Risk Management and Time Series

Time series models are often employed in risk management applications. They can be used to estimate conditional loss distributions which in turn can be used to estimate risk measures such as VaR and CVaR. In these notes we will briefly describe GARCH models which are often used in financial applications. In addition to describing how these models can be used to estimate loss measures based on the conditional loss distribution, we can also use GARCH models for more sophisticated risk management applications. These applications include using GARCH to model the dynamics of *principal components* in a factor model or combining GARCH with *extreme value theory* to estimate the conditional probability of extreme losses in a given portfolio. We note that these methods can of course also be used with other time series models.

1 A Brief Introduction to GARCH Models

Let r_t denote the $\log$-return of a portfolio between periods $t-1$ and t and let $\mathcal{F}_t$ denote the *information filtration* generated by these returns. We assume

$$r_t \;=\; \mu_t + a_t \tag{1}$$

where $\mathsf{E}[r_t \mid \mathcal{F}_{t-1}] = \mu_t$ and

$$\sigma_t^2 \;:=\; \mathsf{Var}\left(r_t \mid \mathcal{F}_{t-1}\right) \;=\; \mathsf{Var}\left(a_t \mid \mathcal{F}_{t-1}\right) \tag{2}$$

Note that μ_t and σ_t are known at time $t-1$ and are therefore *predictable* processes. It is common to assume that μ_t follows a stationary time series model such as an *ARMA*(m,s) process in which case

$$\mu_t \;=\; \phi_0 \;+\; \sum_{i=1}^{m} \phi_i r_{t-i} \;+\; \sum_{j=1}^{s} \gamma_j a_{t-j} \tag{3}$$

where the ϕ_i's and γ_j's are parameters to be estimated. Note that a_t is the *innovation* or random component of the log-return and in practice, the conditional variance of this innovation, σ_t^2, is time varying and stochastic. GARCH models may be used to model this dynamics behavior of conditional variances. In particular, we say σ_t^2 follows a GARCH(p,q) model[1] if

$$a_t \;=\; \sigma_t\, \epsilon_t \tag{4}$$

$$\text{and} \quad \sigma_t^2 \;=\; \alpha_0 \;+\; \sum_{i=1}^{p} \alpha_i a_{t-i}^2 \;+\; \sum_{j=1}^{q} \beta_j \sigma_{t-j}^2 \tag{5}$$

where $\alpha_0 > 0$, $\alpha_i \geq 0$ for $i = 1, \ldots, p$, $\beta_j \geq 0$ for $j = 1, \ldots, q$ and where the ϵ_t's are IID random variables with mean zero and variance one. It should be clear from (4) and (5) that the volatility clustering effect we observe in the market-place is therefore captured by the GARCH(p,q) model. Fat tails are also captured by this model in the sense that the log-returns in (1) have heavier tails (when a_t satisfies (4) and (5)) than the corresponding normal distribution. GARCH models are typically estimated jointly with the conditional mean equation in (3) using maximum likelihood techniques. The fitted model can then be checked for goodness-of-fit using standard diagnostic methods. See, for example, Ruppert (2011) or Tsay (2010).

[1] See Chapter 18 of *Statistics and Data Analysis for Financial Engineering* (2011) by Ruppert or Chapter 3 of *Analysis of Financial Time Series* (2010) by Tsay for a discussion of volatility modeling in general and additional details on the GARCH(p,q) model. Our modal description here follows Tsay. The acronym "GARCH" stands for *generalized autoregressive conditional heteroscedastic.*

1.1 The GARCH$(1,1)$ Model

The GARCH$(1,1)$ is obtained when we take $p = q = 1$ in (5) and it is the most commonly used GARCH model in practice. It is not too difficult to see that the GARCH$(1,1)$ model is stationary if and only $\alpha_1 + \beta_1 < 1$ in which case the *unconditional variance*, θ, is given by

$$\theta \;=\; \frac{\alpha_0}{1 - \alpha_1 - \beta_1}.$$

Exercise 1 *Under what condition(s) is the GARCH(p, q) model stationary? Give an expression for the unconditional variance, θ, when these condition(s) are satisfied.*

Exercise 2 *Assuming $\alpha_1 + \beta_1 < 1$ show that we may write*

$$\sigma_{t+1}^2 = \kappa\theta + (1-\kappa)\left[(1-\lambda)a_t^2 + \lambda\sigma_t^2\right] \tag{6}$$

for some constants κ and λ. What are the values of κ and λ?

