NMK40403 ARTIFICIAL INTELLIGENCE

SUPPORT VECTOR MACHINES (B)

Mohamed Elshaikh

SVM - Understanding the math - Part 1 - The margin

- The sum of two vectors

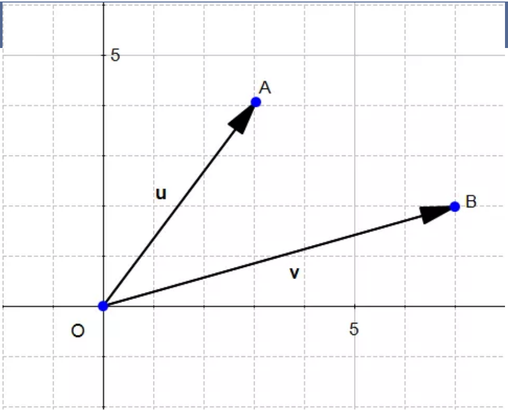


Figure 13

- *Given two vectors $u(u_1,u_2)$ and $v(v_1,v_2)$ then :*

$$u+v=(u_1+v_1,u_2+v_2)$$

- *Which means that adding two vectors gives us a third vector whose coordinate are the sum of the coordinates of the original vectors.*

SVM - Understanding the math - Part 1 - The margin

- Example:

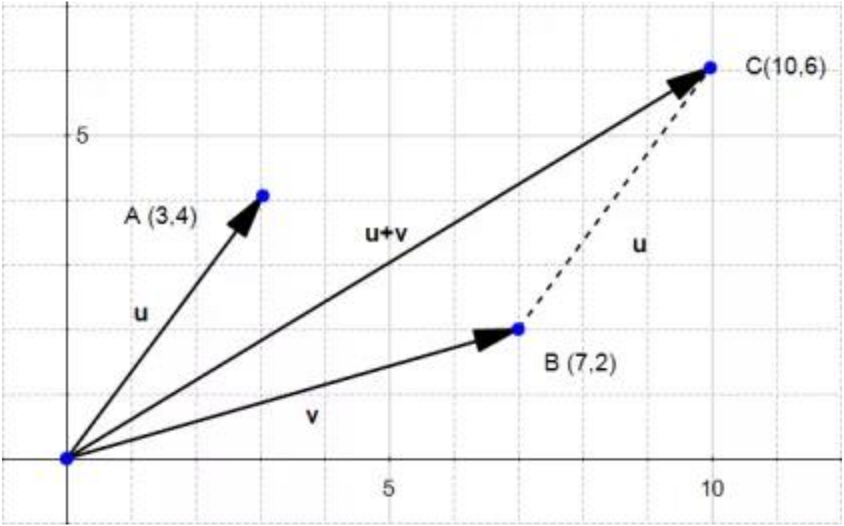


Figure 14: the sum of two vectors

SVM - Understanding the math - Part 1 - The margin

The difference between two vectors

- The difference works the same way :

$$u - v = (u_1 - v_1, u_2 - v_2)$$

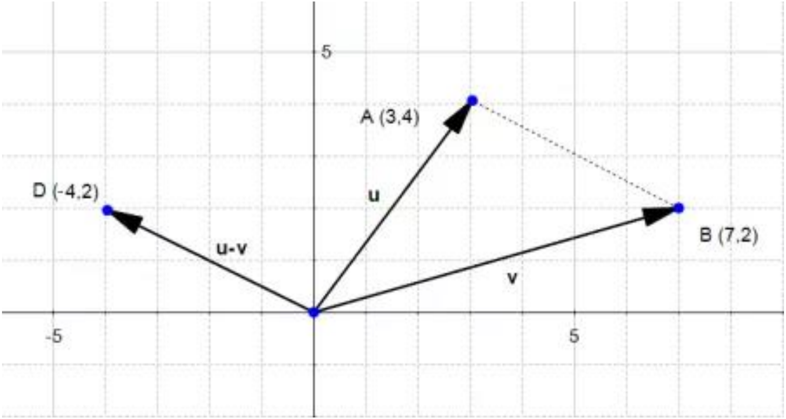


Figure 15: the difference of two vectors

SVM - Understanding the math - Part 1 - The margin

- Since the subtraction is not commutative, we can also consider the other case:

$$v - u = (v_1 - u_1, v_2 - u_2)$$

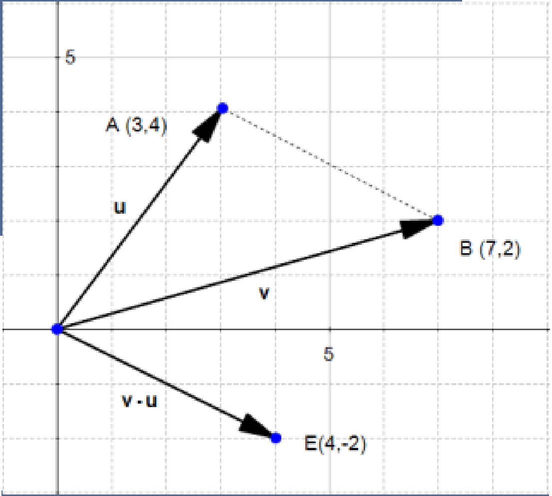


Figure 16: the difference v-u

SVM - Understanding the math - Part 1 - The margin

- The last two pictures describe the "true" vectors generated by the difference of u and v.

- However, since a vector has a magnitude and a direction, we often consider that parallel translate of a given vector (vectors with the same magnitude and direction but with a different origin) are the same vector, just drawn in a different place in space.

- So don't be surprised if you meet the following (Figure 17 (a) & (b) :

SVM - Understanding the math - Part 1 - The margin

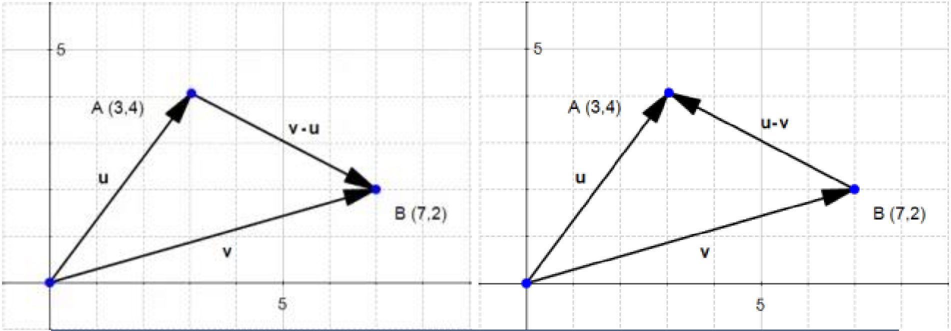


Figure 17 (a) & (b) : another way to view the difference v-u

- If you do the math, it looks wrong, because the end of the vector $u-v$ is not in the right point, but it is a convenient way of thinking about vectors which you'll encounter often.

SVM - Understanding the math - Part 1 - The margin

The dot product

- One very important notion to understand SVM is the dot product.
- Definition:

  Geometrically, it is the product of the Euclidian magnitudes of the two vectors and the cosine of the angle between them

- Which means if we have two vectors x and y and there is an angle θ (theta) between them, their dot product is :

$$x \cdot y = \|x\| \|y\| \cos(\theta)$$

SVM - Understanding the math - Part 1 - The margin

Why ?

- To understand let's look at the problem geometrically.

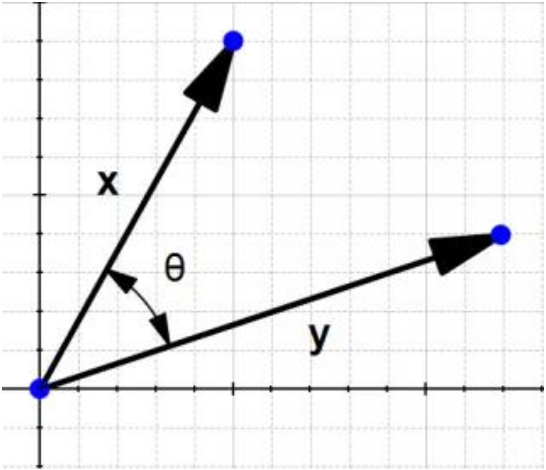


Figure 18

SVM - Understanding the math - Part 1 - The margin

- In the definition, they talk about cos(θ), let's see what it is.

