Stochastic differential equations as jets and an application to filtering

John Armstrong (King's College London)
Damiano Brigo (Imperial)

1/11/2016

Outline

- Part I: Coordinate free SDEs with jets
- Part II: Projection of SDEs
- Part III: PAn application to filtering

Section 1

Coordinate free SDEs

Coordinate free SDEs

Approaches to SDEs on manifolds

- Itô: coordinate based approach.
- Elworthy: Stratonovich calculus
- Schwarz, Emery: second order tangent vectors, diffusors and Schwartz morphism.
- Y. Belopolskaja and Y. Dalecky, Gliklikh: Itô-bundle
- This talk: 2-jets.

Goals:

- Can we give a formulation of SDEs that makes their geometry more apparent?
- Can we understand SDEs using familiar geometric objects?

Applications

What are the applications?

- We can draw a picture of an SDE.
- We obtain new numerical schemes for solving SDEs on manifolds.
- We can define a new, optimal, notion of projection that allows us to approximate high-dimensional SDEs with low dimensional SDEs.

Tangent vectors (hence ODEs on manifolds)

The coordinate based approach:

Definition

Let M^n be an n-dimensional manifold. A tangent vector at a point $x \in M$ is defined to be an equivalence class of pairs:

$$(v, \phi) = ((v^1, v^2, \dots v^n), \phi)$$

where v is a vector in $\mathbb{R}^n$ and ϕ is a chart.

$$(v, \phi) \sim (w, \Phi)$$

$$\text{if and only if} \quad v^j = \sum_i \frac{\partial \tau^j}{\partial x^i} w^i,$$

where $\tau = \Phi \circ \phi^{-1}$ is the transition function.

Pictorial representation

Vector fields are pairs of a components charts that transform correctly from one coordinate system to another.

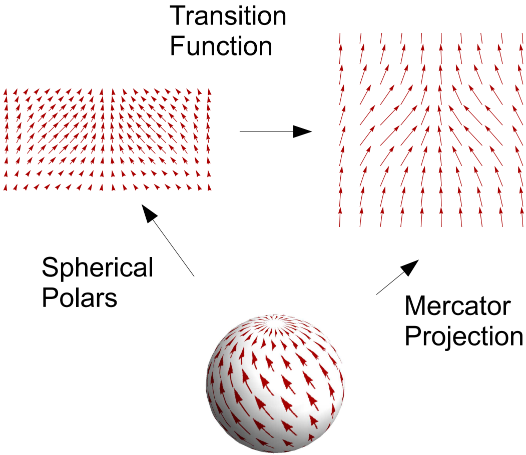


Figure:

SDE on manifold

Itô's approach:

Definition

Let M^n be an n-dimensional manifold. An SDE at a point $x \in M$ is defined to be an equivalence class of quadruples: (W_t, ϕ, a, b)

$$(W_t, \phi, a, b) \sim (V_t, \Phi, A, B) \ \text{ if } \begin{cases} W_t = V_t \\ A^j = a^j \partial_j \tau^i + \frac{1}{2} b^j_\alpha b^k_\beta g^{\alpha\beta} \partial_j \partial_k \tau^i \\ B^j = b^j_\alpha \partial_j \tau^i \end{cases}$$

for the transition function $\tau = \Phi \circ \phi^{-1}$.

Here $g^{\alpha\beta} = [W^\alpha, W^\beta]_t$ denotes the quadratic covariation of W^α and W^β. We are using the Einstein summation convention.

Vector: Operator definition

Derivation:

- A function $D : C^\infty(x) \to \mathbb{R}$ satisfying:
  - $D(a\,f + b\,g) = a\,D(f) + b\,D(g)$ when a, $b \in \mathbb{R}$
  - $D(fg) = f\,, D(g) + g\,D(f)$ when f, $g \in C^\infty(x)$
- where $C^\infty(x)$ is set of germs of smooth functions
- Germ at x: $f \sim g$ if $f(y) = g(y)$ for all y in some neighbourhood $U \ni x$

Example

1. $\frac{\partial}{\partial x}$ is a derivation.
2. Given a vector $V \in \mathbb{R}^n$

$$V(f) := \lim_{h \to 0} \frac{f(x + hV) - f(x)}{h}$$

is a derivation on $\mathbb{R}^n$.

