A brief summary of the univariate and multivariate GARCH-type models that were addressed in the lectures of prof. Michael McAleer – as interpreted by Lennart Hoogerheide

Univariate GARCH-type models

The most popular class of volatility models is the class of Generalized Autoregressive Conditional Heteroskedasticity (GARCH) type models. The first model in this class is the simple Autoregressive Conditional Heteroskedasticity (ARCH) model, introduced by Robert Engle in 1982. Denote the error term of the model by ε_t and let I_{t-1} be the information set that is available at time t-1. Then in the **ARCH(1)** model the conditional variance $h_t = \text{var}(\varepsilon_t \mid I_{t-1})$ is given by

$$h_t = \omega + \alpha \cdot \varepsilon_{t-1}^2$$

with ω, α parameters to be estimated. This model has many flaws, but is useful as a starting point to build better models. Since large 'shocks' (large values of ε_t^2) do not only affect the conditional variance in the next period, but for a much longer period, an appropriate ARCH model would require the inclusion of many lags of ε_{t-k}^2 . The estimation of an unrestricted **ARCH(q)** model with

$$h_t = \omega + \sum_{k=1}^{q} \alpha_k \cdot \varepsilon_{t-k}^2$$

would give problems: there are simply too many coefficients to be estimated. Therefore, one needs to impose restrictions on the coefficients α_k. The most popular way to do this is to use a Generalized Autoregressive Conditional Heteroskedasticity (GARCH) model. In the **GARCH(1,1)** model the conditional variance $h_t = \text{var}(\varepsilon_t \mid I_{t-1})$ is given by

$$h_t = \omega + \alpha \cdot \varepsilon_{t-1}^2 + \beta \cdot h_{t-1} \qquad (*)$$

Substituting

$$h_{t-1} = \omega + \alpha \cdot \varepsilon_{t-2}^2 + \beta \cdot h_{t-2}$$

into (*) leads to

$$h_t = \omega + \alpha \cdot \varepsilon_{t-1}^2 + \beta \cdot (\omega + \alpha \cdot \varepsilon_{t-2}^2 + \beta \cdot h_{t-2}) =$$

$$= (1+\beta)\omega + \alpha \cdot \varepsilon_{t-1}^2 + \beta\alpha \cdot \varepsilon_{t-2}^2 + \beta^2 h_{t-2}$$

If $|\beta|<1$, then repeating such substitutions shows that the GARCH(1,1) model is equivalent with the ARCH(∞) model

$$h_t = \tilde{\omega} + \sum_{k=1}^{\infty} \alpha_k \cdot \varepsilon_{t-k}^2$$

with $\tilde{\omega} = \dfrac{\omega}{1-\beta}$ and $\alpha_k = \alpha \cdot \beta^{k-1}$. The two parameters $\alpha > 0$ (reflecting the direct effect of a 'shock' on the conditional variance) and $\beta \geq 0$ (reflecting the rate at which the exponential decay of a shock's effect takes place) can be easily estimated, if we have enough data at hand (e.g. several years of daily data).

A property of the GARCH(1,1) model, that it shares with ARCH models, is that positive and negative innovations ε_t of the same size have the same effect on the conditional variance $h_t = \mathrm{var}(\varepsilon_t \mid I_{t-1})$, since ε_{t-1} only appears as the square ε_{t-1}^2. In many applications this is considered unrealistic, because 'bad news' typically has larger *impact* than 'good news'.

Therefore, asymmetric GARCH-type models have been developed. The two most well-known asymmetric GARCH-type models are the GJR model (of Glosten, Jagannathan and Runkle) and the Exponential GARCH (EGARCH) model.

