# EECS545 Lecture 9 Quiz Solutions

1. Which of the following can only be used when training data are linearly separable? **Choose all options that apply**:
   (a) Linear hard-margin SVM.
   (b) Linear Logistic Regression.
   (c) Linear Soft-margin SVM.

   **Solution:** (a). Please revisit Slide 12.

2. Choose all true statements regarding soft-margin SVM:
   (a) When the regularization parameter C increases in soft-margin SVM, the number of support vectors will increase.
   (b) When the regularization parameter C increases in soft-margin SVM, the total sum of slack variables will decrease.
   (c) For a fixed linearly-separable dataset, the minimum objective value for soft-margin SVM is at least as small as the minimum objective value for hard-margin SVM.
   (d) When the regularization parameter C increases in soft-margin SVM, the margin of the resulting classifier will increase.

   **Solution:** (b), (c).
   (a), (b), (d): Increasing $C$ would penalize more for the slack variables, decreasing the total sum of slack variables at optimum. So, it will not increase the margin. The number of support vectors is not directly correlated.
   (c): Since the solution for the hard-margin SVM (0 slack variables) is still a feasible point for the soft-margin SVM problem, having more freedom on slack variable can only decrease the minimum for the soft-margin SVM.

3. (True/False) The farthest examples from the decision boundary of a dataset are called "support vectors". (Assume the dataset has extremely large number of examples.)

   **Solution:** False. Examples farthest from the decision boundary will not be the support vector as it will not contribute in determining the margin. Please check the plot in Slide 10.

4. (True/False) After training a linear SVM classifier, you observe that the test error is high while the training error is low. Increasing the value of parameter $C$ is likely to help improve test performance.

**Solution:** False. Decreasing $C$ will improve margin and that might help.

5. (True/False) Consider a supervised learning problem in which the training examples are points in 2-dimensional space. The positive examples are $(1, 1)$ and $(-1, -1)$. The negative examples are $(1, -1)$ and $(-1, 1)$. Note that the positive examples are NOT linearly separable from the negative examples in the original space. Feature transformation $\phi(\mathbf{x}) = [1, x_1, x_2, x_1x_2]$, where $x_1$ and $x_2$ are the first and the second coordinates of a generic example $\mathbf{x}$, could help the SVM to separate the positive and negative examples.

**Solution:** True. $\mathbf{w} = [0, 0, 0, 1]$ satisfies the maximum-margin decision surface separating the positive and negative examples on the feature space.