SDE: Operator definitions

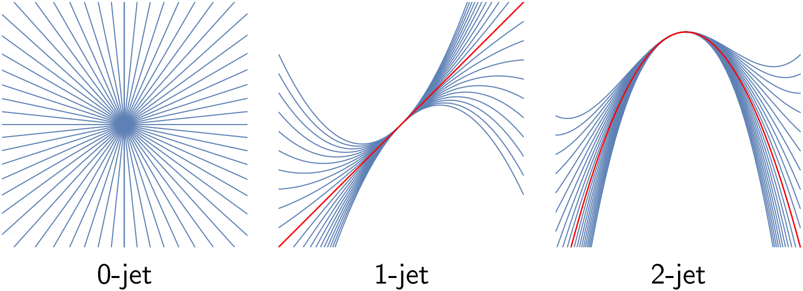
- To an SDE we can associate the forward and backwards diffusion operators acting on, respectively, densities and functions.
- We can read off the coefficients of an SDE from the the coefficients of the operator.

Jets

- A k-jet of a smooth path is defined as an equivalence class of paths with the same Taylor series up to given order.
- Given two smooth functions $f, g : M \to N$ satisfying $f(0) = g(0)$ we say

$$j_k(f) = j_k(g)$$

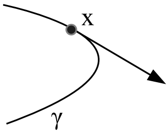
  if f and g have the same Taylor series expansion (in any charts) up to order k.



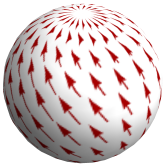
Vectors as jets. Vector fields as infinitesimal diffeomorphisms

Definition

A vector at x is a 1-jet of a path starting at x.



A vector field defines a flow, i.e. a 1-parameter family of diffeomorphisms.



Definitions of tangent vectors and SDEs

Approach	ODE	SDE
Coordinates	Index notation	Itô's definition
Operators	Derivations	Diffusion operators, 2nd order tangent vectors
Jets	1-jets	2-jets
Diffeomorphisms	Vector flows	Stratonovich Calculus

Euler Scheme

- All being well in the limit the Euler scheme

$$\delta X_t = a(X)\,\delta t + b(X)\,\delta W_t$$

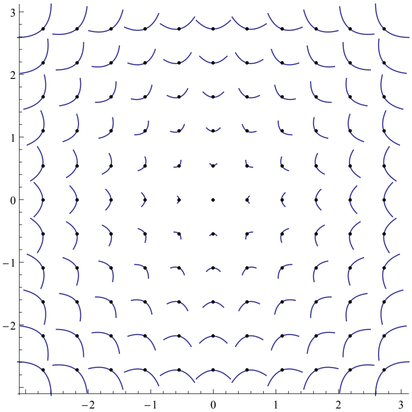
converges to a solution of the SDE

$$\mathrm{d}X_t = a(X)\,\mathrm{d}t + b(X)\,\mathrm{d}W_t$$

- d, δ, $+$ imply vector space structure
- This is highly coordinate dependent

Curved Scheme

Let γ_x be a choice of curve at each point x of M. $\gamma_x(0) = x$.

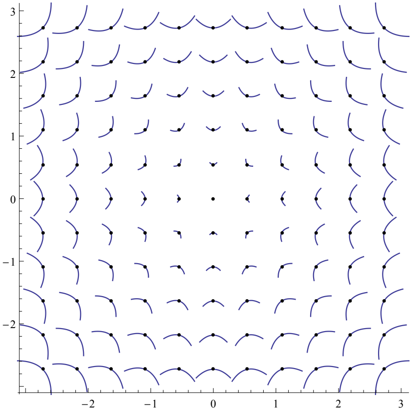


Consider the scheme

$$X_{t+\delta t} = \gamma_{X_t}(\delta W_t) \qquad X_0$$

Concrete example

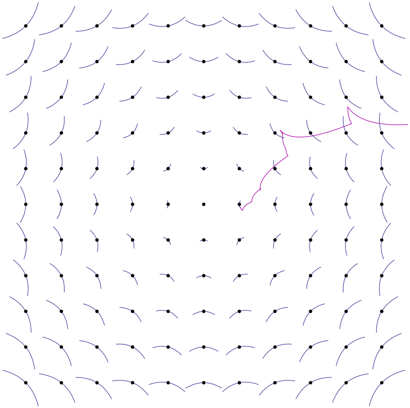
$$\gamma^{E}_{(x_1,x_2)}(s) = (x_1, x_2) + s(-x_2, x_1) + 3s^2(x_1, x_2)$$



