Leture 3: Univariate Volatility Modelling - Part II

FM321: Risk Management and Modelling

Linyan Zhu

11 October 2022

LSE Finance

Volatility modelling

- Univariate volatility modelling (one financial asset)
  - Moving average models
  - ARCH
  - GARCH
  - Model estimation
  - Diagnostics
  - Alternative approaches
- Multivariate volatility modelling

- GARCH: Generalized Autoregressive Conditional Heteroskedasticity

- The defining equation of a GARCH(1,1) process is

$$\sigma_t^2 = \omega + \alpha r_{t-1}^2 + \beta \sigma_{t-1}^2$$

- The general specification GARCH(P,Q) allows for an arbitrary number of lags in both terms:

$$\sigma_t^2 = \omega + \sum_{p=1}^{P} \alpha_p r_{t-p}^2 + \sum_{q=1}^{Q} \beta_q \sigma_{t-q}^2$$

- GARCH(P,Q):

$$\sigma_t^2 = \omega + \sum_{p=1}^{P} \alpha_p \epsilon_{t-p}^2 + \sum_{q=1}^{Q} \beta_q \sigma_{t-q}^2$$

- α_p captures the effect of news on conditional variance.
- β_q correspond to memory, and they lead to persistence in volatility estimates.
- Parameter restrictions: $\omega > 0$, $\alpha_p \geq 0$, $\beta_q \geq 0$, and at least one of the α_p's and β_q's is strictly positive.

GARCH(1,1): unconditional variance

- Assume that r_t is a GARCH(1,1) process with finite unconditional variance σ^2.

$$\sigma_t^2 = \omega + \alpha r_{t-1}^2 + \beta \sigma_{t-1}^2$$

- Take the unconditional expectation of the above equation,

$$E[\sigma_{t+1}^2] = \omega + \alpha E[r_t^2] + \beta E[\sigma_t^2]$$

- Recall that $E[r_t^2] = E[\sigma_t^2] = \sigma^2$.

$$\sigma^2 = \omega + \alpha \sigma^2 + \beta \sigma^2$$

- Solve for σ^2,

$$\sigma^2 = \frac{\omega}{1 - \alpha - \beta}$$

- It is necessary to require $\alpha + \beta < 1$ in order for the unconditional variance to exist.

- Similarly, for a GARCH(P,Q) process with finite unconditional variance σ^2:

$$\sigma_t^2 = \omega + \sum_{p=1}^{P} \alpha_p r_{t-p}^2 + \sum_{q=1}^{Q} \beta_q \sigma_{t-q}^2$$

we have:

$$\sigma^2 = \omega + \sum_{p=1}^{P} \alpha_p \sigma^2 + \sum_{q=1}^{Q} \beta_q \sigma^2$$

- Thus, the unconditional variance is

$$\sigma^2 = \frac{\omega}{1 - \sum_{p=1}^{P} \alpha_p - \sum_{q=1}^{Q} \beta_q}$$

- In order for the unconditional variance to exist, we need

$$\sum_{p=1}^{P} \alpha_p + \sum_{q=1}^{Q} \beta_q < 1$$

- What would you expect σ^2_{t+k} to be when there is a shock to σ^2_{t+1}?
- For a GARCH(1,1) process, we have

$$\sigma^2_{t+k} = \omega + \alpha r^2_{t+k-1} + \beta \sigma^2_{t+k-1}$$

- Take conditional expectations at time t,

$$\begin{aligned} E_t[\sigma^2_{t+k}] &= \omega + \alpha E_t[r^2_{t+k-1}] + \beta E_t[\sigma^2_{t+k-1}] \\ &= \omega + (\alpha + \beta) E_t[\sigma^2_{t+k-1}] \end{aligned}$$

- Recall that $\sigma^2 = \frac{\omega}{1-\alpha-\beta}$. Use it to replace ω above,

$$E_t[\sigma^2_{t+k}] = \sigma^2(1 - \alpha - \beta) + (\alpha + \beta) E_t[\sigma^2_{t+k-1}]$$

- Rearrange,

$$E_t[\sigma^2_{t+k}] - \sigma^2 = (\alpha + \beta) E_t \left[[\sigma^2_{t+k-1}] - \sigma^2 \right]$$

- Iterate backward to the current period,

$$\begin{aligned} E_t[\sigma^2_{t+k}] - \sigma^2 &= (\alpha+\beta)E_t\left[[\sigma^2_{t+k-1}] - \sigma^2\right] \\ &= (\alpha+\beta)^2 E_t\left[[\sigma^2_{t+k-2}] - \sigma^2\right] \\ &\cdots \\ &= (\alpha+\beta)^{k-1} E_t\left[[\sigma^2_{t+1}] - \sigma^2\right] \end{aligned}$$

- By definition, $E_t[\sigma^2_{t+1}] = \sigma^2_{t+1}$, so

$$E_t[\sigma^2_{t+k}] - \sigma^2 = (\alpha+\beta)^{k-1}\left[\sigma^2_{t+1} - \sigma^2\right]$$

- Given $\alpha + \beta < 1$, we can see that conditional variance forecasts are mean reverting to σ^2.

