Testing linear restrictions

In relation to the linear regression model, $y = X\beta + u$, one is often interested in testing hypotheses which take the form of linear restrictions on the parameters. Simple examples include zero restrictions, single or multiple (e.g. $\beta_1 = 0$ or $\beta_3 = \beta_4 = 0$), equality restrictions among the parameters (e.g. $\beta_2 = \beta_3$), and adding-up restrictions (e.g. $\beta_1 + \beta_2 = 1$).

The most general way of expressing a set of such restrictions is

$$R\beta = q \quad \text{or} \quad R\beta - q = 0$$

In this formulation R is a $g \times k$ matrix, where g is the number of restrictions and k is the number of elements in β, and q is a g-vector. Consider a few examples. Suppose we want to test $H_0 : \beta_3 = \beta_4 = 0$, in a model for which $k = 4$. In that case we can use

$$R = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad \text{and} \quad q = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

which gives

$$R\beta = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \beta_1 \\ \beta_2 \\ \beta_3 \\ \beta_4 \end{bmatrix} = \begin{bmatrix} 1 \cdot \beta_3 \\ 1 \cdot \beta_4 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} = q$$

How about $H_0 : \beta_1 + \beta_2 = 1$ (in the same model)? For this case we need

$$R = \begin{bmatrix} 1 & 1 & 0 & 0 \end{bmatrix} \quad \text{and} \quad q = \begin{bmatrix} 1 \end{bmatrix}$$

so that

$$R\beta = \begin{bmatrix} \beta_1 + \beta_2 \end{bmatrix} = \begin{bmatrix} 1 \end{bmatrix}$$

If the null hypothesis, $R\beta = q$, is true, then we would expect that $R\hat{\beta} - q$ should be "small" (where $\hat{\beta}$ is the OLS estimate of β).[1] Our test statistic will be an appropriate measure of the "distance" between $R\hat{\beta}$ and q, such that if this distance is "big enough" we reject the null. How can we construct such a measure?

Well, consider for a moment the simple case of a test on a single regression parameter, e.g., $H_0 : \beta_i = \beta^0$. We know how to construct a test statistic in that case:

$$t = \frac{\hat{\beta}_i - \beta^0}{s.e.(\hat{\beta}_i)}$$

That is, we take the distance between $\hat{\beta}_i$ and the value specified by the null (β^0) and divide it by the appropriate standard error, giving the distance measured in standard errors. And (roughly speaking, for the 5 percent significance level) we reject the null if the absolute value of t is greater than 2, since this means that the probability of observing such a value is less than 0.05 under the null.

Now, as an intermediate step, consider what we get if we square a t statistic:

$$F = t^2 = \frac{\left(\hat{\beta}_i - \beta^0\right)^2}{\text{var}(\hat{\beta}_i)}$$

[1] Due to sampling variability, we do not expect to find that $R\hat{\beta}$ exactly equals q, even if the null is true.

This figure is still meaningful, though it obviously does not follow the t distribution; in fact it follows the F distribution, with one degree of freedom in the numerator and $n - k$ degrees of freedom in the denominator (where n is the number of observations). So we can, if we like, square a t statistic and look up its p-value using a tabulation of the F distribution (though that's a bit roundabout).

Alright, now back to the task in hand. In the general case, we want a comparably standardized measure of the "size" of $R\hat{\beta} - q$, a vector quantity. Following the pattern of the F statistic above, we want something like "the square of the distance divided by its variance". Since all the relevant quantities are matrices we can't get exactly that, but hopefully by now we have a sense of what to do: "squaring" a vector translates to multiplication by its own transpose, and "dividing" by a matrix translates to multiplication by the inverse of that matrix.

OK, but what variance do we need, exactly? The quantity we're looking at is, once again, $R\hat{\beta} - q$. We know how to find the variance of OLS $\hat{\beta}$: where do we go from there? Well, there's a basic property of variance that you should know by now. Let x be a random variable and a and b be constants. What's the variance of $a + bx$? The answer is $b^2\text{var}(x)$. Here, we need the matrix counterpart of that:

$$\text{var}(R\hat{\beta} - q) = R\text{var}(\hat{\beta})R'$$

The constant q term disappears, like the a above, and $\text{var}(\hat{\beta})$ gets sandwiched between R and R', in the matrix counterpart of $b^2\text{var}(x)$.

So, finally, here's our test statistic for the hypothesis $R\beta = q$:

$$\left(R\hat{\beta} - q\right)' \left[R\text{var}(\hat{\beta})R'\right]^{-1} \left(R\hat{\beta} - q\right)$$

or

$$\text{(deviation)}' \ \text{(variance of deviation)}^{-1} \ \text{(deviation)}$$

where "deviation" means the deviation of what we found in the data, via OLS, from what the null hypothesis states.

Well, in fact we're not quite finished. The statistic given above is OK, but it follows the χ^2 (chi-square) distribution. If we want an F statistic we can get one easily: just divide the χ^2 value by g (or in other words, average it over the number of restrictions being tested).

In gretl it is straightforward to test general linear restrictions, without having to formulate everything in matrix terms explicitly (that's done for you in the background). After estimating a model one uses a `restrict` block. Here are a few simple examples:

Test the hypothesis that $\beta_2 = \beta_3$:

```
restrict
  b2 - b3 = 0
end restrict
```

Test $H_0 : \beta_2 + \beta_3 = 1$

```
restrict
  b2 + b3 = 1
end restrict
```

Test the joint null that $\beta_3 = \beta_4 = 0$

```
restrict
  b3 = 0
  b4 = 0
end restrict
```