6.9 Final Remarks

In this chapter, we have dealt only with the estimation of nonlinear regression models by the method of moments and by nonlinear least squares. However, many of the results will reappear, in slightly different forms, when we consider estimation methods for other sorts of models. The NLS estimator is an **extremum estimator**, that is, an estimator obtained by minimizing or maximizing a criterion function. In the next few chapters, we will encounter several other extremum estimators: generalized least squares (Chapter 7), generalized instrumental variables (Chapter 8), the generalized method of moments (Chapter 9), and maximum likelihood (Chapter 10). Most of these estimators, like the NLS estimator, can be derived from the method of moments. All extremum estimators share a number of common features. Similar asymptotic results, and similar methods of proof, apply to all of them.

6.10 Exercises

6.1 Let the expectation of a random variable Y conditional on a set of other random variables $X_1, \ldots, X_k$ be the deterministic function $h(X_1, \ldots, X_k)$ of the conditioning variables. Let Ω be the information set consisting of all deterministic functions of the X_i, $i = 1, \ldots, k$. Show that $\mathrm{E}(Y \mid \Omega) = h(X_1, \ldots, X_k)$. **Hint:** Use the Law of Iterated Expectations for Ω and the information set defined by the X_i.

⋆**6.2** Consider a model similar to (3.20), but with error terms that are normally distributed:

$$y_t = \beta_1 + \beta_2 1/t + u_t, \quad u_t \sim \mathrm{NID}(0, \sigma^2),$$

where $t = 1, 2, \ldots, n$. If the true value of β_2 is β_2^0 and $\hat{\boldsymbol{\beta}}_2$ is the OLS estimator, show that the limit in probability of $\hat{\beta}_2 - \beta_2^0$ is a normal random variable with mean 0 and variance $6\sigma^2/\pi^2$. In order to obtain this result, you will need to use the results that

$$\sum_{t=1}^{\infty} (1/t)^2 = \pi^2/6,$$

and that, if $s(n) = \sum_{t=1}^{n} (1/t)$, then $\lim n^{-1} s(n) = 0$ and $\lim n^{-1} s^2(n) = 0$.

6.3 Show that the MM estimator defined by (6.10) depends on $\boldsymbol{W}$ only through the span $\mathcal{S}(\boldsymbol{W})$ of its columns. This is equivalent to showing that the estimator depends on $\boldsymbol{W}$ only through the orthogonal projection matrix $\boldsymbol{P_W}$.

6.4 Show algebraically that the first-order conditions for minimizing the SSR function (6.28) have the same solutions as the moment conditions (6.27).

6.5 Apply Taylor's Theorem to n^{-1} times the left-hand side of the moment conditions (6.27), expanding around the true parameter vector $\boldsymbol{\beta}_0$. Show that the extra term which appears here, but was absent in (6.20), where the instruments are fixed and we multiply by $n^{-1/2}$, tends to zero as $n \to \infty$. Make clear where and how you use a law of large numbers in your demonstration.