# 1 Problem Set 6 - Problem 1)

We know that the log likelihood of a DCC model with Gaussian shocks nicely factorizes in two parts: one for the volatilities and one for the correlation, see lecture 10. However, we also know that the assumption of joint Gaussianity for financial returns is too restrictive in practice. Can you state three reasons why this is true? Use formal derivations to support your arguments.

First of all we have that in general to assume joint Gaussianity for financial returns is very restrictive in practice. This for example due to financial returns often is not centered and further returns are often skewed differently from the normal distribution, where skewness is equal to 0.

Joint Gaussianity implies implies that each marginal is Gaussian. Specifically if:

$$
\begin{pmatrix} x \\ y \end{pmatrix} \sim N \left( \begin{pmatrix} \mu_x \\ \mu_y \end{pmatrix}, \begin{pmatrix} \sigma_x^2 & \sigma_{xy} \\ \sigma_{xy} & \sigma_0^2 \end{pmatrix} \right)
$$

we have that $x \sim N\left(\mu_x, \sigma_x^2\right)$ and $Y \sim N\left(\mu_x, \sigma_y^2\right)$

- Kurtosis: $x$ and $y$ does not exhibit excess of kurtosis, i.e. $\frac{\mathbb{E}\left[(x-\mu_x)^4\right]}{\mathbb{E}\left[(x-\mu_x)^2\right]^2} = \frac{\mathbb{E}\left[(y-\mu_3)^4\right]}{\mathbb{E}\left[(y-\mu_3)^2\right]^2} = 3$. This is because the kurtosis of any univariate normal distribution is 3, which is very restrictive assumption.

  - Kurtosis is a statistical measure that is used to describe the size of the tails on a distribution. We know that for financial returns the tails will often contain more mass in the tails compared to the normal distribution.
  - Kurtosis is often as a measure of financial risk within finance.

- $x$ and $y$ do not exhibit positive/negative skewness, i.e. $\frac{\mathbb{E}\left[(x-\mu_x)^3\right]}{\mathbb{E}\left[[(x-\mu_x)^2\right]^{3/2}} = \frac{\mathbb{E}\left[(y-\mu_y)^3\right]}{\mathbb{E}\left[(y-\mu_y)^2\right]^{3/2}} = 0$.

  - This is because that we have for the normal distribution that skewness is equal to 0. Again, this is a very restrictive assumption, especially when modelling financial returns. At least for stocks we will often observe right(positive) skewness for returns.

- The third one is concerning the dependence structure when dealing with joint Gaussian assumption. This means that we have no tail dependence, i.e. $\lim_{z \to -\infty} P(x \leq z \mid y \leq z) = \lim_{z \to \infty} P(x \geq z \mid y \geq z) = 0$, which in most cases for financial returns are not valid. Within financial returns we observe empirically that extreme returns of asset $x$ will influence the return of asset $y$. This is seen in times of financial crises and booms.

A possibility to depart from the Gaussian assumption is to assume that returns, conditionally on past observations, are multivariate Student's $t$ distributed. The probability density function of a $p$-dimensional Student's $t$ distribution with location vector $\mu$, scale matrix $\Psi$, and $\nu$ degrees of freedom is given by:

$$
p\left(\mathbf{y}; \boldsymbol{\mu}, \Psi\right) = \frac{\Gamma\left(\frac{\nu+p}{2}\right)}{\Gamma\left(\frac{\nu}{p}\right) \pi^{p/2} \nu^{p/2} |\Psi|^{1/2}} \left(1 + \frac{(\mathbf{y} - \boldsymbol{\mu})' \Psi^{-1} (\mathbf{y} - \boldsymbol{\mu})}{\nu}\right)^{-\frac{\nu+p}{2}}
$$

where $\Gamma(\cdot)$ is the gamma function. The covariance matrix $\Sigma$ is related to the scale by $\Sigma = \Psi \frac{\nu}{\nu-2}$ and can be factorized as $\mathbf{D}^{1/2}\mathbf{R}\mathbf{D}^{1/2}$, where $\mathbf{D}$ is a diagonal matrix with variances at its main diagonal and $\mathbf{R}$ is the correlation matrix.

