UNIVERSITY OF BERGEN

Score matching and sub-Riemannian bridges

Geometric Sciences in Action: from geometric statistics to shape analysis, Luminy

Erlend Grong, erlend.grong@uib.no

Monday, May 27, 2024

Joint work with Karen Habermann (Warwick) and Stefan Sommer (Copenhagen)

Motivation

Motivation

Assume that we have a given Stratonovich SDE

$$dX_t = \sum_{j=1}^{k} \sigma_j(X_t) \circ dW_t^j + \sigma_0(X_t)dt, \qquad X_0 = x_0.$$

Many techniques for simulating paths (Euler-Maryama etc). But what about bridges?

Brownian Bridge processes in $\mathbb{R}^d$

In $\mathbb{R}^d$, we can force a Brownian motion W_t a process to hit the point x_T at time T by adding a non-homogeneous drift

$$dY_t = dW_t + \frac{1}{T-t}(Y_t - x_T)dt.$$

Diffusion with an elliptic generator

$$dX_t = \sum_{j=1}^k \sigma_j(X_t) \circ dW_t^j + \sigma_0(X_t)dt, \qquad X_0 = x_0.$$

Infinitesimal generator $\frac{1}{2}(\sum_{i=1}^k \sigma_j^2 + 2\sigma_0)$, which is assumed to be elliptic. Let $p_t(x, y)$ be its transition densities with respect to some smooth volume measure,
$P(X_t \in U | X_s = x) = \int_M p_{t-s}(x, y) d\mu(x)$.

We need the score $S_t(x, y) = \nabla^y \log p_t(x, y)$.

Uses that the density is positive and smooth. For short time, we have $S_t(x, y) \approx -\frac{1}{t}\Sigma(x)^{-1}(y - x)$, $\Sigma = \sigma\sigma^\top$.

Diffusion with an elliptic generator

$$dX_t = \sum_{j=1}^{k} \sigma_j(X_t) \circ dW_t^j + \sigma_0(X_t)dt, \qquad X_0 = x_0.$$

Let $p_t(x, y)$ be its transition densities with respect to some smooth volume measure.

We need the score $S_t(x_0, y) = \nabla^y \log p_t(x_0, y)$.

If $p_t(x, y)$ is symmetric, then we can find $Y_t = X_t | (X_T = x_T)$ as solution of

$$dY_t = \sum_{j=1}^{k} \sigma_j(Y_t) \circ dW_t^j + \sigma_0(Y_t)dt + S_{T-t}(x_T, Y_t).$$

If the generator is not symmetric, we can define an SDE for $\bar{Y}_t = Y_{T-t}$ which also uses the score.

So everything is fine if we know the score. But what if we do not? Intractable in practice.

Solution: Score matching

$$dX_t = \sum_{j=1}^{k} \sigma_j(X_t) \circ dW_t^j + \sigma_0(X_t)dt, \qquad X_0 = x_0$$

1. Simulate several sample path.
2. Use clever tricks to define a loss function associated to these samples for which (an approximation) of the score is the minimum, but which does not use the score itself.
3. Get a neural network to learn the score.

Assumed elliptic generators.

Why we want to go beyond elliptic generators

$$dX_t = \sum_{j=1}^{k} \sigma_j(X_t) \circ dW_t^j + \sigma_0(X_t)dt, \qquad X_0(x) = x.$$

What if we introduce multiple conditions

$$X_T(x_{0,i}) = x_{T,i}, \qquad i = 1, \ldots, n$$

Can be considered as a bridge process in the landmark space $\mathbb{M} = \{(x_1, \ldots, x_n) \, : \, x_i \neq x_j\}$ of the SDE

$$d\mathsf{X}_t = \sum_{j=1}^{k} \boldsymbol{\sigma}_j(\mathsf{X}_t) \circ dW_t^j + \boldsymbol{\sigma}_0(\mathsf{X}_t)dt, \quad \boldsymbol{\sigma}_j(x_1, \ldots, x_n) = (\sigma_j(x_1), \ldots, \sigma_j(x_n)).$$

Will loose ellipticity when size of landmark space exceeds k.

