COCHRANE'S THEOREM: THE DECOMPOSITION OF A CHI-SQUARE

The standard test of an hypothesis regarding the vector β in the model $N(y; X\beta, \sigma^2 I)$ entails a multi-dimensional version of Pythagoras' Theorem. Consider the decomposition of the vector y into the systematic component and the residual vector. This gives

$$y = X\hat{\beta} + (y - X\hat{\beta}) \quad \text{and}$$

(1)

$$y - X\beta = (X\hat{\beta} - X\beta) + (y - X\hat{\beta}),$$

where the second equation comes from subtracting the unknown mean vector $X\beta$ from both sides of the first. These equations can also be expressed in terms of the projector $P = X(X'X)^{-1}X'$ which gives $Py = X\hat{\beta}$ and $(I - P)y = y - X\hat{\beta} = e$. Using the definition $\varepsilon = y - X\beta$ within the second of the equations, we have

$$y = Py + (I - P)y \qquad \text{and}$$

(2)

$$\varepsilon = P\varepsilon + (I - P)\varepsilon.$$

The reason for rendering the equations in this notation is that it enables us to envisage more clearly the Pythagorean relationship between the vectors. Thus, from the condition that $P = P' = P^2$, which is equivalent to the condition that $P'(I - P) = 0$, it can be established that

$$\varepsilon'\varepsilon = \varepsilon'P\varepsilon + \varepsilon'(I - P)\varepsilon \qquad \text{or}$$

(3)

$$\varepsilon'\varepsilon = (X\hat{\beta} - X\beta)'(X\hat{\beta} - X\beta) + (y - X\hat{\beta})'(y - X\hat{\beta}).$$

The terms in these expressions represent squared lengths; and the vectors themselves form the sides of a right-angled triangle with $P\varepsilon$ at the base, $(I - P)\varepsilon$ as the vertical side and ε as the hypotenuse.

The usual test of an hypothesis regarding the elements of the vector β is based on the foregoing relationships. Imagine that the hypothesis postulates that the true value of the parameter vector is β_0. To test this notion, we compare the value of $X\beta_0$ with the estimated mean vector $X\hat{\beta}$. The test is a matter of assessing the proximity of the two vectors which is measured by the square of the distance which separates them. This is given by $\varepsilon'P\varepsilon = (X\hat{\beta} - X\beta_0)'(X\hat{\beta} - X\beta_0)$. If the hypothesis is untrue and if $X\beta_0$ is remote from the true value of $X\beta$, then the distance is liable to be excessive. The distance can only be assessed in comparison with the variance σ^2 of the disturbance term or with an estimate thereof. Usually, one has to make do with the estimate of σ^2 which is provided by

$$\begin{aligned} \hat{\sigma}^2 &= \frac{(y - X\hat{\beta})'(y - X\hat{\beta})}{T - k} \\ &= \frac{\varepsilon'(I - P)\varepsilon}{T - k}. \end{aligned}$$

(4)

The numerator of this estimate is simply the squared length of the vector $e = (I - P)y = (I - P)\varepsilon$ which constitutes the vertical side of the right-angled triangle.

The test uses the result that

(5)   If $y \sim N(X\beta, \sigma^2 I)$ and if $\hat{\beta} = (X'X)^{-1}X'y$, then

$$F = \left\{ \frac{(X\hat{\beta} - X\beta)'(X\hat{\beta} - X\beta)}{k} \middle/ \frac{(y - X\hat{\beta})'(y - X\hat{\beta})}{T - k} \right\}$$

is distributed as an $F(k, T - k)$ statistic.

