Anders Rahbek
Rasmus Søndergaard Pedersen
University of Copenhagen

October 2024

Part VII
Models for time-varying volatility and their applications

This chapter provides an overview of ARCH and Generalized ARCH (GARCH) models. We state conditions for stationarity and ergodicity of ARCH and GARCH processes, relying on the results for SREs considered in Chapter V, and emphasize that similar conditions can be derived by relying on the drift criterion considered in Chapter I. Moreover, we present limit theory for the ML (and quasi-ML) estimators for ARCH(1) models based on the theory for extremum estimators considered in Chapter VI. As emphasized in Chapter VI, the limit results for the estimators can be derived using either limit theorems for stationary and ergodic processes or for geometrically ergodic Markov chains. As possible extensions of the models, we consider non-linear models, multivariate (MGARCH) models as well as alternative error distributions. We conclude the chapter by presenting two applications of the models in relation to portfolio choice (Section VII.4) and risk quantification (Section VII.5). We emphasize that ARCH models and their extensions have several other applications within finance such as option pricing (Chorro et al., 2015) and empirical asset pricing (e.g., Engle, 2016 and Blasques et al., 2024).

VII.1 ARCH (1)

Consider the ARCH(1) process given by

$$x_t = \sigma_t z_t, \quad t \in \mathbb{Z}, \tag{VII.1}$$

$$\sigma_t^2 = \omega + \alpha x_{t-1}^2, \quad \omega > 0, \alpha \geq 0, \tag{VII.2}$$

where

$$\{z_t\}_{t\in\mathbb{Z}} \text{ is an i.i.d. process with } z_t \overset{D}{=} N(0,1), \tag{VII.3}$$

and z_t and x_{t-1} are independent for all t. From Chapter I we have that (almost surely)

$$\mathbb{E}[x_t|x_{t-1}] = 0 \quad \text{and} \quad \mathbb{V}[x_t|x_{t-1}] = \mathbb{E}[x_t^2|x_{t-1}] = \sigma_t^2,$$

that is, the process has zero conditional mean for all t but time-varying conditional variance. In particular, since $z_t \overset{D}{=} N(0,1)$ and σ_t are independent

$$x_t|x_{t-1} \overset{D}{=} N(0, \sigma_t^2),$$

such that x_t has conditional density

$$f(x_t|x_{t-1}) = \frac{1}{\sqrt{2\pi\sigma_t^2}} \exp\left\{-\frac{x_t^2}{2\sigma_t^2}\right\}.$$

Hence the *conditional* distribution of x_t is Gaussian, while as you may recall from Example V.4.6 in Chapter V, under certain assumptions, the unconditional distribution of x_t has so-called power law tails, meaning that some moments of the distribution are infinite. The remainder of this section summarizes the probabilistic properties of the ARCH(1) found in Chapter V and states the properties of the ML estimator for the model parameters found in Chapter VI. Lastly, we consider the notion of quasi-maximum likelihood estimation.

VII.1.1 Stochastic properties

From Examples V.4.2 and V.4.4 in Chapter V we have the following result.

Corollary VII.1.1 *Consider the ARCH(1) process $\{x_t\}_{t\in\mathbb{Z}}$ given by (VII.1)-(VII.3). If $\alpha < 3.56\ldots$, then $\{x_t\}_{t\in\mathbb{Z}}$ is a stationary and ergodic process. If in addition $\alpha > 0$, then*

$$\mathbb{E}[|x_t|^p] < \infty \quad \textit{if and only if} \quad \mathbb{E}[|\sqrt{\alpha} z_t|^p] < 1,$$

for $p \in [0,\infty)$. If $\alpha = 0$, $\mathbb{E}[|x_t|^p] = \omega^{p/2}\mathbb{E}[|z_t|^p] < \infty$ for all $p \in [0,\infty)$.

The corollary allows us to derive conditions for finite moments and well-defined autocovariances of x_t, as considered in the next example.

Example VII.1.1 *Based on Corollary VII.1.1, if $\alpha < 1$ then $\mathbb{E}[x_t^2] < \infty$ and hence $\mathbb{E}[\sigma_t^2] < \infty$. We then have that*

$$\mathbb{E}[x_t^2] = \mathbb{E}[\sigma_t^2 z_t^2] = \mathbb{E}[\sigma_t^2]\mathbb{E}[z_t^2] = \mathbb{E}[\sigma_t^2] = \mathbb{E}[\omega + \alpha x_{t-1}^2] = \omega + \alpha\mathbb{E}[x_{t-1}^2],$$

where we have used that z_t and σ_t are independent. By stationarity, $\mathbb{E}[x_t^2] = \mathbb{E}[x_{t-1}^2]$, such that

$$\mathbb{E}[x_t^2] = \frac{\omega}{1-\alpha}.$$

In this case, we also have that for any $k > 0$,

$$\mathbb{E}[x_t x_{t-k}] = \mathbb{E}[\mathbb{E}[x_t|x_{t-k}]x_{t-k}] = 0,$$

such that the autocovariance function of x_t is zero at any lag order $k > 0$. Likewise if $\alpha < 1/\sqrt{3}$, then $\mathbb{E}[x_t^4] < \infty$. It holds that

$$\mathbb{E}[x_t^4] = \mathbb{E}[\sigma_t^4 z_t^4] = \mathbb{E}[\sigma_t^4]\mathbb{E}[z_t^4] = 3\mathbb{E}[\sigma_t^4],$$

with

$$\begin{aligned}\mathbb{E}[\sigma_t^4] &= \mathbb{E}[(\omega + \alpha x_{t-1}^2)^2] \\ &= \mathbb{E}[\omega^2 + \alpha^2 x_{t-4}^4 + 2\omega\alpha x_{t-1}^2] \\ &= \omega^2 + \alpha^2 \mathbb{E}[x_t^4] + 2\omega\alpha\mathbb{E}[x_t^2] \\ &= \omega^2 + \alpha^2\mathbb{E}[x_t^4] + \frac{2\omega^2\alpha}{1-\alpha}.\end{aligned}$$

Hence,

$$\mathbb{E}[x_t^4] = 3\left(\omega^2 + \alpha^2\mathbb{E}[x_t^4] + \frac{2\omega^2\alpha}{1-\alpha}\right),$$

such that

$$\mathbb{E}[x_t^4] = 3\frac{\omega^2 + \frac{2\omega^2\alpha}{1-\alpha}}{1-3\alpha^2} = 3\left(\frac{1-\alpha^2}{1-3\alpha^2}\right)\left(\frac{\omega}{1-\alpha}\right)^2.$$

Moreover, under the same assumption,

$$\begin{aligned}\mathbb{E}[x_t^2 x_{t-1}^2] &= \mathbb{E}[z_t^2\sigma_t^2 x_{t-1}^2] \\ &= \mathbb{E}[\sigma_t^2 x_{t-1}^2] \\ &= \mathbb{E}[(\omega + \alpha x_{t-1}^2)x_{t-1}^2] \\ &= \omega\mathbb{E}[x_t^2] + \alpha\mathbb{E}[x_t^4] \\ &= \frac{\omega^2}{1-\alpha} + 3\alpha\left(\frac{1-\alpha^2}{1-3\alpha^2}\right)\left(\frac{\omega}{1-\alpha}\right)^2 \\ &= \left[3\alpha\left(\frac{1-\alpha^2}{1-3\alpha^2}\right) + (1-\alpha)\right]\left(\frac{\omega}{1-\alpha}\right)^2,\end{aligned}$$

such that autocovariance of x_t^2 of order one is given by

$$\begin{aligned}Cov(x_t^2, x_{t-1}^2) &= \mathbb{E}[x_t^2 x_{t-1}^2] - \mathbb{E}[x_t^2]\mathbb{E}[x_{t-1}^2] \\ &= \left[3\alpha\left(\frac{1-\alpha^2}{1-3\alpha^2}\right) + (1-\alpha)\right]\left(\frac{\omega}{1-\alpha}\right)^2 - \left(\frac{\omega}{1-\alpha}\right)^2 \\ &= \left[3\alpha\left(\frac{1-\alpha^2}{1-3\alpha^2}\right) - \alpha\right]\left(\frac{\omega}{1-\alpha}\right)^2 \\ &= 2\alpha\left(\frac{\omega}{1-\alpha}\right)^2,\end{aligned}$$

and the autocorrelation of x_t^2 of order one is

$$\begin{aligned} Corr(x_t^2, x_{t-1}^2) &= \frac{Cov(x_t^2, x_{t-1}^2)}{\sqrt{\mathbb{E}[x_t^4]\mathbb{E}[x_{t-1}^4]}} \\ &= \frac{Cov(x_t^2, x_{t-1}^2)}{\mathbb{E}[x_t^4]} \\ &= \frac{2\alpha}{3}\left(\frac{1-3\alpha^2}{1-\alpha^2}\right) \geq 0. \end{aligned}$$

VII.1.2 ML Estimation

In applications, such as the ones described in Sections VII.4 and VII.5 below, one typically needs estimates of the ARCH model parameters. The estimation is typically carried out by ML, as detailed in this section. Given a sample $\{x_t\}_{t=0}^T$ generated by the ARCH(1) process in (VII.1)-(VII.3), we seek to estimate the parameter vector $\theta = (\omega, \alpha)'$. Recall from Chapter VI that the log-likelihood function is given by

$$\begin{aligned} Q_T(\theta) &= \frac{1}{T}\sum_{t=1}^T q_t(\theta), \qquad \text{(VII.4)} \\ q_t(\theta) &= \log f_\theta(x_t|x_{t-1}), \\ f_\theta(x_t|x_{t-1}) &= \frac{1}{\sqrt{2\pi\sigma_t^2(\theta)}}\exp\left\{-\frac{x_t^2}{2\sigma_t^2(\theta)}\right\}, \\ \sigma_t^2(\theta) &= \omega + \alpha x_{t-1}^2. \end{aligned}$$

We consider estimation over the compact parameter space

$$\Theta = [\omega_L, \omega_U] \times [0, \alpha_U] \qquad \text{(VII.5)}$$

with $0 < \omega_L < \omega_U < \infty$ and $0 < \alpha_U < \infty$, and let $\hat{\theta}_T$ denote the ML estimator given by

$$\hat{\theta}_T = \arg\max_{\theta\in\Theta} Q_T(\theta), \qquad \text{(VII.6)}$$

with $Q_T(\theta)$ given by (VII.4) and Θ given by (VII.5). Moreover, the true value of θ (that is, the value of θ for the DGP) is labelled θ_0. From Examples VI.2.5 and VI.4.2 and Section VI.3.1 we have the following result.

