ML HW #1

Joshua Shapiro 9/30/16

1)

Given Matrix $A = \begin{bmatrix} 2 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 2 \end{bmatrix}$

a) Are the vectors $x = \begin{bmatrix} 2 \\ -1 \\ 0 \end{bmatrix}$, $y = \begin{bmatrix} -1 \\ 2 \\ -1 \end{bmatrix}$, $z = \begin{bmatrix} 0 \\ -1 \\ 2 \end{bmatrix}$ linearly independent? Justify.

Linear independence implies that the only numbers a, b, c solving the following equation are equal to zero. $ax+by+cz=0$. This happens when the determinant of the matrix of these vectors $\neq 0$.

Calculating the determinant:

$$\begin{bmatrix} 2 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 2 \end{bmatrix} = 2(2\cdot 2-(-1\cdot -1)) = 6$$

$$\begin{bmatrix} 2 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 2 \end{bmatrix} = 1(-1\cdot 2 - 0\cdot -1) = -2$$

$$\begin{bmatrix} 2 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 2 \end{bmatrix} = 0(-1\cdot -1 - 0\cdot 2) = 0$$

$$6-2+0 = \underline{4}$$

Since the determinant of the matrix $\neq 0$, the vectors x, y, z are linearly independent.

b) Find the eigenvalues and corresponding eigenvectors of A.

Eigenvalues:

$$\det(A-\lambda I) = \left| \begin{bmatrix} 2 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 2 \end{bmatrix} - \lambda \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \right| = \left| \begin{bmatrix} 2-\lambda & -1 & 0 \\ -1 & 2-\lambda & -1 \\ 0 & -1 & 2-\lambda \end{bmatrix} \right|$$

$$= [(2-\lambda)(4-4\lambda+\lambda^2-1)] + [1(-2+\lambda)] + [0(1-0)] = 6-8\lambda+2\lambda^2-3\lambda+4\lambda^2-\lambda^3-2+\lambda$$

$$= -\lambda^3+6\lambda^2-10\lambda+4 \Rightarrow \lambda = 2,\ 2-\sqrt{2},\ 2+\sqrt{2} \leftarrow \text{eigenvalues}$$

Eigenvectors:

eigenvectors

$\lambda = 2$: $\begin{bmatrix} 0 & -1 & 0 \\ -1 & 0 & -1 \\ 0 & -1 & 0 \end{bmatrix} \Rightarrow \begin{matrix} -y = 0 \\ -x-z=0 \\ -y=0 \end{matrix} \Rightarrow -x = z \Rightarrow V = \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}$

$\lambda = 2-\sqrt{2}$: $\begin{bmatrix} \sqrt{2} & -1 & 0 \\ -1 & \sqrt{2} & -1 \\ 0 & -1 & \sqrt{2} \end{bmatrix} \Rightarrow \begin{matrix} \sqrt{2}x - y = 0 \Rightarrow x = y/\sqrt{2} \\ -x+\sqrt{2}y-z=0 \Rightarrow -y/\sqrt{2}+\sqrt{2}y-y/\sqrt{2}=0 \Rightarrow \sqrt{2}y = 2y/\sqrt{2} \\ -y+\sqrt{2}z=0 \Rightarrow z = y/\sqrt{2} \end{matrix} \Rightarrow V = \begin{bmatrix} 1 \\ \sqrt{2} \\ 1 \end{bmatrix}$

$\lambda = 2+\sqrt{2}$: $\begin{bmatrix} -\sqrt{2} & -1 & 0 \\ -1 & -\sqrt{2} & -1 \\ 0 & -1 & -\sqrt{2} \end{bmatrix} \Rightarrow \begin{matrix} -\sqrt{2}x - y = 0 \Rightarrow x = -y/\sqrt{2} \\ -x-\sqrt{2}y-z=0 \Rightarrow y/\sqrt{2}-\sqrt{2}y+y/\sqrt{2}=0 \Rightarrow \\ -y-\sqrt{2}z=0 \Rightarrow z = -y/\sqrt{2} \end{matrix}\ V = \begin{bmatrix} 1 \\ -\sqrt{2} \\ 1 \end{bmatrix}$

c) Let M be any matrix w/ real entries. M is positive semidefinite if, for any vector x with real components, the dot product of Mx and x is non-negative, $\langle Mx, x \rangle \geq 0$

Let $B = \begin{bmatrix} 1 & 2 \\ -2 & 1 \end{bmatrix}$. Show that B is positive semidefinite.

