**CSE426 Pattern Recognition – Homework Assignment #5**

Yi Luo (yil712)

1. Summary Error Table

Table 1: Total error counts for each of the four parametric classifiers

|  |  |  |
| --- | --- | --- |
| Test set:  Method: | A | B |
| 1 | 350 | 383 |
| 2 | 91 | 102 |
| 3 | 15 | 16 |
| 4 | 12 | 16 |

1. Equations and Comments

**Method 1**

In this method, we assume *identity* covariance matrices for all 10 classes. We use the discriminant function:



We assume is the same, so we only need to consider the likelihood function. The general multivariate normal density in d dimensions is:

,

where for 10 classes. Therefore we ignore the same components and get:

*,*

which is also known as Euclidian distance.

**Method 2**

In this method, we calculate the covariance matrix for each class and average them to get the shared covariance matrix, then we also start from this equation:



By ignoring the same components, we get

*,*

which is also known as Mahalanobis distance.

**Method 3**

The equation in this method is the same as method 1.

*,*

The only difference is that the dimension of the feature increases from 10 to 256. This new feature is pixel-based.

**Method 4**

In this method, we also use pixel-based features to build the classifier. I implement a Bayes classifier assuming features are conditionally ***independent***. For each testing image, we make use of =45 classifiers to judge the class. Each classifier classifies between a pair of classes. The class that wins most often is selected to be the final class.

To achieve it, for each feature in each class, we can get its probability. For example, for features in class 1, we have:



Then we rewrite the likelihood function based on the assumption that features are independent.



So the discriminant function to decide the class between class 1 and class 2 should be:



If , we decide class 1 else we decide class 2.

**Comments on results**

From the error table in Section 1, it is very obvious that method 1 has the poorest performance. It is because we give some improper assumptions such as the identity covariance matrix. Also, the feature has 10 dimensions and may not contain much enough useful information.

Then, by giving a more reasonable assumption – the shared covariance matrix is not an identity one, instead, it's the average covariance matrix – then we get better results in the method 2.

In method 3, we change our features into pixel-based features. Even though we reuse the classifier in method 1 we still get a much better result. It is because we use a more informative feature space. However, this method increases the dimension of the feature, making the computation more time-consuming.

Last, in method 4, we implement a Bayes classifier by assuming features are conditionally independent. We use a bunch of two-class classifiers to generate a multi-class classifier. This method also gives us a good result.

1. Confusion Tables

Table 2: Method 1 – Moment-space minimum-distance classifier (trained on A, tested on A)

|  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | Error Type I |
| 0 | 51 | 7 | 1 | 0 | 4 | 1 | 5 | 0 | 28 | 3 | 49 |
| 1 | 0 | 99 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 1 |
| 2 | 6 | 1 | 55 | 25 | 0 | 6 | 0 | 7 | 0 | 0 | 45 |
| 3 | 0 | 12 | 12 | 70 | 0 | 1 | 0 | 4 | 0 | 1 | 30 |
| 4 | 0 | 7 | 0 | 0 | 92 | 0 | 0 | 0 | 1 | 0 | 8 |
| 5 | 5 | 2 | 6 | 30 | 6 | 47 | 0 | 0 | 4 | 0 | 53 |
| 6 | 1 | 1 | 0 | 0 | 23 | 0 | 56 | 0 | 19 | 0 | 44 |
| 7 | 0 | 1 | 2 | 0 | 0 | 0 | 0 | 79 | 0 | 18 | 21 |
| 8 | 35 | 11 | 1 | 0 | 4 | 0 | 8 | 0 | 41 | 0 | 59 |
| 9 | 0 | 20 | 5 | 1 | 0 | 0 | 0 | 14 | 0 | 60 | 40 |
| Error Type II | 47 | 62 | 27 | 56 | 37 | 8 | 13 | 26 | 52 | 22 | 350 |

