On differential equation driven by Brownian motion

Purba Das and Luca Galimberti

October 10, 2025

Background

1 Literature review

This part will include the basics of Itô-Stratonovich integrals in the same notation of Part 2. Something in the line of https://en.wikipedia.org/wiki/Stratonovich_integral.
This part will also include some details on Girsanov's theorem and Risk-neutral measure.

2 Part 2: Numerical computations

The general theme is– in Q1 the students learns about uniqueness and existence of SDE (might have some overlap with FM04, but not necessary for students to have prior knowledge). This leads to an integral representation, which can be written in either Itô or in Stratonovich sense. This brings to Q2, where they learn how to calculate Statenovich integrals and Itô integrals separately. From this, they can verify the correction term in Itô formula. Then in Q3, under the absence of noise, they learn how to numerically compute the functional form of μ. In Q4 the students will learn how to identify drift and diffusion function from data.

Q1 **Itô-Stratonovich correction:** Something in the line of Ito-Stratonovich change for

$$X_t = X_s + \int_s^t \mu(X_u)du + \int_s^t \sigma(X_u)dB_u^{\mathbb{P}}, \tag{1}$$

For example: Consider the differential equation:

$$dX_t = \mu(t, X_t)dt + \sigma(t, X_t)dB_t, \tag{2}$$

where B is a Brownian motion. Write down a sufficient condition for $\mu, \sigma : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ such that the above SDE has existence and uniqueness for the solution of the form (1).
one such condition can be global Lipschitz + linear growth. They need to define the conditions properly in the project.

Consider now $\mu(t, x) = \sin(tx)$ and $\sigma(t, x) = t+e^{-x^2}$. Write down the corresponding Stratonovich form differential equation.

Ans:

$$dX_t = \sin(tX_t)dt + X_t e^{-X_t^2}(t + e^{-X_t^2})dt + (t + e^{-X_t^2}) \circ dB_t.$$

We can give some more hints from below.
Assume additionally that σ is differentiable. Let X solve

$$X_t = X_0 + \int_0^t \mu(u, X_u)du + \int_0^t \sigma(u, X_u)dB_u.$$

Then X solves the Stratonovich equation

$$X_t = X_0 + \int_0^t \left(\mu(u, X_u) - \frac{1}{2}\sigma(u, X_u)\partial_x\sigma(u, X_u)\right) du + \int_0^t \sigma(u, X_u) \circ dB_u.$$

Q2 **Computing iterated integrals:** Iterated integrals capture how different parts of a path interact over time, and together they form something called 'signature' (these are nothing but a collection of (Stratnovich type) iterated integrals) of that path, which encodes its essential shape and behavior. In finance, signatures can be helpful to describe and analyze how prices evolve, making them useful for modeling, forecasting, and understanding market dynamics. For $k \in \mathbb{N}$, the k-th order signature of a path/process X is defined as follows.

$$S^{(k)}(X)_{s,t} = \int_s^t \int_s^{u_k} \cdots \int_s^{u_2} \circ dX_{u_1} \circ dX_{u_2} \cdots \circ dX_{u_k}.$$

Calculate the $S^{(1)}(X)_{s,t}$, $S^{(2)}(X)_{s,t}$ and $S^{(3)}(X)_{s,t}$ for $X_t = a + bt$.

Ans: Let $X_t = a + bt$, where $a, b \in \mathbb{R}$ and $t \geq 0$. Then $dX_t = b\,dt$.

Since $dX_{u_i} = b\,du_i$, each integral is an ordinary (Riemann) integral, so

$$S^{(k)}(X)_{s,t} = b^k \int_s^t \int_s^{u_k} \cdots \int_s^{u_2} du_1\,du_2 \cdots du_k.$$

The iterated integral of 1 over the simplex $\{s < u_1 < \cdots < u_k < t\}$ equals $\frac{(t-s)^k}{k!}$, hence (the students can simply calculate by hand for the first 3, no need to do this general approach)

$$S^{(k)}(X)_{s,t} = \frac{b^k (t-s)^k}{k!}.$$

In particular,

$$S^{(1)}(X)_{s,t} = b(t-s), \qquad S^{(2)}(X)_{s,t} = \frac{b^2(t-s)^2}{2}, \qquad S^{(3)}(X)_{s,t} = \frac{b^3(t-s)^3}{6}.$$

Now consider $X_t = B_t$, the standard Brownian motion. Calculate $S^{(1)}(X)_{s,t}$ and $S^{(2)}(X)_{s,t}$.

