Dr John Armstrong

Department of Mathematics

Faculty of Natural and Mathematical Sciences

2023.03.14

Dr John Armstrong

Department of Mathematics

KCL Geometry

Itô Stochastic Differentials on Manifolds

Itô Stochastic Differentials on Manifolds

Joint work with:

- Andrei Ionescu (stochastic differentials) and
- Damiano Brigo (jets)

Plan:

- I will describe what a classical stochastic differential equation is using integration
- I will motivate the definiton of stochastic differential
- I will show that once the stochastic differential is defined, the key definitions for stochastic processes on manifolds are extremely simple.

ODEs without differentiation

The ODE

$$\frac{dX}{dt} = f(X_t, t), \quad X_0$$

can be written

$$X_u - X_0 = \int_0^u dX_t = \int_0^u f(X_t, t)\, dt$$

which one might write in shorthand as

$$dX_t = f(X_t, t)\, dt, \quad X_0$$

where dX_t and dt are just formal symbols.

Historically stochastic differential equations (SDEs) have been defined entirely in terms of the Itô integral. Differentials have not been defined.

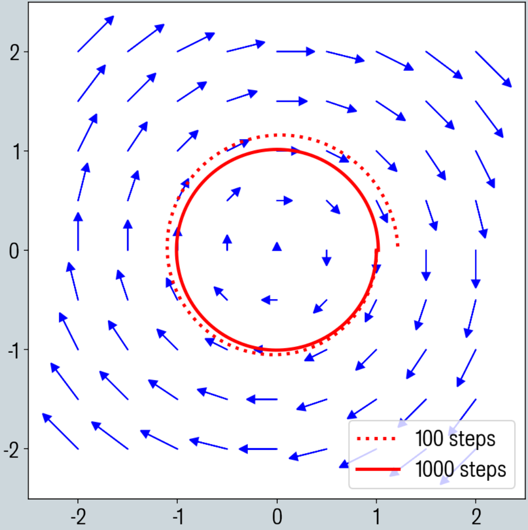
A world without differentiation

- We would not have the notion of two curves being tangent to one another
- We cannot give a meaning to the symbol dX_t at a particular moment in time - ODEs would have to hold in some interval to have any meaning.
- As an integral requires a vector space structure so we are pretty much obliged to define ODEs on manifolds in terms of charts. This is possible, but revolting.

DIfferentiation, visualization and geometry

- Geometrically, ODEs are vector fields. The theory of the tangent space tells us how to think of ODEs in a coordinate free fashion.
- Interpreting a vector as a 1-jet of a curve gives a coordinate free interpretation of Euler-type schemes, just follow the arrows for time δt.

Goal: formally define a stochastic differential and use it to give a coordinate-free treatment of SDEs.



2
1
0
-1
-2
-2
-1
0
1
2
100 steps
1000 steps

History

- Itô was motivated to develop his calculus in order to define Brownian motion on manifolds. He gave a coordinate-based treatment of SDEs on manifolds.
- Itô appears to have been interested in defining a stochastic differential himself
- There is an existing concept called the Nelson derivative which is close to our differential but does not quite achieve what is needed.
- A number of authors have come up with coordinate-free approaches to SDEs including: Stratonovich calculus (Elworthy), Schwarz-Morphisms (Schwarz/Meyer/Emery)

Schwarz: there is nothing “ponctuel” about stochastic differential equations Emery: “existence of the [stochastic differential] is metaphysical and one is free not to believe in it.”

Brownian Motion

In a Stochastic Differential Equation we are interested in the evolution of random processes X_t. The subscript s indicates the value at time t.

The evolution is described in terms of another driving stochastic process, Y_t. For example: Y_t may represent information about the economy and X_t may represent the quantities of different stocks you purchase in response to economic news.

To get started, we will assume that the driving process is a *Brownian motion* W_t.

Definition

W_t is a continuous process in time. The increment $W_{t+\delta t} - W_t$ after time t is independent of any information before time t and is normally distributed with mean 0 and standard deviation $\sqrt{\delta t}$, hence variance δt.