Corollary VII.1.2 *Let $\{x_t\}_{t=0}^T$ follow the ARCH(1) process in (VII.1)-(VII.3) with true value $\theta_0 = (\omega_0, \alpha_0)' \in \Theta$ and $\alpha_0 < 1$. Then the ML estimator $\hat{\theta}_T$ in (VII.6) is consistent for θ_0, that is*

$$\hat{\theta}_T \xrightarrow{p} \theta_0 \quad \text{as } T \to \infty.$$

Suppose in addition that θ_0 lies in the interior of Θ. Then

$$\sqrt{T}(\hat{\theta}_T - \theta_0) \xrightarrow{D} N(0, -\Sigma_0^{-1}),$$

with $-\Sigma_0^{-1}$ a positive definite matrix with

$$\Sigma_0 = \mathbb{E}\left[\frac{\partial^2 q_t(\theta_0)}{\partial\theta\partial\theta'}\right]. \tag{VII.7}$$

A consistent estimator for Σ_0 is given by

$$\hat{\Sigma}_T = \frac{1}{T}\sum_{t=1}^{T}\frac{\partial^2 q_t(\hat{\theta}_T)}{\partial\theta\partial\theta'}.$$

VII.1.3 Quasi-maximum likelihood

Often in applied work, upon estimation, the ARCH(1) model (or its extension as considered below) is found to have non-Gaussian innovations, that is, estimated innovations $\hat{z}_t = x_t/\sigma_t(\hat{\theta}_T)$ do not appear (approximately) standard normal as the theory would suggest; see e.g. Tsay (2010, Chapter 3) and Mikosch and Starica (2000, Section 6) for empirical examples. Fortunately, it shows up that reliable estimation and inference can still be done using the estimator $\hat{\theta}_T$ given by (VII.6). Suppose that the DGP is given by the ARCH(1) process in (VII.1)-(VII.2) under the assumption that

$$\{z_t\}_{t\in\mathbb{Z}} \text{ is an i.i.d. process with } \mathbb{E}[z_t] = 0 \text{ and } \mathbb{E}[z_t^2] = 1. \tag{VII.8}$$

Then x_t^2 satisfies the SRE

$$x_t^2 = \alpha z_t^2 x_{t-1}^2 + \omega z_t^2,$$

and $\{x_t^2\}_{t\in\mathbb{Z}}$ (and hence $\{x_t\}_{t\in\mathbb{Z}}$; see Exercises) is strictly stationary and ergodic provided that $\mathbb{E}[\log(\alpha z_t^2)] < 0$ (Theorem V.4.1). Moreover, from Theorem V.4.2 $\mathbb{E}[x_t^2] < \infty$ if $\mathbb{P}(z_t^2 = 1) < 1$ and $\alpha < 1$. In terms of the estimator $\hat{\theta}_T$ we have the following result that follows by arguments given in Example VI.2.5 and Section VI.3.1.

Corollary VII.1.3 *Suppose that the DGP $\{x_t\}_{t\in\mathbb{Z}}$ is given by the ARCH(1) process in (VII.1)-(VII.2) and (VII.8) with true value $\theta_0 = (\omega_0, \alpha_0)'$ and $\alpha_0 < 1$. Assume that $\theta_0 \in \Theta$ with Θ given by (VII.5) and $\mathbb{P}(z_t^2 = 1) < 1$. Then the estimator $\hat{\theta}_T$ given by (VII.6) satisfies*

$$\hat{\theta}_T \xrightarrow{p} \theta_0, \quad \textit{as } T \to \infty.$$

Suppose in addition that $\mathbb{E}[z_t^4] < \infty$ *and* θ_0 *lies in the interior of* Θ*. Then*

$$\sqrt{T}(\hat{\theta}_T - \theta_0) \xrightarrow{D} N(0, \Sigma_0^{-1}\Omega_0\Sigma_0^{-1}),$$

with Σ_0 *invertible given by (VII.7) and* Ω_0 *positive definite given by*

$$\Omega_0 = \mathbb{E}\left[\frac{\partial q_t(\theta_0)}{\partial \theta}\frac{\partial q_t(\theta_0)}{\partial \theta'}\right].$$

The above result is powerful in the sense that the estimator $\hat{\theta}_T$ is useful (under mild conditions) even if one is not willing to specify a particular distribution for z_t. In this setting the objective function $Q_T(\theta)$ in (VII.4) is not necessarily the true log-likelihood function, and $Q_T(\theta)$ is referred to as the *quasi*-log-likelihood function, and $\hat{\theta}_T$ is the quasi-ML (QML) estimator. The estimator has another limiting variance ($\Sigma_0^{-1}\Omega_0\Sigma_0^{-1}$) than the ML estimator ($-\Sigma_0^{-1}$). It can be shown that $\Omega_0 = \mathbb{E}[(z_t^4 - 1)/2](-\Sigma_0)$, such that $\Omega_0 = -\Sigma_0$ in the case where $z_t \overset{D}{=} N(0,1)$. The matrix Σ_0 can be estimated by $\hat{\Sigma}_T$ provided in Corollary VII.1.2, whereas the quantity $\mathbb{E}[(z_t^4 - 1)/2]$ may be estimated from the standardized residuals, that is, by $T^{-1}\sum_{t=1}^{T}(\hat{z}_t^4 - 1)/2$.

Note that asymptotic normality of the QML estimator relies on the assumption that $\mathbb{E}[z_t^4] < \infty$, and the asymptotic covariance matrix is increasing (entry-wise) in $\mathbb{E}[z_t^4]$. Hence, the QML estimator is potentially imprecise when $\mathbb{E}[z_t^4]$ is large. Moreover, in the case where $\mathbb{E}[z_t^4] = \infty$, the rate of convergence of the QML estimator is slower than $\sqrt{T}$ and the limiting distribution is given in terms of a random vector with a non-Gaussian stable distribution; see Mikosch and Straumann (2006) for technical details and precise assumptions. Consequently, if one expects (or the estimation results suggests) that z_t is heavy-tailed, in the sense that $\mathbb{E}[z_t^4]$ is very large or infinite, it might be desirable to consider a model that directly addresses this feature, such as the model with (scaled) Student's t innovations described in Section VII.2.4 below.

We emphasize that the above considerations about quasi-ML estimation based on the Gaussian pdf applies to any extension, including the multivariate setting, of the ARCH model considered in the following sections.

VII.2 Extensions

The ARCH(1) model can be extended in multiple ways. A general structure is of the form

$$x_t = \mu_t + \sigma_t z_t, \quad t \in \mathbb{Z}, \tag{VII.9}$$

with $\{z_t\}_{t\in\mathbb{Z}}$ an i.i.d. process with $\mathbb{E}[z_t] = 0$ and $\mathbb{E}[z_t^2] = 1$, and z_t independent of $\mathcal{F}_{t-1}$ for all t. Here $\mu_t, \sigma_t \in \mathcal{F}_{t-1}$ with $\mathcal{F}_t$ the natural filtration generated by $\{x_s\}_{s\leq t}$ and potentially other (observable) processes. The conditional mean μ_t may be constant or contain lagged values of x_t, e.g. $\mu_t = \delta + \rho x_{t-1}$ for constants δ and ρ. Likewise, the conditional variance σ_t^2 could include several lags of x_t (ARCH(q)), lags of σ_t^2 (GARCH), non-linear terms (e.g., GJR-GARCH), or explanatory covariates (GARCH-X), all considered in the following sections. For most practical purposes a minimal requirement is that $\mathbb{P}(\sigma_t^2 > 0) = 1$ for all t. We refer to Bollerslev's (2009) "Glossary to ARCH (GARCH" for an overview of potential specifications for σ_t^2.

VII.2.1 ARCH(q)

Similar to the AR(k) extension of the AR(1) model in Chapter II, we have ARCH(q) models of the form

$$
\begin{aligned}
x_t &= \sigma_t z_t, \quad t \in \mathbb{Z}, \\
\sigma_t^2 &= \omega + \alpha_1 x_{t-1}^2 + \cdots + \alpha_q x_{t-q}^2,
\end{aligned}
$$

with the innovations z_t satisfying (VII.3), and $\omega > 0$, $\alpha_1, \ldots, \alpha_q \geq 0$. With $X_t := (x_t^2, x_{t-1}^2, \ldots, x_{t-q+1}^2)'$ we have the SRE

$$
X_t = A_t X_{t-1} + B_t,
$$

with

$$
A_t = \begin{pmatrix} z_t^2\alpha_1 & z_t^2\alpha_2 & \cdots & \cdots & z_t^2\alpha_q \\ 1 & 0 & 0 & \cdots & 0 \\ 0 & 1 & 0 & \cdots & 0 \\ \vdots & & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & 1 & 0 \end{pmatrix}, \quad B_t = \begin{pmatrix} z_t^2\omega \\ 0 \\ \vdots \\ 0 \end{pmatrix},
$$

and conditions for stationarity and ergodicity can be derived from Theorem V.5.1. Estimation can be carried out by means of ML or quasi-ML as above. The motivation for letting $q > 1$ is to have a richer model that provides a better specification of the conditional variance of x_t and that takes into account the fact that squared (or absolute) returns are typically found to have a large order of non-zero autocorrelation; see, e.g., the recent work by Nielsen and Rahbek (2024) that consider ARCH models with $q > 100$.

VII.2.2 GARCH

The most prominent extension of the ARCH(1) model is the Generalized ARCH (GARCH) of Bollerslev (1986) and Taylor (1986) given by

$$\begin{aligned} x_t &= \sigma_t z_t, \quad t \in \mathbb{Z}, \\ \sigma_t^2 &= \omega + \alpha x_{t-1}^2 + \beta \sigma_{t-1}^2, \end{aligned}$$

with the innovations z_t satisfying (VII.3), $\omega > 0$ and $\alpha, \beta \geq 0$. The GARCH model is the workhorse model in classical volatility modelling and is widely used in empirical work.