Ans: In this case, $dX_t = dB_t$. So,

$$S^{(1)}(X)_{s,t} = \int_s^t \circ dB_u = \lim_{N\to\infty} \sum_{i=0}^{N-1} 1 \times (B_{t_{i+1}} - B_{t_i}) = B_t - B_s.$$

Second order:

$$\begin{aligned}
S^{(2)}(X)_{s,t} &= \int_s^t \int_s^{u_2} \circ dB_{u_1} \circ dB_{u_2} = \int_s^t (B_{u_2} - B_s) \circ dB_{u_2}. \\
&= \int_s^t B_{u_2} \circ dB_{u_2} - B_s \int_s^t \circ dB_{u_2} \\
&= \left(\lim_{N\to\infty} \sum_{i=0}^{N-1} \frac{B_{t_i} + B_{t_{i+1}}}{2} \times (B_{t_{i+1}} - B_{t_i}) \right) - B_s(B_t - B_s) \\
&= \cdots = \frac{B_t - B_s}{2}.
\end{aligned}$$

Finally, by calculating the Itô iterated integrals $\int_s^t \int_s^{u_k} \cdots \int_s^{u_2} dX_{u_1}\,dX_{u_2} \cdots dX_{u_k}$ in the above setup. What is the Itô Statenovich correction term in this case?

To be added

Q3 **ODE setup** For simplicity take $\sigma(X_t) \equiv 0$ in (2). So the SDE can be rewritten as

$$\dot{X} = \frac{d}{dt} X(t) = \mu(X(t)), \tag{3}$$

where $X(t) \in \mathbb{R}$ is the state of the system at time t, and the function $\mu(x(t))$ defines the equations of motion and constraints of the system. Let us now say we observe the dataset X at N equidistant points in time $(t_1, t_2, \cdots, t_N)$. The goal is to learn μ from the data. The steps are summarized as follows (Feel free to refer to https://en.wikipedia.org/wiki/Sparse_identification_of_non-linear_dynamics):
First, use the finite difference method to estimate the $\dot{X}$. These can be arranged into matrices of the form

$$\mathbf{X} = \begin{bmatrix} X(t_1) \\ X(t_2) \\ \vdots \\ X(t_{N-1}) \end{bmatrix} \qquad \dot{\mathbf{X}} = \begin{bmatrix} \dot{X}(t_1) \\ \dot{X}(t_2) \\ \vdots \\ \dot{X}(t_{N-1}). \end{bmatrix} \tag{4}$$

Next, a library $\mathbf{\Theta}(\mathbf{X})$ of nonlinear candidate functions of the columns of $\mathbf{X}$ is constructed, which may be constant, polynomial, or more exotic functions (like trigonometric and rational terms, and so on):

$$\mathbf{\Theta}(\mathbf{X}) = \begin{bmatrix} | & | & | & | & & | & | & \\ 1 & \mathbf{X} & \mathbf{X^2} & \mathbf{X^3} & \cdots & \sin(\mathbf{X}) & \cos(\mathbf{X}) & \cdots \\ | & | & | & | & & | & | & \end{bmatrix}$$

