Home Work Assignment 2

CS365

Aayush Mudgal [12008]

A. Linear dimensionality Reduction: Principal Components

Lower Dimensional Embedding (Principal Component Analysis, dimension=2)

Original Image (converted to grayscale)

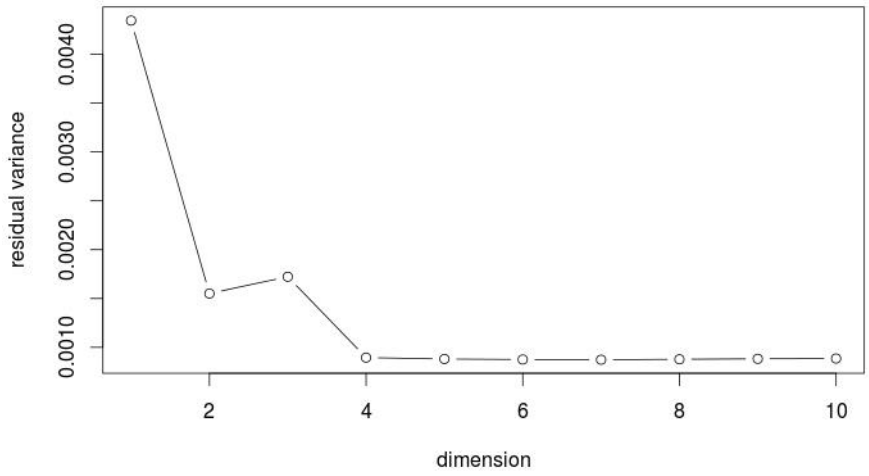
Reconstructed Image with m=2

Reconstructed Image with m=10

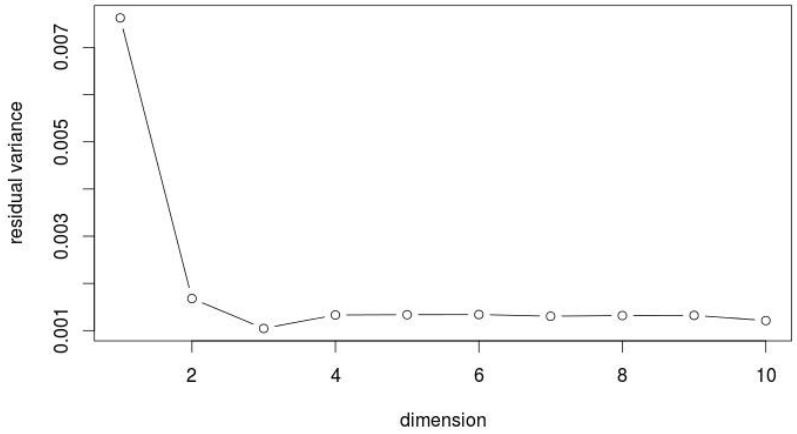
Reconstructed Image with n=30

Reconstructed image with n=80

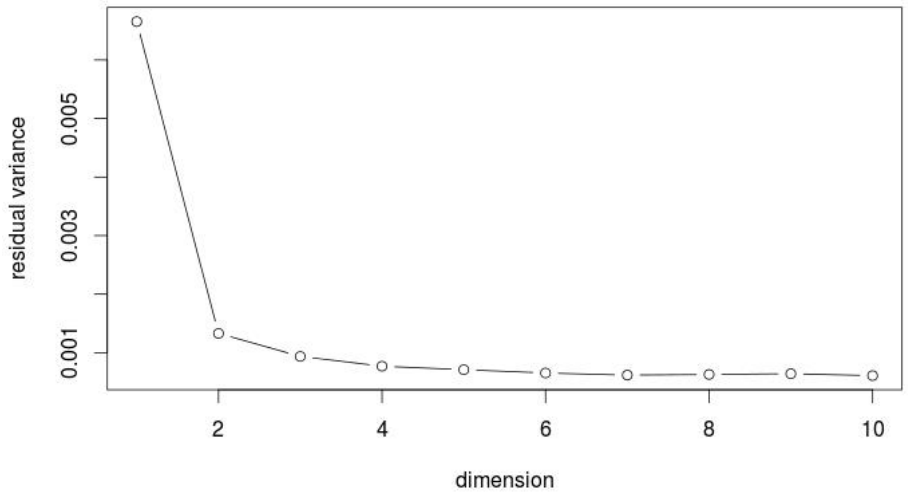
B. Nonlinear Dimensionality Reduction (Isomap)



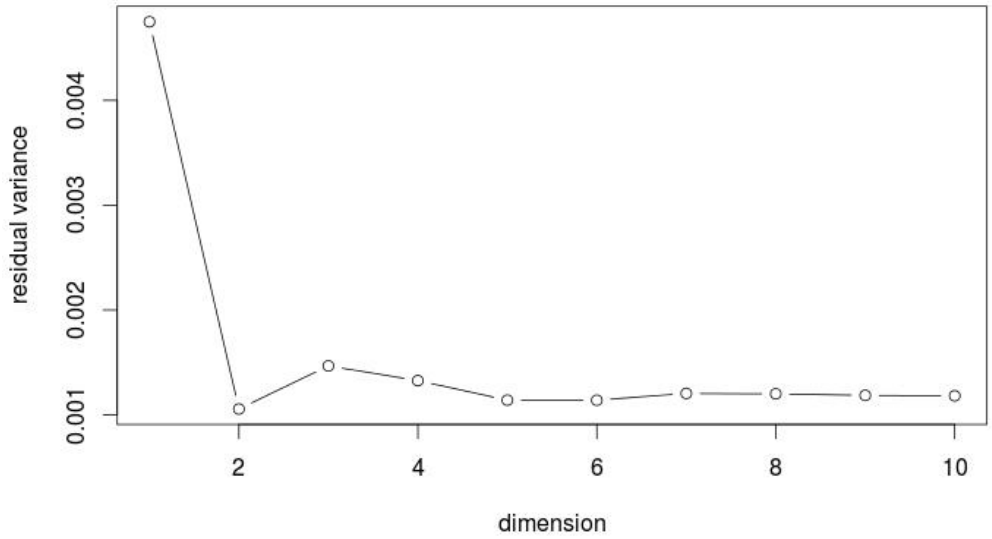
Residual Variation versus number of dimensions with 2 as the nearest distance



