

Support vector machines (SVM)

As explained SVM uses a hyperplane to segregate between classes.

Hyperplane :

For a P -dimensional space, a hyperplane is a flat subspace of dimension $P-1$. For example for a 2 dimensional space, it is 1 line or 1 dimensional. (P dimensions mean P -variables).

Mathematical definition of a hyperplane is

$$\beta_0 + \beta_1 X_1 + \beta_2 X_2 = 0$$

X_1, X_2 are two variables

~~think of X_1 as a row in a matrix of values X_{ij} .~~

If we have p variables, then it becomes a P -dimensional space. The hyperplane there becomes

$$\beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_p X_p = 0$$

$$\text{or } \beta_0 + \sum_{i=1}^p \beta_i X_i = 0$$

$$\beta_0 + \sum_{i=1}^p \beta_i X_i = 0$$

If for a certain i , this does not satisfy. So,

$$\beta_0 + \sum_{i=1}^p \beta_i X_i \text{ is either } > 0 \text{ or } < 0.$$

If > 0 it goes to one class

If < 0 it goes to another class.

Think of 2 classes $[1, -1]$, then

$$\beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \dots + \beta_p X_{ip} > 0 \text{ if } y_i = 1$$

and

$$\beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \dots + \beta_p X_{ip} < 0 \text{ if } y_i = -1.$$

So,

$$y_i (\beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \dots + \beta_p X_{ip}) > 0.$$

Now let's see what is maximum margin classifier.