The number of possible model structures from this library is combinatorially high. $\mu(x(t))$ is then substituted by $\mathbf{\Theta}(\mathbf{X})$ and a vector of coefficients $\mathbf{\Xi} = [\xi_\mathbf{1}\xi_\mathbf{2}\cdots\xi_\mathbf{n}]$ determining the active terms in $\mu(\mathbf{X})$:

$$\dot{\mathbf{X}} = \mathbf{\Theta}(\mathbf{X})\mathbf{\Xi}$$

Because only a few terms are expected to be active at each point in time, an assumption is made that $\mu(\mathbf{x}(t))$ admits a sparse representation in $\mathbf{\Theta}(\mathbf{X})$. This then becomes an optimization problem in finding a sparse $\mathbf{\Xi}$ which optimally embeds $\dot{\mathbf{X}}$. In other words, a parsimonious model is obtained by performing least squares regression on the system (3) as follows.

$$\xi_\mathbf{k} = \arg\min_{\xi'_\mathbf{k}} \left|\left|\dot{\mathbf{X}}_\mathbf{k} - \mathbf{\Theta}(\mathbf{X})\xi'_\mathbf{k}\right|\right|_\mathbf{2} + \lambda \, ||\xi'_\mathbf{k}||_\mathbf{1} \,,$$

where λ is a regularization parameter. Finally, the sparse set of $\xi_\mathbf{k}$ can be used to reconstruct the dynamical system:

$$\dot{x}_k = \mathbf{\Theta}(\mathbf{x})\xi_\mathbf{k}$$

We have flexibility of choosing this Take $\mu(y) = 2y^3$, $t \in [0, 1]$ and $X(0) = \frac{1}{2\sqrt{2}}$. Then $X(t) = \frac{1}{2\sqrt{2-t}}$ solves the ODE. We give the student to the vector (X_{t_i}) for $i = 1, \cdots 10000$ in a Excel file $(t_i = \frac{1}{N})$. We ask the students to use the polynomial library in SINDy package up to order 5, i.e., $\Theta(\mathbf{X}) = \begin{bmatrix}1 & \mathbf{X} & \mathbf{X^2} & \cdots \mathbf{X^5}\end{bmatrix}$. They need to recover the symbolic version of μ using the Python 'pysindy' package (Code available at https://github.com/dynamicslab/pysindy) for two regularization parameters of λ (say 0.001 and 0.5). We ask them to repeat everything for $\mu(y) = \sin y$ (something beyond the basis, to highlight the importance of choosing $\mathbf{\Theta}$ carefully).

Q4 **Learning drift and diffusion** The idea taken from [1]: Consider the Itô stochastic differential equation (2). The function $\mu : \mathbb{R} \to \mathbb{R}$ represents the drift, while $\sigma : \mathbb{R} \to \mathbb{R}$ is the diffusion function, which governs the stochastic force.

We have flexibility for this: We chose $\mu(y) = 0.2y$ and $\sigma(y) = y^2$, something fairly simple. We provide the students with the vector $(X_{t_1}, \cdots, X_{t_N})$ for $N = 10000$ or so. They need to recover symbolic μ and σ using stochastic SINDy. To be more precise, the students needs to do the following step:
Step-1 Use Eq (4.1) and (4.2) of [1] to estimate the drift vector $(\mu(X_{t_1}), \mu(Xt_2), \cdots \mu(X_{t_{N-1}}))$, and the diffusion vector $(\sigma(X_{t_1}), \sigma(X_{t_2}), \cdots \sigma(X_{t_{N-1}}))$.
Step-2: Set-up a SINDy dictionary. Anything reasonable should work (like the one given in

Wikipedia).

Step-3 Run the symbolic regression (4.3) and (4.4) using the 'pysindy' package to get the final symbolic representation of μ and σ.

The following question looks interesting and we feel that our students should become familiar with that. However, currently, we do not know how to fit this specific question in the previous framework. Any ideas?

Q5 **Risk-neutral measure** Consider a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and three independent Brownian motions B_1, B_2, B_3 defined on it. Let $\{\mathcal{F}_t\}_{t\in[0,T]}$ be the filtration generated by the Brownian motions. We consider the following financial market:

$$\begin{cases} Y_0(t) = 1,\ 0 \le t \le T \\ dY_1(t) = \beta_1(t)dt + dB_1(t) + 2dB_2(t) + 3dB_3(t) \\ dY_2(t) = \beta_2(t)dt + dB_1(t) + 2dB_2(t) + 2dB_3(t) \end{cases} \tag{5}$$

where:

- Y_0 represents the price of the risk-free asset,
- $Y_i, i = 1, 2$ represent the log-discounted price of two securities,
- $\beta_i, i = 1, 2$ are bounded adapted processes.