Residual Variation versus number of dimensions with 3 as the nearest distance

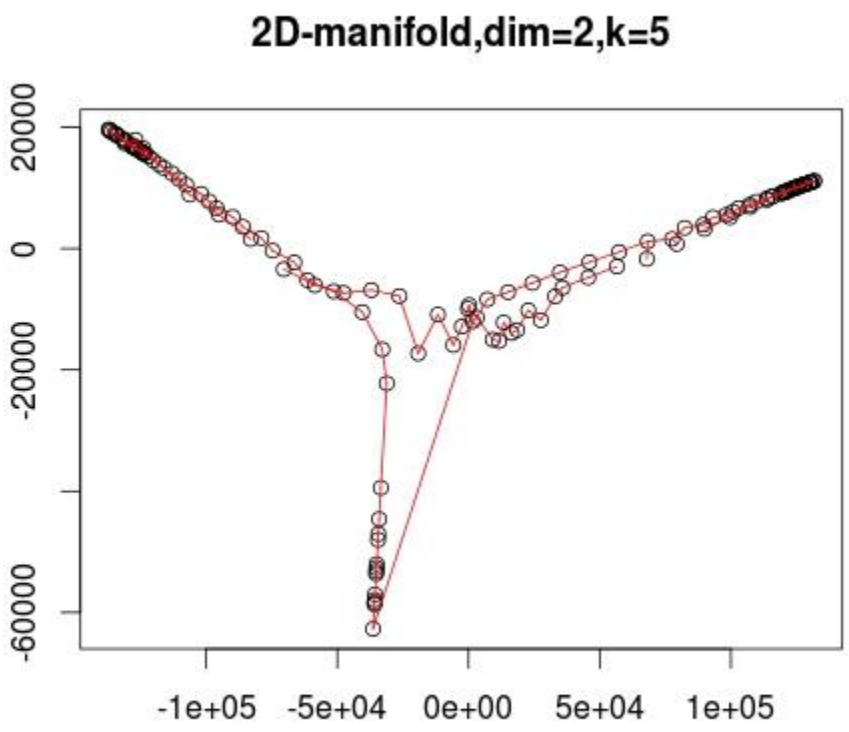


Residual Variation versus number of dimensions with 5 as the nearest distance

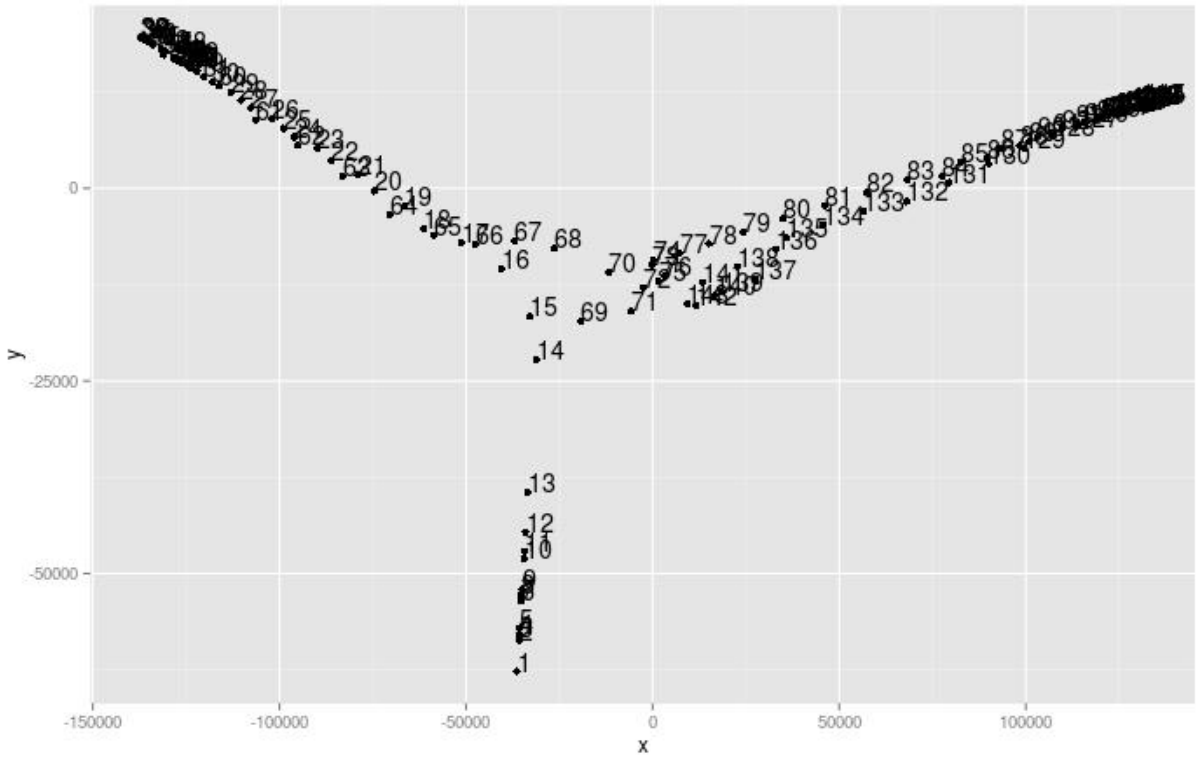


