STAT 802 Multivariate Analysis Oct. 2, 2003

Today: More About Multiple Regression (the multivariate nature of the predictor space).
Multivariate Regression

One thing I skipped over last time:
Leverage and Influence(pp 377-380)

First, we need a bit of theory about residuals: residual vector is n x 1

$\hat{\varepsilon} = y - \hat{y} = (I-H)y$ where $H = Z(Z'Z)^{-1}Z'$ is n x n

$E(\hat{\varepsilon}) = 0$ but $Cov(\hat{\varepsilon}) = \sigma^2$ (□□□) which is not diagonal. (H is called the "hat" matrix.) However, we expect the off-diagonal elements to be small. In the case of the house price data, the 380 (=400-20) off-diagonal elements of I-H have a root mean square of .075 wheras the diagonal elements range from .06 to .23 but the one large one (row 16) is .74. Note that these diagonal elements are all less than 1. A $(1-h_{ii})$ close to 0 (an h_{ii} close to 1) means that the ith row is so far from the rest of the data that the model is forced to come close to its y, so the residual is small. This does not mean that this row is good data – the "residual" here is really an estimated residual, and when the leverage h_{ii} is large, we need to worry about the model being too much influenced by this one row.

When we are evaluating the residuals, we should be allowing for the fact that Z values distant from the other Z values will have residuals that are more poorly estimated (larger variance) than Z values close to others. (Think of simple linear regression to see this – Since the regression line slope might be wrong, the $\hat{y}$ may be far from E(y), and the "true" residual is y-E(y) – these residuals may be expected to have the same variance - but unfortunately we don't know E(y).) Since the estimate of $Var(\hat{\varepsilon}_j) = s^2(1-h_{jj})$, a way of comparing the size of residuals that allows for their different variance is to compute the so-called "studentized residual" which is $\hat{\varepsilon}_j^* = \hat{\varepsilon}_j / (s^2(1-h_{jj}))^{1/2} = \hat{\varepsilon}_j / s(1-h_{jj})^{1/2}$. This $\hat{\varepsilon}_j^*$ is evaluated relative to its approximate N(0,1) distribution.

High leverage means that the row might have a big influence on the estimated model. The actual influence can be assessed by deleting the row and observing the change in the parameter estimate. Usually a row with high leverage (a property of the predictor values only) will result in that row having a big influence on the beta coefficients. It would be a coincidence, but possible, that a high leverage row results in a low influence row.

The usual approach to regression is to use equally weighted rows. This is appropriate only when we have equal faith in each row, or when we want to see the consequences of this assumption. But if the model is unduly dependent on only a few rows, then equal weighting actually results in unequal weighting with respect to our model. Leverage and influence measures help to draw attention to this problem.

Exercise: Modify the house data (row 16) to show that it is possible to have high leverage high influence but also high leverage low influence – the latter is more rare.

Multivariate Regression Model (p 384)

m response variables, r predictors, n cases.

Responses from the n cases are uncorrelated but the m responses within each case may be correlated.

Least Squares estimators look the same is for univariate response (7-31). Z is the design matrix (n x r) – just as before – but now Y is a matrix (n x m) not a vector (n x 1)

Simple example 7.8 p 385 ff

m = 2, r=1, n=5

What is the difference between this analysis and 2 multiple regression analyses?

The regression coefficients of the two multiple regressions are correlated (Result 7.9 p 388)
So inference on the joint distribution of the Beta vectors should use this fact.
I
Example 7.10 gives an example of this. See the result graphically p 398. Note that an outlier of (y_1,y_2) would not necessarily appear as an outlier of one of the component multiple regressions.

Partial Correlation Coefficient (p 406)

Section 7.9: Z fixed? Or Z p-normal? Same linear predictors (p 409)