The conditional variance σ_t^2 obeys the SRE

$$\sigma_t^2 = \omega + (\alpha z_{t-1}^2 + \beta)\sigma_{t-1}^2,$$

such that $\{\sigma_t^2\}_{t\in\mathbb{Z}}$ is stationary and ergodic if

$$\mathbb{E}[\log(\alpha z_t^2 + \beta)] < 0.$$

To conclude that $\{x_t\}_{t\in\mathbb{Z}}$ is stationary and ergodic under the same condition, note that

$$\begin{pmatrix} z_{t-1} \\ \sigma_t^2 \end{pmatrix} = \underbrace{\begin{pmatrix} 0 & 0 \\ 0 & \alpha z_{t-1}^2 + \beta \end{pmatrix}}_{A_t} \begin{pmatrix} z_{t-2} \\ \sigma_{t-1}^2 \end{pmatrix} + \begin{pmatrix} z_{t-1} \\ \omega \end{pmatrix},$$

is an SRE. By Theorem V.5.1 the process $\{(z_{t-1}, \sigma_t^2)'\}_{t\in\mathbb{Z}}$ is strictly stationary and ergodic, if the top Lyapunov exponent, γ, associated with the SRE is negative, which is immediate, as

$$\begin{aligned} \gamma &= \inf_{t\geq 1} \frac{1}{t}\mathbb{E}[\log \| \prod_{i=1}^{t} A_i \|] \\ &= \inf_{t\geq 1} \frac{1}{t}\mathbb{E}\left[\log \left\| \prod_{i=1}^{t} \begin{pmatrix} 0 & 0 \\ 0 & \alpha z_i^2 + \beta \end{pmatrix} \right\| \right] \\ &= \inf_{t\geq 1} \frac{1}{t}\mathbb{E}\left[\log \left\| \begin{pmatrix} 0 & 0 \\ 0 & \prod_{i=1}^{t}(\alpha z_i^2 + \beta) \end{pmatrix} \right\| \right] \\ &= \inf_{t\geq 1} \frac{1}{t}\mathbb{E}[\log \prod_{i=1}^{t}(\alpha z_i^2 + \beta)] \\ &= \inf_{t\geq 1} \frac{1}{t} \sum_{t=1}^{t} \mathbb{E}[\log(\alpha z_t^2 + \beta)] \\ &= \mathbb{E}[\log(\alpha z_t^2 + \beta)] < 0. \end{aligned}$$

Since, $x_t = z_t\sqrt{\sigma_t^2}$ is a measurable function of $\{(z_{t-1}, \sigma_t^2)'\}_{t\in\mathbb{Z}}$ with $P(|x_t| < \infty) = 1$, we have by Theorem V.2.1 that $\{x_t\}_{t\in\mathbb{Z}}$ is stationary and ergodic. Theorem V.4.2 can be used to find conditions for finite moments of x_t. For instance $\mathbb{E}[\sigma_t^2] = \mathbb{E}[x_t^2] < \infty$ if and only if $\alpha + \beta < 1$.

Maximum likelihood estimation is more involved than for ARCH(1) as σ_t^2 is unobservable even if the true values of the model parameters were known. To see this, under the stationarity condition above,

$$\sigma_t^2 = \sum_{i=0}^{\infty} \beta^i(\omega + \alpha x_{t-1-i}^2),$$

such that the conditional variance σ_t^2 depends on returns from the infinite past. In practice, one only has the observations $\{x_t\}_{t=0}^T$ at hand. Consequently, it is customary to let the initial conditional variance for the log-likelihood function be fixed, that is $\sigma_0^2(\theta) := c > 0$ is constant. Then the log-likelihood function $Q_T(\theta)$ is given as in (VII.4) with

$$\sigma_t^2(\theta) = \omega + \alpha x_{t-1}^2 + \beta\sigma_{t-1}^2(\theta), \quad t = 1, \ldots, T.$$

The asymptotic theory for the ML estimator becomes additionally tedious due to (i) the introduction of the fixed initial value c in the log-likelihood function, and (ii) the recursive structure of $\sigma_t^2(\theta)$. To see the latter, note that the first derivative (say, with respect to α) of the log-likelihood contribution involves the derivative

$$\frac{\partial\sigma_t^2(\theta)}{\partial\theta} = x_{t-1}^2 + \beta\frac{\partial\sigma_{t-1}^2(\theta)}{\partial\theta}, \quad t \geq 0.$$

We refer to Francq and Zakoïan (2019, Chapter 7) for detailed arguments.

VII.2.3 Nonlinear models and explanatory covariates

Another strand of extensions of the ARCH and GARCH models consider non-linear dynamics for the conditional variance σ_t^2. The most prominent example is the Glosten-Jagannathan-Runkle (GJR) GARCH model given by

$$\begin{aligned} x_t &= \sigma_t z_t, \quad t \in \mathbb{Z}, \\ \sigma_t^2 &= \omega + \alpha x_{t-1}^2 + \gamma\mathbb{I}(x_{t-1} < 0)x_{t-1}^2 + \beta\sigma_{t-1}^2, \end{aligned}$$

with $\omega > 0$, $\alpha, \gamma, \beta \geq 0$, and

$$\mathbb{I}(x_{t-1} < 0) = \begin{cases} 1 & \text{if } x_{t-1} < 0, \\ 0 & \text{if } x_{t-1} \geq 0. \end{cases}$$

The term $\gamma\mathbb{I}(x_{t-1} < 0)x_{t-1}^2$ allows for the possibility that a negative shock/return at time $t-1$ has a larger impact on the conditional variance at time t than a positive return, which is referred to as a so-called leverage effect; see Glosten et al. (1993, Section II.B) for further economic motivation.

Noting that $\mathbb{I}(x_{t-1} < 0) = \mathbb{I}(z_{t-1} < 0)$, we have that σ_t^2 obeys the SRE

$$\sigma_t^2 = \omega + [\alpha z_{t-1}^2 + \gamma\mathbb{I}(z_{t-1} < 0)z_{t-1}^2 + \beta]\sigma_{t-1}^2,$$

and conditions for stationarity, ergodicity and finite moments can be derived from Theorem V.4.1 and V.4.2.

A last notable extension of the univariate ARCH and GARCH models is the inclusion of explanatory covariates in the conditional variance equation,

$$\sigma_t^2 = \sigma_t^2 = \omega + \alpha x_{t-1}^2 + \beta\sigma_{t-1}^2 + \tau y_{t-1},$$

where $\tau \geq 0$ and y_{t-1} is a non-negative random variable. This class of GARCH models is referred to as GARCH-X. Examples of y_{t-1} include credit spreads, macroeconomic variables, unexpected shocks from other assets or markets, as well as alternative volatility measures. Since the dynamics of y_{t-1} are not directly specified, its inclusion in the model requires careful considerations about the dependence structure between y_{t-1} and (x_{t-1}, z_t). We refer to Han and Kristensen (2014), Francq and Thieu (2019) and Pedersen and Rahbek (2019) for many more details.

VII.2.4 Alternative distributions for z_t

For the ARCH(1) in (VII.1)-(VII.2), instead of having a standard normal distribution, z_t could be assumed to have a more heavy tailed distribution, which may for instance be desirable (and empirically relevant) in terms of risk quantification; see Section VII.5 below. Specifically, assume that $\{z_t\}_{t\in\mathbb{Z}}$ is i.i.d. with z_t scaled Student's t-distributed with $\nu > 2$ degrees of freedom, as considered in Example I.4.3 in Part I. It holds that $\mathbb{E}[z_t] = 0$ and $\mathbb{E}[z_t^2] = 1$, and z_t has pdf given by

$$f(x) = \frac{\Gamma\left(\frac{\nu+1}{2}\right)/\Gamma\left(\frac{\nu}{2}\right)}{\sqrt{(\nu-2)\pi}}\left(1 + \frac{x^2}{(\nu-2)}\right)^{-\left(\frac{\nu+1}{2}\right)}, \quad x \in \mathbb{R}.$$

In this case, consider the SRE for x_t^2,

$$x_t^2 = \sigma_t^2 z_t^2 = \underbrace{\alpha z_t^2}_{A_t} x_{t-1}^2 + \underbrace{\alpha z_t^2}_{B_t},$$

and conditions stationarity and ergodicity are derived along the arguments in Section VII.1.3. In particular, $\{x_t\}_{t\in\mathbb{Z}}$ is stationary and ergodic, if $\mathbb{E}[\log(A_t)] < \infty$, or equivalently

$$\alpha < \exp\left(-\mathbb{E}[\log(z_t^2)]\right).$$

Here the quantity

$$\mathbb{E}[\log(z_t^2)] = \frac{\Gamma\left(\frac{\nu+1}{2}\right)/\Gamma\left(\frac{\nu}{2}\right)}{\sqrt{(\nu-2)\,\pi}} \int_{-\infty}^{\infty} \log(x^2)\left(1+\frac{x^2}{(\nu-2)}\right)^{-\left(\frac{\nu+1}{2}\right)} dx$$

depends on the degrees of freedom $\nu > 2$. For instance, $\mathbb{E}[\log(z_t^2)] = -2$ for $\nu = 3$, such that the stationarity condition becomes $\alpha < \exp(2) = 7.38\ldots$, which is milder than the condition stated in Corollary VII.1.1 for the case of $z_t \overset{D}{=} N(0,1)$.

Assuming that z_t is scaled Student's t-distributed, introduces the additional parameter ν that one would typically estimate, such that the parameter vector is given by $\theta = (\omega, \alpha, \nu)'$. From Example I.4.3, we note that x_t has conditional density given by

$$f(x_t|x_{t-1}) = \frac{\Gamma\left(\frac{\nu+1}{2}\right)/\Gamma\left(\frac{\nu}{2}\right)}{\sqrt{\sigma_t^2\,(\nu-2)\,\pi}} \left(1+\frac{x_t^2}{\sigma_t^2\,(\nu-2)}\right)^{-\left(\frac{\nu+1}{2}\right)},$$

and the log-likelihood function is given by

$$Q_T(\theta) = \frac{1}{T}\sum_{t=1}^{T} \log\left\{\frac{\Gamma\left(\frac{\nu+1}{2}\right)/\Gamma\left(\frac{\nu}{2}\right)}{\sqrt{(\omega+\alpha x_{t-1}^2)\,(\nu-2)\,\pi}} \left(1+\frac{x_t^2}{(\omega+\alpha x_{t-1}^2)\,(\nu-2)}\right)^{-\left(\frac{\nu+1}{2}\right)}\right\}.$$

The ML estimator is found by maximizing $Q_T(\theta)$ over the parameter space $\Theta \subset (0,\infty)\times[0,\infty)\times(2,\infty)$. Limit theory for the ML estimator may be derived from Theorems VI.2.2 and VI.3.1 in Chapter VI. We refer to Tsay (2010, Chapter 3) for alternative distributions for z_t, including the so-called skewed scaled Student's t-distribution that allows for the possibility that z_t has an asymmetric heavy-tailed distribution.

