obtained by differentiating (3.42) with respect to $\boldsymbol{\beta}$ and $\boldsymbol{\lambda}$ and setting the derivatives to zero are

$$-\boldsymbol{X}^\top(\tilde{\boldsymbol{\beta}})\big(\boldsymbol{y}-\boldsymbol{x}(\tilde{\boldsymbol{\beta}})\big)+\boldsymbol{R}^\top\tilde{\boldsymbol{\lambda}}=\boldsymbol{0} \tag{3.43}$$

$$\boldsymbol{R}\tilde{\boldsymbol{\beta}}-\boldsymbol{r}=\boldsymbol{0}, \tag{3.44}$$

where $\tilde{\boldsymbol{\beta}}$ denotes the restricted estimates and $\tilde{\boldsymbol{\lambda}}$ denotes the estimated Lagrange multipliers. From (3.43), we see that

$$\boldsymbol{R}^\top\tilde{\boldsymbol{\lambda}}=\tilde{\boldsymbol{X}}^\top\big(\boldsymbol{y}-\tilde{\boldsymbol{x}}\big), \tag{3.45}$$

where, as usual, $\tilde{\boldsymbol{x}}$ and $\tilde{\boldsymbol{X}}$ denote $\boldsymbol{x}(\tilde{\boldsymbol{\beta}})$ and $\boldsymbol{X}(\tilde{\boldsymbol{\beta}})$. The expression on the right-hand side of (3.45) is minus the k–vector of the derivatives of $\frac{1}{2}SSR(\boldsymbol{\beta})$ with respect to all the elements of $\boldsymbol{\beta}$, evaluated at $\tilde{\boldsymbol{\beta}}$. This vector is often called the **score vector**. Since $\boldsymbol{y}-\tilde{\boldsymbol{x}}$ is simply a vector of residuals, which should converge asymptotically under H_0 to the vector of error terms $\boldsymbol{u}$, it seems plausible that the asymptotic covariance matrix of the vector of scores is

$$\sigma_0^2\boldsymbol{X}^\top(\boldsymbol{\beta}_0)\boldsymbol{X}(\boldsymbol{\beta}_0). \tag{3.46}$$

Subject to certain asymptotic niceties, that is indeed the case, and a more rigorous version of this result will be proved in Chapter 5.

The obvious way to estimate (3.46) is to use $\tilde{s}^2\tilde{\boldsymbol{X}}^\top\tilde{\boldsymbol{X}}$, where $\tilde{s}^2$ is $SSR(\tilde{\boldsymbol{\beta}})/(n-k+r)$. Putting this estimate together with the expressions on each side of (3.45), we can construct two apparently different, but numerically identical, test statistics. The first of these is

$$\tilde{\boldsymbol{\lambda}}^\top\boldsymbol{R}\big(\tilde{s}^2\tilde{\boldsymbol{X}}^\top\tilde{\boldsymbol{X}}\big)^{-1}\boldsymbol{R}^\top\tilde{\boldsymbol{\lambda}}=\frac{1}{\tilde{s}^2}\tilde{\boldsymbol{\lambda}}^\top\boldsymbol{R}\big(\tilde{\boldsymbol{X}}^\top\tilde{\boldsymbol{X}}\big)^{-1}\boldsymbol{R}^\top\tilde{\boldsymbol{\lambda}}. \tag{3.47}$$

In this form, the test statistic is clearly a Lagrange multiplier statistic. Since $\tilde{\boldsymbol{\lambda}}$ is an r–vector, it should not be surprising that this statistic would be asymptotically distributed as $\chi^2(r)$. A proof that this is the case follows from essentially the same arguments used in the case of the Wald test, since (3.47) is a quadratic form similar to (3.37). Of course, the result depends critically on the vector $\tilde{\boldsymbol{\lambda}}$ being asymptotically normally distributed, something that we will prove in Chapter 5.

The second test statistic, which we stress is numerically identical to the first, is obtained by substituting $\tilde{\boldsymbol{X}}^\top(\boldsymbol{y}-\tilde{\boldsymbol{x}})$ for $\boldsymbol{R}^\top\tilde{\boldsymbol{\lambda}}$ in (3.47). The result, which is the **score form** of the LM statistic, is

$$\frac{1}{\tilde{s}^2}(\boldsymbol{y}-\tilde{\boldsymbol{x}})^\top\tilde{\boldsymbol{X}}\big(\tilde{\boldsymbol{X}}^\top\tilde{\boldsymbol{X}}\big)^{-1}\tilde{\boldsymbol{X}}^\top(\boldsymbol{y}-\tilde{\boldsymbol{x}})=\frac{1}{\tilde{s}^2}(\boldsymbol{y}-\tilde{\boldsymbol{x}})^\top\tilde{\boldsymbol{P}}_X(\boldsymbol{y}-\tilde{\boldsymbol{x}}), \tag{3.48}$$

where $\tilde{\boldsymbol{P}}_X\equiv\tilde{\boldsymbol{X}}(\tilde{\boldsymbol{X}}^\top\tilde{\boldsymbol{X}})^{-1}\tilde{\boldsymbol{X}}^\top$. It is evident that this expression is simply the explained sum of squares from the **artificial linear regression**

$$\frac{1}{\tilde{s}}(\boldsymbol{y}-\tilde{\boldsymbol{x}})=\tilde{\boldsymbol{X}}\boldsymbol{b}+\text{residuals}, \tag{3.49}$$