1 Question 1 - Methodology

1.1 Point a) Derive a GAS model for μ_t

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

$$Y_t \mid \mathcal{F}_{t-1} \sim \mathcal{G}a\left(\mu_t, a\right)$$

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

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

We implement the parameterization of the Gamma distribution used by Engle and Gallo (2006) where

$$\mathbb{E}\left[Y_t | \mathcal{F}_{t-1}\right] = \mu_t$$
$$\mathbb{V}\left[Y_t | \mathcal{F}_{t-1}\right] = \frac{\mu_t^2}{a}$$

Derive a GAS model for μ_t and scale the score by the inverse of the square root of the Fisher information quantity for μ_t, i.e. set $d = 1/2$ in slide 19 of Lecture 10. Note that

$$E\left[Y_t^2 \mid \mathcal{F}_{t-1}\right] = \frac{\mu_t^2(1+a)}{a}$$

If you cannot derive the information quantity, use identity scaling, i.e. $d = 0$.

We start by logging the density function to make it easier to take the derivative w.r.t. μ_t

$$\begin{aligned}
p\left(y_t \mid \mathcal{F}_{t-1}\right) &= \frac{1}{\Gamma(a)} a^a y_t^{a-1} \mu_t^{-a} \exp\left(-a \frac{y_t}{\mu_t}\right) \\
\ln\left[p\left(y_t \mid \mathcal{F}_{t-1}\right)\right] &= \ln\left[\frac{1}{\Gamma(a)} a^a y_t^{a-1} \mu_t^{-a} \exp\left(-a \frac{y_t}{\mu_t}\right)\right] \\
\ln\left[p\left(y_t \mid \mathcal{F}_{t-1}\right)\right] &= \underbrace{\ln(1)}_{=0} - \ln\left(\Gamma(a)\right) + a\ln(a) + (a-1)\ln\left(y_t\right) - a\ln\left(\mu_t\right) - a\frac{y_t}{\mu_t} \\
\ln\left[p\left(y_t \mid \mathcal{F}_{t-1}\right)\right] &= -\ln\left(\Gamma(a)\right) + a\ln(a) + (a-1)\ln\left(y_t\right) - a\ln\left(\mu_t\right) - a\frac{y_t}{\mu_t}
\end{aligned}$$

Now we want to find the score of the conditional distribution (∇_t)

$$\begin{aligned}
\nabla_t &= \frac{\partial \ln\left[p\left(y_t \mid \mathcal{F}_{t-1}\right)\right]}{\partial \mu_t} = -\frac{a}{\mu_t} - \left(-\frac{a y_t}{\mu_t^2}\right) \\
\nabla_t &= -\frac{a}{\mu_t} + \frac{a y_t}{\mu_t^2} \\
\nabla_t &= \frac{a y_t}{\mu_t^2} - \frac{a}{\mu_t}
\end{aligned}$$

We know from slides (Lecture 9 slide 17) that the score ∇_t has zero expectation. The derivitation below is pasted from slides, whereas we write ψ.

$$\begin{aligned}
\mathbb{E}_{t-1}\left[\nabla_t\right] &= \int_{\mathcal{Y}} \frac{\partial \log p\left(y_t \mid \mathbf{y}_{1:t-1}; \psi\right)}{\partial \psi_t} p\left(y_t \mid \mathbf{y}_{1:t-1}; \psi\right) \mathrm{d}y \\
&= \int_{\mathcal{Y}} \frac{\partial p\left(y_t \mid \mathbf{y}_{1:t-1}; \psi\right)}{\partial \psi_t} \frac{1}{p\left(y_t \mid \mathbf{y}_{1:t-1}; \psi\right)} p\left(y_t \mid \mathbf{y}_{1:t-1}; \psi\right) \mathrm{d}y \\
&= \int_{\mathcal{Y}} \frac{\partial p\left(y_t \mid \mathbf{y}_{1:t-1}; \psi\right)}{\partial \psi_t} \mathrm{d}y \\
&= \frac{\partial}{\partial \psi_t} \int_{\mathcal{Y}} p\left(y_t \mid \mathbf{y}_{1:t-1}; \psi\right) \mathrm{d}y \\
&= \frac{\partial}{\partial \psi_t} 1 = 0
\end{aligned}$$