- First order term is rotational vector
- Second order term is axial vector

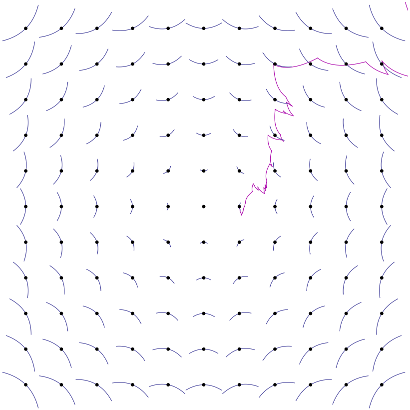
Simulation: Large time step

$$\gamma^E_{(x_1,x_2)}(s) = (x_1, x_2) + s(-x_2, x_1) + 3s^2(x_1, x_2)$$



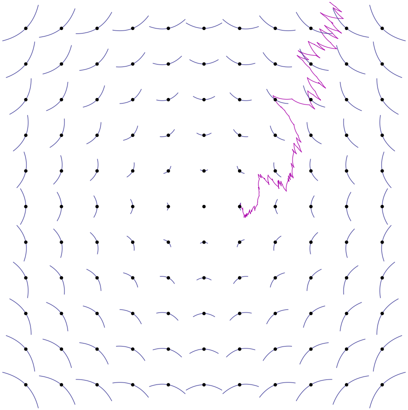
Simulation: Smaller time step

$$\gamma^{E}_{(x_1,x_2)}(s) = (x_1, x_2) + s(-x_2, x_1) + 3s^2(x_1, x_2)$$



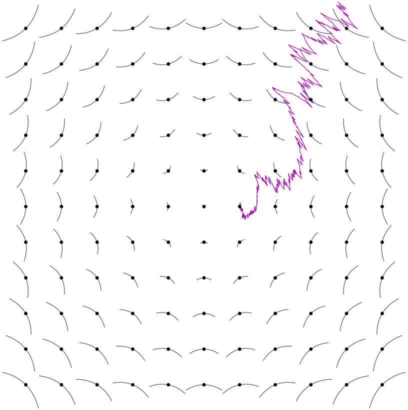
Simulation: Even smaller

$$\gamma^E_{(x_1,x_2)}(s) = (x_1, x_2) + s(-x_2, x_1) + 3s^2(x_1, x_2)$$



Simulation: Convergence

$$\gamma^E_{(x_1,x_2)}(s) = (x_1, x_2) + s(-x_2, x_1) + 3s^2(x_1, x_2)$$



Formal argument

Write:

$$\gamma_x(s) = x + \gamma_x'(0)s + \frac{1}{2}\gamma_x''(0)s^2 + O(s^3)$$

Then:

$$\begin{aligned} X_{t+\delta t} &= \gamma_t(\delta W_t) \\ &= X_t + \gamma_{X_t}'(0)\delta W_t + \frac{1}{2}\gamma'' X_t(0)(\delta W_t)^2 + O\left((\delta W_t)^3\right) \end{aligned}$$

Rearranging:

$$\delta X_t = X_{t+\delta t} - X_t = \gamma_{X_t}'(0)\delta W_t + \frac{1}{2}\gamma'' X_t(0)(\delta W_t)^2 + O\left((\delta W_t)^3\right)$$

Taking the limit:

$$\begin{aligned} \mathrm{d}X_t &= b(X)\mathrm{d}W_t + a(X)(\mathrm{d}W_t)^2 + O\left((\mathrm{d}W_t)^3\right) \\ &= b(X)\mathrm{d}W_t + a(X)\mathrm{d}t \end{aligned}$$

where

$$b(X) = \gamma_X'(0)$$
$$a(X) = \gamma_X''(0)/2$$

Comments

- The curved scheme depends only on the 2-jet of the curve
- SDEs driven by 1-d Brownian motion are determined by 2-jets of curves
- The first derivative determines the volatility term
- The second derivative determines the drift term

ODEs correspond to 1-jets of curves
SDEs correspond to 2-jets of curves

- Rigorous proof of convergence of quadratic scheme can be proved using standard results on Euler scheme

$$\begin{aligned}
\mathrm{d}X_t &= a(X)\mathrm{d}t + b(X)\mathrm{d}W_t \\
&= a(X)\left(\mathrm{d}(W_t^2) - 2W_t\mathrm{d}(W_t)\right) + b(X)\mathrm{d}W_t \\
&\approx a(X)\left(\delta(W_t^2) - 2W_t\delta(W_t)\right) + b(X)\delta W_t \\
&= a(X)\left((\delta W_t)^2\right) + b(X)\delta W_t
\end{aligned}$$

- For general curved schemes some analysis needed.

Itô's lemma

Given a family of curves γ_x we will write:

$$X_t \smile j_2\left(\gamma_x(\mathrm{d}W_t)\right)$$

if X_t is the limit of our scheme.
If

$$X_t \smile j_2\left(\gamma_x(\mathrm{d}W_t)\right)$$

and $f : X \to Y$ then:

$$f(X)_t \smile j_2\left(f \circ \gamma_x(\mathrm{d}W_t)\right)$$

Itô's lemma is simply composition of functions.

