G1 Data Set

Principal components analysis did effectively retain the distinct classes that the G1 data had originally. It is easy to see, even by eye, that the data were and still are in three distinct classes. PCA simply reduced the dimensionality of the data.

Reducing 10 dimensions to 2 dimensions

Reducing 10 dimensions to 1 dimension

G2 Data Set

Principal components analysis did effectively retain the distinct classes that the G1 data had originally. It is easy to see, even by eye, that the G2 data were and still are in three distinct classes. However, reducing the data to only 1-dimension makes it impossible to separate the blue class from the green. In this regard, PCA isn't favorable.

Reducing 3 dimensions to 2 dimensions

Reducing 3 dimensions to 1 dimension

G3 Data Set

In this case, PCA doesn't help much to separate out the three classes of this G3 data set. The inherent structure of the data is well-preserved in that PCA reduced the dimensionality without collapsing the data into a narrower region than before. In other words, PCA simply takes a "bird's-eye view" of the data in this case. But it doesn't help much.

Reducing 3 dimensions to 2 dimensions

Reducing 3 dimensions to 1 dimension

Halfmoons Data Set

PCA again does a fine job at isolating the two data classes from each other in two dimensions, but it fails to reveal the high variability between the red and green halfmoons in dimension 3. It is almost no use to reduce the data to 1-dimension in this case.

Reducing 3 dimensions to 2 dimensions

Reducing 3 dimensions to 1 dimension

Swissroll Data Set

Reducing 3 dimensions to 2 dimensions

Reducing 3 dimensions to 1 dimension

PCA once again does a fine job at retaining the inherent variability in the swissroll data set in 2-dimensions, but there's virtually no information to gain at the 1-dimensional reduction. Concerning the goal of this data set, it is true that farthest neighbors are still kept far from each other. However, nearest neighbors blend in on top of one another. PCA doesn't reveal nearest neighbors!

Seeds Data Set

Reducing 7 dimensions to 2 dimensions

Reducing 7 dimensions to 1 dimension
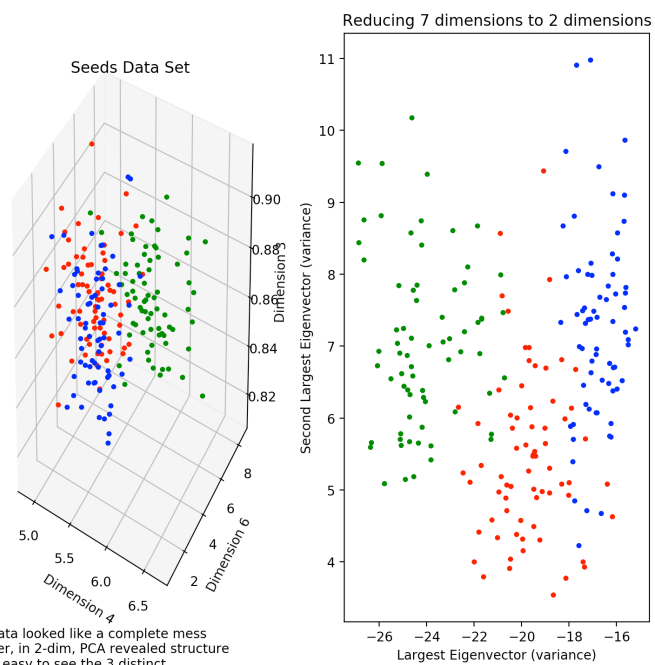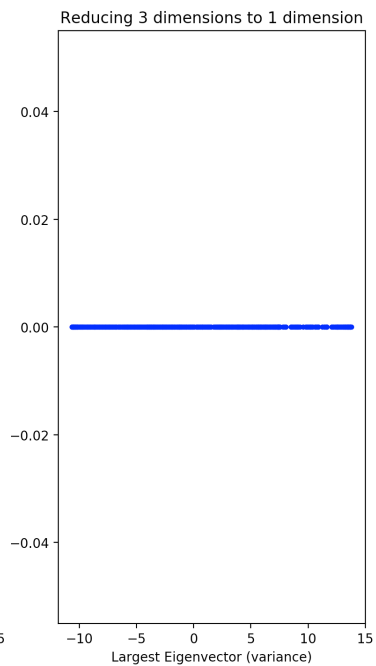
The seeds collection of data looked like a complete mess in the 3-dim plot. However, in 2-dim, PCA revealed structure from the chaos! It is now easy to see the 3 distinct classes.