- By definition we know that in a right-angled triangle:

$$\cos(\theta) = \frac{adjacent}{hypotenuse}$$

- In our example, we don't have a right-angled triangle.

- However if we take a different look Figure 18 we can find two right-angled triangles formed by each vector with the horizontal axis.

SVM - Understanding the math - Part 1 - The margin

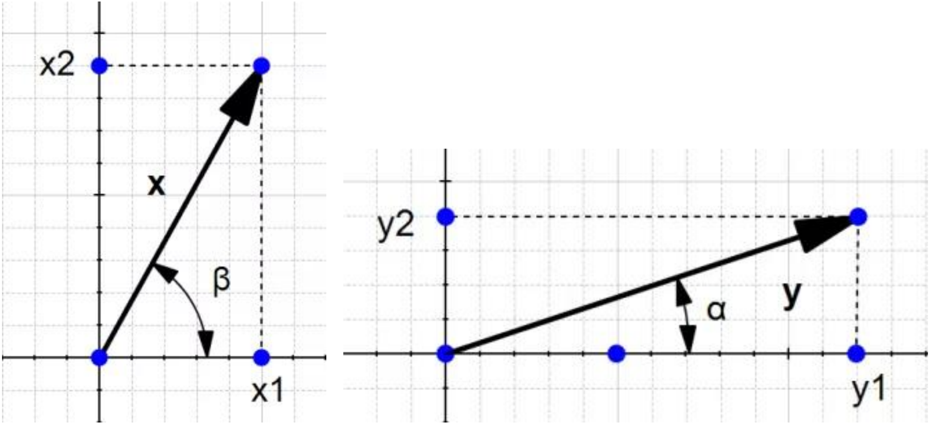


Figure 19 (a) & (b)

SVM - Understanding the math - Part 1 - The margin

- So now we can view our original schema like this:

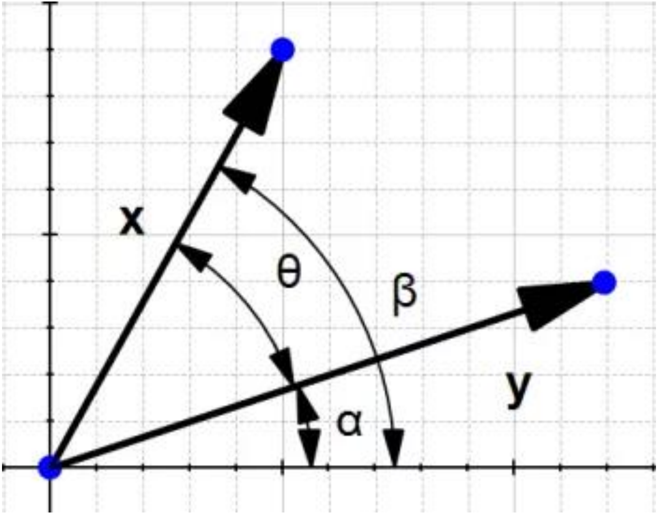


Figure 20

SVM - Understanding the math - Part 1 - The margin

- We can see that

$$\theta = \beta - \alpha$$

- So computing $\cos(\theta)$ is like computing $\cos(\beta - \alpha)$

- There is a special formula called the difference identity for cosine which says that:

$$\cos(\beta - \alpha) = \cos(\beta)\cos(\alpha) + \sin(\beta)\sin(\alpha)$$

SVM - Understanding the math - Part 1 - The margin

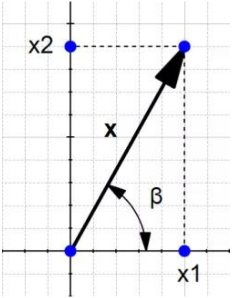
- Let's use this formula!

