Chapter 4

Univariate Volatility Processes

4.1 Objectives[1]

A univariate model assumes only one source of randomness, in the case of volatility models the source of randomness is the conditional returns. Define, under measure P, the conditional returns as

$$\varepsilon_t = ln\frac{S_t}{S_{t-1}}$$

In this chapter two of the main univariate volatility processes are discussed. The Exponentially Weighted Moving Averages (EWMA) process is discussed in section 4.2 and the various GARCH processes is discussed in section 4.3 and further. This chapter includes a discussion on Asymmetric GARCH in section 4.7.

4.2 Exponentially Weighted Moving Averages

Weighing the $MA(q)$ process in equation 3.3, by the sum of its parameters yields

$$Z_t = \frac{\varepsilon_t + \lambda\varepsilon_{t-1} + \lambda^2\varepsilon_{t-2} + .. + \lambda^q\varepsilon_{t-q}}{1 + \lambda + \lambda^2 + .. + \lambda^q} \tag{4.1}$$

where $\theta_i = \lambda^i$ and $\lambda \in (0, 1)$.

Taking the limit of 4.1 to infinite

$$\begin{aligned} \lim_{q\to\infty} Z_t &= \lim_{q\to\infty} \frac{\varepsilon_t + \lambda\varepsilon_{t-1} + \lambda^2\varepsilon_{t-2} + .. + \lambda^q\varepsilon_{t-q}}{1 + \lambda + \lambda^2 + .. + \lambda^q} \\ &= (1-\lambda)\sum_{i=1}^{\infty}\lambda^i\varepsilon_{t-i} \end{aligned} \tag{4.2}$$

[1] Suggested reading: [1], [18] and [23].

since $\lambda \in (0,1)$.

Equation 4.2 is the basis of the EWMA conditional variance process,

$$\begin{aligned}
\hat{\sigma}_t^2 &= (1-\lambda)\sum_{i=1}^{\infty}\lambda^{i-1}\varepsilon_{t-i}^2 \\
&= (1-\lambda)\sum_{i=2}^{\infty}\lambda^{i-1}\varepsilon_{t-i}^2 + (1-\lambda)\,\varepsilon_{t-1}^2 \\
&= \lambda(1-\lambda)\sum_{i=1}^{\infty}\lambda^{i-2}\varepsilon_{(t-1)-i}^2 + (1-\lambda)\,\varepsilon_{t-1}^2 \\
&= \lambda\sigma_{t-1}^2 + (1-\lambda)\,\varepsilon_{t-1}^2 \qquad (4.3)
\end{aligned}$$

with $\lambda \in (0,1)$.

Alexander [1] interprets the smoothing constant λ in the following two ways:

1. The term, $(1-\lambda)\,\varepsilon_{t-1}^2$ determines the *intensity of reaction* of volatility to market events. A low value of λ will give a process highly reactive to shocks. The effect of these shocks will quickly die away. Lower values of λ is mostly used for short term forecasts.

2. Term $\lambda\sigma_{t-1}^2$ determines the *persistence in volatility.* A high λ will give a process that persists at a certain level of volatility, despite recent shocks.

Parameters of the EWMA process can be estimated by minimizing the root mean square error or similar method. The accuracy of forecasts are however difficult to assess.

4.2.1 RiskMetrics

The EWMA model is also the basis of volatility forecasts in the RiskMetrics system by J.P. Morgan. The RiskMetrics model has the following to distinctive features:

1. The parameter λ is fixed, $\lambda = 0.94$.

2. The definition of volatility is different than the standard definition of volatility. Under the assumption of normality, the RiskMetrics volatility is the 95^{th} percentile or 1.65 times the standard deviation.

4.3 Generalized Conditional Autoregressive Conditional Heteroscedasticity

The Autoregressive Conditional Heteroscedastic (ARCH) process was introduced by Engle (1982) [14]. This process allows for the change of conditional volatility over time as a function of past errors.

The Generalized Autoregressive Conditional Heteroscedastic (GARCH) process by Bollerslev (1986) [6] is the most popular and widely used stochastic volatility measure and forecasting method.

The $GARCH(p,q)$ process is discussed in section 4.4 below. It will be shown that this discussion encompasses the ARCH process in a simple way. The GARCH process is also the basis for many subsequent models.