Residual Variation versus number of dimensions with 8 as the nearest distance

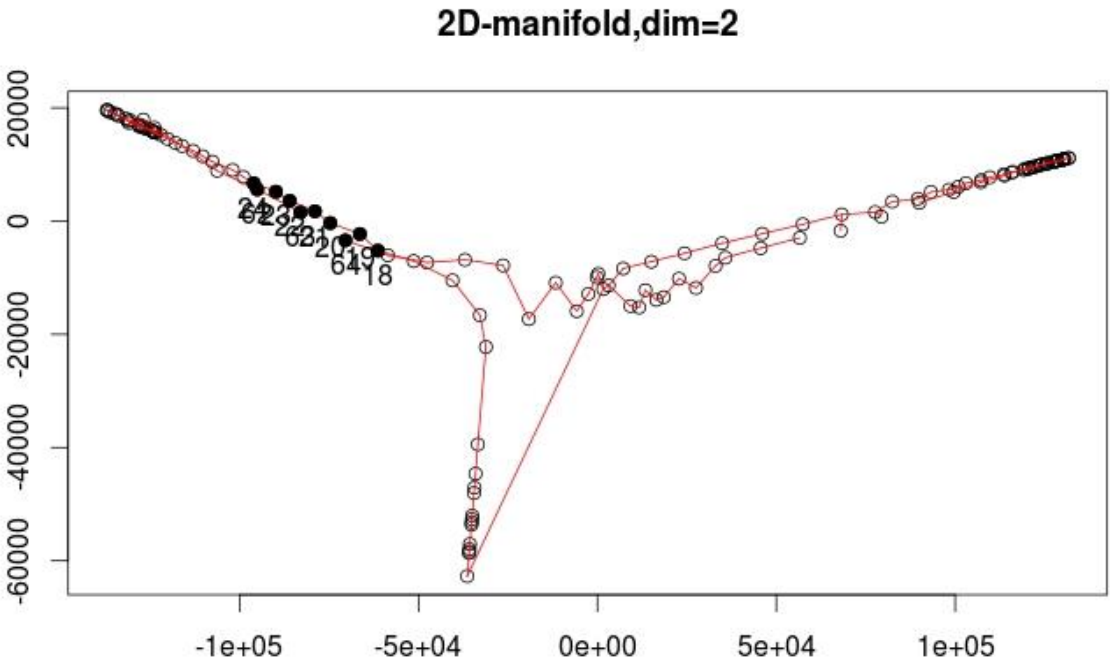
Isomap 2D Plots



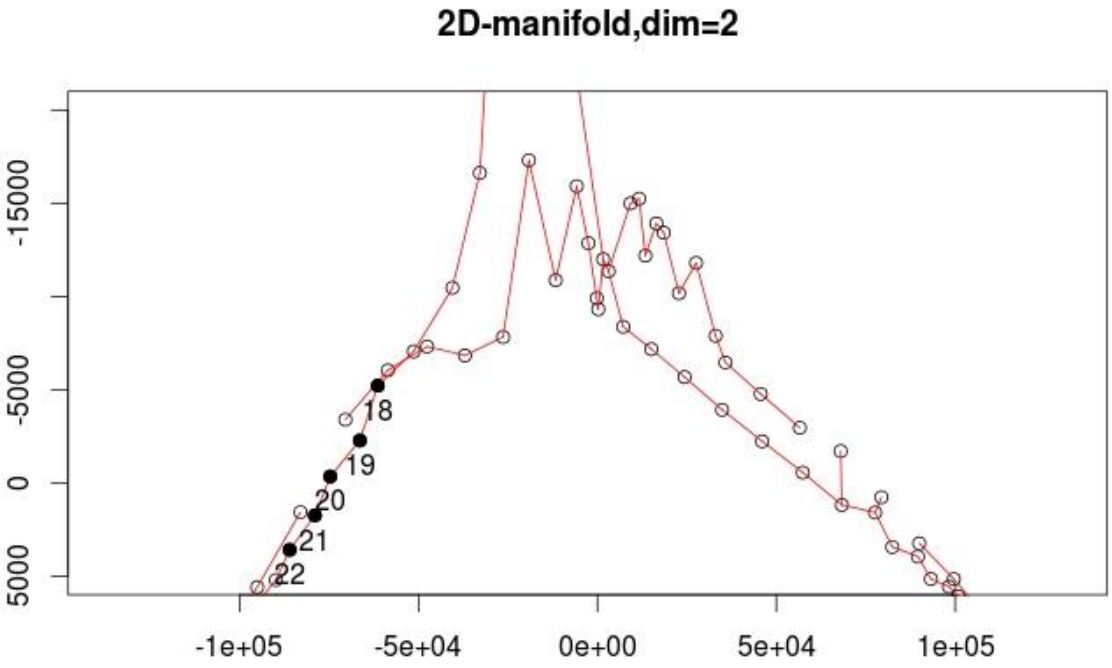
Lower Dimensional Embedding (Isomap, dimension=2)



