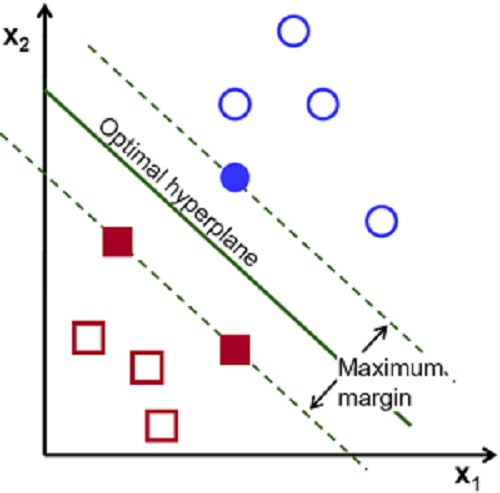
A Support Vector Machine (SVM) is a supervised machine learning algorithm which can be used for both classification and regression problem. The basic SVM formulation is for linearly separable dataset.

Assume there are 2 classes of data, a SVM classifies those different classes by drawing an optimal hyperplane that separates all the classes. It finds a plane that has the maximum margin, i.e the maximum distance between data points of both classes closest to the plane. These data points are called Support vectors



The advantage of SVM, is that it takes into account both experimental data and structural behavior for better generalization capability.

It can be used for non-linear datasets by indirectly mapping the nonlinear inputs into to linear feature space where the maximum margin decision function is approximated. The mapping is done by using a kernel function.

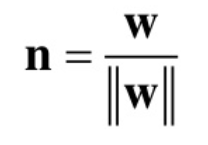
SVMs are learning systems that

* Use a hypothesis space of linear functions
* In a high dimensional feature space - kernel functions
* Trained with a learning algorithm from optimization theory -
* Lagrange
* Implements a learning bias derived from statistical learning
* theory - Generalization

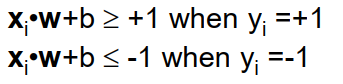
Support vectors - Data points closest to optimal hyperplane.

Concept of SVM - Finding a hyperplane that separates the classes and has the maximum margin between supports vectors of different class.



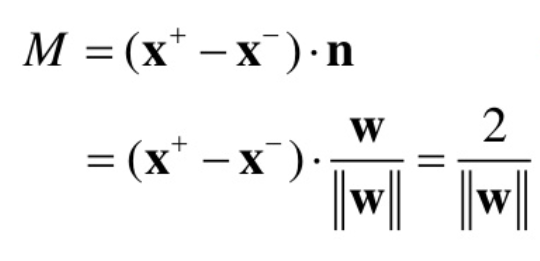
Equation of hyperplane - g(x) =WX + b

Normal vector to hyperplane =

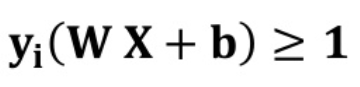
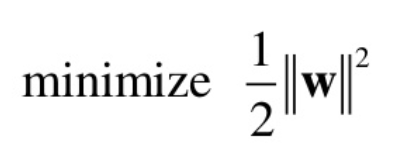
Classifier is constructed as follow 



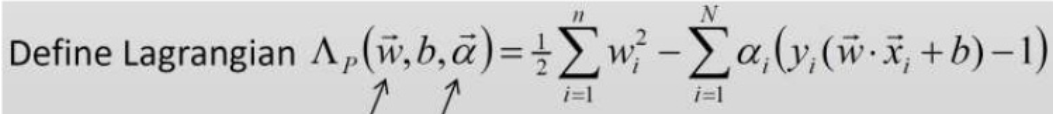
Can be combined into

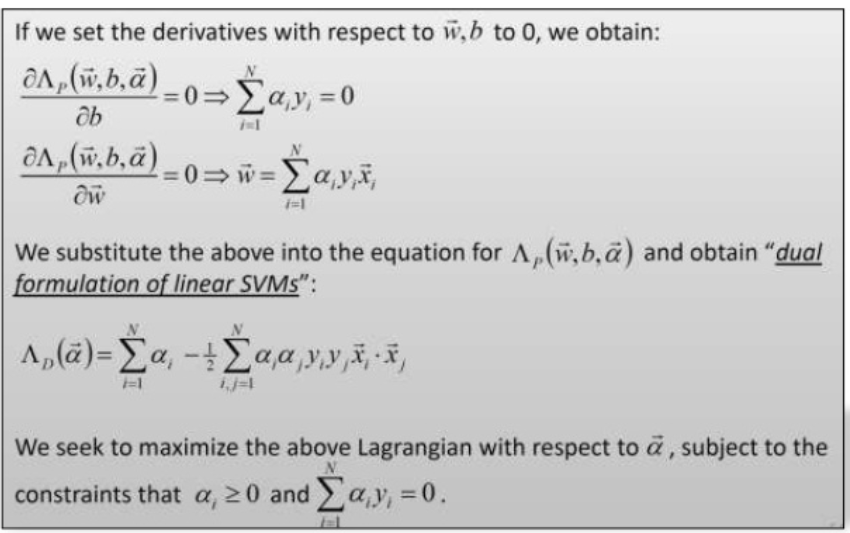


The margin width is



For maximum margin we need to ,with constrain

Applying lagrangian method, we need to maximize the 



Kernel trick

Kernel method in SVM is used when we encounter nonlinear classification

* It involves nonlinear transformation of data into higher dimensional feature space and detects optimal solution in the kernel feature space.