Scaling behaviour

This is the only possible scaling behaviour for independent identical increments with finite variance:

$$\mathrm{Var}(W_{n\delta t} - W_0) = \mathrm{Var}\left(\sum_{i=1}^{n} W_{i\delta t-(i-1)\delta t}\right) = \sum_{i=1}^{n} \mathrm{Var}(W_{i\delta t-(i-1)\delta_t}) = n\,\delta t$$

The Itô Integral

$$\int_0^T a_t \, dW_t \approx \sum_{i=1}^{n} a_{(i-1)\delta t} W_{i\delta t-(i-1)\delta t}, \qquad \delta t = \frac{T}{n}$$

Naively this appears to diverge: we have $T/\delta t$ terms each of size $\sqrt{\delta t}$ so the integral seems to be of the order $\delta t^{-\frac{1}{2}}$. But...

The Itô Integral

$$\int_0^T a_t \, dW_t \approx \sum_{i=1}^{n} a_{(i-1)\delta t} W_{i\delta t-(i-1)\delta t}, \qquad \delta t = \frac{T}{n}$$

Naively this appears to diverge: we have $T/\delta t$ terms each of size $\sqrt{\delta t}$ so the integral seems to be of the order $\delta t^{-\frac{1}{2}}$. But...

$$\begin{aligned} E\left(\int_0^T a_t \, dW_t\right)^2 &\approx \sum_{i=1}^{n}\sum_{j=1}^{n} a_{(i-1)\delta t} a_{(j-1)\delta t} E(W_{i\delta t-(i-1)\delta t} W_{j\delta t-(j-1)\delta t}) \\ &= \sum_{i=1}^{n}\sum_{j=1}^{n} a_{(i-1)\delta t} a_{(j-1)\delta t} E(W_{i\delta t-(i-1)\delta t}) E(W_{j\delta t-(j-1)\delta t}) \end{aligned}$$

using the fact that increments are uncorrelated. So terms where $i \neq j$ vanish.

$$E\left(\int_0^T a_t \, dW_t\right)^2 \approx \sum_{i=1}^{n} a^2_{(i-1)\delta t}(\delta t) = \sum_{i=1}^{n} a^2_{(i-1)\delta t} \frac{T}{n}$$

which no longer looks divergent. The Itô integral is therefore defined using mean-square convergence.

Stochastic Differential Equations

The Itô SDE

$$dX_t = a(X_t, t)dt + b(X_t, t)dW_t, \quad X_0$$

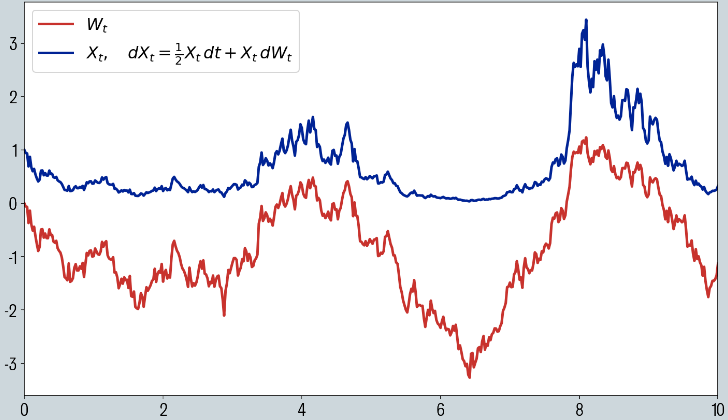
is shorthand for the integral equation

$$X_u - X_0 = \int_0^u a(X_t, t)dt + \int_0^u b(X_t, t)dW_t$$

where the left hand integral is a Riemann integral and the right hand integral is an Itô integral.

An example SDE

$$dX_t = \frac{1}{2} X_t dt + X_t dW_t., \qquad X_0 = 1$$



Idea

The differential is equal to 0 if the process is “small”. For ODEs the correct definition is

$$d_t X = 0$$

$$\text{is defined by } X_{t+\delta t} - X_t = o(\delta t)$$

$$\iff \frac{X_{t+\delta} - X_t}{\delta t} \to 0$$

Idea

The differential is equal to 0 if the process is “small”. For ODEs the correct definition is

$$d_t X = 0$$
$$\text{is defined by } X_{t+\delta t} - X_t = o(\delta t)$$
$$\Longleftrightarrow \frac{X_{t+\delta} - X_t}{\delta t} \to 0$$

We can then define an equivalence relation $\sim_t$ by $X \sim Y$ if $d(X_t - Y_t) = 0$.

The differential d_t of X is the equivalence class of X.

How should we define small for stochastic processes?