Forecasting with the GARCH$(1,1)$ Model

Noting that $\mathsf{E}[r_{t+1}^2 \mid \mathcal{F}_t] = \sigma_{t+1}^2$ we can use (6) to obtain

$$\begin{aligned}\mathsf{E}[\sigma_{t+2}^2 \mid \mathcal{F}_t] &= \kappa\theta + (1-\kappa)\left[(1-\lambda)\mathsf{E}[a_{t+1}^2 \mid \mathcal{F}_t] + \lambda\sigma_{t+1}^2\right] \\ &= \kappa\theta + (1-\kappa)\sigma_{t+1}^2 \end{aligned} \tag{7}$$

and a similar argument then shows that (7) generalizes to

$$\mathsf{E}[\sigma_{t+n}^2 \mid \mathcal{F}_t] \;=\; \kappa\theta\left[1 + (1-\kappa) + \cdots + (1-\kappa)^{n-2}\right] + (1-\kappa)^{n-1}\sigma_{t+1}^2. \tag{8}$$

It follows from (8) (why?) that $\mathbf{E}[\sigma_{t+n}^2 \mid \mathcal{F}_t] \to \theta$ as $n \to \infty$ so that θ is the best estimate of the long-term conditional variance.

Assuming we know σ_1^2 we can use (6) repeatedly together with the observed returns, i.e. the r_i's for $i = 1, \ldots, t$, and the ML estimates of κ, θ, λ and the μ_t's to compute an estimate of σ_{t+1} at time t. In practice we do not know σ_1^2 and so we typically replace it with the ML estimate of the unconditional variance, θ, or an estimate of the sample variance of the a_t's. We can then use (8) to forecast return variances n periods ahead. Forecasts of the mean return n periods ahead can be computed in a similar manner by working with the conditional mean equations of (1) and (3).

1.2 Fitting GARCH Models in *R*

When the conditional mean equation satisfies (1) and (3) and the volatility dynamics are GARCH(p,q) as given by (4) and (5) then we have an *ARMA*(m,s) / GARCH(p,q) model. This model can be fit in *R* using the *garchFit* function which is part of the *fGarch* library. The following code fragment is taken from the *R* script *Ex18.3-18.4Fig18.4.R* that is associated with Chapter 18 of Ruppert's *Statistics and Data Analysis for Financial Engineering* (2011).

A Code Fragment from the Script *Ex18.3-18.4Fig18.4.R*

```
library(fGarch)
data(bmw,package="evir")

bmw.garch_norm = garchFit(~arma(1,0)+garch(1,1),data=bmw,cond.dist="norm")
options(digits=3)
summary(bmw.garch_norm)
```

In this code fragment an *ARMA*$(1,0)$ / GARCH$(1,1)$ model was fit to *BMW* return data. Of course it's important to check how well the fitted model actually fits the data by performing various diagnostics tests. In this example Ruppert ultimately settled on an *ARMA*$(1,1)$ / GARCH$(1,1)$ model where the ϵ_t's had a t-distribution rather than the normal distribution used in the code fragment above. Typing "*? garchFit*" at the *R* prompt will provide further details on the *garchFit* function.

