Solution to Exercise 10.14

⋆10.14 Consider the Wald statistic W, the likelihood ratio statistic LR, and the Lagrange multiplier statistic LM for testing the hypothesis that $\boldsymbol{\beta}_2 = \mathbf{0}$ in the linear regression model (10.106). Since these are asymptotic tests, all the estimates of σ^2 are computed using the sample size n in the denominator. Express these three statistics as functions of the squared norms of the three components of the threefold decomposition (4.37) of the dependent variable $\boldsymbol{y}$. By use of the inequalities

$$x > \log(1+x) > \frac{x}{1+x}, \quad x > 0,$$

show that W > LR > LM.

The threefold decomposition (4.37) is

$$\boldsymbol{y} = \boldsymbol{P}_1\boldsymbol{y} + \boldsymbol{P}_{\boldsymbol{M}_1\boldsymbol{X}_2}\boldsymbol{y} + \boldsymbol{M}_{\boldsymbol{X}}\boldsymbol{y}. \tag{4.37}$$

This tells us that the variation in $\boldsymbol{y}$ can be divided into three orthogonal parts. The first term, $\boldsymbol{P}_1\boldsymbol{y}$, is the part that is explained by $\boldsymbol{X}_1$ alone. The second term, $\boldsymbol{P}_{\boldsymbol{M}_1\boldsymbol{X}_2}\boldsymbol{y}$, is the additional part that is explained by adding $\boldsymbol{X}_2$ to the regression. The final term, $\boldsymbol{M}_{\boldsymbol{X}}\boldsymbol{y}$, is the part that is not explained by the regressors.

The Wald statistic was given in the solution to Exercise 10.13 as

$$\mathrm{W} = \frac{1}{\hat{\sigma}^2}\,\boldsymbol{y}^{\top}\boldsymbol{M}_1\boldsymbol{X}_2(\boldsymbol{X}_2^{\top}\boldsymbol{M}_1\boldsymbol{X}_2)^{-1}\boldsymbol{X}_2^{\top}\boldsymbol{M}_1\boldsymbol{y}, \tag{S10.27}$$

which can be rewritten as

$$\mathrm{W} = n\,\frac{\|\boldsymbol{P}_{\boldsymbol{M}_1\boldsymbol{X}_2}\boldsymbol{y}\|^2}{\|\boldsymbol{M}_{\boldsymbol{X}}\boldsymbol{y}\|^2}.$$

The LM statistic is given by expression (10.74). Since $\boldsymbol{y} - \boldsymbol{X}\tilde{\boldsymbol{\beta}} = \boldsymbol{M}_1\boldsymbol{y}$ in this case, we have

$$\mathrm{LM} = n\,\frac{\boldsymbol{y}^{\top}\boldsymbol{M}_1\boldsymbol{P}_{\boldsymbol{X}}\boldsymbol{M}_1\boldsymbol{y}}{\boldsymbol{y}^{\top}\boldsymbol{M}_1\boldsymbol{y}}.$$

From equation (4.37), we see that $\boldsymbol{M}_1\boldsymbol{y} = \boldsymbol{P}_{\boldsymbol{M}_1\boldsymbol{X}_2}\boldsymbol{y} + \boldsymbol{M}_{\boldsymbol{X}}\boldsymbol{y}$, and so the denominator of LM is $\|\boldsymbol{P}_{\boldsymbol{M}_1\boldsymbol{X}_2}\boldsymbol{y}\|^2 + \|\boldsymbol{M}_{\boldsymbol{X}}\boldsymbol{y}\|^2$. Equation (4.37) also implies that $\boldsymbol{P}_{\boldsymbol{X}} = \boldsymbol{P}_1 + \boldsymbol{P}_{\boldsymbol{M}_1\boldsymbol{X}_2}$, from which we see that $\boldsymbol{M}_1\boldsymbol{P}_{\boldsymbol{X}}\boldsymbol{M}_1 = \boldsymbol{P}_{\boldsymbol{M}_1\boldsymbol{X}_2}$. Thus the numerator of LM is $\|\boldsymbol{P}_{\boldsymbol{M}_1\boldsymbol{X}_2}\boldsymbol{y}\|^2$, and so

$$\mathrm{LM} = n\,\frac{\|\boldsymbol{P}_{\boldsymbol{M}_1\boldsymbol{X}_2}\boldsymbol{y}\|^2}{\|\boldsymbol{P}_{\boldsymbol{M}_1\boldsymbol{X}_2}\boldsymbol{y}\|^2 + \|\boldsymbol{M}_{\boldsymbol{X}}\boldsymbol{y}\|^2}. \tag{S10.30}$$

It follows readily from this that

$$\mathrm{LM}/n = \frac{\mathrm{W}/n}{1 + \mathrm{W}/n}. \tag{S10.31}$$

Now consider the LR statistic. Equation (10.12) gives the maximized value of the concentrated loglikelihood function for a linear regression model. For the unrestricted model, this is

$$-\frac{n}{2}(1 + \log 2\pi - \log n) - \frac{n}{2}\log \boldsymbol{y}^\top \boldsymbol{M_X y},$$

and for the restricted model it is

$$\begin{aligned} &-\frac{n}{2}(1 + \log 2\pi - \log n) - \frac{n}{2}\log \boldsymbol{y}^\top \boldsymbol{M}_1 \boldsymbol{y} \\ &\quad = -\frac{n}{2}(1 + \log 2\pi - \log n) - \frac{n}{2}\log(\boldsymbol{y}^\top \boldsymbol{M_X y} + \boldsymbol{y}^\top \boldsymbol{P}_{\boldsymbol{M}_1 \boldsymbol{X}_2} \boldsymbol{y}). \end{aligned}$$

Therefore,

$$\begin{aligned} \mathrm{LR} &= -n\log \boldsymbol{y}^\top \boldsymbol{M_X y} + n\log(\boldsymbol{y}^\top \boldsymbol{M_X y} + \boldsymbol{y}^\top \boldsymbol{P}_{\boldsymbol{M}_1 \boldsymbol{X}_2} \boldsymbol{y}) \\ &= n\log\left(\frac{\|\boldsymbol{M_X y}\|^2 + \|\boldsymbol{P}_{\boldsymbol{M}_1 \boldsymbol{X}_2} \boldsymbol{y}\|^2}{\|\boldsymbol{M_X y}\|^2}\right) \\ &= n\log(1 + \mathrm{W}/n). \end{aligned}$$

Thus

$$\mathrm{LR}/n = \log(1 + \mathrm{W}/n). \tag{S10.32}$$

The desired inequality now follows directly from (S10.31) and (S10.32). Since the inequalities stated in the question imply that

$$\mathrm{W}/n > \log(1 + \mathrm{W}/n) > \frac{\mathrm{W}/n}{1 + \mathrm{W}/n},$$

we have shown that W > LR > LM.