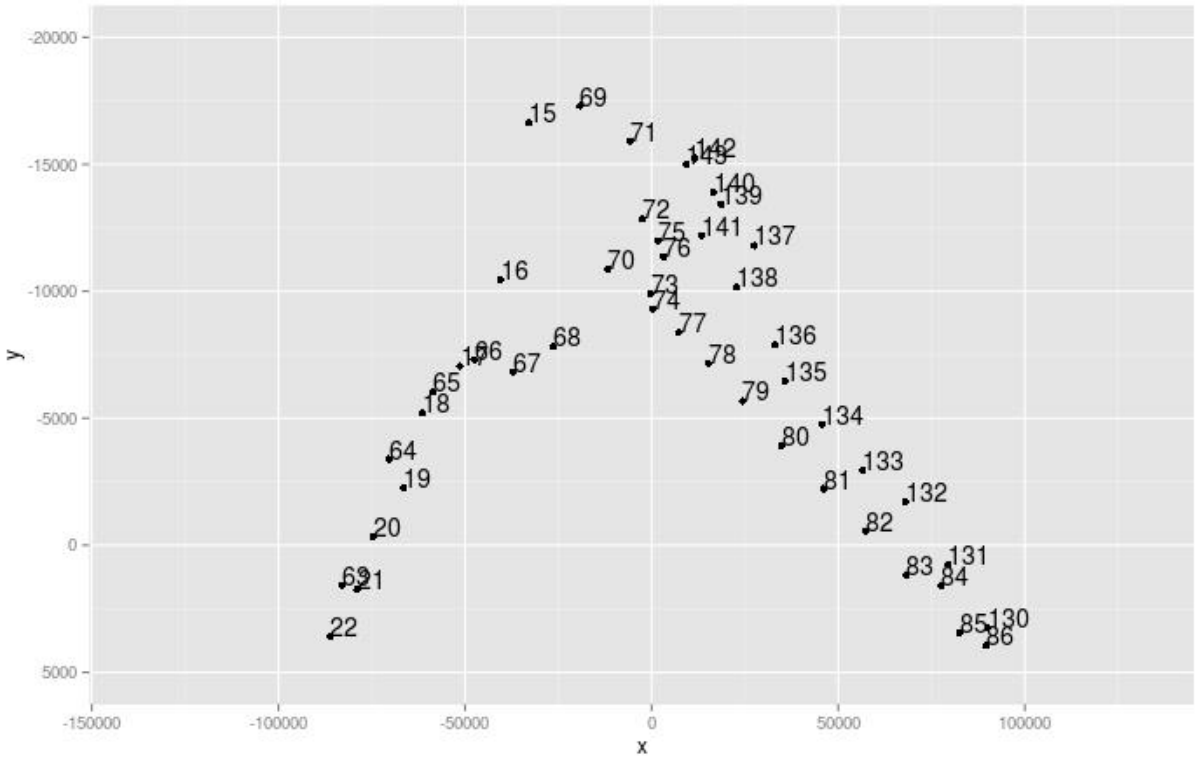
Lower Dimensional Embedding (Isomap, dimension=2) with nodes labelled.



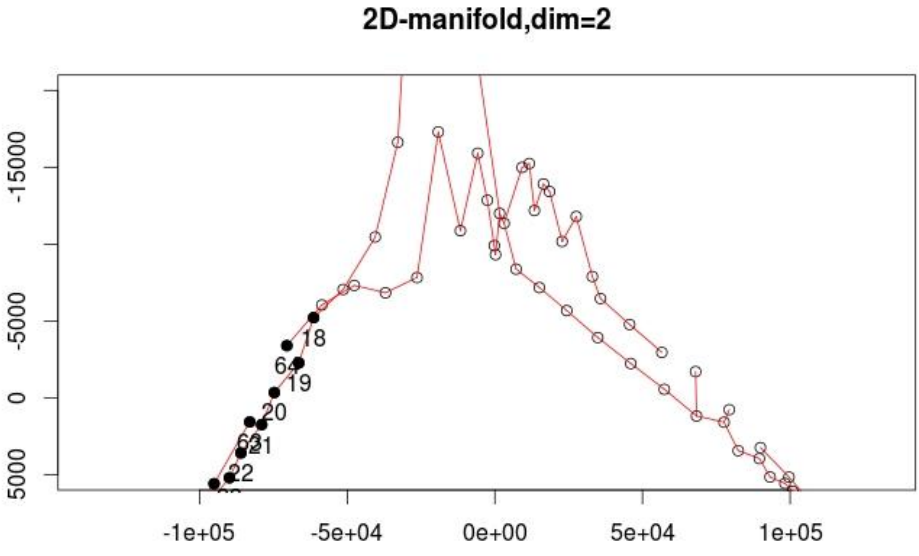
10 nearest neigbhours of image 25 being shown on the **Lower Dimensional Embedding (Isomap, dimension=2)**



5 nearest neighbors of the image 25 on the same Isomap plot (zoomed)



The same plot Zommed



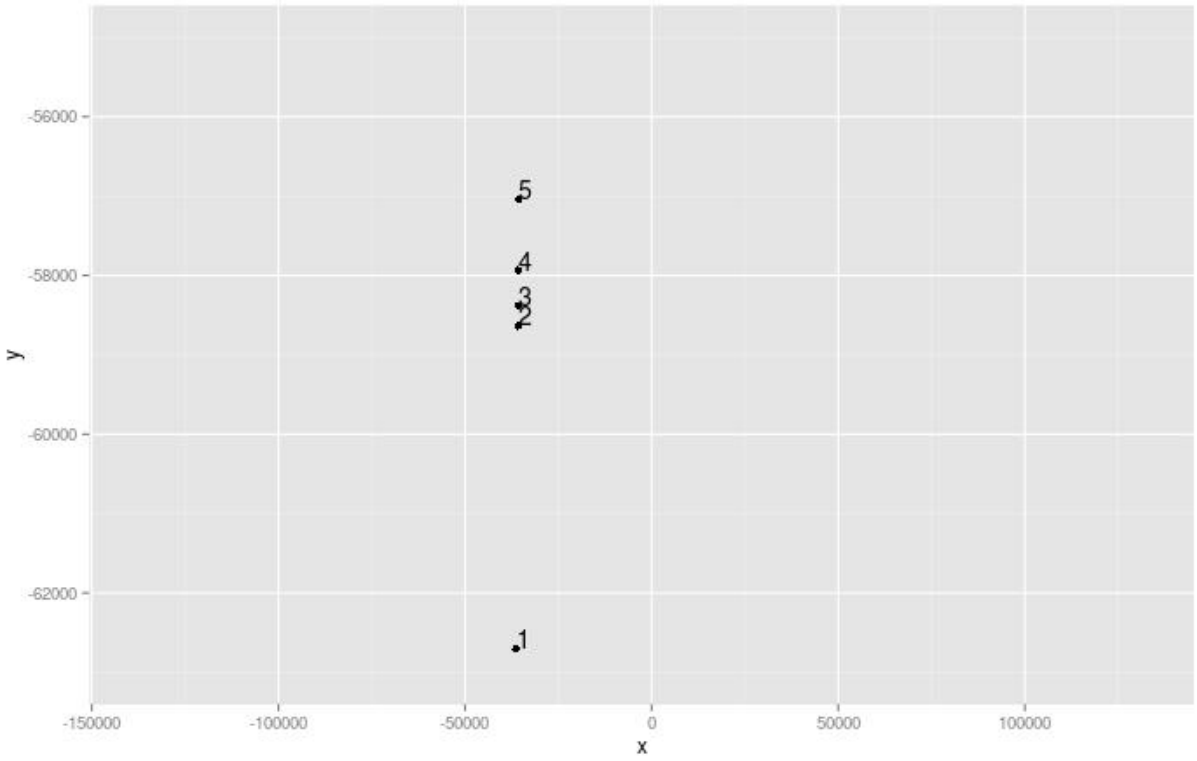
5 nearest neighbors of the image 25 on the same Isomap plot (zoomed)

