# CSC 340 Software Development Project 1

23 January

Due: 11:59 PM, 8 February 2018

Name \_Brenton Belanger\_\_

## Problems:

Consider the data posted to the Web in the file “2018 Spring Project 1 data.txt”, which represent information describing the objects in the two classes illustrated in the chart below. For context, one might assume that the parameters are distances, such as space between bone features as measured on the skeletons of two groups (say pelvic measurements of males and females).

For questions 1-8, suppose that the measurements are multivariate normally distributed for each dimension for each class.

1. Find the *mean* *vectors* **m**1 and **m**2 for each of the classes. (Note: **m**1 and **m**2 are each n-by-1 vectors, where n=2.) Hint: This problem requires the computation of a **sum** of k n-by-1 vectors, and computing a scalar multiple, 1/k \* **sum**, of the **sum**, for each class. Here k is the number of measurement vectors in a given class, and **sum** is a vector that contains the sum of the measurement vectors in a class.

The mean for vector one is [2.12567500418182, 3.18256442081818]

The mean for vector two is [-0.229423759045454, -1.00712048863636]

1. Find the covariance matrices ****1 and ****2 for the classes. This problem requires several steps including:
   1. Subtracting the *mean* vector for each class from each of the measurement vectors in that class;
   2. Multiplying each resulting n-by-1 difference vector by its 1-by-n transpose to obtain an n x n product;
   3. Accumulating the sum of the n-by-n products for the class; and
   4. Multiplying the n-by-n sum for the class by a scalar, 1/k.

The covariance matrix ****1 =

The covariance matrix ****2 =

1. Find and report the ***determinants***, |****1| and |****2|, of the covariance matrices ****1 and ****2 for the classes. The solution to this problem *must* employ the determinant finder derived from the Gauss reduction algorithm.

Determinant of Covariance Matrix 1: 1.0692535232197

Determinant of Covariance Matrix 2: 2.59491542862553

1. Find and report the ***inverses***, ****1-1 and ****2-1, of the covariance matrices ****1 and ****2 for the classes. The solution for this problem *must* employ the inverse finder algorithm derived from the Gauss-Jordan reduction algorithm.

The inverse ****1-1 =

The inverse ****2-1 =

1. Find and report the discriminant functions g1(**x**) and g2(**x**) for the classes? Report these with the right-hand side of each equation ***in matrix form***. Refer to the classifier lecture notes for a detailed example. Feel free to use the equation editor in your word processor to write the equations; alternatively, a neat and legible, handwritten document may be submitted with these equations.

1.0692535232197)

1. Into which classes would your classifier place the points **m**1 and **m**2? Use your matrix tools to evaluate the discriminant functions to support your choices, and be sure to include both the evidence and your interpretation thereof.

My classifier placed m1 in class 1 and m2 in class 2. When the points are put into each matrix, one side, the side the point is on, will equal zero as x-µ will equal zero.

1. Use your personally implemented matrix manipulation tools to determine how many classification errors occur when you apply the discriminant functions to the example data for each class?
   1. List the ***misclassified*** points separately for each class and provide the values of both discriminant functions g1(**x**) and g2(**x**) for each point (example vector) **x**. (Use a table showing **x**, g1(**x**), and g2(**x**) for each misclassified point **x** to organize this response.)

Class 1: 5/110 misclassified

|  |  |  |  |
| --- | --- | --- | --- |
| X | Y | G1 | G2 |
| 0.835257274 | 2.613833101 | -2.40522296363253 | -1.39091378998018 |
| 1.019497222 | 2.126262401 | -1.80688889917104 | -1.57286849638985 |
| 1.058177788 | 2.857649205 | -1.66177257989983 | -1.60155807950255 |
| 0.997512342 | 7.678653712 | -1.65993628371656 | -1.46477915205865 |
| 0.971964211 | 3.610145814 | -1.89879797224067 | -1.50241607096864 |

