Stochastic Integration and Stochastic Differential Equations

Rohan Hitchcock

13 September 2023

In this talk we will discuss how to interpret an equation of the form

$$\frac{dX}{dt} = a(t, X) + b(t, X) \cdot \text{"noise"} \tag{1}$$

whose solution is supposed to be a stochastic process $X = (X(t))_{t\in[0.\infty)}$. What people (usually) mean by (1) is the integral equation

$$X(t) = X(0) + \int_0^t a(s, X(s))\ ds + \int_0^t b(s, X(s))\ dB(s)$$

where $B(t)$ is Brownian motion, and so we need to make sense of the integral

$$\text{“} \int_0^t b(s, X(s))\ dB(s)\text{”} \ .$$

More generally, given two stochastic processes X and Y we wish to define the integral of X with respect to Y: $\int_0^t X(s)\ dY(s)$.

Before continuing we will recall that there are three concepts of integration, all of which coincide for functions $f : \mathbb{R} \to \mathbb{R}$ but are *a priori* distinct [Tao08]. We have:

(1) *Antidifferentiation*: for f above $\int f$ satisfies $f = \frac{d}{dx}\int f$.

(2) *Measure-theoretic integrals*: for f above $\int_{[a,b]} f(x)\ dx$ is the area under the curve. These types of integrals are used to compute things like volume, the mass of an object from its density function, etc.

(3) *Path integrals*: for f above this is $\int_a^b f(x)dx = -\int_b^a f(x)dx$. These types of integrals are used to compute things like the work done by a field on a particle.

In the study of non-stochastic differential equations usually one considers integrals of the first kind, however for stochastic integral equations the right model is to generalise path-like integrals. Lets consider how the path integral $\int_a^b f(x)dx$ is defined. Let $x_0, x_1, x_2, \dots, x_n \in \mathbb{R}$ be points along a continuous path P from $x_0 = a$ to $x_n = b$. Note that we do *not* require that $x_i < x_{i+1}$, so the path can both backtrack and leave the interval $[a, b]$. The path integral of f along P is defined to be the limit of

$$\sum_{k=0}^{n-1} f(x_k)(x_{k+1} - x_k)$$

as the jump sizes $\Delta x_k = x_{k+1} - x_k$ all approach zero. For paths in $\mathbb{R}$ it turns out that the value of this limit depends only on the endpoints of the path a and b (and f) so we denote it by $\int_a^b f(x)dx$, however when generalised to higher dimensions the value of the integral depends on the path P.

For a path $a = x_0 < x_1 < x_2 < \ldots < x_n = b \in \mathbb{R}$ with fixed step size Δx we can generalise the path integral $\int_a^b f(x)dx$ and consider the *Riemann-Stieltjes integral*

$$\int_a^b f(x)dg(x) \coloneqq \lim_{\Delta x \to 0} \sum_{k=0}^{n-1} f(x_i)(g(x_{i+1}) - g(x_i))$$

which exists given conditions on f and g which constrain how wildly they can fluctuate[1].

1 Examples of stochastic integration

A discrete-time example We first consider an example coming from the study of discrete-time martingales, in [Wil91, Chapter 10.6]. Let $X = (X_n)_{n=0}^{\infty}$ be a discrete time real-valued stochastic process representing the price of some asset changing over time. For simplicity suppose at time $n = 0$ we have $X_0 = 0$, so X_n represents the *change* in the price of the asset from a fixed historical point (it may be negative). If we were to buy one unit of the asset at time $n = 0$ then X_n is our profit (or loss) if we were to sell at time n. Suppose we have a strategy for buying and selling this asset over time. We can represent this strategy by a stochastic process[2] $C = (C_n)_{n=0}^{\infty}$, where C_n is the amount of the asset we own — our *stake* — at time n. The change in our position from time $n-1$ to n is $C_{n-1}(X_n - X_{n-1})$. This means our net position at time Y_n is