Since $\mathbb{E}_{t-1}\left[\nabla_t\right]=0$ we can write the Fisher information matrix $\mathcal{I}_t\left(\mu_t\right)$ as

$$
\begin{aligned}
\mathbb{V}\left(\nabla_t\right) &= \mathbb{E}_{t-1}\left[\nabla_t^2\right]=\mathcal{I}_t\left(\mu_t\right) \\
\mathcal{I}_t\left(\mu_t\right) &= \underbrace{\mathbb{E}_{t-1}\left[\left(\frac{a y_t}{\mu_t^2}-\frac{a}{\mu_t}\right)^2\right]}_{\mathbb{E}_{t-1}\left[\nabla_t^2\right]} \\
\mathcal{I}_t\left(\mu_t\right) &= \mathbb{E}_{t-1}\left[\left(\frac{a y_t}{\mu_t^2}-\frac{a}{\mu_t}\right)^2\right] \\
\mathcal{I}_t\left(\mu_t\right) &= \mathbb{E}_{t-1}\left[\left(\frac{a y_t}{\mu_t^2}\right)^2+\left(\frac{a}{\mu_t}\right)^2-2 \cdot \frac{a y_t}{\mu_t^2} \frac{a}{\mu_t}\right] \\
\mathcal{I}_t\left(\mu_t\right) &= \mathbb{E}_{t-1}\left[\frac{a^2 y_t^2}{\mu_t^4}+\frac{a^2}{\mu_t^2}-2 \cdot \frac{a^2 y_t}{\mu_t^3}\right] \\
\mathcal{I}_t\left(\mu_t\right) &= \mathbb{E}_{t-1}\left[\frac{a^2 y_t^2}{\mu_t^4}+\frac{a^2}{\mu_t^2}-2 \cdot \frac{a^2 y_t}{\mu_t^3}\right] \\
\mathcal{I}_t\left(\mu_t\right) &= \mathbb{E}_{t-1}\left[\frac{a^2 y_t^2}{\mu_t^4}\right]+\mathbb{E}_{t-1}\left[\frac{a^2}{\mu_t^2}\right]-\mathbb{E}_{t-1}\left[2 \cdot \frac{a^2 y_t}{\mu_t^3}\right] \\
\mathcal{I}_t\left(\mu_t\right) &= \mathbb{E}_{t-1}\left[\frac{a^2 y_t^2}{\mu_t^4}\right]+\mathbb{E}_{t-1}\left[\frac{a^2}{\mu_t^2}\right]-\mathbb{E}_{t-1}\left[2 \cdot \frac{a^2 y_t}{\mu_t^3}\right]
\end{aligned}
$$

Now we can use the information given in the question, namely

$$
\begin{aligned}
\mathbb{E}\left[Y_t \mid \mathcal{F}_{t-1}\right] &= \mu_t \\
E\left[Y_t^2 \mid \mathcal{F}_{t-1}\right] &= \frac{\mu_t^2(1+a)}{a} \\
\mathbb{V}\left[Y_t \mid \mathcal{F}_{t-1}\right] &= \frac{\mu_t^2}{a}
\end{aligned}
$$

$$
\begin{aligned}
\mathcal{I}_t\left(\mu_t\right) &= \mathbb{E}_{t-1}\left[\frac{a^2 y_t^2}{\mu_t^4}\right]+\mathbb{E}_{t-1}\left[\frac{a^2}{\mu_t^2}\right]-\mathbb{E}_{t-1}\left[2 \cdot \frac{a^2 y_t}{\mu_t^3}\right] \\
\mathcal{I}_t\left(\mu_t\right) &= a^2 \mathbb{E}_{t-1}\left[\frac{y_t^2}{\mu_t^4}\right]+a^2 \mathbb{E}_{t-1}\left[\frac{1}{\mu_t^2}\right]-2 a^2 \mathbb{E}_{t-1}\left[\frac{y_t}{\mu_t^3}\right]
\end{aligned}
$$

We know that all information contained in μ_t is observable given information at time $(t-1)$ namely

