

LINEAR DISCRIMINANT ANALYSIS

Program #3

EE 8993: Speech Processing

Audrey Le

Professor: Dr. Joseph Picone

June 9, 1998



1. PROBLEM DESCRIPTION

In this project we are to implement a program that uses linear discriminant analysis technique to identify which of the data sets, given in Figure 1 below, each of the test vectors belongs.

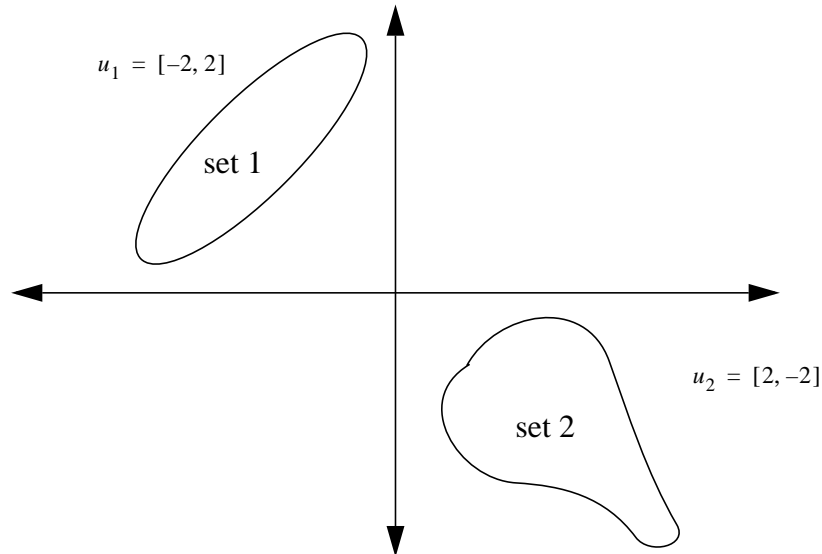


Figure 1: General shape of the data sets.

Each data set is to contain 100 2-D points. Both data sets have a mean approximately $[-2, 2]$ and $[2, -2]$, respectively. The test vectors are given in Table 1. Linear discriminant analysis is used to identify the class of the each of the test vectors. The results are compared to those obtained from Euclidean distance method.

Test vectors	Coordinates
a	-1, -1
b	0, 0
c	0.5, 0.5
d	0.5, -0.5

Table 1: Test vectors.

2. INTRODUCTION

Linear discriminant analysis is one of the numerous data classification techniques. This technique maximizes the class separability by maximizing the ratio of the between-class variance to the within-class variance or the ratio of the overall variance to the within-class variance. Choosing the former or the latter case rests on the type of transformation used. There are two types of transfor-

mation: class-dependent transformation and class-independent transformation.

The class-dependent transformation maximizes the separability of the classes by maximizing the ratio of the variance between the classes to the variance within the classes while the class-independent transformation maximizes the class separability by maximizing the ratio of the overall variance to the variance within the classes [1].

Class-dependent transformation is given by Equation (1):

$$\text{transformed_set_j} = \text{transform_j}^T \times \text{set_j} \quad (1)$$

where set_j is the data set j and transform_j is the transformation matrix obtained by taking the non-redundant eigen vectors of criterion_j which can be obtained by Equation (2):

$$\text{criterion_j} = \text{inv}(\text{cov}_j) \times S_b \quad (2)$$

For class-dependent transformation, there is one transformation matrix obtained from one criterion for each data set. Thus, for j data sets, we will have j criteria and j transformation matrices.

Class-independent transformation is given by Equation (3):

$$\text{transformed_set} = \text{transform}^T \times \text{total_data}^T \quad (3)$$

where total_data is the combined data set (merging j data sets into one) and transform is the transformation matrix obtained by taking the non-redundant eigen vectors of criterion which can be obtained by Equation (4):

$$\text{criterion} = \text{inv}(S_w) \times S_b \quad (4)$$

Unlike the class-dependent transform, class-independent transform has only one criterion and one transformation matrix.

