1) Joint Gaussianity implies that each marginal is Gaussian. Specifically if $\begin{pmatrix} X \\ Y \end{pmatrix} \sim N\left(\begin{pmatrix} \mu_x \\ \mu_y \end{pmatrix}, \begin{pmatrix} \sigma_x^2 & \sigma_{xy} \\ \sigma_{xy} & \sigma_y^2 \end{pmatrix}\right)$ we have that

$$X \sim N(\mu_x, \sigma_x^2) \quad \text{and} \quad Y \sim N(\mu_x, \sigma_y^2)$$

The first two reasons are:

i) X and Y do not exhibit excess of kurtosis, i.e.

$$\frac{\mathbb{E}[(X-\mu_x)^4]}{\mathbb{E}[(X-\mu_x)^2]^2} = \frac{\mathbb{E}[(Y-\mu_y)^4]}{\mathbb{E}[(Y-\mu_y)^2]^2} = 3$$

ii) X and Y do not exhibit positive/negative skewness, i.e.

$$\frac{\mathbb{E}[(X-\mu_x)^3]}{\mathbb{E}[(X-\mu_x)^2]^{3/2}} = \frac{\mathbb{E}[(Y-\mu_y)^3]}{\mathbb{E}[(Y-\mu_y)^2]^{3/2}} = 0$$

The third reason concerns the dependence structure induced by the joint Gaussian assumption. Specifically, X and Y do not display positive/negative tail dependence, i.e.

$$\lim_{z \to -\infty} P(X \leq z \mid Y \leq z) = \lim_{z \to \infty} P(X \geq z \mid Y \geq z) = 0$$

this is in contrast with empirical evidence which suggests that extreme negative/positive returns (Crisis/Booms) of one asset (Soy, Y) have an effect on another asset (Soy, X).

2)

i)

The model is

$$y_t = \Sigma_t^{1/2} z_t \quad , \quad z_t \overset{iid}{\sim} T_\nu(0, I)$$

where $T_\nu(0, I)$ represent a p-variate Student's t distribution with variance I (the identity matrix), mean 0, and $\nu > 2$ degrees of freedom. The variance/covariance matrix of y_t given the past is Σ_t, and is factorized as

$\Sigma_t = D_t^{1/2} R_t D_t^{1/2}$, where D_t is a diagonal matrix with typical element $\sigma_{it}^2 = Var(y_{it} \mid \mathcal{F}_{t-1})$, for example, σ_{it}^2 can follow a GARCH(1,1) process:

$$\sigma_{it}^2 = \omega + \alpha y_{it-1}^2 + \beta \sigma_{it-1}^2 .$$

R_t is a correlation matrix with typical element $\rho_{ijt} = Cor(y_{it}, y_{jt} \mid \mathcal{F}_{t-1})$.

The DCC model assumes that

$R_t = \tilde{Q}_t^{-1/2} Q_t \tilde{Q}_t^{-1/2}$, where $\tilde{Q}_t$ is a diagonal matrix which contains the diagonal elements of Q_t. The matrix Q_t is defined as

$$Q_t = \bar{Q}(1-a-b) + a \eta_{t-1} \eta_{t-1}' + b Q_{t-1}$$

where $\eta_t = D_t^{-1/2} y_t$, and $\bar{Q}$ is fixed to the empirical correlation of η_t.

To derive the log likelihood we need to compute the density of $y_t \mid \mathcal{F}_{t-1}$.

We first recover the density of z_t by setting $\mu = 0$, $\Psi = I\,\frac{\nu-2}{\nu}$ in the equation of the exercise:

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

where we have used $|aX| = a^p|X|$

and $(aX)^{-1} = \frac{1}{a}X^{-1}$, $a \neq 0$

the density of $y_t \mid \mathcal{F}_t$ is then

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

the log density is

$$\begin{aligned}\log p(y_t \mid \mathcal{F}_{t-1}) = {} & \ln\Gamma\left(\frac{\nu+p}{2}\right) - \ln\Gamma\left(\frac{\nu}{2}\right) - \frac{p}{2}\ln\pi \; + \\ & -\frac{p}{2}\ln(\nu-2) - \ln|D_t| - \frac{1}{2}\ln|R_t| \; + \\ & -\frac{\nu+p}{2}\ln\left[1+\frac{y_t'\left(D_t^{1/2}R_tD_t^{1/2}\right)^{-1}y_t}{\nu-2}\right]\end{aligned}$$

where we used the result that

$$|ABC| = |A||B||C| \quad \text{and}$$

$$\log(|AB|) = \log(|A||B|) = \log|A| + \log|B|$$

The log likelihood for a time series of lenght T is

$$\mathcal{L}\mathcal{L}K = T\Big[\ln\Gamma\Big(\frac{v+p}{2}\Big) - \ln\Gamma\Big(\frac{v}{2}\Big) - \frac{p}{2}\ln\pi + - \frac{p}{2}\ln(v-2)\Big] - \sum_{t=1}^{T} \ln|D_t| + - \frac{1}{2}\sum_{t=1}^{T}\ln|R_t| - \frac{v+p}{2}\sum_{t=1}^{T}\ln\Big[1 + \frac{y_t'\big(D_t^{1/2}R_tD_t^{1/2}\big)^{-1}y_t}{v-2}\Big]$$

ii)

The log likelihood function cannot be factorized in two parts because of this term

$$\sum_{t=1}^{T}\ln\Big[1 + \frac{y_t'\big(D_t^{1/2}R_tD_t^{1/2}\big)^{-1}y_t}{v-2}\Big]$$

iii) The model can be estimated in two steps using the Quasi Maximum Estimator for Σ_t, and then the Maximum Likelihood estimator for ν.

The estimator proceeds as follow:

- Estimate Σ_t by maximizing a Gaussian likelihood. We know that we can split this estimation problem in $p+1$ estimations.

- Compute $\hat{z}_t = \hat{\Sigma}_t^{-1/2} y_t$

- Maximize

$$\mathcal{LLK}_\nu = T\left[\ln \Gamma\left(\frac{\nu+p}{2}\right) - \ln \Gamma\left(\frac{\nu}{2}\right) - \frac{p}{2}\ln \pi - \frac{p}{2}\ln(\nu-2)\right] - \frac{\nu+p}{2}\sum_{t=1}^{T} \ln\left[1 + \frac{\hat{z}_t'\hat{z}_t}{\nu-2}\right]$$

with respect to ν.

This estimator is consistent but inefficient, see the discussion in 6.2

of the book.