Solution to Exercise 6.8

⋆6.8 Show that a Taylor expansion to second order of an NLS residual gives

$$\hat{u}_t = u_t - \boldsymbol{X}_t(\boldsymbol{\beta}_0)(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta}_0) - \tfrac{1}{2}(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta}_0)^\top \bar{\boldsymbol{H}}_t(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta}_0), \tag{6.11}$$

where $\boldsymbol{\beta}_0$ is the parameter vector of the DGP, and the $k \times k$ matrix $\bar{\boldsymbol{H}}_t \equiv \boldsymbol{H}_t(\bar{\boldsymbol{\beta}})$ is the matrix of second derivatives with respect to $\boldsymbol{\beta}$ of the regression function $x_t(\boldsymbol{\beta})$, evaluated at some $\bar{\boldsymbol{\beta}}$ that satisfies (6.19).

Define $\boldsymbol{b} \equiv n^{1/2}(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta}_0)$. As $n \to \infty$, $\boldsymbol{b}$ tends to the normal random variable $\operatorname{plim}(n^{-1}\boldsymbol{X}_0^\top\boldsymbol{X}_0)^{-1} n^{-1/2}\boldsymbol{X}_0^\top\boldsymbol{u}$. By expressing equation (6.11) in terms of $\boldsymbol{b}$, show that the difference between $\hat{\boldsymbol{u}}^\top\hat{\boldsymbol{u}}$ and $\boldsymbol{u}^\top\boldsymbol{M}_{\boldsymbol{X}_0}\boldsymbol{u}$ tends to 0 as $n \to \infty$. Here $\boldsymbol{M}_{\boldsymbol{X}_0} \equiv \mathbf{I} - \boldsymbol{P}_{\boldsymbol{X}_0}$ is the orthogonal projection on to $\mathcal{S}^\perp(\boldsymbol{X}_0)$.

Under the specified DGP,

$$\hat{u}_t \equiv y_t - x_t(\hat{\boldsymbol{\beta}}) = x_t(\boldsymbol{\beta}_0) - x_t(\hat{\boldsymbol{\beta}}) + u_t, \tag{S6.12}$$

which depends on $\hat{\boldsymbol{\beta}}$ though $x_t(\hat{\boldsymbol{\beta}})$. A second-order Taylor expansion of the latter around $\boldsymbol{\beta}_0$ is

$$x_t(\hat{\boldsymbol{\beta}}) = x_t(\boldsymbol{\beta}_0) + \boldsymbol{X}_t(\boldsymbol{\beta}_0)(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta}_0) + \tfrac{1}{2}(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta}_0)^\top \bar{\boldsymbol{H}}_t(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta}_0). \tag{S6.13}$$

Substituting (S6.13) into the right-hand side of (S6.12) yields expression (6.11) for $\hat{u}_t$, as required.

We can rewrite (6.11) using the notation $\boldsymbol{b} = n^{1/2}(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta}_0)$ as follows:

$$\hat{u}_t = u_t - n^{-1/2}\boldsymbol{X}_t(\boldsymbol{\beta}_0)\boldsymbol{b} - \tfrac{1}{2}n^{-1}\boldsymbol{b}^\top\bar{\boldsymbol{H}}_t\boldsymbol{b}. \tag{S6.14}$$

The sum of squared residuals, $\hat{\boldsymbol{u}}^\top\hat{\boldsymbol{u}}$, is the sum over all t of the square of the right-hand side of (S6.14). This square is the sum of six terms, of which the first is just u_t^2, which, summed over t, equals $\boldsymbol{u}^\top\boldsymbol{u}$ in matrix notation.