VII.3 Multivariate GARCH

In Chapter II we considered the VAR models as multivariate extensions of the AR models. Likewise, we consider here multivariate extensions of the ARCH and GARCH which we label Multivariate GARCH (MGARCH). Let $X_t = (X_{t,1},\ldots,X_{t,d})' \in \mathbb{R}^d$ be a vector that (for instance) contains returns of $d \geq 1$ different assets (e.g., different stocks, stock indices, currencies, commodities etc.). A simple MGARCH process – corresponding to a d-dimensional ARCH(1) – is given by

$$X_t = \Omega_t^{1/2} Z_t, \quad t \in \mathbb{Z}, \tag{VII.10}$$

$$\Omega_t^{1/2}(\Omega_t^{1/2})' = \Omega_t = g(X_{t-1}), \tag{VII.11}$$

for some measurable matrix function g, satisfying that $g(x)$ is positive definite for all $x \in \mathbb{R}^d$. Moreover

$$\{Z_t\}_{t\in\mathbb{Z}} \text{ is an } i.i.d. \text{ process with } Z_t \overset{D}{=} N(0, I_d), \tag{VII.12}$$

and Z_t and X_{t-1} are independent for all t. Parallel to the univariate ARCH(1), we have that

$$X_t|X_{t-1} \overset{D}{=} N(0, \Omega_t),$$

such that X_t has conditional density given by

$$f(X_t|X_{t-1}) = \frac{1}{\sqrt{(2\pi)^d \det(\Omega_t)}} \exp\left(-\frac{1}{2} X_t' \Omega_t^{-1} X_t\right).$$

The positive definiteness of Ω_t is (for most practical purposes) a minimal requirement for a covariance matrix, parallel to the positivity of σ_t required for univariate models. The matrix square-root $\Omega_t^{1/2}$ of Ω_t is not unique. For instance, $\Omega_t^{1/2}$ could be lower triangular stemming from a Cholesky decomposition of Ω_t, or $\Omega_t^{1/2}$ could be symmetric, obtained via an eigendecomposition of Ω_t. In the following sections we provide examples of specifications of Ω_t.

In terms of estimation, the conditional covariance matrix is assumed to be parametrized by a vector $\theta \in \mathbb{R}^k$ such that $\Omega_t(\theta) = g(X_{t-1}; \theta)$ with g known. The log-likelihood function is then given via the conditional density of X_t, and the properties of the ML estimator may be derived from the results i Chapter VI. Typically, due to the multivariate nature of MGARCH, derivations are more cumbersome compared to the univariate case. We refer to Francq and Zakoïan (2019, Chapter 10.4) for general (high-level) conditions ensuring consistency and asymptotic normality of the ML estimator for MGARCH models. Moreover, as will be clear from the following examples, the MGARCH models may contain many parameters such that k is large. In addition, the parametrization of the MGARCH model should ensure that Ω_t is positive definite. In practice these challenges imply that numerical maximization of the log-likelihood with respect to all parameters simultaneously is tedious (if not infeasible). This has given rise to alternative estimation methods, where estimation is carried out in multiple steps; see e.g. Noureldin et al. (2014), Pedersen and Rahbek (2014) and Francq and Zakoïan (2016).

In the following sections we present two classical MGARCH specifications, namely the so-called Baba-Engle-Kraft-Kroner (BEKK) and the constant conditional correlation (CCC) model. MGARCH models are widely used in empirical work, and we emphasize that several other relevant specifications exist, including dynamic conditional correlation (DCC) models (Engle, 2002), stochastic correlation models (Pelletier, 2006), orthogonal models (see Hetland et al., 2023, and the references therein) and so-called score-driven models (see D'Innocenzo and Lucas, 2024, and the references therein).

VII.3.1 BEKK

As already considered in Example V.3.4 in Chapter V, with $\{X_t\}_{t\in\mathbb{Z}}$ given by (VII.10)-(VII.12), Engle and Kroner (1995) considered the so-called BEKK specification for Ω_t given by

$$\Omega_t = \Omega + AX_{t-1}X_{t-1}'A',$$

for some positive definite matrix Ω and square matrix A. By construction, this specification ensures that Ω_t is positive definite.

Example VII.3.1 *Let $d = 2$ with*

$$\Omega = \begin{pmatrix} \Omega_{11} & \Omega_{12} \\ \Omega_{12} & \Omega_{22} \end{pmatrix}, \quad \text{and} \quad A = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix}.$$

Then the conditional covariance matrix is given by

$$\begin{aligned} \Omega_t &= \begin{pmatrix} \Omega_{t,11} & \Omega_{t,12} \\ \Omega_{t,12} & \Omega_{t,22} \end{pmatrix} \\ &= \begin{pmatrix} \Omega_{11} & \Omega_{12} \\ \Omega_{12} & \Omega_{22} \end{pmatrix} + \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix} \begin{pmatrix} X_{t,1}^2 & X_{t,1}X_{t,2} \\ X_{t,1}X_{t,2} & X_{t,1}^2 \end{pmatrix} \begin{pmatrix} A_{11} & A_{21} \\ A_{12} & A_{22} \end{pmatrix}. \end{aligned}$$

with

$$\begin{aligned} \Omega_{t,11} &= \Omega_{11} + A_{11}^2 X_{t,1}^2 + 2X_{t,2}A_{11}A_{12}X_{t,1} + A_{12}^2 X_{t,1}^2, \\ \Omega_{t,22} &= \Omega_{11} + A_{21}^2 X_{t,1}^2 + 2X_{t,2}A_{21}A_{22}X_{t,1} + A_{22}^2 X_{t,1}^2, \\ \Omega_{t,12} &= \Omega_{12} + A_{11}A_{21}X_{t,1}^2 + A_{12}A_{22}X_{t,1}^2 + A_{11}A_{22}X_{t,1}X_{t,2} + A_{12}A_{21}X_{t,1}X_{t,2}. \end{aligned}$$

Recall from Example V.3.4 that X_t has an SRE representation given by

$$X_t = m_t A X_{t-1} + B_t,$$

with $\{(m_t, B_t')\}_{t\in\mathbb{Z}}$ an i.i.d. process with $(m_t, B_t')'$ a $(d+1)$-dimensional random vector satisfying

$$\begin{pmatrix} m_t \\ B_t \end{pmatrix} \stackrel{D}{=} N\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ 0 & \Omega \end{bmatrix}\right).$$

From Example V.5.2 we have that $\{X_t\}_{t\in\mathbb{Z}}$ is strictly stationary and ergodic if

$$\rho(A) < \sqrt{3.56\ldots},$$

with $\rho(A)$ denoting the spectral radius of A. Furthermore, applications of the drift criterion (see Pedersen and Rahbek, 2014, Appendix C) give that

$$\begin{aligned}
\mathbb{E}[\|X_t\|^2] < \infty \quad & \text{if } \rho(A) < 1, \\
\mathbb{E}[\|X_t\|^4] < \infty \quad & \text{if } \rho(A) < 1/3^{1/4} = 0.75\ldots, \\
\mathbb{E}[\|X_t\|^6] < \infty \quad & \text{if } \rho(A) < 1/15^{1/6} = 0.63\ldots, \\
\mathbb{E}[\|X_t\|^8] < \infty \quad & \text{if } \rho(A) < 1/105^{1/8} = 0.55\ldots,
\end{aligned}$$

parallel to the results for the tail index for univariate ARCH(1) processes in Example V.4.6. We refer to Matsui and Pedersen (2022) for additional results for BEKK processes.

The BEKK process considered above can be extended to a GARCH version given by

$$\Omega_t = \Omega + AX_{t-1}X_{t-1}'A' + B\Omega_{t-1}B',$$

for a square matrix B. The properties of the resulting BEKK GARCH process $\{X_t\}_{t\in\mathbb{Z}}$ are complicated to derive, and the results for SREs are not directly applicable. The process $\{(X_t, \Omega_t)\}_{t\in\mathbb{Z}}$ is a Markov chain, and by carefully taking into account that Ω_t belongs to the space of positive definite matrices, Boussama et al. (2011) used an extended version of the drift criterion to prove that $\{X_t\}_{t\in\mathbb{Z}}$ is stationary and ergodic with $\mathbb{E}[\|X_t\|^2] < \infty$ provided that $\rho(A\otimes A + B\otimes B) < 1$, where $\otimes$ denotes the Kronecker product. This condition is analogous to the condition $\alpha + \beta < 1$ in Section VII.2.2 that ensures stationarity, ergodicity and finite second moment of x_t, when x_t is given by a univariate GARCH process.

ML estimation of BEKK models is considered in the works by Hafner and Preminger (2009) and Avarucci et al. (2012) and references therein.

VII.3.2 CCC

Another much applied specification for Ω_t is the so-called constant conditional correlation (CCC) model of Bollerslev (1990) and Jeantheau (1998).

Here, with $\{X_t\}_{t\in\mathbb{Z}}$ given by (VII.10)-(VII.12),

$$\Omega_t^{1/2} = D_t R^{1/2},$$

where $R^{1/2}$ is a lower triangular matrix arising from a Cholesky decomposition of a correlation matrix R, and D_t is a diagonal matrix with *positive* diagonal elements given by $\sqrt{h_{t,1}}, \ldots, \sqrt{h_{t,d}}$ with

$$h_t = (h_{t,1}, \ldots, h_{t,d})' = \omega + A\left(X_{t-1} \odot X_{t-1}\right). \tag{VII.13}$$

Here ω is a d-dimensional vector with strictly positive entries, A is a $(d \times d)$ matrix, and $\odot$ denotes element-wise multiplication of vectors or matrices of same dimensions, that is,

$$(X_t \odot X_t) = (X_{t,1}^2, \ldots, X_{t,d}^2)'.$$

We note that for $R = I_d$ and A diagonal with non-negative entries, $\{X_t\}_{t\in\mathbb{Z}}$ simply stacks d independent univariate ARCH(1) processes.

Example VII.3.2 *Let $d = 2$ such that the correlation matrix*

$$R = \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix},$$

with correlation $\rho \in (-1, 1)$ and

$$R^{1/2} = \begin{pmatrix} 1 & 0 \\ \rho & \sqrt{1-\rho^2} \end{pmatrix}.$$

With $Z_t = (Z_{t,1}, Z_{t,2})'$, we have that

$$\begin{pmatrix} X_{t,1} \\ X_{t,2} \end{pmatrix} = \begin{pmatrix} \sqrt{h_{t,1}} Z_{t,1} \\ \sqrt{h_{t,1}}\rho Z_{t,1} + \sqrt{h_{t,1}}\sqrt{1-\rho^2} Z_{t,2} \end{pmatrix},$$

with conditional covariance matrix

$$\Omega_t = \begin{pmatrix} h_{t,1} & \rho\sqrt{h_{t,1}h_{t,2}} \\ \rho\sqrt{h_{t,1}h_{t,2}} & h_{t,2} \end{pmatrix}.$$

With

$$\omega = \begin{pmatrix} \omega_1 \\ \omega_2 \end{pmatrix} \quad \text{and} \quad A = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix},$$

it holds that

$$h_t = \begin{pmatrix} h_{t,1} \\ h_{t,2} \end{pmatrix} = \begin{pmatrix} \omega_1 + A_{11}X_{t-1,1}^2 + A_{12}X_{t-1,2}^2 \\ \omega_2 + A_{21}X_{t-1,1}^2 + A_{22}X_{t-1,2}^2 \end{pmatrix}.$$

It holds that X_t has conditional covariance matrix (almost surely)

$$\begin{aligned}
\mathbb{V}[X_t|X_{t-1}] &= \mathbb{E}[D_t R^{1/2} Z_t Z_t' (D_t R^{1/2})' | X_{t-1}] \\
&= D_t R^{1/2} \mathbb{E}[Z_t Z_t' | X_{t-1}] (R^{1/2})' D_t \\
&= D_t R^{1/2} \mathbb{E}[Z_t Z_t'] (R^{1/2})' D_t \\
&= D_t R^{1/2} I_d (R^{1/2})' D_t \\
&= D_t R D_t \\
&= \Omega_t.
\end{aligned}$$

In particular, (almost surely)

$$\mathbb{V}(X_{t,i}|X_{t-1}) = h_{t,i}, \quad i = 1, \dots, d.$$

To ensure that all conditional variances $h_{t,i}$ are positive, all entries of the matrix A are typically assumed to be non-negative (see Conrad and Karanasos, 2010, for additional considerations).

