Scale Mixture Of Normals (SMiN) Representation

In foreign exchange markets such as the EUR/USD, market moves are widely regarded to be symmetric in nature [**tankov2003financial**]. This motivates the selection of the parameter $\beta = 0$. Instead of considering the general α-stable process, we can instead restrict ourselves to a more analytically tractable sub-class, the a Symmetric α-stable (SαS) process instead.

For $\beta = 0$, we have a convenient Scale Mixture of Normals (SMiN) representation based off the product property of α-stable distributions: If X and Y are independent random variables with $\lambda_1 \sim S_{\alpha/2}(1,1,0)$ and $\lambda_2 \sim S_2(1,0,0) = \mathcal{N}(0,1)$, then $\lambda_1\lambda_2 \sim S_\alpha(1,0,0)$.

We can then convert the discretised state space model specified in **??** into SMiN form as specified in Equation 1 below.

$$\underbrace{\begin{bmatrix} x_{1,t+\delta t} \\ x_{2,t+\delta_t} \end{bmatrix}}_{\mathbf{S_{t+\delta t}}} = \underbrace{\begin{bmatrix} 1 & \delta t \\ 0 & e^{\theta \delta t} \end{bmatrix}}_{\mathbf{A}} \underbrace{\begin{bmatrix} x_{1,t} \\ x_{2,t} \end{bmatrix}}_{\mathbf{S_t}} + \underbrace{\begin{bmatrix} 0 \\ \sigma_t \sqrt{\lambda_{1,t-1}} \end{bmatrix}}_{\mathbf{b}} \lambda_{2,t}; \qquad \lambda_{2,t} \sim \mathcal{N}(0,1),\ \lambda_{1,t} \sim S_{\alpha/2}(1,1,0) \tag{1}$$

This means that conditional upon us observing λ_1, λ_2 is gaussian, making inference much more tractable. We exploit this fact by designing a Rao-Blackwellised Particle Filter to sample λ_1 by using a simple particle filter, and propogating the states using a Kalman Filter conditioned upon the sampled λ_1 of the particle.

We take a short detour to properly formulate the inference problem using the SMiN representation here. In the inference problem, we are seeking to infer the state variables $\mathbf{s}_t$ given the observations y_t. We assume model dynamics as follows:

$$\begin{aligned} \mathbf{s}_t &= \mathbf{A}\mathbf{s}_{t-\delta t} + \mathbf{b}\sqrt{\lambda_t}\eta_t \\ y_t &= \mathbf{C}\mathbf{s}_t + d\epsilon_t \\ \lambda_t &\sim S_{\alpha/2}(1,1,0) \\ \eta_t, \epsilon_t &\sim \mathcal{N}(0,1) \end{aligned} \tag{2}$$

It will also be useful to use the following expression for y_t:

$$
\begin{aligned}
y_t &= \mathbf{CA}\mathbf{s}_{t-\delta t} + \underbrace{\begin{bmatrix} \mathbf{Cb}\sqrt{\lambda_t} & d \end{bmatrix}}_{\mathbf{e}_t} \underbrace{\begin{bmatrix} \eta_t \\ \epsilon_t \end{bmatrix}}_{\mathbf{n_t}} \\
\mathbf{n}_t &\sim \mathcal{N}(\mathbf{0}, \mathbf{I}) \\
\lambda_t &\sim S_{\alpha/2}(1, 1, 0)
\end{aligned} \tag{3}
$$

This transforms our state equation form into a (nearly?) α-stable sub-gaussian form. (This is NOT in α-stable sub-gaussian form, see: $<$ α-stable sub-gaussian definition $>$ for reference)

Generic Rao-Blackwellised Particle Filter (RBPF)

For Rao-Blackwellised Particle Filtering, we partition the state vector into gaussian and non-gaussian components. We can then use standard Kalman Filtering to obtain optimal estimates for the gaussian state components, after obtaining estimates for the non-gaussian state components.

We partition the state from Equation 1 as $\vec{s}_t = (\vec{x}_t, \lambda_t)$.

Again, detailed treatment of the Rao-Blackwellised Particle Filter is omitted, but a short summary of the recursive steps taken every time step are presented below.

Adaptive Sampling

One way of attempting to improve the performance of the RBPF is to form a better importance sampling distribution. In this section, we seek to form a better importance sampling distribution for $\pi(\lambda_t | y_{1:t}, \lambda_{0:t-1})$.

For optimality, we want $\pi(\lambda_t | y_{1:t} \lambda_{0:t-1}) \approx p(\lambda_t | y_{1:t}, \lambda_{0:t-1})$.

It is hard to formulate an analytical form for this density directly. To get around this, we can instead form the analytical joint distribution for $p(y_t, \lambda_t | y_{1:t-1}, \lambda_{0:t-1})$. The complete derivation is given in the appendix, with the key steps highlighted below.

