Solution to Exercise 15.6

⋆15.6 Consider a fully parametrized model for which the t^{th} observation is characterized by a conditional density function $f_t(\boldsymbol{y}^t, \boldsymbol{\theta})$, where the vector $\boldsymbol{y}^t$ contains the observations $y_1, \ldots, y_t$ on the dependent variable. The density is that of y_t conditional on $\boldsymbol{y}^{t-1}$. Let the moment function $m_t(\boldsymbol{\theta})$, which implicitly depends on y_t and possibly also on $\boldsymbol{y}^{t-1}$, have expectation zero conditional on $\boldsymbol{y}^{t-1}$ when evaluated at the true parameter vector $\boldsymbol{\theta}_0$. Show that

$$\mathrm{E}\big(m_t(\boldsymbol{\theta}_0)\boldsymbol{G}_t(\boldsymbol{\theta}_0)\big) = -\mathrm{E}\Big(\frac{\partial m_t(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}}\Big)\Big|_{\boldsymbol{\theta}=\boldsymbol{\theta}_0},$$

where $\boldsymbol{G}_t(\boldsymbol{\theta})$ is the row vector of derivatives of $\log f_t(\boldsymbol{y}^t, \boldsymbol{\theta})$, the contribution to the loglikelihood function made by the t^{th} observation, and $\partial m_t / \partial \boldsymbol{\theta}(\boldsymbol{\theta})$ denotes the row vector of derivatives of $m_t(\boldsymbol{\theta})$ with respect to $\boldsymbol{\theta}$. All expectations are taken under the density $f_t(\boldsymbol{y}^t, \boldsymbol{\theta})$. **Hint:** Use the same approach as in Exercise 10.6.

Explain why this result implies equation (15.26) under condition R2 of Section 15.2. **Hint:** Apply a central limit theorem to the appropriate expression.

The expectation $\mathrm{E}\big(m_t(\boldsymbol{\theta})\big)$ is equal to

$$\int_{-\infty}^{\infty} m_t(\boldsymbol{\theta}) f_t(\boldsymbol{y}^t, \boldsymbol{\theta})\, dy_t.$$

If we differentiate this expression with respect to $\boldsymbol{\theta}$ and set the vector of derivatives to zero, which we can do because $\mathrm{E}\big(m_t(\boldsymbol{\theta})\big) = 0$, we find that

$$\int_{-\infty}^{\infty} \boldsymbol{N}_t(\boldsymbol{\theta}) f_t(\boldsymbol{y}^t, \boldsymbol{\theta})\, dy_t + \int_{-\infty}^{\infty} m_t(\boldsymbol{\theta}) \boldsymbol{G}_t(\boldsymbol{\theta}) f_t(\boldsymbol{y}^t, \boldsymbol{\theta})\, dy_t = 0, \tag{S15.11}$$

where $\boldsymbol{N}_t(\boldsymbol{\theta})$ denotes the row vector of derivatives of $m_t(\boldsymbol{\theta})$ with respect to $\boldsymbol{\theta}$. The second term on the left-hand side of this equation uses the fact that

$$\boldsymbol{G}_t(\boldsymbol{\theta}) = \frac{\partial \log f_t(\boldsymbol{y}^t, \boldsymbol{\theta})}{\partial \boldsymbol{\theta}} = \frac{1}{f_t(\boldsymbol{y}^t, \boldsymbol{\theta})} \frac{\partial f_t(\boldsymbol{y}^t, \boldsymbol{\theta})}{\partial \boldsymbol{\theta}}.$$

Rearranging equation (S15.11), we find that

$$\int_{-\infty}^{\infty} m_t(\boldsymbol{\theta}) \boldsymbol{G}_t(\boldsymbol{\theta}) f_t(\boldsymbol{y}^t, \boldsymbol{\theta})\, dy_t = -\int_{-\infty}^{\infty} \boldsymbol{N}_t(\boldsymbol{\theta}) f_t(\boldsymbol{y}^t, \boldsymbol{\theta})\, dy_t.$$

In other words,

$$\mathrm{E}\big(m_t(\boldsymbol{\theta})\boldsymbol{G}_t(\boldsymbol{\theta})\big) = -\mathrm{E}\big(\boldsymbol{N}_t(\boldsymbol{\theta})\big). \tag{S15.12}$$

This is the first result we were required to show.

Equation (S15.12) can be rewritten as

$$\mathrm{E}\big(\boldsymbol{N}_t(\boldsymbol{\theta})\big) + m_t(\boldsymbol{\theta})\boldsymbol{G}_t(\boldsymbol{\theta})\big) = \mathbf{0}.$$

Now consider the expression

$$n^{-1/2}\sum_{t=1}^{n}\big(\boldsymbol{N}_t(\boldsymbol{\theta}_0) + m_t(\boldsymbol{\theta}_0)\boldsymbol{G}(\boldsymbol{\theta}_0)\big). \tag{S15.13}$$

We have just seen that the expectation of each term in this sum is 0. If we assume that we can apply a central limit theorem to it, it follows that the sum is $O_p(1)$. Therefore, dividing everything by $n^{1/2}$, we can write

$$\frac{1}{n}\sum_{t=1}^{n}\boldsymbol{N}_t(\boldsymbol{\theta}_0) = -\frac{1}{n}\sum_{t=1}^{n}m_t(\boldsymbol{\theta}_0)\boldsymbol{G}(\boldsymbol{\theta}_0) + O_p(n^{-1/2}).$$

But this is just equation (15.26) rewritten to use a more compact notation for the vector of derivatives of $m_t(\boldsymbol{\theta})$.