The **GJR** model is given by

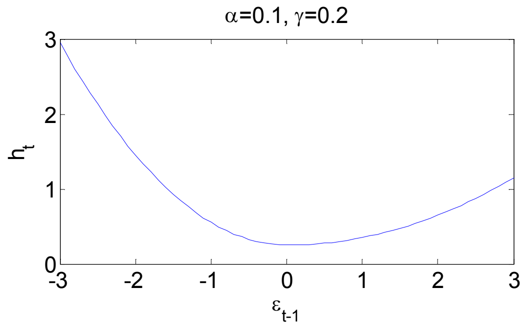
$$h_t = \omega + \alpha \cdot \varepsilon_{t-1}^2 + \gamma \cdot I\{\varepsilon_{t-1}\} \cdot \varepsilon_{t-1}^2 + \beta \cdot h_{t-1}$$

with

$$I\{\varepsilon_{t-1}\} = \begin{cases} 1 & \text{if } \varepsilon_{t-1} < 0 \\ 0 & \text{if } \varepsilon_{t-1} \geq 0 \end{cases}$$

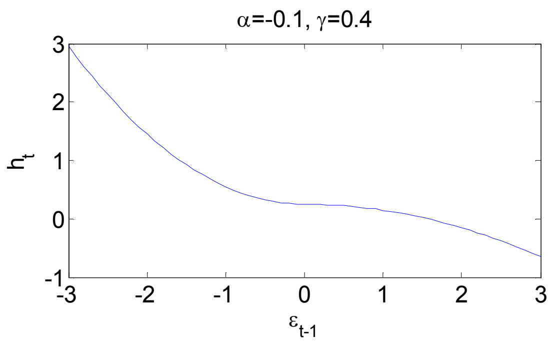
In the case of $\gamma = 0$ we have the original, symmetric GARCH model. If $\gamma \neq 0$ then we have asymmetry.

For example, if we have $\alpha = 0.1$, $\gamma = 0.2$, then the conditional variance h_t as a function of ε_{t-1} (for certain fixed values of h_{t-1}, ω) looks like:



It should be noted at this point that *asymmetry* and *leverage* are different concepts. Leverage is a special case of asymmetry, where larger (i.e. 'less negative' or 'more positive') ε_{t-1} lead to lower h_t than smaller (i.e. 'more negative' or 'less positive') ε_{t-1}. The reason is that for larger ε_{t-1}, the ratio debt/equity decreases (since the denominator equity increases), which reflects a reduction of the risk.

In the GJR model, leverage corresponds to the values $\alpha < 0$ (to make h_t a decreasing function of ε_{t-1} for $\varepsilon_{t-1} > 0$) and $\gamma + \alpha > 0$ (to make h_t a decreasing function of ε_{t-1} for $\varepsilon_{t-1} < 0$). For example, for $\alpha = -0.1$, $\gamma = 0.4$ the function $h_t(\varepsilon_{t-1})$ (for certain fixed values of h_{t-1}, ω) looks like:



The disadvantage is that for large enough positive shocks ε_{t-1}, the conditional variance h_t would become negative, which is obviously a serious flaw of the unrestricted version of this GJR model. We can prevent this by imposing the restrictions that $\alpha > 0$ and $\gamma + \alpha > 0$ (in addition to the obvious assumption of $\omega > 0$). But these restrictions that ensure a positive variance h_t also immediately exclude the possibility of leverage.

A model that ensures the positivity of the variance h_t in a different way is the **EGARCH** model:

$$\log h_t = \omega + \alpha \cdot |\eta_{t-1}| + \gamma \cdot \eta_{t-1} + \beta \cdot \log h_{t-1}$$

with η_{t-1} the 'so-called' standardized residual

$$\eta_{t-1} = \frac{\varepsilon_{t-1}}{\sqrt{h_{t-1}}}.$$

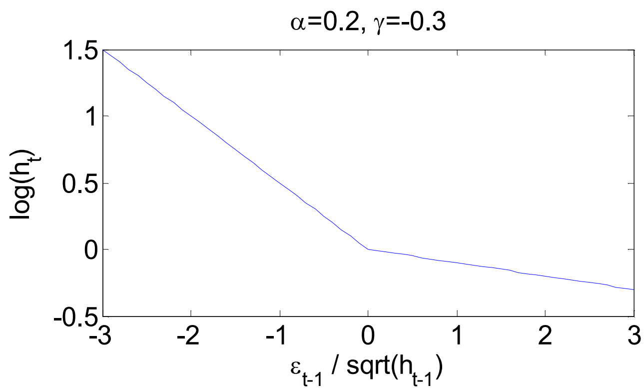
The equivalent equation

$$h_t = \exp(\omega + \alpha \cdot |\eta_{t-1}| + \gamma \cdot \eta_{t-1} + \beta \cdot \log h_{t-1})$$