4.4 GARCH(p,q)

The GARCH(p,q) process under conditionally normal, discrete time errors, is defined by

$$\begin{aligned} \varepsilon_t &\mid \mathcal{F}_{t-1} \sim N\left(0,\sigma_t^2\right) \\ \sigma_t^2 &= \alpha_0 + \sum_{i=1}^{q} \alpha_i \varepsilon_{t-i}^2 + \sum_{i=1}^{p} \beta_i \sigma_{t-i}^2 \end{aligned} \tag{4.4}$$

where p,q are nonnegative integers, α_i, β_j are nonnegative real numbers for every applicable i,j and α_0 is a positive real.

For $p,q=0$, the GARCH process is simple white noise. For $p=0$, $q\neq 0$ the process is an ARCH process. Thus, the GARCH process is to volatility what the ARMA process is to the AR process, for means.

Any GARCH(p,q) process can be defined as a GARCH$(1,1)$ process. Define

$$\sigma_t^2 = \alpha_0 + A(L)\,\varepsilon_t^2 + B(L)\,\sigma_t^2$$

where for lag operator L,

$$\begin{aligned} A(L) &= \sum_{i=1}^{q} \alpha_i L^i \\ B(L) &= \sum_{i=1}^{p} \beta_i L^i \end{aligned}$$

4.4.1 Stationarity

Theorem 4.4.1 *A $GARCH(p,q)$ process is stationary, with (long-term) variance*

$$E\left[\sigma_t^2\right] = \frac{\alpha_0}{1 - A(1) - B(1)}$$

for any t if and only if $A(1) + B(1) < 1$.

Proof. For any t

$$\begin{aligned} E\left[\sigma_t^2\right] &= E\left[var\left[\varepsilon_t \mid \mathcal{F}_{t-1}\right]\right] \\ &= E\left[E\left[\varepsilon_t^2 \mid \mathcal{F}_{t-1}\right]\right] \end{aligned}$$

since we assume that $E\left[\varepsilon_t \mid \mathcal{F}_{t-1}\right] = 0$. It follows that

$$E\left[\sigma_t^2\right] = E\left[\varepsilon_t^2\right]$$

by the tower property of conditional expectation. Since ε_t is white noise, we have that

$$var\left[\varepsilon_t\right] = E\left[\varepsilon_t^2\right] = \sigma^2$$

for all t, where σ^2 is the long-term variance of ε_t. It follows directly then that

$$E\left[\varepsilon_t^2\right] = E\left[\varepsilon_{t-1}^2\right]$$

and

$$E\left[\sigma_t^2\right] = E\left[\sigma_{t-1}^2\right]$$

The expected value of the $GARCH(p, q)$ process

$$\sigma_t^2 = \alpha_0 + \sum_{i=1}^{q} \alpha_i \varepsilon_{t-i}^2 + \sum_{i=1}^{p} \beta_i \sigma_{t-i}^2$$

is

$$\begin{aligned} \sigma^2 &= E\left[\sigma_t^2\right] \\ &= \alpha_0 + \sum_{i=1}^{q} \alpha_i E\left[\varepsilon_{t-i}^2\right] + \sum_{i=1}^{p} \beta_i E\left[\sigma_{t-i}^2\right] \\ &= \alpha_0 + \sum_{i=1}^{q} \alpha_i \sigma^2 + \sum_{i=1}^{p} \beta_i \sigma^2 \end{aligned}$$

It follows that

$$\sigma^2 \left(1 - \sum_{i=1}^{q} \alpha_i - \sum_{i=1}^{p} \beta_i\right) = \alpha_0$$

or

$$\sigma^2 = \frac{\alpha_0}{\left(1 - \sum_{i=1}^{q} \alpha_i - \sum_{i=1}^{p} \beta_i\right)}$$

For σ^2 to be finite it's required that

$$\sum_{i=1}^{q} \alpha_i + \sum_{i=1}^{p} \beta_i < 1$$

■

4.4.2 Stylized Facts

In the financial literature four properties of returns series have been coined, stylized facts. These stylized facts are volatility clustering, mean reversion, excess kurtosis and the leverage effect. The leverage effect is discussed in section 4.7.

