1. For hard-SVM, there is perfectly separable data in the example. The hyperplane found by the hard-SVM will correctly classify all points without any misclassifications.
   For soft-SVM, choosing C is important. When the dataset is linearly separable, setting C to a high value will result in a similar solution to hard-SVM, pushing the solution to correctly classify all points. However, it isn't guaranteed that a value of C will always result in the same w* as the hard-SVM for every possible linearly separable dataset because the soft-SVM's optimization problem includes an additional term $C\sum \xi_i$ which might lead to a different margin if C is not large enough.
   While in many cases high values of C in the soft-SVM can approximate the hard-SVM solution closely, the statement that there exists a fixed C that makes sure the soft-SVM and hard-SVM solutions are always identical for every linearly separable dataset is false. In an example of n=2, d=1, with points (a,1) and (-a, -1), the hard-SVM solution would be a vector w such that the hyperplane w*x+b=0 separates the points with the maximal margin. For soft-SVM with a large C, we expect a similar separating hyperplane. However, for smaller values of C, the soft-SVM might choose a smaller margin if it reduces the objective function due to a less severe penalty on the slack variables. The soft-SVM solution might not have a maximal margin, resulting in a different w than the hard SVM solution.