The between-class scatter, S_b , and the within-class scatter, S_w , in the above equations can be found using Equation (5) and Equation (6).

$$S_b = \sum_j (\mu_j - \mu) \times (\mu_j - \mu)^T \quad (5)$$

$$S_w = \sum_j p_j \times \text{cov}_j \quad (6)$$

where μ_j is the mean of the data set j , μ is the mean of the combined data set, and cov_j is the covariance of data set j . The covariance cov_j can be found using Equation (7) and the mean μ of the combined data set can be found using Equation (8).

$$cov_j = \sum_j (x_j - \mu_j) \times (x_j - \mu_j)^T \quad (7)$$

$$\mu = \sum_j p_j \times \mu_j \quad (8)$$

where p_j is the apriori probability of the class j .

3. PROCEDURES

Two data sets were created following the required specifications. The specifications are 1) both data sets must have a general shape as given in Figure 1, 2) each set contains 100 2-D points, and 3) one set must have a mean of $[-2, 2]$, and the other must have a mean of $[2, -2]$.

Xmgr, a Unix graphing tool, was used to create the data sets following the general shape as given in Figure 1. The data sets were readjusted so that their means met the specifications, $\mu_1 = [-2, 2]$ and $\mu_2 = [2, -2]$. The result is given in Figure 2.

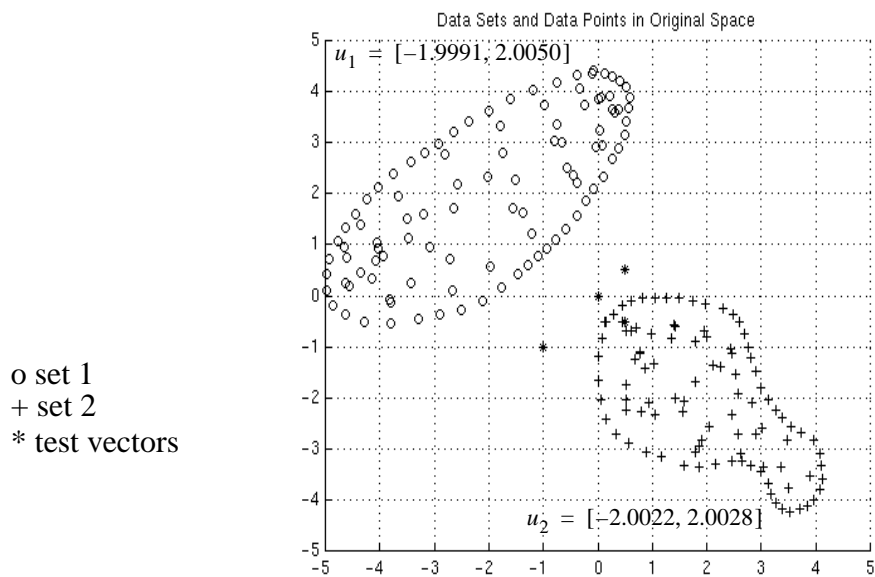


Figure 2: Data sets and test vectors in original space.

Both data sets were merged to create a combined data set. The mean of the combined data set, μ , was computed using Equation (8). No information was given for the apriori probability factors so we assumed them to be 0.5.

The covariance of set 1 and that of set 2 were calculated using Equation (7).

Having the means and the covariances, the between-class scatter and the within-class scatter were computed using Equation (5) and Equation (6).

For the class-dependent transform, the optimizing criterion for each class was calculated using Equation (2) and for the class-independent transform, the overall optimizing criterion was calculated using Equation (4). Thus, for the class-dependent transform, we had j criteria for j data sets with each set representing a class.

Next, the eigen vectors and eigen values of the optimizing criteria were calculated. The non-redundant eigen vectors formed the transformation matrix. To find the non-redundant eigen vectors, we discarded the redundant eigen vectors were discarded using the rule: an eigen vector is redundant if its corresponding eigen value in the diagonal matrix is zero.

The data sets and the test vectors were transformed using Equation (1) for class-dependent transform and Equation (2) for class-independent transform.

The means of the transformed data sets were calculated, and the Euclidean distance between each transformed test vectors and each of the means of the transformed data sets was calculated. The shortest distance k among the j distances classifies the test vector as belonging to class k . The results are given in the next section.

4. RESULTS

The test vectors were first classified using Euclidean distance in original space, in transformed space using class-independent and class-dependent transforms. The results are given in Table 2.