A stationary GARCH process captures these stylized facts in the following ways:

1. Volatility clustering is described in section 2.7.3 as strong autocorrelation of squared returns. Thus if σ_{t-1}^2 is high (low), then σ_t^2 will probably also be high (low). The long-term variance of a $GARCH\,(p,q)$ process was provided in theorem 4.4.1. The long-term variance of a $GARCH\,(1,1)$ process is

$$E\left[\sigma_t^2\right] = \frac{\alpha_0}{1-\alpha-\beta} \equiv V \tag{4.5}$$

thus

$$\alpha_0 \equiv V\left(1-\alpha-\beta\right)$$

and

$$\sigma_t^2 = V\left(1-\alpha-\beta\right) + \alpha\varepsilon_{t-1}^2 + \beta\sigma_{t-1}^2$$

equivalently,

$$\sigma_t^2 - V = \alpha\left(\varepsilon_{t-1}^2 - V\right) + \beta\left(\sigma_{t-1}^2 - V\right)$$

Taking expected value yields

$$\begin{aligned} E\left[\sigma_t^2 - V \mid \mathcal{F}_{t-2}\right] &= E\left[\alpha\left(\varepsilon_{t-1}^2 - V\right) + \beta\left(\sigma_{t-1}^2 - V\right) \mid \mathcal{F}_{t-2}\right] \\ &= \alpha E\left[\varepsilon_{t-1}^2 - V \mid \mathcal{F}_{t-2}\right] + \beta\left(\sigma_{t-1}^2 - V\right) \\ &= \left(\alpha+\beta\right)\left(\sigma_{t-1}^2 - V\right) \qquad (4.6) \end{aligned}$$

since $E\left[\varepsilon_{t-1} \mid \mathcal{F}_{t-2}\right] = 0$ and $Var\left[\varepsilon_{t-1} \mid \mathcal{F}_{t-2}\right] = \sigma_{t-1}^2$. This equation can be rewritten as

$$E\left[\sigma_t^2 \mid \mathcal{F}_{t-2}\right] = V + \left(\alpha+\beta\right)\left(\sigma_{t-1}^2 - V\right)$$

thus if σ_{t-1}^2 is large (small) then it's expected for σ_t^2 also to be large (small).

2. Mean reversion is the gradual return of variance levels, after a shock, to a long-term variance level. Equation 4.6 can be rewritten as

$$E\left[\sigma_{t+k}^2 - V \mid \mathcal{F}_t\right] = \left(\alpha+\beta\right) E\left[\sigma_{t+k-1}^2 - V \mid \mathcal{F}_t\right]$$

By repeating this relationship yields

$$E\left[\sigma_{t+k}^2 - V \mid \mathcal{F}_t\right] = (\alpha+\beta)^k \left(\sigma_t^2 - V\right)$$

or

$$E\left[\sigma_{t+k}^2 \mid \mathcal{F}_t\right] = V + (\alpha+\beta)^k \left(\sigma_t^2 - V\right) \tag{4.7}$$

Since the GARCH process is stationary, $\alpha + \beta < 1$. This means that the second term of equation 4.7 tends to zero, as k tends to infinity. Thus the expected value of the conditional variance tends to the long-term variance level, V.

3. Excess kurtosis in returns series can be described as kurtosis, see section 2.5.2, larger than that of the normal distribution. In theorem 4.4.1 above, we proved that for the $GARCH\,(1,1)$ process

$$\begin{aligned} E\left[\varepsilon_t^2\right] &= E\left[\sigma_t^2\right] \\ &= \frac{\alpha_0}{1-\alpha-\beta} \end{aligned}$$

Bollerslev, see [6], proved that if $3\alpha^2 + 2\alpha\beta + \beta^2 < 1$ the stationary fourth moment of ε exists,

$$E\left[\varepsilon_t^4\right] = \frac{3\alpha_0^2 (1+\alpha+\beta)}{(1-\alpha-\beta)\left(1-\beta^2-2\alpha\beta-3\alpha^2\right)}$$

The stationary kurtosis is

$$K = \frac{E\left[\varepsilon_t^4\right]}{E\left[\varepsilon_t^2\right]^2} = \frac{3\left(1-(\alpha+\beta)^2\right)}{1-\beta^2-2\alpha\beta-3\alpha^2} > 3$$

thus the GARCH process is heavy-tailed (leptokurtic).

