Maximum Likelihood Estimate of 3-dimensional SDEs

There is an 3-dimensional SDEs:

$$d\vec{X}(t) = A \cdot Y(t)dt + \vec{b} \cdot d\vec{W}(t) \tag{1}$$

Where $\vec{X}(t) = \begin{bmatrix} X_1(t) \\ X_2(t) \\ X_3(t) \end{bmatrix}$;

$Y(t) = \begin{bmatrix} \vec{Y_1(t)} \\ \vec{Y_2(t)} \\ \vec{Y_3(t)} \end{bmatrix} = \begin{bmatrix} X_1(t) & X_2(t) & X_3(t) & X_2^2(t) & X_3^2(t) & X_1(t)X_2(t) & X_1(t)X_3(t) \\ X_1(t) & X_2(t) & X_3(t) & X_1^2(t) & X_3^2(t) & X_1(t)X_2(t) & X_2(t)X_3(t) \\ X_1(t) & X_2(t) & X_3(t) & X_1^2(t) & X_2^2(t) & X_1(t)X_3(t) & X_2(t)X_3(t) \end{bmatrix}^\intercal$;

$\vec{W}(t) = \begin{bmatrix} W_1(t) \\ W_2(t) \\ W_3(t) \end{bmatrix}$ are independent Wiener process;

$A = [a_{i,j}]_{3\times 7}$ is an 3×7 real matrix;

$\vec{b} = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix}^\intercal$ is an 1×3 real vector;

Consider the 3-dimensional discrete time series $\{X(t)\}_{3\times n} = \begin{bmatrix} \vec{X(1)}, \dots, \vec{X(n)} \end{bmatrix} = \begin{bmatrix} X_1(1), \dots, X_1(n) \\ X_2(1), \dots, X_2(n) \\ X_3(1), \dots, X_3(n) \end{bmatrix}$.

Each $\vec{X}(k)$,$k = 1, \dots, n$ is an 3 dimensional vector. $\{X\}_{3\times n}$ consist of n observation of the sample path $\vec{X}(t)$ at discrete time $t = k * \Delta t$,$k = 1, \dots, n$.

In this case, there are totally 24 unknown parameters, that we want to get the MLE of them.

To be more specific in the analysis and calculation ,we rewrite the SDEs in detail as the following:

$$\begin{aligned} dX_1(t) &= [a_{1,1}X_1(t) + a_{1,2}X_2(t) + a_{1,3}X_3(t)]dt \\ &+ [a_{1,4}X_2^2(t) + a_{1,5}X_3^2(t) + a_{1,6}X_1(t)X_2(t) + a_{1,7}X_1(t)X_3(t)]dt \\ &+ b_1 dW_1(t) \end{aligned} \tag{2}$$

$$\begin{aligned} dX_2(t) &= [a_{2,1}X_1(t) + a_{2,2}X_2(t) + a_{2,3}X_3(t)]dt \\ &+ [a_{2,4}X_1^2(t) + a_{2,5}X_3^2(t) + a_{2,6}X_1(t)X_2(t) + a_{2,7}X_2(t)X_3(t)]dt \\ &+ b_2 dW_2(t) \end{aligned} \tag{3}$$

$$\begin{aligned} dX_3(t) &= [a_{3,1}X_1(t) + a_{3,2}X_2(t) + a_{3,3}X_3(t)]dt \\ &+ [a_{3,4}X_1^2(t) + a_{3,5}X_2^2(t) + a_{3,6}X_1(t)X_3(t) + c_{3,7}X_2(t)X_3(t)]dt \\ &+ b_3 dW_3(t) \end{aligned} \tag{4}$$

The Euler approximation for equation (2) :

$$\begin{aligned} X_1(t+\Delta t) - X_1(t) &= [a_{1,1}X_1(t) + a_{1,2}X_2(t) + a_{1,3}X_3(t)]\Delta t \\ &+ [a_{1,4}X_2^2(t) + a_{1,5}X_3^2(t) + a_{1,6}X_1(t)X_2(t) + a_{1,7}X_1(t)X_3(t)]\Delta t \\ &+ b_1 \Delta W_1(t) \end{aligned}$$

Notice that $s_1\Delta W_1(t)$ follows Gaussian distribution $N(0, b_1^2\Delta t)$.