- One can show that the unconditional kurtosis of a GARCH(1,1) process is given by

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

- We need $(\alpha+\beta)^2 + 2\alpha^2 < 1$ for this to be finite. If it is, the process exhibits unconditionally heavy tails.

- GARCH carries the same benefits of ARCH, and in addition leads to more stable volatility estimates.

- This comes at a price in terms of an increase in complexity of the estimation algorithms, but usually this isn't serious.

- In fact, since one can typically use much shorter lag lengths with GARCH, the number of parameters is smaller and the estimation is faster.

- We use the maximum likelihood to estimate the model parameters based on the data.

- Maximum likelihood asks the following question: which set of values for the parameters is more likely to have generated the observed data?

- Example: suppose we have the following observed i.i.d. sample from a normally distributed random variable:

$$-0.2, -0.2, -0.1, 0, 0, 0.2, 0.2, 0.3$$

Which of the following sets of parameters is more likely to have led to this sample being observed?

	Mean	Std. Dev.
A	0	0.2
B	0	1
C	2	0.2

- Normal density: recall that, if X is a random variable with a normal distribution $N(\mu, \sigma^2)$, its density is given by

$$f_X(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left[-\frac{(x-\mu)^2}{2\sigma^2}\right]$$

- How can we estimate (μ, σ^2) in the previous example?

- Now suppose we have four return observations, r_0, r_1, r_2, and r_3. How can we estimate ARCH(1) given $z_t \sim N(0,1)$?

- Recall that $r_t = \sigma_t z_t$, where $z_t \sim N(0,1)$ and $\sigma_t^2 = \omega + \alpha r_{t-1}^2$.
- The conditional density for one observation is given by

$$f(r_t | r_{t-1}) = \frac{1}{\sqrt{2\pi(\omega + \alpha r_{t-1}^2)}} \exp\left[-\frac{r_t^2}{2(\omega + \alpha r_{t-1}^2)}\right]$$

- The joint density of the observations except the first one is given by

$$\prod_{t=2}^{T} f(r_t | r_{t-1})$$

- This expression needs to be maximized with respect to ω and α.

ARCH(1): log likelihood

- There are a number of theoretical reasons to work with log-likelihood instead of the original likelihood.

- For ARCH(1), this means

$$log\mathcal{L} = -\frac{T-1}{2}log(2\pi) - \frac{1}{2}\sum_{t=2}^{T} log(\omega + \alpha r_{t-1}^2) - \frac{1}{2}\sum_{t=2}^{T}\left(\frac{r_t^2}{\omega+\alpha r_{t-1}^2}\right)$$

- Suppose we're trying to estimate a single parameter θ, which belongs to a parameter space Θ.

- Typically, no closed-form solution exists, which implies we need to use a numerical algorithm to find the optimal value of θ.

- Often, black boxes exist to do so (say, in R or Matlab).

- Numerical problems often arise in this process, rendering the solution uncertain.

- Suppose we're trying find θ that maximizes $f(\theta)$.

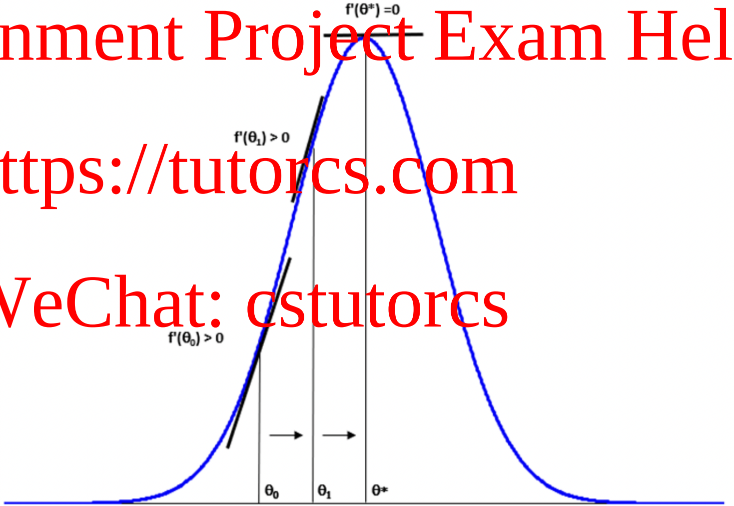
- The general process works as follows:

  - choose an initial value θ^0 for the parameters

  - compute $f'(\theta^0)$
  - if $f'(\theta^0) > 0$, should increase θ; if $f'(\theta^0) < 0$, should decrease θ

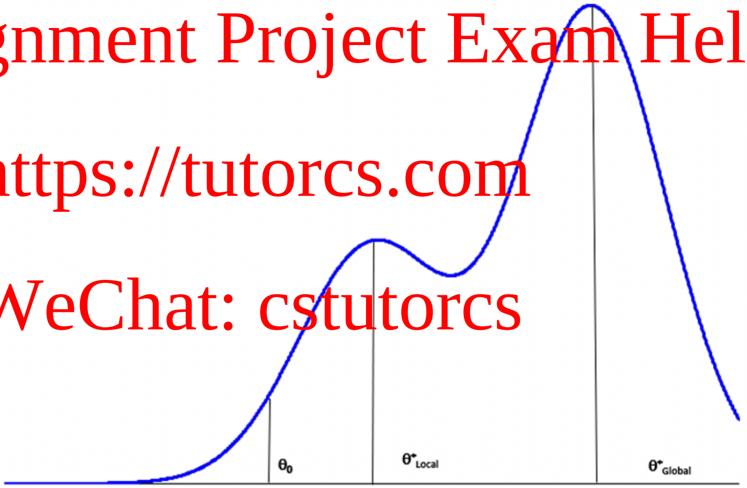
  - increment θ^0 accordingly (details depends on algorithm) to find θ^1
  - iterate until convergence

- With a smooth function that has a clear unique global maximum, the process usually works well:

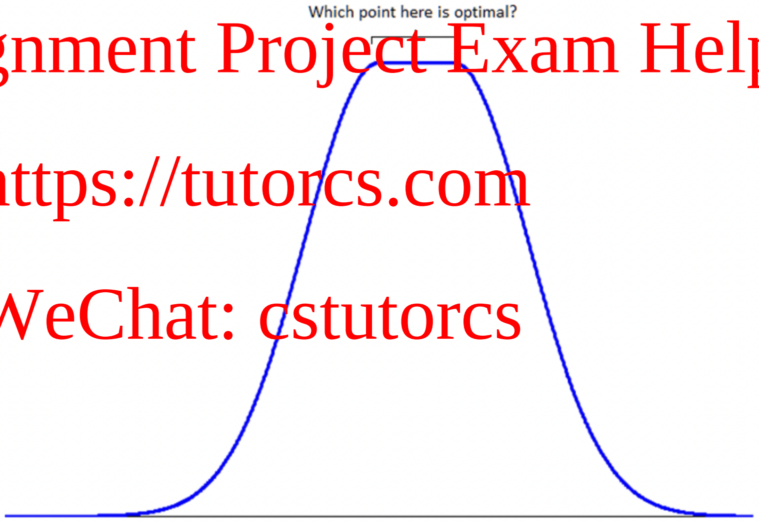


Optimization in one dimension: issues

- If local maxima are not unique, problems can arise:



- If the objective function is flat near the maximum, we can also have problems.



Numerial optimization considerations

- There are no general solutions to these types of issues.

- The simpler the model, the less likely you'll run into problems (GARCH of low orders is usually OK)

- For multivariate models (next topic), problems are common.

- This problem makes it particulary important to make sure that estimates are sensible.

- There are a number of tools one can use to evaluate models:

  - Parameter tests

  - Likelihood ratios

  - Residual analysis

- Consider two nested models (that is, one is a restricted version of the other)

- Example: ARCH(1) is a restricted version of GARCH(1,1) in which $\beta = 0$.

- Which specification is better?

- Estimation output usually produces a test statistic (valid asymptotically) that can be used to test the significance of individual parameters.

Likelihood ratio tests

- Denote the likelihood ratio values for the restricted and unrestricted models $\mathcal{L}_R$ and $\mathcal{L}_U$, respectively.