$x = \begin{bmatrix} a \\ b \end{bmatrix}$ where a & b are real numbers.

$$Bx^T x = \begin{bmatrix} 1 & 2 \\ -2 & 1 \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix} \cdot \begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} a+2b \\ -2a+b \end{bmatrix}^T \begin{bmatrix} a \\ b \end{bmatrix} = a(a+2b) + b(-2a+b) = a^2 + 2ab - 2ab + b^2 = a^2 + b^2$$

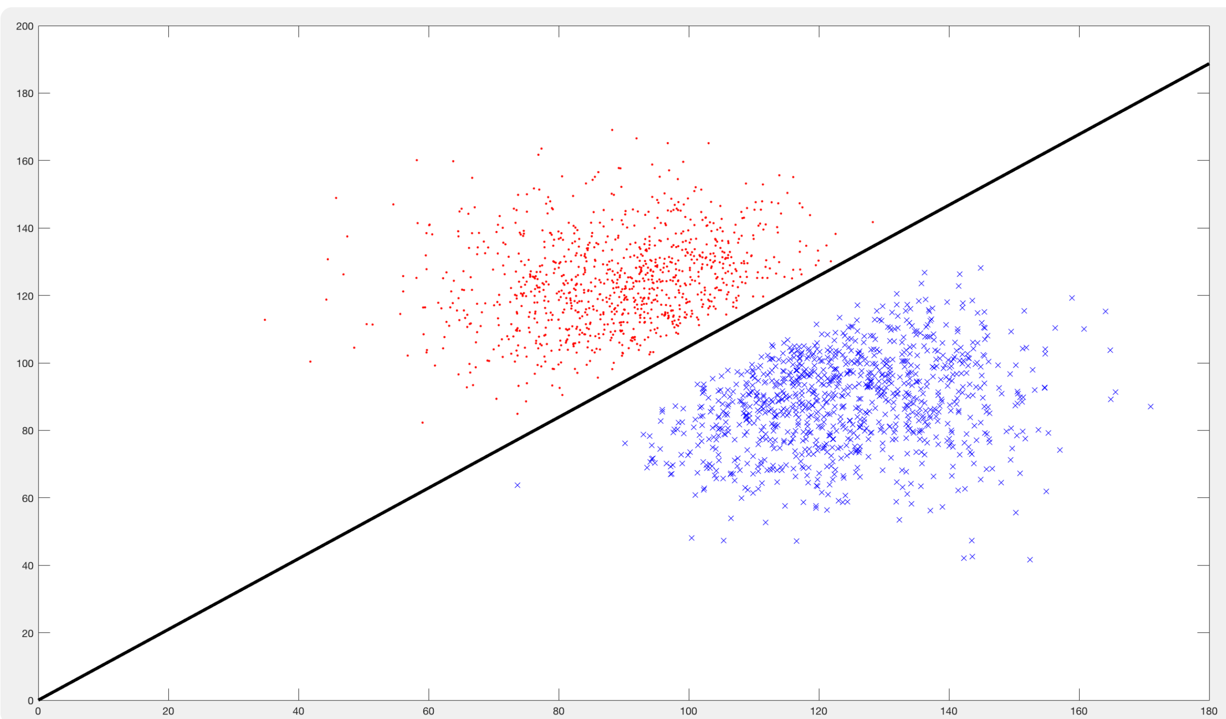
Since the square of any real number is non-negative, $a^2 + b^2$ will always be non-negative. This means $Bx^T x$ is non-negative which means B is positive semidefinite.

d) A symmetric matrix H is positive semidefinite iff the eigenvalues of H are all non-negative. Is matrix A positive semidefinite? Why?

