Michael Miller

Akshay Shetty

CS 498 – HW4

3/1/16

**Problem 3.4**

**a)**

Iris data is plotted in a 2d scatterplot matrix (Figure 1).

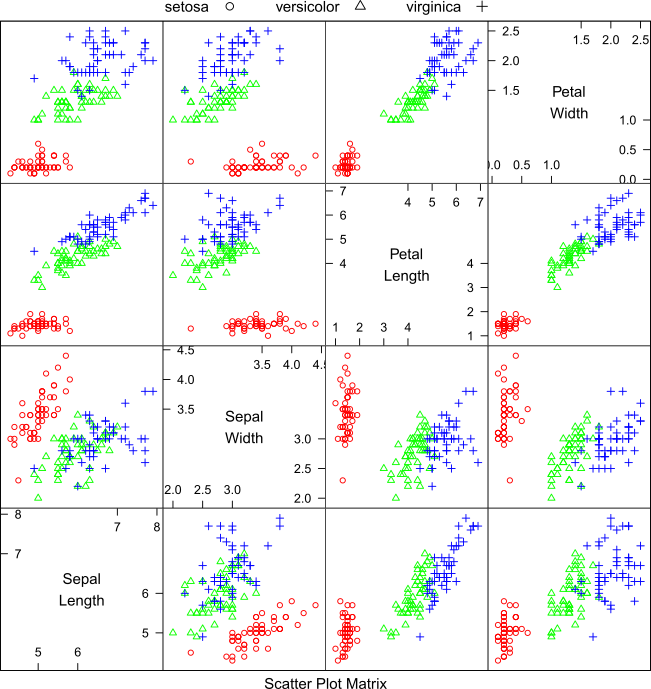
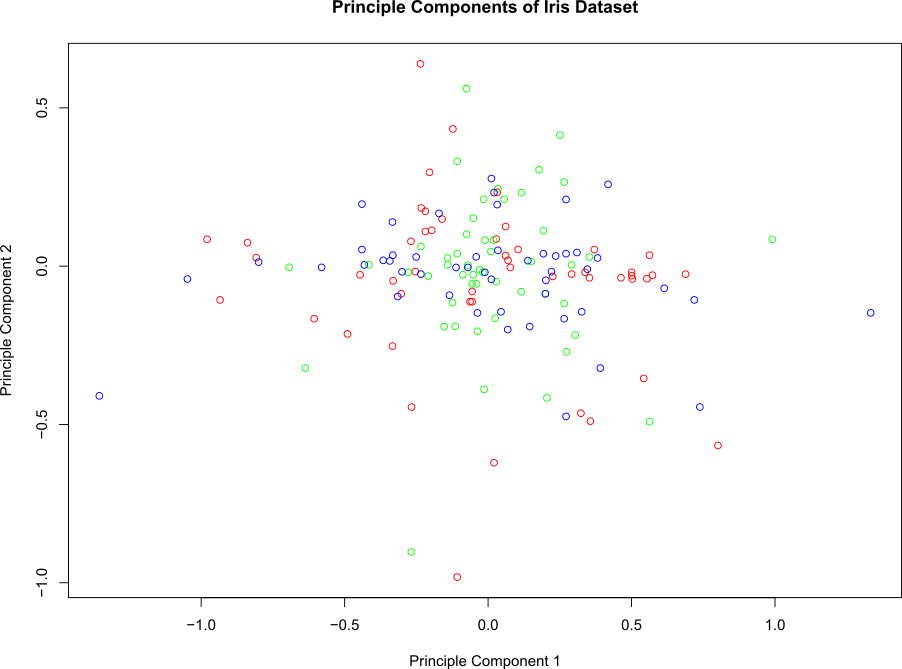


Figure 1. Iris data potted in a two-dimensional scatter plot

**b)**

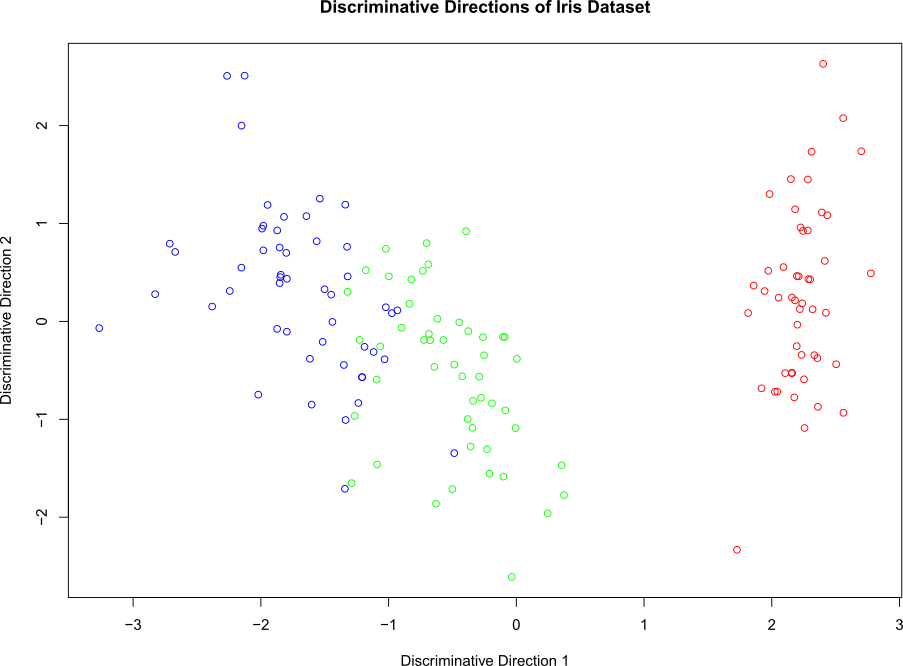
Principle components were found using the R function princomp() in the lattice package. The first two principle components are plotted blow (Figure 2). Significant distortion is introduced into the data. In particular, the mostly distinct Iris types now overlap.



**Figure 2. Principle components of Iris data**

**c)**

Discriminative direction were calculated using the pls2\_nipals function of the chemometrics R package. Figure 3 displays plots of the first and second Discriminative Direction. The plot looks significantly better than the plot in Figure 2. The data is more separated, meaning it will be easier to cluster or classify.



**Figure 3. Discriminative Directions of Iris data**