Exam 2020

1 Methodology

Let $Y_t$be the VIX at time t adn consider the following model:

$$Y_t | F_{t-1} \sim \mathcal{G}(\mu_t, a)$$

Where $\mathcal{G}(\mu_t, a)$ is the Gamma distribution with mean $\mu_t > 0$ and scale $a > 0$ with probability density function given by:

$$p(y_t \mid \mathcal{F}_{t-1}) = \frac{1}{\Gamma(a)} a^a y_t^{a-1} \mu_t^{-a} exp\left(-a\frac{y_t}{\mu_t}\right)$$

Here we implement the parameterization of the Gamma distrubution used by Engle and Gallo for which $E[Y_t | F_{t-1}] = \mu_t$ and $Var[Y_t | F_{t-1}] = \frac{\mu_t^2}{a}$
Form L9S15 we got. GAS implements this filter as or an updating score as:

$$\begin{aligned}\psi_t &= \psi(y_{1:t-1}) \\ &= \omega + \alpha u_{t-1} + \beta \psi_{t-1}\end{aligned}$$

Where ψ_t is the variable of interest, and the one we want to filter out. As we have that $d = \frac{1}{2}$, from L9S19, we have that the choice of $S_t = I_t^{-1}$ will then become the inverse square root scaling $\mu_y = I_t^{-\frac{1}{2}} \nabla_t$ such that $Var(u_t) = 1$. Where I is the Fischer information matrix and further more from L10S15 $I_t^{-1} = E_{t-1}\left[\nabla_t^2\right]^{-1}$. ∇_t ís the unscaled score of the conditional distribution $\nabla_t = \frac{\partial logp(y_t | y_{1:t-1}, \varphi_t, a)}{\partial \mu_t}$. The approach is to use the score in order to 'force' the direction of the updating step, which in our case has been scaled. We will introduce a link function to accomodate the specifications of μ, i.e. that it is always positive, and from L9S24, we have that when $d \neq 1$ we employ an exponenttial link function

$$\mu_t = exp(\tilde{\mu}_t) \tag{1}$$

$$\tilde{\mu}_t = \omega + \alpha \tilde{u}_{t-1} + \beta \tilde{\mu}_{t-1} \tag{2}$$

As we have $d = \frac{1}{2}$,we have that $u_t = \tilde{u}_t = I_t^{-\frac{1}{2}} \nabla_t$ i.e $\rightarrow$ We would like the score, which is defined as the derivative of the log density

$$\nabla_t = \frac{\partial logp(y_t | \mathbf{y}_{1:t-1}, \varphi_t, a)}{\partial \mu_t}$$

$$S_t = I_t^{-\frac{1}{2}} = E_{t-1}\left[\nabla_t^2\right]^{-\frac{1}{2}}$$

Note, the condition that $d = \frac{1}{2}$ imply that $u = \tilde{u}$ arises from the fact that our reparameterized model require a reparameterized u_t, i.e.:

$$
\begin{aligned}
u = \tilde{u} &= \tilde{S}_t \tilde{\nabla}_t \\
&= \tilde{I}_t^{-\frac{1}{2}} \tilde{\nabla}_t \\
&= \tilde{I}_t^{-\frac{1}{2}} \frac{\partial logp\left(y_t \middle| \mathbf{y}_{1:t-1}, \varphi_t, a\right)}{\partial \tilde{\mu}_t} \\
&= E_{t-1}\left[\tilde{\nabla}_t^2\right]^{-\frac{1}{2}} \frac{\partial logp\left(y_t \middle| \mathbf{y}_{1:t-1}, \varphi_t, a\right)}{\partial \tilde{\mu}_t} \\
&= E_{t-1}\left[\left(\frac{\partial logp\left(y_t \middle| \mathbf{y}_{1:t-1}, \varphi_t, a\right)}{\partial \tilde{\mu}_t}\right)^2\right]^{-\frac{1}{2}} \frac{\partial logp\left(y_t \middle| \mathbf{y}_{1:t-1}, \varphi_t, a\right)}{\partial \tilde{\mu}_t} \\
&= E_{t-1}\left[\left(\frac{\partial logp\left(y_t \middle| \mathbf{y}_{1:t-1}, \varphi_t, a\right)}{\partial \mu_t} * \frac{\partial \mu_t}{\partial \tilde{\mu}_t}\right)^2\right]^{-\frac{1}{2}} \frac{\partial logp\left(y_t \middle| \mathbf{y}_{1:t-1}, \varphi_t, a\right)}{\partial \mu_t} \frac{\partial \mu_t}{\partial \tilde{\mu}_t} \\
&= E_{t-1}\left[\left(\nabla_t * \frac{\partial \mu_t}{\partial \tilde{\mu}_t}\right)^2\right]^{-\frac{1}{2}} \nabla_t \frac{\partial \mu_t}{\partial \tilde{\mu}_t} \\
&= E_{t-1}\left[\nabla_t^2 * \left(\frac{\partial \mu_t}{\partial \tilde{\mu}_t}\right)^2\right]^{-\frac{1}{2}} \nabla_t \frac{\partial \mu_t}{\partial \tilde{\mu}_t} \\
&= E_{t-1}\left[\nabla_t^2\right]^{-\frac{1}{2}} \left(\frac{\partial \mu_t}{\partial \tilde{\mu}_t}\right)^{2*-\frac{1}{2}} \nabla_t \frac{\partial \mu_t}{\partial \tilde{\mu}_t} \\
&= E_{t-1}\left[\nabla_t^2\right]^{-\frac{1}{2}} \left(\frac{\partial \mu_t}{\partial \tilde{\mu}_t}\right)^{-\frac{1}{1}} \nabla_t \frac{\partial \mu_t}{\partial \tilde{\mu}_t} \\
&= E_{t-1}\left[\nabla_t^2\right]^{-\frac{1}{2}} \nabla_t \frac{\partial \mu_t}{\partial \tilde{\mu}_t} \frac{\partial \tilde{\mu}_t}{\partial \mu_t} \\
&= E_{t-1}\left[\nabla_t^2\right]^{-\frac{1}{2}} \nabla_t \\
u = \tilde{u} &= S_t \nabla_t
\end{aligned}
$$