Since R is a correlation matrix, the diagonal of D_t^2 is the diagonal of Ω_t, such that the *conditional* correlation matrix of X_t (given X_{t-1}) is

$$D_t^{-1} \Omega_t D_t^{-1} = R.$$

In terms of stationarity and ergodicity, let

$$\tilde{Z}_t = (\tilde{Z}_{t,1}, \dots, \tilde{Z}_{t,d})' := R^{1/2} Z_t,$$

such that $\{\tilde{Z}_t\}_{t \in \mathbb{Z}}$ is an i.i.d. process with $\tilde{Z}_t \overset{D}{=} N(0, R)$. Then, using that D_t^2 is a diagonal matrix with diagonal given by h_t,

$$X_t \odot X_t = (D_t \tilde{Z}_t) \odot (D_t \tilde{Z}_t) = \begin{pmatrix} \tilde{Z}_{t,1}^2 h_{t,1} \\ \vdots \\ \tilde{Z}_{t,d}^2 h_{t,d} \end{pmatrix} = K_t h_t.$$

with K_t a diagonal matrix with diagonal given by $(\tilde{Z}_t \odot \tilde{Z}_t)$. Using (VII.13), we have that $(X_t \odot X_t)$ obeys the SRE

$$(X_t \odot X_t) = \underbrace{K_t \omega}_{=B_t} + \underbrace{K_t A}_{=A_t} (X_{t-1} \odot X_{t-1}),$$

and conditions for stationarity and ergodicity can be derived using Theorem V.5.1. Sufficient conditions for finite moments of $\|X_t\|$ are given in Pedersen

(2017). We refer to Pedersen (2016) and Damek et al. (2019) for considerations about the tail index of the unconditional distributions of $\|X_t\|$ and $|X_{t,i}|$.

Similar to the BEKK model above, a GARCH extension of the CCC model is given by

$$h_t = \omega + A\left(X_{t-1} \odot X_{t-1}\right) + Bh_{t-1}.$$

Asymptotic theory for the ML estimator for CCC GARCH models is provided in Francq and Zakoïan (2012) and Pedersen (2017).

VII.4 (*) Application: Portfolio Choice

Consider an investor who can invest in a *risky asset* and a *risk-free asset* with returns x_{t+1} and $x_{t+1}^{(f)}$, respectively, from time t to $t+1$. The risk-free asset is risk-free in the sense that, given some information set $\mathcal{F}_t$, $x_{t+1}^{(f)} \in \mathcal{F}_t$. The objective of the investor at time t is to decide how much to invest in the risky asset relative to the risk-free asset. Specifically, let $w_t \in \mathcal{F}_t$ denote the weight put on the risky asset and $(1-w_t)$ the weight on the risk-free asset, that is, at time t $(w_t \times 100)\%$ of the money is invested in the risky asset. Then the one-period return of the portfolio, $x_{t+1}^{(p)}$, is given by

$$\begin{aligned} x_{t+1}^{(p)} &= w_t x_{t+1} + (1-w_t)x_{t+1}^{(f)} \\ &= w_t \tilde{x}_{t+1} + x_{t+1}^{(f)}, \end{aligned}$$

with $\tilde{x}_{t+1} = x_{t+1} - x_{t+1}^{(f)}$ denoting the *excess return* of the risky asset. The choice of ω_t depends on the investor's objective. For instance, $w_t = 0$ yields a deterministic return equal to the risk-free rate. This is desirable, if the investor is completely risk averse. In a more general setting, the investor may seek to balance (expected) reward and risk. The reward may be measured in terms of the expected portfolio return, $\mathbb{E}[x_{t+1}^{(p)}|\mathcal{F}_t]$, and the risk may be measured in terms of the conditional variance, $\mathbb{V}[x_{t+1}^{(p)}|\mathcal{F}_t]$ (or some other risk measure, such as Value-at-Risk or Expected Shortfall considered in Section VII.5 below). Consequently, and following e.g. Neely et al. (2014, Section 4), we suppose that the investor seeks to maximize the following utility function at time t,

$$\mathrm{U}_t = \mathbb{E}[x_{t+1}^{(p)}|\mathcal{F}_t] - \frac{\gamma}{2}\mathbb{V}[x_{t+1}^{(p)}|\mathcal{F}_t],$$

where $\gamma > 0$ is a risk aversion parameter. Note that a higher risk aversion γ, the lower utility for a given level of risk.

Using that $w_t, x_{t+1}^{(f)} \in \mathcal{F}_t$

$$\mathbb{E}[x_{t+1}^{(p)}|\mathcal{F}_t] = \mathbb{E}[w_t\tilde{x}_{t+1} + x_{t+1}^{(f)}|\mathcal{F}_t] = w_t\mathbb{E}[\tilde{x}_{t+1}|\mathcal{F}_t] + x_{t+1}^{(f)},$$

and

$$\mathbb{V}[x_{t+1}^{(p)}|\mathcal{F}_t] = w_t^2\mathbb{V}[\tilde{x}_{t+1}|\mathcal{F}_t],$$

so we have that

$$\mathrm{U}_t = w_t\mathbb{E}[\tilde{x}_{t+1}|\mathcal{F}_t] + x_{t+1}^{(f)} - \frac{\gamma}{2}w_t^2\mathbb{V}[\tilde{x}_{t+1}|\mathcal{F}_t].$$

Maximizing U_t with respect to w_t gives the first-order condition

$$\frac{\partial \mathrm{U}_t}{\partial w_t} = \mathbb{E}[\tilde{x}_{t+1}|\mathcal{F}_t] - \gamma w_t \mathbb{V}[\tilde{x}_{t+1}|\mathcal{F}_t] = 0,$$

with solution

$$\begin{aligned} w_t^* &= \frac{1}{\gamma}\left(\frac{\mathbb{E}[\tilde{x}_{t+1}|\mathcal{F}_t]}{\mathbb{V}[\tilde{x}_{t+1}|\mathcal{F}_t]}\right) \\ &= \frac{1}{\gamma}\left(\frac{\mathbb{E}[x_{t+1}|\mathcal{F}_t] - x_{t+1}^{(f)}}{\mathbb{V}[x_{t+1}|\mathcal{F}_t]}\right). \end{aligned}$$

The optimal weight w_t^* depends on the conditional mean and conditional variance of the risky asset. These two quantities may be specified in terms of econometric models of the type (VII.9).

Note that the above considerations can be extended to a multivariate setting where the investor can invest in d different risky assets (e.g., different stock indices, sectors, different asset classes), with returns given by the $(d \times 1)$ vector X_{t+1}. In this case, the weights on the risky assets are given by the $(d \times 1)$ vector w_t and the risk-free asset is given weight $1 - w_t'\iota_d$, where ι_d is a $(d \times 1)$ vector of ones. Then with $\tilde{X}_{t+1} = X_{t+1} - \iota_d x_{t+1}^{(f)}$ the $(d \times 1)$ vector of excess returns of the risky assets, the utility is given by

$$\mathrm{U}_t = w_t'\mathbb{E}[\tilde{X}_{t+1}|\mathcal{F}_t] + x_{t+1}^{(f)} - \frac{\gamma}{2} w_t'\mathbb{V}[\tilde{X}_{t+1}|\mathcal{F}_t] w_t,$$

where $\mathbb{E}[\tilde{X}_{t+1}|\mathcal{F}_t]$ is the vector of conditional mean excess returns and $\mathbb{V}[\tilde{X}_{t+1}|\mathcal{F}_t]$ the conditional covariance matrix of the excess returns. The vector of optimal weights are then given by

$$w_t^* = \frac{1}{\gamma}\left(\mathbb{V}[\tilde{X}_{t+1}|\mathcal{F}_t]\right)^{-1}\mathbb{E}[\tilde{X}_{t+1}|\mathcal{F}_t] = \frac{1}{\gamma}\left(\mathbb{V}[X_{t+1}|\mathcal{F}_t]\right)^{-1}\left(\mathbb{E}[X_{t+1}|\mathcal{F}_t] - \iota_d x_{t+1}^{(f)}\right),$$

provided that $\mathbb{V}[X_{t+1}|\mathcal{F}_t] = \Omega_{t+1}$ is invertible. Determining the optimal weight typically requires modelling the joint dynamics of the risky asset returns, which may be done in terms of VAR and/or multivariate GARCH models.

VII.5 (*) Application: Value-at-Risk (VaR) and beyond

An important measure for quantifying risk is the so-called *Value-at-Risk* (VaR) risk measure. VaR is important, for instance, from a regulatory perspective, as many financial institutions are obliged to disclose their estimates of such risk measures in relation to their holdings of risky assets (see, e.g., the Basel Committee on Banking Supervision). Importantly, a bank may be forced to set aside additional capital if their actual losses exceed their estimated VaR. We here discuss how VaR can be computed if the returns (or losses) are determined from an ARCH process. We emphasize that the computation of VaR boils down to computing a quantile of some (conditional) distribution. For the ARCH processes, on the other hand, conditional distributions are (in general) intractable, and we discuss how one may circumvent this issue by means of simulation-based estimation. We also discuss how the estimation uncertainty of the estimated VaR is addressed.

