Solution to Exercise 13.22

⋆13.22 Consider the following artificial regression in connection with the model with GARCH(1,1) errors considered in the preceding exercise. Each real observation corresponds to two artificial observations. For observation t, the two corresponding elements of the regressand are

$$u_t/\sigma_t \quad \text{and} \quad (u_t^2-\sigma_t^2)/(\sigma_t^2\sqrt{2}).$$

The elements of the regressors that correspond to the elements of $\boldsymbol{\beta}$ are the elements of

$$\frac{\boldsymbol{X}_t}{\sigma_t} \quad \text{and} \quad -\frac{\alpha_1\sqrt{2}}{\sigma_t^2}\sum_{s=1}^{t-1}\delta_1^{s-1}\boldsymbol{X}_{t-s}u_{t-s}.$$

Similarly, the elements of the regressor that corresponds to α_0 are 0 and

$$\frac{1}{\sigma_t^2\sqrt{2}}\Big(\frac{1-\delta_1^{t-1}}{1-\delta_1}+\frac{\delta_1^{t-1}}{1-\alpha_1-\delta_1}\Big),$$

and those of the regressor that corresponds to α_1 are 0 and

$$\frac{1}{\sigma_t^2\sqrt{2}}\bigg(\sum_{s=1}^{t-1}\delta_1^{s-1}u_{t-s}^2+\frac{\alpha_0\delta_1^{t-1}}{(1-\alpha_1-\delta_1)^2}\bigg).$$

Finally, the elements of the regressor that corresponds to δ_1 are 0 and

$$\begin{aligned}\frac{1}{\sigma_t^2\sqrt{2}}\bigg(&-\frac{\alpha_0(t-1)\delta_1^{t-2}}{1-\delta_1}+\frac{\alpha_0(1-\delta_1^{t-1})}{(1-\delta_1)^2}\\&+\alpha_1\sum_{s=1}^{t-1}(s-1)\delta_1^{s-2}u_{t-s}^2+\frac{\alpha_0(t-1)\delta_1^{t-2}}{1-\alpha_1-\delta_1}+\frac{\alpha_0\delta_1^{t-1}}{(1-\alpha_1-\delta_1)^2}\bigg).\end{aligned}$$

Show that, when the regressand is orthogonal to the regressors, the sums over all the observations of the contributions (13.97) to the gradient of the loglikelihood are zero.

Let $\boldsymbol{R}(\boldsymbol{\beta},\boldsymbol{\theta})$ denote the $2n\times(k+3)$ matrix of the regressors, and let $\hat{\boldsymbol{\beta}}$ and $\hat{\boldsymbol{\theta}}$ denote the ML estimates. Then show that $\boldsymbol{R}^{\top}(\hat{\boldsymbol{\beta}},\hat{\boldsymbol{\theta}})\boldsymbol{R}(\hat{\boldsymbol{\beta}},\hat{\boldsymbol{\theta}})$ is the information matrix, where the contribution from observation t is computed as an expectation conditional on the information set Ω_t.

We have to show that gradient of the loglikelihood is zero when the regressand is orthogonal to the regressors. In order to do so, we show that, for each observation and for each parameter, the product of the first element of the regressand with the first element of the corresponding regressor, plus the product of the second element of the regressand with the second element of the regressor, is equal to the partial derivative of ℓ_t with respect to the parameter, as given by equations (13.97). The sum over t of these partial derivatives is an element of the gradient. Orthogonality of the regressand and the regressor thus implies that this element of the gradient vanishes.

First, for the elements of the parameter vector $\boldsymbol{\beta}$, we form the expression

$$\begin{aligned}
&\frac{u_t}{\sigma_t}\frac{\boldsymbol{X}_t}{\sigma_t} - \frac{u_t^2-\sigma_t^2}{\sigma_t^2\sqrt{2}}\frac{\alpha_1\sqrt{2}}{\sigma_t^2}\sum_{s=1}^{t-1}\delta_1^{s-1}\boldsymbol{X}_{t-s}u_{t-s} \\
&\quad = \frac{\boldsymbol{X}_t u_t}{\sigma_t^2} - \frac{\alpha_1(u_t^2-\sigma_t^2)}{\sigma_t^4}\sum_{s=1}^{t-1}\delta_1^{s-1}\boldsymbol{X}_{t-s}u_{t-s} = \frac{\partial\ell_t}{\partial\boldsymbol{\beta}},
\end{aligned}$$

by the first equation of (13.97).