$$Y_n = (C \bullet X)_n \coloneqq \sum_{k=0}^{n-1} C_k(X_{k+1} - X_k)$$

This is the discrete analogue of the stochastic Itô integral of C with respect to X. There is an obvious continuous time analogue: suppose the price of the asset $X = (X(t))_{t\in[0,\infty)}$ now varies continuously in time t, and that we can also continuously vary our trading strategy $C = (C(t))_{t\in[0,\infty)}$. Let $Y(t)$ be our net position at time t. We might hope that we could approximate $Y(t)$ by discretising the interval $[0,t]$ and computing "discrete" integrals using this discretisation, and furthermore that finer discretisations would produce increasingly good approximations of $Y(t)$. That is

$$Y(t) = \lim_{n\to\infty} \sum_{k=0}^{n-1} C(t_k)(X(t_{k+1}) - X(t_k)) \eqqcolon \int_0^t C(s)dX(s)$$

where $t_k = \frac{k}{n}t$. This is, by definition, the *Itô integral of C with respect to X*. Before explaining when this limit exists and commenting on its uniqueness we will consider another example.

[1] A sufficient condition is that f is α-Hölder continuous and g is β-Hölder continuous where $\alpha + \beta > 1$.

[2] For this to be a valid strategy we need to impose some conditions on C. In particular C_n cannot depend on any X_m where $m > n$ since this would mean we could see into the future. To formalise this we consider the σ-algebras $\mathcal{F}_n = \sigma(X_0, \ldots, X_n)$, giving an ascending sequence of σ-algebras $\mathcal{F}_0 \subseteq \mathcal{F}_1 \subseteq \cdots$ (a *filtration*). One thinks about $\mathcal{F}_n$ as representing the information revealed at time n. We impose the condition that C_n must be $\mathcal{F}_n$-measurable (C is *predictable* with respect to the filtration), which is equivalent to $C_n = f(X_0, \ldots, X_n)$ for some measurable function f. **While in the discrete time case the condition that C is predictable only serves to restrict ourselves to physically plausible strategies, in the continuous time case it is a necessary condition for the Itô integral to converge.**

Algorithm 1 Metropolis-Hastings sampler

Require: Initial x_0, a distribution $q(x)$ where we can compute $q(x')/q(x)$ easily.

$x \leftarrow x_0$
loop
  Draw x' from $\mathcal{N}(x, \sigma^2)$ ▷ *Generate a proposal for the next state*
  $\alpha_t \leftarrow \min(1, q(x')/q(x))$
  With probability α_t output x' and set $x \leftarrow x'$ ▷ *Accept/reject the proposal*

A sampling example Recall the *Metropolis-Hastings* (MH) algorithm for Markov Chain Monte Carlo (MCMC) sampling (Algorithm 1). Suppose we have a distribution $q(x)$ which we wish to sample from. It is frequently the case that, while it may be hard to compute $q(x)$ exactly, it is fairly easy to compute $q(x)$ up to a fixed multiplicative constant and hence $q(x')/q(x)$ may be easily computed.

Let $B(t)$ be standard Brownian motion and consider discrete time intervals $\tau_1, \tau_2, \ldots$ where $\sigma^2 = \tau_{n+1} - \tau_n$ is fixed. It is a property of Brownian motion that each $B(\tau_{n+1}) - B(\tau_n) \sim \mathcal{N}(0, \sigma^2)$ and is independent from any other increment. Let X_n be the 'position' of the MH-sampler at time τ_n, where we consider the sampler to stay at its current position if a proposal is rejected. Then $X_0 = x_0$ and

$$X_n = \begin{cases} X_{n-1} + N \text{ where } N \sim \mathcal{N}(0, \sigma^2) & \text{w.p. } \alpha_1 \\ X_{n-1} & \text{w.p. } 1 - \alpha_1 \end{cases}$$

for $n \geq 1$. Let

$$A(X_n) = \begin{cases} 1 & \text{w.p. } \alpha_n \\ 0 & \text{w.p. } 1 - \alpha_n \end{cases}$$

