1. Describe two advantages that the softmax cross-entropy loss has over the SVM loss.

1. The softmax cross-entropy loss provides more information on the similarities between the label and the training image since the score of every object, no matter how small it is, will be accounted during the loss computation and back propagation. However, SVM loss does not take the difference into account if the y_label - y_j > 1, which makes softmax cross-entropy loss better with finer granularity.
2. The softmax cross-entropy loss adds a log to the score and is continuous, whereas the svm loss is not continuous nor convex, which made softmax cross-entropy loss a better option.

2. Why is the bias term needed in linear classifiers?

The bias term is needed since we would otherwise always pass through the origin in the projection plane, which is not good for a wide variety of cases. The bias term solve this issue and made the curve of our projection function closer to what we want and possible reduce the loss.