VII.5.1 VaR for ARCH processes

Let x_{t+1} denote the log-return of some asset from t to $t+1$. We shall also need the h-period return, $h \geq 1$, which by definition is given by

$$x_{t+1,h} = \sum_{i=1}^{h} x_{t+i}.$$

The 1*-period* VaR *at risk level* $\kappa \in (0,1)$ (or, in short, the VaR) is denoted VaR_t^{κ} and satisfies

$$\mathbb{P}(x_{t+1} < -\text{VaR}_t^{\kappa} | \mathcal{F}_t) = \kappa, \quad \text{VaR}_t^{\kappa} \in \mathcal{F}_t,$$

where $\mathcal{F}_t$ denotes some information set available at time t (e.g. the series of previous returns). Note that $\mathbb{P}(-x_{t+1} \leq \text{VaR}_t^{\kappa} | \mathcal{F}_t) = 1 - \kappa$, so that the VaR measures the maximum loss $(-x_{t+1})$ not exceeded with probability $1-\kappa$, or equivalently,VaR is the $1-\kappa$ percentile of the conditional loss distribution.[1] Note that, by construction, the VaR depends on the return process, the

[1] Note that the definition above implicitly assumes that the VaR exists, which is indeed the case whenever the conditional return distribution is continuous. A more general definition that ensures that the VaR always exists is that $\text{VaR}_t^{\kappa} = \inf\{y \in \mathbb{R} : P(-x_{t+1} \leq y | \mathcal{I}_t) \geq 1-\kappa\}$. Some textbooks, such as the one by Francq and Zakoïan (2019), make the convention that the VaR must be non-negative, such that the VaR is given by $\max[0, \inf\{y \in \mathbb{R} : P(-x_{t+1} \leq y | \mathcal{I}_t) \geq 1-\kappa\}]$.

information set $\mathcal{F}_t$, as well as the *confidence level* $1-\kappa$. Typical values of κ in applications are 1%, 2.5%, and 5%. Throughout, we assume that the information set contains only past values of the returns, i.e. $\mathcal{F}_t = \{x_i : i \leq t\}$. We emphasize that one could include additional variables to the information set, which would lead to careful considerations, and assumptions, about how these variables are related to the return process.

Example VII.5.1 (Gaussian returns) *Suppose that $\{x_t\}_{t\in\mathbb{Z}}$ is an i.i.d. process with $x_t \overset{D}{=} N(0,1)$. Then, with $\Phi(\cdot)$ the cdf of the $N(0,1)$ distribution and using that x_{t+1} is independent of $\mathcal{F}_t$,*

$$\mathbb{P}(x_{t+1} < -\mathrm{VaR}_t^\kappa | \mathcal{F}_t) = \Phi(-\mathrm{VaR}_t^\kappa) = \kappa$$

Hence,

$$\mathrm{VaR}_t^\kappa = -\Phi^{-1}(\kappa),$$

i.e. the VaR *is (negative) the κ percentile of the standard normal distribution. Likewise, if instead $x_t \overset{D}{=} N(\mu, \sigma^2)$,*

$$\mathrm{VaR}_t^\kappa = -\mu - \sigma\Phi^{-1}(\kappa).$$

Example VII.5.2 (ARCH returns) *Suppose that the returns are given by the stationary ARCH(1) process*

$$\begin{aligned} x_t &= \sigma_t z_t, \quad t \in \mathbb{Z}, \\ \sigma_t^2 &= \omega + \alpha x_{t-1}^2, \end{aligned}$$

with $\{z_t\}_{t\in\mathbb{Z}}$ an i.i.d. process with $z_t \overset{D}{=} N(0,1)$ and $\omega > 0$, $\alpha \geq 0$. Then using that $\sigma_{t+1}^2 \in \mathcal{F}_t$ and z_{t+1} and $\mathcal{F}_t$ are independent,

$$\begin{aligned} \kappa &= \mathbb{P}\left(x_{t+1} < -\mathrm{VaR}_t^\kappa | \mathcal{F}_t\right) \\ &= \mathbb{P}\left(\sigma_{t+1} z_{t+1} < -\mathrm{VaR}_t^\kappa | \mathcal{F}_t\right) \\ &= \mathbb{P}\left(z_{t+1} < -\mathrm{VaR}_t^\kappa / \sigma_{t+1} | \mathcal{F}_t\right) \\ &= \Phi(-\mathrm{VaR}_t^\kappa / \sigma_{t+1}). \end{aligned}$$

Hence,

$$\mathrm{VaR}_t^\kappa = -\sigma_{t+1}\Phi^{-1}(\kappa). \tag{VII.14}$$

Note that the above considerations apply to any $0 < \sigma_{t+1} \in \mathcal{F}_t$, e.g. GARCH(1,1) processes.

The above definition of 1-period VaR easily extends to the h-period VaR, $\mathrm{VaR}_{t,h}^\kappa$, given by

$$\mathbb{P}\left(x_{t+1,h} < -\mathrm{VaR}_{t,h}^\kappa | \mathcal{F}_t\right) = \kappa. \tag{VII.15}$$

Example VII.5.3 (Gaussian returns, ctd.) *Proceeding with the case of i.i.d. returns with* $x_t \overset{D}{=} N(\mu, \sigma^2)$, *it follows that* $x_{t+1,h} = \sum_{i=1}^{h} x_{t+i} \overset{D}{=} h\mu + \sigma\sqrt{h}z$ *where* $z \sim N(0,1)$ *and independent of* $\mathcal{F}_t$. *Hence,*

$$\begin{aligned}
\mathbb{P}\left(x_{t+1,h} < -\mathrm{VaR}_{t,h}^{\kappa} | \mathcal{F}_t\right) &= \mathbb{P}\left(h\mu + \sigma\sqrt{h}z < -\mathrm{VaR}_{t,h}^{\kappa}\right) \\
&= \mathbb{P}\left(z < -\frac{\mathrm{VaR}_{t,h}^{\kappa} + h\mu}{\sigma\sqrt{h}}\right) \\
&= \Phi\left(-\frac{\mathrm{VaR}_{t,h}^{\kappa} + h\mu}{\sigma\sqrt{h}}\right) = \kappa,
\end{aligned}$$

such that

$$\mathrm{VaR}_{t,h}^{\kappa} = -h\mu - \sigma\sqrt{h}\Phi^{-1}(\kappa).$$

Example VII.5.4 (ARCH returns, ctd.) *For the ARCH(1) we have* $x_{t+1} = \sqrt{\omega + \alpha x_t^2} z_{t+1}$, *with* $z_{t+1} \overset{D}{=} N(0,1)$. *Since the factor* $\sqrt{\omega + \alpha x_t^2} \in \mathcal{F}_t$, *and* z_{t+1} *is independent of* $\mathcal{F}_t$, *it holds that* $x_{t+1}|\mathcal{F}_t \overset{D}{=} N(0, \omega + \alpha x_t^2)$ *which we exploited in Example VII.5.2 to find the 1-period* VaR. *Now, suppose that we want to compute the two-period ahead* VaR. *This relies on computing the* κ *percentile of the conditional loss distribution, i.e. the conditional distribution of* $-x_{t+1,2}$ *given* $\mathcal{F}_t$ *with* $x_{t+1,2} = (x_{t+1} + x_{t+2})$. *By recursions,*

$$x_{t+2} = \sqrt{\omega + \alpha x_{t+1}^2} z_{t+2} = \sqrt{\omega + \alpha(\omega + \alpha x_t^2) z_{t+1}^2} z_{t+2}.$$

Clearly, $x_{t+2}|\mathcal{F}_t \overset{D}{=} N(0, \omega + \alpha x_{t+1}^2)$, *but the conditional distribution of* x_{t+2} *(given* $\mathcal{F}_t$*) is non-Gaussian, since the factor* $\sqrt{\omega + \alpha(\omega + \alpha x_t^2) z_{t+1}^2}$ *does not belong to the information set* $\mathcal{F}_t$. *In fact, it can be shown that conditional distribution of* x_{t+2} *is a so-called normal variance mixture whenever* $\alpha > 0$ *(see, e.g., the recent work of Abadir et al., 2023). Consequently, the conditional distribution of* $x_{t+1,2}$ *is a sum of dependent normal variance mixture random variables, and may be viewed as effectively intractable. In such a case it is customary to quantify (or approximate) the* VaR *using the algorithm stated below.*

Algorithm VII.5.1 *Let* $(\omega, \alpha)'$ *and* x_t *be known and fixed.*

1. *For* $i = 1, \ldots, M$ *(with* $(1-\kappa)M \geq 1$*) draw* $z_{t+1}^{(i)}$ *and* $z_{t+2}^{(i)}$ *independently from* $N(0,1)$, *and compute*

$$x_{t+1,2}^{(i)} = (x_{t+1}^{(i)} + x_{t+2}^{(i)}),$$

with

$$x_{t+1}^{(i)} = \sqrt{\omega + \alpha x_t^2} z_{t+1}^{(i)},$$
$$x_{t+2}^{(i)} = \sqrt{\omega + \alpha (x_{t+1}^{(i)})^2} z_{t+2}^{(i)}.$$

2. *Consider the ordered returns* $x_{t+1,2}^{[M]} \leq \ldots \leq x_{t+1,2}^{[1]}$. *Using the definition of* VaR *in (VII.15), obtain the approximate* VaR *as the* $(1-\kappa)$ *empirical quantile of the simulated losses, i.e.*

$$\mathrm{VaR}_{t,2}^{\kappa,\mathrm{sim}} = -(x_{t+1,2}^{[\lfloor (1-\kappa)M \rfloor]}),$$

where $\lfloor y \rfloor$ *denotes the integer part of* $y \in \mathbb{R}$.

VII.5.2 VaR Inference

In practice, the VaR depends on unknown parameters, e.g., $\theta = (\mu, \sigma^2)'$ in Example VII.5.1 and $\theta = (\omega, \alpha)'$ in Example VII.5.2. As already considered, these parameters may be estimated by ML (or other estimation methods), leading to estimators of VaR. Provided that the ML estimators for θ are consistent and asymptotically normal (cf., Theorems VI.2.2 and VI.3.1), we likewise have that the VaR estimators are consistent and asymptotically normal.

