Sheet 2_LDA

February 27, 2019

0.1 Name: Amr Ashraf Ibrahim Elzawawy ID: 3788

0.1.1 Question 1 on Orthogonal Projection

1) For the given vectors u1 and u2 below

$$u1 = \begin{pmatrix} 3 \\ 4 \\ 0 \end{pmatrix}, u2 = \begin{pmatrix} -4 \\ 3 \\ 0 \end{pmatrix}$$

---

a) Verify u1 and u2 are orthogonal
Ans:

$$u1^T u2 = \begin{pmatrix} 3 & 4 & 0 \end{pmatrix} \begin{pmatrix} -4 \\ 3 \\ 0 \end{pmatrix} = (3 * -4) + (4 * 3) + 0 = 0$$

Thus, u1 and u2 are orthogonal.

---

b) Find the projection of the point $y = [6, 3, -2]^T$ on u1 and u2
Ans:
"y point projected on u1"

$$y_1' = \frac{u_1^T y}{u_1^T u_1} u_1 = \frac{30}{25} u_1 = 1.2 u_1 = \begin{pmatrix} 3.6 \\ 4.8 \\ 0 \end{pmatrix}$$

"y point projected on u2"

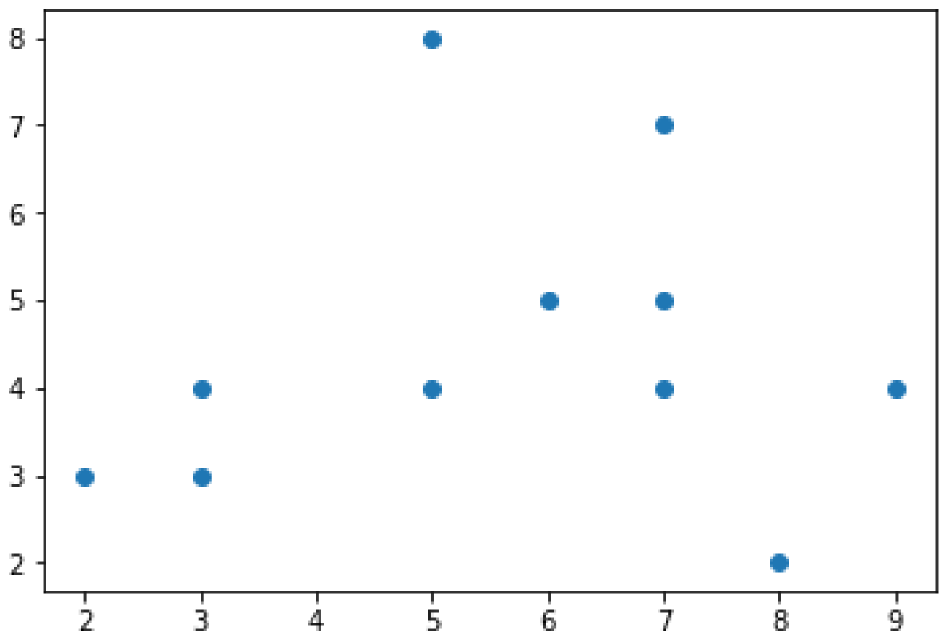
$$y_2' = \frac{u_2^T y}{u_2^T u_2} u_2 = \frac{-15}{25} u_2 = -0.6 u_2 = \begin{pmatrix} 2.4 \\ -1.8 \\ 0 \end{pmatrix}$$

0.1.2 Question 2 on LDA

For the data on two class problem shown below

In [8]:
```
import numpy as np
import matplotlib.pyplot as plt
data = np.array([[2,3],[3,3],[3,4],[5,8],[7,7],[5,4],[6,5],[7,5],[7,4],[8,2],[9,4]])
dim1 = np.array([2,3,3,5,7,5,6,7,7,8,9])
dim2 = np.array([3,3,4,8,7,4,5,5,4,2,4])
plt.scatter(dim1,dim2)
```

Out[8]: <matplotlib.collections.PathCollection at 0x7f544fb9fd68>



a) Compute μ_1 and $\mu 1$, and B, the between-class scatter matrix.
Ans:

$$\mu_1 == (\frac{2+3+3+5+7}{5}, \frac{3+3+4+8+7}{5}) = \begin{pmatrix} 4 \\ 5 \end{pmatrix},$$

$$\mu_{-}1 = (\frac{5+6+7+7+8+9}{6}, \frac{4+5+5+4+2+4}{6}) = \begin{pmatrix} 7 \\ 4 \end{pmatrix},$$

$$B = (\mu_1 - \mu_2)(\mu_1 - \mu_2)^T = \begin{pmatrix} 3 \\ -1 \end{pmatrix} \begin{pmatrix} -3 & 1 \end{pmatrix} = \begin{pmatrix} 9 & -3 \\ -3 & 1 \end{pmatrix}.$$

b) Find the best direction w that discriminates between the classes and sketch it. Given $S^{-}1 = \begin{pmatrix} 0.056 & -0.029 \\ -0.029 & 0.056 \end{pmatrix}$