This is the forcing variable. Note, I have used the fact that $\tilde{\mu}_t$ and μ_t is adopted to the filtration at time $t-1$, hence one may just take it out of the expectation sign.
This is for the more general case. For the model of the question:
Taking logs to the probability density function of the model we get.

$$
\begin{aligned}
log\left(p\left(y_t \mid \mathcal{F}_{t-1}\right)\right) &= log\left(\frac{1}{\Gamma(a)} a^a y_t^{a-1} \mu_t^{-a} exp\left(-a\frac{y_t}{\mu_t}\right)\right) \\
&= log\left(\frac{1}{\Gamma(a)}\right) + log\left(a^a\right) + log\left(y_t^{a-1}\right) + log\left(\mu_t^{-a}\right) + log\left(exp\left(-a\frac{y_t}{\mu_t}\right)\right) \\
&= -log\left(\Gamma\left(a\right)\right) + alog\left(a\right) + \left(a-1\right) log\left(y_t\right) + \left(-a\right) log\left(\mu_t\right) - a\frac{y_t}{\mu_t}
\end{aligned}
$$

Taking the score of this by: $\nabla_t = \frac{\partial logp\left(y_t \middle| y_{1:t-1}, \varphi_t, a\right)}{\partial \mu_t}$

$$
\begin{aligned}
\nabla_t &= \frac{\partial}{\partial \mu_t}\left(logp\left(y_t \middle| \mathbf{y}_{1:t-1}, \varphi_t, a\right)\right) \\
&= \frac{\partial}{\partial \mu_t}\left(-log\left(\Gamma\left(a\right)\right) + alog\left(a\right) + \left(a-1\right) log\left(y_t\right) + \left(-a\right) log\left(\mu_t\right) - a\frac{y_t}{\mu_t}\right) \\
&= -a\frac{1}{\mu_t} + a\frac{y_t}{\mu_t^2}
\end{aligned}
$$

This being unscaled score of the model, for this question, we have scaled the model by ∇_t^2 and thereby

$$
\begin{aligned}
\nabla_t^2 &= \left(\frac{ay_t}{\mu_t^2} - \frac{a}{\mu_t}\right)^2 \\
&= \left(\frac{ay_t}{\mu_t^2}\right)^2 - \left(\frac{a}{\mu_t}\right)^2 - 2 * \left(\frac{ay_t}{\mu_t^2}\frac{a}{\mu_t}\right) \\
&= \left(\frac{ay_t}{\mu_t^2}\right)^2 - \left(\frac{a}{\mu_t}\right)^2 - \frac{2a^2y_t}{\mu_t^3} \\
&= \frac{a^2y_t^2}{\mu_t^4} + \frac{a^2}{\mu_t^2} - \frac{2a^2y_t}{\mu_t^3} \\
&= a^2\left(\frac{y_t^2}{\mu_t^4} + \frac{1}{\mu_t^2} - \frac{2y_t}{\mu_t^3}\right)
\end{aligned}
$$

As on L9S18, we have that $Var\left(\nabla_t\right) = E_{t-1}\left[\nabla_t^2\right] = I_t$, and where we use what is in the question about the $E_t\left[Y_t \middle| F_{t-1}\right] = \mu_t$ and $E_t\left[Y_t^2 \middle| F_{t-1}\right] = \frac{\mu_t^2(1+a)}{a}$

$$
\begin{aligned}
E_{t-1}\left[\nabla_t^2\right] &= E_{t-1}\left[a^2\left(\frac{y_t^2}{\mu_t^4} + \frac{1}{\mu_t^2} - \frac{2y_t}{\mu_t^3}\right)\right] \\
&= a^2 E_{t-1}\left[\frac{y_t^2}{\mu_t^4} + \frac{1}{\mu_t^2} - \frac{2y_t}{\mu_t^3}\right] \\
&= a^2\left(E_{t-1}\left[\frac{y_t^2}{\mu_t^4}\right] + E_{t-1}\left[\frac{1}{\mu_t^2}\right] - E_{t-1}\left[\frac{2y_t}{\mu_t^3}\right]\right) \\
&= a^2\left(\frac{1}{\mu_t^4}\frac{\mu_t^2\left(1+a\right)}{a} + \frac{1}{\mu_t^2} - \frac{2}{\mu_t^3}\mu_t\right) \\
&= a^2\left(\frac{\mu_t^2\left(1+a\right)}{a\mu_t^4} - \frac{1}{\mu_t^2}\right) \\
&= \left(\frac{a+a^2}{\mu_t^2} - \frac{a^2}{\mu_t^2}\right) \\
&= \left(\frac{a}{\mu_t^2} + \frac{a^2}{\mu_t^2} - \frac{a^2}{\mu_t^2}\right) \\
&= \frac{a}{\mu_t^2}
\end{aligned}
$$

where the fact that μ_t is constant given the information at time $t-1$ - it is adopted to the filtration at time $t-1$ - hence we may treat is a constant and just take it out of the expectation sign.
Next, one need to recall that we need to find $\mathcal{I}^{-\frac{1}{2}}$, hence we get:

$$
\mathcal{I} = \frac{a}{\mu_t^2} \rightarrow \mathcal{I}^{-\frac{1}{2}} = \frac{\mu_t}{a^{\frac{1}{2}}}
$$

$$
\mathcal{I}^{-\frac{1}{2}} = \frac{\mu_t}{\sqrt{a}}.
$$

Now putting the pieces together we get: $\rightarrow$ forccing variable is then Eq.