Example VII.5.5 (Gaussian returns, ctd.) *Consider Example VII.5.1 with* $x_t \overset{D}{=} N(0, \sigma^2)$. *Given a set of observations* $\{x_t\}_{t=1,\ldots,T}$, *the ML estimator for* σ^2 *is given by*

$$\hat{\sigma}_T^2 = \frac{1}{T} \sum_{t=1}^{T} x_t^2,$$

and an estimator for the h*-period* VaR *is given by*

$$\widehat{\mathrm{VaR}}_{t,h}^{\kappa} = -\hat{\sigma}_T \sqrt{h} \Phi^{-1}(\kappa).$$

Trivially, the i.i.d. DGP $\{x_t\}_{t \in \mathbb{Z}}$ *is stationary and ergodic with* $\mathbb{E}[x_t^2] < \infty$. *By Theorem V.2.2* $\hat{\sigma}_T^2 \xrightarrow{p} \sigma_0^2$ *as* $T \to \infty$, *and hence* $\widehat{\mathrm{VaR}}_{t,h}^{\kappa}$ *is consistent for* $\mathrm{VaR}_{t,h}^{\kappa}$, *that is*

$$\widehat{\mathrm{VaR}}_{t,h}^{\kappa} \xrightarrow{p} \mathrm{VaR}_{t,h}^{\kappa}.$$

Likewise, noting that $\mathbb{E}[x_t^2 - \sigma_0^2 | \mathcal{F}_t] = \mathbb{E}[x_t^2 - \sigma_0^2] = 0$ *almost surely, and* $\Sigma := \mathbb{E}[(x_t^2 - \sigma_0^2)^2] = 2\sigma_0^4 < \infty$, *we have by Theorem V.2.3,* $\sqrt{T}(\hat{\sigma}_T^2 - \sigma_0^2) \xrightarrow{D} N(0, \Sigma)$. *Moreover, since* $x \mapsto \sqrt{x}$ *is continuously differentiable on* *the positive real axis, we have that* $\sqrt{T}(\hat{\sigma}_T - \sigma_0) \xrightarrow{D} N(0, \sigma_0^2/2)$ *[by the* Δ*-method]. Hence,*

$$\sqrt{T}\left(\widehat{\mathrm{VaR}}^{\kappa}_{t,h} - \mathrm{VaR}^{\kappa}_{t,h}\right) = -\sqrt{h}\Phi^{-1}(\kappa)\sqrt{T}(\hat{\sigma}_T - \sigma_0) \xrightarrow{D} N(0, h\sigma_0^2\Phi^{-1}(\kappa)^2/2).$$

Hence, we may apply the approximation

$$\widehat{\mathrm{VaR}}^{\kappa}_{t,h} \overset{D}{\approx} N(\mathrm{VaR}^{\kappa}_{t,h}, h\sigma_0^2\Phi^{-1}(\kappa)^2/(2T)),$$

and one may report, for instance, 95% error bands of the VaR *as* $\widehat{\mathrm{VaR}}^{\kappa}_{t,h} \pm \Phi^{-1}(0.975)\sqrt{h/(2T)}|\Phi^{-1}(\kappa)|\hat{\sigma}_T$. *In order to take into account the estimation uncertainty, or the additional "estimation risk", one may for instance use the upper band,*

$$\widehat{\mathrm{VaR}}^{\kappa}_{t,h} + 1.96\sqrt{h/(2T)}|\Phi^{-1}(\kappa)|\hat{\sigma}_T,$$

as the "estimation risk-adjusted VaR *measure".*

Example VII.5.6 (ARCH returns, ctd.) *From Part VI, we can estimate the parameters* $\theta = (\omega, \alpha)'$ *by ML estimation* $\hat{\theta}_T = (\hat{\omega}_T, \hat{\alpha}_T)'$. *Assume that the DGP satisfies* $\alpha_0 \in (0, 1)$. *Then from Example VI.2.5 and Section VI.3.1, respectively,*

$$\hat{\theta}_T \xrightarrow{p} \theta_0 \quad \text{and} \quad \sqrt{T}(\hat{\theta}_T - \theta_0) \xrightarrow{D} N(0, -\Sigma_0^{-1}), \tag{VII.16}$$

for some positive definite matrix $-\Sigma_0$. *Based on the estimator* $\hat{\theta}_T$, *we obtain an estimator for the conditional volatility, given by*

$$\hat{\sigma}_{t+1} = \sqrt{\hat{\omega}_T + \hat{\alpha}_T x_t^2},$$

and, using (VII.14), we have the VaR *estimator,*

$$\widehat{\mathrm{VaR}}^{\kappa}_{t} = -\hat{\sigma}_{t+1}\Phi^{-1}(\kappa).$$

Notice that (unlike the case of Gaussian returns in the previous examples) VaR^{κ}_t *is random as it depends on* x_t. *In order to analyze the statistical properties of the* VaR *estimator, it is customary to consider* x_t *as fixed and setting it equal to some fixed value,* $x_t = x$. *We then have that*

$$\widehat{\mathrm{VaR}}^{\kappa}_{t} - \mathrm{VaR}^{\kappa}_{t} = -(\hat{\sigma}_{t+1} - \sigma_{t+1})\Phi^{-1}(\kappa),$$

with $\hat{\sigma}^2_{t+1} = \hat{\omega}_T + \hat{\alpha}_T x^2$ *and* $\sigma^2_{t+1} = \omega + \alpha x^2$. *By (VII.16), a first-order Taylor expansion (up to a negligible remainder term),*

$$\hat{\sigma}_{t+1} - \sigma_{t+1} = \frac{1}{2\sqrt{\theta_0' w}} w'(\hat{\theta}_T - \theta_0),$$

with $w = (1, x^2)'$, and we conclude that

$$\widehat{\text{VaR}}_t^\kappa \xrightarrow{p} \text{VaR}_t^\kappa,$$

and

$$\sqrt{T}(\widehat{\text{VaR}}_t^\kappa - \text{VaR}_t^\kappa) \xrightarrow{D} N\left(0, \frac{\Phi^{-1}(\kappa)^2}{4\theta_0' w} w'(-\Sigma_0^{-1})w\right).$$

Similar to Example VII.5.5, one may use this result to construct error bands for the estimated VaR, *for instance by relying on the estimator $\hat{\Psi}_T$ for $(-\Sigma_0^{-1})$ provided in Example VI.4.2.*

VII.5.3 (*) Extensions and alternative risk measures

We end this section by providing some directions for potential extensions and additional details.

VII.5.3.1 Extending Algorithm VII.5.1

The Algorithm VII.5.1 easily extends to any finite horizon $h \geq 2$, other specifications for σ_t, as well as other distributions for z_t. Moreover, even if the distribution of z_t is assumed to be unknown one may incorporate draws from the empirical distribution of the standardized returns $\hat{z}_t = x_t/\hat{\sigma}_t$. This approach is widely used in applications, as is typically referred to as so-called *filtered historical simulation* (Barone-Adesi et al., 1999).

One may also extend the algorithm to incorporate estimation uncertainty. This may for instance be done by making draws of θ from the limiting distribution of the ML estimator in (VII.16), and compute the VaR for each draw (e.g., Blasques et al., 2016). Alternatively, one may instead draw θ from bootstrapped distributions of the ML estimator (see e.g. Beutner et al., 2024 and Cavaliere et al., 2018).

VII.5.3.2 VaR Backtesting

In practice, one may typically want to evaluate if a given (estimated) econometric model, such as the ARCH(1), does well in terms of quantifying VaR. One way of doing so is to test for so-called *unconditional correct coverage*, which can be viewed as a misspecification test. Given a set of observations $\{x_t\}_{t=1}^T$ and their associated VaR, $\{\text{VaR}_{t-1,1}^\kappa\}_{t=1}^T$, define the *hit* sequence

$$\text{Hit}_t = \begin{cases} 1 & \text{if } -x_t > \text{VaR}_{t-1,1}^\kappa \\ 0 & \text{otherwise} \end{cases}, \quad t = 1, 2, \ldots, T.$$

Given correct model specification, it holds that the process $\{\text{Hit}_t\}_{t=1}^T$ is i.i.d. Bernoulli(κ), such that

$$E[\text{Hit}_t] = \kappa.$$

Kupiec suggested to model $\{\text{Hit}_t\}_{t=1}^T$ as an *i.i.d.* Bernoulli(p) sequence, and test the hypothesis $p = \kappa$ against $p \neq \kappa$. The ML estimator for p is given by

$$\hat{p} = T^{-1} \sum_{t=1}^{T} \text{Hit}_t,$$

and the LR statistic for the hypothesis is given by

$$LR_T(p = \kappa) = -2 \log \left[\frac{(1-\kappa)^{T_0} \kappa^{T_1}}{(1-\hat{p})^{T_0} \hat{p}^{T_1}} \right],$$

with $T_1 = \sum_{t=1}^{T} \text{Hit}_t$ and $T_0 = T - T_1$. Under the hypothesis, it holds that

$$LR_T(p = \kappa) \xrightarrow{D} \chi_1^2, \quad \text{as } T \to \infty,$$

which can be proved by verifying the conditions of Theorem VI.4.1

Note that as in Example VII.5.6, the VaR is typically estimated, based on estimated parameters. Hence, in practice one has the sequence $\{\widehat{\text{Hit}}_t\}_{t=1}^T$ based on estimates of $\text{VaR}_{t-1,1}^{\kappa}$. Inherently, the LR statistic $LR_T(p = \kappa)$ depends on the ML estimator for the model parameters $\hat{\theta}_T = (\hat{\omega}_T, \hat{\alpha}_T)'$, and the χ_1^2 limiting distribution is potentially unreliable. This issue has been addressed by Escanciano and Olmo (2010).

Lastly, one may note that there exist various extensions of Kupiec's test, for instance allowing for alternatives where violations of the value-at-risk, that is, the events $\text{Hit}_t = 1$ appear in consecutive time periods; see e.g. the much applied test by Christoffersen (1998).

VII.5.3.3 Expected Shortfall (ES)

Although VaR is widely used in practice, some researchers and practitioners has raised their concern that VaR is unreliable risk measure in certain applications. Specifically, it can be shown that VaR is a so-called incoherent risk measure. Specifically, under certain (heavy tailed) asset return distributions, it can be shown that VaR discourages risk diversification and hence may be an undesirable measure with respect to managing risk; see e.g. Ibragimov (2009). Moreover, recall that VaR is the maximum loss not exceeded with a given probability $1 - \kappa$. The risk measure does not tell us how much we lose (or may expect to lose) given that the loss exceeds the VaR (which happens with probability κ). These concerns have led to the so-called Expected Shortfall (ES) risk measure that, by definition, quantifies the expected loss given that the loss exceeds the VaR. *The* 1*-period* ES *at risk level* $\kappa \in (0,1)$, is given by

$$\mathrm{ES}_{t,1}^{\kappa} = \mathbb{E}[-x_{t+1} | x_{t+1} < -\mathrm{VaR}_t^{\kappa}, \mathcal{F}_t],$$

and clearly $\mathrm{ES}_t^{\kappa} \geq \mathrm{VaR}_t^{\kappa}$. If x_t follows an ARCH(1) process as in Example VII.5.2, it can be shown that

$$\mathrm{ES}_{t,1}^{\kappa} = \kappa^{-1} \sigma_{t+1} \phi(-\Phi^{-1}(\kappa)),$$

where ϕ is the pdf of the $N(0,1)$ distribution. Similar to VaR, one may address the estimation uncertainty when estimating ES. Likewise, parallel to Algorithm VII.5.1, one may have to compute multiple period ES by means of simulations.