4.4.3 Estimation of GARCH Regression Model

This section focusses on the maximum likelihood estimation (MLE) of the GARCH regression model. The GARCH model in equation 4.4 may be written in terms of the following nonlinear regression model

$$\varepsilon_t = y_t - \mathbf{x}_t \mathbf{b}$$

which is the means process of the error ε_t, which is conditionally normal

$$\varepsilon_t \mid \mathcal{F}_{t-1} \sim N\left(0, \sigma_t^2\right)$$

where

$$\sigma_t^2 = \mathbf{z}_t'\boldsymbol{\omega}$$

is the GARCH(p,q) process. The vector

$$\mathbf{z}_t' = \left(1, \varepsilon_{t-1}^2, \ldots, \varepsilon_{t-q}^2, \sigma_{t-1}^2, \ldots, \sigma_{t-p}^2\right)$$

and parameter vector

$$\boldsymbol{\omega}' = \left(\alpha_0, \alpha_1, \ldots, \alpha_q, \beta_1, \ldots, \beta_p\right)$$

Define Θ as a compact subspace of a Euclidean space, with $\boldsymbol{\theta} = (\mathbf{b}', \boldsymbol{\omega}') \in \Theta$. Denote the true parameter values of by θ_0, where $\theta_0 \in int\,\Theta$.

The likelihood function of ε_t is the pdf of the error process ε_t, written in terms of its parameters

$$f^*\left(0, \sigma_t^2\right) = \prod_{t=1}^{T} \frac{1}{\sqrt{2\pi\sigma_t^2}} e^{-[\varepsilon_t/\sigma_t]^2/2} \tag{4.8}$$

since the conditional mean is zero and the process follows GARCH variance. There are T observations.

It's computationally easier to take the ln of equation 4.8. The loglikelihood function is

$$f\left(0, \sigma_t^2\right) = \sum_{t=1}^{T} -\frac{1}{2}\ln\sigma_t^2 - \frac{1}{2}\frac{\varepsilon_t^2}{\sigma_t^2} + \text{constants} \tag{4.9}$$

The constants will have no effect on later results, thus redefine

$$\begin{aligned} f\left(0, \sigma_t^2\right) &= \sum_{t=1}^{T} -\frac{1}{2}\ln\sigma_t^2 - \frac{1}{2}\frac{\varepsilon_t^2}{\sigma_t^2} \\ &= \sum_{t=1}^{T} l_t(\boldsymbol{\theta}) \end{aligned} \tag{4.10}$$

where $l_t(\boldsymbol{\theta})$ is the likelihood function of observation t.

Differentiating $l_t(\boldsymbol{\theta})$ with respect to the variance parameters yields

$$\begin{aligned} \frac{\partial l_t}{\partial \omega} &= -\frac{1}{2}\sigma_t^{-2}\frac{\partial\sigma_t^2}{\partial\omega} + \frac{1}{2}\varepsilon_t^2\left(\sigma_t^2\right)^{-2}\frac{\partial\sigma_t^2}{\partial\omega} \\ &= \frac{1}{2\sigma_t^2}\frac{\partial\sigma_t^2}{\partial\omega}\left(\frac{\varepsilon_t^2}{\sigma_t^2} - 1\right) \end{aligned}$$

the second derivative

$$\frac{\partial l_t}{\partial\omega\partial\omega'} = \left(\frac{\varepsilon_t^2}{\sigma_t^2} - 1\right)\frac{\partial\sigma_t^2}{\partial\omega}\left[\frac{1}{2\sigma_t^2}\frac{\partial\sigma_t^2}{\partial\omega}\right] - \frac{1}{2\left(\sigma_t^2\right)^2}\frac{\partial\sigma_t^2}{\partial\omega}\frac{\partial\sigma_t^2}{\partial\omega'}\frac{\varepsilon_t^2}{\sigma_t^2}$$

where

$$\frac{\partial \sigma_t^2}{\partial \omega} = z_t + \sum_{i=1}^{p} \beta_i \frac{\partial h_{t-i}}{\partial \omega}$$