We first augment the probability density given above with $x_{0:t-1}$. This allows us to use results from the previously-performed Kalman Filter step.

$$
\begin{aligned}
& p(y_t.\lambda_t, x_{0:t-1}|y_{1:t-1}, \lambda_{0:t-1}) \\
\propto \;& p(y_t|y_{1:t-1}, \lambda_{0:t}, x_{0:t-1})p(x_{0:t-1}|y_{1:t-1}, \lambda_{0:t})p(\lambda_t|\lambda_{0:t-1}) \\
\approx \;& p(y_t|y_{1:t-1}, \lambda_{0:t}, x_{0:t-1})p(x_{0:t-1}|y_{1:t-1}, \lambda_{0:t-1})p(\lambda_t|\lambda_{0:t-1}) \\
= \;& \mathcal{N}(y_t|\mathbf{CAx}_{t-1}, \mathbf{ee}^T)\mathcal{N}(\mathbf{x}_{0:t-1}|\mu_k, \mathbf{\Sigma}_k)S_{\alpha/2}(1,1,0) \\
= \;& \mathcal{N}\left(\begin{bmatrix} y_t \\ \vec{x}_{t-1}\end{bmatrix} \mid \begin{bmatrix} C\vec{A}x_{t-1}\end{bmatrix}, \begin{bmatrix} \vec{e}\vec{e}^T + (C\vec{A})\vec{\Sigma}_k(C\vec{A})^T & C\vec{A}\Sigma_k \\ \vec{\Sigma}_k(C\vec{A})^T & \vec{\Sigma}_k\end{bmatrix}\right) S_{\alpha/2}(1,1,0)
\end{aligned}
$$

We can then obtain the required density $p(y_t, \lambda_t|y_{1:t-1}, \lambda_{0:t-1})$ by marginalising out $x_{0:t-1}$.

Option 2: It is hard to formulate an analytical form for this density directly. To get around this, we can instead form the analytical joint distribution for $p(y_t, \lambda_t|y_{1:t-1}, \lambda_{0:t-1})$:

$$
p(y_t, \lambda_t|y_{1:t-1}, \lambda_{0:t-1}) = \mathcal{N}(y_t|\mathbf{CA}\mu_k, (\mathbf{Cb})(\mathbf{Cb})^T\lambda_t + d^2)S_{\alpha/2}(\lambda_t|1,1,0) \tag{4}
$$

By fixing the value of y_t in Equation 4, we can obtain the desired distribution $p(\lambda_t|y_{1:t}, \lambda_{0:t-1})$. In general, there is no closed form expression for this distribution.

This formulation of the joint distribution also gives us an alternative interpretation for the sampling density $p(\lambda_t|y_{1:t}, \lambda_{0:t-1})$.

We can reinterpret this as finding the posterior density of λ_t We can interpret this density as the posterior density of the variance of a normal distribution (λ_t)

However, there are a variety of methods to draw samples from this distribution.

1. **Rejection Sampling**

   We adapt the methods of Godsill and Kuruoglu, 1999 to draw samples from this distribution.

   The target distribution for the rejection sampling scheme is given by:

   $$
   f(y_t, \lambda_t|y_{1:t-1}, \lambda_{0:t-1}) = \mathcal{N}(y_t|\mathbf{CA}\mu_k, (\mathbf{Cb})(\mathbf{Cb})^T\lambda_t + d^2)S_{\alpha/2}(\lambda_t|1,1,0)
   $$

   Option 1:

   Using a proposal distribution $g(\lambda_t|y_{1:t}, \lambda_{0:t-1}) = S_{\alpha/2}(1,1,0)$, we can bound the probability density from above using:

$$
\begin{aligned}
M &\geq \frac{f(y_t, \lambda_t | y_{1:t-1}, \lambda_{0:t-1})}{g(y_t, \lambda_t | y_{1:t-1}, \lambda_{0:t-1})} \\
&= \frac{1}{\sqrt{2\pi v_t^2}} \exp(-0.5)
\end{aligned}
\tag{5}
$$

Option 2:

Using a proposal distribution $g(\lambda_t | y_{1:t}, \lambda_{0:t-1}) = S_{\alpha/2}(1, 1, 0)$, we can use the likelihood as a valid rejection function as it is bounded from the above:

$$
\begin{aligned}
\mathcal{N}(...) &\geq \frac{f(y_t, \lambda_t | y_{1:t-1}, \lambda_{0:t-1})}{g(y_t, \lambda_t | y_{1:t-1}, \lambda_{0:t-1})} \\
&= \frac{1}{\sqrt{2\pi v_t^2}} \exp(-0.5)
\end{aligned}
\tag{6}
$$

This gives a suitable rejection sampler as:

(a) Draw $\lambda_t \sim S_{\alpha_2}(1, 1, 0)$

(b) Draw $u \sim U\left(0, \frac{1}{\sqrt{2\pi v_t^2}} \exp(-0.5)\right)$

(c) If $u > \mathcal{N}(....)$, reject λ_t and go to Step (a)

$$
p(y_t, \lambda_t | y_{1:t-1}, \lambda_{0:t-1}) = \mathcal{N}(y_t - \mathbf{CA}\mu_k | 0, \lambda_t') S_{\frac{\alpha}{2}}(\lambda_t' | (\mathbf{Cb})(\mathbf{Cb})^T, 1, \mathbf{dd}^T + \mathbf{CA}\mathbf{\Sigma}_k(\mathbf{CA})^T) \tag{7}
$$

In general, we cannot sample from Equation 7 exactly.