**Theory Questions**

1.

**<a>**

*Compute the mean:*

*Compute the Standard deviation:*

*Standardize data:*

*Find covariance matrix:*

*Find eigenvalues:*

*Find eigenvectors:*

for :

for :

**<b>**

*Project the data:*

2.

**<a>**

*Compute for feature 1:*

*Compute for feature 2:*

**<b>**

*Compare two features:*

we should prioritize feature 1.

**<c>**

*Standardize all the data:*

Using the same computing method in previous question, find out

, , ,

Then, the data becomes

and

*Compute the mean for each class:*

,

*Compute scatter matrices for each class:*

Using formula , where

*Within class scatter matrix:*

, by using row reduction

*Preform eigen-decomposition:*

*Eigen-values and eigen-vector:*

Compute for eigen-values, got the result .

Compute eigen-vector by using non-zero eigen-value : .

**<d>**

*Project the data of Class 1:*

*Project the data of Class 2:*

**<e>**

Consider the projected data as feature 3 and compute the information gain for that feature.

, , which is greater than .

For this feature, most data in class 1 are smaller and most data in class 2 are larger.

I conclude that the projection I performed seem to provide a good class separation.

**Dimensionality Reduction via PCA**