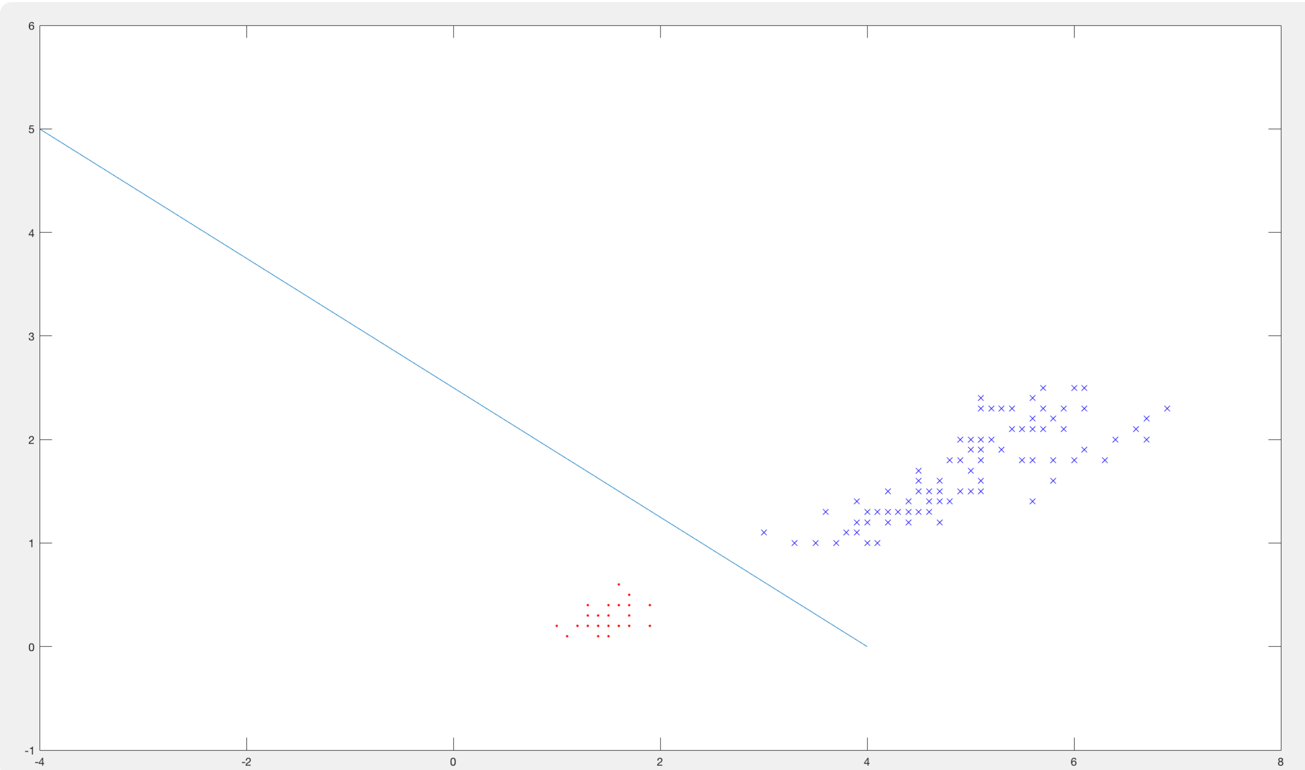
Yes, matrix A is positive semidefinite because it is symmetric and as shown in part b its eigenvalues are all non-negative.

2)

Simply execute the matlab script 'perceptronTest.m' to execute the code and print out the data requested. The script does not output the plot, so I've attached the plots below (as well as some of the data measurements requested).



The angle between theta and [1,0]T is 136.3558 degrees.
The number of updates over the training data is 9.
The margin is 2.0246.



The angle between theta and [1,0,0]T when dealing with IRIS data is 103.0680 degrees.
The number of updates over the IRIS data is 3.
The margin is 0.1135.

For both datasets, the error rate when testing is 0.

