ISYE 8803 HDDA

Exam 2

June 27th, 2021

1.

Multinomial Logistic Regression

Provided was the output of the coefficients for the first model where alpha and lambda = 0:

```
$`1`
16 x 1 sparse Matrix of class "dgCMatrix"
                                  s0
Area                   -0.0001518537
Perimeter              -0.0672444842
MajorAxisLength         0.0579029492
MinorAxisLength         0.5163178131
AspectRation          -26.3192666966
Eccentricity          -29.0661341989
ConvexArea             -0.0001970225
EquivDiameter          -0.1196069538
Extent                  9.9435753216
Solidity             -399.4886523771
roundness             -80.2497338373
Compactness          -362.7905849062
ShapeFactor1        -2701.9336760038
ShapeFactor2         5872.2369055598
ShapeFactor3           19.3050598152
ShapeFactor4          650.3621865493

$`2`
16 x 1 sparse Matrix of class "dgCMatrix"
                                  s0
Area                    0.0003079825
Perimeter              -0.0354750125
MajorAxisLength         0.2327111916
MinorAxisLength         0.1173886847
AspectRation          -63.6883633694
Eccentricity          -13.5440118872
ConvexArea             -0.0003392002
EquivDiameter           0.1844685751
```

```
Extent              -12.2074197519
Solidity            572.2548906164
roundness            49.7578845326
Compactness          20.0540584979
ShapeFactor1      15998.1103533640
ShapeFactor2      -3065.7676464205
ShapeFactor3        -37.8950713030
ShapeFactor4       -567.9180043295

$`3`
16 x 1 sparse Matrix of class "dgCMatrix"
                                s0
Area                 -0.0001561288
Perimeter             0.1027194967
MajorAxisLength      -0.2906141408
MinorAxisLength      -0.6337064978
AspectRation         90.0076300659
Eccentricity         42.6101460861
ConvexArea            0.0005362227
EquivDiameter        -0.0648616213
Extent                2.2638444303
Solidity           -172.7662382394
roundness            30.4918493047
Compactness         342.7365264083
ShapeFactor1     -13296.1766773603
ShapeFactor2      -2806.4692591393
ShapeFactor3         18.5900114878
ShapeFactor4        -82.4441822198
```

The coefficients represent 3 seperate logistic models for each indepedent class variable, we can see that the ridge regression component has pushed several variable close to 0, while other variables seem to have a significant impact on the predicitor variable.

We can see the confusion matrix output as the following:

```
         True
Predicted   1   2   3 Total
```

```
    1      358   0   3   361
    2        0 136   0   136
    3        4   0 570   574
    Total  362 136 573  1071

 Percent Correct:  0.9935
```

We can see that this model performed well with a score of 99.35%, where we can see that the actual was 1 ('Barbunya') but the model predicted 3 ('Seker'), 4 times.In addition the actual was 3 ('Seker'), but the model predicted ('Barbunya'), 3 times.

2.

Multinomial Logistic Ridge Regression

```
Area                  -0.0001518537
Perimeter             -0.0672444842
MajorAxisLength        0.0579029492
MinorAxisLength        0.5163178131
AspectRation         -26.3192666966
Eccentricity         -29.0661341989
ConvexArea            -0.0001970225
EquivDiameter         -0.1196069538
Extent                 9.9435753216
Solidity            -399.4886523771
roundness            -80.2497338373
Compactness         -362.7905849062
ShapeFactor1       -2701.9336760038
ShapeFactor2        5872.2369055598
ShapeFactor3          19.3050598152
ShapeFactor4         650.3621865493

$`2`
16 x 1 sparse Matrix of class "dgCMatrix"
                              s0
Area                   0.0003079825
Perimeter             -0.0354750125
MajorAxisLength        0.2327111916
MinorAxisLength        0.1173886847
```

```
AspectRation       -63.6883633694
Eccentricity       -13.5440118872
ConvexArea          -0.0003392002
EquivDiameter        0.1844685751
Extent             -12.2074197519
Solidity           572.2548906164
roundness           49.7578845326
Compactness         20.0540584979
ShapeFactor1     15998.1103533640
ShapeFactor2     -3065.7676464205
ShapeFactor3       -37.8950713030
ShapeFactor4      -567.9180043295

$`3`
16 x 1 sparse Matrix of class "dgCMatrix"
                               s0
Area                -0.0001561288
Perimeter            0.1027194967
MajorAxisLength     -0.2906141408
MinorAxisLength     -0.6337064978
AspectRation        90.0076300659
Eccentricity        42.6101460861
ConvexArea           0.0005362227
EquivDiameter       -0.0648616213
Extent               2.2638444303
Solidity          -172.7662382394
roundness           30.4918493047
Compactness        342.7365264083
ShapeFactor1    -13296.1766773603
ShapeFactor2     -2806.4692591393
ShapeFactor3        18.5900114878
ShapeFactor4       -82.4441822198
```

