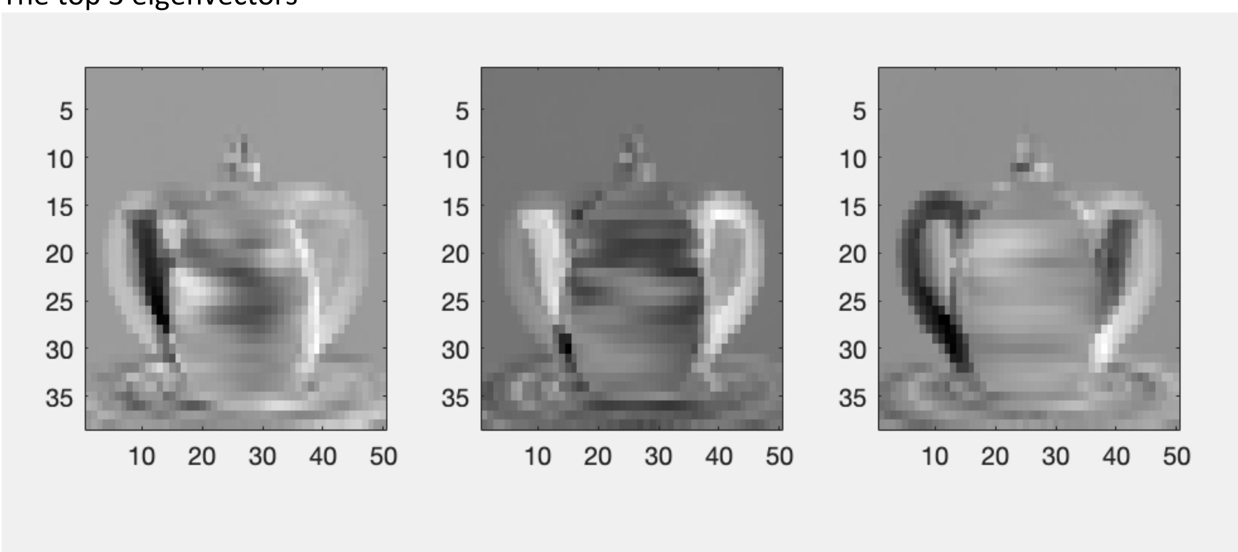
Jincheng Tian assignment 4

Problem. 1. (20 points)

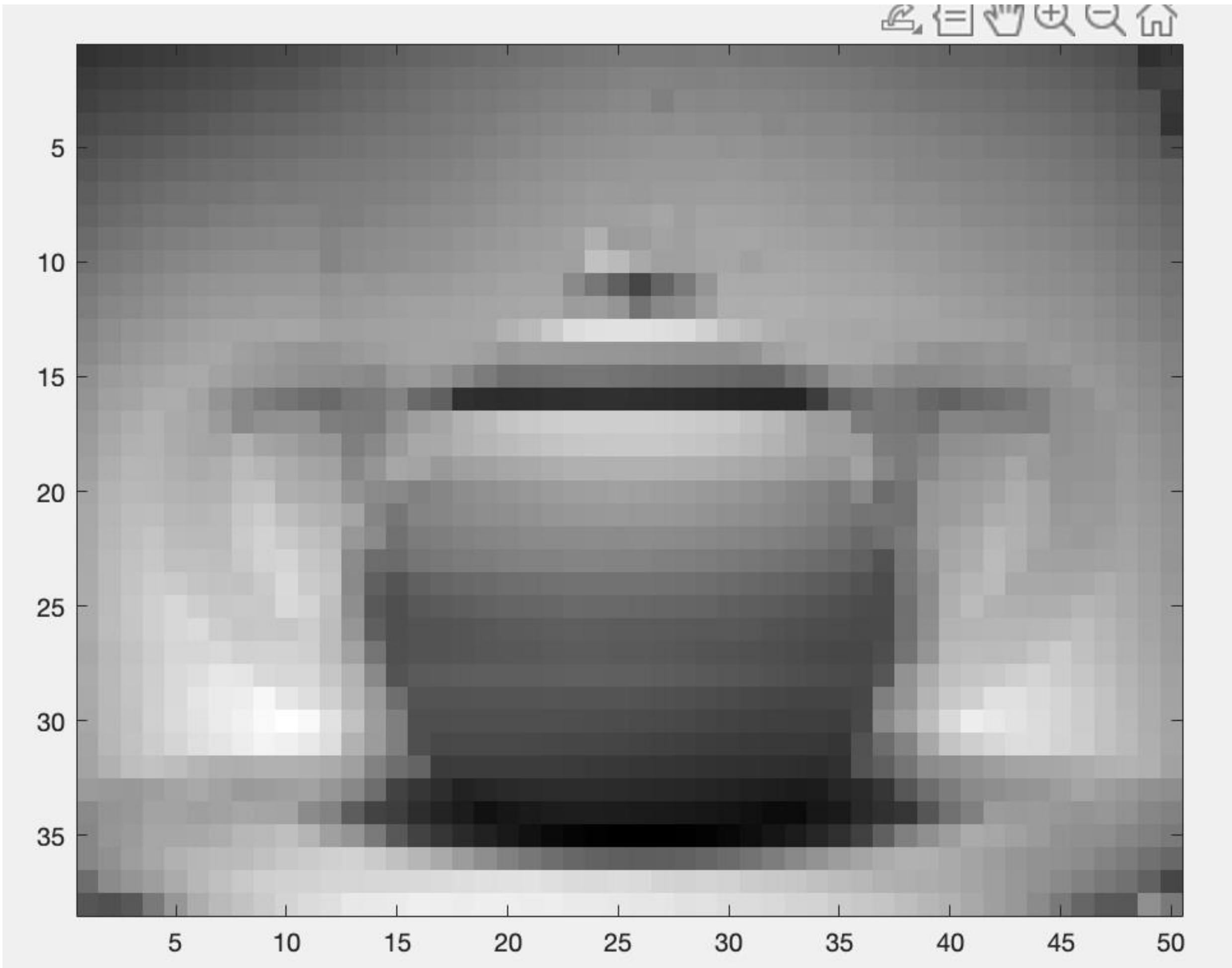
Download the “teapots.mat” data set containing 100 images of teapots of size 38×50. To view an image, say the second one in the data set type:
imagesc(reshape(teapotImages(2,:),38,50));
colormap gray;
Compute the data mean and top 3 eigenvectors of the data covariance matrix and show them as images. Reconstruct the data using PCA with least squares error using only the mean and a linear combination of the top 3 eigenvectors. Show 10 different images before and after reconstruction. Discuss results.