Ans:

$$(S^{-}1B)w = \lambda w$$

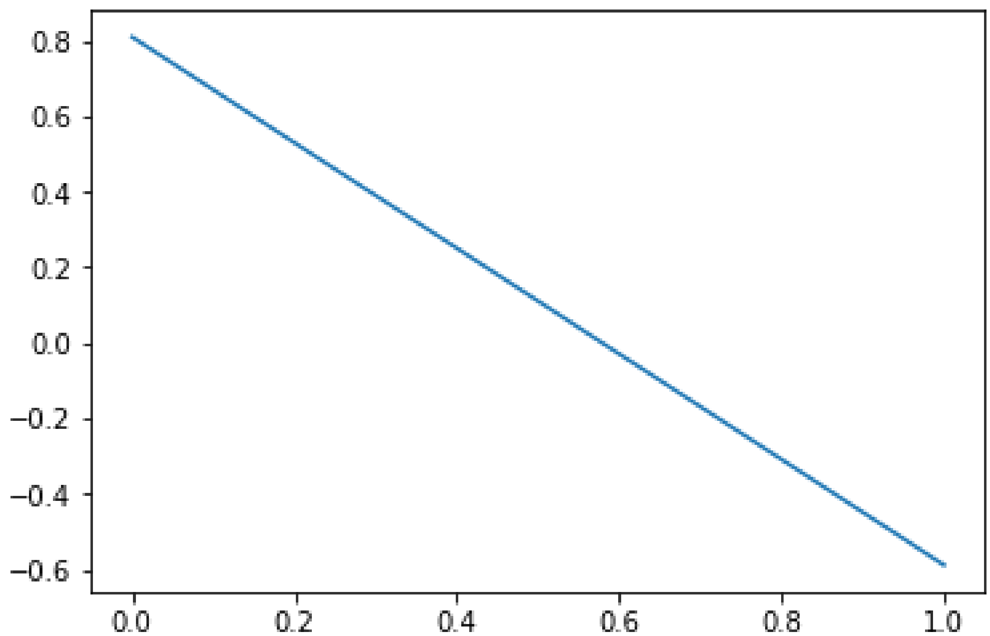
This is simply an eigen value-vector problem solved below by python solver

```
In [2]: sInvMat = np.array([[0.056,-0.029],[-0.029,0.056]])
        bMat = np.array([[9,-3],[-3,1]])
        np.linalg.eig(np.dot(sInvMat,bMat))

Out[2]: (array([0.734, 0.   ]), array([[ 0.80926868,  0.31622777],
                [-0.58743868,  0.9486833 ]]))

In [3]: w = np.array([ 0.80926868,  -0.58743868])
        plt.plot(w)

Out[3]: [<matplotlib.lines.Line2D at 0x7f544fcff2e8>]
```



0.1.3 Question 3 on LDA

For the data on two class problem below

i	xi	yi
x1	(4,2.9)	+1
x2	(3.5,4)	+1
x3	(2.5,1)	-1
x4	(2,2.1)	-1

---

a) Compute μ_1 and $\mu 1$, and B, the between-class scatter matrix.

Ans:

$$\mu+1 == (\frac{4+3.5}{2}, \frac{2.9+4}{2}) = \begin{pmatrix} 3.75 \\ 3.45 \end{pmatrix},$$

$$\mu 1 = (\frac{2.5+2}{2}, \frac{1+2.1}{2}) = \begin{pmatrix} 2.25 \\ 1.55 \end{pmatrix},$$

$$B = (\mu_1 - \mu_2)(\mu_1 - \mu_2)^T = \begin{pmatrix} 1.5 \\ 1.9 \end{pmatrix} \begin{pmatrix} 1.5 & 1.9 \end{pmatrix} = \begin{pmatrix} 2.25 & 2.85 \\ 2.85 & 3.61 \end{pmatrix}.$$