$\lambda_{ridge} = 0.04476533$

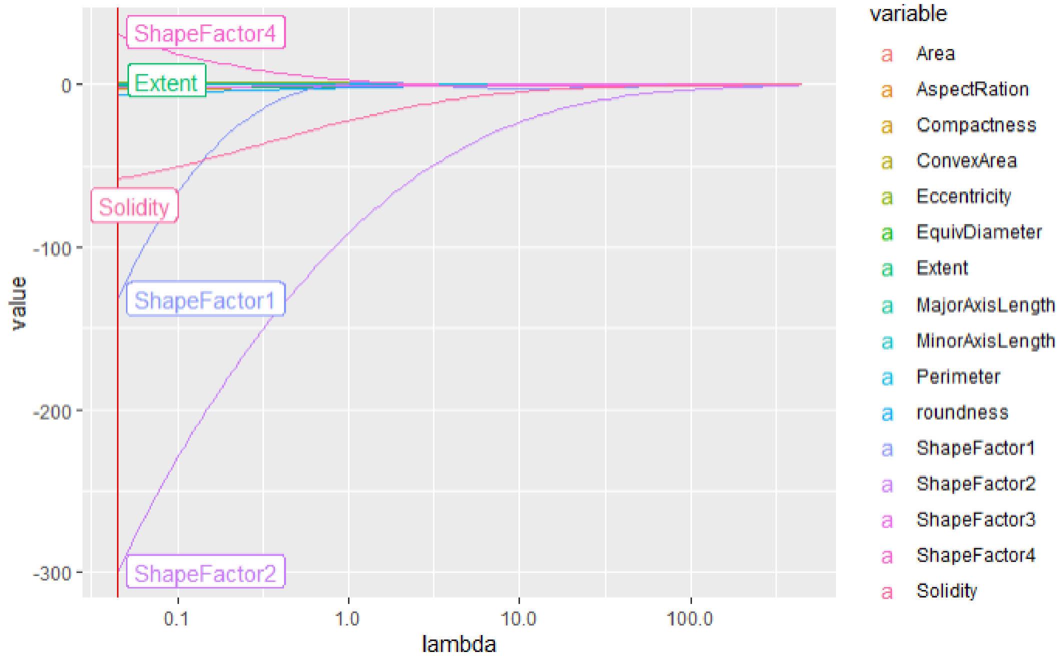


Figure 1: Coefficient Path Barbunya Ridge

We can see from our first plot that all variables seem to be significant accept for Area to MinorAxisLength, and ConvexArea to EquivDiameter. Some other coefficients have a value of < 10, so they will not be as pronounced in the above plot. Our significant variables also all seem to be negatively stastically significant.

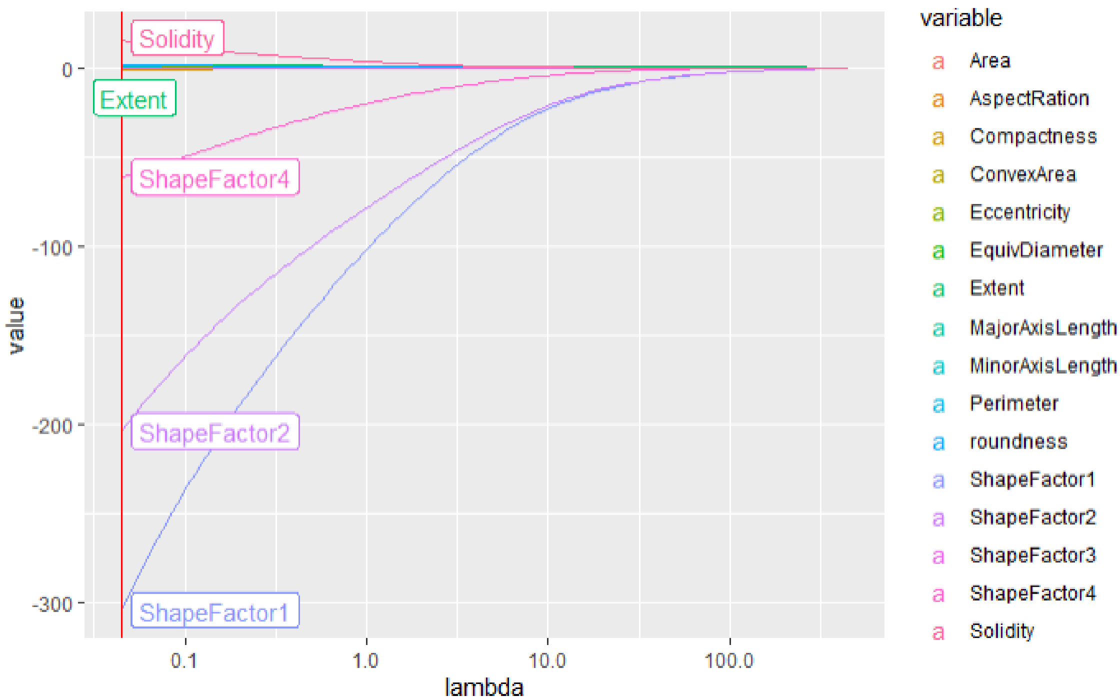


Figure 2: Coefficient Path Bombay Ridge

Very similar to our first plot, all variables seem to be significant accept for Area to MinorAxisLength, and ConvexArea to EquivDiameter.