All the other terms have a factor which is a negative power of n. When these terms are summed over t, the result may therefore tend to zero as $n \to \infty$. Unfortunately, this is true for some terms, but not all, and so we must consider each of them separately. We look first at the square of the second term on the right-hand side of (S6.14). It can be written as $n^{-1}\boldsymbol{b}^\top\boldsymbol{X}_t^\top(\boldsymbol{\beta}_0)\boldsymbol{X}_t(\boldsymbol{\beta}_0)\boldsymbol{b}$, and the sum over t is therefore $n^{-1}\boldsymbol{b}^\top\boldsymbol{X}_0^\top\boldsymbol{X}_0\boldsymbol{b}$. As $n \to \infty$, the limit of this expression is the same as that of

$$n^{-1/2}\boldsymbol{u}^\top\boldsymbol{X}_0(n^{-1}\boldsymbol{X}_0^\top\boldsymbol{X}_0)^{-1}n^{-1}\boldsymbol{X}_0^\top\boldsymbol{X}_0(n^{-1}\boldsymbol{X}_0^\top\boldsymbol{X}_0)^{-1}n^{-1/2}\boldsymbol{X}_0^\top\boldsymbol{u}, \tag{S6.15}$$

where we use the fact that the k–vectors $\boldsymbol{b}$ and $(n^{-1}\boldsymbol{X}_0^\top\boldsymbol{X}_0)^{-1}n^{-1/2}\boldsymbol{X}_0^\top\boldsymbol{u}$ tend to the same limit. On cancelling the powers of n and using the definition of the projection matrix $\boldsymbol{P}_0$, we see that (S6.15) is equal to $\boldsymbol{u}^\top\boldsymbol{P}_0\boldsymbol{u}$.

The square of the third term on the right-hand side of equation (S6.14), summed over t, gives the contribution

$$\frac{1}{4n^2}\sum_{t=1}^{n}(\boldsymbol{b}^\top\bar{\boldsymbol{H}}_t\boldsymbol{b})^2.$$

This contribution does tend to zero as $n \to \infty$, since, by a law of large numbers, the sum divided by n has a finite, deterministic, limit. When this limit is divided once more by n, the limit is 0. A similar argument shows that the sum of the product of the second and third terms on the right-hand side of (S6.14) also tends to 0.

Summing twice the product of the first and third terms over t yields

$$-\tfrac{1}{n}\boldsymbol{b}^\top\Big(\sum_{t=1}^{n}u_t\bar{\boldsymbol{H}}_t\Big)\boldsymbol{b}. \tag{S6.16}$$

Since the regression functions $x_t(\boldsymbol{\beta})$ are assumed to be predetermined with respect to the error terms u_t, the same must be true of the second derivatives of the regression functions, which implies that $\mathrm{E}(u_t\bar{\boldsymbol{H}}_t) = \mathbf{O}$. Therefore, by a law of large numbers, the limit of the contribution (S6.16) is zero.

Finally, we find that twice the product of the first two terms on the right-hand side of (S6.14), summed over t, gives a contribution of

$$-2n^{-1/2}\sum_{t=1}^{n}u_t\boldsymbol{X}_t(\boldsymbol{\beta}_0)\boldsymbol{b} = -2n^{-1/2}\boldsymbol{u}^\top\boldsymbol{X}_0\boldsymbol{b}.$$

If we replace $\boldsymbol{b}$ by $(n^{-1}\boldsymbol{X}_0^\top\boldsymbol{X}_0)^{-1}n^{-1/2}\boldsymbol{X}_0^\top\boldsymbol{u}$ and cancel powers of n, we obtain just $-2\boldsymbol{u}^\top\boldsymbol{P}_{\boldsymbol{X}_0}\boldsymbol{u}$. Thus we have proved that

$$\hat{\boldsymbol{u}}^\top\hat{\boldsymbol{u}} = \boldsymbol{u}^\top\boldsymbol{u} + \boldsymbol{u}^\top\boldsymbol{P}_{\boldsymbol{X}_0}\boldsymbol{u} - 2\boldsymbol{u}^\top\boldsymbol{P}_{\boldsymbol{X}_0}\boldsymbol{u} = \boldsymbol{u}^\top\boldsymbol{u} - \boldsymbol{u}^\top\boldsymbol{P}_{\boldsymbol{X}_0}\boldsymbol{u} = \boldsymbol{u}^\top\boldsymbol{M}_{\boldsymbol{X}_0}\boldsymbol{u},$$

plus some additional terms that tend to zero as $n \to \infty$. This is just what we were asked to prove.