Requirements

- $dX_t = 0$ for all $t \in [0, T]$ should imply X_s is a constant random variable.
- The fundamental theorem of calculus should hold

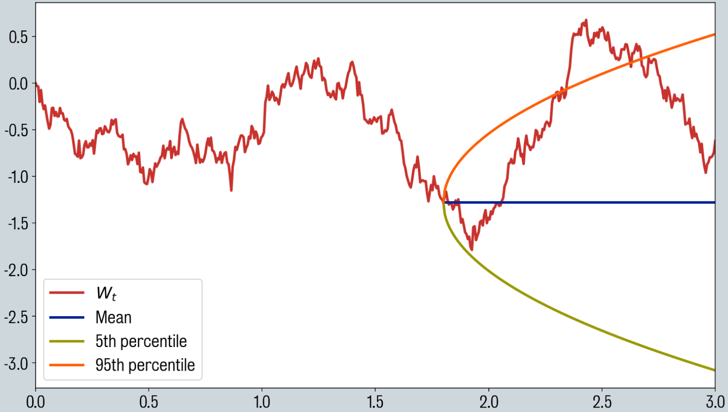
$$d_t \left(\int_0^s dX_u \right) = d_t X$$

  where the integral is an Itô integral.
- The space of differentiable processes X should be as large as possible

Martingales

A martingale is a stochastic process with

- $E(|X_t|)$ finite for all t
- $E(X_{t+\delta t} \,|\, \text{Information up to time t}) = E(X_t)$



Etymology?

Nearly correct definition

e.g. Martingales have uncorrelated increments and so the nice cancellation we saw before occurs for Martingales.

If M_s is a Martingale we might think of defining

$$d_t^E M = 0$$

if

$$E(|M_{t+\delta t} - M_t|^2) = o(\delta t)$$

Modes of convergence

- $X_n \to X$ *in* L^p if $E(|X_n - X|^p) \to 0$ as $n \to \infty$
- $X_n \to X$ *almost surely* if $P(X_n \to X) = 1$
- $X_n \to X$ *in probability* if $P(|X_n - X| > \epsilon) \to 0$ for all $\epsilon > 0$

Also

- $X_n \to X$ in distribution if the distribution function of X_n converges to the distribution of X away from jumps.

Modes of convergence

- $X_n \to X$ *in* L^p if $E(|X_n - X|^p) \to 0$ as $n \to \infty$
- $X_n \to X$ *almost surely* if $P(X_n \to X) = 1$
- $X_n \to X$ *in probability* if $P(|X_n - X| > \epsilon) \to 0$ for all $\epsilon > 0$

Also

- $X_n \to X$ in distribution if the distribution function of X_n converges to the distribution of X away from jumps.

If X_n takes values in a topological space, convergence in expectation depends on metric at infinity, convergence in probability does not.

If we want a diffeomorphism invariant differential we will need convergence in probability. But this is then too weak and $d_t X = 0$ for all t will not imply X is constant.

Quadratic variation

The quadratic covariation of two process is defined by:

$$[X, Y]_t := \underset{|\mathcal{P}|->0}{\text{Plim}} \sum_{k=1}^{m} (X_{t_{k+1}} - X_{t_k})(Y_{t_{k+1}} - Y_{t_k})$$

where $\mathcal{P} = \{x_1, x_2, \ldots x_n\}$ is a partition of $[0, t]$.

The quadratic varation is defined by:

$$[X]_t = [X, X]_t$$

Example:

$$[W]_t = t$$

Birkholder Davis Gundy

Theorem

If X_t is a martingale with $X_0 = 0$ then there are constants c_1, c_2 such that

$$c_1 E([X]_t) \leq E(\sup_{s \in [0,t]} X_s^2) \leq c_2 E([X]_t)$$

Birkholder Davis Gundy

Theorem

If X_t is a martingale with $X_0 = 0$ then there are constants c_1, c_2 such that

$$c_1 E([X]_t) \leq E(\sup_{s\in[0,t]} X_s^2) \leq c_2 E([X]_t)$$

Using the right hand inequality, we can prove uniqueness from a bound on quadratic variation: A martingale with quadratic variation 0 must be constant

Using the left hand inequality, we can get bounds on the quadratic variation from the maximum of the increments