Figure 3, Figure 4, and Figure 5 show the decision regions of various classification approaches. One can see that the class-independent transform is good with generalization but poor with discrimination, and vice versa is true with class-dependent transform.

Test Vectors	Original Space	Class-Ind. LDA	Class-Dep. LDA
a	set 2	set 2	set 1
b	set 1	set 1	set 1
c	set 1	set 1	set 1
d	set 2	set 2	set 1

Table 2: Classes of test vectors based on different types of classification approaches.

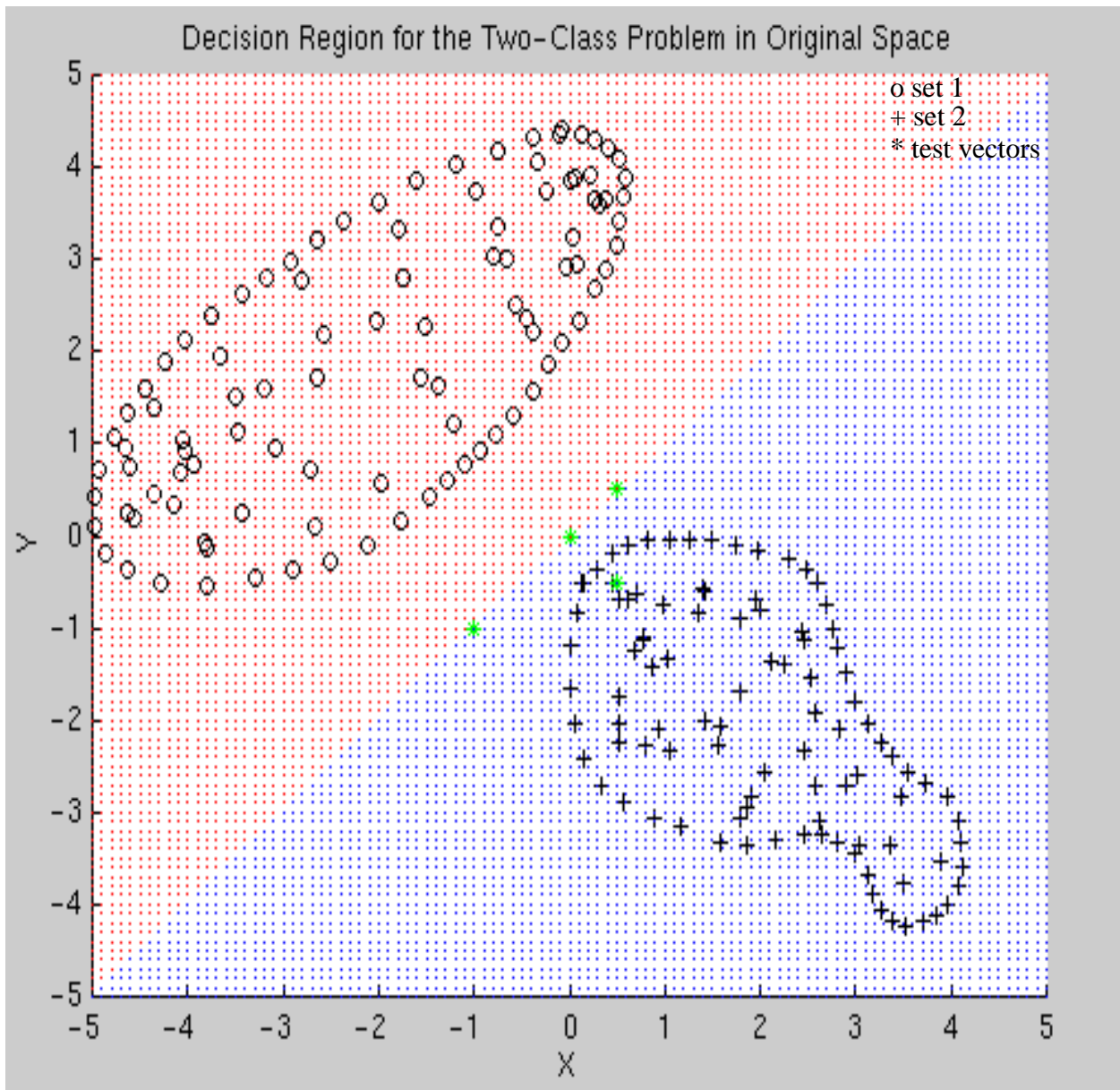


Figure 3: Decision region in original space.

