The kernel trick is used to give the effect of projecting our data vectors into higher dimensions – without actually computing the vectors in those higher dimensional spaces. One such kernel is the RBF kernel.

What is the dimension of the vector space that the RBF kernel projects to? Justify mathematically. Given this, what is the theoretical maximum training accuracy one can expect for a binary classification problem using the RBF kernel with SVMs?

What role do the slack variables and 'C' play in soft-margin SVM?
How is the optimization problem for soft-margin SVM different
from that of hard-margin SVM?
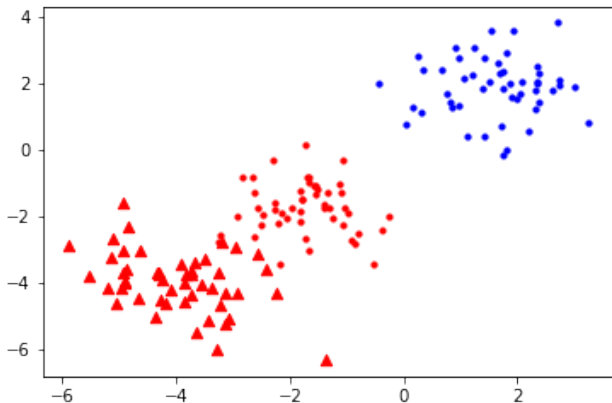
If we use hard-margin SVM, will the separating plane be different if
we remove the triangle points? Justify your answer mathematically.