The stochastic differential

Definition

If $X_t = A^1_t - A^2_t + M_t$ where $A^1 + t$ and A^2_t are increasing and M_t is a martingale then we say

$$d_t X = 0$$

iff

$$A^1_{t+\delta t} - A^1_t = o_P(\delta t)$$

$$A^2_{t+\delta t} - A^2_t = o_P(\delta t)$$

$$\sup_{s \in [t, t+\delta t]} (M_s - M_t)^2 = o_P(\delta t)$$

Where $X_h = o_P(f(h))$ if $\frac{X_h}{f(h)}$ converges in probability

Operations on differentials

- Addition

$$d_tX + d_tY := d_t(X + Y)$$

- Multiplication by a random scalar X_t

$$X_t d_t Y = d_t(X_t Y) \quad \text{where } (X_t Y)_s := X_t Y_s$$

- Expectation

$$E_t(d_t X) := d_t(E_t X)$$

Product and quadratic variation

Since

$$X_{t+\delta t}Y_{t+\delta t} - X_tY_t = X_t(Y_{t+\delta t} - Y_t) + Y_t(X_{t+\delta t} - X_t) + (X_{t+\delta t} - X_t)(Y_{t+\delta t} - Y_t)$$

Definition

$$d_t(XY) = X_t d_t Y + Y_t d_t X + d_t X\, d_t Y$$

Lemma

$$d_t[X,Y]_t = d_t X\, d_t Y$$

Stochastic differential on manifolds

Definition

If X, Y are $\mathbb{R}^n$ valued processes $d_t^M X = d_t^M Y$ iff $X_t = Y_t$ almost surely and $d_t X = d_t Y$. If X, Y are manifold valued processes $d_t^M X = d_t^M Y$ iff $d_t^M f(X) = d_t^M f(Y)$ for all smooth $f : M \to \mathbb{R}$.

Stochastic differential on manifolds

Definition

If X, Y are $\mathbb{R}^n$ valued processes $d_t^M X = d_t^M Y$ iff $X_t = Y_t$ almost surely and $d_t X = d_t Y$.
If X, Y are manifold valued processes $d_t^M X = d_t^M Y$ iff $d_t^M f(X) = d_t^M f(Y)$ for all smooth $f : M \to \mathbb{R}$.

Definition

(Pushforward)
If $F : M \to N$ is smooth and X_t is an M valued process we may define

$$F_*(d_t^M X) := d_t^M F(X)$$

it only depends upon the 2-jet of F.

Stochastic differential equations on manifolds

Given a field of 2-jets $\gamma_x : \mathbb{R}^d \to M$ with $\gamma_x(0) = x$ we can associate an SDE by

$$d_t X = (\gamma_x)_* d_t W$$

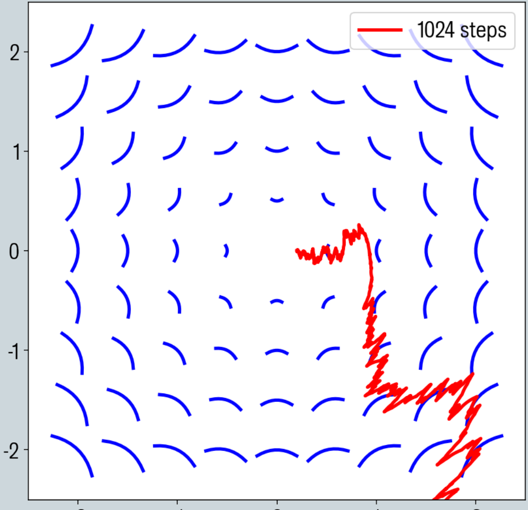
To write this out in classical notation:

- Expand γ_x as a Taylor series to order 2
- Replace the product $d_t W^\alpha \, d_t W^\beta$ with $g^{\alpha\beta} dt$ where $g^{\alpha\beta}$ is the Euclidean metric on $\mathbb{R}^d$

Example

Define a field of 2-jets of curves on $\mathbb{R}^2$ by $\gamma_{(x,y)} : \mathbb{R} \to \mathbb{R}^2$ by

$$\gamma_{(x,y)}(s) = (x, y) + (y, -x)s + 3(x, y)s^2$$



Itô's Lemma

In this notation, the transformation rule for SDEs under a change of coordinates is trivial, but it is opaque when written classically