Differentiating $l_t(\boldsymbol{\theta})$ with respect to the mean parameters yields

$$\frac{\partial l_t}{\partial b} = \frac{\varepsilon_t x_t}{\sigma_t^2} + \frac{1}{2(\sigma_t^2)^2} \frac{\partial \sigma_t^2}{\partial b} \left(\frac{\varepsilon_t^2}{\sigma_t^2} - 1 \right)$$

the second derivative

$$\begin{aligned} \frac{\partial l_t}{\partial b \partial b'} &= -\frac{1}{\sigma_t^2} x_t x_t' - \frac{1}{2} \frac{1}{(\sigma_t^2)^2} \frac{\partial \sigma_t^2}{\partial b} \frac{\partial \sigma_t^2}{\partial b'} \left(\frac{\varepsilon_t^2}{\sigma_t^2} \right) \\ & \quad -2 \frac{1}{(\sigma_t^2)^2} \varepsilon_t x_t \frac{\partial \sigma_t^2}{\partial b} + \left(\frac{\varepsilon_t^2}{\sigma_t^2} - 1 \right) \frac{\partial}{\partial b'} \left[\frac{1}{2\sigma_t^2} \frac{\partial \sigma_t^2}{\partial b} \right] \end{aligned}$$

where

$$\frac{\partial \sigma_t^2}{\partial b} = -2 \sum_{j=1}^{q} \alpha_j x_{t-j} \varepsilon_{t-j} + \sum_{j=1}^{q} \beta_j \frac{\partial \sigma_{t-j}^2}{\partial b}$$

4.5 Integrated GARCH

The Integrated GARCH or I-GARCH process is defined as the standard $GARCH(p,q)$ process defined in equation 4.4 where $\alpha_1 + \beta_1 = 1$, thus if we put $\beta_1 = \lambda$ then

$$\sigma_t^2 = \alpha_0 + (1-\lambda)\varepsilon_{t-1}^2 + \lambda \sigma_{t-1}^2$$

where $\varepsilon_t \mid \mathcal{F}_{t-1} \sim N(0, \sigma_t^2)$ and clearly $\lambda \in [0,1]$.

From the stationary variance of the $GARCH(1,1)$ process defined in equation 4.5, it's clear that the stationary variance of the I-GARCH process doesn't exists. I-GARCH processes are often encountered in foreign exchange and commodity markets.

When the constant term $\alpha_0 = 0$ then the I-GARCH process is an EWMA process.

The I-GARCH process can however by strictly stationary, this result follows from Nelson (see [18]). For the $GARCH(1,1)$ process

$$\begin{aligned} \sigma_t^2 &= \alpha_0 + \alpha \varepsilon_{t-1}^2 + \beta \sigma_{t-1}^2 \\ &= \alpha_0 + \alpha \varepsilon_{t-1}^* \sigma_{t-1}^2 + \beta \sigma_{t-1}^2 \end{aligned}$$

where $\varepsilon_t^* \mid \mathcal{F}_{t-1} \sim N(0,1)$. Further

$$\begin{aligned}
\sigma_t^2 &= \alpha_0 + \left(\alpha\varepsilon_{t-1}^* + \beta\right)\sigma_{t-1}^2 \\
&= \alpha_0 + \left(\alpha\varepsilon_{t-1}^* + \beta\right)\left(\alpha_0 + \left(\alpha\varepsilon_{t-2}^* + \beta\right)\sigma_{t-2}^2\right) \\
&= \alpha_0\left(1 + \left(\alpha\varepsilon_{t-1}^* + \beta\right)\right) + \left(\alpha\varepsilon_{t-1}^* + \beta\right)\left(\alpha\varepsilon_{t-2}^* + \beta\right)\sigma_{t-2}^2 \\
&\vdots \\
&= \alpha_0 + \left(1 + \sum_{i=1}^{t-1}\prod_{j=1}^{i}\left(\alpha\varepsilon_{t-j}^* + \beta\right)\right) + \prod_{j=1}^{i}\left(\alpha\varepsilon_{t-i}^* + \beta\right)\sigma_0^2
\end{aligned}$$

where σ_0^2 is the first conditional variance. Nelson proved that the process is strictly stationary if

$$E\left[\ln\left(\alpha\varepsilon_{t-i}^* + \beta\right)\right] < 1$$

for every applicable i.