Usual formulation

$$X_t \smile j_2\left(\gamma_x(\mathrm{d}W_t)\right)$$

Is equivalent to:

$$\mathrm{d}X_t = a(X)\mathrm{d}t + b(X)\mathrm{d}W_t, \quad a(X) = \frac{1}{2}\gamma_X''(0), \quad b(X) = \gamma_X'(0)$$

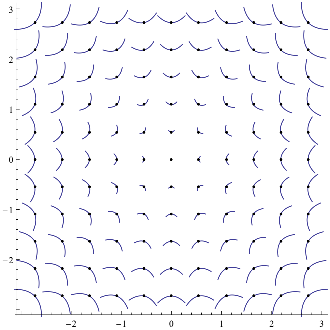
We calculate the first two derivatives of $f \circ \gamma_X$:

$$\begin{aligned}
(f \circ \gamma_X)'(t) &= \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(\gamma_X(t))\frac{\mathrm{d}\gamma_X}{\mathrm{d}t} \\
(f \circ \gamma_X)''(t) &= \sum_{j=1}^{n}\sum_{i=1}^{n} \frac{\partial^2 f}{\partial x_i \partial x_j}(\gamma_X(t))\frac{\mathrm{d}\gamma_X^i}{\mathrm{d}t}\frac{\mathrm{d}\gamma_X^j}{\mathrm{d}t} \\
&\quad + \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(\gamma_X(t))\frac{\mathrm{d}^2\gamma_X}{\mathrm{d}t^2}
\end{aligned}$$

So $f(X_t) \smile j_2\left(f \circ \gamma_x(\mathrm{d}W_t)\right)$ is equivalent to standard Itô's formula

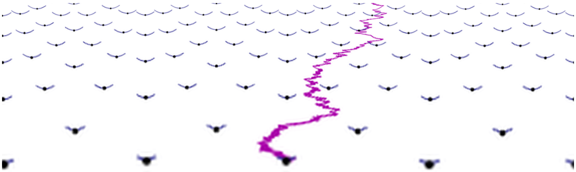
Example

$$\gamma^E_{(x_1,x_2)}(s) = (x_1, x_2) + s(-x_2, x_1) + 3s^2(x_1, x_2)$$

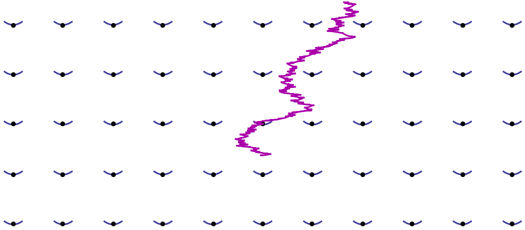


Clearly polar coordinates might be a good idea. So consider the transformation $\phi : \mathbb{R}^2/\{0\} \to [-\pi, \pi] \times \mathbb{R}$ by:

$$\phi(\exp(s)\cos(\theta), \exp(s)\sin(\theta)) = (\theta, s),$$



The process $j_2(\phi \circ \gamma^E)$ plotted using image manipulation software



The process $j_2(\phi \circ \gamma^E)$ plotted by applying Itô's lemma

$$\mathrm{d}(\theta, s) = \left(0, \frac{7}{2}\right) \mathrm{d}t + (1, 0)\, \mathrm{d}W_t.$$

Drawing SDEs

The following diagram commutes:

$$
\begin{array}{ccc}
\text{SDE for X} & \xrightarrow{\text{Itô's lemma}} & \text{SDE for } f(X) \\
\text{Draw}\big\downarrow & & \text{Draw}\big\downarrow \\
\text{Picture of SDE for X in } \mathbb{R}^n & \xrightarrow{f} & f(\text{Picture of SDE for X})
\end{array}
$$

Stratonovich formulation

- Let $\overline{a}$ and b be vector fields on M.
- Define
$$\gamma_x(s) = \Phi^{\overline{a}}_{s^2}\left(\Phi^{b}_{s}(x)\right)$$
where Φ^X_s is the flow associated with a vector field X.
- This defines a field of curves and hence an SDE
- This is a geometric interpretation of the relation between Stratonovich and Itô calculus.
- Application: following these flows should give numerical approximations to SDEs which stay closer to an embedded manifold than the Euler scheme.