2 Applications to Risk Management

Time series models can be used to estimate conditional loss distributions as well as risk measures associated with these loss distributions. Towards this end, we need to be able to construct a time series of historical returns on a portfolio assuming the *current* portfolio composition. Whether or not this is possible will depend on whether historical return data is available for the individual securities in the portfolio. More generally, we need historical data for the changes in the risk factors, $\mathbf{X}$, that drive the portfolio loss distribution. This data is usually available for equities, currencies, futures and government bonds. It is also often available for exchange-traded derivatives but is generally not available for OTC securities, structured products and other exotic securities.

2.1 Estimating Risk Measures

If we assume that historical data for the changes in the risk factors, $\mathbf{X}$, are available then we can use this data to construct a time series of portfolio returns assuming the *current* portfolio composition. We can then fit a time series model to these portfolio returns and use the fitted model to estimate risk measures such as VaR or ES. Suppose for example that today is date t and we wish to estimate the portfolio VaR over the period t to $t+1$. Our fitted time series model will then take the form

$$r_{t+1} = \widehat{\mu}_{t+1} + a_{t+1} \tag{9}$$
$$a_{t+1} = \widehat{\sigma}_{t+1}\, \epsilon_{t+1} \tag{10}$$

where $\widehat{\mu}_{t+1}$ and $\widehat{\sigma}_{t+1}$ are the time t estimates of the next periods mean log-return and volatility. They are obtained of course from the MLE estimates of the fitted time series model. In the setup of Section 1, we would obtain $\widehat{\mu}_{t+1}$ and $\widehat{\sigma}_{t+1}$ from the fitted versions of (3) and (5), respectively. We could now use (9) and (10) to obtain the desired risk measures. For example, we could estimate the current period's $\widehat{\mathsf{VaR}}_\alpha$ and $\widehat{\mathsf{ES}}_\alpha$ according to

$$\widehat{\mathsf{VaR}}_\alpha^t = \widehat{\mu}_{t+1} + \widehat{\sigma}_{t+1}\, q_\alpha(\epsilon_{t+1}) \tag{11}$$
$$\widehat{\mathsf{ES}}_\alpha^t = \widehat{\mu}_{t+1} + \widehat{\sigma}_{t+1}\, \mathsf{ES}_\alpha(\epsilon_{t+1}). \tag{12}$$

The distribution of ϵ_{t+1} will have been estimated when the time-series of portfolio returns was fitted. Note that (11) and (12) are estimates of loss measures based on the *conditional* loss distribution. We therefore expect them to be considerably more accurate than estimates based on the *unconditional* loss distribution, particularly over short horizons.

Figure 1 displays the daily price level and daily returns of the S&P 500 from January 2006 to July 2011. Figure 2 then displays the daily $\mathsf{VaR}_{.99}$ violations for the S&P 500 over this period where the daily VaR was estimated using one of four possible methods: (1) historical Monte-Carlo (2) a normal approximation (3) a t_6 approximation and (4) a GARCH$(1,1)$ model. Each method used a rolling window of one year's worth of daily returns to estimate the VaR. Note that only the GARCH model attempts to estimate the conditional loss distribution. Over this time period the percentages of VaR violations for the four methods were 1.86%, 2.48%, 1.95% and 1.06%, respectively. Clearly the GARCH model performs best according to this metric. Equally importantly we see that the GARCH violations are not clustered and appear much closer to an IID sequence of Bernoulli$(p=.01)$ trials than the VaR violations of the unconditional methods. (Standard statistical tests could be used to test this observation more formally.) This is particularly noticeable during the height of the global financial crisis in the latter half of 2009. The unconditional methods were clearly underestimating VaR in this extremely volatile period.

