ECE 8560 Takehome 1

Rohan Dani (rdani)

February 14, 2017

1 Engineering Rationale

The training data in the takehome file has 5000 instances of each class with each instance having 4 features. In this case, Gaussian classifier was determined to be a good option for the following reasons.

- Since there are 5000 instances of each class in the training set, the apriori probability is equal
- Plotting the histogram of features for each class reveal that the PDFs are Gaussian in nature
- There being 4 features, the covariance matrix is very small (4×4)
- Gaussian classifier is computationally very efficient

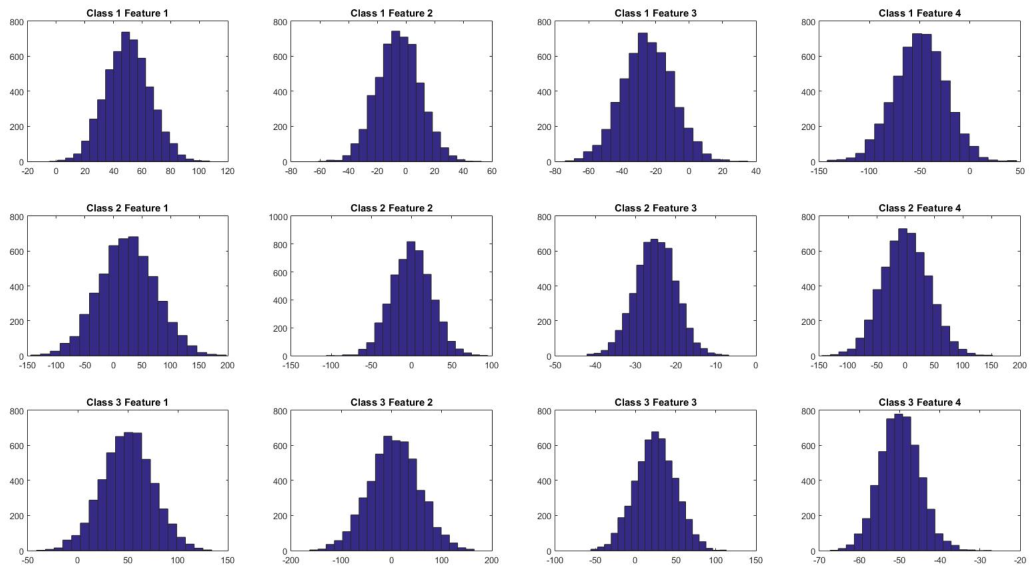

Figure 1: Histogram plots

```
>> corrcoef(data)
ans =
    1.0000   -0.0113    0.1386   -0.2204
   -0.0113    1.0000    0.0854   -0.0024
    0.1386    0.0854    1.0000   -0.2643
   -0.2204   -0.0024   -0.2643    1.0000

>> var(data)
ans =
   1.0e+03 *
    1.2589    1.1286    0.8370    1.3206
```

The covariance matrix does not show any significant correlation between features however the variances of features are not equal.

```
>> var(dataset_1)
ans =
  236.1959  218.9983  224.4166  648.6541
>> var(dataset_2)
ans =
   1.0e+03 *
    2.4548    0.6268    0.0256    1.6109
>> var(dataset_3)
ans =
   1.0e+03 *
    0.6429    2.4866    0.6287    0.0246
```

Thus the decision boundaries are going to be formed based on the mean for the classes.

2 Discriminant Function Form

Since a Gaussian case with unequal variances is given, formula (2-14) is used for each of the classes.

$$g_1(\underline{x}) = -\frac{1}{2}||\underline{x} - \underline{\mu_1}||_{\Sigma_i^{-1}} - \frac{1}{2}\log|\Sigma_1| + \log\{P(w_1)\}$$
$$g_2(\underline{x}) = -\frac{1}{2}||\underline{x} - \underline{\mu_2}||_{\Sigma_i^{-1}} - \frac{1}{2}\log|\Sigma_2| + \log\{P(w_2)\}$$
$$g_3(\underline{x}) = -\frac{1}{2}||\underline{x} - \underline{\mu_3}||_{\Sigma_i^{-1}} - \frac{1}{2}\log|\Sigma_3| + \log\{P(w_3)\}$$

The third term is a class-independent bias and can be eliminated since its equal in all the classes.

3 Probability of Error

Using the mentioned discriminant functions, the probability of error for the training data is 9.0533%