Assignment 7

due Thursday Oct. 30, 2008 at the beginning of the class.

The following questions from the book refer to the dataset cheese which can be found on the book cd in the appendix folder of the pcdatasets. For your convenience a .csv version of the file is attached with this assignment. I would advise you to use R for this assignment.

1. Answer exercises 11.43, 11.44 on page 715.

2. Answer exercises 11.45, 11.46, 11.47 and 11.48.

3. Perform a multicollinearity test. Does is appear that the variables are linearly related?

4. Carry out a multiple regression analysis. Perform a selection stepwise to determine the best set of variables to be included in the model.

5. Perform residual diagnostics. Based on partial regression plots and general additive model plots does it appear that higher order terms are necessary in the model? Are the residuals independent? Are they normal? Do they seem to have the same variance? Are there influential outliers in the model? Based on all these previous points find the best possible model and use it to predict the taste for cheese that tests as having $(6, 2.5, 1.5)$ for the three variables: "vinegar-ness", H_2O, and "milky-ness" content respectively.

BONUS Test for interaction effects present in the model (i.e. if the product terms are significant).