$$
\begin{aligned}
u_t = \tilde{u}_t &= \mathcal{I}^{-\frac{1}{2}}\nabla_t \\
&= \frac{\mu_t}{\sqrt{a}} * \left(\frac{ay_t}{\mu_t^2} - \frac{a}{\mu_t}\right) \\
&= \frac{1}{\sqrt{a}} * \left(\mu_t\frac{ay_t}{\mu_t^2} - \mu_t\frac{a}{\mu_t}\right) \\
&= \frac{1}{a^{\frac{1}{2}}} * \left(\frac{ay_t}{\mu_t} - a\right) \\
&= a^{-\frac{1}{2}} * a\left(\frac{y_t}{\mu_t} - 1\right) \\
&= a^{\frac{1}{2}}\left(\frac{y_t}{\mu_t} - 1\right) \\
u_t &= a^{\frac{1}{2}}\left(\frac{y_t}{exp\left(\tilde{\mu_t}\right)} - 1\right) \qquad (3)
\end{aligned}
$$

Where $Eq.$ now can be interpretede as the the updating step. We now have the Gamma-GAS updating step as

$$\mu_t = exp\left(\tilde{\mu_t}\right) \tag{4}$$
$$\tilde{\mu}_t = \omega + \alpha\tilde{u}_{t-1} + \beta\tilde{\mu}_{t-1} \tag{5}$$
$$\tilde{\mu}_t = a^{\frac{1}{2}}\left(\frac{y_t}{exp\left(\tilde{\mu_t}\right)} - 1\right) \tag{6}$$

From L9S22 we have that when $d = \frac{1}{2}$ $\tilde{\mu}_t = \mu_t$ and from L9S and when $d = \frac{1}{2}$: Inverse square root scaling $u_t = I_t^{-\frac{1}{2}}\nabla_t$ such that $Var\left(u_y\right) = Var\left(\tilde{u}_t\right) = 1$ i.e. taking the variance to $\tilde{\mu_t}$ should equal 1.

$$\begin{aligned}
Var\left(\tilde{\mu}_t\right) &= Var\left(a^{\frac{1}{2}}\left(\frac{y_t}{exp\left(\tilde{\mu}_t\right)} - 1\right)\right) \\
&= a^{\frac{1}{2}*2}Var\left(\frac{y_t}{exp\left(\tilde{\mu}_t\right)} - 1\right) \\
&= a * \left(\frac{1}{exp\left(\tilde{\mu}_t\right)}\right)^2 Var\left(y_t\right) \\
&= a * \frac{1}{exp\left(\tilde{\mu}_t\right)^2}\frac{\mu_t^2}{a} \\
&= \frac{a}{\mu_t^2}\frac{\mu_t^2}{a} \\
Var\left(\tilde{\mu}_t\right) &= 1
\end{aligned}$$

where I have used that the variance to a constant is zero and that $\mu_t = exp\left(\tilde{\mu}_t\right)$. We may therefore confirm that we have found the right $u_t = \tilde{u}_t$, as we have $Var\left(u_t\right) = Var\left(\tilde{u}_t\right) = 1$.
For the loglikelihood of the model we have
In exercise 1.1, the individual log-likelihood contributions has been derived as:

$$logp\left(y_t \middle| \mathbf{y}_{1:t-1}, \mu_t, a\right) = -log\left(\Gamma\left(a\right)\right) + alog\left(a\right) + \left(a-1\right)log\left(y_t\right) - alog\left(\mu_t\right) - a\frac{y_t}{\mu_t}. \tag{7}$$

Adding up, recalling that μ_t is a function of (ω, α, β), one obtain for length T the loglikelihood as:

$$logL\left(\omega, \alpha, \beta, a\right) = -\sum_{t=1}^{T}\left[log\left(\Gamma\left(a\right)\right) - alog\left(a\right) - \left(a-1\right)log\left(y_tFor\right) + alog\left(\mu_t\left(\omega, \alpha, \beta\right)\right) + a\frac{y_t}{\mu_t\left(\omega, \alpha, \beta\right)}\right] \tag{8}$$
$$= T\left(-ln\eth\left(a\right) + aln\left(a\right)\right) + \sum_{t=1}^{T}\left(\left(a-1\right)log\left(y_tFor\right) + alog\left(\mu_t\left(\omega, \alpha, \beta\right)\right) + a\frac{y_t}{\mu_t\left(\omega, \alpha, \beta\right)}\right)$$