$$
\mathbb{E}\left[Y_t \mid \mathcal{F}_{t-1}\right]=\mu_t
$$

whereas we can take it outside the expectations operator, thus yielding

$$
\begin{aligned}
\mathcal{I}_t\left(\mu_t\right) &= \frac{a^2}{\mu_t^4} \mathbb{E}_{t-1}\left[y_t^2\right]+\frac{a^2}{\mu_t^2}-\frac{2 a^2}{\mu_t^3} \mathbb{E}_{t-1}\left[y_t\right] \\
\mathcal{I}_t\left(\mu_t\right) &= a^2\left(\frac{1}{\mu_t^4} \mathbb{E}_{t-1}\left[y_t^2\right]+\frac{1}{\mu_t^2}-\frac{2}{\mu_t^3} \mathbb{E}_{t-1}\left[y_t\right]\right) \\
\mathcal{I}_t\left(\mu_t\right) &= a^2\left(\frac{1}{\mu_t^4} \frac{\mu_t^2(1+a)}{a}+\frac{1}{\mu_t^2}-\frac{2}{\mu_t^3} \mu_t^2\right) \\
\mathcal{I}_t\left(\mu_t\right) &= a^2\left(\frac{1}{\mu_t^4} \frac{\mu_t^2(1+a)}{a}+\frac{1}{\mu_t^2}-2 \frac{1}{\mu_t^2}\right) \\
\mathcal{I}_t\left(\mu_t\right) &= a^2\left(\frac{(1+a)}{\mu_t^2 a}-\frac{1}{\mu_t^2}\right) \\
\mathcal{I}_t\left(\mu_t\right) &= \frac{a^2+a^3}{\mu_t^2 a}-\frac{a^2}{\mu_t^2} \\
\mathcal{I}_t\left(\mu_t\right) &= \frac{a+a^2}{\mu_t^2}-\frac{a^2}{\mu_t^2} \\
\mathcal{I}_t\left(\mu_t\right) &= \frac{a}{\mu_t^2}+\cancel{\frac{a^2}{\mu_t^2}-\frac{a^2}{\mu_t^2}} \\
\mathcal{I}_t\left(\mu_t\right) &= \frac{a}{\mu_t^2}
\end{aligned}
$$

We are given $d = 1/2$ whereas we have inverse square root scaling - thus we can now write up an expression for u_t

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

We know from slides (Lecture 9 slide 22) that for $d = 1/2$ we have $\widetilde{u}_t = u_t$. We introduce an exponential link function as y_t (VIX index) only can take on positive values, thus we have

$$\mu_t = \exp\left(\widetilde{\mu}_t\right)$$

Thus we can derive the Gamma GAS updating equation as the following and thus we have the requested GAS model for μ_t.

$$
\begin{aligned}
\widetilde{\mu}_t &= \omega + \alpha\widetilde{u}_{t-1} + \beta\widetilde{\mu}_{t-1} \\
\widetilde{\mu}_t &= \omega + \alpha\left[\sqrt{a}\left(\frac{y_{t-1}}{\mu_{t-1}} - 1\right)\right] + \beta\widetilde{\mu}_{t-1} \\
\widetilde{\mu}_t &= \omega + \alpha\left[\sqrt{a}\left(\frac{y_{t-1}}{\exp\left(\widetilde{\mu}_{t-1}\right)} - 1\right)\right] + \beta\widetilde{\mu}_{t-1}
\end{aligned}
$$

1.2 Point b) Derive a GAS model for μ_t

We know from point a) that the individual likelihood contributions can be written as

$$\ln\left[p\left(y_t \mid \mathbf{y}_{t:t-1}, \mu_t, a\right)\right] = -\ln\left(\Gamma\left(a\right)\right) + a\ln\left(a\right) + \left(a-1\right)\ln\left(y_t\right) - a\ln\left(\mu_t\right) - a\frac{y_t}{\mu_t}$$

We take the sum (log-product) to derive the full log-likelihood function

$$
\begin{aligned}
\ln\left[L\left(y_t \mid \mathbf{y}_{t:t-1}, \mu_t, a\right)\right] &= \sum_{t=1}^{T} -\ln\left(\Gamma\left(a\right)\right) + a\ln\left(a\right) + \left(a-1\right)\ln\left(y_t\right) - a\ln\left(\mu_t\right) - a\frac{y_t}{\mu_t} \\
\ln\left[L\left(y_t \mid \mathbf{y}_{t:t-1}, \mu_t, a\right)\right] &= T\left[-\ln\left(\Gamma\left(a\right)\right) + a\ln\left(a\right)\right] + \sum_{t=1}^{T}\left[\left(a-1\right)\ln\left(y_t\right) - a\ln\left(\mu_t\right) - a\frac{y_t}{\mu_t}\right]
\end{aligned}
$$

2 Question 1 - Computational Part

2.1 Point a)

Write a function to estimate the GAMMA-GAS model of the previous point using the Maximum Likelihood estimator. The function should accept a vector of observations and return the estimated parameters, the filtered means μ_t, for $t = 1, \ldots, T$, and the log likelihood evaluated at its maximum value. Assume that ω, α, and β are the intercept, score coefficient and autoregressive coefficient of the GAS process. Impose the following constraints during the optimization: $\omega \in [-0.5, 0.5], \alpha \in [0.001, 1.5], \beta \in [0.01, 0.999]$, $a \in [0.1, 300]$.

.

.

.

