Adrian Beer
1997304

Exam B

Task B.1

Let $\{x_t\}_{t \in \{1,\dots,T\}}$ be the returns of the Apple stocks.

Assuming they follow a AR(1)-T-GARCH(2,2) the data generating process is

$$x_{t+1} = \hat{x}_{t+1} + \cancel{\hat{\sigma}_{t+1}} \, \varepsilon_{t+1}$$

where we model $\hat{x}_{t+1} := \phi_0 + \phi_1 x_t$, i.e. the conditional mean of x_{t+1} $\mathbb{E}[x_{t+1} \mid \mathcal{F}_t]$ as an AR(1)-process. $\mathcal{F}_t$ is the information set available at time t.

We model the conditional variance of ε_{t+1}, $\mathbb{V}(\varepsilon_{t+1} \mid \mathcal{F}_t)$ with the T-GARCH(2,2), i.e.

$\varepsilon_{t+1} = \hat{\sigma}_{t+1} \eta_{t+1}$ where η_{t+1} follows a white noise process, i.e. WNP(0,1) with variance of 1 and mean 0 and.

$$\hat{\sigma}^2_{t+1} = \mathbb{V}(\varepsilon_{t+1} \mid \mathcal{F}_t) =$$

$$\alpha_0 + \left(\alpha_1 + \theta_1 \mathbb{1}_{(\varepsilon_t < 0)}\right) \cdot \varepsilon_t^2 + \left(\alpha_2 + \theta_2 \mathbb{1}_{(\varepsilon_{t-1} < 0)}\right) \varepsilon_{t-1}^2 \quad (*)$$

$$+ \beta_1 \hat{\sigma}_t + \beta_2 \hat{\sigma}_{t-1}.$$

$\mathbb{1}$ is the indicator function: $\mathbb{1}_{(\varepsilon_t < 0)} = \begin{cases} 1 & \varepsilon_t < 0, \\ 0 & \text{sonst}. \end{cases}$

We can assume that the ~~mean~~ cond. mean of ε_{t+1} is 0, because of the AR(1) term $\hat{x}_{t+1}$.

Adrian Beer
1997304

Now for the joint log-likelihood function:

We need to assume some distribution of η. Let f be the probability density function of η. It must be a WNP(0,1).

The following describes the log-likelihood function $l(\lambda|X)$ where $X = \{x_t\}_{t\in\{1,..,T\}}$ and λ is the parameter vector which we want to optimize with a maximum likelihood estimation (MLE). The parameter vector λ consists of

$$(\phi_0, \phi_1, \theta_1, \theta_2, \alpha_1, \alpha_2, \beta_1, \beta_2, \underbrace{\hat{\sigma}_2, \hat{\sigma}_3}_{\text{these are needed because of } (*)})$$

$$\cancel{\{\hat{x}\}_t} \quad \{\hat{x}_t\}_{t\in\{2,..,T\}} = \phi_0 + \phi_1 \{x_t\}_{t\in\{1,...,T-1\}},$$

so in vector notation

$$\underset{\cancel{\text{T-1 vector}}\ (T-1\times 1)}{\hat{x}} = \underset{(T-1\times 1)}{\phi_0} + \underset{\text{scalar}}{\phi_1} \underset{(T-1\times 1)}{X_{[1:T-1]}},$$

We can now obtain $\varepsilon = \hat{x} - X_{[1:T-1]}. = \{\varepsilon_t\}_{t\in\{2,..,T\}}$

For the T-GARCH we need a for loop aswell as initial estimates for $\hat{\sigma}_2^2, \hat{\sigma}_3$. (*)

for $t = 4, ..., T$ (the first two epsilons are available only at $t=2, t=3$ and they are required for the calculation)

$\hat{\sigma}_{t+1}^2 = \quad \cdots \quad$ see (*)

Then $l(\lambda|X) = \sum_{t=4}^{T} \log f\left(\frac{x_t - \hat{x}_t}{\hat{\sigma}_t}\right)$ is the log likelihood function.

We can then maximize $l(\lambda|X)$ or minimize $-l(\lambda|X)$ w.r.t. λ using e.g. gradient descent or simulated annealing or some other algorithm for finding minima. We can restrict search space by imposing restrictions on the search space, e.g. finite unconditional mean $\phi_1 \overset{!}{<} 1$ or positive unconditional mean: $\phi_1 > 0$ & $\phi_0 > 0$. This increases the speed of convergence.

Adrian Beer
1997304

We can obtain the time series of return variance of apple via $\{\sigma_t^2\}_{t\in\{2,\dots,T\}} = \{\varepsilon_t^2\}_{t\in\{2,\dots,T\}}$ since

$$\mathbb{E}[\varepsilon_t] = 0,\ \mathbb{V}(\varepsilon_t) = \mathbb{E}[\varepsilon_t^2] - \mathbb{E}[\varepsilon_t]^2 = \mathbb{E}[\varepsilon_t^2].$$

If our model is a good fit, then $\hat{\sigma}_t$ and σ_t should not deviate a lot. Furthermore we can check if ~~we decorrelated the variance~~ $\{\sigma_t^2 - \hat{\sigma}_t^2\}_{t\in\{2,\dots,T\}}$ still has any serial correlation with e.g. Ljung-Box test. If we modeled the heteroscedasticity correctly there should be none.