The figure after the PCA reconstruct is better and clearer. However, the figure before reconstruction is easily demonstrating the directions but the one after reconstruction is hard to see.

The top 3 eigenvectors



Data mean figure



10 different comparision before and after

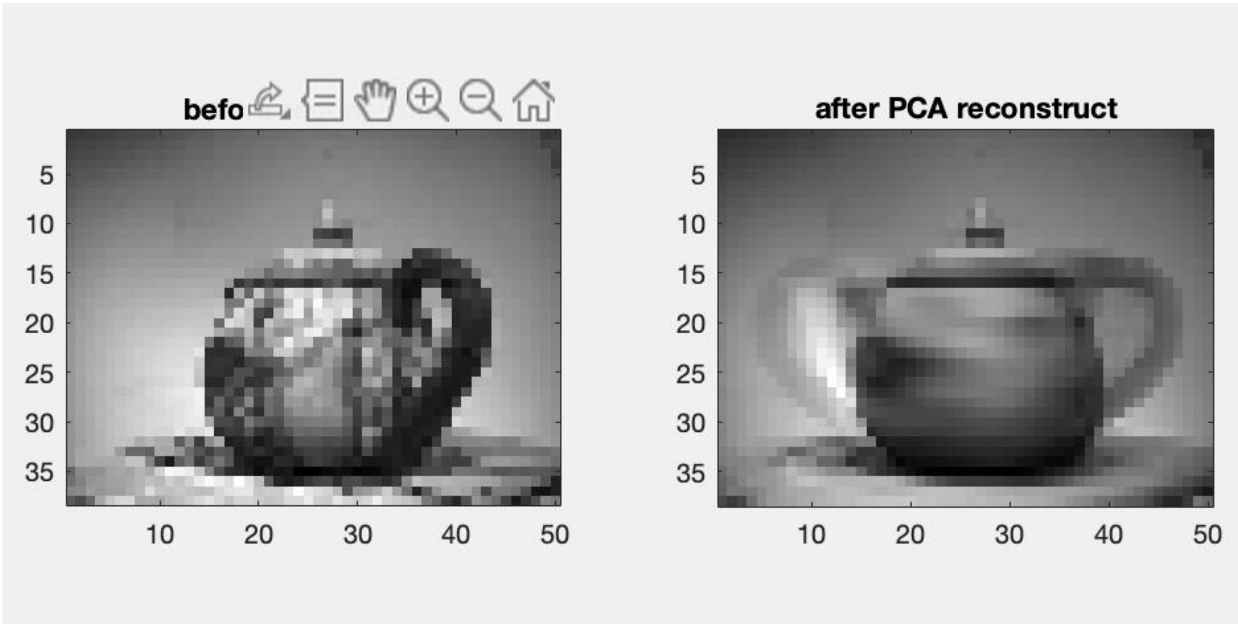


Figure-10

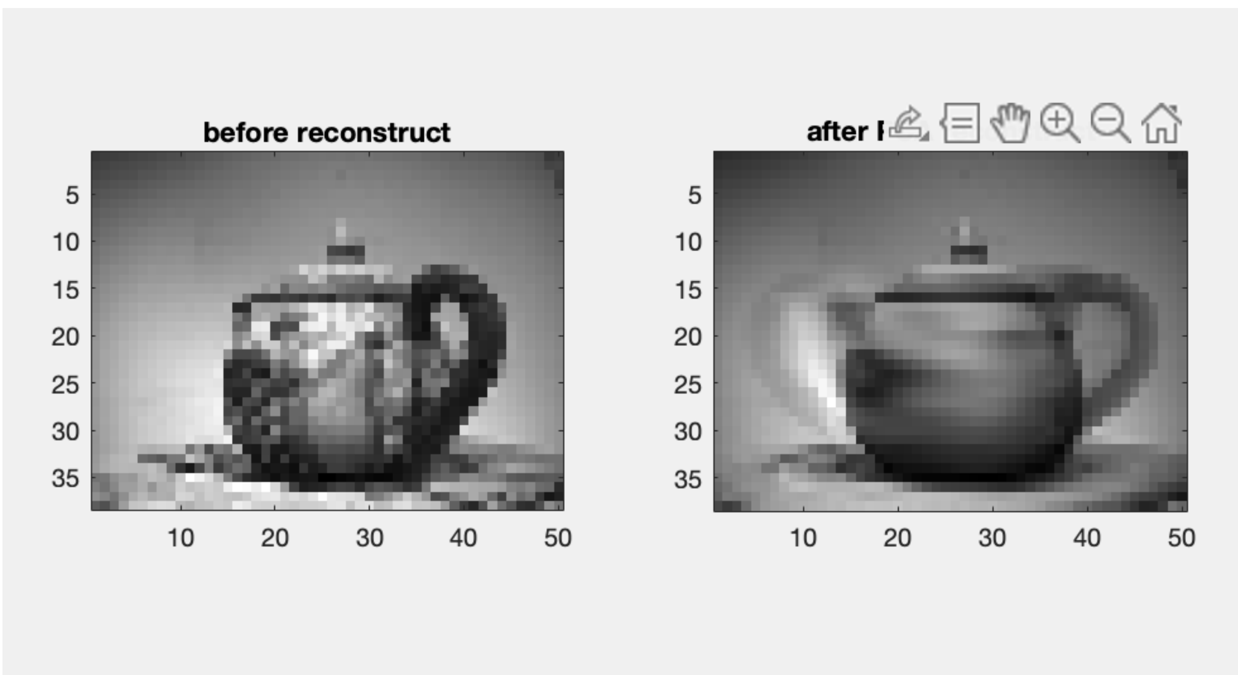


Figure -9

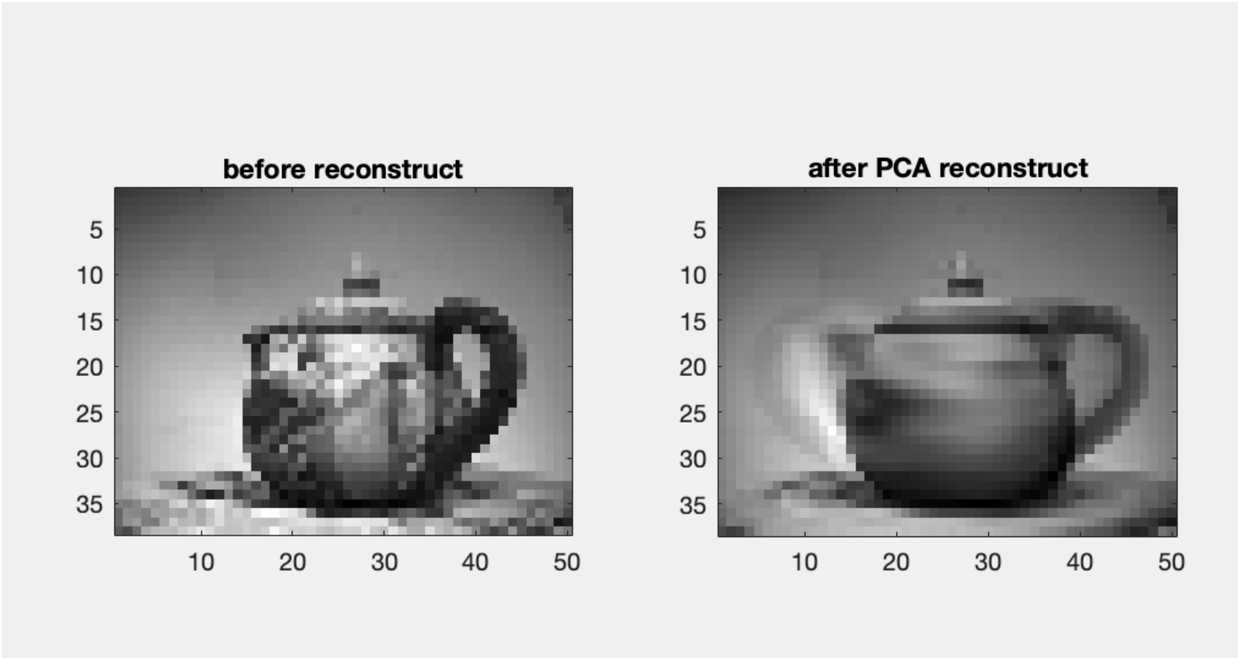


Figure -8

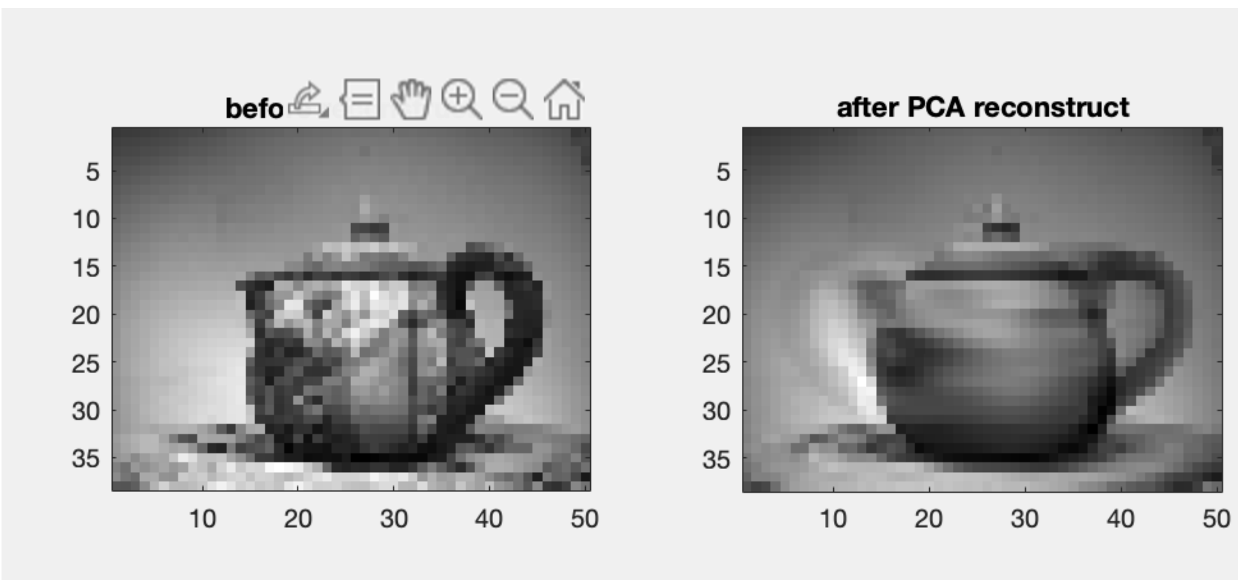


Figure -7

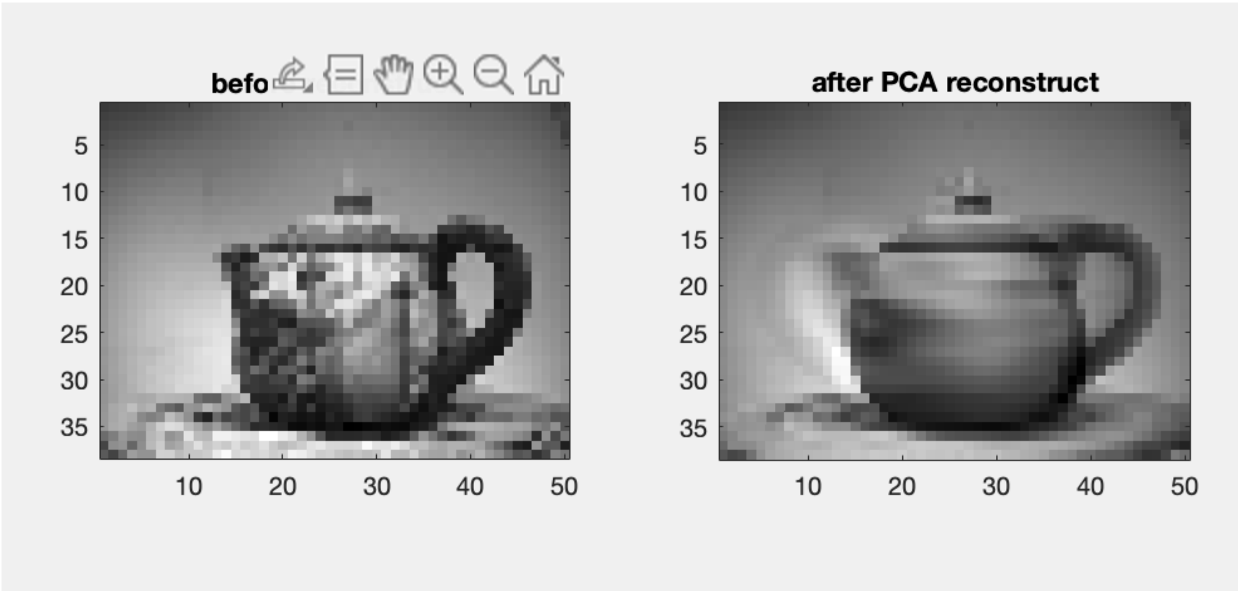