In this talk, I will present score matching for bridge processes, where we do not necessarily have an elliptic generator

G, Habermann, Sommer
Score matching for sub-Riemannian bridge sampling.
arXiv:2404.15258

More on score matching and geometry:

- De Bortoli et al., **Riemannian Score-Based Generative Modelling.** Advances in Neural Information Processing Systems, 2022.
- Heng et al., **Simulating diffusion bridges with score matching.** arXiv:2111.07243.

In this talk, I will present score matching for bridge processes, where we do not necessarily have an elliptic generator

G, Habermann, Sommer
Score matching for sub-Riemannian bridge sampling.
arXiv:2404.15258

More on landmarks and geometry:

- Sommer, Arnaudon, Kuhnel, and Joshi. **Bridge Simulation and Metric Estimation on Landmark Manifolds.** Graphs in Biomedical Image Analysis, Computational Anatomy and Imaging Genetics, 2017.
- G., Vega-Molino, **Controllability of shapes through Landmark Manifolds.** arXiv:2403.08090.
- G., Sommer, **Most probable paths for developed processes,** arXiv:2211.15168.

Stochastic processes and sub-Riemannian geometry

SDEs and geometry

$$dX_t = \sum_{j=1}^{k} \sigma_j(X_t) \circ dW^j + \sigma_0(X_t)dt, \qquad X_0 = x_0.$$

- If $g^* = \sum_{j=1}^{k} \sigma_j(x) \otimes \sigma_j(x)$ is an invertible then we can define a corresponding Riemannian metric g.
- Sub-Riemannian: that $\sigma_1, \ldots, \sigma_k$ along with their iterated brackets span the tangent bundle.

For any collection of vector fields, around a generic point, we can reduce our consideration to one of the above cases by restricting to a submanifold.

Similarities and differences

	Rie.	Sub-Rie.
Associated distance d_g	✓	✓
Associated gradient	✓	✓
Smooth probability density	✓	✓
Positive probability density	✓	✓
$-2t \log p_t(x, y) = d_g(x, y)^2 + O(t)$	✓	✓
$g^* = \Sigma = \sigma\sigma^\top$ invertible	✓	
Parallel transport	✓	
d_g Lipchitz equiv. to $\|\cdot\|$	✓	

Similarities and differences

Simplified: if $\sigma_1, \ldots, \sigma_k$ linearly independent, then we can define it as an orthonormal basis of a subbundle E with inner product $\langle \cdot, \cdot \rangle_g$ and define a distance

$$d_g(x,y) = \inf_{\substack{c \text{ from } x \text{ to } y \\ c \text{ tangent to } E}} \int_0^T \langle \dot{c}(t), \dot{c}(t) \rangle_g^{1/2} \, dt$$

- If g is Riemannian, $d_g(x_0, y) \asymp \|x_0 - y\|$
- If g is sub-Riemannian,

$$\begin{aligned} d_g(x_0,y) \asymp |x_0^1 - y^1| + \cdots + |x_0^k - y^k| \\ + |x_0^{k+1} - y^{k+1}|^{1/2} + \cdots + |x_0^{k_2} - y^{k_2}|^{1/2} + \\ + \cdots + |x_0^{k_{s-1}+1} - y^{k_{s-1}+1}|^{1/s} + \cdots + |x_0^{k_s} - y^{k_s}|^{1/s} \end{aligned}$$

More references

On the distance

- Bellaïche. The tangent space in sub-Riemannian geometry. 1996.
- Montgomery. A tour of subriemannian geometries, their geodesics and applications, 2002.

Smooth, positive density with short-time asymptotics.

- Hörmander. Hypoelliptic second order differential equations. 1967.
- Stroock, Varadhan. On the support of diffusion processes with applications to the strong maximum principle. 1972.
- Léandre. Majoration/Minoration en temps petit de la densité d'une diffusion dégénéré. 1987.