2.2 Combining Time Series Models with Factor Models

Let $\mathbf{X}_{t+1}$ represent the changes in n risk factors between periods t and $t+1$. Suppose also that we have a factor model of the form

$$\mathbf{X}_{t+1} = \mu + \mathbf{B}\,\mathbf{F}_{t+1} + \epsilon_{t+1} \tag{13}$$

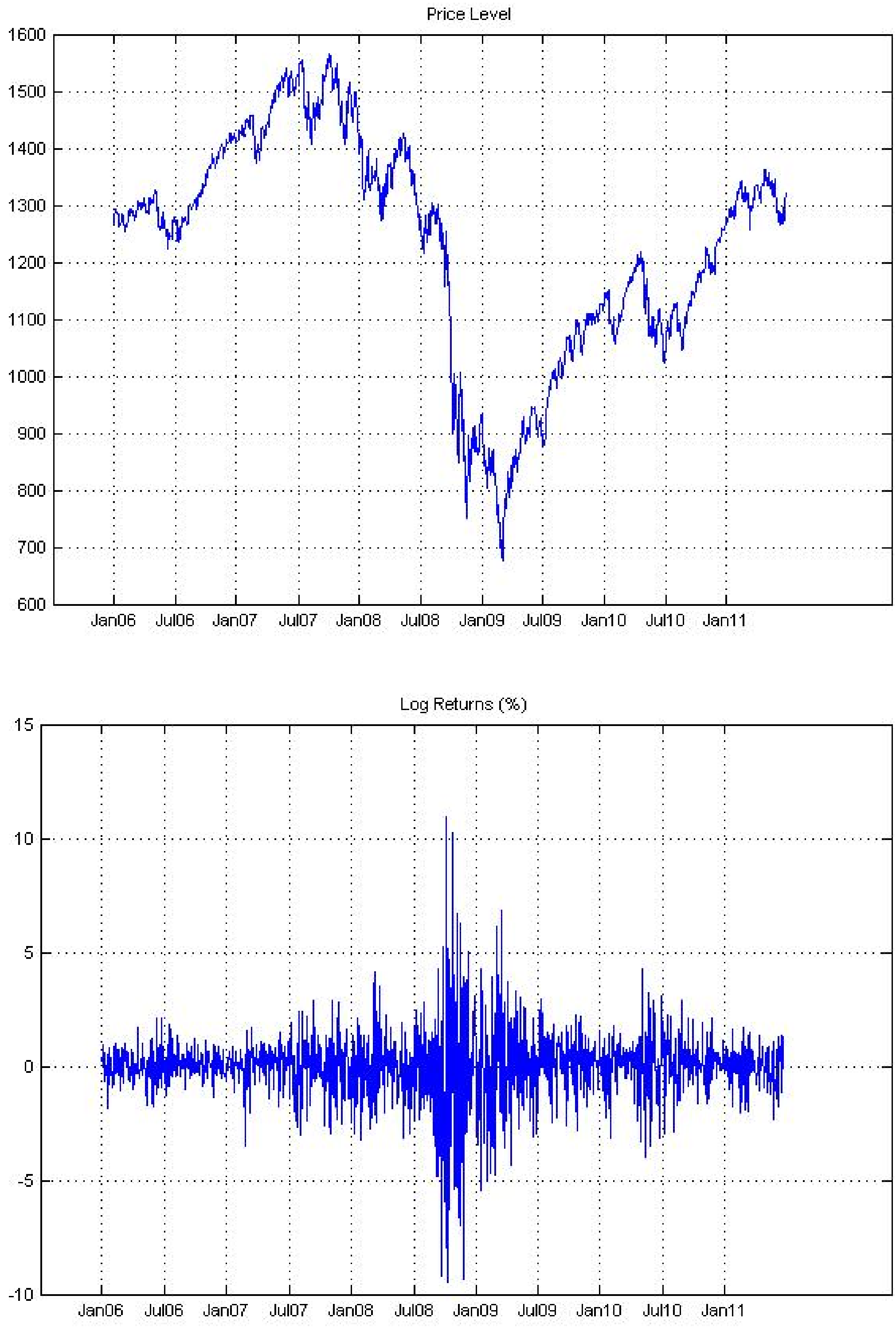

Figure 1: Price Level and Returns for the S&P 500

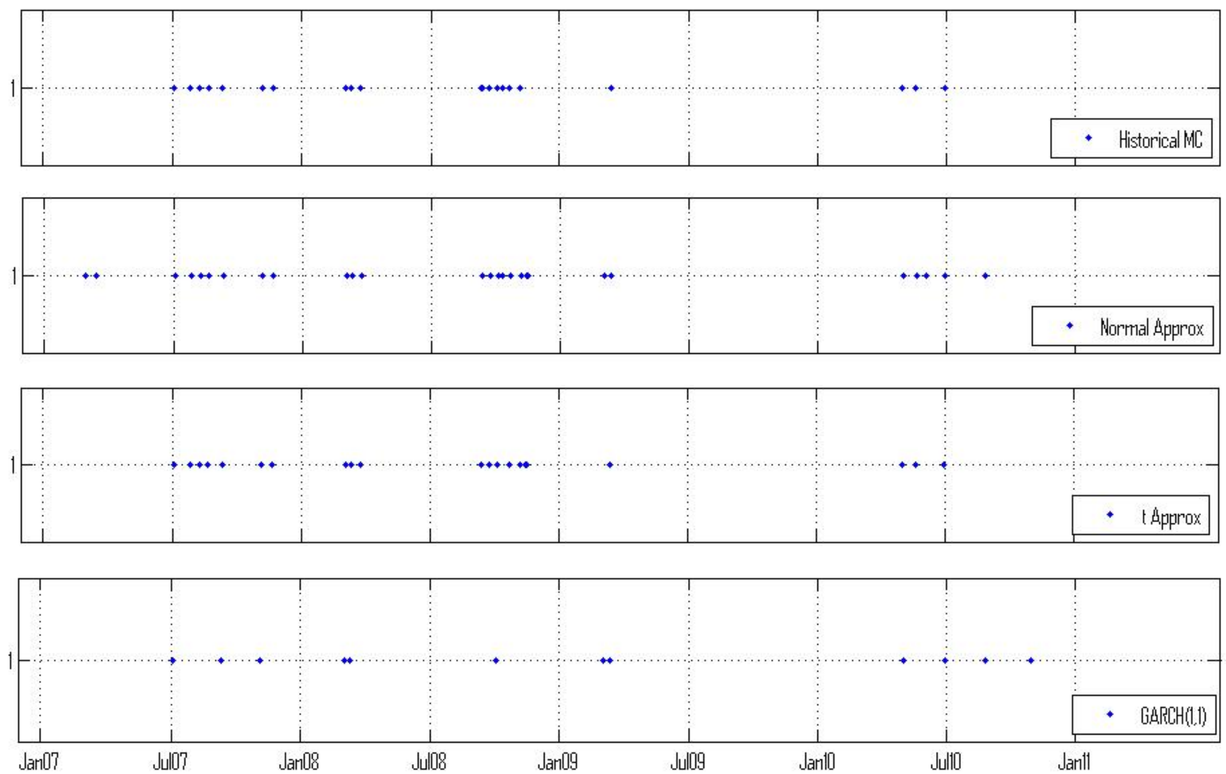

Figure 2: $\text{VaR}_{.99}$ violations for the S&P500. The first three methods (Historical MC, Normal Approx and t Approx) all estimate the VaR using a rolling window of one year's worth of daily closing prices. They are therefore largely based on approximating the **unconditional** loss distribution. In the case of the t approximation we simply set the degrees of freedom equal to 6. The fourth method (GARCH$(1,1)$) is based on approximating the **conditional** loss distribution with the VaR estimated via (11) with $\widehat{\mu}_{t+1}$ simply taken to be the mean daily return over the previous year.

where $\mathbf{B}$ is an $n \times k$ matrix of factor loadings, $\mathbf{F}_{t+1}$ is a $k \times 1$ vector of factor returns and ϵ_{t+1} is an $n \times 1$ random vector of idiosyncratic error terms which are uncorrelated and have mean zero. We also assume $k < n$, that $\mathbf{F}_{t+1}$ has a positive-definite covariance matrix and that each component of $\mathbf{F}_{t+1}$ is uncorrelated with each component of ϵ_{t+1}. All of these statements are conditional upon the information available at time t. Equation (13) is then a factor model for $\mathbf{X}$.