When considering the GAMMA-GAS model, I have written a few codes; 1) to derive the log-pdf of the gamma distribution, 2) that filter the means, μ_t, and 3) that uses maximum likelihood to estimate the parameter values in the filter, i.e. $(\omega, \alpha, \beta, a)$. When initialising the filter, I have taken the unconditional expectation to $\tilde{\mu}_t$, i.e.:

$$\begin{aligned}
E\left[\tilde{\mu}_t\right] &= \omega + \alpha\underbrace{\sqrt{a}\left(\frac{E\left[y_t\right]}{exp\left(\tilde{\mu}_t\right)} - 1\right)}_{Eq.6} + \beta E\left[\mu_{t-1}\right] \\
&= \omega + \alpha\sqrt{a}\left(\frac{\mu_t}{\mu_t} - 1\right) + \beta E\left[\mu_{t-1}\right] \\
&= \omega + \alpha\sqrt{a}\underbrace{\left(\frac{\mu_t}{\mu_t} - 1\right)}_{=0} + \beta E\left[\mu_{t-1}\right] \\
&= \frac{\omega}{1-\beta}
\end{aligned} \tag{9}$$

which will be the starting condition for $\tilde{\mu}_t$, i.e. $\tilde{\mu}_1$. Note, we have assumed stationarity in order for this to hold, i.e. $E\left[\tilde{\mu}_t\right] = E\left[\tilde{\mu}_{t-1}\right]$. Furthermore, from the exercise, the parameter constraints has been specified for the coefficients $(\omega, \alpha, \beta, \alpha)$.

Multiplicative Error Model

After considering the GAMMA-GAS model - constrained and unconstrained - one moves onto considering the Multiplicative Errror Model by Engle and Gallo (2006). Notice it is formulated in the following way:

$$\begin{aligned} Y_t | \mathcal{F}_{t-1} &\sim \mathcal{G}a\left(\mu_t, a\right), \\ \mu_t &= \kappa + \eta y_{t-1} + \phi \mu_{t-1}. \end{aligned} \tag{10}$$

For the estimation of the Multiplicative Error Model, I have created two functions, 1) that filter the means, μ_t, and 2) that uses maximum likelihood to estimate the parameter values in the filter, i.e. (κ, η, ϕ, a). Furthermore it compute the BIC as well as the loglikelihood.

When estimating the filter, I have taken the unconditional expectation to μ_t, i.e.:

$$\begin{aligned} E\left[\mu_t\right] &= \kappa + \eta \underbrace{E\left[y_{t-1}\right]}_{\mu_{t-1}} + \phi \underbrace{E\left[\mu_{t-1}\right]}_{\mu_t} \\ &= \kappa + \eta \mu_{t-1} + \phi E\left[\mu_t\right] \\ E\left[\mu_t\right] &= \kappa + \mu_t\left(\eta + \phi\right) \\ \kappa &= \mu_t\left(1 - \eta - \phi\right) \\ &= \frac{\kappa}{1 - \eta - \phi}, \end{aligned} \tag{11}$$

where I have used two things; 1) $E\left[y_t\right] = \mu_t$ and stationarity, i.e. $E\left[\mu_t\right] = E\left[\mu_{t-1}\right]$. Notice that this further imply one condition, which has not been imposed in the exercise, i.e. that

$$\eta + \phi < 1,$$

Question 2

Methodology

Bivariate Gaussian DCC model - model, constraints and factorization of the likelihood

We consider a bivariate random vector of the returns of GSPC and DJI returns. As one notices from the exercise, we assume that the distribution at time t, conditional on the filtration at time $t-1$, is bivarate gaussian according to:

$$\mathbf{Y}_t | \mathcal{F}_{t-1} \sim N\left(\mathbf{0}, \Sigma_t\right). \tag{12}$$

Leopoldo further specifies the model as so:

$$\begin{aligned} \boldsymbol{Y}_t &= \left(y_{1,t}, ..., y_{N,t}\right)' \\ y_{i,t} &= \sigma_{i,t}\epsilon_t \\ \sigma_{i,t}^2 &= \omega + \alpha_i y_{i,t-1}^2 + \beta\sigma_{i,t-1}^2 \end{aligned}$$

$\sigma_{i,t}^2$ are the marginals GARCH, put into the DCC framework

$$Y_y = \Sigma^{\frac{1}{2}}\epsilon_t \rightarrow \Sigma^{\frac{1}{2}}\Sigma^{\frac{1}{2}} = \boldsymbol{D}_t\boldsymbol{R}_t\boldsymbol{D}_t$$

For when considering the DCC model we turn to L10S24 / L10S30, were we note that the conditional covariance matriz can be factorized as;

$$\boldsymbol{H}_t = \boldsymbol{D}_t^{\frac{1}{2}}\boldsymbol{R}_t\boldsymbol{D}_t^{\frac{1}{2}} \tag{13}$$

where $\boldsymbol{D_t}$ is a matrix with the variance on the diagonol i.e,

$$\boldsymbol{D}_t = \begin{bmatrix} \sigma_{1,t}^2 & 0 & 0 \\ 0 & \sigma_{2,t}^2 & 0 \\ 0 & 0 & \sigma_{3,t}^2 \end{bmatrix} \tag{14}$$

and where $\sigma_{i,t}^2 = var\left(y_{i,t}|\mathcal{F}_{t-1}\right)$. $\boldsymbol{R}_t$ is the evaluation matrix with typical element $\rho_{i;j,t} = cor\left(y_{i,t}, y_{j,t}|\mathcal{F}_{t-1}\right)$ i.e

$$\boldsymbol{R}_t = \begin{bmatrix} 1 & \rho_{1.2} & \rho_{1.3} \\ \rho_{2,1} & 1 & \rho_{2.3} \\ \rho_{3.1} & \rho_{3.2} & 1 \end{bmatrix} \tag{15}$$