Figure -6

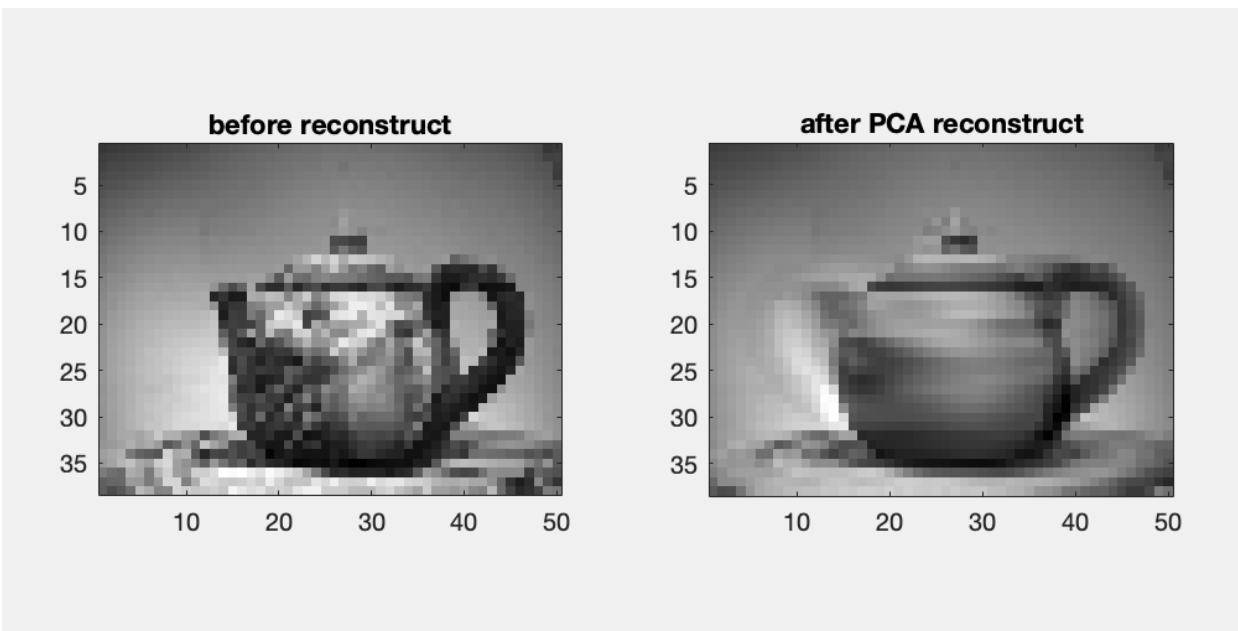


Figure -5

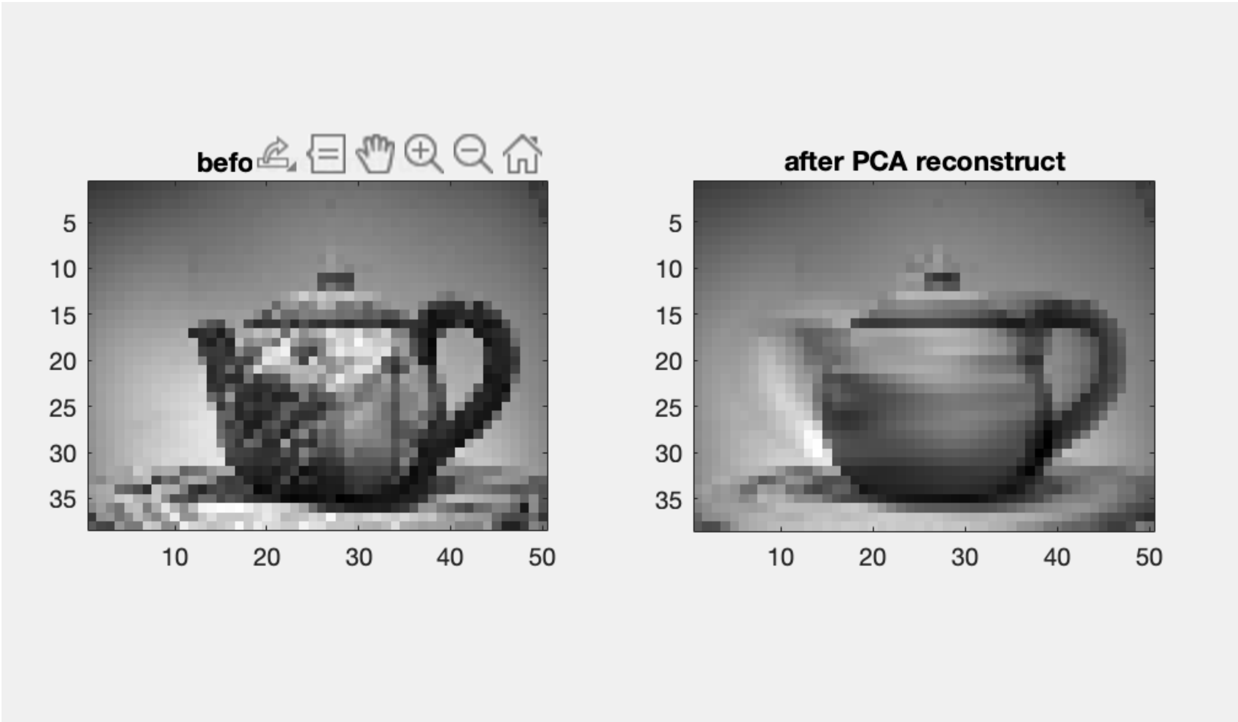


Figure -4

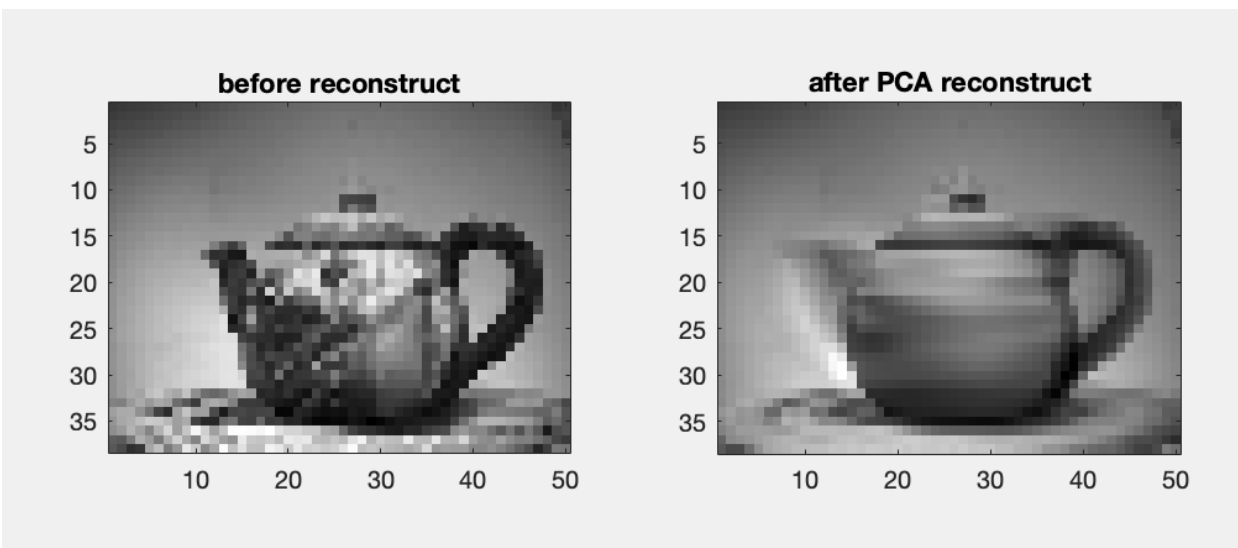


Figure -3

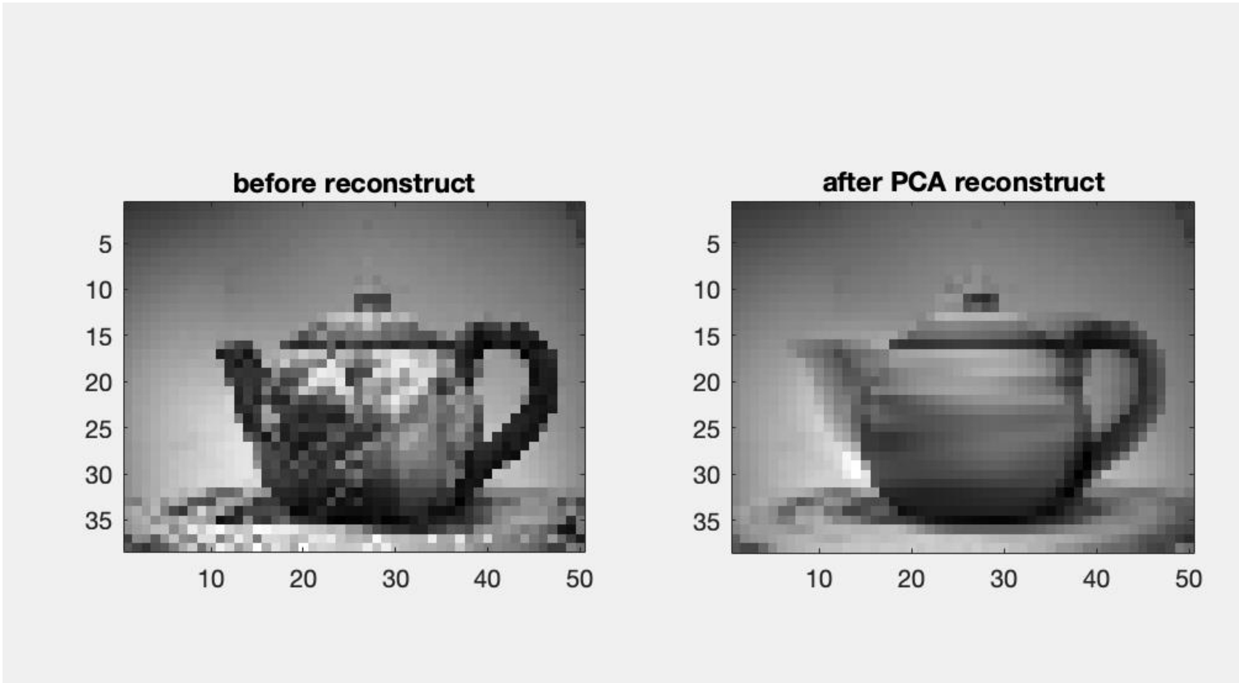


Figure -2

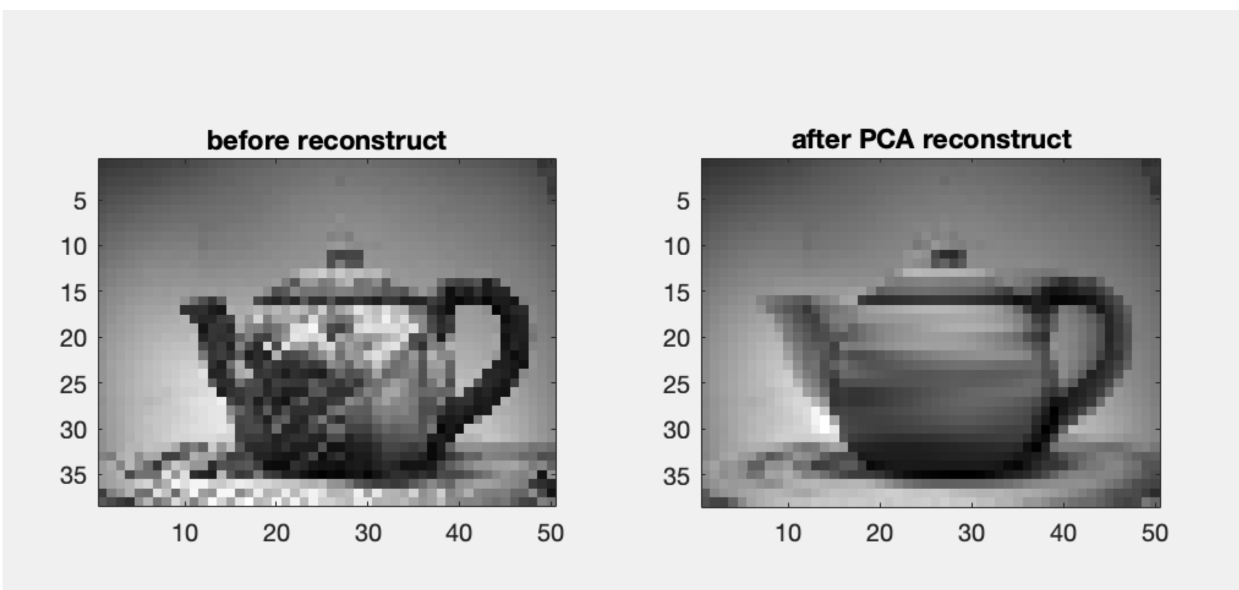


Figure -1

Problem. 2. (10 points)

Suppose we have a box containing 8 apples and 4 oranges and a second box containing 10 apples and 2 oranges. One of the boxes is chosen at random (with equal probability of choosing either box) and an item is selected from the box and found