Definitions of tangent vectors and SDEs

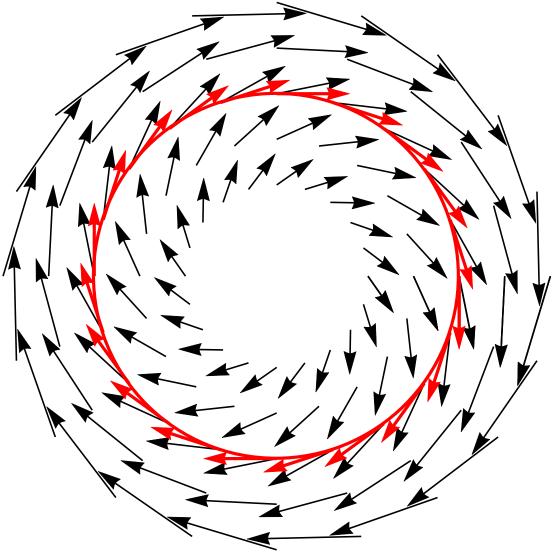
Approach	ODE	SDE
Coordinates	Index notation	Itô's definition
Operators	Derivations	Diffusion operators, 2nd order tangent vectors
Jets	1-jets	2-jets
Diffeomorphisms	Vector flows	Stratonovich Calculus

I've only discussed SDEs driven by 1-d Brownian motion.
Considering 2-jets of maps $\mathbb{R}^k \to M$ gives a similar theory for higher dimensional drivers.

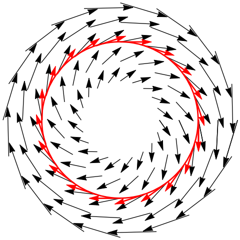
Section 2

Projection

Idea: Projection



Idea: Projection

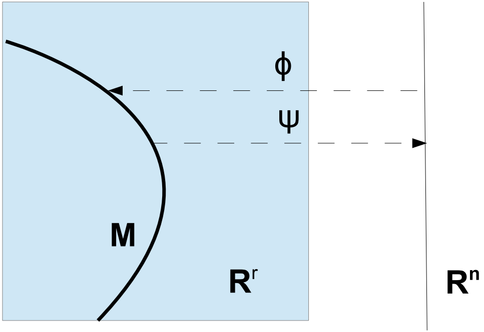


- Projection gives a method of systematically reducing the dimension of an ODE
- Projection onto a linear subspace is the standard numerical method for solving PDEs
- Projecting onto a curved manifold may be more effective if we know the solution is close to this manifold
- e.g. perhaps the known soliton solutions to the KdV equation might give good approximations to solutions to a pertubed KdeV equation?

Projecting SDEs

- Question: How should the notion of projection be extended to stochastic differential equations?
- Answer:
  - There is a Stratonovich Projection which is best understood using Stratonovich calculus.
  - There is an Itô-vector Projection which is best understood using Itô's coordinate formulation.
  - There is an Itô-jet Projection which is best understood by using 2-jets.

Setup

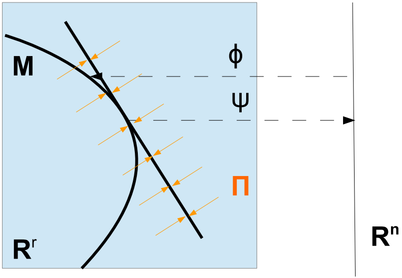


- M is a submanifold of $\mathbb{R}^r$
- $\psi : U \to \mathbb{R}^n$ is a chart for M
- $\phi = \psi^{-1}$
- We have an SDE on $\mathbb{R}^r$

$$\mathrm{d}X_t = a\,\mathrm{d}t + \sum_\alpha b_\alpha\,\mathrm{d}W_t^\alpha, \qquad X_0$$

and want to approximate this using an SDE on $\mathbb{R}^n$.

Definition: Stratonovich projection



1. Write the SDE in Stratonovich form

$$\mathrm{d}X_t = \overline{a(X_t)}\,\mathrm{d}t + \sum_\alpha b_\alpha(X_t) \circ \mathrm{d}W_t^\alpha, \qquad X_0$$

2. Apply the projection operator Π to each coefficient to obtain an SDE on M

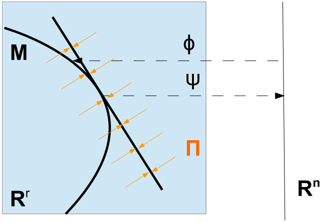
$$\mathrm{d}X_t = \Pi_{X_t}\overline{a(X_t)}\,\mathrm{d}t + \sum_\alpha \Pi_{X_t} b_\alpha(X_t) \circ \mathrm{d}W_t^\alpha, \qquad \psi(X_0)$$

Justifications

What are the justifications for using the Stratonovich projection?

- It is clearly a well defined SDE. (Contrast with projecting Itô coefficients)
- It is clearly generalizes projection of ODEs - i.e. when $b = 0$ we get ODE projection.
- It gives good numerical results when applied to the filtering problem
- It generalizes the Galerkin method which can be interpreted as projection onto a linear subspace.