$Eq.15$ is also what Leopoldo calls $\bar{\boldsymbol{R}}_t = cor(\eta_t)$
From here we turn to L10S31 $\rightarrow$ As specified in the exercise, in the following, we will allow for the marginal processes or the σ^2 to follow a $GARCH(1,1)$ process of below form, and further, one needs to incorporate the assumption about $\mathbf{R}_t$:

$$\sigma_{i,j}^2 = \omega + \alpha y_{i,t-1}^2 + \beta\sigma_{i.t-1}^2 \tag{16}$$

$$\mathbf{R}_t = \tilde{\mathbf{Q}}_t^{-\frac{1}{2}}\mathbf{Q}_t\tilde{\mathbf{Q}}_t^{-\frac{1}{2}} \tag{17}$$

Where $\tilde{\boldsymbol{Q}}_t$ is a diagonol matrix with typical elements of $\boldsymbol{Q}_t$, which we can write out as on L10S31.

$$\boldsymbol{Q}_t = \bar{\boldsymbol{Q}}\left(1-a-b\right) + a\left(\boldsymbol{\eta}_{t-1}\boldsymbol{\eta}_{t-1}^{'}\right) + b\left(\boldsymbol{Q}_{t-1}\right) \tag{18}$$

Leopoldo alsp has $\rightarrow \boldsymbol{\eta}_t = \left(\eta_{1,t},...,\eta_{N,T}\right)^{'}, \eta_{i,t} = \frac{y_{i,t}}{\sigma_{i,t}}$
Where $\boldsymbol{\eta}_t = \boldsymbol{D}_t^{-\frac{1}{2}}\boldsymbol{y}_t$ are standarized returns, and $\bar{\boldsymbol{Q}}$ is fixed to the emperical evaluation of $\boldsymbol{\eta}_t\bar{\boldsymbol{Q}}_t = \bar{\boldsymbol{R}}_t = cor\left(z_t\right)$ used in the CCC r
in above formulation, we have the following constraints for our model. this is to insure both that the variances are positive but $\beta+\alpha\leq 1$, is also for weak stationarity

$$positivity\ of\ \sigma_{i,t}^2 = \begin{cases} \omega_i > 0 \\ \alpha_i > 0 \\ \beta_i > 0 \end{cases}$$

$$weak\ stationarity\ of\ y_{i,t} = \{\alpha_i + \beta_i < 1$$

$$\begin{aligned} 0 \leq a, b \leq 1, &\qquad a+b \leq 1, \\ 0 \leq \alpha, \beta < 1, &\qquad \alpha+\beta < 1. \end{aligned} \tag{19}$$

These constraints do not only cause weak stationarity of the our GARCH(1,1) process[1], but also causes the conditional covariance matrix, $\mathbf{Q}_t$, to be positive definite, which in terms imply that $\mathbf{R}_t$ is positive definite. As oppose to the CCC model, $\mathbf{Q}_t$ allows for $\mathbf{R}_t$ to change over time, but a drawback of the general procedure is that a and b drive the dynamics of the correlations, which has led to more recent extensions of the DCC model, which overcomes this feature.

Next step is to derive the loglikelihood, hence one consider the density of $\mathbf{y}_t|\mathcal{F}_{t-1}$; note $\mu = 0$, hence it reduces to: Simply taking from wiki, and with $\sigma^2 = \Sigma$

$$p\left(\mathbf{y};0,\Sigma\right) = \frac{1}{2\pi\left|\Sigma_t\right|^{\frac{1}{2}}}exp\left\{-\frac{1}{2}\mathbf{y}_t{}'\Sigma_t^{-1}\mathbf{y}_t\right\}. \tag{20}$$

Taking logs yields:

---

[1]Note, the GARCH model has been initialised with its unconditional variance, i.e.

$$\begin{aligned} E\left[\sigma_{i,t}^2\right] &= \omega + \alpha\underbrace{E\left[y_{i,t-1}^2\right]}_{\sigma_{i,t-1}^2} + \beta\left[\sigma_{i,t-1}^2\right] \\ &= \frac{\omega}{1-\alpha-\beta}. \end{aligned}$$

$$\begin{aligned}
log\left(p\left(\mathbf{y};0,\Sigma\right)\right) &= log\left(\frac{1}{2\pi\left|\Sigma_t\right|^{\frac{1}{2}}}exp\left\{-\frac{1}{2}\mathbf{y}_{\mathrm{t}}{}'\Sigma_t^{-1}\mathbf{y}_{\mathrm{t}}\right\}\right)\\
&= log\left(2\pi\left|\Sigma_t\right|^{\frac{1}{2}}\right)^{-1}exp\left\{-\frac{1}{2}\mathbf{y}_{\mathrm{t}}{}'\Sigma_t^{-1}\mathbf{y}_{\mathrm{t}}\right\}\\
&= -log\left(2\pi\right)-log\left(\left|\Sigma_t\right|^{\frac{1}{2}}\right)exp\left\{-\frac{1}{2}\mathbf{y}_{\mathrm{t}}{}'\Sigma_t^{-1}\mathbf{y}_{\mathrm{t}}\right\}\\
&= -log\left(2\pi\right)-\frac{1}{2}log\left(\left|\Sigma_t\right|\right)-\frac{1}{2}\mathbf{y}_{\mathrm{t}}{}'\Sigma_t^{-1}\mathbf{y}_{\mathrm{t}}
\end{aligned}\tag{21}$$