- We always have $\mathcal{L}_R \leq \mathcal{L}_U$; if the constraints actually hold in the data, we would have $\mathcal{L}_R = \mathcal{L}_U$.

- Even if the constraint holds for the parameters, in the sample we can observe a discrepancy due to statistical variation.

- The difference between $\mathcal{L}_R$ and $\mathcal{L}_U$ in the sample can be used as the basis for a statistical test of the validity of the hypothesis.

- Under the null hypothesis that the constraints hold, one can show that asymptotically

$$LR = -2(\log \mathcal{L}_R - \log \mathcal{L}_U) \sim \chi^2_{(k)}$$

where k is the number of constraints involved in the null hypothesis.

- For example, if we're comparing ARCH(1) with GARCH(1,1), then $k = 1$ (as the only constraint is $\beta = 0$).

- Result holds asymptotically, so the more data we have the more precise the test is.

Residual analysis

- We usually make a distributional assumption to be able to estimate our models.

- For instance, standardized residuals are normal.

- Once we estimate parameters, we can compute the model's estimates for conditional variance.

- For instance, with GARCH(1,1) we have:

$$\hat{\sigma}_1^2 = \frac{\hat{\omega}}{1 - \hat{\alpha} - \hat{\beta}}$$

$$\hat{\sigma}_t^2 = \hat{\omega} + \hat{\alpha} r_{t-1}^2 + \hat{\beta} \hat{\sigma}_{t-1}^2 \text{ for } t = 2, 3, \ldots, T$$

- With a series for $\hat{\sigma}_t^2$, we can compute the model's standardized residuals:

$$\hat{z}_t = \frac{r_t}{\hat{\sigma}_t}$$

- Question: does $\{\hat{z}_t\}$ satisfy our distributional assumptions?

- Can apply the methods from Lecture 1 to evaluate this question:
  - Ljung-Box test for serial correlation

  - Jarque-Bera test for normality
  - QQ-Plots can help determine the distribution of $\{\hat{z}_t\}$

Extensions of GARCH

- GARCH models tend to produce reasonable estimates for conditional volatility, and for this reason it is often used as a workhorse model.

- GARCH has been extended in many different ways, to account for a number of economic, financial and statistical effects.

- We'll go over some of those extensions below.

- Non-Normal GARCH: Instead of assuming that z_t is conditionally normal, we can adopt other specifications.

- For example, we could have $z_t \sim t_\nu$.

- The parameter ν would need to be estimated, which typically requires lots of observations.

- Note that we assume that $E_{t-1}[z_t^2] = 1$; since the Student-t distribution does not have unit variance, the z_t is a normalized version of that distribution.

- GARCH-in-Mean: accounts for the possibility that in times of higher risk expected returns may be higher.

- In some cases, the zero mean assumption isn't appropriate.

- One possible specification is

$$
\begin{aligned}
r_t &= \mu_t + \sigma_t z_t \\
\sigma_t^2 &= \omega + \alpha r_{t-1}^2 + \beta \sigma_{t-1}^2 \\
\mu_t &= \delta \sigma_t^2
\end{aligned}
$$

where δ is an additional parameter to estimate.

Extensions of GARCH: GJR-GARCH

- Leverage Effects: in many settings, it is observed that positive and negative shocks have different impact on conditional volatility.

- Effectively, this means that the ARCH term should read

$$
\begin{array}{ll}
\alpha r_{t-1}^2 & \text{if } r_{t-1} > 0 \\
\alpha' r_{t-1}^2 & \text{if } r_{t-1} < 0
\end{array}
$$

  where $\alpha' > \alpha > 0$.

- Using the indicator function $\mathcal{I}_{[r_{t-1}<0]}$ (which equals 1 if $r_{t-1} < 0$ and 0 otherwise), we can write the model as

$$\sigma_t^2 = \omega + \left(\alpha + \gamma \mathcal{I}_{[r_{t-1}<0]}\right) r_{t-1}^2 + \beta \sigma_{t-1}^2$$

  where $\gamma = \alpha' - \alpha$.

- This specification is known as the GJR-GARCH model.

Model choice

- There is no cookbook recipe for how to choose models.
- Researchers make choices of models based on:
  - Statistical considerations
  - Economic performance characteristics
  - Knowledge of factors outside the model that may be relevant
- When choosing a model, bear in mind that:
  - Models with more parameters or more complex specification are harder to estimate, and require more data.
  - More parameters usually imply lower precision in parameter estimates, and more model uncertainty.
  - Data from a long time ago isn't as representative of how markets operate today.