STAT 8230 — Applied Nonlinear Regression
Homework 2 – Due Tuesday, Sept. 20

Homework Guidelines:

- Homework is due by 4:30 on the due date specified above. You may turn it in at the beginning of class or place it in my mailbox in the Statistics Building. **No late homeworks will be accepted without permission granted prior to the due date.**
- Use only standard (8.5×11 inch) paper and use only one side of each sheet.
- Homework should show enough detail so that the reader can clearly understand the procedures of the solutions.
- Problems should appear in the order that they were assigned.

Assignment:

1. On the course web site is a data set (cherry.dat) containing measurements of the following variables on 31 black cherry trees:

$$V = \text{Volume of usable wood (cubic feet)}$$
$$H = \text{Height of tree (feet)}$$
$$D = \text{Diameter at breast height (inches)}$$

From these data we'd like to develop a model to predict wood volume from diameter and height. If we believe that the tree bole (trunk) is shaped roughly like a cone, then the formula for the volume of a cone, $V = \pi d^2 h/12$, can be used to motivate the following linear regression model:

$$y_i = \beta_1 + \beta_2 x_{i2} + \beta_3 x_{i3} + e_i, \quad i = 1, \ldots, 31, \qquad (*)$$

where $e_1, \ldots, e_{31} \overset{iid}{\sim} N(0, \sigma^2)$, y_i is the log volume of the ith tree, and x_{i2} and x_{i3} are $\log(D)$ and $\log(H)$ for the ith tree.

   a. Fit model (*) via ordinary least squares and summarize the fitted model.

   b. Alternatively, we can try to fit the volume data on their original scale (without linearizing by taking logs) with a nonlinear model:

$$V_i = \theta_1 D_i^{\theta_2} H_i^{\theta_3} + e_i, \quad \mathbf{e} \sim N(\mathbf{0}, \sigma^2 \mathbf{I}) \qquad (**)$$

   Fit model (**) using nonlinear least squares using starting values obtained from the fitted model in part (a). Summarize the fitted model.

c. Compute the studentized residuals from models (*) and (**) and plot them versus the fitted values. For which model do you think the assumption of homoscedastic errors is more appropriate?

2. Do problem 2.3 from chapter 2 of Bates and Watts' book (distributed in class).

3. Do problem 2.5 from chapter 2 of Bates and Watts' book (distributed in class).

4. Consider the nonlinear regression model

$$y_i = 1 - \exp(-\theta x_i) + \varepsilon_i, \quad i = 1, 2,$$

where $x_1 = 1, x_2 = 5$ and suppose $y_1 = y_2 = .5$ are observed. Calculate the least-squares estimate $\hat{\theta}$ of θ by performing two iterations of the Gauss-Newton algorithm by hand. Use a starting value of $\hat{\theta}^0 = .25$ and show each iteration graphically. To plot the iterations you may want to use a computer (e.g., using S-PLUS) to plot the expectation surface; the points $\boldsymbol{\eta}(\hat{\theta}^0)$, $\boldsymbol{\eta}(\hat{\theta}^1)$, and $\boldsymbol{\eta}(\hat{\theta}^2)$ on this expectation surface; and the point $\mathbf{y} = (y_1, y_2)^T$. It may be easiest to draw in other details by hand.