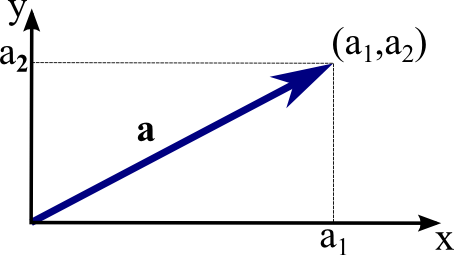
**Mathematical Intuition Behind Support Vector Machine**

What is a vector?

A vector is an object that has both magnitude and direction.



Point a(a1,a2) is plotted in R2 plane. Vector is starting at the origin and ending at a.

**Norm of a vector**

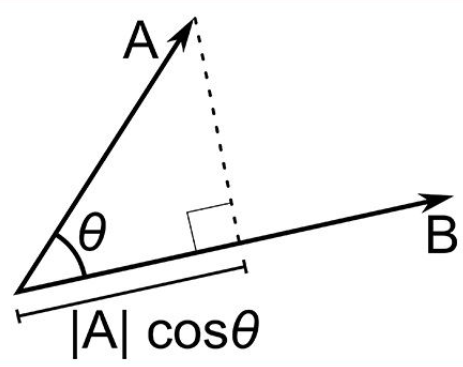
The magnitude or length of a vector is written as ||a|| and is called its **norm**.

**Unit Vector**

A vector that has a magnitude of 1 is a **unit vector**.

Just like numbers we can use mathematical operations such as addition, multiplication on vectors. In this section, we will try to learn about the multiplication of vectors which can be done in two ways, **dot product** and **cross product**. The difference is only that the dot product is used to get a scalar value as a resultant whereas cross-product is used to obtain a vector again.

**The dot product can be defined as the projection of one vector along with another**, multiply by the product of another vector.



Here a and b are 2 vectors, to find the dot product between these 2 vectors we first find the magnitude of both the vectors and to find magnitude we use the Pythagorean theorem or the distance formula.

After finding the magnitude we simply multiply it with the cosine angle between both the vectors. Mathematically it can be written as:

**A.B = |A|** **cosθ \* |B|**

Where, |A| cosθ is the projection of A on B

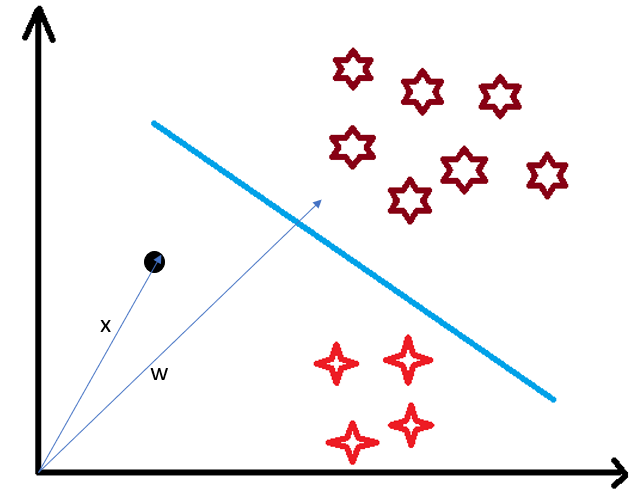
And |B| is the magnitude of vector B

Now in SVM we just need the projection of A not the magnitude of B, I’ll tell you why later. To just get the projection we can simply take the unit vector of B because it will be in the direction of B but its magnitude will be 1. Hence now the equation becomes:

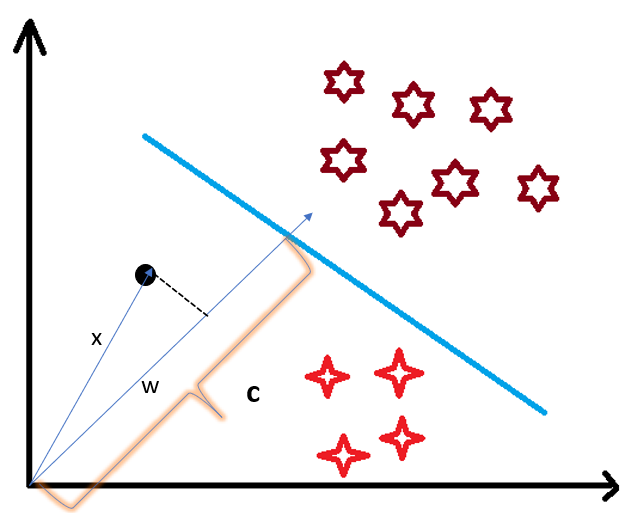
**A.B = |A|** **cosθ \* unit vector of B**

#### **Use of Dot Product in SVM**

Consider a random point X and we want to know whether it lies on the right side of the plane or the left side of the plane (positive or negative).



To find this first we assume this point is a vector (X) and then we make a vector (w) which is perpendicular to the hyperplane. Let’s say the distance of vector w from origin to decision boundary is ‘c’. Now we take the projection of X vector on w.

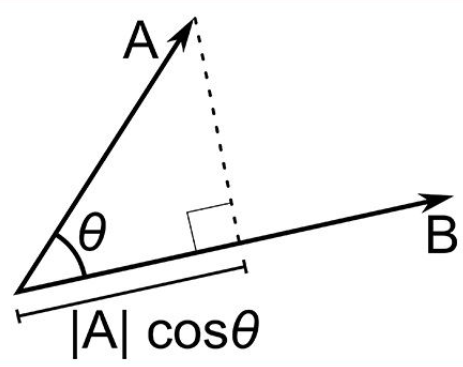


## Mathematical Intuition Behind Support Vector Machine

### **Understanding Dot-Product**

We all know that a vector is a quantity that has magnitude as well as direction and just like numbers we can use mathematical operations such as addition, multiplication. In this section, we will try to learn about the multiplication of vectors which can be done in two ways, dot product, and cross product. The difference is only that the dot product is used to get a scalar value as a resultant whereas cross-product is used to obtain a vector again.

The dot product can be defined as the projection of one vector along with another, multiply by the product of another vector.



Here a and b are 2 vectors, to find the dot product between these 2 vectors we first find the magnitude of both the vectors and to find magnitude we use the Pythagorean theorem or the distance formula.

After finding the magnitude we simply multiply it with the cosine angle between both the vectors. Mathematically it can be written as:

**A . B = |A|** **cosθ \* |B|**

Where |A| cosθ is the projection of A on B

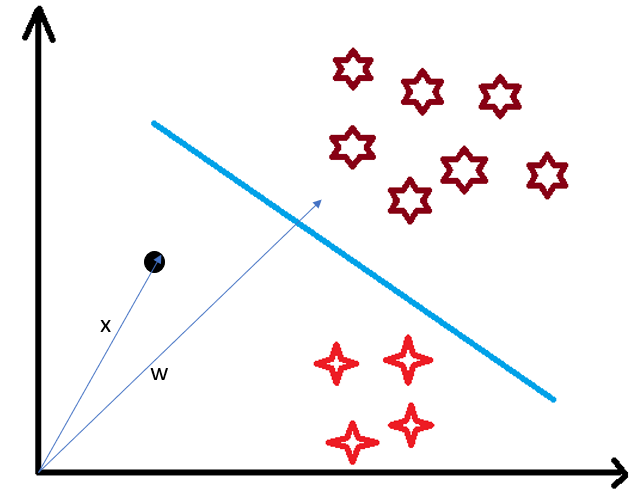
And |B| is the magnitude of vector B

Now in SVM we just need the projection of A not the magnitude of B, I’ll tell you why later. To just get the projection we can simply take the unit vector of B because it will be in the direction of B but its magnitude will be 1. Hence now the equation becomes:

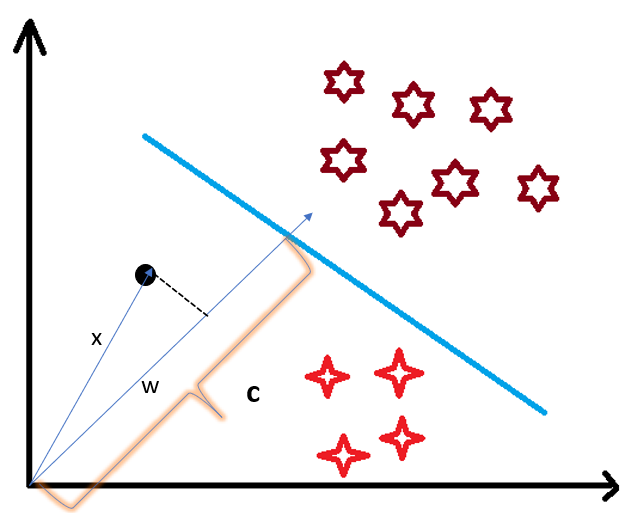
**A.B = |A| cosθ \* unit vector of B**

#### **Use of Dot Product in SVM**

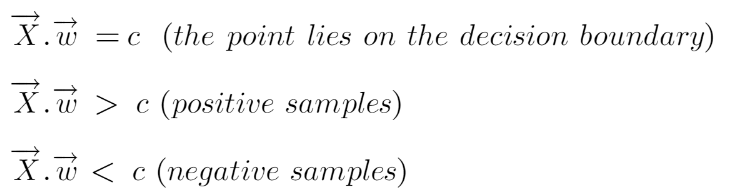
Consider a random point X and we want to know whether it lies on the right side of the plane or the left side of the plane (positive or negative).



To find this first we assume this point is a vector (X) and then we make a vector (w) which is perpendicular to the hyperplane. Let’s say the distance of vector w from origin to decision boundary is ‘c’. Now we take the projection of X vector on w.

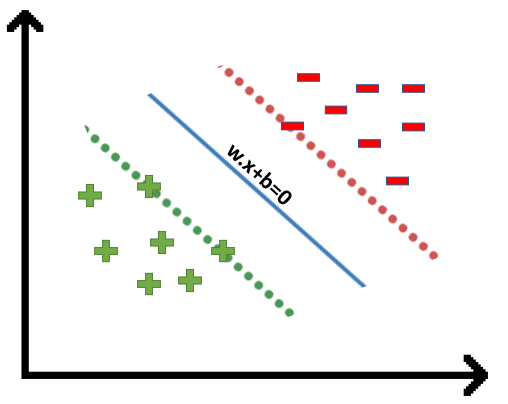


We already know that projection of any vector or another vector is called dot-product. Hence, we take the dot product of x and w vectors. If the dot product is greater than ‘c’ then we can say that the point lies on the right side. If the dot product is less than ‘c’ then the point is on the left side and if the dot product is equal to ‘c’ then the point lies on the decision boundary.

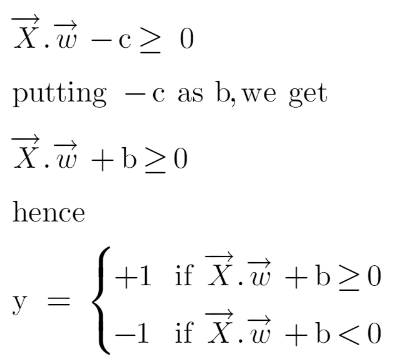


## Margin in Support Vector Machine

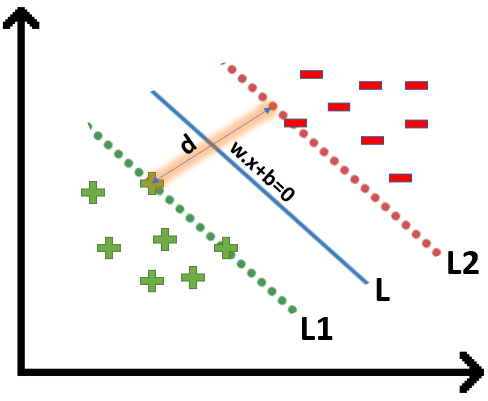
We all know the equation of a hyperplane is w.x+b=0 where w is a vector normal to hyperplane and b is an offset or bias.



