Stochastic Volatility Derivations

Baseline SV Model

$$
\begin{aligned}
y_t &= \exp\left\{(h_t + \mu)/2\right\} \epsilon_t && (1)\\
h_t &= \phi h_{t-1} + \sigma v_t && (2)\\
\begin{bmatrix} \epsilon_t \\ v_t \end{bmatrix} &\sim N\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}\right) && (3)
\end{aligned}
$$

Baseline SV Model with Correlated Innovations

Days where returns go down tend to also be days where volatility goes up. Unlike predicted return distributions of the baseline model, predicted return distributions of this model have a negative skew that is more representative of what we see in real life. If you fit this model in Stan the posterior of ρ is heavily concentrated away from zero towards negative values.

This model is misleadingly called a leverage term in the SV literature but don't confuse with the leverage that comes from holding a 3x vs 1x ETF).

$$
\begin{aligned}
y_t &= \exp\left\{(h_t + \mu)/2\right\} \epsilon_t && (4)\\
h_t &= \phi h_{t-1} + \sigma v_t && (5)\\
\begin{bmatrix} \epsilon_t \\ v_t \end{bmatrix} \Bigg| \rho &\sim N\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 & \rho \\ \rho & 1 \end{bmatrix}\right) && (6)
\end{aligned}
$$

Fouque Multi Time Scale Volatiltiy Model

Original model specification from Fouque with added correlation between return innovation and volatility innovations. This model says that the volatility is actually a sum of K AR1 processes with different time scales of their mean reverse parametners. h_t, v_t, and ρ become vectors now.

$$
\begin{aligned}
y_t &= \exp\left\{(1' h_t + \mu)/2\right\} \epsilon_t && (7)\\
h_t &= \Phi h_{t-1} + \Sigma^{1/2} v_t && (8)\\
\begin{bmatrix} \epsilon_t \\ v_t \end{bmatrix} \Bigg| \rho &\sim N\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 & \rho^T \\ \rho & I_K \end{bmatrix}\right) && (9)
\end{aligned}
$$

We now derive the likelihood. First we use the last two equations to get the likelihood of $\epsilon_t, h_t | \rho, \Phi, h_{t-1}, \Sigma$. This consists of just applying the transformation

$$
\begin{bmatrix} 1 & 0^T \\ 0 & \Sigma^{1/2} \end{bmatrix} \begin{bmatrix} \epsilon_t \\ v_t \end{bmatrix} + \begin{bmatrix} 0 \\ \Phi h_{t-1} \end{bmatrix} \tag{10}
$$

We abbreviate the model parameters to be estimated as $\Omega := (\rho, \Phi, \Sigma)$. Since this is just an affine transformation of a multivariate Gaussian we have

$$\begin{bmatrix}\epsilon_t \\ h_t\end{bmatrix}\Bigg| h_{t-1}, \Omega \quad \sim \quad N\left(\begin{bmatrix}0 \\ \Phi h_{t-1}\end{bmatrix}, \begin{bmatrix}1 & \rho^T \Sigma^{1/2} \\ \Sigma^{1/2}\rho & \Sigma\end{bmatrix}\right). \tag{11}$$

This gives us the joint distribution $p(\epsilon_t, h_t | h_{t-1}, \Omega)$. The entire likelihood is $p(h_0|\Omega)\prod_{t=1}^{T} p(\epsilon_t, h_t|h_{t-1}, \Omega)$. We next factorize the individual terms in to

$$\begin{aligned} p(\epsilon_t, h_t|h_{t-1}, \Omega) &= p(\epsilon_t|h_t, h_{t-1}, \Omega)p(h_t|h_{t-1}, \Omega) & (12) \\ &= p(\epsilon_t|h_t, \Omega)p(h_t|h_{t-1}, \Omega). & (13) \end{aligned}$$

Because the joint distribution is Gaussian, we have these densities in closed form:

$$\begin{aligned} \epsilon_t|h_t, \Omega &\sim N(\rho^T\Sigma^{-1/2}(h_t - \Phi h_{t-1}), 1 - \rho^T\rho) & (14) \\ h_t|h_{t-1}, \Omega &\sim N(\Phi h_{t-1}|\Sigma). & (15) \end{aligned}$$

The entire likelihood of all time points now becomes

$$\begin{aligned} p(\epsilon_0, \cdots, \epsilon_T, h_0, \cdots, h_T|\Omega) &= p(h_0|\Omega)\prod_{t=1}^{T} p(\epsilon_t, h_t|h_{t-1}, \Omega) & (16) \\ &= p(h_0|\Omega)\prod_{t=1}^{T} p(\epsilon_t|h_t, \Omega)p(h_t|h_{t-1}, \Omega) & (17) \\ &= \left(\prod_{t=1}^{T} p(\epsilon_t|h_t, \Omega)\right)\left(\prod_{t=1}^{T} p(h_t|h_{t-1}, \Omega)\right)p(h_0|\Omega). & (18) \end{aligned}$$

Lastly, since y_t is a linear transformation of ϵ_t, and we know that $\epsilon_t|h_t, \Omega$ is normally distributed, we get the distribution of $y_t|h_t, \Omega$. If we define $s_t := \exp\{(1'h_t + \mu)/2\}$, it is simply

$$y_t|h_t, \Omega \quad \sim \quad N\left(\rho^T\Sigma^{-1/2}(h_t - \Phi h_{t-1})s_t, (1 - \rho^T\rho)s_t^2\right). \tag{19}$$

This gives us the final likelihood

$$p(y_0, \cdots, y_T, h_0, \cdots, h_T|\Omega) = \left(\prod_{t=1}^{T} p(y_t|h_t, \Omega)\right)\left(\prod_{t=1}^{T} p(h_t|h_{t-1}, \Omega)\right)p(h_0|\Omega). \tag{20}$$

ARMA Extension

$$h_t = \Phi h_{t-1} + \Theta v_{t-1} + \Sigma^{1/2} v_t \tag{21}$$

Higher-Order ARMA Extension

$$h_t = \sum_{i=1}^{p} (\Phi h_{t-i} + \Theta v_{t-i}) + \Sigma^{1/2} v_t \quad (22)$$