**i)** Can you derive a DCC model with multivariate Student's $t$ shocks? Write down the log-likelihood of this model.

**DCC: dynamic conditional correlation**

First we write up the expression for $y$ which is given by:

$$
y_t = \Sigma_t^{1/2} z_t, \quad z_t \sim T(0, I)
$$

where $z_t$ follows a $\mathcal{T}$ distribution parameterized by $\boldsymbol{\mu} = 0$ and $I$ (identity matrix). As written in exercise description we have the covariance matrix $\Sigma$ is related to the scale by $\Sigma = \Psi \frac{\nu}{\nu-2}$ and can be factorized as $\mathbf{D}^{1/2}\mathbf{R}\mathbf{D}^{1/2}$, where $\mathbf{D}$ is a diagonal matrix with variances at its main diagonal and $\mathbf{R}$ is the correlation matrix.

To derive the log-likelihood function, we need the density of $y_t$ conditional on past information. We set $\mu = 0$ and $\Psi = I\frac{\nu-2}{\nu}$ in the expression given in the exercise.

$$p(\mathbf{y}; \boldsymbol{\mu}, \Psi) = \frac{\Gamma\left(\frac{\nu+p}{2}\right)}{\Gamma\left(\frac{\nu}{p}\right)\pi^{p/2}\nu^{p/2}|\Psi|^{1/2}}\left(1 + \frac{(\mathbf{y}-\boldsymbol{\mu})'\Psi^{-1}(\mathbf{y}-\boldsymbol{\mu})}{\nu}\right)^{-\frac{\nu+p}{2}}$$

$$p(z) = \frac{\Gamma\left(\frac{\nu+p}{2}\right)}{\Gamma\left(\frac{\nu}{p}\right)\pi^{p/2}\nu^{p/2}|I\frac{\nu-2}{\nu}|^{1/2}}\left(1 + \frac{\mathbf{y}'\left(I\frac{\nu-2}{\nu}\right)^{-1}\mathbf{y}}{\nu}\right)^{-\frac{\nu+p}{2}}$$

$$= \frac{\Gamma\left(\frac{\nu+p}{2}\right)}{\Gamma\left(\frac{\nu}{p}\right)\pi^{p/2}\nu^{p/2}\frac{\nu-2}{\nu}^{p/2}|I|^{1/2}}\left(1 + \frac{\mathbf{y}'I^{-1}\mathbf{y}}{\nu-2}\right)^{-\frac{\nu+p}{2}}$$

$$= \frac{\Gamma\left(\frac{\nu+p}{2}\right)}{\Gamma\left(\frac{\nu}{p}\right)\pi^{p/2}(\nu-2)^{p/2}|I|^{1/2}}\left(1 + \frac{\left(\Sigma_t^{\frac{1}{2}}z_t\right)'I^{-1}\boldsymbol{\Sigma}_t^{\frac{1}{2}}z_t}{\nu-2}\right)^{-\frac{\nu+p}{2}}$$

$$= \frac{\Gamma\left(\frac{\nu+p}{2}\right)}{\Gamma\left(\frac{\nu}{p}\right)\pi^{p/2}(\nu-2)^{p/2}|I|^{1/2}}\left(1 + \frac{\left(I^{\frac{1}{2}}z_t\right)'I^{-1}I^{\frac{1}{2}}z_t}{\nu-2}\right)^{-\frac{\nu+p}{2}}$$

$$= \frac{\Gamma\left(\frac{\nu+p}{2}\right)}{\Gamma\left(\frac{\nu}{p}\right)\pi^{p/2}(\nu-2)^{p/2}|I|^{1/2}}\left(1 + \frac{z_t'z_t}{\nu-2}\right)^{-\frac{\nu+p}{2}}$$

In the derivation we use that we use $|ax| = a^p|x|$ and $(ax)^{-1} = \frac{1}{a}x^{-1}$
and that inserting $\Psi = I\frac{\nu-2}{\nu}$ in $\Sigma = \Psi\frac{\nu}{\nu-2}$ yields $\Sigma = I$

Then using $y_t = \Sigma_t^{\frac{1}{2}}z_t \Rightarrow z_t = \Sigma_t^{-\frac{1}{2}}y_t$ and $\Sigma = I$
then we get

$$p(z) = \frac{\Gamma\left(\frac{\nu+p}{2}\right)}{\Gamma\left(\frac{\nu}{p}\right)\pi^{p/2}(\nu-2)^{p/2}|I|^{1/2}}\left(1 + \frac{z_t'z_t}{\nu-2}\right)^{-\frac{\nu+p}{2}}$$

$$p\left(y_t|\mathcal{F}_{t-1}\right) = \frac{\Gamma\left(\frac{\nu+p}{2}\right)}{\Gamma\left(\frac{\nu}{p}\right)\pi^{p/2}(\nu-2)^{p/2}|\Sigma|^{1/2}}\left(1 + \frac{\left(\Sigma_t^{-\frac{1}{2}}y_t\right)'\Sigma_t^{-\frac{1}{2}}y_t}{\nu-2}\right)^{-\frac{\nu+p}{2}}$$

$$= \frac{\Gamma\left(\frac{\nu+p}{2}\right)}{\Gamma\left(\frac{\nu}{p}\right)\pi^{p/2}(\nu-2)^{p/2}|\Sigma|^{1/2}}\left(1 + \frac{(y_t)'\Sigma_t^{-1}y_t}{\nu-2}\right)^{-\frac{\nu+p}{2}}$$

