

# LINEAR REGRESSION MODELS W4315

## HOMEWORK 6 QUESTIONS

March 4, 2010

Due: 03/09/2010

Instructor: Frank Wood

**1. (15 points)** Refer to the design matrix given in “hw6p1.dat” on the course website. Read it into MATLAB ( the first **1** column is already added). Use “load hw6p1.dat” to load the file. What is the most complex model in terms of number of parameters that one could fit to this data?

Extra credit: If you fit a model with this number of parameters, how could you figure out which should be non-zero?

**2. (45 points)** Suppose  $X_1, \dots, X_n$  are i.i.d. samples from  $N(0, \sigma^2)$ . Denote  $\bar{X}$  as the sample mean. Prove  $S = \sum_{i=1}^n (X_i - \bar{X})^2 \sim \sigma^2 \chi^2(n-1)$  following the steps below using Cochran’s theorem:

a. Remember that we have the decomposition

$$\sum_{i=1}^n X_i^2 = \sum_{i=1}^n (X_i - \bar{X})^2 + n\bar{X}^2 \tag{1}$$

Show the matrices corresponding to all the three quadratic terms in (3).

- b. Derive the rank of each matrix above. (Hint: Recall problem 3 in homework 5.)
- c. Use Cochran’s theorem to prove  $S \sim \sigma^2 \chi^2(n-1)$ .