A justification for ODE projection



- Consider an ODE on $\mathbb{R}^r$

$$\frac{\mathrm{d}X}{\mathrm{d}t} = a(X), \qquad X_0$$

- Look for an ODE on $\mathbb{R}^n$ of the form

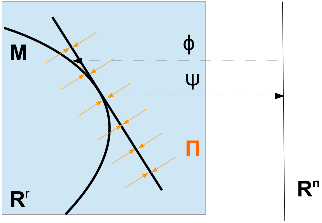
$$\frac{\mathrm{d}x}{\mathrm{d}t} = a(x), \qquad \psi(X_0)$$

  such that

$$|\phi(x_t) - X_t|^2$$

  is as small as possible.

A justification for ODE projection

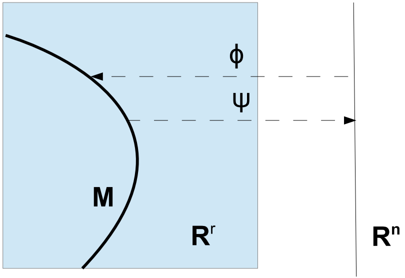


- Compute Taylor expansion to see that leading term is minimized when:

$$a(\psi(x_0)) = \psi_* \Pi_{X_0} A(x_0)$$

- Therefore ODE projection is the unique asymptotically optimal ODE approximating the original ODE at all points on M.
- (Linear projection operator gives solution to a quadratic optimization problem)

Repeat idea for SDEs



Equation in larger space $\mathbb{R}^r$:
$\mathrm{d}X = a(X,t)\,\mathrm{d}t + b(X,t)\,\mathrm{d}W_t$

Equation in chart:
$\mathrm{d}x = A(x,t)\,\mathrm{d}t + B(x,t)\,\mathrm{d}W_t$

We have Itô Taylor series estimates (Kloeden and Platen):

$$E(|X_t - \phi(x_t)|) = |b_0 - \phi_* B_0|\sqrt{t} + O(t)$$

$$|E(X_t - \phi(x_t))| = \left| a_0 - \phi_* A_0 - \frac{1}{2}(\nabla_{B_{\alpha,0}}\phi_*)B_{\beta,0}[W^\alpha, W^\beta] \right| t + O(t^2)$$

Itô-Vector Projection

To minimize first estimate:

$$\phi_* B = \Pi b$$

If we define B like this for whole chart, second estimate is minimized when:

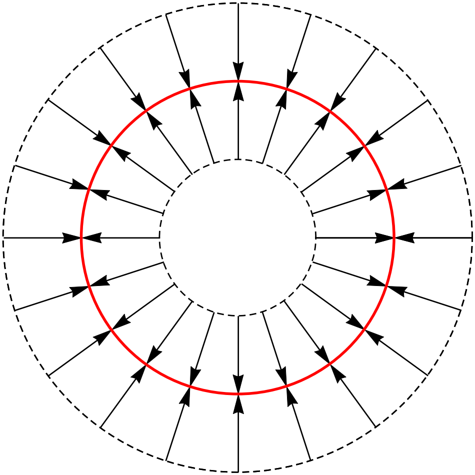
$$\phi_* A = \Pi a - \frac{1}{2}\Pi(\nabla_{B_\alpha}\phi_*)B_\beta[W^\alpha, W^\beta]$$

- Given ϕ, define A and B using these equations
- This defines an SDE on the manifold
- We call this the <u>Itô-vector projection</u>
- It is different from the Stratonovich projection

Alternative

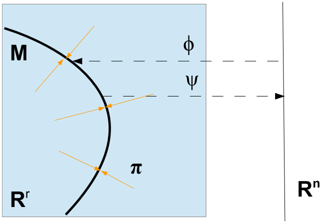
- The use of a weak estimate seems somewhat unsatisfactory.
- An alternative derivation is to compute the strong Itô-Taylor series to one extra order and to try to minimize the coefficient of t.
- This again yields the Itô-vector projection
- Note that this is also somewhat unsatisfactory: whey minimize a term of order t if you can't get the term of order $t^{\frac{1}{2}}$ to vanish?

Metric projection map



Let π denote the smooth map defined on a tubular neighbourhood of M that projects $\mathbb{R}^r$ onto M along geodesics.

An alternative justification for ODE projection



- Consider an ODE on $\mathbb{R}^r$

$$\frac{\mathrm{d}X}{\mathrm{d}t} = a(X), \qquad X_0$$

- Look for an ODE on $\mathbb{R}^n$ of the form

$$\frac{\mathrm{d}x}{\mathrm{d}t} = a(x), \qquad \psi(X_0)$$

such that

$$\cancel{|\phi(x_t) - X_t|^2}$$

$$|\phi(x_t) - \pi(X_t)|^2$$

is as small as possible.