VII.5.3.4 Alternative risk measures

Recently, risk quantification has been given wide attention in relation to so-called systemic risk. For instance Adrian and Brunnermeier (2016), see also Banulescu-Radu et al. (2021), consider VaR **co**nditional on a particular event, labelled CoVaR. As an example, let $-x_{t+1}$ denote the loss of a share of Company 1 and $-y_{t+1}$ the loss of a share of Company 2. Then the CoVaR is given by

$$\mathbb{P}(-x_{t+1} > \mathrm{CoVaR}_{t,1}^{\kappa} | -y_{t+1} > \mathrm{VaR}_t^{Y,\kappa}, \mathcal{F}_t) = \kappa,$$

where $\mathrm{VaR}_t^{Y,\kappa}$ is the VaR for Company 2. Likewise, Brownlees and Engle (2017) have considered the notion of so-called long-run marginal expected shortfall, LRMES: Let $-x_{t+1,h}$ denote the h-period loss of the share of some bank (say, Goldman Sachs), and let $-y_{t+1,h}$ denote the loss of the entire financial sector (e.g., as mention as the loss of an equity index consisting of financial companies). Then the LRMES is given by

$$\mathrm{LRMES}_{t,h}^{\kappa} = \mathbb{E}[-x_{t+1,h} | -y_{t+1,h} > 40\%],$$

where the event $-y_{t+1,h} > 40\%$ is interpreted as a major financial crisis. Note that the computation of CoVaR and LRMES requires a specification of the joint conditional distribution of (x_{t+1}, y_{t+1}), which may be done in terms of a multivariate GARCH model.

The notion of VaR may be used outside of financial risk management. For instance the International Monetary Fund (IMF), as well as several central banks, make use of the so-called Growth-at-Risk (GaR) measure that quantifies the smallest possible (that is, the worst-case scenario) GDP growth rate at a given confidence level $(1-\kappa)$; see e.g. Brownlees and Souza (2021). For a given country, let Y_t denote the GDP growth rate in, say, quarter t, then the 1-step ahead GaR at risk level κ, $\mathrm{GaR}_{t,1}^{\kappa}$, is defined by

$$\mathbb{P}(Y_{t+1} < \mathrm{GaR}_{t,1}^{\kappa} | \mathcal{F}_t) = \kappa, \quad \mathrm{GaR}_{t,1}^{\kappa} \in \mathcal{F}_t,$$

for some information set $\mathcal{F}_t$, potentially containing various macroeconomic variables and indicators.

References

Abadir, K.M., Luati, A., & Paruolo, P., 2023, "GARCH density and functional forecasts", *Journal of Econometrics*, Vol. 235, pp. 470–483.

Adrian, T., & Brunnermeier, M.K., 2016, "CoVaR", *American Economic Review*, Vol. 106, pp. 1705–1741.

Avarucci, M., Beutner, E., & Zaffaroni, P., 2012, "On moment conditions for quasi-maximum likelihood estimation of multivariate ARCH models", *Econometric Theory*, Vol.29, pp. 545–566.

Banulescu-Radu, D., Hurlin, C., Leymarie, J., & Scaillet, O., 2021, "Backtesting Marginal Expected Shortfall and Related Systemic Risk Measures", *Management Science*, Vol. 67, pp. 5730-5754.

Barone-Adesi, G., Giannopoulos, K., & Vosper, L., 1999, "VaR without correlations for portfolios of derivative securities", *The Journal of Futures Markets*, Vol. 19, pp. 583–602.

Beutner, E., Heinemann, A., & Smeekes, S., 2024, "A residual bootstrap for conditional Value-at-Risk", *Journal of Econometrics*, Vol. 238, 105554.

Blasques, F., Francq, C., & Laurent, S., 2024, "Autoregressive conditional betas", *Journal of Econometrics*, Vol. 238, 105630.

Blasques, F., Koopmann, S.J. , Lasak, K., & Lucas, A., 2016, "In-sample confidence bands and out-of-sample forecast bands for time-varying parameters in observation-driven models", *International Journal of Forecasting*, Vol. 32, pp. 875–887.

Bollerslev, T., 1986, "Generalized autoregressive conditional heteroskedasticity", *Journal of Econometrics*, Vol. 31, pp. 307–327.

Bollerslev, T., 1990, "Modelling the coherence in short-run nominal exchange rates: A multivariate Generalized ARCH model Arch Model", *Review of Economics and Statistics*, Vol. 72, pp. 498–505.

Bollerslev, T., 2010, "Glossary to ARCH (GARCH)" in *Volatility and Time Series Econometrics: Essays in Honor of Robert Engle*, Oxford University Press.

Bougerol, P., & Picard, N., 1992, "Strict stationarity of generalized autoregressive processes", *The Annals of Statistics*, Vol. 20, pp. 1714–1730.

Boussama, F., Fuchs, F., & Stelzer, R., 2011, "Stationarity and geometric ergodicity of BEKK multivariate GARCH models", *Stochastic Processes and their Applications*, Vol. 121, pp. 2331–2360.

Brownlees, C., & Engle, R., 2016, "SRISK: A Conditional Capital Shortfall Measure of Systemic Risk", *Review of Financial Studies*, *Vol. 30, pp. 48–79.*

Brownlees, C., & Souza, A.B.M., 2021, "Backtesting global Growth-at-Risk", *Journal of Monetary Economics*, Vol. 118, pp. 312–330.

Cavaliere, G., Pedersen, R.S., & Rahbek, A., 2018, "The fixed volatility bootstrap for a class of ARCH processes", *Journal of Time Series Analysis*, vol. 39, pp. 920-941.

Chorro, C., Guégan, D., & Ielpo, F., 2015, *A Time Series Approach to Option Pricing*, Springer.

Christoffersen, P., 1998, "Evaluating interval forecasts", *International Economic Review*, Vol. 39, pp. 841–862.

Conrad, C., & Karanasos, M., 2010, "Negative volatility spillovers in the unrestricted ECCC-GARCH model", *Econometric Theory*, Vol. 26, pp. 838–862.

Damek, E., Matsui, M., & Swiatkowski, W., (2019), "Componentwise different tail solutions for bivariate stochastic recurrence equations with application to GARCH (1,1) processes", *Colloquium Mathematicum*, Vol. 155, pp. 227–254.

D'Innocenzo, E., & Lucas, A., 2024, "Dynamic partial correlation models", *Journal of Econometrics*, Vol. 241, 105747.

Engle, R.F., 1982, "Autoregressive conditional heteroskedasticity with estimates of the variance of United Kingdom inflation", *Econometrica*, Vol. 50, pp. 987–1008.

Engle, R.F., 2002, "Dynamic conditional correlation: A simple class of multivariate generalized autoregressive conditional heteroskedasticity models", *Journal of Business & Economic Statistics*, Vol. 20, pp. 339–350.

Engle, R.F., 2016, "Dynamic conditional beta", *Journal of Financial Econometrics*, Vol. 14., pp. 643–667.

Engle, R.F., & Kroner, K.F., 1995, "Multivariate simultaneous generalized ARCH", *Econometric Theory*, Vol. 11, pp. 122–150.

Escanciano, J.C., & Olmo, J., 2010, "Backtesting parametric Value-at-Risk with estimation risk", *Journal of Business & Economic Statistics*, Vol. 28, pp. 36–51.

Francq, C. & Thieu, L.Q., 2019, "QML inference for volatility models with covariates", *Econometric Theory*, Vol. 35, pp. 37–72.

Francq, C. & Zakoïan, J.-M., 2012, "QML estimation of a class of multivariate asymmetric GARCH models", *Econometric Theory*, Vol. 28, pp. 179–206.

Francq, C. & Zakoïan, J.-M., 2016, "Estimating multivariate volatility models equation by equation", *Journal of the Royal Statistical Society: Series B*, Vol. 78, pp. 613–635.

Francq, C. & Zakoïan, J.-M., 2019, *GARCH Models: Structure, Statistical Inference and Financial Applications*, 2nd edition. Wiley.

Glosten, L.R., Jagannathan, R., & Runkle, D.E., 1993, "On the relation between expected value and the volatility of the nominal excess return on stocks", *The Journal of Finance*, Vol. XLVIII, pp. 1779–1801.

Hafner, C.M., & Preminger, A., 2009, "On asymptotic theory for multivariate GARCH models", *Journal of Multivariate Analysis*, Vol. 100, pp. 2044–2054.

Han, H, & Kristensen, D., 2014, "Asymptotic theory for the QMLE in GARCH-X models with stationary and nonstationary covariates", *Journal of Business & Economic Statistics*, Vol. 32, pp. 416–429.

Hetland, S., Pedersen, R.S., & Rahbek, A., 2023, "Dynamic conditional eigenvalue GARCH", *Journal of Econometrics*, Vol. 237, 105175.

Ibragimov, R., 2009, "Portfolio diversification and value at risk under thick-tailedness", *Quantitative Finance*, Vol. 9, pp. 565–580.

Jeantheau, T., 1998, "Strong consistency of estimators for multivariate ARCH models", *Econometric Theory*, Vol. 14, pp. 70–86.

Matsui, M., & Pedersen, R.S., 2022, "Characterization of the tail behavior of a class of BEKK processes: A stochastic recurrence equation approach", *Econometric Theory*, Vol. 38, pp. 1–34.

Mikosch, T., & Starica, C., 2000, "Limit theory for the sample autocorrelations and extremes of a GARCH(1,1) process“, *The Annals of Statistics*, Vol. 28, pp. 1427–1451.

Mikosch, T., & Straumann, D., 2006, "‘Stable limits of martingale transforms with application to the estimation of GARCH parameters“, *The Annals of Statistics*, Vol. 34, pp. 493–522.

Neely, C.J., Rapach, D.E., Tu, J., & Zhou, G., 2014, "Forecasting the equity risk premium: The role of technical indicators“, *Management Science*, Vol. 60, pp. 1772–1791.

Nielsen, H.B., & Rahbek, A., 2024, "Penalized quasi-likelihood estimation and model selection with parameters on the boundary of the parameter space“, *Econometrics Journal*, Vol. 27, pp. 107–125.

Noureldin, D., Shephard, N., & Sheppard, K., 2014, "Multivariate rotated ARCH models“, *Journal of Econometrics*, Vol. 179, pp. 16–30.

Pedersen, R.S., (2016), "Targeting estimation of CCC-GARCH models with infinite fourth moments“, *Econometric Theory*, Vol. 32, pp. 498–531.

Pedersen, R.S., (2017), "Inference and testing on the boundary in extended constant conditional correlation GARCH models“, *Journal of Econometrics*, Vol. 196, pp. 23–36.

Pedersen, R.S., & Rahbek, A., 2014, "Multivariate variance targeting in the BEKK-GARCH model“, *Econometrics Journal*, Vol. 17, pp. 24–55.

Pedersen, R.S., & Rahbek, A., 2019, "Testing GARCH-X type models“, *Econometric Theory*, Vol. 35, pp. 1012–1047.

Pelletier, D., 2006, "Regime switching for dynamic correlations“, *Journal of Econometrics*, Vol. 131, pp. 445-473.

Taylor, S.J., 1986, *Modelling Financial Time Series*, Wiley.

Tsay, R.S., 2010, *Analysis of Financial Time Series*, 3rd edition, Wiley.