Similarly, for α_0, we have

$$\frac{u_t^2-\sigma_t^2}{\sigma_t^2\sqrt{2}}\frac{1}{\sigma_t^2\sqrt{2}}\Big(\frac{1-\delta_1^{t-1}}{1-\delta_1} + \frac{\delta_1^{t-1}}{1-\alpha_1-\delta_1}\Big) = \frac{\partial\ell_t}{\partial\alpha_0}.$$

For α_1, we have

$$\frac{u_t^2-\sigma_t^2}{\sigma_t^2\sqrt{2}}\frac{1}{\sigma_t^2\sqrt{2}}\Big(\sum_{s=1}^{t-1}\delta_1^{s-1}u_{t-s}^2 + \frac{\alpha_0\delta_1^{t-1}}{(1-\alpha_1-\delta_1)^2}\Big) = \frac{\partial\ell_t}{\partial\alpha_1}.$$

Finally, for δ_1, we have

$$\begin{aligned}
&\frac{u_t^2-\sigma_t^2}{\sigma_t^2\sqrt{2}}\frac{1}{\sigma_t^2\sqrt{2}}\Big(-\frac{\alpha_0(t-1)\delta_1^{t-2}}{1-\delta_1} + \frac{\alpha_0(1-\delta_1^{t-1})}{(1-\delta_1)^2} \\
&\quad + \alpha_1\sum_{s=1}^{t-1}(s-1)\delta_1^{s-2}u_{t-s}^2 + \frac{\alpha_0(t-1)\delta_1^{t-2}}{1-\alpha_1-\delta_1} + \frac{\alpha_0\delta_1^{t-1}}{(1-\alpha_1-\delta_1)^2}\Big) = \frac{\partial\ell_t}{\partial\delta_1}.
\end{aligned}$$

Consider the contribution $\boldsymbol{I}_t(\boldsymbol{\beta},\boldsymbol{\theta})$ made by observation t to the information matrix. According to the definition proposed in the exercise, the element corresponding to the parameters β_i and β_j can be calculated using the first equation of (13.97) as

$$\begin{aligned}
\mathrm{E}\Big(\frac{\partial\ell_t}{\partial\beta_i}\frac{\partial\ell_t}{\partial\beta_j}\,\Big|\,\Omega_t\Big) &= \frac{x_{ti}x_{tj}}{\sigma_t^2} \\
&\quad + \frac{2\alpha_1^2}{\sigma_t^4}\Big(\sum_{s=1}^{t-1}\delta_1^{s-1}x_{(t-s)i}u_{t-s}\Big)\Big(\sum_{s=1}^{t-1}\delta_1^{s-1}x_{(t-s)j}u_{t-s}\Big),
\end{aligned}\tag{S13.31}$$

where x_{ti} is the ti^{th} element of $\boldsymbol{X}$, and we have used the facts that $\mathrm{E}(u_t^2)=\sigma_t^2$, $\mathrm{E}(u_t^2-\sigma_t^2)=0$, $\mathrm{E}\big(u_t(u_t^2-\sigma_t^2)\big)=0$, and $\mathrm{E}\big((u_t^2-\sigma_t^2)^2\big)=2\sigma_t^4$, all of which follow directly from the fact that $u_t\sim \mathrm{N}(0,\sigma_t^2)$. Note also that the lagged error terms belong to the information set Ω_t, and are therefore equal to their expectations conditional on Ω_t.

The contribution made by observation t to the element of the cross-product matrix $\boldsymbol{R}^\top(\boldsymbol{\beta},\boldsymbol{\theta})\boldsymbol{R}(\boldsymbol{\beta},\boldsymbol{\theta})$ that corresponds to β_i and β_j is the sum of two terms. The first is the product of the first element of the regressor for parameter β_i and observation t with the first element of the regressor for β_j and the same observation. The second is the analogous product with the second elements of the regressors. The sum of these terms is therefore

$$\frac{x_{ti}\,x_{tj}}{\sigma_t^2}+\frac{2\alpha_1^2}{\sigma_t^4}\Big(\sum_{s=1}^{t-1}\delta_1^{s-1}x_{(t-s)i}\,u_{t-s}\Big)\Big(\sum_{s=1}^{t-1}\delta_1^{s-1}x_{(t-s)j}\,u_{t-s}\Big),$$

which is equal to the right-hand side of equation (S13.31). Summing over t shows that the elements of the information matrix and those of the matrix $\boldsymbol{R}^\top(\boldsymbol{\beta},\boldsymbol{\theta})\boldsymbol{R}(\boldsymbol{\beta},\boldsymbol{\theta})$ corresponding to the parameters $\boldsymbol{\beta}$ are equal.