Itô-jet projection

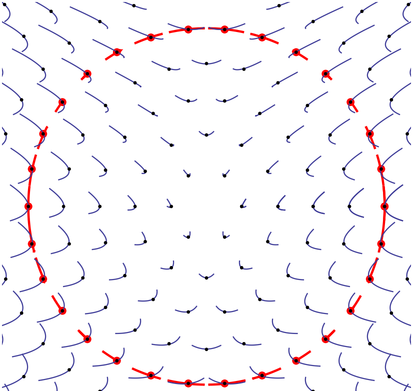
Definition

If original SDE is:

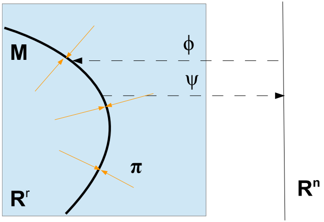
$$X_t \smile j_2\left(\gamma_x(\mathrm{d}W_t)\right)$$

then intrinsic Itô projection is:

$$x_t \smile j_2\left(\pi \circ \gamma_x(\mathrm{d}W_t)\right)$$



Itô-jet projection



Repeating the ideas used to derive the Itô-vector projection:

Theorem

The Itô-jet projection is the best approximation to $\pi(X_t)$ in the sense that it asymptotically minimizes the coefficients in the Taylor series for:

$$E(|\phi(x_t) - \pi(X_t)|)$$

Note that the term of $O(t^{\frac{1}{2}})$ can be made to vanish. You get the same result if distance is measured using geodesic distance on M in the induced metric.

Local coordinate formulation

Calculate Taylor series for π to second order to compute Itô-jet projection in local coordinates:

$$\mathrm{d}x = A\,\mathrm{d}t + B_\alpha\,\mathrm{d}W_t^\alpha, \qquad x_0$$

where:

$$B_\alpha^i = (\pi_*)^i_\beta b^\beta_\alpha$$

and:

$$\begin{aligned}
A^i = {} & (\pi_*)^i_\alpha a^\alpha + \\
& \left(-\frac{1}{2}\frac{\partial^2 \phi^\gamma}{\partial x^\alpha \partial x^\beta}(\pi_*)^a_\gamma (\pi_*)^\alpha_\delta (\pi_*)^\beta_\epsilon \right. \\
& \left. + \frac{\partial^2 \phi^\epsilon}{\partial x^\alpha \partial x^\beta}(\pi_*)^\beta_\delta h^{a\alpha} - \frac{\partial^2 \phi^\gamma}{\partial x^\alpha \partial x^\beta}(\pi_*)^\beta_\epsilon (\pi_*)^\eta_\gamma (\pi_*)^\zeta_\delta h_{\eta\zeta} h^{a\alpha} \right) \\
& \times b^\delta_\kappa b^\epsilon_\iota [W^\kappa, W^\iota]_t.
\end{aligned}$$

h is the induced metric tensor. π_* is the first order projection operator.

Discussion

All three projections are distinct. Which is better?

Lemma

Suppose that S is an SDE for X on $\mathbb{R}^r$ such that $\pi(X)$ solves an SDE S' on M then the Stratonovich and Itô-jet projections are both equal to S'. However, the Itô-vector projection may be different.

Example

The "cross diffusion" SDE S on $\mathbb{R}^2$

$$\begin{aligned} \mathrm{d}X_t &= \sigma Y_t \, \mathrm{d}W_t \\ \mathrm{d}Y_t &= \sigma X_t \, \mathrm{d}W_t \end{aligned}$$

In polar coordinates, solutions satisfy:

$$\mathrm{d}\theta = -\frac{1}{2}\sigma^2 \sin(4\theta) \, \mathrm{d}t + \sigma \cos(2\theta) \, \mathrm{d}W_t$$

Section 3

An application to filtering

The filtering problem

The state of a system evolves according to an SDE:

$$\mathrm{d}X_t = f(X_t, t)\,\mathrm{d}t + \sigma(X_t, t)\,\mathrm{d}W_t$$

with X_0 drawn from some prior distribution. We can only observe

$$\mathrm{d}Y_t = b(X_t, t)\,\mathrm{d}t + \mathrm{d}V_t$$

then, if the coefficients are nice enough, the conditional probability density p satisfies:

$$\mathrm{d}p = \mathcal{L}^* p \; \mathrm{d}t + p[b - E_p(b)]^T[\mathrm{d}Y - E_p(b)\mathrm{d}t].$$

This is the Kushner–Stratonovich equation.
If f and b are linear in X and σ is a deterministic function of time then this is called a linear filter. We can find exact solution given by Gaussians, the so-called Kalman filter.