To estimate the Gamma GAS model using Maximum Likelihood I first derive the unconditional value of μ_t for initializing the model

$$
\begin{aligned}
&& \widetilde{\mu}_t &= \omega + \alpha \widetilde{u}_{t-1} + \beta \widetilde{\mu}_{t-1} \\
&\underset{\text{Recursive substitution}}{\Longrightarrow} & \widetilde{\mu}_t &= \omega + \alpha \widetilde{u}_{t-1} + \beta \left(\omega + \alpha \widetilde{u}_{t-2} + \beta \widetilde{\mu}_{t-2} \right) \\
&& \widetilde{\mu}_t &= \omega + \alpha \widetilde{u}_{t-1} + \beta \left(\omega + \alpha \widetilde{u}_{t-2} + \beta \left[\omega + \alpha \widetilde{u}_{t-3} + \beta \widetilde{\mu}_{t-3} \right] \right) \\
&& \widetilde{\mu}_t &= \omega + \alpha \widetilde{u}_{t-1} + \left(\beta\omega + \beta\alpha \widetilde{u}_{t-2} + \beta^2 \omega + \beta^2 \alpha \widetilde{u}_{t-3} + \beta^3 \widetilde{\mu}_{t-3} \right) \\
&& \widetilde{\mu}_t &= \omega + \beta\omega + \beta^2\omega + \alpha \left(\widetilde{u}_{t-1} + \beta \widetilde{u}_{t-2} + \beta^2 \widetilde{u}_{t-3} \right) + \beta^3 \widetilde{\mu}_{t-3} \\
&& \widetilde{\mu}_t &= \omega + \beta\omega + \beta^2\omega + \alpha \left(\widetilde{u}_{t-1} + \beta \widetilde{u}_{t-2} + \beta^2 \widetilde{u}_{t-3} \right) + \beta^3 \widetilde{\mu}_{t-3} \\
&& \widetilde{\mu}_t &= \omega \left(1 + \beta + \beta^2 \right) + \alpha \left(\widetilde{u}_{t-1} + \beta \widetilde{u}_{t-2} + \beta^2 \widetilde{u}_{t-3} \right) + \beta^3 \widetilde{\mu}_{t-3}
\end{aligned}
$$

Intermezzo, geometric series

$$
a + ar + ar^2 + ar^3 + ar^4 + \cdots = \sum_{k=0}^{\infty} ar^k = \frac{a}{1-r}, \ \textbf{for } |r| < 1 \tag{Geo 1}
$$

Thus we can use Equation (Geo 1) to reduce the first part, as $|\beta| < 1$ by construction

$$
\widetilde{\mu}_t = \frac{\omega}{1-\beta} + \alpha \left(\widetilde{u}_{t-1} + \beta \widetilde{u}_{t-2} + \beta^2 \widetilde{u}_{t-3} \right) + \beta^3 \widetilde{\mu}_{t-3}
$$

As $t \to \infty$ the part $\left[\beta^3 \widetilde{\mu}_{t-3} \right]$ goes becomes infinitesimal and thus can be ignored.

$$
\begin{aligned}
\widetilde{\mu}_t &= \frac{\omega}{1-\beta} + \alpha \left(\widetilde{u}_{t-1} + \beta \widetilde{u}_{t-2} + \beta^2 \widetilde{u}_{t-3} \right) + \underbrace{\beta^3 \widetilde{\mu}_{t-3}}_{=0, \text{ as } t \to \infty} \\
\widetilde{\mu}_t &= \frac{\omega}{1-\beta} + \alpha \left(\widetilde{u}_{t-1} + \beta \widetilde{u}_{t-2} + \beta^2 \widetilde{u}_{t-3} \right) \\
\widetilde{\mu}_t &= \frac{\omega}{1-\beta} + \alpha \sum_{s=0}^{\infty} \beta^s \widetilde{u}_{t-1-s}
\end{aligned}
$$

The sum can be interpreted as,

$$
\begin{aligned}
s = 0 &\quad \Longrightarrow \quad \widetilde{u}_{t-1} \\
s = 1 &\quad \Longrightarrow \quad \beta \widetilde{u}_{t-2} \\
s = 2 &\quad \Longrightarrow \quad \beta^2 \widetilde{u}_{t-3} \\
\vdots &
\end{aligned}
$$

Now we can take the unconditional expectation

$$
\begin{aligned}
\mathbb{E}\left[\widetilde{\mu}_t \right] &= \mathbb{E}\left[\frac{\omega}{1-\beta} + \alpha \sum_{s=0}^{\infty} \beta^s \widetilde{u}_{t-1-s} \right] \\
\mathbb{E}\left[\widetilde{\mu}_t \right] &= \mathbb{E}\left[\frac{\omega}{1-\beta} \right] + \mathbb{E}\left[\alpha \sum_{s=0}^{\infty} \beta^s \widetilde{u}_{t-1-s} \right]
\end{aligned}
$$