The calculations needed in order to extend this result to all the parameters are similar, but tedious. For the elements of $\boldsymbol{\beta}$ and α_0, we see that

$$\mathrm{E}\Big(\frac{\partial\ell_t}{\partial\beta_i}\frac{\partial\ell_t}{\partial\alpha_0}\;\Big|\;\Omega_t\Big)=-\frac{\alpha_1}{\sigma_t^4}\Big(\sum_{s=1}^{t-1}\delta_1^{s-1}x_{(t-s)i}\,u_{t-s}\Big)\Big(\frac{1-\delta_1^{t-1}}{1-\delta_1}+\frac{\delta_1^{t-1}}{1-\alpha_1-\delta_1}\Big),$$

which is the contribution from observation t to the element of $\boldsymbol{R}^\top(\boldsymbol{\beta},\boldsymbol{\theta})\boldsymbol{R}(\boldsymbol{\beta},\boldsymbol{\theta})$ corresponding to β_i and α_0. For $\boldsymbol{\beta}$ and α_1,

$$\mathrm{E}\Big(\frac{\partial\ell_t}{\partial\beta_i}\frac{\partial\ell_t}{\partial\alpha_1}\;\Big|\;\Omega_t\Big)=$$
$$-\frac{\alpha_1}{\sigma_t^4}\Big(\sum_{s=1}^{t-1}\delta_1^{s-1}x_{(t-s)i}\,u_{t-s}\Big)\Big(\sum_{s=1}^{t-1}\delta_1^{s-1}u_{t-s}^2+\frac{\alpha_0\delta_1^{t-1}}{(1-\alpha_1-\delta_1)^2}\Big),$$

which is again the appropriate contribution to $\boldsymbol{R}^\top(\boldsymbol{\beta},\boldsymbol{\theta})\boldsymbol{R}(\boldsymbol{\beta},\boldsymbol{\theta})$. For β and δ_1,

$$\mathrm{E}\Big(\frac{\partial\ell_t}{\partial\beta_i}\frac{\partial\ell_t}{\partial\delta_1}\;\Big|\;\Omega_t\Big)=$$
$$-\frac{\alpha_1}{\sigma_t^4}\Big(\sum_{s=1}^{t-1}\delta_1^{s-1}x_{(t-s)i}\,u_{t-s}\Big)\Big(-\frac{\alpha_0(t-1)\delta_1^{t-2}}{1-\delta_1}+\frac{\alpha_0(1-\delta_1^{t-1})}{(1-\delta_1)^2}$$
$$+\,\alpha_1\sum_{s=1}^{t-1}(s-1)\delta_1^{s-2}u_{t-s}^2+\frac{\alpha_0(t-1)\delta_1^{t-2}}{1-\alpha_1-\delta_1}+\frac{\alpha_0\delta_1^{t-1}}{(1-\alpha_1-\delta_1)^2}\Big),$$

as it should be.

For the specifically GARCH parameters, we calculate for α_0 with itself that

$$\mathrm{E}\Big(\frac{\partial\ell_t}{\partial\alpha_0}\frac{\partial\ell_t}{\partial\alpha_0}\;\Big|\;\Omega_t\Big)=\frac{1}{2\sigma_t^4}\Big(\frac{1-\delta_1^{t-1}}{1-\delta_1}+\frac{\delta_1^{t-1}}{1-\alpha_1-\delta_1}\Big)^2,$$

for α_0 with α_1 that

$$\mathrm{E}\Big(\frac{\partial\ell_t}{\partial\alpha_0}\frac{\partial\ell_t}{\partial\alpha_1}\;\Big|\;\Omega_t\Big) = \frac{1}{2\sigma_t^4}\Big(\frac{1-\delta_1^{t-1}}{1-\delta_1}+\frac{\delta_1^{t-1}}{1-\alpha_1-\delta_1}\Big)\Big(\sum_{s=1}^{t-1}\delta_1^{s-1}u_{t-s}^2+\frac{\alpha_0\delta_1^{t-1}}{(1-\alpha_1-\delta_1)^2}\Big),$$

