Solution to Exercise 9.17

⋆9.17 Suppose the n–vector $\boldsymbol{f}(\boldsymbol{\theta})$ of elementary zero functions has a covariance matrix $\sigma^2\mathbf{I}$. Show that, if the instrumental variables used for GMM estimation are the columns of the $n \times l$ matrix $\boldsymbol{W}$, the GMM criterion function is

$$\frac{1}{\sigma^2}\boldsymbol{f}^\top(\boldsymbol{\theta})\boldsymbol{P_W}\boldsymbol{f}(\boldsymbol{\theta}). \tag{9.124}$$

Next, show that, whenever the instruments are predetermined, the artificial regression

$$\boldsymbol{f}(\boldsymbol{\theta}) = -\boldsymbol{P_W}\boldsymbol{F}(\boldsymbol{\theta})\boldsymbol{b} + \text{residuals}, \tag{9.125}$$

where $\boldsymbol{F}(\boldsymbol{\theta})$ is defined as usual by (9.63), satisfies all the requisite properties for hypothesis testing. These properties, which are spelt out in detail in Exercise 8.20 in the context of the IVGNR, are that the regressand should be orthogonal to the regressors when they are evaluated at the GMM estimator obtained by minimizing (9.124); that the OLS covariance matrix from (9.125) should be a consistent estimate of the asymptotic variance of that estimator; and that (9.125) should admit one-step estimation.

In this case, the vector of empirical moments is $\boldsymbol{W}^\top\boldsymbol{f}(\boldsymbol{\theta})$. Since the covariance matrix of the vector of elementary zero functions is $\sigma^2\mathbf{I}$, the covariance matrix of the vector of empirical moments is $\sigma^2\boldsymbol{W}^\top\boldsymbol{W}$. Therefore, the appropriate criterion function is

$$\frac{1}{\sigma^2}\boldsymbol{f}^\top(\boldsymbol{\theta})\boldsymbol{W}(\boldsymbol{W}^\top\boldsymbol{W})^{-1}\boldsymbol{W}^\top\boldsymbol{f}(\boldsymbol{\theta}) = \frac{1}{\sigma^2}\boldsymbol{f}^\top(\boldsymbol{\theta})\boldsymbol{P_W}\boldsymbol{f}(\boldsymbol{\theta}),$$

which is identical to (9.124).

The first property the artificial regression (9.125) has to satisfy is that the regressand must be orthogonal to the regressors when they are evaluated at the GMM estimator. Because the first-order conditions that correspond to (9.124) are equivalent to the equations

$$\boldsymbol{F}^\top(\hat{\boldsymbol{\theta}})\boldsymbol{P_W}\boldsymbol{f}(\hat{\boldsymbol{\theta}}) = \mathbf{0}, \tag{S9.37}$$

where $\hat{\boldsymbol{\theta}}$ denotes the GMM estimator, it is clear that (9.125) does indeed satisfy this property.

The OLS covariance matrix from (9.125) is

$$s^2\big(\boldsymbol{F}^\top(\hat{\boldsymbol{\theta}})\boldsymbol{P_W}\boldsymbol{F}(\hat{\boldsymbol{\theta}})\big)^{-1},$$

where s^2 is the OLS estimate of the variance of the artificial regression. Since the regression has no explanatory power, $s^2 = \boldsymbol{f}^\top(\hat{\boldsymbol{\theta}})\boldsymbol{f}(\hat{\boldsymbol{\theta}})/(n-k)$. Because $\hat{\boldsymbol{\theta}}$ is consistent, it must be the case that

$$\operatorname*{plim}_{n\to\infty}\frac{1}{n-k}\boldsymbol{f}^\top(\hat{\boldsymbol{\theta}})\boldsymbol{f}(\hat{\boldsymbol{\theta}}) = \sigma^2.$$

This result and the consistency of $\hat{\boldsymbol{\theta}}$ then imply that

$$\operatorname*{plim}_{n\to\infty} s^2\Big(\frac{1}{n}\boldsymbol{F}^\top(\hat{\boldsymbol{\theta}})\boldsymbol{P_W}\boldsymbol{F}(\hat{\boldsymbol{\theta}})\Big)^{-1} = \sigma^2 \operatorname*{plim}_{n\to\infty}\Big(\frac{1}{n}\boldsymbol{F}^\top(\boldsymbol{\theta}_0)\boldsymbol{P_W}\boldsymbol{F}(\boldsymbol{\theta}_0)\Big)^{-1}.$$

The left-hand side of this equation is the plim of n times the covariance matrix from the artificial regression (9.125), evaluated at $\hat{\boldsymbol{\theta}}$, and the right-hand side, as was seen in the discussion following (9.77), is the asymptotic covariance matrix of $n^{1/2}(\hat{\boldsymbol{\theta}}-\boldsymbol{\theta}_0)$. Thus we have established that the artificial regression satisfies the second property.

For the one-step property, we need to show that, when (9.125) is evaluated at any consistent estimator $\acute{\boldsymbol{\theta}}$, the OLS parameter estimates $\acute{\boldsymbol{b}}$ are such that

$$n^{1/2}(\hat{\boldsymbol{\theta}}-\acute{\boldsymbol{\theta}}) \stackrel{a}{=} n^{1/2}\acute{\boldsymbol{b}}. \tag{S9.38}$$

The right-hand side of (S9.38) is easy to obtain. By the standard formula for the OLS estimator, introducing the factors of n that are needed for asymptotic analysis, we find that

$$n^{1/2}\acute{\boldsymbol{b}} = -(n^{-1}\acute{\boldsymbol{F}}^\top\boldsymbol{P_W}\acute{\boldsymbol{F}})^{-1}n^{-1/2}\acute{\boldsymbol{F}}^\top\boldsymbol{P_W}\acute{\boldsymbol{f}}, \tag{S9.39}$$

where $\acute{\boldsymbol{f}} \equiv \boldsymbol{f}(\acute{\boldsymbol{\theta}})$ and $\acute{\boldsymbol{F}} \equiv \boldsymbol{F}(\acute{\boldsymbol{\theta}})$.

For the left-hand side of (S9.38), we begin by performing a first-order Taylor expansion of the first-order conditions (S9.37) around $\acute{\boldsymbol{\theta}}$. The result is

$$\acute{\boldsymbol{F}}^\top\boldsymbol{P_W}\acute{\boldsymbol{f}} + \acute{\boldsymbol{F}}^\top\boldsymbol{P_W}\acute{\boldsymbol{F}}(\hat{\boldsymbol{\theta}}-\acute{\boldsymbol{\theta}}) \stackrel{a}{=} \boldsymbol{0}. \tag{S9.40}$$

There is no term involving the derivatives of $\boldsymbol{F}(\boldsymbol{\theta})$ because it can be shown that this term is asymptotically negligible relative to the second term on the left-hand side of (S9.40); recall the discussion centering on (6.61). Solving (S9.40), and introducing appropriate factors of n, yields the result that

$$n^{1/2}(\hat{\boldsymbol{\theta}}-\acute{\boldsymbol{\theta}}) \stackrel{a}{=} -(n^{-1}\acute{\boldsymbol{F}}^\top\boldsymbol{P_W}\acute{\boldsymbol{F}})^{-1}n^{-1/2}\acute{\boldsymbol{F}}^\top\boldsymbol{P_W}\acute{\boldsymbol{f}}. \tag{S9.41}$$

Since the right-hand side of (S9.41) is identical to the right-hand side of (S9.39), we have established that the artificial regression (9.125) has the one-step property (S9.38). Notice that this is the reason for the minus sign in that regression.