4.6 GARCH-in-Mean

The ARCH-in-Mean (GARCH-M) process was introduced by Engle, Lilien & Robins in 1987. In this process the connection between returns and risk, represented by AR and GARCH processes respectively, is set. Risk averse investors are expected to demand higher returns on risky assets than on less risky ones. The GARCH process in this model is therefore fixed to a risk premium. This risk premium can be seen as the positive correlation between current return and conditional covariance.

An example of an GARCH-M process is

$$y_t = \phi_0 + \phi_1 y_{t-1} + \ldots + \phi_p y_{t-p} + g\left(\sigma_t, \lambda\right) + \varepsilon_t \tag{4.11}$$

where the ϕ-parameters are AR parameters and g is a function of a GARCH process, σ_t and the risk premia, λ. The function is mostly taken as the identity or square root function of σ_t multiplied with λ.

The GARCH-M process by Duan, discussed in chapter 6, is

$$S_t = S_{t-1}\exp\left(r\Delta t - \frac{1}{2}\sigma_t^2 + \lambda\sigma_t + \varepsilon_t\right) \tag{4.12}$$

or

$$\ln\frac{S_t}{S_{t-1}} = r\Delta t - \frac{1}{2}\sigma_t^2 + \lambda\sigma_t + \varepsilon_t$$

where, for an annual risk-free rate r and daily volatility measurements t, $\Delta t = 1/252$, since we assume 252 trading days in a year.

GARCH-M process can be extended by any other GARCH process.

4.7 Asymmetric GARCH and the Leverage Effect

The leverage effect was reviewed in section 2.7.4. The jest of the leverage effect is: markets tend to react more volatile to negative information than to positive information. Symmetric GARCH processes react equally to positive and negative news.

Asymmetric GARCH processes have an extra parameter, denoted by γ in this dissertation, that skew returns information to market reaction. Here follow a few Asymmetric GARCH processes:

4.7.1 Exponential GARCH

The Exponential GARCH (EGARCH) was introduced by Nelson (1991). The EGARCH process is given by

$$\ln \sigma_t^2 = \alpha_0 + \beta_1 \ln \sigma_{t-1}^2 + \beta_2 \left(|\varepsilon_{t-1}| - \gamma \varepsilon_{t-1} \right)$$

where $\beta_2, \gamma > 0$.

The upside of EGARCH is that it generally fits empirical financial data well, but the downside is that EGARCH has no analytic form for its term structure.

4.7.2 Asymmetric GARCH

The Asymmetric GARCH (AGARCH) process is by Engle and Ng (1993). The AGARCH process is as follows

$$\sigma_t^2 = \alpha_0 + \alpha \left(\varepsilon_{t-1} - \gamma \right)^2 + \beta \sigma_{t-1}^2$$

where $\alpha_0 > 0$ and $\alpha, \beta, \gamma \geq 0$.

The parameters of the AGARCH process is easier to estimate than that of the EGARCH process, and it possesses an analytical term structure.

4.7.3 Glosten, Jagannathan and Runkle GARCH

The Glosten, Jagannathan and Runkle GARCH (GJR) process (1993), is named after its founders. The process is

$$\sigma_t^2 = \alpha_0 + \beta \sigma_{t-1}^2 + \alpha \varepsilon_{t-1}^2 + \gamma \max \left(-\varepsilon_t, 0 \right)^2$$

where $\gamma > 0$.

4.8 Limitations of the GARCH Process

The GARCH processes have the following limitations:

1. The GARCH processes perform best under stable market conditions. This process often fails to capture highly unexpected shocks, like market crashes. Except for the direct effect of a sudden shock, it may also cause structural changes in the market.

2. It's often hard to decide which GARCH process fits empirical data the best. There is no single GARCH process that can adequately model all conditional volatility processes. The conditional volatility structure of underlying assets also occasionally changes, which necessitates the using a different process.

3. The GARCH processes presented here depends on normal innovations. These processes often fail to fully capture the heavy tails observed in return series. Student's t-distribution and distributions like the Normal Inverse Gaussian distribution are often used as sources of innovation.

4. Investment decisions mustn't be solely based on the results of the GARCH processes. Other sources of information and models must also be used to make such decisions.