Week 8 homework problems

1. Use the data `lathe1` from problem **5.12** and do the following:

   a. As described in **5.12.2**, fit the full second-order model using log(`Life`) as the response and `Speed` and `Feed` as the predictors. A full second-order model is one which includes both predictors, second-order polynomial terms in both predictors, and the interaction term between the predictors. Report your fitted model.

   b. Obtain the effect plot for the interaction between `Speed` and `Feed` in the full second-order model. You should see a series of quadratic curves which, owing to the insignificance of the interaction term, essentially match one another except separated by vertical shifts.

   c. Test the interaction term in your model from parts a and b. You can either use the coefficient's t test or perform a partial F test using `anova()`. Does your result agree with what the effect plot showed?

   d. Remove the interaction term and refit the model. Test the quadratic terms in `Speed` and `Feed` to determine if the first-order model is adequate.

2. Use the data frame `lidar` in the `SemiPar` library, which you will likely have to install. After doing so, open the library. Next, open the data set by doing:

   ```
   data(lidar)
   ```

   This data frame contains the response `logratio` and predictor `range` obtained from 221 observations of a light detection and ranging experiment. Do the following:

   a. Fit splines to the data with three, four, five, and six degrees of freedom. Plot the data with the fitted splines overlaid on top, in a single plot.

   b. Use a spline with three degrees of freedom to predict the logratio for an observation with range equal to 900. Do the same for a spline with four degrees of freedom. Which prediction do you trust more, and why?

3. Use the data frame `BigMac2003` in `alr4`, which contains the response `BigMac`, or the price of a Big Mac in various world cities in 2003, expressed in minutes of labor. It also contains quantitative predictors `Bread`, `Rice`, `Bus`, and `Apt`, which are prices of other goods, as well as economic indicators `FoodIndex`, `TeachGI`, `TeachNI`, `TaxRate`, and `TeachHours`.

   a. Use principal components analysis on the nine predictors. Report your scree plot and a biplot.

   b. How many principal components are necessary to use in order to account for at least 90% of the variance in the predictors?

   c. Fit the MLR model with the response `BigMac` and the first four principal components as regressors. Then fit the model that contains all nine original predictors. Compare the coefficients of determination for the two models. Are you satisfied that the model with fewer regressors fits sufficiently well compared to the full model?