We know from slide 18 of Lecture 9 that the forcing variable u_t has zero expectation. Thus we now have an expression for the unconditional expectation

$$
\begin{aligned}
\mathbb{E}\left[\widetilde{\mu}_t \right] &= \mathbb{E}\left[\frac{\omega}{1-\beta} \right] + \underbrace{\mathbb{E}\left[\alpha \sum_{s=0}^{\infty} \beta^s \widetilde{u}_{t-1-s} \right]}_{=0} \\
\mathbb{E}\left[\widetilde{\mu}_t \right] &= \frac{\omega}{1-\beta}
\end{aligned}
$$

3 Question 1 - Empirical Analysis

3.1 Point d)

Engle and Gallo (2006) introduced a series of Multiplicative Error Models to model positive valued series like the VIX. In their simpler specification, they assume that:

$$Y_t \mid \mathcal{F}_{t-1} \sim \mathcal{G}a\left(\mu_t, a\right)$$

and

$$\mu_t = \kappa + \eta y_{t-1} + \phi \mu_{t-1}$$

- Write a function to estimate the MEM model of Engle and Gallo (2006). Impose these constraints on the parameters of the MEM model: $\kappa \in [0.1, 10], \eta \in [0.01, 0.99], \phi \in [0.01, 0.99]$, and $a \in [0.1, 300]$.

.
.
.
.
.

To estimate this model using the Maximum Likelihood estimator I need to derive the unconditional mean of the model following the procedure in Point a) of Computational part.

$$
\begin{aligned}
& \mu_t = \kappa + \eta y_{t-1} + \phi \mu_{t-1} \\
\underset{\text{Recursive substitution}}{\Longrightarrow} \quad & \mu_t = \kappa + \eta y_{t-1} + \phi \left[\kappa + \eta y_{t-2} + \phi \mu_{t-2}\right] \\
& \mu_t = \kappa + \eta y_{t-1} + \phi \left[\kappa + \eta y_{t-2} + \phi \left(\kappa + \eta y_{t-3} + \phi \mu_{t-3}\right)\right] \\
& \mu_t = \kappa + \eta y_{t-1} + \phi \left[\kappa + \eta y_{t-2} + \phi \kappa + \phi \eta y_{t-3} + \phi^2 \mu_{t-3}\right] \\
& \mu_t = \kappa + \eta y_{t-1} + \phi \kappa + \phi \eta y_{t-2} + \phi^2 \kappa + \phi^2 \eta y_{t-3} + \phi^3 \mu_{t-3} \\
\underset{\text{Re-arrange}}{\Longrightarrow} \quad & \mu_t = \kappa + \phi \kappa + \phi^2 \kappa + \eta \left[y_{t-1} + \phi y_{t-2} + \phi^2 y_{t-3}\right] + \phi^3 \mu_{t-3} \\
& \mu_t = \kappa \left(1 + \phi + \phi^2\right) + \eta \left[y_{t-1} + \phi y_{t-2} + \phi^2 y_{t-3}\right] + \phi^3 \mu_{t-3} \\
& \mu_t = \underbrace{\frac{\kappa}{1 - \phi}}_{\text{Using Geo 1}} + \eta \left[y_{t-1} + \phi y_{t-2} + \phi^2 y_{t-3}\right] + \phi^3 \mu_{t-3} \\
& \mu_t = \frac{\kappa}{1 - \phi} + \eta \left[y_{t-1} + \phi y_{t-2} + \phi^2 y_{t-3}\right] + \underbrace{\phi^3 \mu_{t-3}}_{=0, \text{ as } t \to \infty} \\
& \mu_t = \frac{\kappa}{1 - \phi} + \eta \left[y_{t-1} + \phi y_{t-2} + \phi^2 y_{t-3}\right] \\
& \mu_t = \frac{\kappa}{1 - \phi} + \eta \sum_{s=0}^{\infty} \phi^s y_{t-s-1}
\end{aligned}
$$