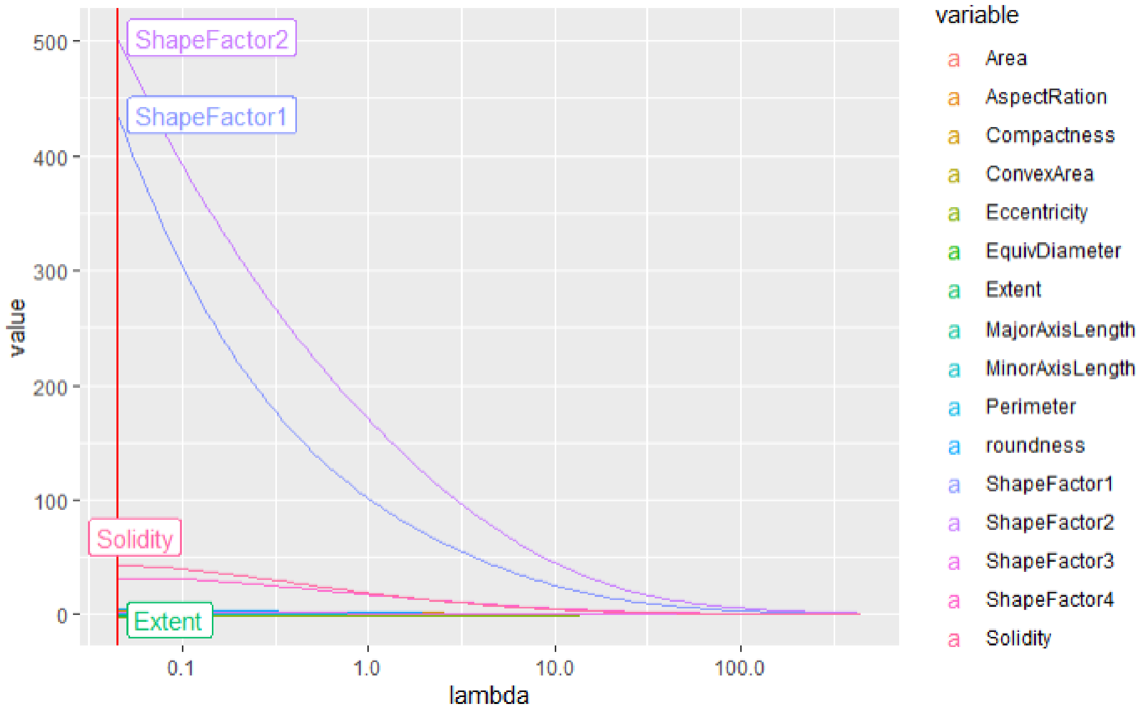


Figure 3: Coefficient Path Seker Ridge

For our last plot, were similar results to our above two plots, however our significant variables our positively statistically significant.

However, the same variables seem to be stastisically signficant across all different dependent variable plots.

Provided is the confusion matrix output:

```
          True
Predicted    1    2    3  Total
    1      359    0    5    364
    2        0  136    0    136
    3        3    0  568    571
    Total  362  136  573   1071

 Percent Correct:  0.9925
```

We can see that this model performed well with a score of 99.25%, where we can see that the actual was 1 ('Barbunya') but the model predicted 3 ('Seker'), 3 times.In addition the actual was 3 ('Seker'), but the model predicted ('Barbunya') 1, 5 times.

3. Multinomial Logistic Lasso Regression

Provided is the coefficient output for Lasso regression:

```
$`1`
16 x 1 sparse Matrix of class "dgCMatrix"
                               s0
Area                  .
Perimeter             .
MajorAxisLength       .
MinorAxisLength       .
AspectRation          .
Eccentricity          .
ConvexArea            .
EquivDiameter         .
Extent                0.01242462
Solidity          -213.87173804
roundness          -17.12624273
Compactness           .
ShapeFactor1          .
ShapeFactor2          .
ShapeFactor3          .
ShapeFactor4       254.29586360

$`2`
16 x 1 sparse Matrix of class "dgCMatrix"
                              s0
Area            0.0001423756
Perimeter       .
MajorAxisLength .
MinorAxisLength 0.0307786246
AspectRation    .
Eccentricity    .
ConvexArea      .
EquivDiameter   .
Extent          .
Solidity        .
```

```
roundness        4.1577019238
Compactness      .
ShapeFactor1     .
ShapeFactor2     .
ShapeFactor3     .
ShapeFactor4     .

$`3`
16 x 1 sparse Matrix of class "dgCMatrix"
                           s0
Area                  .
Perimeter             .
MajorAxisLength       .
MinorAxisLength   -0.09674459
AspectRation          .
Eccentricity          .
ConvexArea            .
EquivDiameter         .
Extent            -7.85914798
Solidity              .
roundness             .
Compactness           .
ShapeFactor1          .
ShapeFactor2    8692.04645534
ShapeFactor3          .
ShapeFactor4          .
```

We can see that the Lasso penalizes a lot of coefficients, and also seems to not handle multicolinearity well since some of the significant ShapeFactor variables have been pushed to 0 compared to the results for ridge regressions' coefficients.

