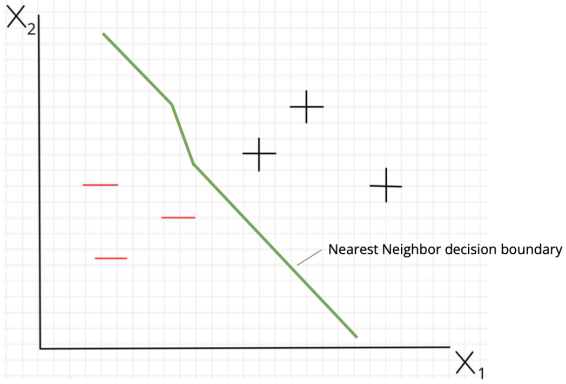
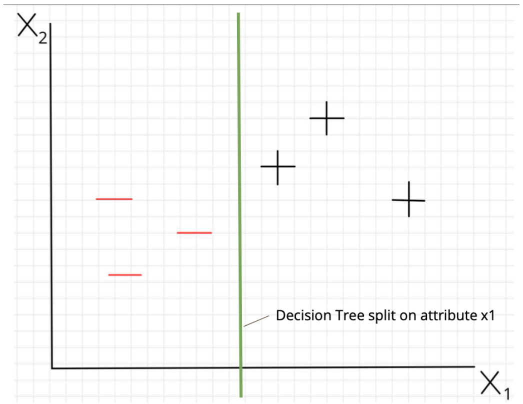
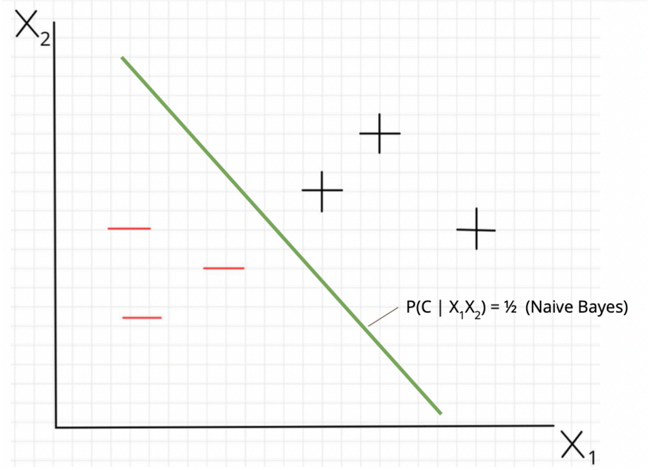
CAS CS 506
Lec 14

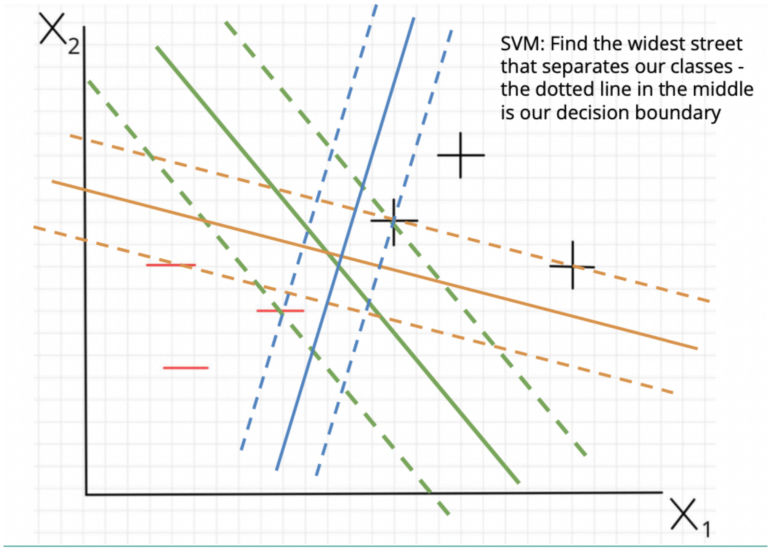
Support Vector Machines

1. Introduction
    a. Given dataset, we can perform classification in various ways
    b. 

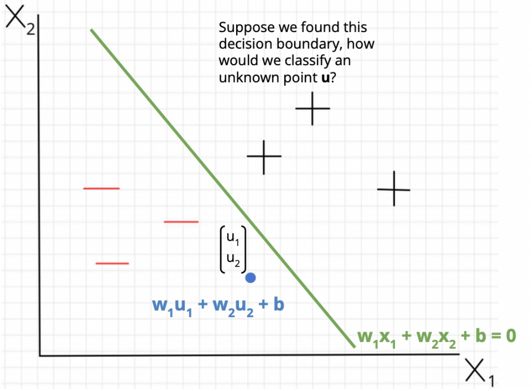
    c. 

