Solution to Exercise 10.16

⋆10.16 Show that the artificial OPG regression (10.73) possesses all the properties needed for hypothesis testing in the context of a model estimated by maximum likelihood. Specifically, show that

- the regressand $\boldsymbol{\iota}$ is orthogonal to the regressors $\boldsymbol{G}(\boldsymbol{\theta})$ when the latter are evaluated at the MLE $\hat{\boldsymbol{\theta}}$;
- the estimated OLS covariance matrix from (10.73) evaluated at $\hat{\boldsymbol{\theta}}$, when multiplied by n, consistently estimates the inverse of the asymptotic information matrix;
- the OPG regression (10.73) allows one-step estimation: If the OLS parameter estimates $\acute{\boldsymbol{c}}$ from (10.73) are evaluated at $\boldsymbol{\theta} = \acute{\boldsymbol{\theta}}$, where $\acute{\boldsymbol{\theta}}$ is any root-n consistent estimator of $\boldsymbol{\theta}$, then the one-step estimator $\grave{\boldsymbol{\theta}} \equiv \acute{\boldsymbol{\theta}} + \acute{\boldsymbol{c}}$ is asymptotically equivalent to $\hat{\boldsymbol{\theta}}$, in the sense that $n^{1/2}(\grave{\boldsymbol{\theta}} - \boldsymbol{\theta}_0)$ and $n^{1/2}(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}_0)$ tend to the same random variable as $n \to \infty$.

The first result is very easy to demonstrate. The inner product of the regressand $\boldsymbol{\iota}$ with the matrix of regressors $\hat{\boldsymbol{G}} \equiv \boldsymbol{G}(\hat{\boldsymbol{\theta}})$ is

$$\boldsymbol{\iota}^\top \hat{\boldsymbol{G}} = \boldsymbol{g}(\hat{\boldsymbol{\theta}}) = \boldsymbol{0},$$

by the first-order conditions (10.14). Thus we see that the regressand is indeed orthogonal to the matrix of regressors when the latter is evaluated at $\hat{\boldsymbol{\theta}}$.

The second result is also easily shown. The OLS covariance matrix from the artificial regression (10.73) evaluated at $\hat{\boldsymbol{\theta}}$ is

$$s^2(\hat{\boldsymbol{G}}^\top \hat{\boldsymbol{G}})^{-1}, \text{ where } s^2 = \frac{n}{n-k}. \tag{S10.37}$$

Clearly, s^2 tends to 1 as $n \to \infty$. Moreover, as we saw in Section 10.4, $n^{-1}\hat{\boldsymbol{G}}^\top \hat{\boldsymbol{G}}$ consistently estimates $\mathfrak{I}(\boldsymbol{\theta})$. It follows that n times expression (S10.37) must consistently estimate $\mathfrak{I}^{-1}(\boldsymbol{\theta})$.

For the final result, we start with the usual expression for the OLS estimates from regression (10.73) evaluated at $\acute{\boldsymbol{\theta}}$, multiplying each factor by the appropriate powers of n for asymptotic analysis:

$$n^{1/2}\acute{\boldsymbol{c}} = (n^{-1}\acute{\boldsymbol{G}}^\top \acute{\boldsymbol{G}})^{-1} n^{-1/2} \acute{\boldsymbol{G}}^\top \boldsymbol{\iota}. \tag{S10.38}$$

A Taylor expansion of the right-hand side of (S10.38) around $\boldsymbol{\theta}_0$ yields the result that

$$\begin{aligned} n^{1/2}\acute{\boldsymbol{c}} \overset{a}{=}\; & (n^{-1}\boldsymbol{G}_0^\top \boldsymbol{G}_0)^{-1} n^{-1/2} \boldsymbol{G}_0^\top \boldsymbol{\iota} \\ & + (n^{-1}\boldsymbol{G}_0^\top \boldsymbol{G}_0)^{-1} n^{-1} \boldsymbol{H}_0 n^{1/2}(\acute{\boldsymbol{\theta}} - \boldsymbol{\theta}_0), \end{aligned} \tag{S10.39}$$

where $\boldsymbol{G}_0 \equiv \boldsymbol{G}(\boldsymbol{\theta}_0)$ and $\boldsymbol{H}_0 \equiv \boldsymbol{H}(\boldsymbol{\theta}_0)$. There should be an additional term in (S10.39), but it is asymptotically negligible and is therefore ignored. Taking probability limits of both sides, and using the information matrix equality, we see that

$$\operatorname*{plim}_{n\to\infty} n^{1/2}\acute{\boldsymbol{c}} = \mathfrak{I}^{-1}(\boldsymbol{\theta}_0) \operatorname*{plim}_{n\to\infty} n^{-1/2}\boldsymbol{g}_0 - \operatorname*{plim}_{n\to\infty} n^{1/2}(\acute{\boldsymbol{\theta}} - \boldsymbol{\theta}_0),$$

which can be rearranged to yield

$$\operatorname*{plim}_{n\to\infty} n^{1/2}(\acute{\boldsymbol{\theta}} + \acute{\boldsymbol{c}} - \boldsymbol{\theta}_0) = \mathfrak{I}^{-1}(\boldsymbol{\theta}_0) \operatorname*{plim}_{n\to\infty} n^{-1/2}\boldsymbol{g}_0. \tag{S10.40}$$

The right-hand side of (S10.40) is the probability limit of the right-hand side of the asymptotic equality (10.38), the left-hand side of which is $n^{1/2}(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}_0)$. Therefore, since $\grave{\boldsymbol{\theta}} = \acute{\boldsymbol{\theta}} + \acute{\boldsymbol{c}}$, we conclude that

$$\operatorname*{plim}_{n\to\infty} n^{1/2}(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}_0) = \operatorname*{plim}_{n\to\infty} n^{1/2}(\grave{\boldsymbol{\theta}} - \boldsymbol{\theta}_0),$$

which establishes the one-step property for the artificial regression (10.73).