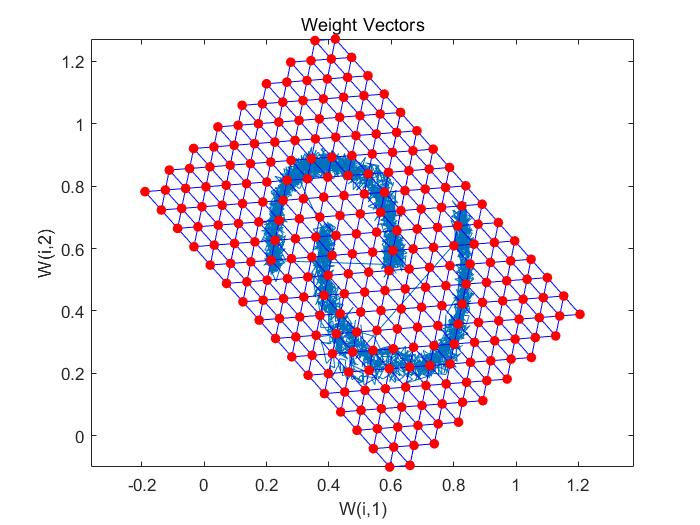
# 3.1 SELF-ORGANIZING MAP

Exercise1

The prototypes distribute in the feature space before training:



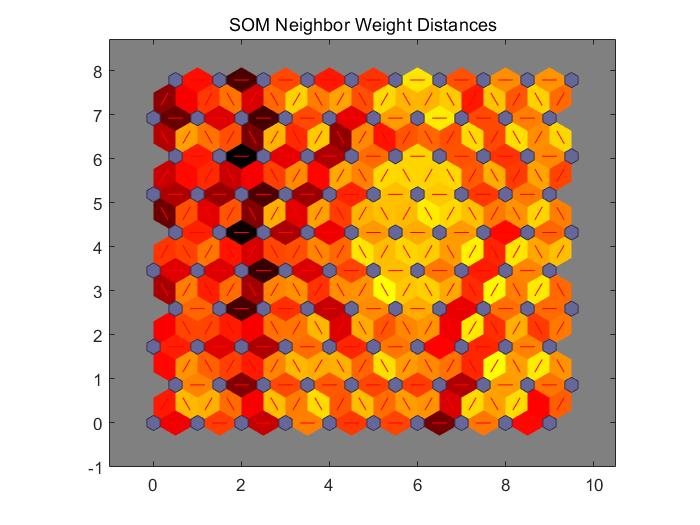
the final placement:

|  |
| --- |
|  |

Exercise2

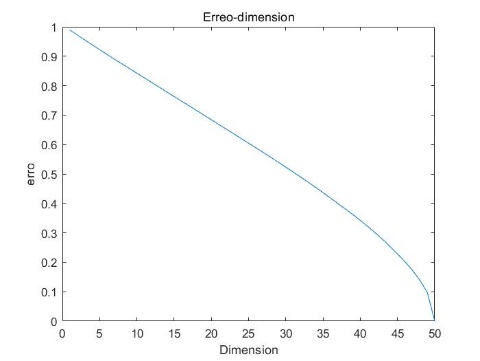
|  |  |  |  |
| --- | --- | --- | --- |
| grid size | topology | epochs | ARI |
| 10\*10 | hextop | 10 | 0.005 |
|  |  |  |  |
|  |  |  |  |

According to SOM Neighbor Distances, we can know distance between adjacent neurons. The deeper the colour, the longer the distance. The light colour line segment and the dark colour line segment on the graph can form an obvious boundary, which can be used as a reference for clustering.



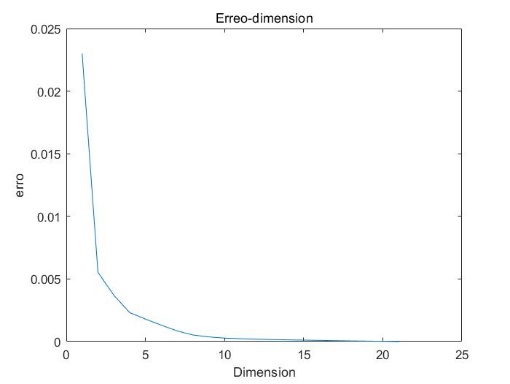
**3.2 Principal Component Analysis**

Exercise 1



With the increase of dimension, the error root mean square difference between the reconstructed and the original data) decrees gradually. The straight-line downward trend is relatively gentle and there is no rapid decline at a certain point because the data is generated randomly.

Exercise 2



In sharp contrast to the previous graph, the curve of error decline shows a rapid decline in some dimensions because the data in this dataset is highly correlated. This comparison also shows that PCA needs data to be related, and the effect of PCA with uncorrelated data will not be very significant.

# 3.3AUTOENCODER

Exercise 1

|  |  |  |
| --- | --- | --- |
|  | Number of neurons | number of training epochs |
| Autoenc1 | 50 | 1000 |
| Autoenc2 | 50 | 1500 |
| Autoenc3 | 100 | 1000 |
| Autoenc4 | 100 | 1500 |
| Autoenc5 | 200 | 1000 |
| Autoenc6 | 200 | 1500 |