Now we take the unconditional expectation

$$\mathbb{E}\left[\mu_t\right] = \mathbb{E}\left[\frac{\kappa}{1-\phi} + \eta\sum_{s=0}^{\infty}\phi^s y_{t-s-1}\right]$$
$$\mathbb{E}\left[\mu_t\right] = \mathbb{E}\left[\frac{\kappa}{1-\phi}\right] + \mathbb{E}\left[\eta\sum_{s=0}^{\infty}\phi^s y_{t-s-1}\right]$$
$$\mathbb{E}\left[\mu_t\right] = \frac{\kappa}{1-\phi} + \mathbb{E}\left[\eta\sum_{s=0}^{\infty}\phi^s y_{t-s-1}\right]$$
$$\mathbb{E}\left[\mu_t\right] = \frac{\kappa}{1-\phi} + \mathbb{E}\left[\eta\sum_{s=0}^{\infty}\phi^s y_{t-s-1}\right]$$
$$\mathbb{E}\left[\mu_t\right] = \frac{\kappa}{1-\phi} + \mathbb{E}\left[\eta\sum_{s=0}^{\infty}\phi^s \underbrace{y_{t-s-1}}_{\mathbb{E}[y_{t-s-1}]=\mathbb{E}[\mu_t]\text{ uncon.}}\right]$$
$$\mathbb{E}\left[\mu_t\right] = \frac{\kappa}{1-\phi} + \eta\sum_{s=0}^{\infty}\phi^s\mathbb{E}\left[\mu_t\right]$$
$$\mathbb{E}\left[\mu_t\right] = \frac{\kappa}{1-\phi} + \underbrace{\frac{\eta}{1-\phi}\mathbb{E}\left[\mu_t\right]}_{\text{Geo. series}}$$
$$\mathbb{E}\left[\mu_t\right] - \frac{\eta}{1-\phi}\mathbb{E}\left[\mu_t\right] = \frac{\kappa}{1-\phi}$$
$$\mathbb{E}\left[\mu_t\right]\left(1-\frac{\eta}{1-\phi}\right) = \frac{\kappa}{1-\phi}$$
$$\mathbb{E}\left[\mu_t\right] = \frac{\frac{\kappa}{1-\phi}}{\left(1-\frac{\eta}{1-\phi}\right)}$$
$$\underset{\text{substitute }\frac{(1-\phi)}{(1-\phi)}=1}{\Longrightarrow} \quad \mathbb{E}\left[\mu_t\right] = \frac{\frac{\kappa}{1-\phi}}{\left(\frac{(1-\phi)}{(1-\phi)}-\frac{\eta}{1-\phi}\right)}$$
$$\mathbb{E}\left[\mu_t\right] = \frac{\frac{\kappa}{1-\phi}}{\frac{1-\phi-\eta}{1-\phi}}$$
$$\mathbb{E}\left[\mu_t\right] = \frac{\kappa}{1-\phi-\eta}$$

4 Question 2 - Methodology

4.1 Point a)

Consider the bivariate random vector $\mathbf{Y}_t = (Y_{1,t}, Y_{2,t})'$, where $Y_{1,t}$ and $Y_{2,t}$ are the GSPC and DJI returns at time t, respectively.

- Assume that $\mathbf{Y}_t \mid \mathcal{F}_{t-1}$ is bivariate Gaussian with mean $\mathbf{0}$ and variance covariance matrix Σ_t
  - Derive a DCC model for Σ_t assuming that each marginal process is GARCH $(1,1)$.
  - Clearly state the constraints of the model.
  - Detail how the likelihood factorizes and what this implies for the estimation of the model parameters.
  - Obtain the Constant correlation model (CCC) as a special case of the DCC

.
.
.
.

4.2 Point a, 1) Derive the DCC model

We have that $\mathbf{Y}_t \mid \mathcal{F}_{t-1}$ follows

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

We can write up the standardized returns as

$$\boldsymbol{\eta}_t = \mathbf{D}_t^{-1/2}\mathbf{y}_t$$

Where $\mathbf{D}_t$ is a diagonal matrix with elements modeled from univariate GARCH processes (variances)

$$\mathbf{D}_t = \operatorname{diag}\left(\operatorname{Var}_{t-1}\left(y_{1,t}\right), \ldots, \operatorname{Var}_{t-1}\left(y_{N,t}\right)\right)$$

We know that we can write the conditional GARCH process as

$$\begin{aligned}
\mathbf{Y}_t &= \left(y_{1,t}, \ldots y_{N,t}\right) \\
y_{i,t} &= \sigma_{i,t}\varepsilon_t \\
\sigma_{i,t}^2 &= \omega + \alpha_i y_{i,t-1}^2 + \beta\sigma_{i,t-1}^2
\end{aligned}$$

Thus we can also write the diagonal matrix more explicitly as

$$\mathbf{D}_t = \operatorname{diag}\left(\sigma_{i,t}^2, \ldots, \sigma_{i,N}^2\right)$$

We know that in the DCC framework we can write

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

Where $\mathbf{R}_t$ is an evaluation matrix with elements $\{\rho_{ij,t}\}$. We can write this as

$$E_{t-1}\left(\boldsymbol{\eta}_t\boldsymbol{\eta}_t'\right) = \mathbf{D}_t^{-1/2}\mathbf{H}_t\mathbf{D}_t^{-1/2} = \mathbf{R}_t = \{\rho_{ij,t}\}$$

