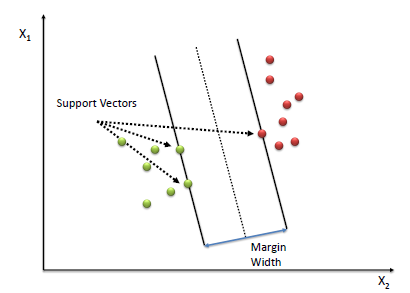
**show that changing the vectors other than the support vectors has no effect on the decision boundary**

In the second graph plot a linear discriminative classifier would attempt to draw a straight line separating the two sets of data, and thereby create a model for classification. For two dimensional data like that shown here, this is a task we could do by hand. But immediately we see a problem: there is more than one possible dividing line that can perfectly discriminate between the two classes.



### **Soft-margin**

To extend SVM to cases in which the data are not linearly separable, we introduce the *[hinge loss](https://en.wikipedia.org/wiki/Hinge_loss" \o "Hinge loss)* function,

{\displaystyle \max \left(0,1-y\_{i}({\vec {w}}\cdot {\vec {x}}\_{i}-b)\right).}IMG_256

Note that {\displaystyle y\_{i}}IMG_257 is the *i*-th target (i.e., in this case, 1 or −1), and {\displaystyle {\vec {w}}\cdot {\vec {x}}\_{i}-b}IMG_258 is the *i*-th output.

This function is zero if the constraint in (1) is satisfied, in other words, if {\displaystyle {\vec {x}}\_{i}}IMG_259 lies on the correct side of the margin. For data on the wrong side of the margin, the function's value is proportional to the distance from the margin.

We then wish to minimize

{\displaystyle \left[{\frac {1}{n}}\sum \_{i=1}^{n}\max \left(0,1-y\_{i}({\vec {w}}\cdot {\vec {x}}\_{i}-b)\right)\right]+\lambda \lVert {\vec {w}}\rVert ^{2},}IMG_260

where the parameter {\displaystyle \lambda }IMG_261 determines the trade-off between increasing the margin size and ensuring that the {\displaystyle {\vec {x}}\_{i}}IMG_262 lie on the correct side of the margin. Thus, for sufficiently small values of {\displaystyle \lambda }IMG_263, the second term in the loss function will become negligible, hence, it will behave similar to the hard-margin SVM, if the input data are linearly classifiable, but will still learn if a classification rule is viable or not.

Hence

**As in the SVM or first plot the value is random for the sepration of vectors. No matter what values we give the decission boundries goes through only supportive vectors. Thus that proves changing the vectors other than the support vectors has no effect on the decision boundary**.