For simple expression, denoting the mean as

$$\begin{aligned} \mu_1(t+\Delta t) = X_1(t) &+ [a_{1,1}X_1(t) + a_{1,2}X_2(t) + a_{1,3}X_3(t)]\Delta t \\ &+ [a_{1,4}X_2^2(t) + a_{1,5}X_3^2(t) + a_{1,6}X_1(t)X_2(t) + a_{1,7}X_1(t)X_3(t)]\Delta t \end{aligned}$$

Denoting the collection of all unknown parameters in the equation (1) then

$$X_1(t+\Delta t)|X_1(t), X_2(t), X_3(t) \sim\ N(\mu_1(t+\Delta t), b_1^2\Delta t)$$

The transition probability density:

$$p_\theta 1(X_1(t+\Delta t)|X_1(t), X_2(t), X_3(t)) = \frac{1}{\sqrt{2\pi b_1^2\Delta t}} exp\{-\frac{[X_1(t+\Delta t) - \mu_1(t)]^2}{2b_1^2\Delta t}\} \tag{5}$$

Similarly, let

$$\mu_2(t+\Delta t) = X_2(t) + [a_{2,1}X_1(t) + a_{2,2}X_2(t) + a_{2,3}X_3(t)]\Delta t$$
$$+ [a_{2,4}X_1^2(t) + a_{2,5}X_3^2(t) + a_{2,6}X_1(t)X_2(t) + a_{2,7}X_2(t)X_3(t)]\Delta t$$

$$\mu_3(t+\Delta t) = X_3(t) + [a_{3,1}X_1(t) + a_{3,2}X_2(t) + a_{3,3}X_3(t)]\Delta t$$
$$+ [a_{3,4}X_1^2(t) + a_{3,5}X_2^2(t) + a_{3,6}X_1(t)X_3(t) + a_{3,7}X_2(t)X_3(t)]\Delta t$$

We have transition probability for X_2, X_3:

$$p_{\theta 2}(X_2(t+\Delta t)|X_1(t), X_2(t), X_3(t)) = \frac{1}{\sqrt{2\pi b_2^2 \Delta t}} exp\{-\frac{[X_2(t+\Delta t) - \mu_2(t+\Delta t)]^2}{2b_2^2\Delta t}\} \tag{6}$$

$$p_{\theta 3}(X_3(t+\Delta t)|X_1(t), X_2(t), X_3(t)) = \frac{1}{\sqrt{2\pi b_3^2 \Delta t}} exp\{-\frac{[X_3(t+\Delta t) - \mu_3(t+\Delta t)]^2}{2b_3^2\Delta t}\} \tag{7}$$

where $\theta 2$ and $\theta 3$ are the collection of unknown parameters in the equation (3) and (4) , respectively.

The log-likelihood functions:

$$\ln L(\theta 1|\vec{X}(1), \ldots, \vec{X}(n)) = -\frac{1}{2b_1^2\Delta t}\sum_{i=1}^{n-1}[X_1(i+1) - \mu_1(i+1)]^2 \tag{8}$$

$$\ln L(\theta 2|\vec{X}(1), \ldots, \vec{X}(n)) = -\frac{1}{2b_2^2\Delta t}\sum_{i=1}^{n-1}[X_2(i+1) - \mu_2(i+1)]^2 \tag{9}$$

$$\ln L(\theta 3|\vec{X}(1), \ldots, \vec{X}(n)) = -\frac{1}{2b_3^2\Delta t}\sum_{i=1}^{n-1}[X_3(i+1) - \mu_3(i+1)]^2 \tag{10}$$