We then have that

$$X_n = x_0 + \sum_{k=0} A_k(X_k)(B(\tau_{k+1}) - B(\tau_k))$$

since $B(\tau_{k+1}) - B(\tau_k) \sim \mathcal{N}(0, \sigma^2)$. That is, the path of the sampler is a solution to the discrete stochastic integral equation

$$X = x_0 + A(X) \bullet B \ .$$

2 The Itô and Stratonovich integrals

Let $X = (X(t))_{t \in [0,\infty)}$ and $Y = (Y(t))_{t \in [0,\infty)}$ be stochastic processes.

Definition 1 (Itô integral)**.** We define the *Itô integral* of X with respect to Y as the stochastic process $Z(t)$

$$Z(t) := \int_0^t X(s)\ dY(s) := \lim_{n \to \infty} \sum_{k=0}^{n-1} X(t_k) Y((t_{k+1}) - Y(t_k))$$

provided the limit on the right-hand-side exists as a limit in probability.

Definition 2 (Stratonovich integral)**.** We define the *Stratonovich integral* of X with respect to Y as the stochastic process $\mathring{Z}(t)$

$$\mathring{Z}(t) := \int_0^t X(s) \circ dY(s) := \lim_{n \to \infty} \sum_{k=0}^{n-1} X(\tfrac{t_k + t_{k+1}}{2}) Y((t_{k+1}) - Y(t_k))$$

provided the limit on the right-hand-side exists as a limit in probability.

Sufficient conditions for the existence of the Itô and Stratonovich integrals are difficult to state briefly given we aren't assuming a background in stochastic processes. The Itô integral $\int_0^t X(s)\, dY(s)$ exists if both X and Y are *semimartingales adapted to the same filtration*:

- A *filtration* $\mathcal{F} = (\mathcal{F}_t)_{t\in[0,\infty)}$ is an ascending sequence of σ-algebras. The σ-algebra $\mathcal{F}_t$ represents the information revealed at time t. Typically we might take $\mathcal{F}_t = \sigma(\bigcup_{s\leq t} X(s))$. A stochastic process X is *adapted* to $\mathcal{F}$ if $X(t)$ is $\mathcal{F}_t$-measurable.

- An adapted stochastic process X is a *martingale* if $X(t) = \mathbf{E}(X(t+\Delta t) \mid \mathcal{F}_t)$ for all $t, \Delta t > 0$. The condition of being a martingale means that the average behaviour of the process in the future is predicated by its history.

- An adapted stochastic process X is a *local martingale* if there exists an sequence of random[3] times $(\tau_n)_{n=1}^{\infty}$ which are (almost surely) strictly increasing and diverging to infinity such that, for all τ_n, if we stop[4] X at time $t = \tau_n$ it is a martingale. This has the effect of 'averaging out' large but rare fluctuations away from martingale behaviour.

- An adapted stochastic process is a *semimatingale* if it can be written as the sum of a local martingale and a process of finite variation, and if its paths are right continuous with left limits (*càdlàg*).

The class of semimartingales arose from an attempt to find the most general class of integrators for the Itô integral and can, in some sense, be characterised in those terms. See [Pro86, Definition 3.3, Theorem 10.1]. It is a fairly broad class; in particular all adapted *Lévy processes* [Pro05, Chapter I.4] are semimartingales [Pro05, Chapter II.3].

Conditions for the existence of the Stratonovich integral are difficult to state succinctly, and Definition 2 is not the most general description of the Stratonovich integral [Man13, Section IV.D]. The Stratonovich integral is the correct, coordinate independent generalisation of stochastic integration to manifolds: see [Eme89, Chapter VII]. One can usually translate between Itô and Stratonovich integrals using a correction term [Pro05, p. 82].