Given time series data on $\mathbf{X}_t$ we could simply use the factor model to compute a univariate time series of portfolio losses and then estimate risk measures as described in Section 2.1. An alternative approach, however, would be to fit separate time series models to each factor, i.e. each component of $\mathbf{X}$. We could still use the fitted time series models to estimate conditional loss measures as in (11) and (12) but we could also use the factor model to perform a scenario analysis, however. In particular, we could use the fitted time series models to provide guidance on the range of plausible factor stresses.

For example, we could use a principal components analysis to construct a factor model and then fit a separate GARCH model to the time series of each principal component. The estimated conditional variance of each principal component could then be used to determine the range of factor stresses. This is in contrast to the method of using the eigen values to determine the range of factor stresses. Since the eigen values are estimates of the *unconditional* factor variances we would generally prefer the GARCH approach to construct plausible scenarios.

2.3 Computing Dynamic Risk Measures Using Extreme Value Theory

We can also combine *extreme value theory* (EVT) with time series methods to estimate the tails of conditional loss distributions. Suppose for example, that we have fitted a model such as a GARCH or ARMA / GARCH model to a time series of portfolio returns and have used to fitted model to compute risk measures as in (11) and (12). These estimates will have been obtained assuming that the ϵ_t's are IID with some fixed distribution such as the normal or t distribution.

We can use the fitted time series to compute the residuals, $\widehat{\epsilon}_t$, which should be approximately IID if the fitted time series model is a good fit. We can then apply EVT to the fitted residuals. In particular, we can fit the *generalized Pareto distribution* (GPD) to the tails of the fitted residuals and estimate the corresponding risk measures to again obtain

$$\widehat{\mathsf{VaR}}_\alpha^t = \widehat{\mu}_{t+1} + \widehat{\sigma}_{t+1}\, q_\alpha(\epsilon_{t+1}^{\mathsf{GPD}}) \tag{14}$$

$$\widehat{\mathsf{ES}}_\alpha^t = \widehat{\mu}_{t+1} + \widehat{\sigma}_{t+1}\, \mathsf{ES}_\alpha(\epsilon_{t+1}^{\mathsf{GPD}}) \tag{15}$$

where the superscript *GPD* in (14) and (15) is used to emphasize that the ϵ_t's (or their tails), follow a GPD distribution. Note that there appears to be an inconsistency here in that the original time series model was fitted using one set of assumptions for the ϵ_t's, i.e. that they are normally or t distributed, and that a different assumption is used in (14) and (15), i.e. that they have a GPD distribution. This does not present a problem due to the theory of *quasi-maximum likelihood estimation (QMLE)* which effectively states that $\widehat{\mu}_{t+1}$ and $\widehat{\sigma}_{t+1}$ are still *consistent* estimators of μ_{t+1} and σ_{t+1} even though the distributions of the ϵ_t's were misspecified.

Section 7.2.6 (which refers to results in Section 2.3.6) of *Quantitative Risk Management* by *McNeil, Frey and Embrechts* describes several numerical experiments that compare different methods for estimating VaR or ES. They conclude that methods based on GARCH-EVT, i.e. estimators such as (14) and (15), are most accurate. Note that we could also have used the *Hill estimator* as an alternative to the GPD distribution when estimating $\widehat{\mathsf{VaR}}_\alpha^t$ and $\widehat{\mathsf{ES}}_\alpha^t$. Section 19.6 of *Ruppert* provides some examples and should be consulted for further details.