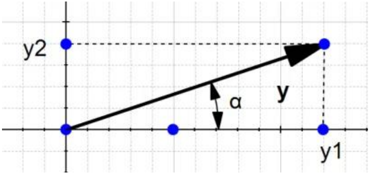
$$\cos(\beta) = \frac{adjacent}{hypotenuse} = \frac{x_1}{||x||}$$

$$\sin(\beta) = \frac{opposite}{hypotenuse} = \frac{x_2}{||x||}$$

$$\cos(\alpha) = \frac{adjacent}{hypotenuse} = \frac{y_1}{||y||}$$

$$\sin(\alpha) = \frac{opposite}{hypotenuse} = \frac{y_2}{||y||}$$





SVM - Understanding the math - Part 1 - The margin

- So if we replace each term

$$\cos(\theta)=\cos(\beta-\alpha)=\cos(\beta)\cos(\alpha)+\sin(\beta)\sin(\alpha)$$

$$\cos(\theta)=\frac{x_1}{\|x\|}\frac{y_1}{\|y\|}+\frac{x_2}{\|x\|}\frac{y_2}{\|y\|}$$

$$\cos(\theta)=\frac{x_1y_1+x_2y_2}{\|x\|\|y\|}$$

- If we multiply both sides by ||x|| ||y|| we get:

$$\|x\|\|y\|\cos(\theta)=x_1y_1+x_2y_2$$

- Which is the same as :

$$\|x\|\|y\|\cos(\theta)=x\cdot y$$

- This is the geometric definition of the dot product !

SVM - Understanding the math - Part 1 - The margin

- Eventually from the two last equations we can see that :

$$x \cdot y = x_1 y_1 + x_2 y_2 = \sum_{i=1}^{2} (x i y i)$$

- This is the algebraic definition of the dot product !

A few words on notation

- The dot product is called like that because we write a dot between the two vectors.
- Talking about the dot product $x \cdot y$ is the same as talking about
  - the ***inner product*** $\langle x, y \rangle$ (in linear algebra)
  - ***scalar product*** because we take the product of two vectors and it returns a scalar (a real number)

SVM - Understanding the math - Part 1 - The margin

The orthogonal projection of a vector

- Given two vectors x and y, we would like to find the orthogonal projection of x onto y.