Namely, in a vectorial notation:

$$d\boldsymbol{Y}(t) = \begin{pmatrix} \beta_1(t) \\ \beta_2(t) \end{pmatrix} dt + \boldsymbol{\Theta} \begin{pmatrix} dB_1(t) \\ dB_2(t) \\ dB_3(t) \end{pmatrix} \tag{6}$$

where

$$\boldsymbol{Y}(t) = \begin{pmatrix} Y_1(t) \\ Y_2(t) \end{pmatrix}, \quad \boldsymbol{\Theta} = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 2 & 2 \end{pmatrix}$$

Show that there are infinitely many equivalent martingale measures $\mathbb{Q}$ for $\boldsymbol{Y}$.

`Hint:` First, fix an arbitrary and adapted bounded process u_1. Find then two other bounded adapted stochastic processes u_2 and u_3 such that, by setting $\boldsymbol{u}(t) := (u_1(t), u_2(t), u_3(t))^\top$ it holds

$$\boldsymbol{\Theta} \cdot \boldsymbol{u}(t) = \begin{pmatrix} \beta_1(t) \\ \beta_2(t) \end{pmatrix}.$$

Define

$$M_t^{\boldsymbol{u}} = \exp\left\{ -\int_0^t \boldsymbol{u}(s)^\top \cdot d\boldsymbol{B}(s) - \frac{1}{2}\int_0^t ||\boldsymbol{u}(s)||^2 ds \right\}, \quad 0 \le t \le T$$

where

$$d\boldsymbol{B}(s) = \begin{pmatrix} dB_1(s) \\ dB_2(s) \\ dB_3(s) \end{pmatrix}$$

and conclude.

Ans: By solving the linear system above, we see that the constituent parts in the expression of $M_t^{\boldsymbol{u}}$ are given by

$$\int_0^t \boldsymbol{u}(s)^\top \cdot d\boldsymbol{B}(s) = \int_0^t u_1(s) dB_1(s) + \int_0^t \frac{3\beta_2(s) - 2\beta_1(s) - u_1(s)}{2} dB_2(s) + \int_0^t (\beta_1(s) - \beta_2(s)) dB_3(s)$$

and

$$\int_0^t ||\boldsymbol{u}(s)||^2 ds = \int_0^t \frac{5u_1^2(s) + 13\beta_2^2(s) + 8\beta_1^2(s) - 6\beta_2(s)u_1(s) - 20\beta_1(s)\beta_2(s) + 4u_1(s)\beta_1(s)}{4} ds$$

In virtue of the boundedness assumptions for β_1, β_2 and u_1, it is easy to check that the Novikov condition holds good, and therefore $(M_t)_t$ is a martingale wrt to $\mathbb{P}$. By Girsanov Thm,

$$\hat{\boldsymbol{B}}(t) = \int_0^t \boldsymbol{u}(s)^\top ds + \boldsymbol{B}(t)$$

is a $\mathbb{Q}$-BM, where $d\mathbb{Q} = M_T^{\boldsymbol{u}} d\mathbb{P}$ on $\mathcal{F}_T$. Besides,

$$d\boldsymbol{Y}(t) = \boldsymbol{\Theta} d\hat{\boldsymbol{B}}(t)$$

is a $\mathbb{Q}$-martingale, namely $\mathbb{Q}$ is an equivalent martingale measure for $\boldsymbol{Y}$ for any possible choice of u_1.

References

[1] M. Wanner and I. Mezić, *On higher order drift and diffusion estimates for stochastic sindy*, SIAM Journal on Applied Dynamical Systems, 23 (2024), pp. 1504–1539.