The Maximum Likelihood Estimate :

$$\hat{\theta 1} = argmax(\ln L(\theta 1|\vec{X}(1), \ldots, \vec{X}(n)))$$

$$\hat{\theta 2} = argmax(\ln L(\theta 2|\vec{X}(1), \ldots, \vec{X}(n)))$$

$$\hat{\theta 3} = argmax(\ln L(\theta 3|\vec{X}(1), \ldots, \vec{X}(n)))$$

To maximizing the log-likelihood function, we have

$$\begin{cases} \frac{\partial \ln L(\theta 1|\vec{X}(1),\ldots,\vec{X}(n))}{\partial a_{1,1}} = \frac{1}{b_1^2}\sum_{i=1}^{n-1}\{[X_1(i+1)-\mu_1(i+1)]*X_1(i)\}=0 \\ \frac{\partial \ln L(\theta 1|\vec{X}(1),\ldots,\vec{X}(n))}{\partial a_{1,2}} = \frac{1}{b_1^2}\sum_{i=1}^{n-1}\{[X_2(i+1)-\mu_1(i+1)]*X_1(i)\}=0 \\ \frac{\partial \ln L(\theta 1|\vec{X}(1),\ldots,\vec{X}(n))}{\partial a_{1,3}} = \frac{1}{b_1^2}\sum_{i=1}^{n-1}\{[X_1(i+1)-\mu_1(i+1)]*X_3(i)\}=0 \\ \frac{\partial \ln L(\theta 1|\vec{X}(1),\ldots,\vec{X}(n))}{\partial a_{1,4}} = \frac{1}{b_1^2}\sum_{i=1}^{n-1}\{[X_1(i+1)-\mu_1(i+1)]*X_2^2(i)\}=0 \\ \frac{\partial \ln L(\theta 1|\vec{X}(1),\ldots,\vec{X}(n))}{\partial a_{1,5}} = \frac{1}{b_1^2}\sum_{i=1}^{n-1}\{[X_1(i+1)-\mu_1(i+1)]*X_3^2(i)\}=0 \\ \frac{\partial \ln L(\theta 1|\vec{X}(1),\ldots,\vec{X}(n))}{\partial a_{1,6}} = \frac{1}{b_1^2}\sum_{i=1}^{n-1}\{[X_1(i+1)-\mu_1(i+1)]*X_1(i)X_2(i)\}=0 \\ \frac{\partial \ln L(\theta 1|\vec{X}(1),\ldots,\vec{X}(n))}{\partial a_{1,7}} = \frac{1}{b_1^2}\sum_{i=1}^{n-1}\{[X_1(i+1)-\mu_1(i+1)]*X_1(i)X_3(i)\}=0 \end{cases}$$

Solving these 7 linear equations, we get the MLE of 7 parameters, such that

$$\hat{\theta 1}_a = [\hat{a_{1,1}}, \hat{a_{1,2}}, \hat{a_{1,3}}, \hat{a_{1,4}}, \hat{a_{1,5}}, \hat{a_{1,6}}, \hat{a_{1,7}}]^\intercal = C_1^{-1}\cdot D_1 \tag{11}$$

$$\hat{b_1} = \pm\sqrt{\sum_{i=1}^{n-1}\frac{[X_1(i+1)-X_1(i)-\hat{\theta 1}_a\cdot Y_1(i)]^2}{(n-1)\Delta t}} \tag{12}$$

Where

$C_1 = \sum_{i=1}^{n-1} Y_1(i)\cdot Y_1(i)^\intercal \Delta t$ is a 7×7 symmetric matrix

$D_1 = \sum_{i=1}^{n-1}[X_1(i+1)-X_1(i)]Y_1(i)$ is a 7×1 vector

$Y_1(i) = [X_1(i), X_2(i), X_3(i), X_2^2(i), X_3^2(i), X_1(i)X_2(i), X_1(i)X_3(i)]^\intercal$

Similarly, we have the MLE for the rest parameters

$$\hat{\theta 2}_a = [\hat{a_{2,1}}, \hat{a_{2,2}}, \hat{a_{2,3}}, \hat{a_{2,4}}, \hat{a_{2,5}}, \hat{a_{2,6}}, \hat{a_{2,7}}]^\intercal = C_2^{-1}\cdot D_2 \tag{13}$$

$$\hat{b_2} = \pm\sqrt{\sum_{i=1}^{n-1}\frac{[X_2(i+1)-X_2(i)-\hat{\theta 2}_a\cdot Y_2(i)]^2}{(n-1)\Delta t}} \tag{14}$$