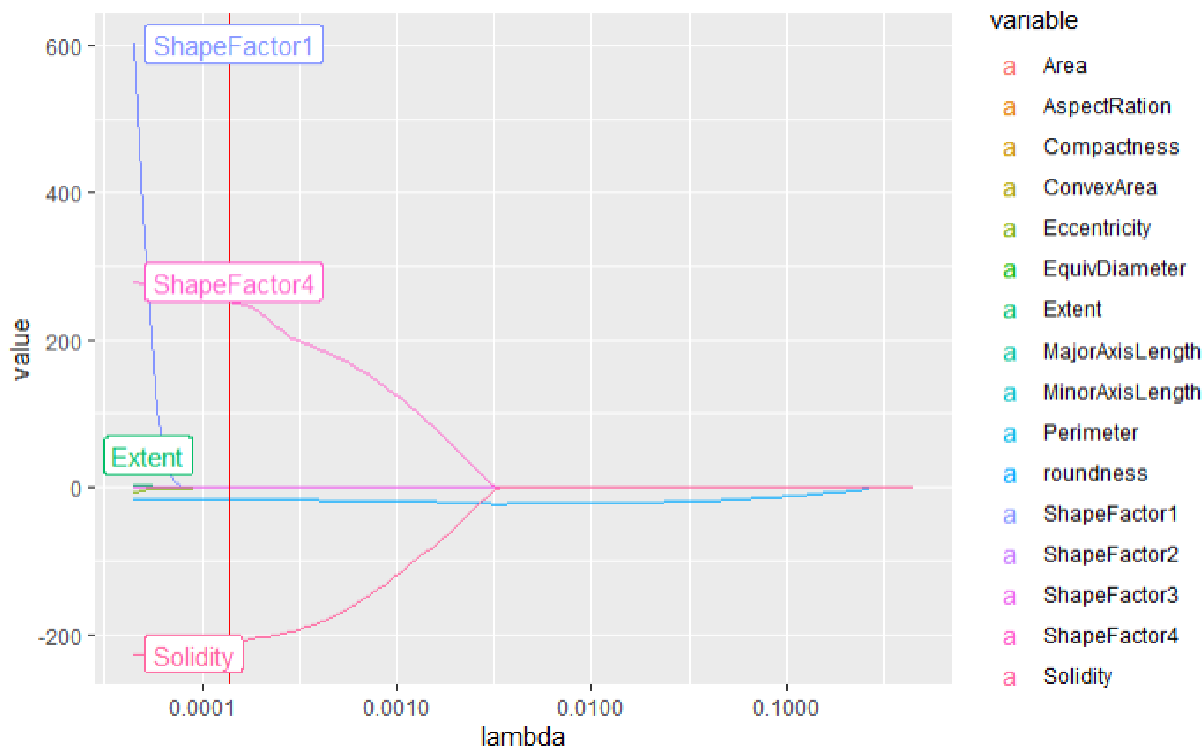


Figure 4: Coefficient Path Barbunya Lasso

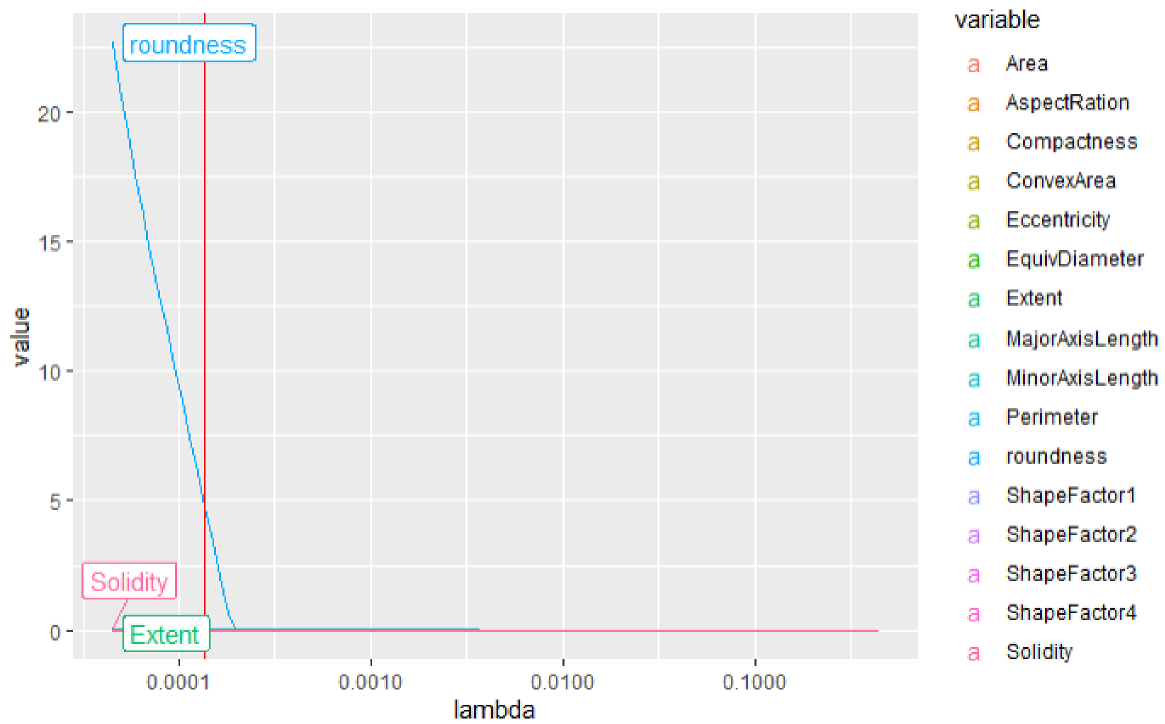


Figure 5: Coefficient Path Bombay Lasso

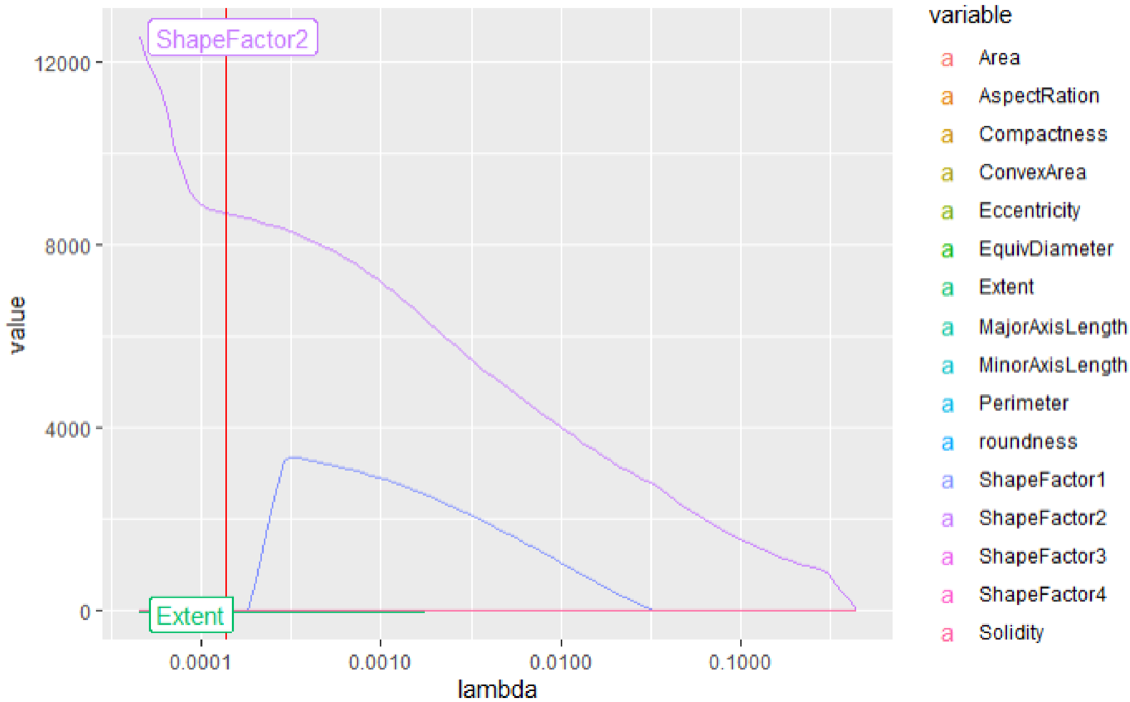


Figure 6: Coefficient Path Seker Lasso

We can see that all statistically signfiicant variable seem to be positive, and are either related to roundness or ShapeFactor.

$\lambda_{lasso} = 0.0001367068$

```
          True
Predicted    1    2    3  Total
    1      358    0    3    361
    2        0  136    0    136
    3        4    0  570    574
    Total  362  136  573   1071

 Percent Correct:  0.9935
```

We can see that this model performed well with a score of 99.35%, where we can see that the actual was 1 ('Barbunya') but the model predicted 3 ('Seker'), 4 times.In addition the actual was 3 ('Seker'), but the model predicted ('Barbunya'), 3 times.