Table 3: Method 1 – Moment-space minimum-distance classifier (trained on A, tested on B)

|  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | Error Type I |
| 0 | 44 | 6 | 0 | 0 | 4 | 0 | 9 | 1 | 33 | 3 | 56 |
| 1 | 0 | 100 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 2 | 10 | 2 | 44 | 20 | 0 | 9 | 0 | 12 | 0 | 3 | 56 |
| 3 | 0 | 11 | 16 | 72 | 0 | 0 | 0 | 1 | 0 | 0 | 28 |
| 4 | 0 | 5 | 0 | 0 | 94 | 0 | 0 | 0 | 1 | 0 | 6 |
| 5 | 7 | 3 | 3 | 35 | 2 | 43 | 2 | 3 | 2 | 0 | 57 |
| 6 | 1 | 5 | 0 | 0 | 25 | 0 | 48 | 0 | 21 | 0 | 52 |
| 7 | 0 | 4 | 1 | 0 | 0 | 0 | 0 | 72 | 0 | 23 | 28 |
| 8 | 33 | 12 | 0 | 0 | 3 | 1 | 8 | 1 | 42 | 0 | 58 |
| 9 | 0 | 25 | 5 | 0 | 0 | 0 | 0 | 12 | 0 | 58 | 42 |
| Error Type II | 51 | 73 | 25 | 55 | 34 | 10 | 19 | 30 | 57 | 29 | 383 |

Table 4: Method 2 – Moment-space classifier with identical covariances (trained on A, tested on A)

|  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | Error Type I |
| 0 | 70 | 3 | 0 | 0 | 0 | 0 | 0 | 0 | 27 | 0 | 30 |
| 1 | 0 | 100 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 2 | 0 | 4 | 96 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 4 |
| 3 | 0 | 13 | 0 | 85 | 0 | 1 | 0 | 0 | 0 | 1 | 15 |
| 4 | 0 | 1 | 0 | 0 | 98 | 0 | 1 | 0 | 0 | 0 | 2 |
| 5 | 0 | 2 | 0 | 1 | 0 | 97 | 0 | 0 | 0 | 0 | 3 |
| 6 | 0 | 1 | 0 | 0 | 0 | 0 | 99 | 0 | 0 | 0 | 1 |
| 7 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 99 | 0 | 0 | 1 |
| 8 | 27 | 3 | 0 | 0 | 0 | 2 | 0 | 0 | 68 | 0 | 32 |
| 9 | 0 | 3 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 97 | 3 |
| Error Type II | 27 | 31 | 0 | 1 | 0 | 3 | 1 | 0 | 27 | 1 | 91 |

Table 5: Method 2 – Moment-space classifier with identical covariances (trained on A, tested on B)

|  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | Error Type I |
| 0 | 67 | 4 | 1 | 0 | 0 | 0 | 0 | 0 | 28 | 0 | 33 |
| 1 | 0 | 100 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 2 | 0 | 4 | 93 | 0 | 1 | 2 | 0 | 0 | 0 | 0 | 7 |
| 3 | 0 | 13 | 0 | 86 | 0 | 1 | 0 | 0 | 0 | 0 | 14 |
| 4 | 0 | 0 | 0 | 0 | 100 | 0 | 0 | 0 | 0 | 0 | 0 |
| 5 | 0 | 3 | 0 | 1 | 0 | 96 | 0 | 0 | 0 | 0 | 4 |
| 6 | 0 | 1 | 0 | 0 | 0 | 0 | 99 | 0 | 0 | 0 | 1 |
| 7 | 0 | 4 | 0 | 0 | 0 | 0 | 0 | 96 | 0 | 0 | 4 |
| 8 | 33 | 2 | 0 | 0 | 0 | 0 | 0 | 0 | 65 | 0 | 35 |
| 9 | 0 | 3 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 96 | 4 |
| Error Type II | 33 | 34 | 1 | 1 | 1 | 3 | 0 | 1 | 28 | 0 | 102 |

