Assignment 3

(1) Under our usual assumptions for the linear model with multivariate normal errors, for testing the hypothesis $H_0 : D\beta = 0$ with $D \in \mathbb{R}^{q\times p}, r(D) = q \leq p$, let

$$F = \frac{\|\widehat{Y} - \widehat{Y}_{M_0}\|^2/q}{\|Y - \widehat{Y}\|^2/(n-p)}$$

where $\widehat{Y}_{M_0} = P_{M_0}Y$, with P_{M_0} is the orthogonal projection onto the subspace M_0 equals $\mathcal{R}(X)\cap\mathcal{N}(D)$. Show that when H_0 is true then $F \sim F_{q,n-p}$.

(2) Show that the F statistic in the previous problem may be written as

$$F = \frac{(D\widehat{\beta} - c)^T[D(X^TX)^{-1}D^T]^{-1}(D\widehat{\beta} - c)}{qS^2},$$

with $c = 0$.

(3) Compute, the F statistic for testing the hypothesis that $H_0 : \beta_2 = 0$ in the simple linear regression setting: $Y = \beta_1 + \beta_2 x_2 + \epsilon$. Interpret the resulting formula, and give intuition as to how the data needs to behave in order that this F be large.

(4) Same as the problem above, for the test of $\beta_1 = \beta_2$ for the model $Y = \beta_1 x_1 + \beta_2 x_2 + \epsilon$.

(5) Reduce the problem of least squares estimation under the constraint $D\beta = c$ to one under the constraint $D\beta = 0$, and use that reduction to give the least squares estimate of the β parameter in general.

(6) Use the results of the previous problem to show (2).

(7) Show how the independent two sample t test with possibly unequal sample sizes can be derived as a special instance of the F test. Can one generalize this result to the natural k sample extension? (No need to work any computations to their endpoints. If an extension exists, show how one effects the generalization, and if not explain why it cannot be done.)

(8) Having made n_1, n_2 and n_3 measurements on the three angles of a triangle, each estimate unbiased, and having errors of equal variance, find the least squares estimates of the three angles subject to the constraint that they sum to π.