Question 1

Using Kernel 0



Using Kernel 1



Using Kernel 2



Using Kernel 3



Question 2

a)

Referring to the equation, we want to minimize 0.5||w||^2, which implies we should minimize w. Using the constraint that it is subjected to, to minimize w we have:

y(w.x + w0) = 1

as bias b = 0,

y(w.x ) = 1

Using the points given we have and setting our w as [w1 w2]:

For x1 = [1 1] transpose and y1 = 1:

1 \* [w1 w2] . [1 1].T = 1

w1 + w2 = 1

For x2 = [1 0] transpose and y2 = -1:

(-1) \* [w1 w2] . [1 0].T = 1

w1 = -1

Therefore we get w1 = -1, w2 = 2

Thus our **w is [-1 2]**

To get , we use y(w.x + w0)/||w||

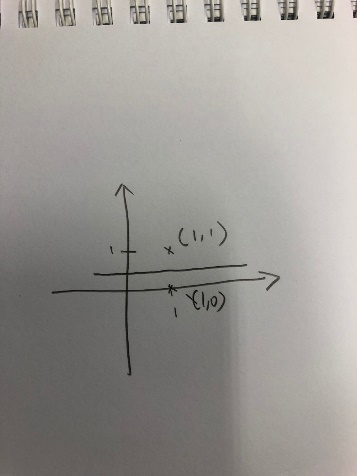
For x1 = [1 1] transpose and y1 = 1:

1 \* ([-1 2] . [1 1].T + 0)/(-12 + 22)0.5

Hence,  **= 1/50.5**

b)

If bias b is allowed to be non-zero, this would allow the classifier to be optimized as it is no longer restricted and results in optimal and w values. Given this, since the support vectors x1 and x2 are exactly 1 unit apart as seen in the following drawing (sorry I was rushing and could not plot a better one so I drew it out PS), we can find the best fit line as shown:



The margin can hence be deduced to be ½

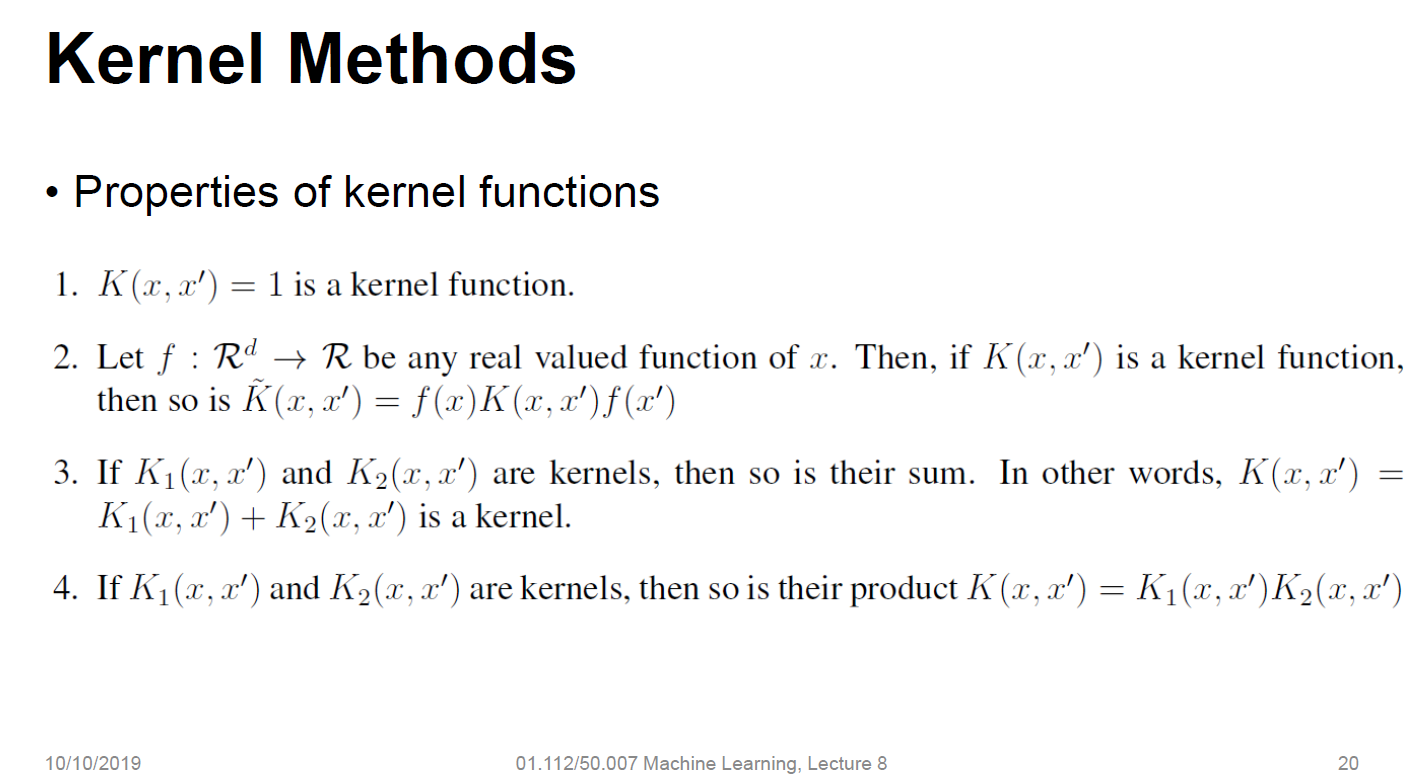
Hence, **=0.5**

Using the above graph, to calculate the gradient of the line we can take 2 points A(0, 0.5) and B(1, 0.5). in this case, dy/dx would be 0/1.

For every change in dx by 1, dy would change by 0. Resulting in the tangent vector being [1 0]. To get w, we can find the normal of this vector which would be [0 1], hence

**w = [0 1]**

Question 3



The above is reference for answering the following questions.

1. K(x, z) = K1(x, z)K2(x, z) is a Kernel

**Proof:**

Using the 4th property from the above and given that:

K1(x; z) and K2(x; z) are kernels over RnxRn.

since K1(x, z) and K2(x, z) are Kernels, their product K(x, z) will also be a Kernel.

1. K(x; z) = aK1(x; z) + bK2(x; z), where a; b > 0 are real numbers, is also considered a Kernel.

**Proof:**

Using the 3rd property from above and given that:

K1(x; z) and K2(x; z) are kernels over RnxRn.

,As aK1(x, z) and bK2(x, z) are Kernels and a, b are real numbers greater than 0, the summation of these two Kernels which results in K(x, z) is a Kernel as well.

1. K(x; z) = aK1(x; z) - bK2(x; z), where a; b > 0 are real numbers

In this case, K(x, z) is not a kernel as there is the potential for it to be less than 0. Given that **Kernels have the property of being non-negative**, the next part will be to prove that this is indeed not a Kernel

**Proof:**

As a and b are given to be real numbers that are greater than 0, we can take any non negative values. In this case lets take **a = 9** and **b = 17.**

Also given that

K1(x; z) and K2(x; z) are kernels over RnxRn.

Using the 1st property in the properties list above, K1(x, z) and K2(x, z) can be taken to be equal to 1.

Plugging these values into the given equation of:

K(x, z) = aK1(x, z) – bK2(x, z)

We get

K(x, z) = (9 \* 1) – (17 \* 1)

This would result in K(x, z) = -8 which is a value lesser than 0.

As this does not follow the property of non-negativity, we can prove that in this case, K is not a Kernel.