Under the null hypothesis, as we will now demonstrate, this test statistic follows the F distribution with r and $n-k$ degrees of freedom. Not surprisingly, it is called an **$\boldsymbol{F}$ statistic**.

The restricted SSR is $\boldsymbol{y}^\top\boldsymbol{M}_1\boldsymbol{y}$, and the unrestricted one is $\boldsymbol{y}^\top\boldsymbol{M}_{\boldsymbol{X}}\boldsymbol{y}$. One way to obtain a convenient expression for the difference between these two expressions is to use the FWL Theorem. By this theorem, the USSR is the SSR from the FWL regression

$$\boldsymbol{M}_1\boldsymbol{y} = \boldsymbol{M}_1\boldsymbol{X}_2\boldsymbol{\beta}_2 + \text{residuals}. \tag{4.31}$$

The total sum of squares from (4.31) is $\boldsymbol{y}^\top\boldsymbol{M}_1\boldsymbol{y}$. The explained sum of squares can be expressed in terms of the orthogonal projection on to the r–dimensional subspace $\mathcal{S}(\boldsymbol{M}_1\boldsymbol{X}_2)$, and so the difference is

$$\text{USSR} = \boldsymbol{y}^\top\boldsymbol{M}_1\boldsymbol{y} - \boldsymbol{y}^\top\boldsymbol{M}_1\boldsymbol{X}_2(\boldsymbol{X}_2^\top\boldsymbol{M}_1\boldsymbol{X}_2)^{-1}\boldsymbol{X}_2^\top\boldsymbol{M}_1\boldsymbol{y}. \tag{4.32}$$

Therefore,

$$\text{RSSR} - \text{USSR} = \boldsymbol{y}^\top\boldsymbol{M}_1\boldsymbol{X}_2(\boldsymbol{X}_2^\top\boldsymbol{M}_1\boldsymbol{X}_2)^{-1}\boldsymbol{X}_2^\top\boldsymbol{M}_1\boldsymbol{y},$$

and the F statistic (4.30) can be written as

$$F_{\boldsymbol{\beta}_2} = \frac{\boldsymbol{y}^\top\boldsymbol{M}_1\boldsymbol{X}_2(\boldsymbol{X}_2^\top\boldsymbol{M}_1\boldsymbol{X}_2)^{-1}\boldsymbol{X}_2^\top\boldsymbol{M}_1\boldsymbol{y}/r}{\boldsymbol{y}^\top\boldsymbol{M}_{\boldsymbol{X}}\boldsymbol{y}/(n-k)}. \tag{4.33}$$

In general, $\boldsymbol{M}_{\boldsymbol{X}}\boldsymbol{y} = \boldsymbol{M}_{\boldsymbol{X}}\boldsymbol{u}$. Under the null hypothesis, $\boldsymbol{M}_1\boldsymbol{y} = \boldsymbol{M}_1\boldsymbol{u}$. Thus, under this hypothesis, the F statistic (4.33) reduces to

$$\frac{\boldsymbol{\varepsilon}^\top\boldsymbol{M}_1\boldsymbol{X}_2(\boldsymbol{X}_2^\top\boldsymbol{M}_1\boldsymbol{X}_2)^{-1}\boldsymbol{X}_2^\top\boldsymbol{M}_1\boldsymbol{\varepsilon}/r}{\boldsymbol{\varepsilon}^\top\boldsymbol{M}_{\boldsymbol{X}}\boldsymbol{\varepsilon}/(n-k)}, \tag{4.34}$$

where, as before, $\boldsymbol{\varepsilon} \equiv \boldsymbol{u}/\sigma$. We saw in the last subsection that the quadratic form in the denominator of (4.34) is distributed as $\chi^2(n-k)$. Since the quadratic form in the numerator can be written as $\boldsymbol{\varepsilon}^\top\boldsymbol{P}_{\boldsymbol{M}_1\boldsymbol{X}_2}\boldsymbol{\varepsilon}$, it is distributed as $\chi^2(r)$. Moreover, the random variables in the numerator and denominator are independent, because $\boldsymbol{M}_{\boldsymbol{X}}$ and $\boldsymbol{P}_{\boldsymbol{M}_1\boldsymbol{X}_2}$ project on to mutually orthogonal subspaces: $\boldsymbol{M}_{\boldsymbol{X}}\boldsymbol{M}_1\boldsymbol{X}_2 = \boldsymbol{M}_{\boldsymbol{X}}(\boldsymbol{X}_2 - \boldsymbol{P}_1\boldsymbol{X}_2) = \mathbf{O}$. Thus it is apparent that the statistic (4.34) follows the $F(r, n-k)$ distribution under the null hypothesis.

A Threefold Orthogonal Decomposition

Each of the restricted and unrestricted models generates an orthogonal decomposition of the dependent variable $\boldsymbol{y}$. It is illuminating to see how these two decompositions interact to produce a threefold orthogonal decomposition. It turns out that all three components of this decomposition have useful interpretations. From the two models, we find that

$$\boldsymbol{y} = \boldsymbol{P}_1\boldsymbol{y} + \boldsymbol{M}_1\boldsymbol{y} \quad \text{and} \quad \boldsymbol{y} = \boldsymbol{P}_{\boldsymbol{X}}\boldsymbol{y} + \boldsymbol{M}_{\boldsymbol{X}}\boldsymbol{y}. \tag{4.35}$$