Image 25: 

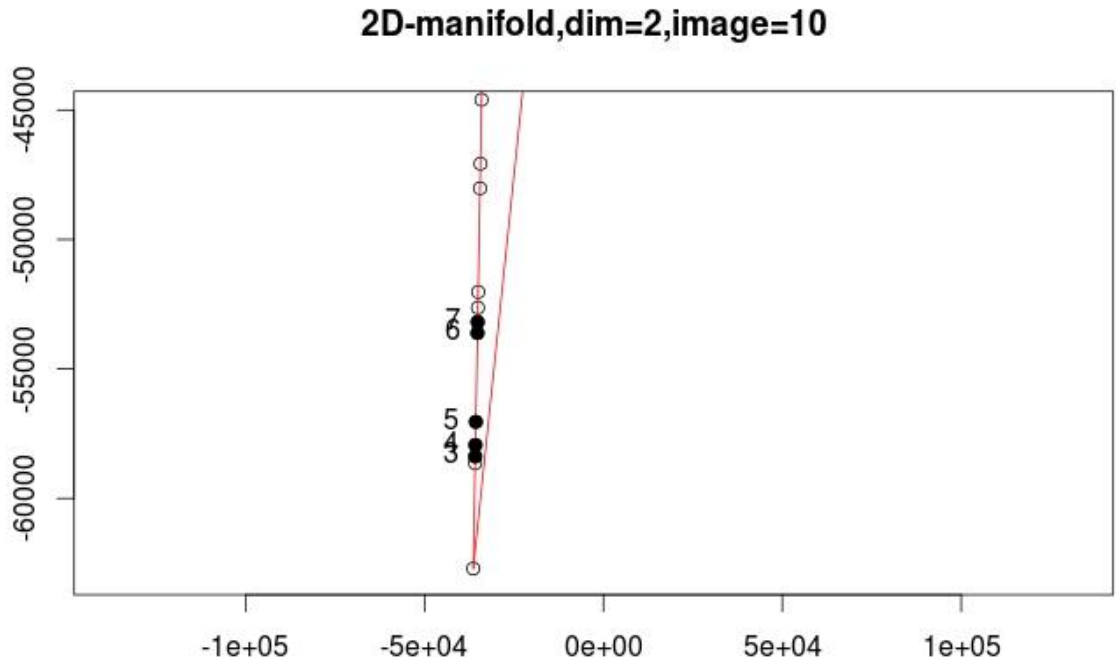
5 nearest neighbors of image 25:

26 24 27 23 (indexing is as per the code)

Decreasing Similarity



Zoomed around image 10 (serial number =5)



5 nearest neighbors of the image 10 on the same Isomap plot (zoomed)

Image: 10: 