Table 6: Method 3 – Pixel-space minimum-distance classifier (trained on A, tested on A)

|  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | Error Type I |
| 0 | 98 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 2 | 2 |
| 1 | 0 | 98 | 0 | 1 | 0 | 0 | 0 | 1 | 0 | 0 | 2 |
| 2 | 0 | 0 | 99 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 1 |
| 3 | 0 | 9 | 0 | 90 | 1 | 0 | 0 | 0 | 0 | 0 | 10 |
| 4 | 0 | 0 | 0 | 0 | 100 | 0 | 0 | 0 | 0 | 0 | 0 |
| 5 | 0 | 0 | 0 | 0 | 0 | 100 | 0 | 0 | 0 | 0 | 0 |
| 6 | 0 | 0 | 0 | 0 | 0 | 0 | 100 | 0 | 0 | 0 | 0 |
| 7 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 100 | 0 | 0 | 0 |
| 8 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 100 | 0 | 0 |
| 9 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 100 | 0 |
| Error Type II | 0 | 9 | 0 | 2 | 1 | 0 | 0 | 1 | 0 | 2 | 15 |

Table 7: Method 3 – Pixel-space minimum-distance classifier (trained on A, tested on B)

|  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | Error Type I |
| 0 | 100 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 1 | 0 | 99 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 1 |
| 2 | 0 | 1 | 97 | 2 | 0 | 0 | 0 | 0 | 0 | 0 | 3 |
| 3 | 0 | 8 | 0 | 92 | 0 | 0 | 0 | 0 | 0 | 0 | 8 |
| 4 | 0 | 0 | 0 | 0 | 100 | 0 | 0 | 0 | 0 | 0 | 0 |
| 5 | 0 | 2 | 0 | 1 | 1 | 96 | 0 | 0 | 0 | 0 | 4 |
| 6 | 0 | 0 | 0 | 0 | 0 | 0 | 100 | 0 | 0 | 0 | 0 |
| 7 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 100 | 0 | 0 | 0 |
| 8 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 100 | 0 | 0 |
| 9 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 100 | 0 |
| Error Type II | 0 | 11 | 0 | 4 | 1 | 0 | 0 | 0 | 0 | 0 | 16 |

Table 8: Method 4 – Pixel-space Bayes classifier assuming class-conditionally independent features (trained on A, tested on A)

|  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | Error Type I |
| 0 | 99 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 1 |
| 1 | 0 | 98 | 0 | 1 | 0 | 0 | 0 | 1 | 0 | 0 | 2 |
| 2 | 0 | 0 | 99 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 1 |
| 3 | 0 | 8 | 0 | 92 | 0 | 0 | 0 | 0 | 0 | 0 | 8 |
| 4 | 0 | 0 | 0 | 0 | 100 | 0 | 0 | 0 | 0 | 0 | 0 |
| 5 | 0 | 0 | 0 | 0 | 0 | 100 | 0 | 0 | 0 | 0 | 0 |
| 6 | 0 | 0 | 0 | 0 | 0 | 0 | 100 | 0 | 0 | 0 | 0 |
| 7 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 100 | 0 | 0 | 0 |
| 8 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 100 | 0 | 0 |
| 9 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 100 | 0 |
| Error Type II | 0 | 8 | 0 | 2 | 0 | 0 | 0 | 1 | 0 | 1 | 12 |

Table 9: Method 4 – Pixel-space Bayes classifier assuming class-conditionally independent features (trained on A, tested on B)

|  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | Error Type I |
| 0 | 100 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 1 | 0 | 99 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 1 |
| 2 | 0 | 1 | 96 | 3 | 0 | 0 | 0 | 0 | 0 | 0 | 4 |
| 3 | 0 | 7 | 0 | 93 | 0 | 0 | 0 | 0 | 0 | 0 | 7 |
| 4 | 0 | 0 | 0 | 0 | 100 | 0 | 0 | 0 | 0 | 0 | 0 |
| 5 | 0 | 1 | 0 | 1 | 1 | 96 | 0 | 0 | 1 | 0 | 4 |
| 6 | 0 | 0 | 0 | 0 | 0 | 0 | 100 | 0 | 0 | 0 | 0 |
| 7 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 100 | 0 | 0 | 0 |
| 8 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 100 | 0 | 0 |
| 9 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 100 | 0 |
| Error Type II | 0 | 9 | 0 | 5 | 1 | 0 | 0 | 0 | 1 | 0 | 16 |