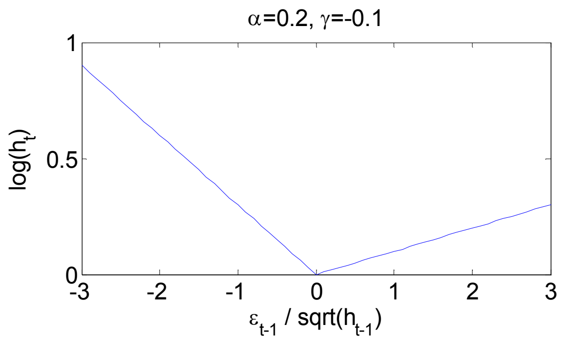
immediately shows that the exponent ensures the positivity of h_t, for any values of the past shocks and parameters. Therefore, the EGARCH model allows for the possibility of leverage without worrying about negative variance values. No *a priori* parameter restrictions are required for this purpose.

For $\gamma < 0$ and $\gamma < \alpha < -\gamma$ (in which case log h_t is a decreasing function of η_{t-1}, and hence h_t is a decreasing function of ε_{t-1} for a given value of h_{t-1}), we observe leverage.

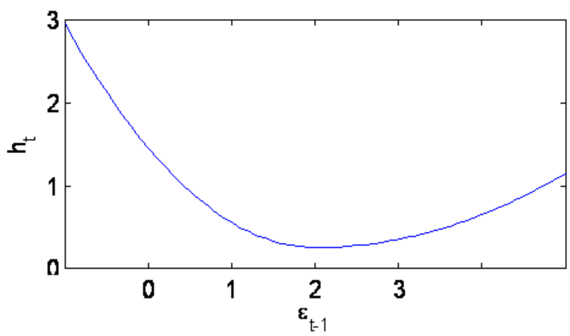
For example, for $\alpha = 0.2$, $\gamma = -0.3$ the log h_t as a function of $\eta_{t-1} = \varepsilon_{t-1} / \sqrt{h_{t-1}}$ (for certain fixed values of h_{t-1}, ω) looks like:



Alternatively, the EGARCH model also allows for asymmetry without the presence of leverage. For example, for $\alpha = 0.2$, $\gamma = -0.1$ the log h_t as a function of $\eta_{t-1} = \varepsilon_{t-1} / \sqrt{h_{t-1}}$ (for certain fixed values of h_{t-1}, ω) looks like:



In many applications, asymmetry is a plausible property of the GARCH model. Whether a plausible model should display leverage, is an issue of debate. One may also assume that for small positive returns the leverage effect is present (in the sense that small positive returns lead to a lower variance h_t than a return of 0), whereas large positive returns increase the conditional variance h_t, since one may then fear a ‘correction’ of the market. Then the h_t as a function of ε_{t-1} (for fixed values of h_{t-1}, ω) may look like:



Such a behavior could be obtained by a GJR-type model

$$h_t = \omega + \alpha \cdot (\varepsilon_{t-1} - \delta)^2 + \gamma \cdot I\{\varepsilon_{t-1}\} \cdot (\varepsilon_{t-1} - \delta)^2 + \beta \cdot h_{t-1}$$

with

$$I\{\varepsilon_{t-1}\} = \begin{cases} 1 & if \ \varepsilon_{t-1} < \delta \\ 0 & if \ \varepsilon_{t-1} \geq \delta \end{cases}$$

with an extra threshold parameter δ that allows the minimum of h_t to be taken for a positive value of ε_{t-1} (and with $\alpha > 0$ and $\gamma + \alpha > 0$ ensuring positivity of the variance).

It should be mentioned that the EGARCH model does have a relative disadvantage, as compared with the GJR model. There is no doubt on the asymptotic properties of the (maximum likelihood) estimator in the GJR model, consistency and asymptotic normality, so that one can perform a reliable test using the t-ratio in the GJR model (given that we have ‘enough’ data observations). It is doubtful whether the pseudo quasi maximum likelihood estimator (PQMLE) in the EGARCH model has these asymptotic properties. (In the Bayesian framework, the EGARCH model does not have this disadvantage, since roughly stated, we can simulate a set of draws from the posterior distribution in both models in basically the same way.)