Equations for bridge-processes

Diffusions and conditions

$$dX_t = \sum_{j=1}^{k} \sigma_j(X_t) \circ dW^j + \sigma_0(X_t)dt, \qquad X_0 = x_0.$$

The generator of this process is $\frac{1}{2}L$ with

$$L = \sum_{j=1}^{k} \sigma_j^2 + 2\sigma_0.$$

We can define a horizontal gradient of a function $\nabla^E f = \sum_{j=1}^{k}(\sigma_j f)\sigma_j$. If $d\mu$ is a volume density, we can define

$$\Delta f = \Delta_{E,g,d\mu} f = \operatorname{div}_{d\mu} \nabla^E f.$$

Symmetric with respect to $d\mu$. We can write $L = \Delta + 2Z$.

Diffusions and conditions

$$dX_t = \sum_{j=1}^{k} \sigma_j(X_t) \circ dW^j + \sigma_0(X_t)dt, \qquad X_0 = x_0.$$

The generator $\frac{1}{2}L = \frac{1}{2}(\Delta + 2Z)$.

Lemma

Let $p_t(x, y)$ be the heat kernel with respect to $d\mu$. Let Y_t be the result of conditioning $X.$ on $X_T = x_T$. Then Y_t has generator

$$\frac{1}{2}L + \nabla^{x,E} \log p_{T-t}(\cdot, x_T).$$

Idea of proof: Doob h-transform, $h_t = p_{T-t}(x, x_T)$, transition density of Y_t is $p^h_{s,t}(x, y) = p_{t-s}(x, y)\frac{h_t(y)}{h_s(x)}$.

Diffusions and conditions

$$dX_t = \sum_{j=1}^{k} \sigma_j(X_t) \circ dW^j + \sigma_0(X_t)dt, \qquad X_0 = x_0.$$

The generator $\frac{1}{2}L = \frac{1}{2}(\Delta + 2Z)$.

Lemma

Let $p_t(x, y)$ be the heat kernel with respect to $d\mu$. Let Y_t be the result of conditioning $X_.$ on $X_T = x_T$. Then $\overline{Y}_t = Y_{T-t}$ has generator

$$\frac{1}{2}\Delta - Z + \nabla^{y,E} \log p_{T-t}(x_0, \cdot).$$

Note that we now only need the score from the initial point, so we are free to change the final point.

Finding the loss function

We define the score as

$$S_t(x, y) = \nabla^{y,E} \log p_t(x, y).$$

We define the loss functions

$$\mathscr{E}(\theta) = \int_0^T \mathbb{E}^{x_0} \left(\|S_t^\theta(X_t) - S_t(x_0, X_t)\|_g^2 \right) \, dt$$

where S^θ is an estimated score depending on parameters θ in the neural network.

How do we minimize this loss without knowing the true score?

Method 1: Divergence method

Lemma

$\mathscr{E}(\theta) = \int_0^T \mathbb{E}^{x_0} \left(\|S_t^\theta(X_t)\|_g^2 + 2 \operatorname{div} S_t^\theta(X_t) \right) \, dt + C$

Idea of proof: Integration by parts, and being a bit careful.

Method 1: Divergence method

Lemma

$\mathscr{E}(\theta) = \int_0^T \mathbb{E}^{x_0} \left(\|S_t^\theta(X_t)\|_g^2 + 2\,\mathrm{div}\, S_t^\theta(X_t) \right)\, dt + C$

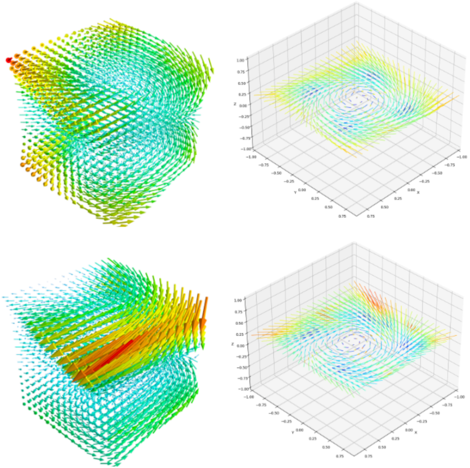
Discretized: Sample paths $X^{(1)}, \dots, X^{(K)}$, $t_i = i \cdot \Delta t$,

$$\mathscr{E}(\theta) - C \approx \frac{\Delta t}{K} \sum_{l=1}^{K} \sum_{i=1}^{n} \left(\left\| S_{t_i}^\theta(X_{t_i}^{(l)}) \right\|_g^2 + 2(\mathrm{div}_{d\mu}\, S_{t_i}^\theta)(X_{t_i}^{(l)}) \right).$$