---

b) Compute S+1 and S1, and S, the within-class scatter matrix.
Ans:

$$S_1 = \sum_{x_i-->D_1} (x_i - \mu_1)(x_i - \mu_1)^T = (x_1 - \mu_1)(x_1 - \mu_1)^T + (x_2 - \mu_1)(x_2 - \mu_1)^T =$$

$$\begin{pmatrix} 0.25 \\ -0.55 \end{pmatrix} \begin{pmatrix} 0.25 & -0.55 \end{pmatrix} + \begin{pmatrix} -0.25 \\ 0.55 \end{pmatrix} \begin{pmatrix} -0.25 & 0.55 \end{pmatrix} =$$

$$\begin{pmatrix} 0.0625 & -0.1375 \\ -0.1375 & 0.3025 \end{pmatrix} + \begin{pmatrix} 0.0625 & -0.1375 \\ -0.1375 & 0.3025 \end{pmatrix} = \begin{pmatrix} 0.125 & -0.275 \\ -0.275 & 0.605 \end{pmatrix}$$

$$S_-1 = \sum_{x_i-->D_-1} (x_i - \mu_-1)(x_i - \mu_-1)^T = (x_3 - \mu_-1)(x_3 - \mu_-1)^T + (x_4 - \mu_-1)(x_4 - \mu_-1)^T =$$

$$\begin{pmatrix} 0.25 \\ -0.55 \end{pmatrix} \begin{pmatrix} 0.25 & -0.55 \end{pmatrix} + \begin{pmatrix} -0.25 \\ 0.55 \end{pmatrix} \begin{pmatrix} -0.25 & 0.55 \end{pmatrix} =$$

$$\begin{pmatrix} 0.0625 & -0.1375 \\ -0.1375 & 0.3025 \end{pmatrix} + \begin{pmatrix} 0.0625 & -0.1375 \\ -0.1375 & 0.3025 \end{pmatrix} = \begin{pmatrix} 0.125 & -0.275 \\ -0.275 & 0.605 \end{pmatrix}$$

$$S = \begin{pmatrix} 0.125 & -0.275 \\ -0.275 & 0.605 \end{pmatrix} + \begin{pmatrix} 0.125 & -0.275 \\ -0.275 & 0.605 \end{pmatrix} = \begin{pmatrix} 0.25 & -0.55 \\ -0.55 & 1.21 \end{pmatrix}$$

---

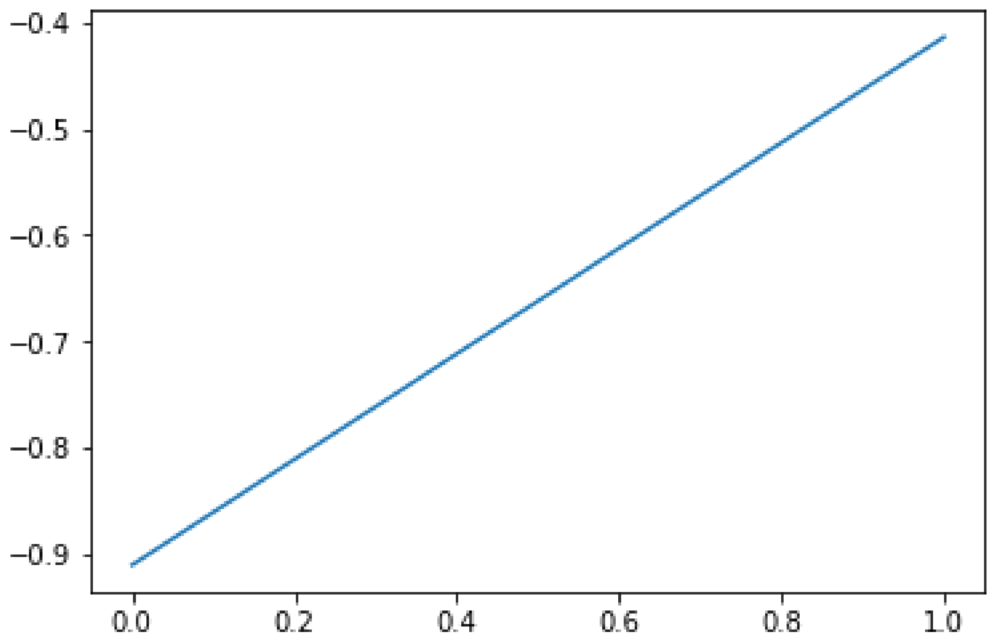
c) Find the best direction w that discriminates between the classes.

```
In [4]: matS = np.array([[0.25,-0.55],[-0.55,1.21]])
        invMatS = np.linalg.inv(matS)
        matB = np.array([[2.25,2.85],[2.85,3.61]])
        np.linalg.eig(np.dot(invMatS,matB))
```

```
Out[4]: (array([-1.55825123e+17,  0.00000000e+00]), array([[-0.91036648,  0.78488277],
                [-0.41380294, -0.61964429]]))
```

```
In [5]: w = np.array([ -0.91036648, -0.41380294])
        plt.plot(w)
```

```
Out[5]: [<matplotlib.lines.Line2D at 0x7f544fc5fe48>]
```



d) Having found the direction w, find the point on w that best separates the two classes.