Multivariate GARCH-type models

Many univariate GARCH-type models have been developed. In practice, however, we often need a model for the (conditional) covariance or correlation matrix of a group of asset returns within a portfolio, rather than the conditional variance of a single asset’s return. Two typical applications are: (1) the daily computation of the Value-at-Risk for a given portfolio, (2) the monitoring or optimization of the portfolio every few weeks. Note that the latter is not done on a daily basis, since we can not sell large numbers of stocks all the time.

The two most famous multivariate GARCH models are the BEKK model (of Baba, Engle, Kraft and Kroner) and the Dynamic Conditional Correlation (DCC) model.

The univariate GARCH(1,1) model’s equation

$$h_t = \omega + \alpha \cdot \varepsilon_{t-1}^2 + \beta \cdot h_{t-1}$$

can be rewritten as

$$h_t = \sqrt{\omega}\sqrt{\omega} + \sqrt{\alpha} \cdot \varepsilon_{t-1} \cdot \varepsilon_{t-1}\sqrt{\alpha} \cdot + \sqrt{\beta} \cdot h_{t-1} \cdot \sqrt{\beta}\,. \qquad (**)$$

In the **BEKK** model the $m \times m$ conditional covariance matrix $Q_t = \mathrm{var}(\varepsilon_t \mid I_{t-1})$ of the $m \times 1$ vector ε_t is given by

$$Q_t = Q \cdot Q' + A \cdot \varepsilon_{t-1} \cdot \varepsilon_{t-1}' \cdot A' + B \cdot Q_{t-1} \cdot B' \qquad (***)$$

with $m \times m$ parameter matrices Q, A, B. From (**) and (***) it is clear that the BEKK model is clearly a generalization of the GARCH(1,1) model to the multivariate case.

The BEKK model has two famous restricted cases:

(1) the scalar BEKK model

$$Q_t = Q \cdot Q' + \alpha \cdot \varepsilon_{t-1} \cdot \varepsilon_{t-1}' + \beta \cdot Q_{t-1}$$

with α and β scalars.

(2) the diagonal BEKK model where A, B are diagonal matrices.

In the unrestricted BEKK model, the number of parameters grows very rapidly with the dimension m. This causes huge computational problems: for large portfolios the time required for the estimation of the model is typically too long for any practical purposes. Moreover, the estimator may not have the key asymptotic properties of consistency and asymptotic normality. Therefore, a t-ratio can not be used since we do not know the appropriate critical value.

The scalar and diagonal BEKK model do not suffer from these disadvantages. These can be estimated in much less computing time, and the desired asymptotic properties of consistency and asymptotic normality are present, so that we can for example perform appropriate t-tests.

In the **DCC** model we focus directly on the conditional correlation matrix, rather than the conditional covariance matrix. In the DCC model we have

$$\Gamma_t = (1 - \theta_1 - \theta_2) \cdot \Gamma + \theta_1 \cdot \eta_{t-1} \cdot \eta_{t-1}' + \theta_2 \cdot \Gamma_{t-1}$$

for a ‘quasi conditional correlation matrix’ Γ_t of ε_t, which has to be scaled to become the true conditional correlation matrix $\Gamma_t^* = corr(\varepsilon_t \mid I_{t-1})$:

$$\Gamma_t^* = (diag(\Gamma_t))^{-1/2} \cdot \Gamma_t \cdot (diag(\Gamma_t))^{-1/2}$$

The $m \times 1$ vector η_{t-1} contains the ‘standardized residuals’. The parameters are Γ ($m \times m$) and θ_1, θ_2 (scalars). The DCC model suffers from serious flaws, including that the estimator of its parameters does not have the key asymptotic properties of consistency and asymptotic normality. Therefore, a t-ratio can not be used since we do not know the appropriate critical value.

In sum, it seems that the scalar and diagonal BEKK models should be preferred over the unrestricted BEKK model and the DCC model. The restrictions of the scalar and diagonal matrix imply that we are not really estimating a '*truly multivariate*' GARCH model: the restricted BEKK models can be considered as a 'set of univariate models'. It can be said that the estimation and analysis of '*truly multivariate*' GARCH models is not possible.