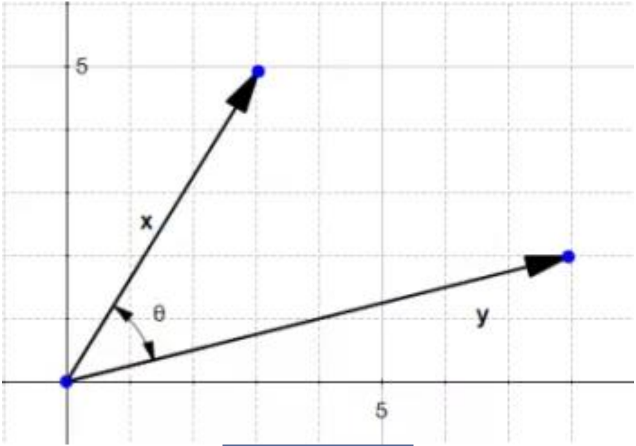


Figure 21

SVM - Understanding the math - Part 1 - The margin

- To do this we project the vector x onto y

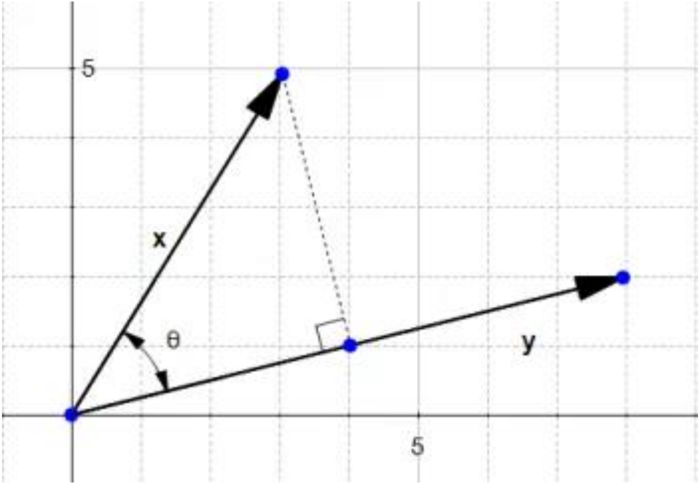


Figure 22

SVM - Understanding the math - Part 1 - The margin

- This give us the vector z

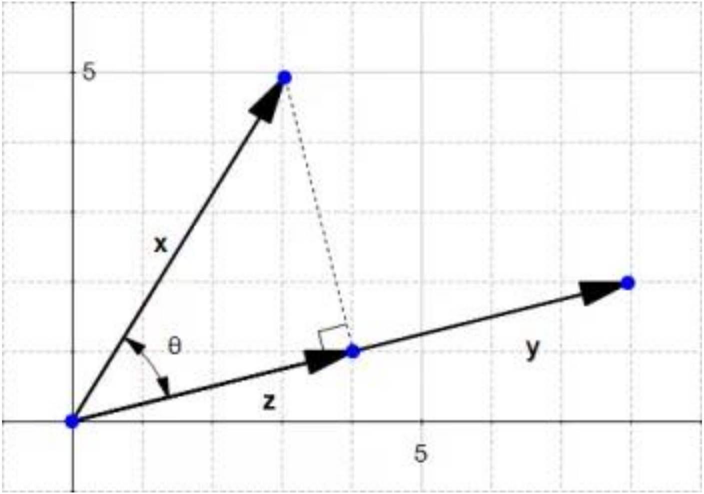


Figure 23 : z is the projection of x onto y

SVM - Understanding the math - Part 1 - The margin

- By definition :

$$\cos(\theta)=\frac{\|z\|}{\|x\|}$$

$$\|z\|=\|x\|\cos(\theta)$$

- We saw in the section about the dot product that

$$\cos(\theta)=\frac{x.y}{\|x\|\|y\|}$$

- So we replace cos(θ) in our equation:

$$\|z\|=\|x\|\frac{x.y}{\|x\|\|y\|}$$

$$\|z\|=\frac{x.y}{\|y\|}$$

SVM - Understanding the math - Part 1 - The margin

- If we define the vector u as the direction of y then

  $u=\frac{y}{\|y\|}$ and $\|z\|=u.x$

- We now have a simple way to compute the norm of the vector z.
- Since this vector is in the same direction as y it has the direction u

  $u=\frac{z}{\|z\|}$

  $z=\|z\|u$

- And we can say :

  *The vector **z=(u·x)u** is the orthogonal projection of **x** onto y.*

SVM - Understanding the math - Part 1 - The margin

- Why are we interested by the orthogonal projection ? Well in our example, it allows us to compute the distance between x and the line which goes through y.

- We see that this distance is ||x−z||

  $$\|x-z\|=\sqrt{(3-4)^2+(5-1)^2}$$
  $$=\sqrt{17}$$

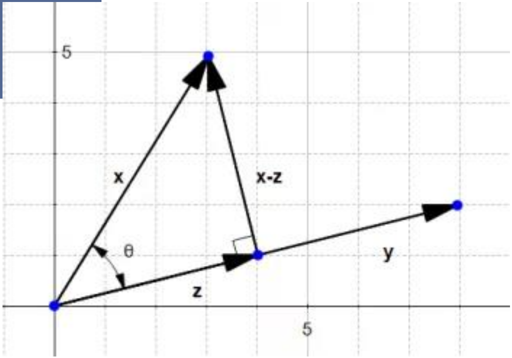


Figure 24

SVM - Understanding the math - Part 1 - The margin

The SVM hyperplane

Understanding the equation of the hyperplane

- You probably learnt that an equation of a line is : $y=ax+b$. However when reading about hyperplane, you will often find that the equation of an hyperplane is defined by :

$$\mathbf{w}^T\mathbf{x}=0$$

- How does these two forms relate ?
- In the hyperplane equation you can see that the name of the variables are in bold. Which means that they are vectors ! Moreover, $\mathbf{w}^T\mathbf{x}$ is how we compute the inner product of two vectors, and if you recall, the inner product is just another name for the dot product !

SVM - Understanding the math - Part 1 - The margin

- Note that $y=ax+b$
- is the same thing as $y-ax-b=0$
- Given two vectors $w \begin{pmatrix} -b \\ -a \\ 1 \end{pmatrix}$ and $x \begin{pmatrix} 1 \\ x \\ y \end{pmatrix}$

$$w^T x = -b \times (1) + (-a) \times x + 1 \times y$$
$$w^T x = y - ax - b$$

- The two equations are just different ways of expressing the same thing.
- It is interesting to note that w_0 is $-b$, which means that this value determines the intersection of the line with the vertical axis.

SVM - Understanding the math - Part 1 - The margin

- Why do we use the hyperplane equation $w^T x$ instead of $y = ax + b$?