Taking the log

$$
\begin{aligned}
\ln p\left(y_t | \mathcal{F}_{t-1}\right) = {} & \ln \Gamma\left(\frac{\nu+p}{2}\right) - \ln \Gamma\left(\frac{\nu}{p}\right) - \frac{p}{2}\ln \pi - \frac{p}{2}\ln(\nu-2) - \frac{1}{2}\ln|\Sigma| \\
& - \frac{\nu+p}{2}\ln\left(1 + \frac{(y_t)'\Sigma_t^{-1}y_t}{\nu-2}\right) \\
= {} & \ln \Gamma\left(\frac{\nu+p}{2}\right) - \ln \Gamma\left(\frac{\nu}{p}\right) - \frac{p}{2}\ln \pi - \frac{p}{2}\ln(\nu-2) - \frac{1}{2}\ln|D^{\frac{1}{2}}RD^{\frac{1}{2}}| \\
& - \frac{\nu+p}{2}\ln\left(1 + \frac{(y_t)'\left(D^{1/2}RD^{1/2}\right)^{-1}y_t}{\nu-2}\right) \\
= {} & \ln \Gamma\left(\frac{\nu+p}{2}\right) - \ln \Gamma\left(\frac{\nu}{p}\right) - \frac{p}{2}\ln \pi - \frac{p}{2}\ln(\nu-2) - \frac{1}{2}\left(\ln|D| + \ln|R|\right) \\
& - \frac{\nu+p}{2}\ln\left(1 + \frac{y_t'\left(D^{1/2}RD^{1/2}\right)^{-1}y_t}{\nu-2}\right) \\
= {} & \ln \Gamma\left(\frac{\nu+p}{2}\right) - \ln \Gamma\left(\frac{\nu}{p}\right) - \frac{p}{2}\ln \pi - \frac{p}{2}\ln(\nu-2) - \frac{1}{2}\ln|D| - \frac{1}{2}\ln|R| \\
& - \frac{\nu+p}{2}\ln\left(1 + \frac{y_t'\left(D^{1/2}RD^{1/2}\right)^{-1}y_t}{\nu-2}\right)
\end{aligned}
$$

We use $\Sigma = D^{\frac{1}{2}}RD^{\frac{1}{2}}$ and get the log-likelihood function

$$
\begin{aligned}
LLK = {} & T\left[\ln \Gamma\left(\frac{\nu+p}{2}\right) - \ln \Gamma\left(\frac{\nu}{p}\right) - \frac{p}{2}\ln \pi - \frac{p}{2}\ln(\nu-2)\right] - \frac{1}{2}\sum_{t=1}^{T}\ln|D_t| \\
& - \frac{\nu+p}{2}\sum_{t=1}^{T}\ln\left(1 + \frac{y_t'\left(D^{1/2}RD^{1/2}\right)^{-1}y_t}{\nu-2}\right)
\end{aligned}
$$

---

**ii)** Show that the log-likelihood of this model does not factorize in two parts as in the Gaussian case.

---

We cannot factorize the likelihood due to the last term in the log-likelihood function

$$
\sum_{t=1}^{T}\ln\left(1 + \frac{y_t'\left(D^{1/2}RD^{1/2}\right)^{-1}y_t}{\nu-2}\right)
$$

Which cannot be factorized as we are taking the log of a sum.

---

**iii)** How can we still estimate the Student's $t$ DCC model in two steps? At which costs?

---

We can use a quasi-maximum likelihood approach where we first estimate $\Sigma$ using the gaussian likelihood function and find $\widehat{z}_t = \widehat{\Sigma}_t^{1/2}y_t$. Note that we use the Gaussian distribution even though we are aware that this is not the true distribution, we will obtain a consistent estimator of $Sigma$.
In the next step we use maximum likelihood to estimate $\nu$. We estimate this $\nu$ by using LLK for $z$. We obtained the density for $z$ previously in this exercise - to obtain the LLK we just have log this density and then take sum over the time-period from 1 up till $T$.
This will yield consistent but inefficient estimates.