    d. 

e. SVM: Find the widest street that separates our classes - the dotted line in the middle is the decision boundary

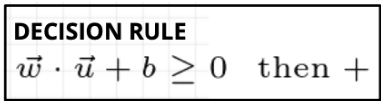


SVM: Find the widest street that separates our classes - the dotted line in the middle is our decision boundary

f. How do we define the street?
   i. What is the format of the equation of the line and decision boundary?
      1. w1x1 + w2x2 + b = 0
      2. w^T*x + b = 0
   ii. Suppose we found the decision boundary, how would we classify an unknown point u?
      1. 
X_2
X_1
$\begin{bmatrix} u_1 \\ u_2 \end{bmatrix}$
$w_1u_1 + w_2u_2 + b$
$w_1x_1 + w_2x_2 + b = 0$

Suppose we found this decision boundary, how would we classify an unknown point **u**?

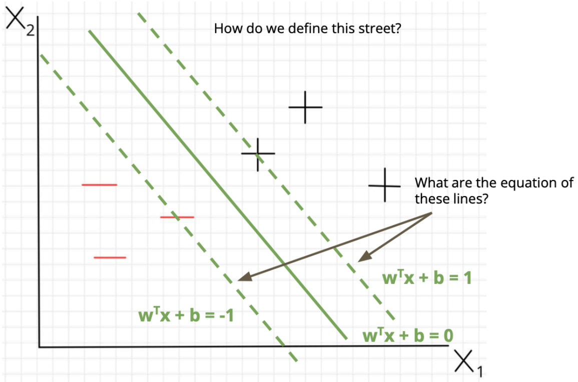
      2. Decision Rule:



DECISION RULE

$\vec{w} \cdot \vec{u} + b \geq 0 \quad \text{then} \; +$

iii. What are the equation of the lines?

1. 
X_2
How do we define this street?
What are the equation of these lines?
$w^Tx + b = 1$
$w^Tx + b = -1$
$w^Tx + b = 0$
X_1

iv. There are many w's and b's since multiplying it by constant c wouldn't move the line

v. What happens if c > 1?

1. A: retracting the line (distance is shortened)

vi. What happens if 0 < c < 1?

1. A: line expands (distance gets longer)

vii. But the decision boundary stays the same

2. Find the Widest Street

a. Means finding the smallest w and b that separates the classes

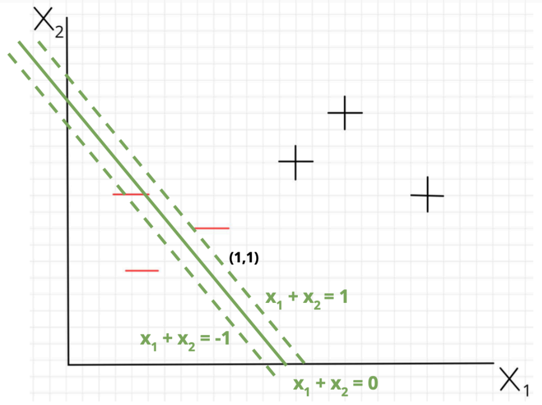
b.

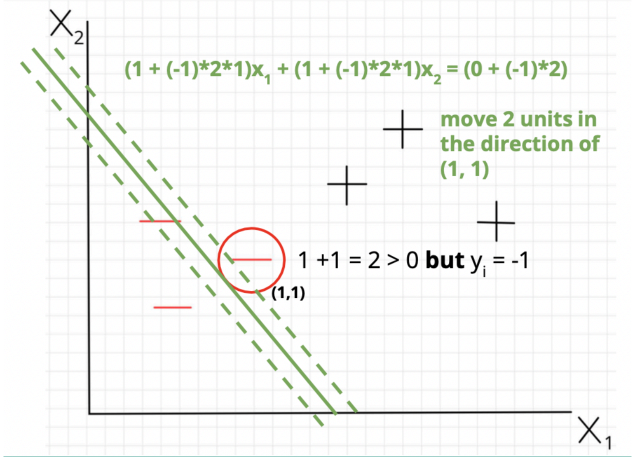
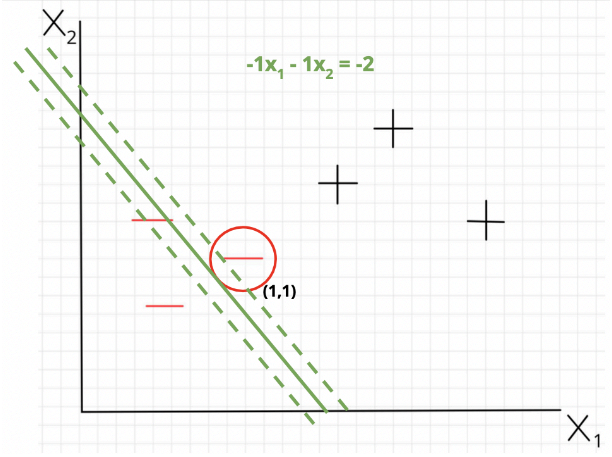
$$y_i = \begin{cases} +1 & \text{if } x_i \text{ is a } + \text{sample} \\ -1 & \text{if } x_i \text{ is a } - \text{sample} \end{cases}$$