- For two reasons:
    - it is easier to work in more than two dimensions with this notation,
    - the vector w will always be normal to the hyperplane (Note: w will always be normal because we use this vector to define the hyperplane, so by definition it will be normal. As you can see, when we define a hyperplane, we suppose that we have a vector that is orthogonal to the hyperplane)

- And this last property will come in handy to compute the distance from a point to the hyperplane.

SVM - Understanding the math - Part 1 - The margin

Compute the distance from a point to the hyperplane

- In Figure 25 we have an hyperplane, which separates two group of data.

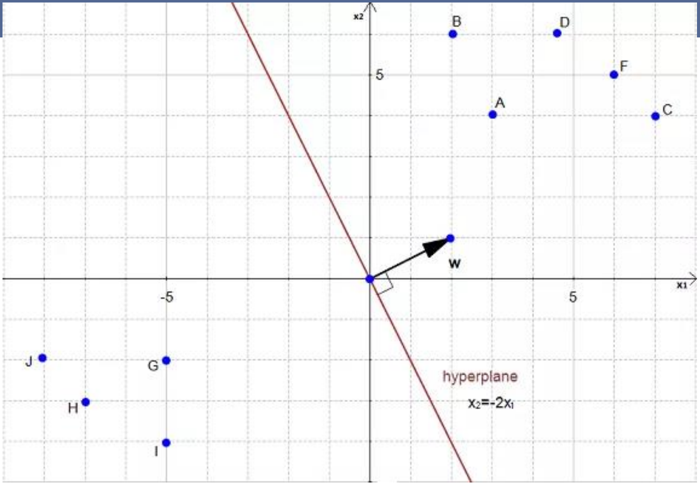


Figure 25

SVM - Understanding the math - Part 1 - The margin

- To simplify this example, we have set w_0=0.

- As you can see on the Figure 25, the equation of the hyperplane is :

$$x_2 = -2x_1$$

- which is equivalent to

$$w^T x$$

- with w$\begin{pmatrix} 2 \\ 1 \end{pmatrix}$ and x$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$
- Note that the vector w is shown on the Figure 25. (w is not a data point)

SVM - Understanding the math - Part 1 - The margin

- We would like to compute the distance between the point A(3,4) and the hyperplane.
- This is the distance between A and its projection onto the hyperplane

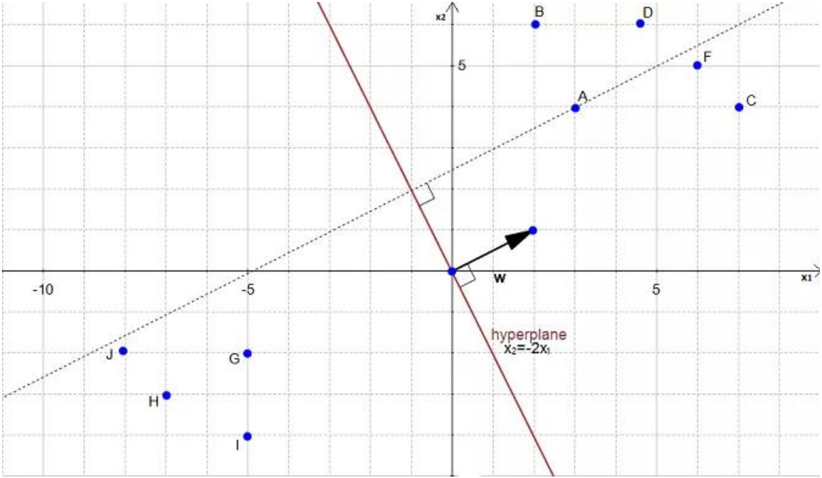


Figure 26

SVM - Understanding the math - Part 1 - The margin

- We can view the point A as a vector from the origin to A.
- If we project it onto the normal vector w