The point that best separates the two classes maybe the midpoint of the eculidian distance between the two means of the two classes when they are projected onto w.

μ_+1 projected on w = -4.841 , μ_-1 projected on w = -2.9628 **Then, the point on w that best separtes the two classes is -3.9019**

0.1.4 Question 4: Midterm Question Fall 2017

For the data on two class problem below

i	xi	yi
x1	(1,1)	+1
x2	(2,1)	+1
x3	(1,2)	+1
x4	(2,2)	-1
x5	(3,2)	-1

---

a) Compute μ_1 and $\mu 1$ and B, the between-class scatter matrix.

Ans:

$$\mu+1 == (\frac{1+2+1}{3}, \frac{1+1+2}{3}) = \begin{pmatrix}1.33\\1.33\end{pmatrix},$$

$$\mu 1 = (\frac{2+3}{2}, \frac{2+2}{2}) = \begin{pmatrix}2.5\\2\end{pmatrix},$$

$$B = (\mu_1-\mu_2)(\mu_1-\mu_2)^T = \begin{pmatrix}-1.17\\-0.67\end{pmatrix}\begin{pmatrix}-1.17 & -0.67\end{pmatrix} = \begin{pmatrix}1.3689 & 0.7839\\0.7839 & 0.4489\end{pmatrix}.$$

---

b) Compute S+1 and S1, and S, the within-class scatter matrix.
Ans:

$$S_1 = \sum_{x_i-->D_1}(x_i-\mu_1)(x_i-\mu_1)^T = (x_1-\mu_1)(x_1-\mu_1)^T + (x_2-\mu_1)(x_2-\mu_1)^T + (x_3-\mu_1)(x_3-\mu_1)^T =$$

$$\begin{pmatrix}-0.33\\-0.33\end{pmatrix}\begin{pmatrix}-0.33 & -0.33\end{pmatrix} + \begin{pmatrix}0.67\\-0.33\end{pmatrix}\begin{pmatrix}0.67 & -0.33\end{pmatrix} + \begin{pmatrix}-0.33\\0.67\end{pmatrix}\begin{pmatrix}-0.33 & 0.67\end{pmatrix} =$$

$$\begin{pmatrix}0.1089 & 0.1089\\0.1089 & 0.1089\end{pmatrix} + \begin{pmatrix}0.4489 & -0.221\\-0.221 & 0.1089\end{pmatrix} + \begin{pmatrix}0.1089 & -0.221\\-0.221 & 0.4489\end{pmatrix} = \begin{pmatrix}0.667 & -0.33\\-0.33 & 0.667\end{pmatrix}$$

$$S\_1 = \sum_{x_i-->D\_1}(x_i-\mu\_1)(x_i-\mu\_1)^T = (x_4-\mu\_1)(x_4-\mu\_1)^T + (x_5-\mu\_1)(x_5-\mu\_1)^T =$$

$$\begin{pmatrix}-0.5\\0\end{pmatrix}\begin{pmatrix}-0.5 & 0\end{pmatrix} + \begin{pmatrix}0.5\\0\end{pmatrix}\begin{pmatrix}0.5 & 0\end{pmatrix} =$$

$$\begin{pmatrix}0.25 & 0\\0 & 0\end{pmatrix} + \begin{pmatrix}0.25 & 0\\0 & 0\end{pmatrix} = \begin{pmatrix}0.5 & 0\\0 & 0\end{pmatrix}$$

$$S = \begin{pmatrix}0.667 & -0.33\\-0.33 & 0.667\end{pmatrix} + \begin{pmatrix}0.5 & 0\\0 & 0\end{pmatrix} = \begin{pmatrix}1.167 & -0.33\\-0.33 & 0.667\end{pmatrix}$$

---

c) Visually sketch the best direction that splits the data into the two classes

In [6]:
```
matrixS = np.array([[1.167,-0.33],[-0.33,0.667]])
invMatrixS = np.linalg.inv(matrixS)
matrixB = np.array([[1.3689,0.7839],[0.7839,0.4489]])
np.linalg.eig(np.dot(invMatrixS,matrixB))
```

Out[6]: (array([2.91908695e+00, -2.22044605e-16]), array([[0.65092464, -0.49693748],
       [0.75914235, 0.86778635]]))

In [7]:
```
w = np.array([ 0.65092464, 0.75914235])
plt.plot(w)
```

Out[7]: [<matplotlib.lines.Line2D at 0x7f544fc48278>]