Results: Model tries to move derivatives away from sample values.

Method 1: Divergence method

Results: Model tries to move derivatives away from sample values. Low row: divergence method. Upper row: Our final method.



Method 2: Euler-Maruyama approximation

How do we generate our samples. Write X_t on Itô form,

$$dX_t = \sum_{j=1}^{k} \sigma_j(X_t)\, dW_t^j + \tau(X_t)\, dt, \qquad X_0 = x_0$$

We approximate with $t_i = i \cdot \Delta t$, $W_{s,t} = W_t - W_s$,

$$\hat{X}_{t_{i+1}} = \hat{X}_{t_i} + \sum_{j=1}^{k} \sigma_j(\hat{X}_{t_i}) W_{t_i,t_{i+1}} + \tau(\hat{X}_{t_i}) \cdot \Delta t.$$

If we define $\Sigma = \sigma\sigma^\top$, then

$$(\hat{X}_{t_{i+1}} | \hat{X}_{t_i} = x) \sim N\Big(x + \Delta t \cdot \tau(x), \Delta t \cdot \Sigma(x)\Big).$$

Method 2: Euler-Maruyama approximation

$$(\hat{X}_{t_{i+1}} | \hat{X}_{t_i} = x) \sim N\Big(x + \Delta t \cdot \tau(x), \Delta t \cdot \Sigma(x)\Big).$$

In the Riemannian case ($\Sigma(x)$ full rank, $d = k$), then $\hat{X}_t$ hat transitional densities $\hat{p}_t(x, y)$, for $t \leq \Delta t$,

$$\hat{p}_t(x, y + x + t \cdot \tau(x)) = \frac{1}{t^{d/2}\sqrt{2\pi \det(\Sigma(x))}} \cdot \exp\left(-\frac{1}{2t}\langle y, \Sigma(x)^{-1} y\rangle\right)$$

Method 2: Euler-Maruyama approximation

$$(\hat{X}_{t_{i+1}} | \hat{X}_{t_i} = x) \sim N\Big(x + \Delta t \cdot \tau(x), \Delta t \cdot \Sigma(x)\Big).$$

In the Riemannian case ($\Sigma(x)$ full rank). then $\hat{X}_t$ transitional densities $\hat{p}_t(x, y)$. Approximation of score

$$S_t(x, y) \approx \nabla^y \hat{p}_t(x, y) = -\frac{1}{t} \sum_{j=1}^{n} \left\langle \sigma_j(y), y - x - t\tau(x) \right\rangle_g \sigma_j(y)$$

$$\mathscr{E}(\theta) = \sum_{i=0}^{n-1} \int_{t_i}^{t_{i+1}} \mathbb{E}^{x_0} \left(\left\langle S_t^\theta(X_t), S_t^\theta(X_t) - 2S_{t-t_i}(X_{t_i}, X_t) \right\rangle \right) dt + C$$

Method 2: Euler-Maruyama approximation

$$(\hat{X}_{t_{i+1}}|\hat{X}_{t_i} = x) \sim N\Big(x + \Delta t \cdot \tau(x), \Delta t \cdot \Sigma(x)\Big).$$