and for α_0 with δ_1 that

$$\begin{aligned}\mathrm{E}\Big(\frac{\partial\ell_t}{\partial\alpha_0}\frac{\partial\ell_t}{\partial\delta_1}\;\Big|\;\Omega_t\Big) = {} & \frac{1}{2\sigma_t^4}\Big(\frac{1-\delta_1^{t-1}}{1-\delta_1}+\frac{\delta_1^{t-1}}{1-\alpha_1-\delta_1}\Big)\Big(-\frac{\alpha_0(t-1)\delta_1^{t-2}}{1-\delta_1}+\frac{\alpha_0(1-\delta_1^{t-1})}{(1-\delta_1)^2}\\ & +\alpha_1\sum_{s=1}^{t-1}(s-1)\delta_1^{s-2}u_{t-s}^2+\frac{\alpha_0(t-1)\delta_1^{t-2}}{1-\alpha_1-\delta_1}+\frac{\alpha_0\delta_1^{t-1}}{(1-\alpha_1-\delta_1)^2}\Big).\end{aligned}$$

All of these are equal to the corresponding contributions to $\boldsymbol{R}^\top(\boldsymbol{\beta},\boldsymbol{\theta})\boldsymbol{R}(\boldsymbol{\beta},\boldsymbol{\theta})$. For α_1 with itself we have that

$$\mathrm{E}\Big(\frac{\partial\ell_t}{\partial\alpha_1}\frac{\partial\ell_t}{\partial\alpha_1}\;\Big|\;\Omega_t\Big) = \frac{1}{2\sigma_t^4}\Big(\sum_{s=1}^{t-1}\delta_1^{s-1}u_{t-s}^2+\frac{\alpha_0\delta_1^{t-1}}{(1-\alpha_1-\delta_1)^2}\Big)^2,$$

and for α_1 with δ_1 that

$$\begin{aligned}\mathrm{E}\Big(\frac{\partial\ell_t}{\partial\alpha_1}\frac{\partial\ell_t}{\partial\delta_1}\;\Big|\;\Omega_t\Big) = {} & \frac{1}{2\sigma_t^4}\Big(\sum_{s=1}^{t-1}\delta_1^{s-1}u_{t-s}^2+\frac{\alpha_0\delta_1^{t-1}}{(1-\alpha_1-\delta_1)^2}\Big)\Big(-\frac{\alpha_0(t-1)\delta_1^{t-2}}{1-\delta_1}+\frac{\alpha_0(1-\delta_1^{t-1})}{(1-\delta_1)^2}\\ & +\alpha_1\sum_{s=1}^{t-1}(s-1)\delta_1^{s-2}u_{t-s}^2+\frac{\alpha_0(t-1)\delta_1^{t-2}}{1-\alpha_1-\delta_1}+\frac{\alpha_0\delta_1^{t-1}}{(1-\alpha_1-\delta_1)^2}\Big).\end{aligned}$$

Finally, for δ_1 with itself, the conditional expectation

$$\begin{aligned}\mathrm{E}\Big(\frac{\partial\ell_t}{\partial\delta_1}\frac{\partial\ell_t}{\partial\delta_1}\;\Big|\;\Omega_t\Big) = {} & \frac{1}{2\sigma_t^4}\Big(-\frac{\alpha_0(t-1)\delta_1^{t-2}}{1-\delta_1}+\frac{\alpha_0(1-\delta_1^{t-1})}{(1-\delta_1)^2}\\ & +\alpha_1\sum_{s=1}^{t-1}(s-1)\delta_1^{s-2}u_{t-s}^2+\frac{\alpha_0(t-1)\delta_1^{t-2}}{1-\alpha_1-\delta_1}+\frac{\alpha_0\delta_1^{t-1}}{(1-\alpha_1-\delta_1)^2}\Big)^2\end{aligned}$$

is the proper contribution to $\boldsymbol{R}^\top(\boldsymbol{\beta},\boldsymbol{\theta})\boldsymbol{R}(\boldsymbol{\beta},\boldsymbol{\theta})$.