Lemma

(Itô) If $F : M \to N$ is smooth and X satisfies

$$d_t X = (\gamma_X)_* d_t W$$

then

$$d_t(F \circ X) = (F \circ \gamma_{X_t})_* d_t W$$

Classical Ito's Lemma

Lemma

(Classical Ito's Lemma)
If

$$dX_t^i = a^i(X,t)dt + \sum_{\alpha=1}^{d} b_\alpha^i(X,t)dW_t^\alpha$$

then

$$d(F \circ X)_t^u = \left(\frac{\partial F^u}{\partial x^i} a^i(X,t) + \sum_{\alpha=1}^{d} \frac{1}{2} b_\alpha^i(X,t) b_\alpha^j(X,t) \frac{\partial^2 F^u}{\partial x^i \partial x^j} \right) dt + \sum_{\alpha=1}^{d} \frac{\partial F^u}{\partial x^i} b_\alpha^i dW_t^\alpha$$

Operators associated with stochastic processes

Definition

The *backward diffusion operator* $\mathcal{L}^X$ associated to a stochastic process X is the operator of order 2 acting on 2-jets of functions $f : M \to R$ by:

$$\mathcal{L}^{\mathcal{X}}_{\sqcup} f := \frac{E_t(f_* d_t X)}{dt}$$

The *forward diffusion operator* is its formal adjoint and acts on densities.

Theorem

(Feynman-Kac)
If X is a diffusion and $f : M \to \mathbb{R}$ and $v_t := E_t(f(X_T))$ then

$$\frac{dv}{dt} = \mathcal{L}^{\mathcal{X}}_{\sqcup} v$$

2-jets and connections

Definition

An invariant chart at a point $x \in M$ defined on a neighbourhood $U \ni x$ is a smooth bijection $\phi : T_xM \to U$.

Lemma

Torsion free connections on the tangent bundle of M correspond invariantly to 2-jets of invariant charts.

Since the tangent space has a vector space structure, we can define expectations on the tangent space.

Definition

Given a torsion free connection ∇ define

$$E_t^{\nabla}(d_tX) = (\phi_*)d_t(E\phi_X^{-1}(X))$$

Product structure as a tensor

Let X_s, Y_s be processes on M and suppose that $X_t = Y_t$ almost surely.

Let η^1, η^2 be differential forms defined on M - equivalently fields 1-jets of functions mapping x to 0.

Extend η^1, η^2 arbitrarily to fields of 2-jets and define:

$$(d_t X d_t Y)(\eta^1, \eta^2) := (\eta^1_* d_t X)(\eta^2_* d_t Y)$$

Since the product of differentials only depends on the martingale term, this is well-defined.

Product structure as a tensor

Let X_s, Y_s be processes on M and suppose that $X_t = Y_t$ almost surely.

Let η^1, η^2 be differential forms defined on M - equivalently fields 1-jets of functions mapping x to 0.

Extend η^1, η^2 arbitrarily to fields of 2-jets and define:

$$(d_t X d_t Y)(\eta^1, \eta^2) := (\eta^1_* d_t X)(\eta^2_* d_t Y)$$

Since the product of differentials only depends on the martingale term, this is well-defined.

$$\frac{d_t X d_t Y}{dt}$$

is a 2-tensor. For any differentiable X_s

$$\frac{(d_t X)^2}{dt}$$

is a symmetric 2-tensor.

Brownian motion on a manifold

Definition

Brownian motion, X_s, on the Riemannian manifold (M, g) is a stochastic process satisfying:

$$\frac{(d_t X)^2}{dt} = g_{X_t}$$

and

$$E_t^{\nabla^{\mathrm{LC}}}(d_t X) = 0$$

where ∇^{LC} is the Levi-Civita connection.

Brownian motion on a manifold

Definition

Brownian motion, X_s, on the Riemannian manifold (M, g) is a stochastic process satisfying:

$$\frac{(d_t X)^2}{dt} = g_{X_t}$$

and

$$E_t^{\nabla^{\mathrm{LC}}}(d_t X) = 0$$

where ∇^{LC} is the Levi-Civita connection.

It was already known how to define Brownian motion on a manifold: what is nice is that this is a “stochastic differential equation’’ characterising Brownian motion but not a classical one.

Brownian motion on a manifold

Definition

Brownian motion, X_s, on the Riemannian manifold (M, g) is a stochastic process satisfying:

$$\frac{(d_t X)^2}{dt} = g_{X_t}$$

and

$$E_t^{\nabla^{\mathrm{LC}}}(d_t X) = 0$$

where ∇^{LC} is the Levi-Civita connection.

It was already known how to define Brownian motion on a manifold: what is nice is that this is a "stochastic differential equation'' characterising Brownian motion but not a classical one.

It is locally modelled on Brownian motion, with the 2-jets given by the 2-jets of the exponential map.

Thank You!