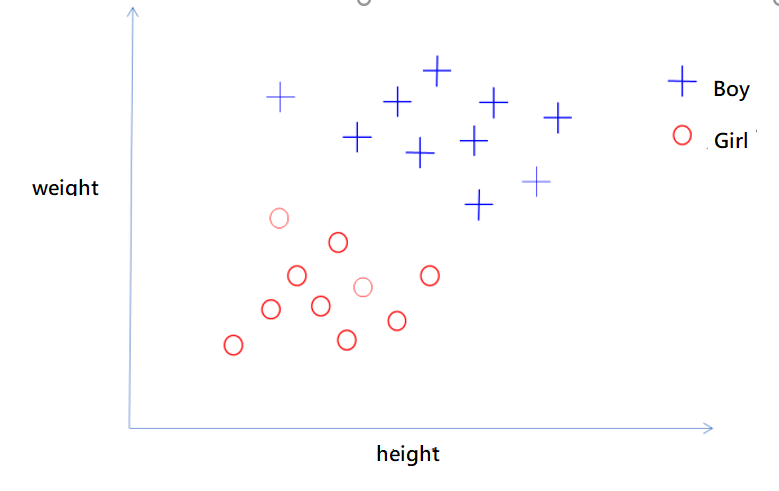
SVM-Detail derivation

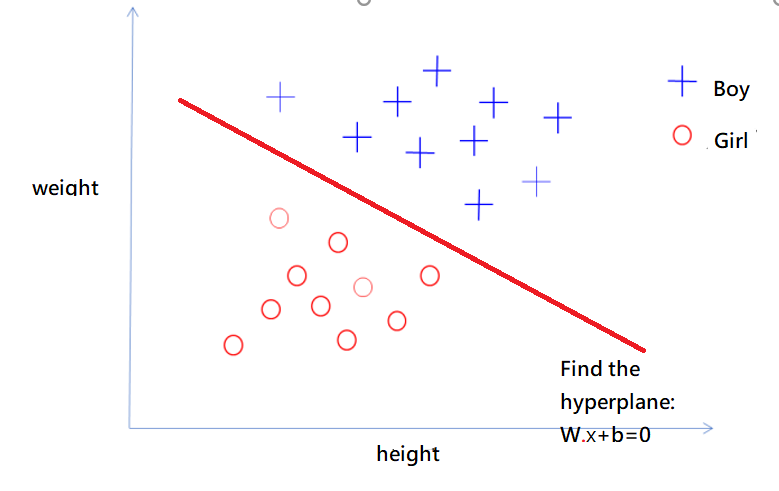
[Tommy Huang](https://medium.com/@chih.sheng.huang821?source=post_page-----c320098a3d2e----------------------)

Follow

[Mar 16, 2018](https://medium.com/@chih.sheng.huang821/%E6%A9%9F%E5%99%A8%E5%AD%B8%E7%BF%92-%E6%94%AF%E6%92%90%E5%90%91%E9%87%8F%E6%A9%9F-support-vector-machine-svm-%E8%A9%B3%E7%B4%B0%E6%8E%A8%E5%B0%8E-c320098a3d2e?source=post_page-----c320098a3d2e----------------------) · 7 min read



*Find a hyperplane that separate the two types of data points.*



The goal of SVM is to find the hyperplane to separate the two types of data points.

SVM: *wT****x***+*b=0*

To find the **w** vector and b constant so that the margin (gap) in between two types of data points is maximized.

**SVM Equations:**

Denote the dataset with the following notation

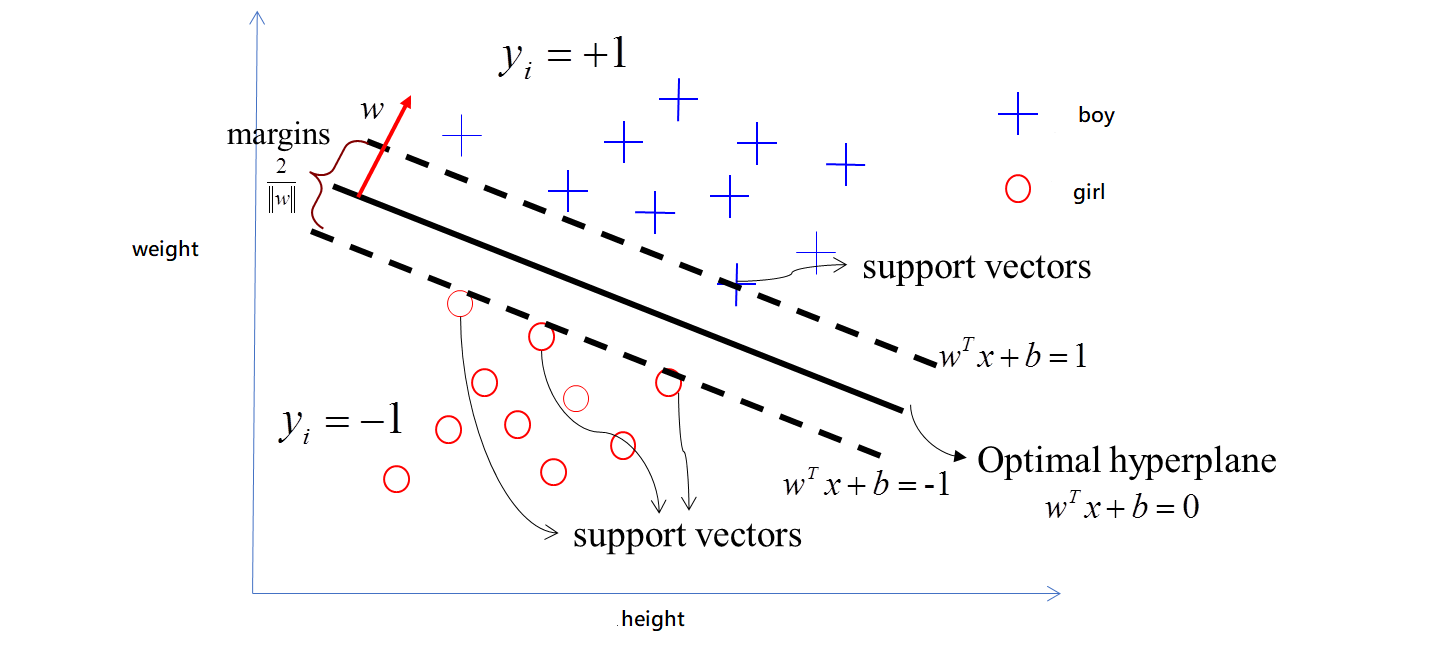
https://miro.medium.com/max/631/1*C6X11SmIAYoSS27fiu4rhA.png