We require that $\mathbf{R}_t$ is positively defined we can employ following transformation

$$\mathbf{R}_t = \widetilde{\mathbf{Q}}_t^{-1/2}\mathbf{Q}_t\widetilde{\mathbf{Q}}_t^{-1/2}$$

where $\widetilde{\mathbf{Q}}_t$ is a diagonal matrix with elements $\tilde{q}_{ii,t} = q_{ii,t}$. This implies that the conditional correlations are

$$\rho_{ij,t} = \frac{q_{ij,t}}{\sqrt{\widetilde{q}_{ii,t}\widetilde{q}_{jj,t}}}$$

Where $q_{ij,t}$ are assumed to follow a GARCH$(1,1)$ model

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

$\overline{\boldsymbol{Q}}$ is fixed to the emperical evaluation of $\boldsymbol{\eta}_t\overline{\boldsymbol{Q}}_t = \overline{\boldsymbol{R}}_t = \operatorname{cor}\left(z_t\right)$ used in the CCC model

.

To derive the log-likelihood we need to compute the density of $\mathbf{Y}_t \mid \mathcal{F}_{t-1}$. This bivariate form of normal distribution comes from wiki.

$$\begin{aligned}
& p\left(\mathbf{Y}_t \mid \mathcal{F}_{t-1}\right) = \frac{1}{2\pi\sqrt{|\Sigma_t|}}\exp\left\{-\frac{1}{2}\left[\left(\mathbf{Y}_t-\mu\right)'\Sigma_t^{-1}\left(\mathbf{Y}_t-\mu\right)\right]\right\} \\
\underset{\mu=0}{\Longrightarrow}\quad & p\left(\mathbf{Y}_t \mid \mathcal{F}_{t-1}\right) = \frac{1}{2\pi\sqrt{|\Sigma_t|}}\exp\left\{-\frac{1}{2}\left[\left(\mathbf{Y}_t\right)'\Sigma_t^{-1}\left(\mathbf{Y}_t\right)\right]\right\}
\end{aligned}$$

Writing up the likelihood function and taking the log

$$\begin{aligned}
\mathcal{L} &= \prod_{t=1}^{T}\frac{1}{2\pi\sqrt{|\Sigma_t|}}\exp\left\{-\frac{1}{2}\left[\left(\mathbf{Y}_t\right)'\Sigma_t^{-1}\left(\mathbf{Y}_t\right)\right]\right\} \\
\ln\{\mathcal{L}\} &= \sum_{t=1}^{T} -\ln(2\pi) - \frac{1}{2}\ln\left(|\Sigma_t|\right) - \frac{1}{2}\left(\mathbf{Y}_t\right)'\Sigma_t^{-1}\left(\mathbf{Y}_t\right) \\
\ln\{\mathcal{L}\} &= -\frac{1}{2}\sum_{t=1}^{T} 2\ln(2\pi) + \ln\left(|\Sigma_t|\right) + \left(\mathbf{Y}_t\right)'\Sigma_t^{-1}\left(\mathbf{Y}_t\right)
\end{aligned}$$

Now we want to factorize the model

$$\ln\{\mathcal{L}\} = -\frac{1}{2}\sum_{t=1}^{T} 2\ln(2\pi) + \ln\left(|\underbrace{\Sigma_t}_{D_t^{1/2}R_tD_t^{1/2}}|\right) + (\mathbf{Y}_t)'\,\Sigma_t^{-1}(\mathbf{Y}_t)$$

$$\ln\{\mathcal{L}\} = -\frac{1}{2}\sum_{t=1}^{T} 2\ln(2\pi) + \ln\left(\left|\mathbf{D}_t^{1/2}\mathbf{R}_t\mathbf{D}_t^{1/2}\right|\right) + (\mathbf{Y}_t)'\,\mathbf{D}_t^{-1/2}\mathbf{R}_t^{-1}\mathbf{D}_t^{-1/2}(\mathbf{Y}_t)$$

$$\ln\{\mathcal{L}\} = -\frac{1}{2}\sum_{t=1}^{T} 2\ln(2\pi) + \frac{1}{2}\ln(|\mathbf{D}_t|) + \ln(|\mathbf{R}_t|) + \frac{1}{2}\ln(|\mathbf{D}_t|) + \underbrace{(\mathbf{Y}_t)'\,\mathbf{D}_t^{-1/2}}_{=\boldsymbol{\eta}_t'}\mathbf{R}_t^{-1}\underbrace{\mathbf{D}_t^{-1/2}(\mathbf{Y}_t)}_{=\boldsymbol{\eta}_t}$$

$$\ln\{\mathcal{L}\} = -\frac{1}{2}\sum_{t=1}^{T} 2\ln(2\pi) + \ln(|\mathbf{D}_t|) + \ln(|\mathbf{R}_t|) + \boldsymbol{\eta}_t'\mathbf{R}_t^{-1}\boldsymbol{\eta}_t$$