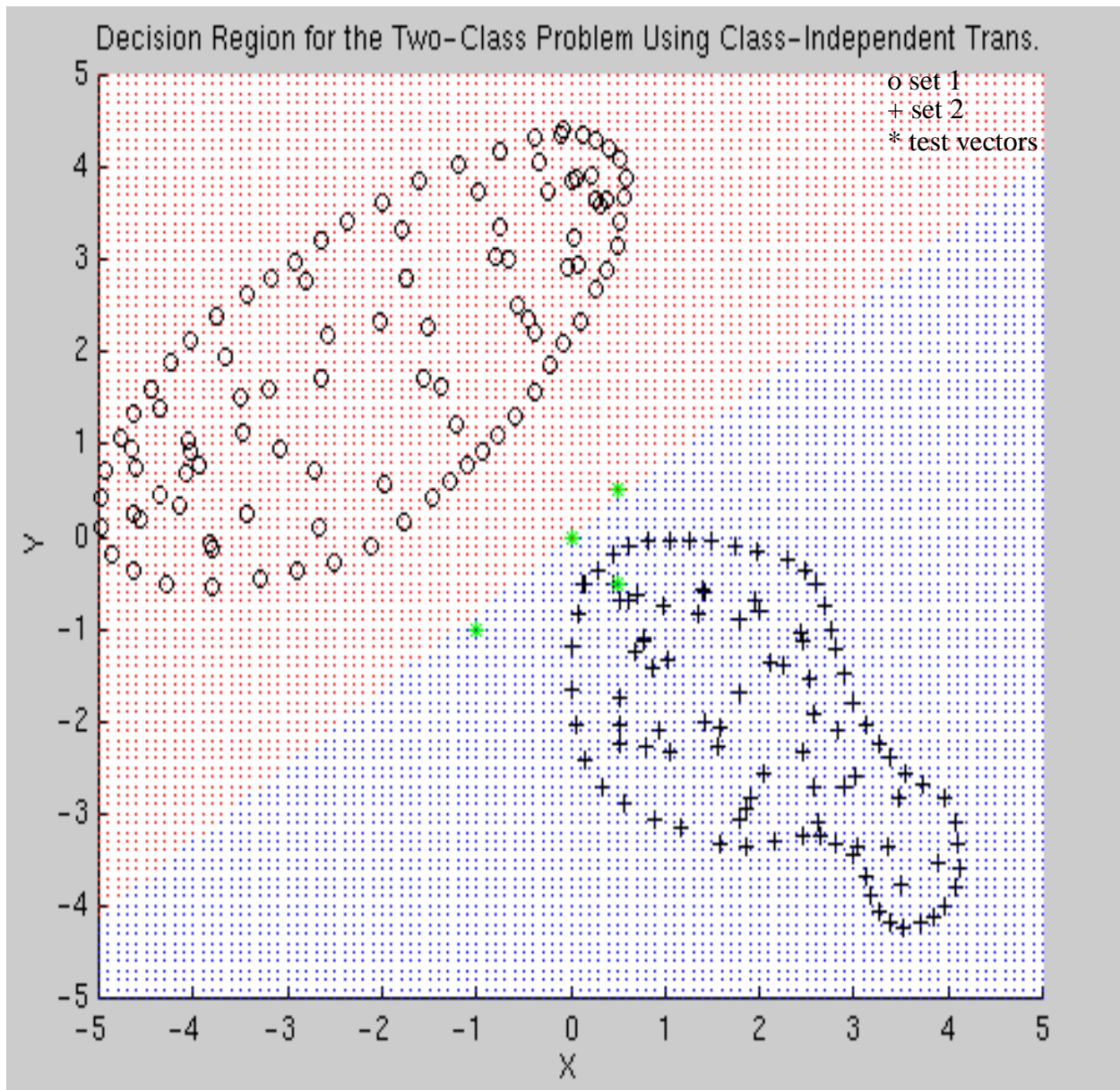


Figure 4: Decision region in transformed space using class-independent transform.