And considering of lenght T

$$\begin{aligned}
logL_t &= \sum_{t=1}^{T}-log\left(2\pi\right)-\frac{1}{2}log\left(\left|\Sigma_t\right|\right)-\frac{1}{2}\mathbf{y}_{\mathrm{t}}{}'\Sigma_t^{-1}\mathbf{y}_{\mathrm{t}}\\
&= -\frac{1}{2}\sum_{t=1}^{T}+2log\left(2\pi\right)+log\left(\left|\Sigma_t\right|\right)+\mathbf{y}_{\mathrm{t}}{}'\Sigma_t^{-1}\mathbf{y}_{\mathrm{t}}\\
&= -\frac{1}{2}\sum_{t=1}^{T}\left(2log\left(2\pi\right)+log\left(\left|\underbrace{\mathbf{D}_t^{\frac{1}{2}}\mathbf{R}_t\mathbf{D}_t^{\frac{1}{2}}}_{\Sigma_t}\right|\right)+\mathbf{y}_t'\left(\underbrace{\mathbf{D}_t^{\frac{1}{2}}\mathbf{R}_t\mathbf{D}_t^{\frac{1}{2}}}_{\Sigma_t}\right)^{-1}\mathbf{y}_t\right)\\
&= -\frac{1}{2}\sum_{t=1}^{T}\left(2log\left(2\pi\right)+log\left(\left|\mathbf{R}_t\right|\right)+log\left(\left|\mathbf{D}_t\right|\right)+\underbrace{\mathbf{y}_t'\mathbf{D}_t^{-\frac{1}{2}}}_{\boldsymbol{\eta}_t}\mathbf{R}_t^{-1}\underbrace{\mathbf{D}_t^{-\frac{1}{2}}\mathbf{y}_t}_{\boldsymbol{\eta}_t}\right)\\
&= -\frac{1}{2}\sum_{t=1}^{T}\left(2log\left(2\pi\right)+log\left(\left|\mathbf{R}_t\right|\right)+log\left(\left|\mathbf{D}_t\right|\right)+\boldsymbol{\eta}_t'\mathbf{R}_t^{-1}\boldsymbol{\eta}_t\right)
\end{aligned}\tag{22}$$

As introduced in L10S34, we add and substract $\mathbf{Y}_t'\mathbf{D}_t^{-\frac{1}{2}}\mathbf{D}_t^{-\frac{1}{2}}\mathbf{Y}_t=\boldsymbol{\eta}_t'\boldsymbol{\eta}\rightarrow_t\mathbf{Y}_t'\mathbf{D}_t^{-\frac{1}{2}}\mathbf{D}_t^{-\frac{1}{2}}\mathbf{Y}_t-\boldsymbol{\eta}_t'\boldsymbol{\eta}$

$$\begin{aligned}
logL_T &= -\frac{1}{2}\sum_{t=1}^{T}\left(2log\left(2\pi\right)+log\left(\left|\mathbf{D}_t\right|\right)+\mathbf{y}_t'\mathbf{D}_t^{-\frac{1}{2}}\mathbf{D}_t^{-\frac{1}{2}}\mathbf{y}_t-\boldsymbol{\eta}_t'\boldsymbol{\eta}_t+log\left(\left|\mathbf{R}_t\right|\right)+\boldsymbol{\eta}_t'\mathbf{R}_t^{-1}\boldsymbol{\eta}_t\right)\\
&= \underbrace{-\frac{1}{2}\sum_{t=1}^{T}\left(2log\left(2\pi\right)+log\left(\left|\mathbf{D}_t\right|\right)+\mathbf{y}_t'\mathbf{D}_t^{-1}\mathbf{y}_t\right)}_{logL_{V,T}(\theta)}\underbrace{-\frac{1}{2}\sum_{t=1}^{T}\left(\boldsymbol{\eta}_t'\mathbf{R}_t^{-1}\boldsymbol{\eta}_t-\boldsymbol{\eta}_t'\boldsymbol{\eta}_t+log\left(\left|\mathbf{R}_t\right|\right)\right)}_{logL_{C,T}(\theta,\phi)},
\end{aligned}\tag{23}$$

$$logL_{V,T}\left(\theta\right)=-\frac{1}{2}\sum_{t=1}^{T}\left(2log\left(2\pi\right)+log\left(\left|\mathbf{D}_t\right|\right)+\mathbf{y}_t'\mathbf{D}_t^{-1}\mathbf{y}_t\right)\tag{24}$$

$$ogL_{C,T}\left(\theta,\phi\right)=-\frac{1}{2}\sum_{t=1}^{T}\left(\boldsymbol{\eta}_t'\mathbf{R}_t^{-1}\boldsymbol{\eta}_t-\boldsymbol{\eta}_t'\boldsymbol{\eta}_t+log\left(\left|\mathbf{R}_t\right|\right)\right)\tag{25}$$

The procedure is then to seperately maximizing the likelihoods and then sum them in order to report the total likelihood of the model. The procedure is following a two-step approach, where one initially maximizes the likelihood with regards to the volatility, hence obtain $\hat{\theta}$, and then one maximizes the likelihood of the correlation part with regards to ϕ in order to obtain $\hat{\phi}$. This approach is possible due to the nice factorization of the likelihood functions.