This result depends upon Cochrane's Theorem concerning the decomposition of a chi-square random variate. The following is a statement of the theorem which is attuned to our present requirements:

(6)   Let $\varepsilon \sim N(0, \sigma^2 I_T)$ be a random vector of T independently and identically distributed elements. Also let $P = X(X'X)^{-1}X'$ be a symmetric idempotent matrix, such that $P = P' = P^2$, which is constructed from a matrix X of order $T \times k$ with $\mathrm{Rank}(X) = k$. Then

$$\frac{\varepsilon'P\varepsilon}{\sigma^2} + \frac{\varepsilon'(I - P)\varepsilon}{\sigma^2} = \frac{\varepsilon'\varepsilon}{\sigma^2} \sim \chi^2(T),$$

which is a chi-square variate of T degrees of freedom, represents the sum of two independent chi-square variates $\varepsilon'P\varepsilon/\sigma^2 \sim \chi^2(k)$ and $\varepsilon'(I - P)\varepsilon/\sigma^2 \sim \chi^2(T - k)$ of k and $T - k$ degrees of freedom respectively.

To prove this result, we begin by finding an alternative expression for the projector $P = X(X'X)^{-1}X'$. First consider the fact that $X'X$ is a symmetric positive-definite matrix. It follows that there exists a matrix transformation T such that $T(X'X)T' = I$ and $T'T = (X'X)^{-1}$. Therefore $P = XT'TX' = C_1C_1'$, where $C_1 = XT'$ is a $T \times k$ matrix comprising k orthonormal vectors such that $C_1'C_1 = I_k$ is the identity matrix of order k.

Now define C_2 to be a complementary matrix of $T - k$ orthonormal vectors. Then $C = [C_1, C_2]$ is an orthonormal matrix of order T such that

$$CC' = C_1C_1' + C_2C_2' = I_T \qquad \text{and}$$

(7)
$$C'C = \begin{bmatrix} C_1'C_1 & C_1'C_2 \\ C_2'C_1 & C_2'C_2 \end{bmatrix} = \begin{bmatrix} I_k & 0 \\ 0 & I_{T-k} \end{bmatrix}.$$

The first of these results allows us to set $I - P = I - C_1C_1' = C_2C_2'$. Now, if $\varepsilon \sim N(0, \sigma^2 I_T)$ and if C is an orthonormal matrix such that $C'C = I_T$, then it follows that $C'\varepsilon \sim N(0, \sigma^2 I_T)$. In effect, if ε is a normally distributed random vector with a density function which is centred on zero and which has spherical contours, and if C is the matrix of a rotation, then nothing is altered by applying the rotation to the random vector. On partitioning $C'\varepsilon$, we find that

(8)
$$\begin{bmatrix} C_1'\varepsilon \\ C_2'\varepsilon \end{bmatrix} \sim N \left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \sigma^2 I_k & 0 \\ 0 & \sigma^2 I_{T-k} \end{bmatrix} \right),$$

2

which is to say that $C_1'\varepsilon \sim N(0, \sigma^2 I_k)$ and $C_2'\varepsilon \sim N(0, \sigma^2 I_{T-k})$ are independently distributed normal vectors. It follows that

(9)
$$\frac{\varepsilon' C_1 C_1' \varepsilon}{\sigma^2} = \frac{\varepsilon' P \varepsilon}{\sigma^2} \sim \chi^2(k) \quad \text{and}$$

$$\frac{\varepsilon' C_2 C_2' \varepsilon}{\sigma^2} = \frac{\varepsilon'(I - P)\varepsilon}{\sigma^2} \sim \chi^2(T - k)$$

are independent chi-square variates. Since $C_1 C_1' + C_2 C_2' = I_T$, the sum of these two variates is

(10)
$$\frac{\varepsilon' C_1 C_1' \varepsilon}{\sigma^2} + \frac{\varepsilon' C_2 C_2' \varepsilon}{\sigma^2} = \frac{\varepsilon' \varepsilon}{\sigma^2} \sim \chi^2(T);$$

and thus the theorem is proved.

The statistic under (5) can now be expressed in the form of

(11)
$$F = \left\{ \frac{\varepsilon' P \varepsilon}{k} \middle/ \frac{\varepsilon'(I - P)\varepsilon}{T - k} \right\}.$$

This is manifestly the ratio of two chi-square variates divided by their respective degrees of freedom; and so it has an F distribution with these degrees of freedom. This result provides the means for testing the hypothesis concerning the parameter vector β.