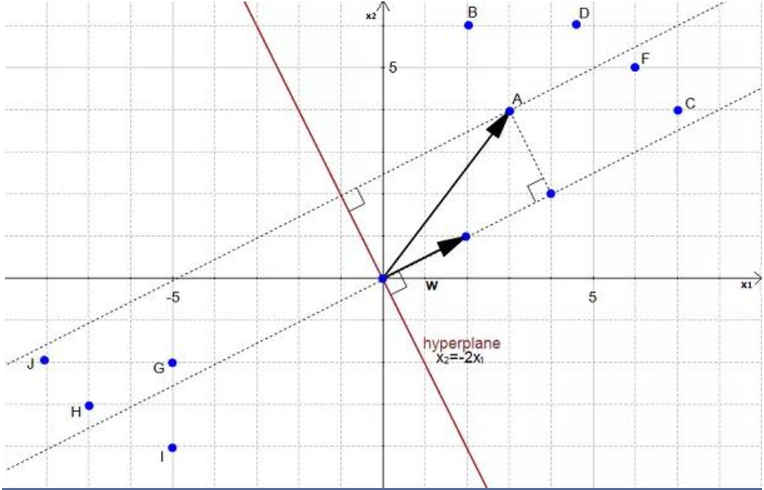


Figure 27: projection of a onto w

SVM - Understanding the math - Part 1 - The margin

- We get the vector p

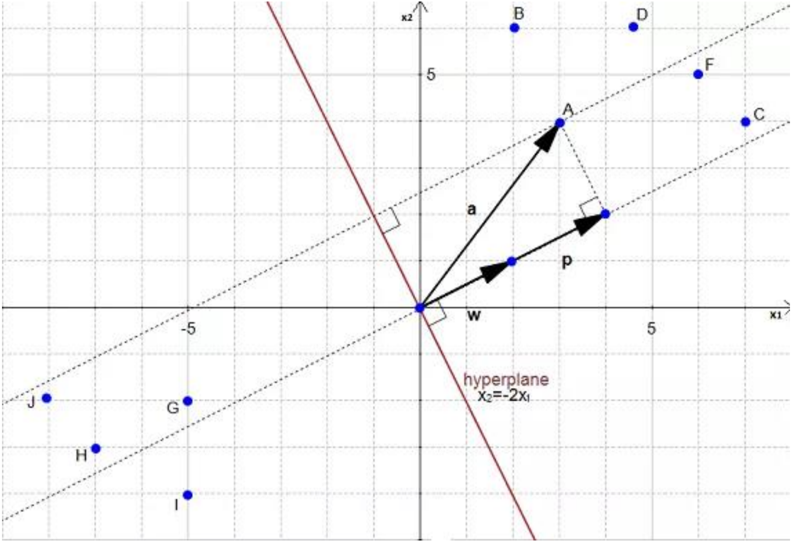


Figure 28: p is the projection of a onto w

SVM - Understanding the math - Part 1 - The margin

- Our goal is to find the distance between the point A(3,4) and the hyperplane. We can see in Figure 28 that this distance is the same thing as ||p||.
- Let's compute this value.

- We start with two vectors, w=(2,1) which is normal to the hyperplane, and a=(3,4) which is the vector between the origin and A.

$$||w|| = \sqrt{2^2 + 1^2} = \sqrt{5}$$

- Let the vector u be the direction of w

$$u = \left(\frac{2}{\sqrt{5}}, \frac{1}{\sqrt{5}}\right)$$

SVM - Understanding the math - Part 1 - The margin

- p is the orthogonal projection of a onto w so :

$$p = (u.a)u$$

$$p = \left(3 \times \frac{2}{\sqrt{5}} + 4 \times \frac{1}{\sqrt{5}}\right)u$$

$$p = \left(\frac{6}{\sqrt{5}} + \frac{4}{\sqrt{5}}\right)u$$

$$p = \frac{10}{\sqrt{5}}u$$

$$p = \left(\frac{10}{\sqrt{5}} \times \frac{2}{\sqrt{5}}, \frac{10}{\sqrt{5}} \times \frac{1}{\sqrt{5}}\right)$$

$$p = \left(\frac{20}{\sqrt{5}}, \frac{10}{\sqrt{5}}\right)$$

$$p = (4,2)$$

$$||p|| = \sqrt{4^2 + 2^2} = 2\sqrt{5}$$

SVM - Understanding the math - Part 1 - The margin

Compute the margin of the hyperplane

- Now that we have the distance ||p|| between A and the hyperplane, the margin is defined by :

$$margin = 2\|p\| = 4\sqrt{5}$$

- We did it ! We computed the margin of the hyperplane !


END