For the volatility part of the likelihood, one may write:

$$\begin{aligned}
logL_{V,T}\left(\theta\right) &= -\frac{1}{2}\sum_{t=1}^{T}\left(2log\left(2\pi\right)+log\left(\left|\mathbf{D}_t\right|\right)+\mathbf{y}_t'\mathbf{D}_t^{-1}\mathbf{y}_t\right)\\
&= -\frac{1}{2}\sum_{t=1}^{T}\sum_{i=1}^{N}\left(2log\left(2\pi\right)+log\left(h_{i,t}\right)+\frac{r_{i,t}^2}{h_{i,t}}\right),
\end{aligned}\tag{26}$$

which simply imply that we may split the problem into 'severals' i.e. one may maximize each GARCH likelihood seperately and then sum them at the end, which is exactly what our function in R will do with the data on GSPC and DJI.

The succes of the procedure crucially depends on consistency of the first estimate, i.e. $\hat{\theta}$ - if it is consistent, then the estimate of phi, $\hat{\phi}$, will be consistent assuming continuity of the loglikelihood function around the true value of ϕ.

As we have specified the DCC model, one may consider the special case with $\mathbf{R}_t = \mathbf{R}$, which is the CCC model. For the CCC model we set $a = b = 0$ in $\boldsymbol{Q}_t$, CCC is a special case of the DCC model.

Constant Conditional Correlation (CCC) Model

When considering the Constant Conditional Correlation model, we will make use of most parts of above in regards to the DCC model, but the crucial point is the definition of $\mathbf{R}$, where the CCC model make use of a time-invariant correlation matrix defined as:

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

hence we NO longer need the part regarding $\mathbf{Q}_t$ in our estimation procedure. This assumption imply a factorization of Σ_t according to $\mathbf{D}_t^{\frac{1}{2}}\mathbf{R}\mathbf{D}_t^{\frac{1}{2}}$, hence our estimation procedure becomes marginally easier as we no longer need to derive $\mathbf{R}_t$ for each time period. Instead of filtering out the correlation matrix across time - as it is now constant - one now only needs to filter out the covariances across time, which may do by recalling the formula for correlation:

$$\rho_{i;j} = \frac{\sigma_{i;j,t}}{\sigma_{i,t}\sigma_{j_t}} \Rightarrow \sigma_{i;j,t} = \rho_{i;j}\sigma_{i,t}\sigma_{j_t}.$$

Note, this is obviously for the simplified case, in practice, we consider the factorization and obtain

$$\Sigma_t = \mathbf{D}_t^{\frac{1}{2}}\mathbf{R}\mathbf{D}_t^{\frac{1}{2}}$$

One may therefore realise that modelling the conditional variances will be sufficient as they explain all variation in Σ_t.

Considering the CCC model, it has been widely used in the litterature due to its simplicity, but in general, it may face difficulties when considering financial returns as the assumption of constant correlation seldom is supported by data - one should note that due to the quite high and quite constant (except for a few spikes) correlation among the indexes we consider, GSPC and DJI, the CCC model will still perform quite well.

Considering the constraints of the model, as it is also based on the GARCH(1,1) marginal processes, we must require that (as we do with the DCC model)

$$0 \leq \alpha, \beta < 1, \qquad \alpha + \beta < 1,$$

in order to have weak stationarity. Furthermore, we have that if the conditional variances, $\mathbf{D}_t$, is all positive, $\mathbf{R}$ is positive definite, then Σ_t will be guarenteed to be positive definite.

DCC vs CCC

In order to visualize the difference between the two models, one may consider the following graph, which indeed describe the main difference between the models, i.e. the correlation matrix for the two models, i.e. $\mathbf{R}_t$ for DCC and $\mathbf{R}$ for CCC.

A figure make disappear

Hence as described in words; the DCC model will allow for the correlation to vary across time, whereas the CCC model will keep correlation constant.

Empirical Analysis

Data Series

Note, as stated initially, for question 2, I have made use of the updated dataset, i.e. data_new.csv, which features the returns for the indexes GSPC and DJI.

As an initial step, I have started by plotting the returns against time and against each other.
A figure make disappear
As one instantly notices, the returns of the two indexes follow very much the same pattern and considering returns against returns, one clearly find a strong positive correlation between the indexes. This seems quite intuitive - the indexes features a variety of companies and if the indexes contain similar companies, one would have indexes that would be positively correlated. This is indeed the case and considering the correlation between the indexes over the entire time interval, it is close to 1.

The strong positive correlation may help us in the following, as it will allow us to form expectations about the results we obtain, including:

1. One would expect that the portfolio weights in the minimum variance portfolio would vary quite a lot over time due to the postive correlation.
2. When computing the Conditional Value at Risk, CoVaR, it will (most likely) be negative.

Minimum Variance Portfolio (MVP)

Considering the minimum variance portfolio, we seek to compute the weight, ω_t associated to the MVP constructed using GSPC and DJI at each point in time, i.e.

$$y_t = \omega_t y_t^{GSPC} + (1 - \omega_t)\, y_t^{DJI}.$$

In the following, we have illustrated the weights in both indexes for both models. Note that short selling as well as fractional holdings of either index is allowed in our setting.
A figure make disappear
One notices instantly that short selling is used to a large extend and that the portfolio weights vary a great deal over the time horizon. It it is a bit difficult to state a clear pattern, but in the first couple of years, our DCC model tends to short the GSPC index more than the CCC model does - it tends to favour larger positions in DJI than the CCC model. The pattern gets less clear as time goes by and the two models look more similar in terms of weights from 2014 and forward.