5- Nearest neighbors of Image 10 are

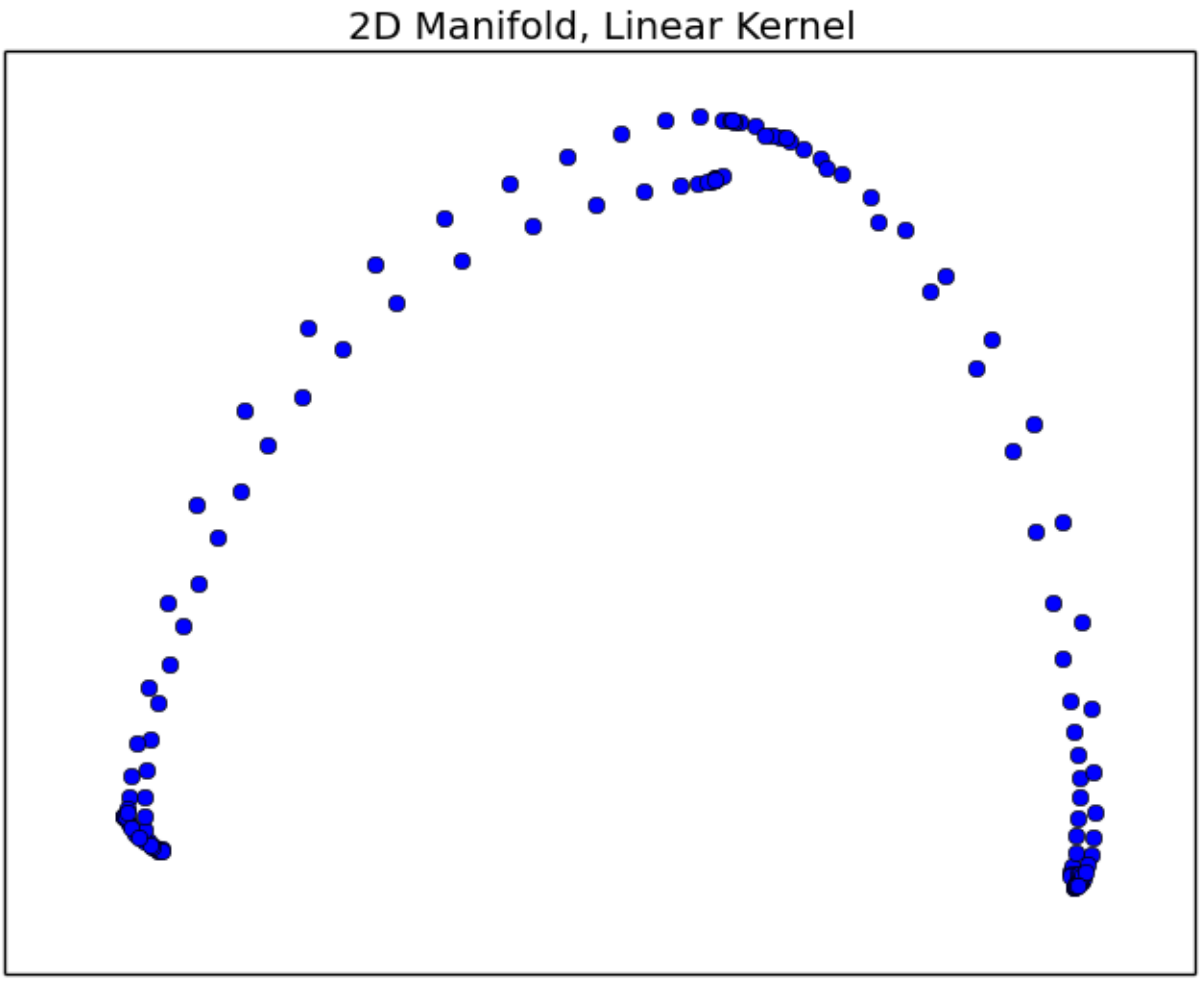
Decreasing Similarity

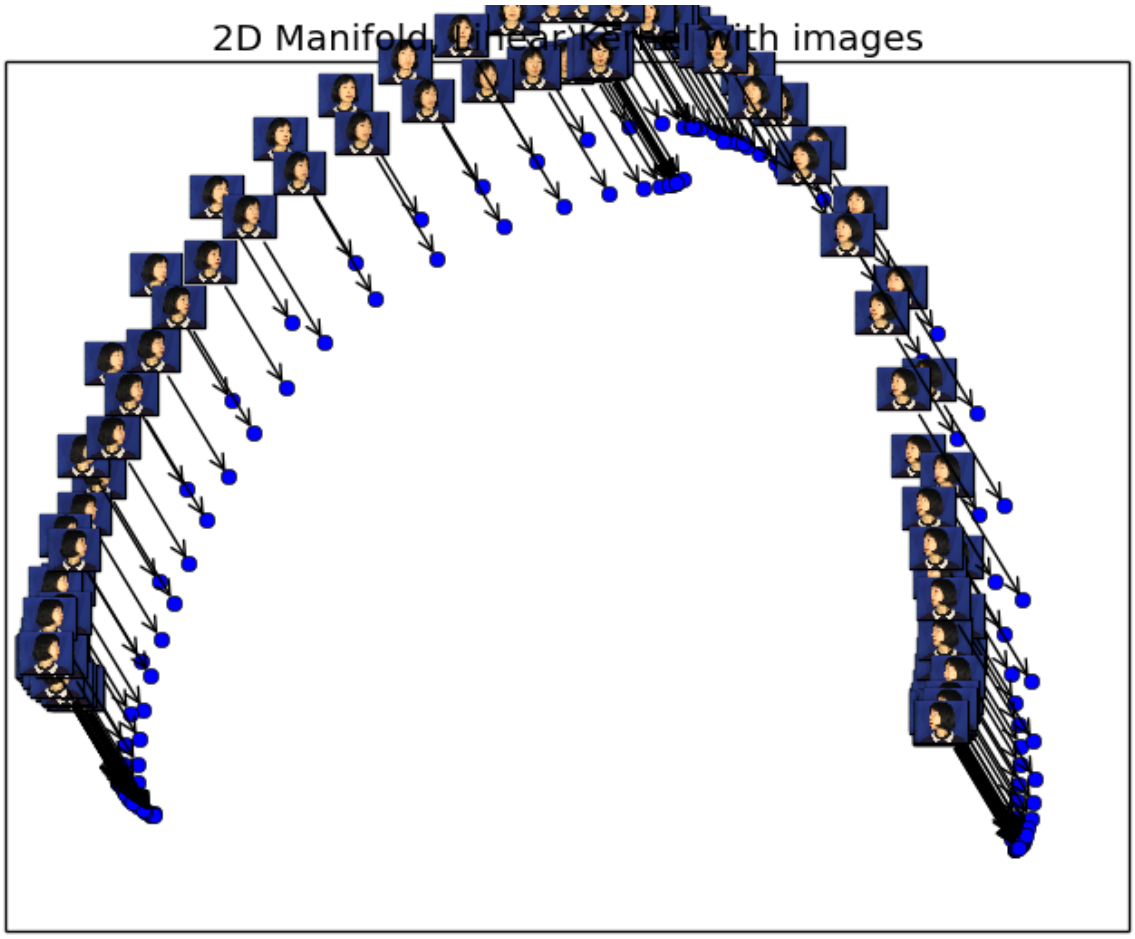
Nature of Manifold: Observations and Explanations

The image dataset is such that it is a continuous rotation of the head of the lady subject. Thus ideally the manifold would have been one dimensional (the angular rotation of the head). Isomap is a non-linear dimensionality reduction technique and represents the data in a two dimensional embedding, retaining the features of the dataset. The image that are more similar are located close to each other and are far separated from the images that are different. The manifold plot shows three different and distinct lines. The vertical line corresponds to the images wherein the head is still, the other line captures the point when the head is moved right while the other directional line captures the images when the head is turned left. Thus it gives a good two dimensional embedding for the given high dimensional image dataset.