Adrian Beer
1087304

2.2. Task B.2

Let $\{x_t\}_{t \in \{1,\dots,T\}} = \begin{bmatrix} x_1 \\ \vdots \\ x_T \end{bmatrix}$ be the returns. When it follows a Gaussian ARMA(1,1) - GARCH(1,1) model the DGP is

$$x_t = \hat{x}_t + \varepsilon_t \quad \text{where}$$

$$\hat{x}_t = \phi_0 + \phi_1 x_{t-1} + \theta_1 \varepsilon_{t-1} = \mathbb{E}[x_t \mid \mathcal{F}_{t-1}] \quad \text{and}$$

$$\varepsilon_t = \hat{\sigma}_t \eta_t \quad \text{where} \quad \eta \sim \mathcal{N}(0,1) \text{ and}$$

$$\hat{\sigma}_t^2 = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \beta_1 \hat{\sigma}_{t-1}^2 = \mathbb{V}(\varepsilon_t \mid \mathcal{F}_{t-1}).$$

ε_{t-1}^2 is the realized variance at time $t-1$: σ_{t-1}^2.

It follows that $x_t \mid \mathcal{F}_{t-1} \sim \mathcal{N}(\hat{x}_t, \hat{\sigma}_t^2)$ so the joint log-likelihood can be written as

$$\ell(\lambda \mid X) = \sum_{t=3}^{T} \log f(x_t \mid \lambda, \mathcal{F}_{t-1}) + \cancel{\log f(x_1 \mid \lambda) + f(\hat{x}_2 \mid \lambda)}$$

will not use x_1, because we would have to estimate x_0 and ε_0, etc. too many additional parameters

where $f(x_t \mid \lambda, \mathcal{F}_{t-1}) = \frac{1}{\sqrt{2\pi}\,\hat{\sigma}_t} \cdot e^{-\frac{(x_t - \hat{x}_t)^2}{2\hat{\sigma}_t^2}}$ is the probability density function of x_t. The first two observations are special because no ~~$\hat{\sigma}_t$~~ conditional variance estimation & cond. mean can be calculated for these observations, because we lack the preceding data necessary.* We can either include $\hat{\sigma}_1^2$, ε_1 in the parameter vector λ, so that they will be estimated or use the unconditional ~~mean~~ values, i.e. $\hat{\sigma}_1^2 = \mathbb{V}(\sigma_\varepsilon^2) = \frac{\alpha_0}{1-\alpha_1-\beta_1}$ and $\mathbb{E}[\varepsilon] = 0 = \varepsilon_1$. This way we can „salvage" x_2 and use it for the MLE.

* namely ε_1 and $\hat{\sigma}_1^2$

We can ignore and not use $f(x_1)$ and $f(x_2)$ in the $\ell(\lambda)$ function at all, if we have enough data, however we need estimates for ε_1 and $\hat{\sigma}_1^2$, because of their recursive nature/definition.

Adrian Beer
1097304

Conditional return density 2-step ahead. For this we need $\mathbb{E}[x_{t+2} | \mathcal{F}_t]$ and $\mathbb{V}(x_{t+2} | \mathcal{F}_t)$.

First the conditional mean:

$$\mathbb{E}[x_{t+2} | \mathcal{F}_t] = \mathbb{E}[\mathbb{E}[x_{t+2} | \mathcal{F}_{t+1}] | \mathcal{F}_t] \quad \text{(law of iterated expectations)}$$

$$= \mathbb{E}[\phi_0 + \phi_1 x_{t+1} + \theta_1 \varepsilon_{t+1} + \underbrace{\varepsilon_{t+2}}_{\mathbb{E}[\varepsilon_{t+2}|\mathcal{F}_t]=0} | \mathcal{F}_t]$$

$$= \mathbb{E}[\phi_0 + \phi_1(\phi_0 + \phi_1 x_t + \theta\varepsilon_t + \varepsilon_{t+1}) + \theta_1 \varepsilon_{t+1} + \varepsilon_{t+2} | \mathcal{F}_t]$$

$$= \phi_0 + \phi_1\phi_0 + \phi_1^2 x_t + \theta\varepsilon_t + 0 + \theta_1 \cdot 0 + 0 = \phi_0(1+\phi_1) + \phi_1^2 x_t + \theta\varepsilon_t$$

because x_t and ε_t are known in $\mathcal{F}_t$ and $\mathbb{E}[\varepsilon_{t+h} | \mathcal{F}_t] = 0 \ \forall h$.