In general, one may conclude that the DCC model is extreme in terms of the position in the indexes, which must be caused by the time-varying correlation, i.e. the model is less restricted, hence become more 'extreme' in terms of portfolio weights.

CoVaR of GSPC given DJI

As a last step, we woud like to obtain the Conditional Value at Risk - CoVaR for our two models; DCC and CCC. The Conditional Value at Risk is defined in the following way as in L14S4:

$$P\left(Y_{1,t} \leq CoVaR_t\left(\alpha\right) | Y_{2,t} \leq VaR_t\left(\alpha\right), \mathcal{F}_{t-1}\right) = \alpha,$$

hence the size we compute is essentially the α% chance of returns below $VaR_{GSPC,t}\left(\alpha\right)$ conditioning on the α% chance of returns below $VaR_{DJI,t}\left(\alpha\right)$. To clarify with an example; if the $CoVaR_{GSPC,t} = VaR_{DJI,t}$ is a 2% return, then $\alpha = 0.01$ imply that we have 1% chance of realising returns smaller than or equal to 2% on our GSPC index, conditioning on the fact that we have a 1% chance of realising returns smaller than or equal to 2% for the DJI index.

As we have already considered the data series and noted that the indexes are very positive correlated, one would therefore naturally assume that the $CoVaR$ is negative. Note, if the two indexes where uncorrelated, $CoVaR_{GSPC,t} = VaR_{GSPC,t}$.

To take a step deeper into the theory, I have decided to follow the approach by Giulio Girardi and A. Tolga Ergün (Journal of Banking & Finance, Volume 37, Issue 8, August 2013, Pages 3169-3180).
Using their definition and reframing it according to our problem, the CoVaR is defined as the VaR for index i, GSPC, conditional on index j, DJI, being in 'financial distress', which will allow to show some of the risk spillovers that may exists between the indexes.

In general we consider two versions of the $CoVaR$;

1. $CoVaR =$ which imply that the event we are conditioning on is exactly at its VaR, i.e.

$$P\left(Y_{1,t} \leq CoVaR_t\left(\alpha\right) \middle| Y_{2,t} = VaR_t\left(\alpha\right), \mathcal{F}_{t-1}\right) = \alpha,$$

   (a) This version has been modelled at the end of the code (but is NOT graphed below, if one would like to compare.

2. $CoVaR =$ which imply that the event we are conditioning on is at most at its VaR, i.e.

$$P\left(Y_{1,t} \leq CoVaR_t\left(\alpha\right) \middle| Y_{2,t} \leq VaR_t\left(\alpha\right), \mathcal{F}_{t-1}\right) = \alpha.$$

   (a) This version is graphed below.

In the following we will consider the second version and in order to estimate the size, we follow the three-step procedure, which reads:

1. Derive $P\left(Y_{2,t} \leq VaR_t\left(\alpha\right)\right) = \alpha$ as it will constitute a upper bound in the integral at point 3.
2. Estimation of the bivariate model one consider; i.e. note it is the bivariate gaussian DCC (and CCC) model.
3. Computation of $P\left(Y_{1,t} \leq CoVaR_t\left(\alpha\right) \middle| Y_{2,t} \leq VaR_t\left(\alpha\right), \mathcal{F}_{t-1}\right) = \alpha$ by the integral:

$$\int_{-\infty}^{\mathrm{CoVaR}_t(\alpha)} \int_{-\infty}^{\mathrm{VaR}_t(\alpha)} pdf_t\left(x, y\right) dy dx = \alpha^2,$$

   which one derives by the uniroot procedure, where we solve for the unknown $\mathrm{CoVaR}_t\left(\alpha\right)$.

The formal steps to get to the integral is from the fact that we have:

$$\begin{aligned}
P\left(Y_{1,t} \leq CoVaR_t\left(\alpha\right) \middle| Y_{2,t} \leq VaR_t\left(\alpha\right), \mathcal{F}_{t-1}\right) &= \alpha \\
\frac{P\left(Y_{1,t} \leq CoVaR_t\left(\alpha\right), Y_{2,t} \leq VaR_t\left(\alpha\right) \middle| \mathcal{F}_{t-1}\right)}{\underbrace{P\left(Y_{2,t} \leq VaR_t\left(\alpha\right) \middle| \mathcal{F}_{t-1}\right)}_{=\alpha}} &= \alpha \\
P\left(Y_{1,t} \leq CoVaR_t\left(\alpha\right), Y_{2,t} \leq VaR_t\left(\alpha\right) \middle| \mathcal{F}_{t-1}\right) &= \alpha \underbrace{P\left(Y_{2,t} \leq VaR_t\left(\alpha\right) \middle| \mathcal{F}_{t-1}\right)}_{=\alpha} \\
\int_{-\infty}^{\mathrm{CoVaR}_t(\alpha)} \int_{-\infty}^{\mathrm{VaR}_t(\alpha)} pdf_t\left(x, y\right) dy dx &= \alpha^2.
\end{aligned}$$

We obtain the $CoVaR_{GSPC,t}\left(\alpha\right)$ for $\alpha \in [0.01, 0.05]$ as:

A figure make disappear

As expected, the $CoVaR_{GSPC}$ is indeed negative for both models. As one notices, the smaller the α, i.e. percentage chance of extreme event, the effect of that event becomes larger, which intuitive makes great sense; the more extreme event one consider, the smaller the probability for that event to occur. Hence moving from $\alpha = 0.05$ to $\alpha = 0.01$ simply shifts the curve downward.