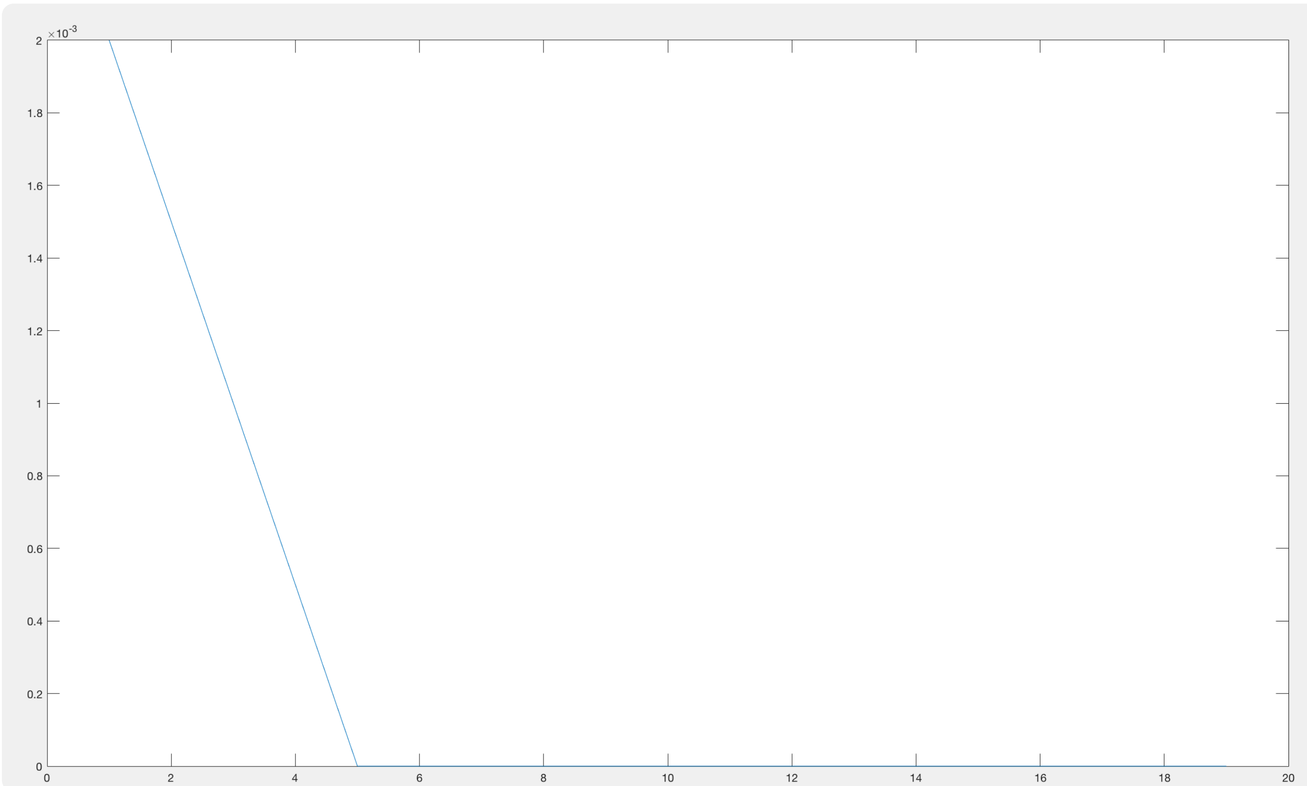
3)

Simply execute the matlab script 'knnPlots.m' to execute the code and print out the data requested. The script does not output both plots, so I've attached both below (as well as some of the data measurements requested).

For k values 1 3 5 7 9 11 13 15 17 19,the error rates are shown below. They have also been graphed in the plot below.

K:	1	3	5	7	9	11	13	15	17	19
Error:	.002	.001	0	0	0	0	0	0	0	0

When using the means instead of the training data, K could only be set to 1. This resulted in an error of 0. The line for this plot has been commented out in the code as it is somewhat pointless.



Regarding why the euclidean distance alternative function works:

We know $\|p\| = \sqrt{p_1^2+p_2^2+\ldots+p_n^2} = \sqrt{p\cdot p}$ $\quad \|q-p\| = \sqrt{(q-p)\cdot(q-p)} = \sqrt{\|p\|^2+\|q\|^2-2p\cdot q}$

We can rewrite $\|q-p\|$ as $\|q-p\|^2 = \|p\|^2+\|q\|^2-2p\cdot q$

When we calculate sqlength, we are calculating $\|x\|^2$ for every row of X. This leaves us with a column vector sqlength that includes the squared euclidean distance for every row-vector in X. Each element in this vector is effectively a $\|p\|^2$ value in the last formula above.

The $-2p\cdot q$ from the formula is also handled in the alternative function provided by the $-2\,X\,z'$ part of dist = sqlength - 2 X z';. So the only piece of the formula missing is $\|q\|^2$. In our case this implies we should be calculating $\|z\|^2$. But this value is just a scalar, and we are adding it, so it would be increasing all of our distances by a constant. Therefore, the minimum values of our dist calculation will be the same regardless of if we incorporate this term.