The Matrix-Based Approach to the General Linear Model - continued more

Back to the Least-Squares Approach: Statistical Inference for Several Coefficients

If we have k covariates in our model, suppose we wish to test if q of them are not important. Suppose they are the first q coefficients, so we wish to test $H_0 : \beta_1 = \beta_2 = \cdots = \beta_q = 0$. Recall that we can express the null hypothesis in terms of the constraint $\boldsymbol{L}\boldsymbol{\beta} = \mathbf{c}$, and that the estimator under this constraint is

$$\mathbf{b}_0 = \mathbf{b} + (\mathbf{X}'\mathbf{X})^{-1}\boldsymbol{L}'[\boldsymbol{L}(\mathbf{X}'\mathbf{X})^{-1}\boldsymbol{L}']^{-1}(\mathbf{c} - \boldsymbol{L}\mathbf{b}).$$

The estimator $\mathbf{b}_0$ is both the least-squares and maximum-likelihood estimator of $\boldsymbol{\beta}$ under the constraint $\boldsymbol{L}\boldsymbol{\beta} = \mathbf{c}$. To test $H_0 : \beta_1 = \beta_2 = \cdots = \beta_q = 0$ with this approach, we can calculate the reduction in regression sums of squares under H_0, which is equivalent to calculating the increase in error sums of squares. Again using the notation that model 0 is constrained under H_0, and RSS and $\widehat{\mathbf{y}}_1$ are from the unconstrained model, we have:

$$RSS_0 - RSS = (\mathbf{y} - \widehat{\mathbf{y}}_0)'(\mathbf{y} - \widehat{\mathbf{y}}_0) - (\mathbf{y} - \widehat{\mathbf{y}}_1)'(\mathbf{y} - \widehat{\mathbf{y}}_1).$$

We can identify $\widehat{\mathbf{y}}_0$ as $\widehat{\mathbf{y}}_0 = \mathbf{X}\mathbf{b}_0$ using the constrained least-squares solution above (with $\mathbf{c} = \mathbf{0}$ under H_0), to get

$$\begin{aligned}\widehat{\mathbf{y}}_0 &= \mathbf{X}\mathbf{b}_0 = \mathbf{X}\mathbf{b} - \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\boldsymbol{L}'[\boldsymbol{L}(\mathbf{X}'\mathbf{X})^{-1}\boldsymbol{L}']^{-1}\boldsymbol{L}\mathbf{b} \\ &= \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y} - \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\boldsymbol{L}'[\boldsymbol{L}(\mathbf{X}'\mathbf{X})^{-1}\boldsymbol{L}']^{-1}\boldsymbol{L}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y} \\ &= (\mathbf{P} - \mathbf{P_C})\mathbf{y} = \mathbf{P_0}\mathbf{y},\end{aligned}$$

where $\mathbf{P} = \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'$ and $\mathbf{P_C} = \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\boldsymbol{L}'[\boldsymbol{L}(\mathbf{X}'\mathbf{X})^{-1}\boldsymbol{L}']^{-1}\boldsymbol{L}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'$. It can be verified that $\mathbf{P}, \mathbf{P_C}$, and $\mathbf{P_0}$ are all idempotent, and further that $\mathbf{P}\mathbf{P}_C = \mathbf{P}_C\mathbf{P} = \mathbf{P}_C$. Thus we have

$$(\mathbf{y} - \widehat{\mathbf{y}}_0) = \mathbf{y} - \mathbf{P_0}\mathbf{y} = (\mathbf{I}_n - \mathbf{P_0})\mathbf{y},$$

and recalling that $(\mathbf{y} - \widehat{\mathbf{y}}_1) = \mathbf{y} - \mathbf{P}\mathbf{y} = (\mathbf{I}_n - \mathbf{P})\mathbf{y}$, so

$$
\begin{aligned}
RSS_0 - RSS &= (\mathbf{y}-\widehat{\mathbf{y}}_0)'(\mathbf{y}-\widehat{\mathbf{y}}_0) - (\mathbf{y}-\widehat{\mathbf{y}}_1)'(\mathbf{y}-\widehat{\mathbf{y}}_1) \\
&= [(\mathbf{I}_n-\mathbf{P_0})\mathbf{y}]'[(\mathbf{I}_n-\mathbf{P_0})\mathbf{y}] - [(\mathbf{I}_n-\mathbf{P})\mathbf{y}]'[(\mathbf{I}_n-\mathbf{P})\mathbf{y}] \\
&= \mathbf{y}'(\mathbf{I}_n-\mathbf{P_0})'(\mathbf{I}_n-\mathbf{P_0})\mathbf{y} - \mathbf{y}'(\mathbf{I}_n-\mathbf{P})'(\mathbf{I}_n-\mathbf{P})\mathbf{y} \\
&= \mathbf{y}'(\mathbf{I}_n-\mathbf{P_0})\mathbf{y} - \mathbf{y}'(\mathbf{I}_n-\mathbf{P})\mathbf{y} \\
&= \mathbf{y}'\mathbf{y} - \mathbf{y}'\mathbf{P_0}\mathbf{y} - (\mathbf{y}'\mathbf{y} - \mathbf{y}'\mathbf{P}\mathbf{y}) \\
&= \mathbf{y}'\mathbf{P}\mathbf{y}-\mathbf{y}'\mathbf{P_0}\mathbf{y} \\
&= \mathbf{y}'(\mathbf{P} - \mathbf{P_0})\mathbf{y}
\end{aligned}
$$

Also, since $RSS = \mathbf{y}'(\mathbf{I}_n - \mathbf{P})\mathbf{y}$ and $(\mathbf{I}_n - \mathbf{P})(\mathbf{P} - \mathbf{P_0}) = \mathbf{P} - \mathbf{P_0} - \mathbf{P}^2 + \mathbf{P}\mathbf{P_0} = \mathbf{P} - \mathbf{P_0} - \mathbf{P} + \mathbf{P_0} = \mathbf{0}$, we know from Result 5 that the reduction in error sum of squares $RSS_0 - RSS$ is independent of RSS, so the ratio of these two quadratic forms divided by their degrees of freedom leads to:

$$F_0 = \frac{(RSS_0 - RSS)/q}{RSS/(n-k-1)},$$

which is distributed as an F statistic with q and $n-k-1$ degrees of freedom.