In the above example:

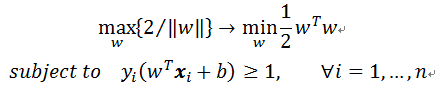
hard-margin SVM:

Combine into one constraint:

In an SVM, we find the optimal hyperplane such that the gap (**2/|w|**) in between the two types of data points is maximized

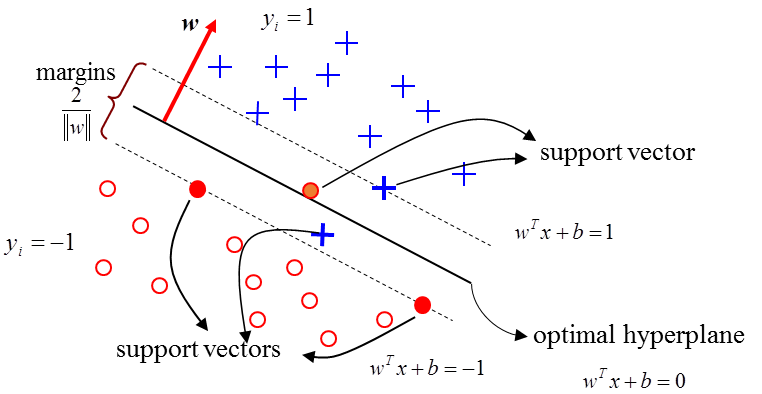


Hard-margin SVM:



In a real-life application, it is difficult to have all the data points fall outside of the supporting planes; therefore, we need the soft-margin SVM model.

Soft-margin SVM



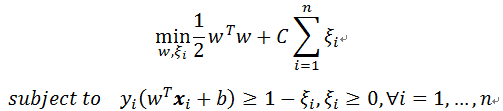
In a soft-margin SVM model，data points are allowed to be located in between the supporting plane with a penalty.

The data points in between the supporting plane do not satisfy the condition of

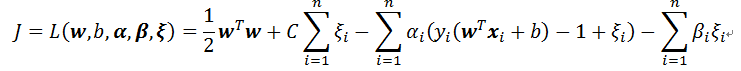
We use a slack variable , to allow the data point to slightly cross the boundary ( But we do not favor large )

*Note:*

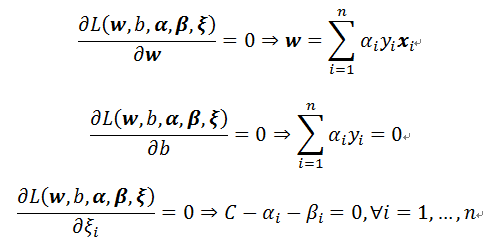
* *data points located in between the supporting planes are called the support vectors*
* *Penalty is given to the data points that cross the supporting planes. C is a positive number to adjust the weight on penalty.*



Applying Lagrangian method to solve the problem.



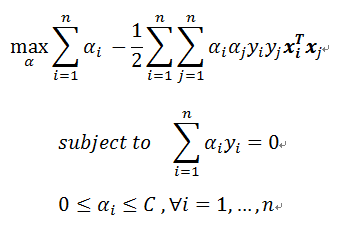
Based on Karush-Kuhn and Tucker (KKT) condition:



*αi*>=0,

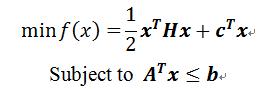
*βi*>=0

We have,



The problem has been reduced to a quadratic programming problem.

**Quadratic Programming**



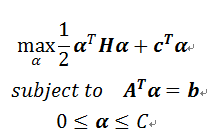
H is the Hessian matrix, which is symmetric matrix:

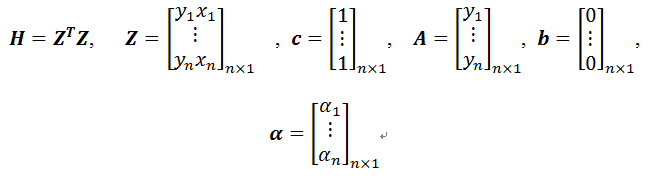
· if H is semi positive-definite, f(x) is a convex function

· if H is positive-definite, the minimum value is unique

· if H=0, the problem is reduced to a linear programming

A general quadratic optimization problem is as follows:





**Conclusion**

In finding an SVM, we need to solve a quadratic optimization problem. Usually, we use numerical method to find the solution. In solving an SVM, only the support vectors (whose ***αi>0*)** are needed.

**Notes:**