4. Adapative Lasso

```
$`1`
16 x 1 sparse Matrix of class "dgCMatrix"
                      s0
Area                .
Perimeter           .
MajorAxisLength     .
MinorAxisLength     .
AspectRation        .
Eccentricity        .
ConvexArea          .
EquivDiameter       .
Extent              .
Solidity            .
roundness         -15.55279
Compactness         .
ShapeFactor1        .
ShapeFactor2        .
ShapeFactor3        .
ShapeFactor4        .

$`2`
16 x 1 sparse Matrix of class "dgCMatrix"
                             s0
Area            0.00002974898291
Perimeter       .
MajorAxisLength .
MinorAxisLength 0.00286839396827
AspectRation    .
Eccentricity    .
ConvexArea      0.00000006596783
EquivDiameter   .
Extent          .
Solidity        .
roundness       .
Compactness     .
ShapeFactor1    .
```

```
ShapeFactor2    .
ShapeFactor3    .
ShapeFactor4    .

$`3`
16 x 1 sparse Matrix of class "dgCMatrix"
                      s0
Area                   .
Perimeter              .
MajorAxisLength        .
MinorAxisLength        .
AspectRation           .
Eccentricity           .
ConvexArea             .
EquivDiameter          .
Extent                 .
Solidity               .
roundness              .
Compactness            .
ShapeFactor1           .
ShapeFactor2    2035.001
ShapeFactor3           .
ShapeFactor4           .
```