For further reading, [Man13] gives a survey on stochastic integration, [Øks13, Chapter 3.1] constructs the Itô integral in the special case where Brownian motion is the integrator, [Pro86] gives an introduction to Itô integrals with technical details in a more general setting, and [Pro05] is a fully detailed reference (see in particular Chapter II Theorem 21 for the existence of Itô integrals).

3 Stochastic differential equations

We now explain the typical interpretation of

$$\frac{dX}{dt} = a(t, X) + b(t, X) \cdot \text{“noise”} \tag{2}$$

as a stochastic integal equation, following the motivation given in [Øks13, Chapter 3.1]. Let $\eta(t)$ be the noise term. One might desire the following properties of $\eta(t)$:

(1) $\eta(t)$ and $\eta(s)$ are independent for $t \neq s$.

[3] The fact that these times are stochastic is critical: it is easy to show that if there exist *deterministic* localising sequence of stopping times τ_n then X is a martingale.

[4] We define the *stopped process* $X^{\tau_n}(t) := X(\min(t, \tau_n))$.

(2) The joint distribution $(\eta(t_1+t), \eta(t_1+t), \dots, \eta(t_n+t))$ does not depend on t, for any sequence $t_1, \dots, t_n$. (*stationary*)

(3) $\mathbf{E}\eta(t) = 0$ for all t.

The existence of such a suitable stochastic process η is dubious: any process satisfying (1) and (2) cannot have continuous paths and if we additionally require $\mathbf{E}(\eta(t)^2) = 1$ then $\eta(t)$ cannot be measurable in t [Øks13, p. 21].

Instead, we can consider $0 = t_0 < t_1 < \cdots < t_n = t$ and re-write a discretised version of (2) as

$$X(t_{k+1}) - X(t_k) = a(t_k, X(t_k))\Delta t_k + b(t, X(t_k))\eta(t_k)\Delta t_k$$

where $\Delta t_k = t_{k+1} - t_k$. We now replace $\eta(t_k)\Delta t_k$ by a stationary stochastic process with independent increments and mean zero. People usually choose Brownian motion $B(t)$ since it is the only such process with continuous paths, although many Lévy processes also satisfy these properties and their paths are right continuous with left limits (càdlàg). We then find

$$X(t) = X(0) + \sum_{k=0}^{n-1} a(t_k, X(t_k))\Delta t_k + \sum_{k=0}^{n-1} b(t_k, X(t_k))(B(t_{k+1} - B(t_k))) \ .$$

Taking $\Delta t_k \to 0$ and assuming appropriate limits exist this yields the stochastic integral equation

$$X(t) = X(0) + \int_0^t a(s, X(s)) \ ds + \int_0^t b(s, X(s)) \ dB(s) \ .$$

References

[Eme89] Michel Emery. *Stochastic Calculus in Manifolds*. Universitext. Springer Berlin Heidelberg, 1989.

[Man13] Jonathan H. Manton. 'A Primer on Stochastic Differential Geometry for Signal Processing'. In: *IEEE Journal of Selected Topics in Signal Processing* 7.4 (Aug. 2013), pp. 681–699. arXiv: `1302.0430 [math]`.

[Øks13] Bernt Øksendal. *Stochastic Differential Equations: An Introduction with Applications*. 6th ed. Universitext. Springer, 2013.

[Pro05] Philip E. Protter. *Stochastic Integration and Differential Equations*. Vol. 21. Stochastic Modelling and Applied Probability. Springer Berlin Heidelberg, 2005.

[Pro86] Philip Protter. 'Stochastic Integration Without Tears (with Apology to P. A. Meyer)'. In: *Stochastics* 16.3-4 (Mar. 1986), pp. 295–325.

[Tao08] Terence Tao. 'Differential Forms and Integration'. In: Timothy Gowers, June Barrow-Green and Imre Leader. *The Princeton Companion to Mathematics*. Princeton University Press, 2008, pp. 175–180.

[Wil91] David Williams. *Probability with Martingales*. Cambridge University Press, 1991.