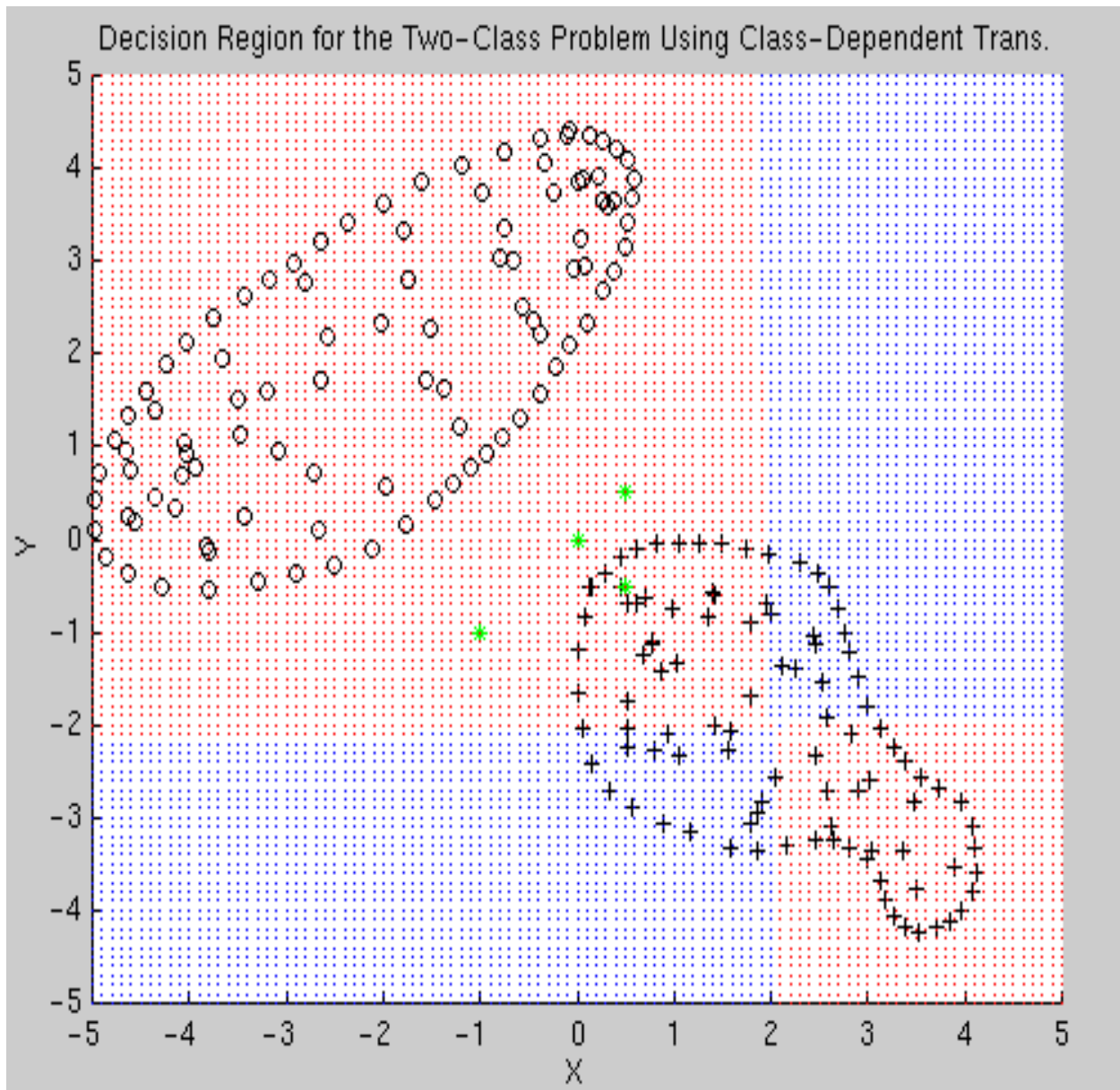


Figure 5: Decision region in transformed space using class-dependent transform.

5. REFERENCES

- [1] S. Balakrishnama and A. Ganapathiraju, "Linear Discriminant Analysis—A Brief Tutorial," Mississippi State Univ., Starkville, MS, 1998.
- [2] K. Fukunaga, *Introduction to Statistical Pattern Recognition*, Academic Press, San Diego, CA, 1990.