Numerical example

- The linear filtering problem has solutions given by Gaussian distributions
- Maybe approximately linear filtering problems can be well approximated by Gaussian distributions?
- Heuristic algorithms:
  - Extended Kalman Filter
  - Itô Assumed Density Filter
  - Stratonovich Assumed Density Filter
  - Stratonovich Projection Filter
- Algorithms based on optimization arguments:
  - Itô-vector Projection Filter
  - Itô-jet Projection Filter

L^2 projection

- Suppose that the density can be shown to lie in $L^2(\mathbb{R})$
- Consider the 2-d submanifold of $L^2(\mathbb{R})$ given by the family of Gaussian distributions.
- This is a curved family of distributions. The induced metric is the hyperbolic metric.
- Idea: project the infinite-dimensional Kushner-Stratonovich equation onto the family of Gaussian distributions.
- Generalizations: higher dimensional Gaussians, project onto higher dimensional submanifolds of L^2 to obtain more accurate approximations, e.g. mixture families or exponential families.

Hellinger projection

- Define the Hellinger distance between two probability measures P and Q by:

$$H(P,Q)^2 = \frac{1}{2}\int \left(\sqrt{\frac{\mathrm{d}P}{\mathrm{d}\lambda}} - \sqrt{\frac{\mathrm{d}Q}{\mathrm{d}\lambda}}\right)^2 \mathrm{d}\lambda$$

  where λ is a measure s.t. both P and Q are absolutely continuous w.r.t. λ.
- If P and Q have densities p and q then:

$$H(P,Q)^2 = \frac{1}{2}\int (\sqrt{p(x)} - \sqrt{q(x)})^2 \,\mathrm{d}x = |\sqrt{p} - \sqrt{q}|_2^2$$

- We can compute projection w.r.t. Hellinger metric

Example: a cubic sensor

State equation:

$$\mathrm{d}X_t = \mathrm{d}W_t.$$

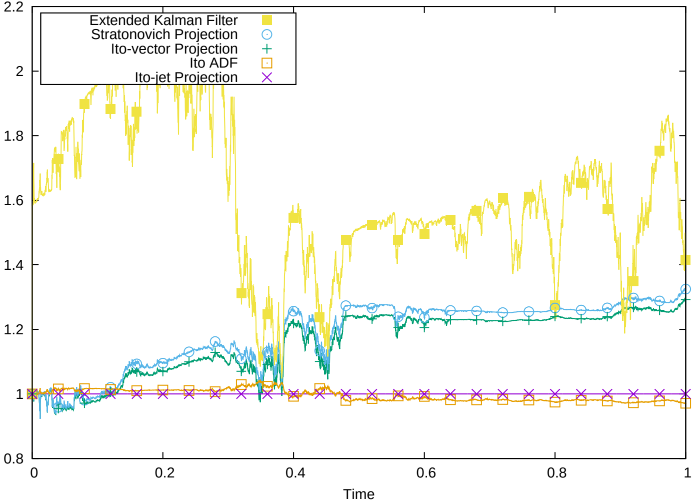
Measurement equation:

$$\mathrm{d}Y_t = (X_t + \epsilon X_t^3)\,\mathrm{d}t + \mathrm{d}V_t$$

ϵ is small ($\epsilon = 0.05$)

Relative performance (Hellinger Residuals)

All projections performed w.r.t. the Hellinger metric.



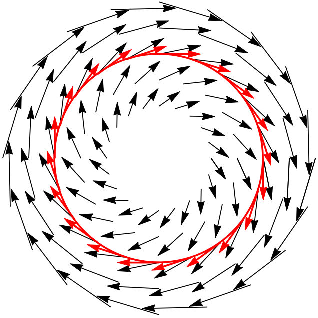
Summary - projection methods

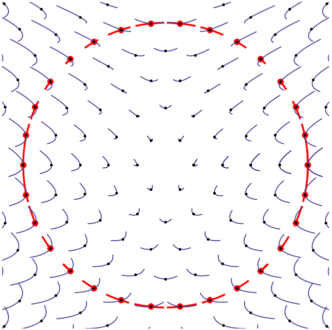
	Ito-vector	Ito-jet	Stratonovich
Optimal?	Yes	Yes	
SDE fibres over π	Surprising	Expected	Expected
Aesthetics		Elegant	
Practice	Best short term	Best medium term	Acceptable

- Note that our notions of optimal are based on expectation of squared residuals

Summary - 2 jets

- 2-jets allow you to draw pictures of SDEs
- They provide an intuitive and elegant reformulation of Itô's lemma
- They provide an alternative route to coordinate free stochastic differential geometry to operator approaches and have found concrete applications.





Higher dimensional jets

$$\gamma_0^A(x,y) = x(1,0) + 2y(0,1) + 2x^2(1,0),$$
$$\gamma_0^B(x,y) = x(1,0) + 2y(0,1) + 2y^2(1,0),$$
$$\gamma_0^C(x,y) = x(1,0) + 2y(0,1) + (x^2+y^2)(1,0).$$