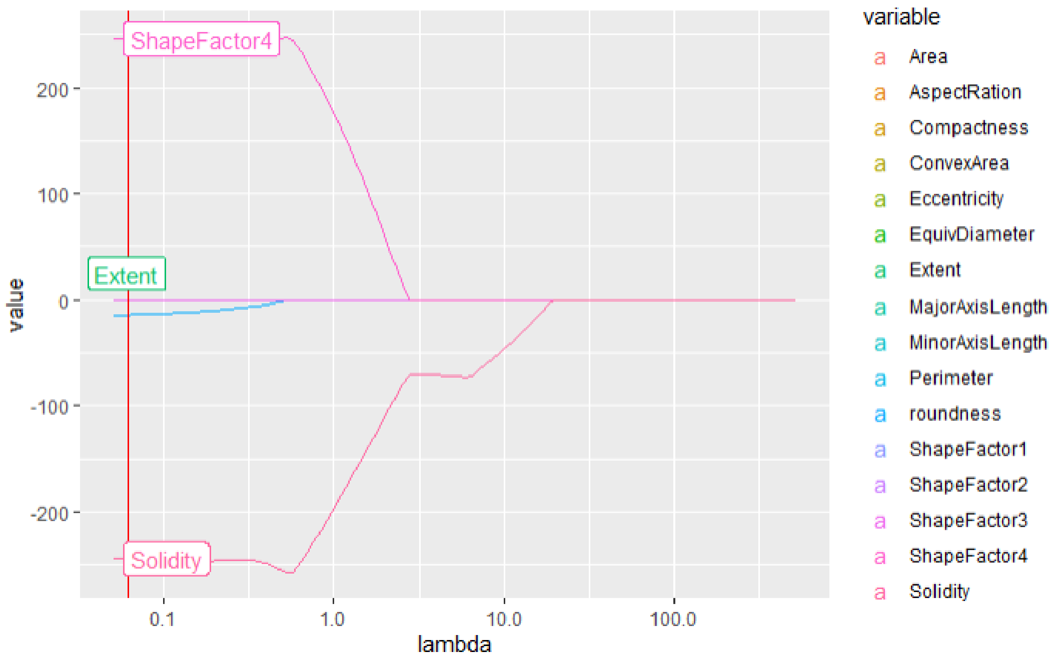


Figure 7: Coefficient Path Barbunya Adaptive Lasso

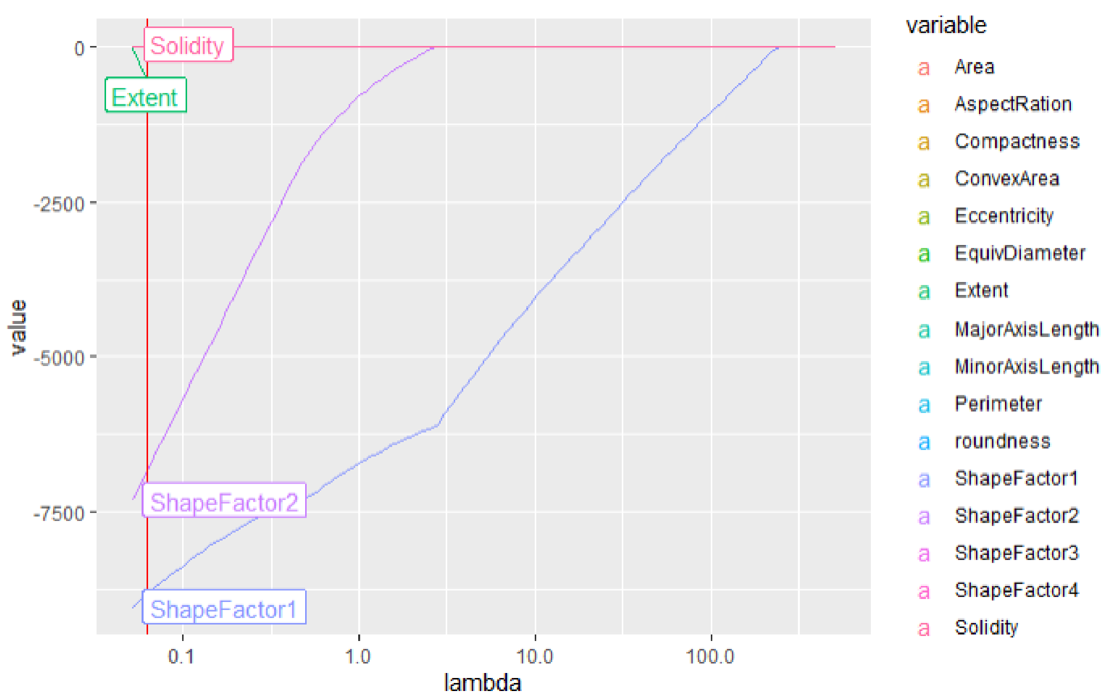


