Homework 6
Patricia Hoffman, PhD.

1) In a past homework, you performed ridge regression on the wine quality data set. Now use a support vector machine to classify these data.

1a) First classify the data treating the last column as an ordered factor (the wine tasters score). Next treat the last column as a numeric. Which SVM implementation is better? Why do you think it is better?

1b) Using the best version choose two attributes and a slice through the data to plot. Choose a different set of attributes and another set of slices to plot.

1c) Compare and contrast the best version of the SVM with the ridge regression model.

2) Classify the sonar data set.

2a) Use a support vector machine to classify the sonar data set. First tune an SVM employing radial basis function (default). Next tune an SVM employing a linear kernel. Compare the results.

2b) In past homework, trees were used to classify the sonar data. Compare the best result using trees with the best result using SVM.

3) The in class example (svm1.r) used the glass data set. Use the Random Forest technique on the glass data. Compare the Random Forest results with the results obtained in class with SVM.

4) Choose a new data set which we haven't used in class yet (suggestion: choose one from http://archive.ics.uci.edu/ml/.) Use SVM to classify the data set. Try different kernels. Does changing the kernel make a difference? Which kernel resulted in the smallest error? Use another technique to classify the data set. Which resulted in the better model? (Make sure you describe the data set)

5) Use SVM with kernel = "linear" to create regression predictions on the data set created using these lines of code:

```
x <- seq(0.1, 5, by = 0.05)          # the observed feature
y <- log(x) + rnorm(x, sd = 0.2)   # the target for the observed feature
```

Next try various kernels and added features with SVM.  Can you improve the model by adding an extra feature which might be a function of the first feature? Compare both lm.ridge and svm.  Which method produced a better model? (don't forget to tune your models)