to be an apple. Use Bayes' rule to find the probability that the apple came from the first box.

Based on the baye's theorem

We have selecting box's probability P(event1) = P (event2) = ½

Then we have the P(select apple from box | event1) = 8/12 = 2/3

We have p(select Apple | event) = 10/12

Therefore, we have

P(event1 | select Apple) = P(event1) * p(select Apple | event1) / p(Event1) * P (select | E) + P (event2)* P(select apple | event2)

= (½ * 2/3) /(1/2 * 2/3 + ½ * 10/12) = 4/9

Problem. 3. (20 points)

Suppose you have a 2-class classification problem, where each class is Gaussia Let $\theta = \{\alpha, \mu_1, \Sigma_1, \mu_2, \Sigma_2\}$ denote the set of model parameters. Suppose th class probability $p(y|\theta)$ is modelled via the Bernoulli distribution, i.e. $p(y|\theta) =$ $\alpha^y(1-\alpha)^{1-y}$, and the probability of the data $p(x|y,\theta)$ is modelled as $p(x|y,\theta) =$ $\mathcal{N}(x|\mu_y, \Sigma_y)$. Recover the parameters of the model from maximum likelihoo approach. Assum the data are i.i.d. and N is the number of data samples. Show all derivation Next, suppose you want to make a classification decision by assigning prope label y to a given data point x. You decide the label based on Bayes optima decision $y = \arg\max_{\hat{y}=\{0,1\}} p(\hat{y}|x)$. Prove that the decision boundary is linea when covariances Σ_1 and Σ_2 are equal and otherwise the boundary is quadrati Illustrate on 2d example the decision boundary fo the case when covariances are not equal clearly indicating which class is mor concentrated around its mean.

① withe the given parameters $\theta = \{\alpha, \mu_0, \Sigma_0, \mu_1, \Sigma_1\}$

