ECE 8560 Takehome 3

Rohan Dani (rdani)

April 20, 2017

1 SVM Software

The SVM software used was `libsvm` because it was easily available to install on RStudio as a package. `Libsvm` is computationally efficient and has facility to change the kernel. It is also available on MATLAB but its trickier to install on it. It has documentation for available functions and it does not require a lot of coding to get it working.

2 SVM with Linear Kernel

Support vectors
The number of support vectors were 4580.

Hyperplane parameters
The parameters for linear hyperplane (variable w_l) were

```
            V1         V2          V3        V4
[1,] 0.7295572 -0.1858186 0.01255235 -1.498935
```

Classification performance
The result of using a linear kernel was as follows

```
Confusion Matrix and Statistics

          Reference
Prediction    1    2
         1 4383 1246
         2  617 3754

               Accuracy : 0.8137
                 95% CI : (0.8059, 0.8213)
    No Information Rate : 0.5
    P-Value [Acc > NIR] : < 2.2e-16
```

Thus we get an accuracy of 81.37% in other words, a probability of error of 18.63%.

Comparison with previous takehomes

Method	P(error)
Bayesian	8.81%
Ho-Kashyap	18.27%
1-NNR	13.25%
3-NNR	11.39%
5-NNR	10.59%
PCA	25.63%
SVM Linear	18.63%

A few reasons why Linear SVM runs poorly can be the fact that only 2 classes are considered and it does better than PCA for same amount of dimensions. Also the data may not be linearly separable at all.

3 SVM with Radial Kernel

Support vectors
The number of support vectors were 2322.

Hyperplane parameters
The parameters for linear hyperplane (variable w_r) were

```
           V1        V2       V3       V4
[1,] 10.56599 -2.621675 5.928734 -19.1017
```

Classification performance
The result of using a radial kernel was as follows

```
Confusion Matrix and Statistics

          Reference
Prediction    1    2
         1 4699  623
         2  301 4377

               Accuracy : 0.9076
                 95% CI : (0.9018, 0.9132)
    No Information Rate : 0.5
    P-Value [Acc > NIR] : < 2.2e-16
```

Thus we get an accuracy of 90.76% in other words, a probability of error of 9.24%.

Comparison with previous takehomes

Method	P(error)
Bayesian	8.81%
Ho-Kashyap	18.27%
1-NNR	13.25%
3-NNR	11.39%
5-NNR	10.59%
PCA	25.63%
SVM Radial	9.24%

SVM with radial kernel does better due to the fact that it maps data onto higher dimensions where the data might be linearly separable.

4 Crisp C-means

Crisp c-means is used to partition data into c number of clusters. The initial centroid points are assigned randomly but eventually using squared euclidean and city block distance measures we end up with same clustering even with random initialization of the centroids. Squared Euclidean distances put greater emphasis on objects which are far apart whereas city block in general gives similar results but the effect of objects farther away is dampened.

For 2 clusters:

Distance Measure	Cluster 1	Cluster 2
Squared Euclidean	11204	3796
City Block	10378	4622

For 3 clusters:

Distance Measure	Cluster 1	Cluster 2	Cluster 3
Squared Euclidean	8470	3528	3002
City Block	7660	4472	2868

For 4 clusters:

Distance Measure	Cluster 1	Cluster 2	Cluster 3	Cluster 4
Squared Euclidean	7242	2331	2101	3326
City Block	7248	2342	2019	3391

For 5 clusters:

Distance Measure	Cluster 1	Cluster 2	Cluster 3	Cluster 4	Cluster 5
Squared Euclidean	2046	2369	2093	6186	2306
City Block	2020	2314	2090	6261	2315

Means of known clusters are as follows:

```
 50.1171   -4.9704  -24.8118  -49.8120 <---- Class 1
 24.1311   -0.1184  -25.0483    0.2915 <---- Class 2
 49.7022    5.4015   24.5501  -49.9373 <---- Class 3
```

For k = 2, we get centroids as

```
centr =
   23.0565   -1.0703  -25.8566   10.0801 <--- Similar to Class 2
   49.4722    0.6289   -0.6566  -52.4612 <--- Similar to Class 3
```

For k = 3, we get centroids as

```
centr =
   47.5987   -4.5547  -22.8624  -48.5556 <---- Similar to Class 1 mean
   16.9294    0.4262  -25.0998   20.8929 <---- Similar to Class 2 mean
   50.2301    8.6871   33.2424  -50.0889 <---- Similar to Class 3 mean
```

For k = 4, we get centroids as

```
centr =
   60.6739   -0.5672  -25.2066   16.1090 <---- Similar to Class 1 mean
   48.9515   -5.0798  -22.4219  -52.5452 <---- Similar to Class 1 mean
   50.0833    8.8591   33.4163  -49.8991 <---- Similar to Class 3 mean
  -20.8088    0.6412  -24.9355    5.2054 <---- Similar to Class 2 mean
```

For k = 5 we get centroids as

```
centr =
  -24.3084    1.0820  -24.8619    2.5367 <---- Similar to Class 2 mean
   50.0725   -0.2792  -24.9694   37.8907 <---- Similar to Class 1 mean
   44.3112   -7.3566  -22.6520  -63.8895 <---- Similar to Class 1 mean
   57.4439   -1.5492  -25.1025  -26.1063 <---- Similar to Class 3 mean
   50.4641    8.1799   31.3214  -49.9479 <---- Similar to Class 3 mean
```

All the means are similar to one or more original classes so we can say that we have essentially the right clustering and thus we have a natural clustering.