Class 2: 13/110 misclassified

|  |  |  |  |
| --- | --- | --- | --- |
| X | Y | G2 | G1 |
| 1.098929784 | 0.196824202 | -1.68969363024903 | -1.63697075255806 |
| 1.799171771 | 0.473503489 | -2.67949790904609 | -0.245534168984998 |
| 2.287938993 | 0.076076615 | -3.63524778083622 | -0.0861524925571798 |
| 1.250504112 | -3.062515699 | -1.93045990922867 | -1.31929290132735 |
| 1.289376169 | 0.909520062 | -1.90559576429313 | -1.10397693232222 |
| 1.716910758 | 0.110969495 | -2.54975961134179 | -0.342340718523946 |
| 1.308178739 | -3.198566347 | -2.00916737440893 | -1.17510751620331 |
| 1.571012791 | 0.006267474 | -2.31989716009126 | -0.555037754847254 |
| 1.289671778 | -4.033869387 | -2.00036183938319 | -1.24576315774243 |
| 1.59817949 | -0.715604314 | -2.37837538586161 | -0.524356823986307 |
| 1.878387575 | -0.997947609 | -2.85744027751282 | -0.187545762898445 |
| 1.809383063 | -1.09224725 | -2.73723257714498 | -0.252538957908825 |
| 1.117740331 | -2.139168832 | -1.75053349114156 | -1.66305650025874 |

* 1. Summarize your findings by presenting a table of the tallies of correctly and incorrectly classified items. The table should contain one row and one column for each class. In the table, the entry in row j and column k should report the number of objects in class j that the classifier indicates would be in class k.)
     1. How many examples are correctly identified for each class?
     2. How many examples are incorrectly identified for each class?

1. Estimate and plot the boundary contour generated by the classifier.
   1. Notice that along the exact boundary contour g1(**x**) = g2(**x**) so that for an estimation of the boundary, | g1(**x**) - g2(**x**)| ≈ 0.0.
   2. Subdivide the area of interest into a grid and evaluate the magnitude of the difference of discriminant function values, |g1(**x**) - g2(**x**)| at the grid points.
   3. Choose and report a small value ε.
   4. When | g1(**x**) - g2(**x**)| < ε, report (so that you can plot) a boundary marker point.
2. Linear systems:
   1. If one a solution exists, use your implementation of Gauss-Jordan Elimination Algorithm to ***estimate the*** ***solution*** for the following linear system:  
      

Please supply your response with the variables in the order [x, y, z, w, a, b, c, d].

x= 5.04275955571779, y= 3.83626218908067, z= -1.14585180163738, w= 5.52291807894874, a=- 21.4025134275434, b= -0.281013714345309, c= -7.61730197632581, d= -5.24049642801272

* 1. What is the determinant of the coefficient matrix A?

The determinant is 19177.

* 1. If they exist, what are
     1. The inverse of the coefficient matrix A-1,

[0.5, -0.0833333333333333, -0.0526315789473684, 0.106995884773663, 0.127214170692432, 0.00455580865603644, 0.0133239382486708, -0.731031965375189]

[0, 0.166666666666667, -0.0526315789473684, 0.0288065843621399, 0.109500805152979, -0.189066059225513, 0.178207674075972, -0.517755644782813]

[0, 0, -0.157894736842105, 0.0864197530864197, -0.0628019323671498, 0.0136674259681094, -0.0725129716225738, 0.299629764822444]

[0, 0, 0, 0.156378600823045, 0.035426731078905, -0.0842824601366743, -0.0215232848632374, -0.822912864368775]

[0, 0, 0, 0, -0.391304347826087, 0.109339407744875, -0.0458010377298059, 3.3778484643062]

[0, 0, 0, 0, 0, 0.314350797266515, -0.289859714304016, 0.16650153830109]

[0, 0, 0, 0, 0, 0, 0.253090769329319, 1.10726390989206]

[0, 0, 0, 0, 0, 0, 0, 0.814048078427283]

* + 1. The determinant of A-1, and

The determinant of A-1 is 5.21457996558377E-05

* + 1. The product of the determinants of A and A-1?  
         
       The product is 1
  1. If A-1 exists, check your system solution results by performing the appropriate matrix multiplication and reporting the results.

1. If it exists, what is the ***condition number*** for the coefficient matrix for the system given in problem 9?

The condition number for the coefficient matrix is 164.576836835793.