$\Rightarrow$ iid data from $p(x, y|\theta) = p(y|\theta)p(x|y, \theta)$

$\Rightarrow$ Recover the parameters from data using maximum likelihood

$$l(\theta) = \log p(\text{data}|\theta) = \sum_{i=1}^{N} \log p(x_i, y_i|\theta)$$

$$= \sum_{i=1}^{N} \log p(y_i|\alpha) + \sum_{y_i=0} \log p(x_i|\mu_0, \Sigma_0) + \sum_{y_i=1} \log p(x_i|\mu_1, \Sigma_1)$$

$$\mu_0 = \frac{1}{N_0} \sum_{y_i=0} x_i \qquad \Sigma_0 = \frac{1}{N_0} \sum_{y_i=0} (x_i - \mu_0)(x_i - \mu_0)^T$$

$$\mu_1 = \frac{1}{N_1} \sum_{y_i=1} x_i \qquad \Sigma_1 = \frac{1}{N_1} \sum_{y_i=1} (x_i - \mu_1)(x_i - \mu_1)^T$$

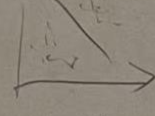
with the Bayes optimal Decision $\hat{y} = \operatorname{argmax}_{y=\{0,1\}} P(y|x)$

To get conditional

$$p(y|x) = \frac{P(x, y)}{p(x)} = \frac{P(x, y)}{\sum_y p(x, y)} = \frac{P(x, y)}{P(x, y=0) + p(x, y=1)}$$

$$= \frac{\alpha N(x|\mu_1, \Sigma_1)}{(1-\alpha) N(x|\mu_0, \Sigma_0) + \alpha N(x|\mu_1, \Sigma_1)}$$

If the covariance are equal
linear decision



If covariances are differt
quadratic decision