c. Assuming the data is linearly separable, we want to impost the constraint that none of the samples can be in the street. That is

$$\vec{w} \cdot \vec{x}_+ + b \geq 1$$

$$\vec{w} \cdot \vec{x}_- + b \leq -1$$

3. Learning w and b
    a. 

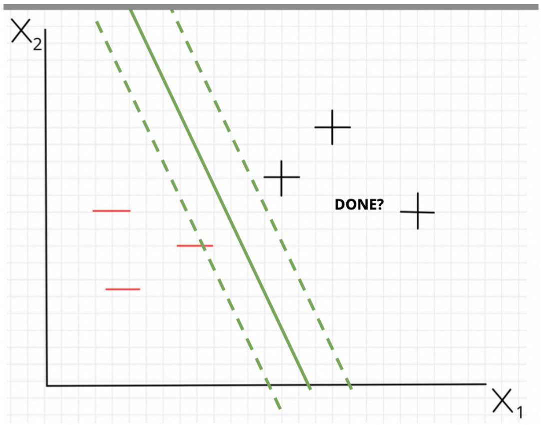
        i. Misclassified the data
        ii. What to do
            1. move the line to the direction of the misclassified point
    b. 
X_2
$(1 + (-1)*2*1)x_1 + (1 + (-1)*2*1)x_2 = (0 + (-1)*2)$
move 2 units in the direction of (1, 1)
$1 + 1 = 2 > 0$ **but** $y_i = -1$
(1,1)
X_1
        i. Therefore, we can move 2 units in the direction of (1,1)
    c. 

4. To Move the Street in the Direction of a Point
    a. Pick a step size a, in order to move a step in the direction of x

$$\mathbf{w}_{new} = \mathbf{w}_{old} + y_i * x * a$$

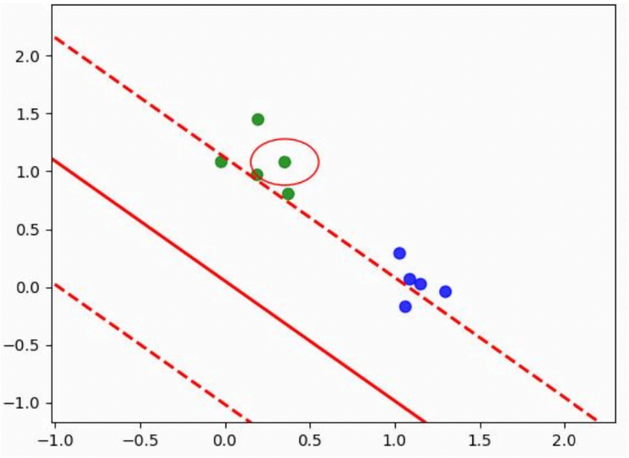
b. $$\mathbf{b}_{new} = \mathbf{b}_{old} + y_i * a$$

c. We also want to find the widest street

5. Now We Know How to Move the Street in the Right Direction - but…
    a. 

        i. Not done here
6. Full Algorithm (Perceptron Algorithm)
    a. Start with random line $w_1x_1 + w_2x_2 + b = 0$
    b. Define
        i. A total number of iterations (ex: 100)
        ii. A learning rate a (not too big not too small)
        iii. An expanding rate c (<1 but not too close to 1)
    c. Repeat number of iterations times:
        i. Pick a point (x_i, y_i) from the dataset
        ii. If correctly classified: do nothing
        iii. If incorrectly classified:
            1. Adjust w_1 by adding $(y_i * a * x_1)$, w_2 by adding $(y_i * a * x_2)$, and b by adding $(y_i * a)$
        iv. Expand or retract the width by c (multiply the new line by c)

d.


2.0
1.5
1.0
0.5
0.0
−0.5
−1.0
−1.0
−0.5
0.0
0.5
1.0
1.5
2.0