Part C: Kernel PCA

1. **Linear Kernel**



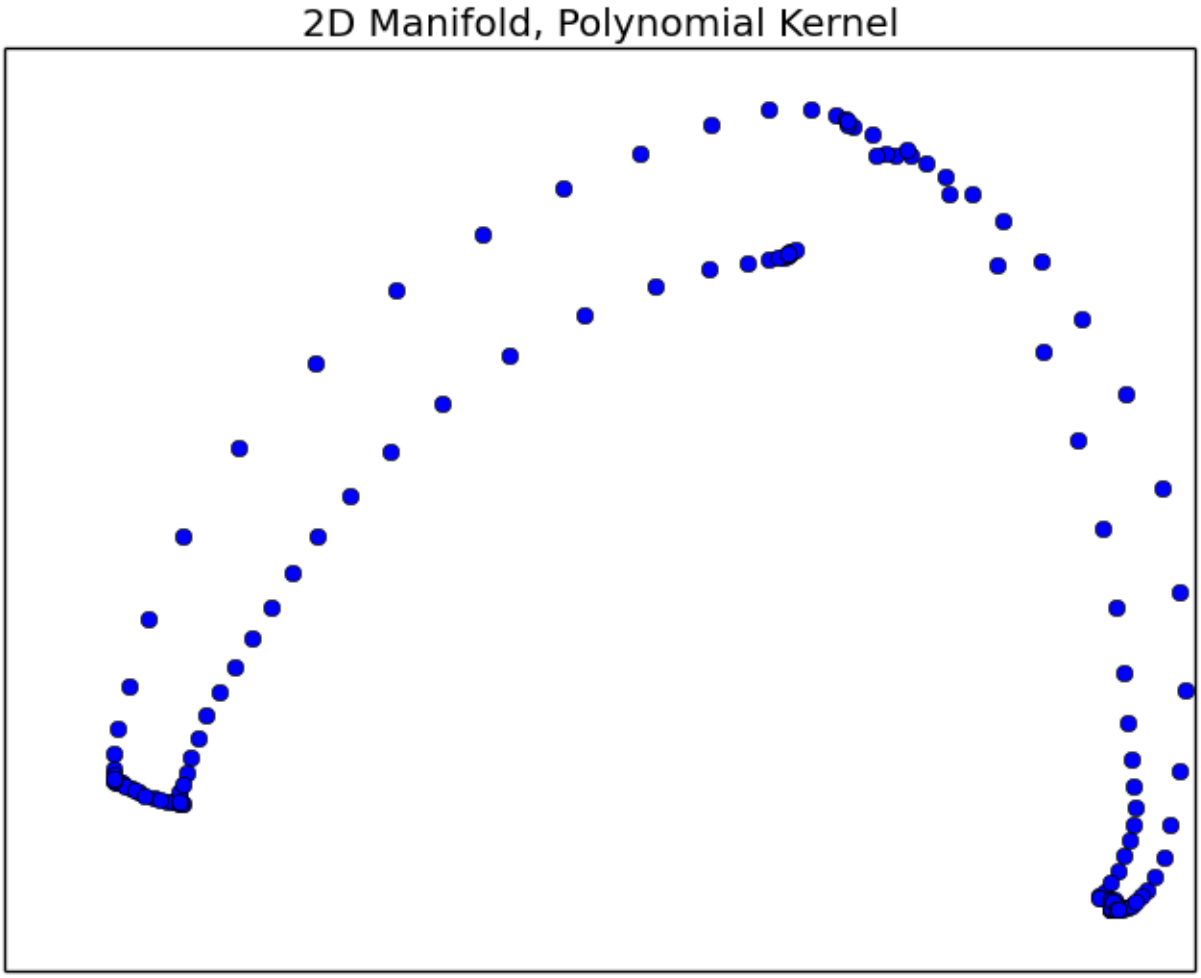


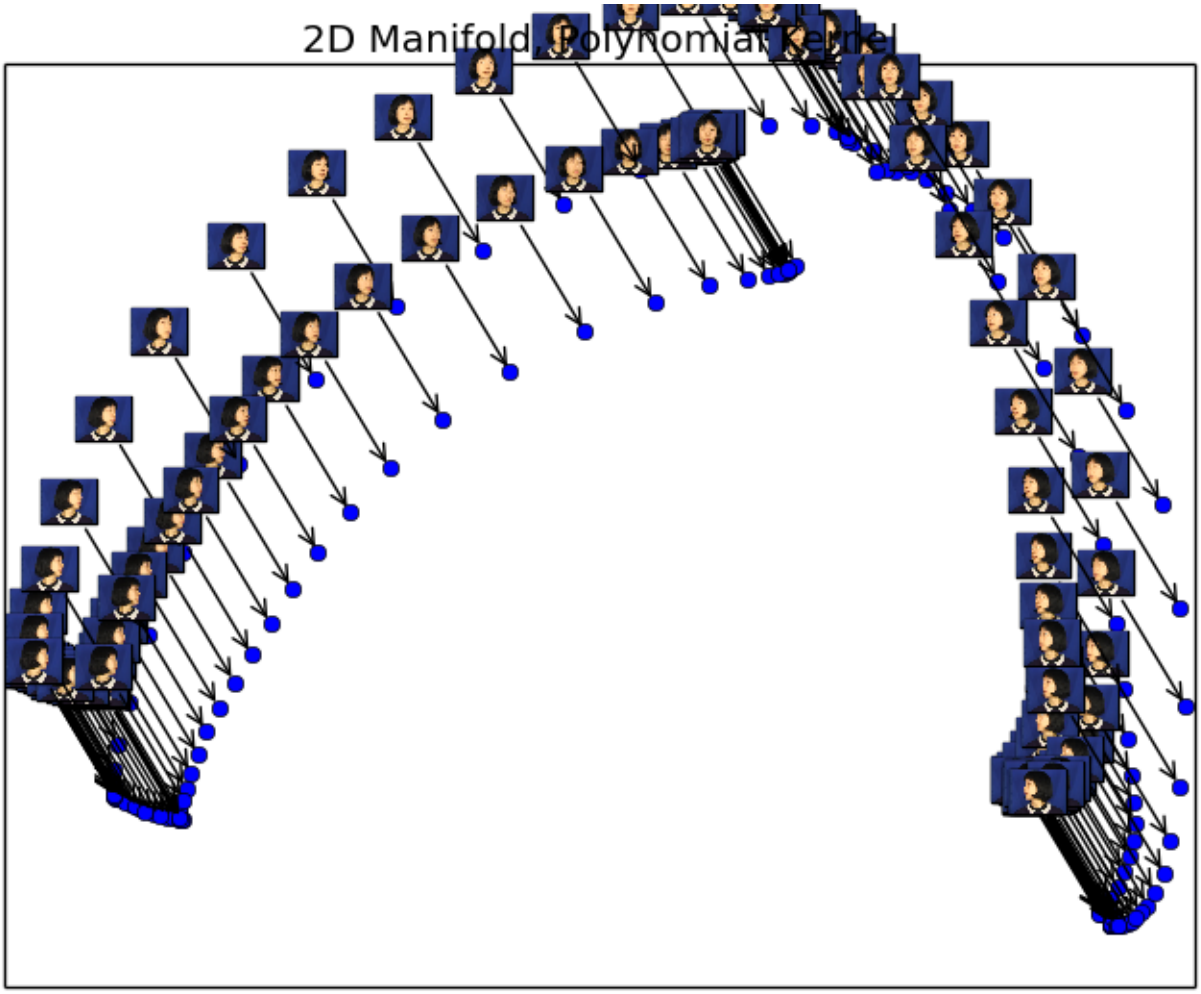
5 nearest neighbors for image 25:

Original Image 25: 

Increasing Distance

2. Polynomial Kernel

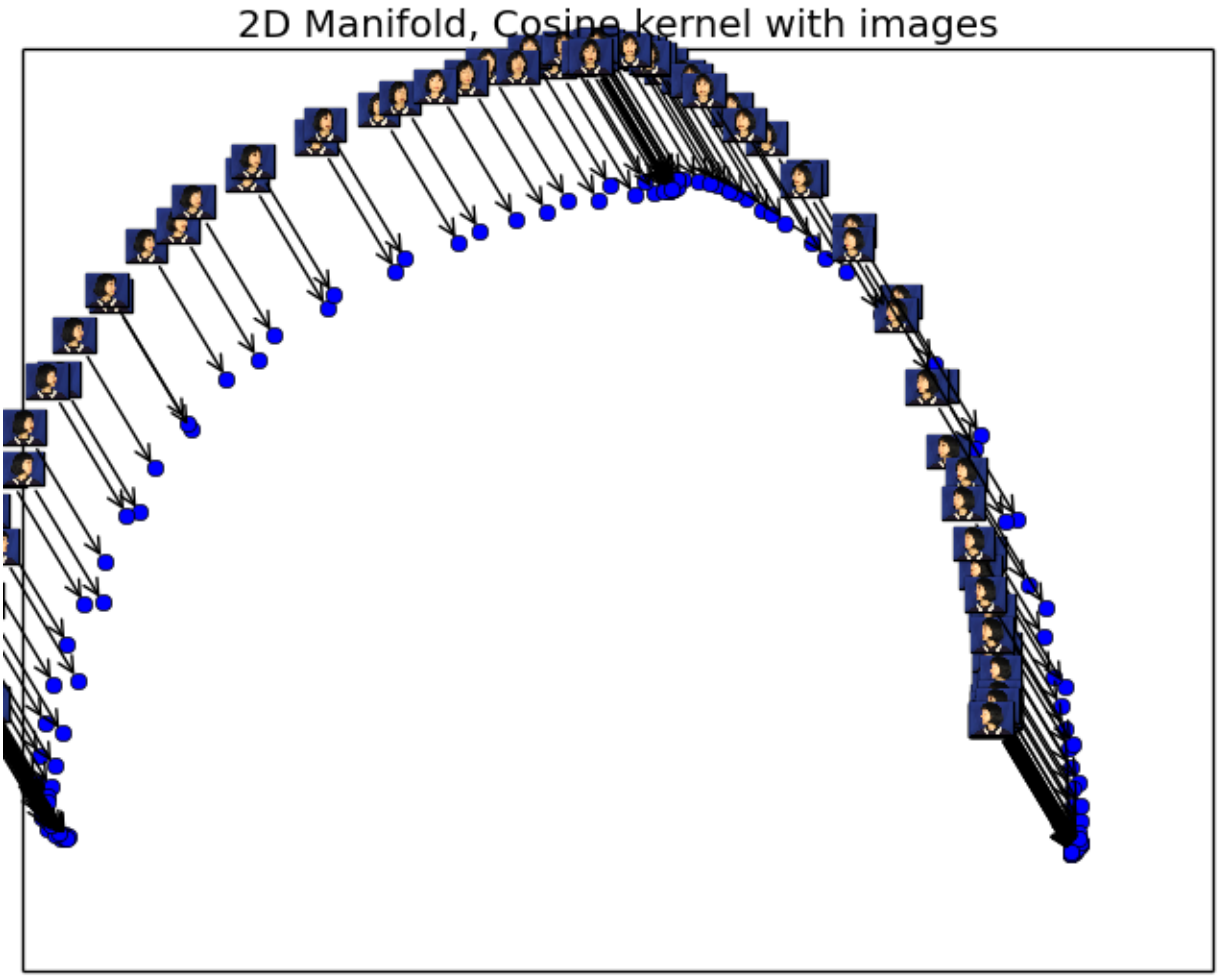




Original Image 25:

Increasing Distance

3. **Cosine Kernel**



Original Image 25:

Increasing Distance

Observations and Explanations

The linear, polynomial and cosine kernels, the lower dimensional captures the similarity across the image dataset. KNN (k nearest neighbors) algorithm also proves that the lower dimensional embedding during various Kernel PCA (using different kernels) preserves the similarity in the database. Thus it very effectively reduces the multi-dimensional dataset into lower dimensional space.