Figure 8: Coefficient Path Bombay Adaptive Lasso

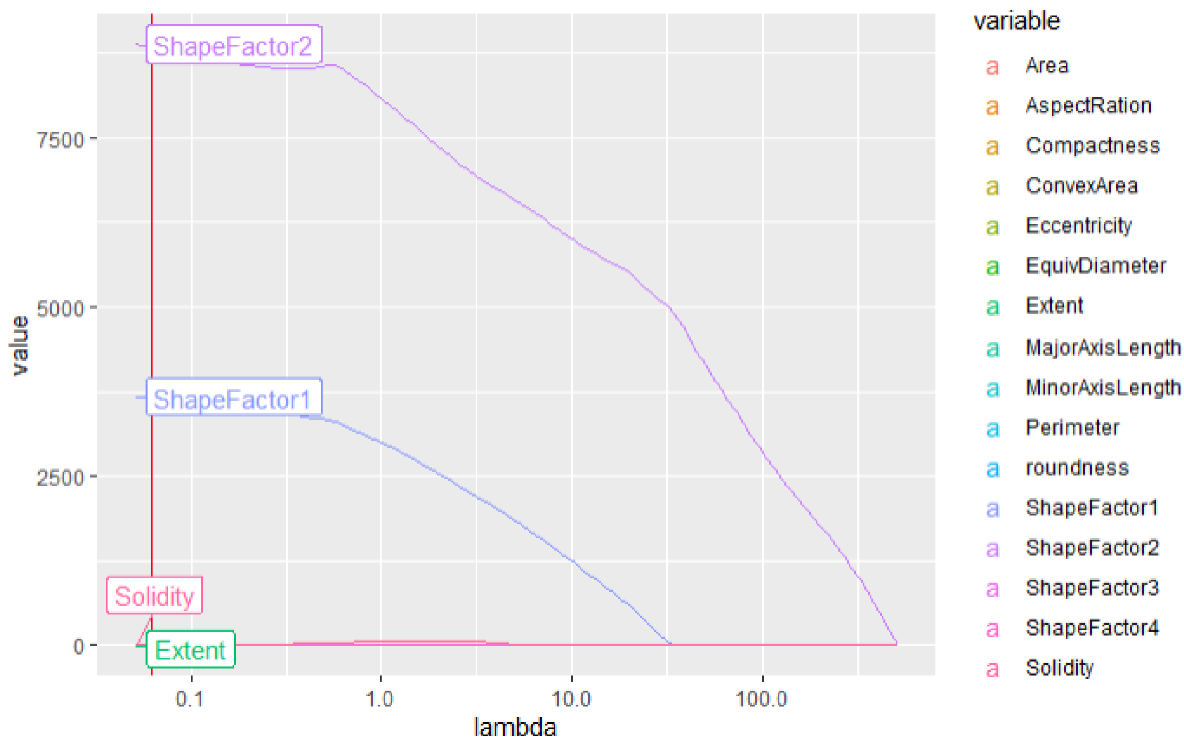
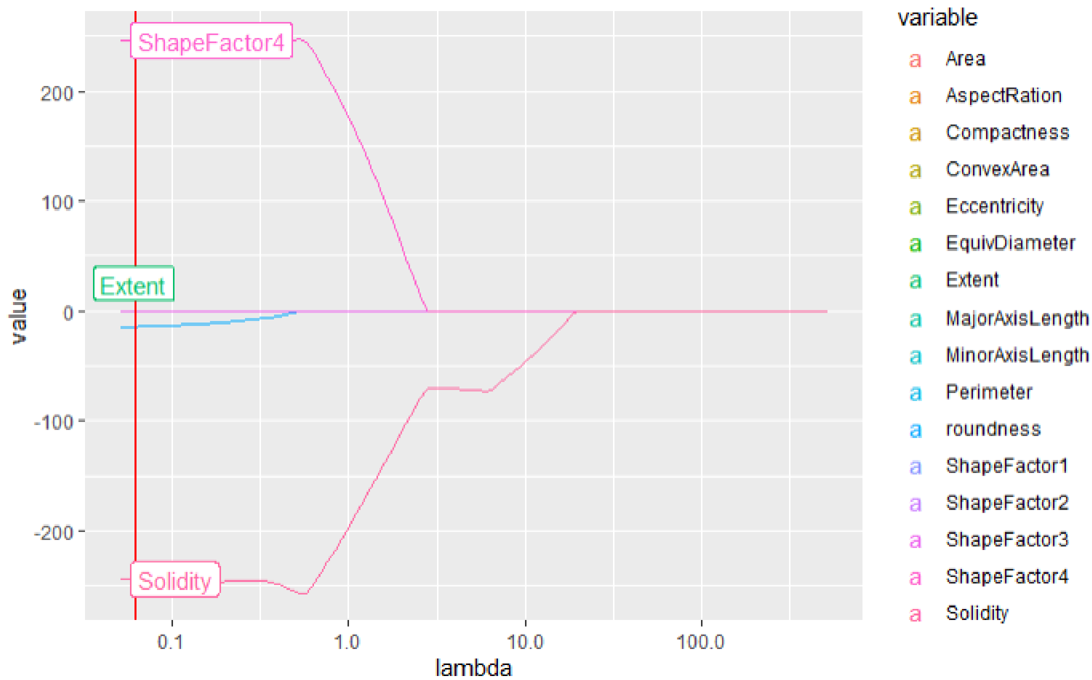
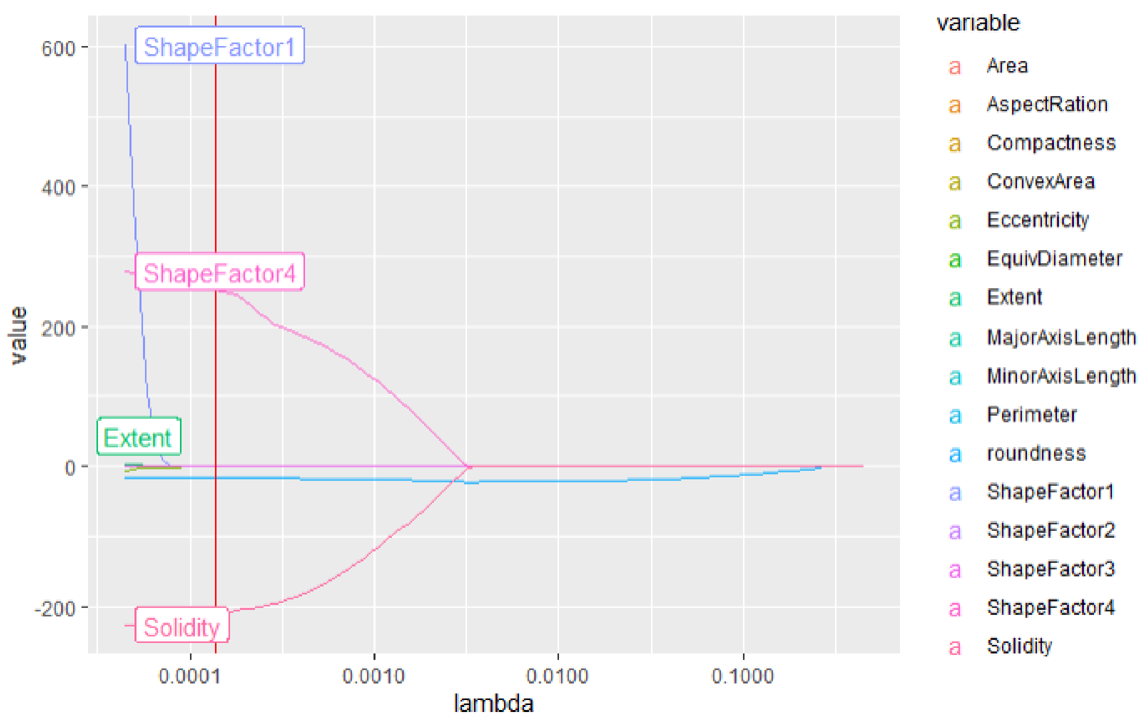