To classify a point as negative or positive we need to define a decision rule. We can define decision rule as:

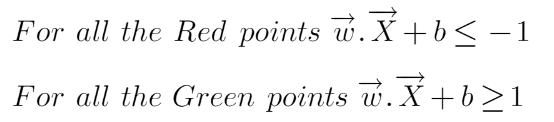


if the value of w.x+b>0 then we can say it is a positive point otherwise it is a negative point. Now we need (w,b) such that the margin has a maximum distance. Let’s say this distance is ‘d’



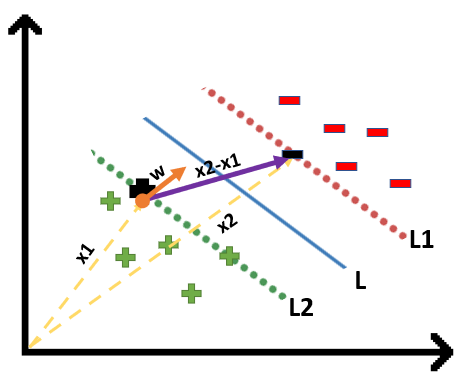
To calculate ‘d’ we need the equation of L1 and L2. For this, we will take few assumptions that the equation of L1 is w.x+b=1 and for L2 it is w.x+b=-1.

**We’ll calculate the distance (d) in such a way that no positive or negative point can cross the margin line”.**Let’s write these constraints mathematically:



We assume that negative classes have **y=-1** and positive classes have **y=1.**

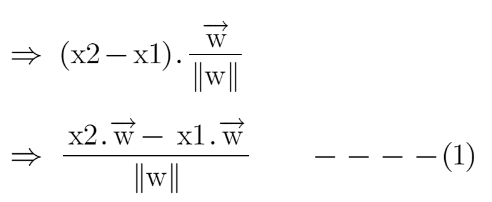
We will take 2 support vectors, 1 from the negative class and 2nd from the positive class. The distance between these two vectors x1 and x2 will be (x2-x1) vector. What we need is, the shortest distance between these two points which can be found using a trick we used in the dot product. We take a vector ‘w’ perpendicular to the hyperplane and then find the projection of (x2-x1) vector on ‘w’. Note: this perpendicular vector should be a unit vector.



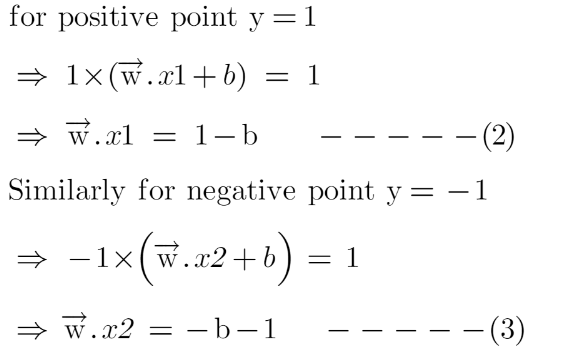
**Finding Projection of a Vector on Another Vector Using Dot Product**

We already know how to find the projection of a vector on another vector.

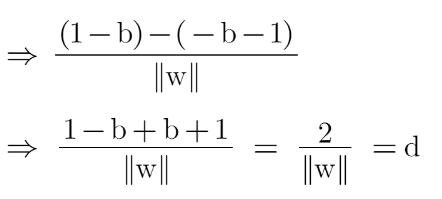
We do this by dot-product of both vectors.



Since x2 and x1 are support vectors and they lie on the hyperplane, hence they will follow yi\* (2.x+b) =1 so we can write it as:



Putting equations (2) and (3) in equation (1) we get:



When, ||w|| = 1, then d=2

||w|| = 2, then d = 1

||w|| = 4, then d = 0.5

It should be noted bigger the norm, the smaller the margin. Our goal is to maximize the margin. Among all the possible hyperplanes meeting the constraints, we will choose the hyperplane with the smallest ||w|| because it will have the biggest margin.

**Numerical problems on linear and non-linear SVM – done in class. Practise.**