~~Now the conditional variance:~~

~~$\mathbb{V}[x_{t+2} | \mathcal{F}_t] = \mathbb{E}[(x_{t+2} - \mathbb{E}[x_{t+2}])^2]$ with the law of iterated expec. and the calculation from above we arrive at $\mathbb{V}(x_{t+2} | \mathcal{F}_t) = \mathbb{E}[((\theta_1 + \phi_1)\varepsilon_{t+1} + \varepsilon_{t+2})^2]$ because the variance of everything known at time t is zero.~~

~~Since~~ $\varepsilon_{t+1} \perp \varepsilon_{t+2}$

We model the conditional variance with GARCH, so

follows from above & law of iterated expect.

$$\mathbb{V}_t(x_{t+2} | \mathcal{F}_t) = \cancel{\mathbb{V}(\varepsilon_t | \mathcal{F}_t)} \ \mathbb{E}_t[(x_{t+2} - \mathbb{E}[x_{t+2}])^2] = \mathbb{E}_t[(\phi_1\varepsilon_{t+1} + \theta\varepsilon_{t+1} + \varepsilon_{t+2})^2]$$

$$\cancel{= \mathbb{E}[\alpha_0 + \alpha_1 \varepsilon_{t+1}^2 + \beta_1 \hat{\sigma}_{t+1}^2 | \mathcal{F}_t] \quad \text{(GARCH)}}$$

$$\cancel{= \mathbb{E}_t[\alpha_0 + \alpha_1 \varepsilon_{t+1}^2 + \beta_1(\alpha_0 + \alpha_1 \varepsilon_t^2 + \beta_1 \hat{\sigma}_t^2)]}$$

$$\cancel{= \alpha_0 + \underbrace{\alpha_1\alpha_0 + \alpha_1^2\varepsilon_t^2 + \alpha_1\beta_1\hat{\sigma}_t^2}_{} + \beta_1\alpha_0 + \beta_1\alpha_1\varepsilon_t^2 + \beta_1^2\hat{\sigma}_t^2}$$

$$\cancel{\text{because } \mathbb{E}[\varepsilon_{t+1}^2 | \mathcal{F}_t] = \mathbb{V}(\varepsilon_{t+1} | \mathcal{F}_t) = \alpha_0 + \alpha_1\varepsilon_t^2 + \beta_1\hat{\sigma}_t^2 \quad \text{(GARCH)}}$$

$$= \mathbb{E}_t[\phi_1^2\varepsilon_{t+1}^2 + 2\phi_1\theta_1\varepsilon_{t+1}^2 + \theta^2\varepsilon_{t+1}^2 + \underbrace{\varepsilon_{t+2}^2}_{}] \quad \text{because } \varepsilon_{t+1} \perp \varepsilon_{t+2}$$

$\boxed{\varepsilon_{t+1} \perp \varepsilon_{t+2}}$

$$\mathbb{E}_t[\mathbb{E}_{t+1}[\varepsilon_{t+2}^2]] = \mathbb{E}_t[\alpha_0 + \alpha_1\varepsilon_{t+1}^2 + \beta_1\hat{\sigma}_{t+1}^2]$$

$$\mathbb{E}_t[\hat{\sigma}_{t+1}^2] = \mathbb{E}_t[\varepsilon_{t+1}^2]$$

Adrian Beer
1897304

$$\Rightarrow \mathbb{E}_t[\varepsilon_{t+2}^2] = \mathbb{E}_t[\alpha_0 + (\alpha_1 + \beta_1)\varepsilon_{t+1}^2]$$

since $\mathbb{E}_t[\varepsilon_{t+1}] = \alpha_0 + \alpha_1 \varepsilon_t^2 + \beta_1 \hat{\sigma}_t^2$ the equation becomes:

$$\mathbb{V}_t(x_{t+2}) = \phi_1^2(\alpha_0 + \alpha_1\varepsilon_t^2 + \beta_1\hat{\sigma}_t^2) + 2\phi_1\theta_1(\alpha_0 + \alpha_1\varepsilon_t^2 + \beta_1\hat{\sigma}_t^2)$$
$$+ \theta^2(\alpha_0 + \alpha_1\varepsilon_t^2 + \beta_1\hat{\sigma}_t^2) + \alpha_0 + \alpha_1\varepsilon_t^2 + \beta_1\hat{\sigma}_t^2$$

$$\Rightarrow (\phi_1^2 + 2\phi_1\theta_1 + \theta^2 + 1)(\alpha_0 + \alpha_1\varepsilon_t^2 + \hat{\sigma}_t^2)$$

Having determined $\underbrace{\mathbb{V}_t(x_{t+2})}_{}$ (Σ:) and $\mathbb{E}_t(x_{t+2})$ (μ:) we can write the conditional return density (2-step ahead) as

$$f(x_{t+2} \mid \mathcal{F}_t) = \frac{1}{\sqrt{\pi 2 \Sigma}} \cdot e^{-\frac{(x_{t+2}-\mu)^2}{2\Sigma}}.$$

Note: for the calculation of the conditional variance we can neglect η, because its variance is always assumed to be 1.