Figure 9: Coefficient Path Seker Adaptive Lasso

Adapative lasso works similarily to lasso regression, however we apply a weight to penalize the regression coefficients to account for their magntitude. For lasso regression, the coefficients are penalized in a constant manner, whereas for adaptive, the coefficeints are penalized based on their magnitude, and are penalized less for larger coefficients. When comparing the coefficient paths below for lasso and adaptive lasso, we can see that the coefficient paths are much smoother for solidity and ShapeFactor variables for adaptive lasso when compared to lasso.





$\lambda_{alasso} = \ 0.06183462$

However, due to the higher sparsity of coefficients for adaptive lasso, it does not seem to perform as well compared to the other models:

```
         True
Predicted   1   2   3 Total
    1     356   0   5   361
    2       0 136   0   136
    3       6   0 568   574
    Total 362 136 573  1071

 Percent Correct:  0.9897
```

The overall classification accuracy is 98.97% for adaptive lasso.

5)
We can see from the outputs of the classification accuracries are the following:

$$Accuracy_{mlr} = 99.35\%$$
$$Accuracy_{ridge} = 99.25\%$$
$$Accuracy_{lasso} = 99.35\%$$
$$Accuracy_{alasso} = 98.97\%$$

From these results we can conclude that either regular multi-nominal logistc or multi-nominal lasso logistic perform equally as well compared to the other models. In addition, adaptive lasso seems to produce the model that leads to the sparsest coefficients.