Where

$C_2 = \sum_{i=1}^{n-1} Y_2(i)\cdot Y_2(i)^\intercal \Delta t$ is a 7×7 symmetric matrix

$D_2 = \sum_{i=1}^{n-1}[X_2(i+1)-X_2(i)]Y_2(i)$ is a 7×1 vector

$Y_2(i) = [X_1(i), X_2(i), X_3(i), X_1^2(i), X_3^2(i), X_1(i)X_2(i), X_2(i)X_3(i)]^\intercal$

And

$$\hat{\theta 3}_a = [\hat{a_{3,1}}, \hat{a_{3,2}}, \hat{a_{3,3}}, \hat{a_{3,4}}, \hat{a_{3,5}}, \hat{a_{3,6}}, \hat{a_{3,7}}]^\intercal = C_3^{-1} \cdot D_3 \tag{15}$$

$$\hat{b_3} = \pm\sqrt{\sum_{i=1}^{n-1} \frac{[X_3(i+1) - X_3(i) - \hat{\theta 3}_a \cdot Y_3(i)]^2}{(n-1)\Delta t}} \tag{16}$$

Where

$C_3 = \sum_{i=1}^{n-1} Y_3(i) \cdot Y_3(i)^\intercal \Delta t$ is a 7×7 symmetric matrix

$D_3 = \sum_{i=1}^{n-1} [X_3(i+1) - X_3(i)] Y_3(i)$ is a 7×1 vector

$Y_3(i) = [X_1(i), X_2(i), X_3(i), X_1^2(i), X_2^2(i), X_1(i)X_3(i), X_2(i)X_3(i)]^\intercal$

In conclusion, the MLE of (1)

$$\boxed{\begin{aligned}
\hat{\theta 1}_a = [\hat{a_{1,1}}, \hat{a_{1,2}}, \hat{a_{1,3}}, \hat{a_{1,4}}, \hat{a_{1,5}}, \hat{a_{1,6}}, \hat{a_{1,7}}]^\intercal &= [\sum_{i=1}^{n-1} Y_1(i) \cdot Y_1(i)^\intercal \Delta t]^{-1} \cdot [\sum_{i=1}^{n-1} [X_1(i+1) - X_1(i)] Y_1(i)] \\
\hat{b_1} &= \pm\sqrt{\sum_{i=1}^{n-1} \frac{[X_1(i+1) - X_1(i) - \hat{\theta 1}_a \cdot Y_1(i)]^2}{(n-1)\Delta t}} \\
\hat{\theta 2}_a = [\hat{a_{2,1}}, \hat{a_{2,2}}, \hat{a_{2,3}}, \hat{a_{2,4}}, \hat{a_{2,5}}, \hat{a_{2,6}}, \hat{a_{2,7}}]^\intercal &= [\sum_{i=1}^{n-1} Y_2(i) \cdot Y_2(i)^\intercal \Delta t]^{-1} \cdot [\sum_{i=1}^{n-1} [X_2(i+1) - X_2(i)] Y_2(i)] \\
\hat{b_2} &= \pm\sqrt{\sum_{i=1}^{n-1} \frac{[X_2(i+1) - X_2(i) - \hat{\theta 2}_a \cdot Y_2(i)]^2}{(n-1)\Delta t}} \\
\hat{\theta 3}_a = [\hat{a_{3,1}}, \hat{a_{3,2}}, \hat{a_{3,3}}, \hat{a_{3,4}}, \hat{a_{3,5}}, \hat{a_{3,6}}, \hat{a_{3,7}}]^\intercal &= [\sum_{i=1}^{n-1} Y_3(i) \cdot Y_3(i)^\intercal \Delta t]^{-1} \cdot [\sum_{i=1}^{n-1} [X_3(i+1) - X_3(i)] Y_3(i)] \\
\hat{b_3} &= \pm\sqrt{\sum_{i=1}^{n-1} \frac{[X_3(i+1) - X_3(i) - \hat{\theta 3}_a \cdot Y_3(i)]^2}{(n-1)\Delta t}}
\end{aligned}}$$