Adding and subtracting

$$(\mathbf{Y}_t)'\,\mathbf{D}_t^{-1/2}\mathbf{D}_t^{-1/2}(\mathbf{Y}_t) = \boldsymbol{\eta}_t'\boldsymbol{\eta}_t$$

Thus yielding

$$\ln\{\mathcal{L}\} = -\frac{1}{2}\sum_{t=1}^{T} 2\ln(2\pi) + \ln(|\mathbf{D}_t|) + (\mathbf{Y}_t)'\,\mathbf{D}_t^{-1/2}\mathbf{D}_t^{-1/2}(\mathbf{Y}_t) - \boldsymbol{\eta}_t'\boldsymbol{\eta}_t + \ln(|\mathbf{R}_t|) + \boldsymbol{\eta}_t'\mathbf{R}_t^{-1}\boldsymbol{\eta}_t$$

$$\ln\{\mathcal{L}\} = -\frac{1}{2}\sum_{t=1}^{T} 2\ln(2\pi) + \ln(|\mathbf{D}_t|) + (\mathbf{Y}_t)'\,\mathbf{D}_t^{-1}(\mathbf{Y}_t) - \boldsymbol{\eta}_t'\boldsymbol{\eta}_t + \ln(|\mathbf{R}_t|) + \boldsymbol{\eta}_t'\mathbf{R}_t^{-1}\boldsymbol{\eta}_t$$

$$\ln\{\mathcal{L}\} = -\frac{1}{2}\sum_{t=1}^{T}\left\{2\ln(2\pi) + \ln(|\mathbf{D}_t|) + (\mathbf{Y}_t)'\,\mathbf{D}_t^{-1}(\mathbf{Y}_t)\right\} - \frac{1}{2}\sum_{t=1}^{T}\left\{\boldsymbol{\eta}_t'\mathbf{R}_t^{-1}\boldsymbol{\eta}_t - \boldsymbol{\eta}_t'\boldsymbol{\eta}_t + \ln(|\mathbf{R}_t|)\right\}$$

Now we can decompose this into a Volatility component and a Correlation component

$$\ln\{\mathcal{L}_V(\theta)\} = -\frac{1}{2}\sum_{t=1}^{T}\left\{2\ln(2\pi) + \ln(|\mathbf{D}_t|) + (\mathbf{Y}_t)'\,\mathbf{D}_t^{-1}(\mathbf{Y}_t)\right\}$$

$$\ln\{\mathcal{L}_C(\theta,\phi)\} = -\frac{1}{2}\sum_{t=1}^{T}\left\{\boldsymbol{\eta}_t'\mathbf{R}_t^{-1}\boldsymbol{\eta}_t - \boldsymbol{\eta}_t'\boldsymbol{\eta}_t + \ln(|\mathbf{R}_t|)\right\}$$

Where θ denotes the parameters in $\mathbf{D}_t$ and ϕ the parameters in $\mathbf{R}_t$

4.3 Point a, 2) Clearly state the constraints of the model

We require $\boldsymbol{Q}_t$ to be positive-definite. This ensures that when mapping to $\mathbf{R}_t$ this will also be positive definite.

TODO: Er alpha + beta < 1 en condition her?

4.4 Point a, 3) Detail how the likelihood factorizes and what this implies for the estimation of the model parameters.

We can factorize the likelihood function into a volatility and correlation component. This implies that we can estimate θ and ϕ in a two step process.

$$\ln\{\mathcal{L}_V(\theta)\} = -\frac{1}{2}\sum_{t=1}^{T}\left\{2\ln(2\pi) + \ln(|\mathbf{D}_t|) + (\mathbf{Y}_t)'\,\mathbf{D}_t^{-1}(\mathbf{Y}_t)\right\}$$

$$\ln\{\mathcal{L}_C(\theta,\phi)\} = -\frac{1}{2}\sum_{t=1}^{T}\left\{\boldsymbol{\eta}_t'\mathbf{R}_t^{-1}\boldsymbol{\eta}_t - \boldsymbol{\eta}_t'\boldsymbol{\eta}_t + \ln(|\mathbf{R}_t|)\right\}$$

4.5 Point a, 4) Obtain the Constant correlation model (CCC) as a special case of the DCC

We know that $\mathbf{R}_t$ is constant $\mathbf{R}_t = \mathbf{R}$ in the CCC model. Thus we can just generalize the DCC model to account for a constant $\mathbf{R}$.