$\hat{p}_t(x, y)$, Approximation of score

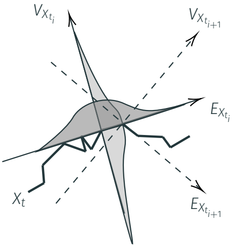
$$S_t(x, y) \approx \nabla^y \hat{p}_t(x, y) = -\frac{1}{t}\sum_{j=1}^{n} \left\langle \sigma_j(y), y - x - t\tau(x)\right\rangle_{g(x)} \sigma_j(y)$$

$$\begin{aligned}\mathscr{E}(\theta) &= \sum_{i=0}^{n-1} \int_{t_i}^{t_{i+1}} \mathbb{E}^{x_0}\left(\Big\langle S_t^\theta(X_t), S_t^\theta(X_t) - 2S_{t-t_i}(X_{t_i}, X_t)\Big\rangle\right) dt + C \\ &\approx \frac{\Delta t}{K} \sum_{l=1}^{K} \sum_{i=0}^{n-1} \sum_{j=1}^{d} S_t^{\theta,j}(X_{t_{t+1}}^{(l)}) \\ &\quad \cdot \left(S^{\theta,j}(X_{t_{i+1}}^{(l)}) + \frac{1}{\Delta t}\sum_{r=1}^{d} \langle \sigma_j(X_{t_{i+1}}^{(l)}), \sigma_r(X_{t_i}^{(l)})\rangle_{g(x_{t_i})} W_{t_i,t_{i+1}}^{(l),r}\right)\end{aligned}$$

Method 2: Euler-Maruyama approximation

$$(\hat{X}_{t_{i+1}} | \hat{X}_{t_i} = x) \sim N\Big(x + \Delta t \cdot \tau(x), \Delta t \cdot \Sigma(x)\Big).$$

If X_t is sub-Riemannian , then $(\hat{X}_{t_{i+1}} | \hat{X}_{t_i} = x)$ is supported on $\Delta t \tau(x) + E_x$ with E_x span of σ_j. Hence $\nabla^{y,E} \hat{p}_t(x, y)$ is not defined in general.



We use $\frac{d}{ds} \sum_{j=1}^{k} \hat{p}_t(x_0, y + s\sigma_j(x))\sigma_j(y)|_{s=0}$, instead but not ideal.

Method 3: Stochastic Taylor expansion

$$dX_t = \sum_{j=1}^{k} \sigma_j(X_t) \circ dW_t^j + \sigma_0(X_t)dt.$$

Define $W_t^0 = t$, $\boldsymbol{\sigma}_\alpha^j = \sigma_{\alpha_1}\sigma_{\alpha_2}\cdots\sigma_{\alpha_l}x^j$,

$$J_{s,t}^\alpha = \int_{s<t_1<\cdots<t_l<t} \circ dW_{t_1}^{\alpha_1} \circ \cdots \circ dW_{t_l}^{\alpha_l}.$$

Define $l(\alpha)$ = length and $n(\alpha)$ = number of zeros. If

$$\hat{X}_{t_i+t} = \sum_{l(\alpha)+n(\alpha)\leq 2\gamma} J_{t_i,t_i+t}^\alpha \cdot \boldsymbol{\sigma}_\alpha(\hat{X}_{t_i}), \qquad \textit{for } t \in [0, \Delta t].$$

Then under appropriate growth conditions on $\boldsymbol{\sigma}_\alpha$,

$$\mathbb{E}^{X_0}\left(\sup_{0\leq t\leq T} \|X_t - \hat{X}_t\|\right) \leq C(\Delta t)^\gamma.$$

Method 3: Stochastic Taylor expansion ($\gamma = 1$)

For $t \in [t_i, t_{i+1}]$, we obtain

$$\begin{aligned}
\hat{X}_t &= \hat{X}_{t_i} + (t - t_i)\,\boldsymbol{\sigma}_0(\hat{X}_{t_i}) + \sum_{j=1}^{k} W^j_{t_i,t}\,\boldsymbol{\sigma}_j(\hat{X}_{t_i}) \\
&\qquad + \sum_{j,l=1}^{k} \int_{t_i}^{t} W^j_{t_i,s} \circ dW^l_s\,\boldsymbol{\sigma}_{(j,l)}(\hat{X}_{t_i}) \\
&= \hat{X}_{t_i} + (t - t_i)\,\boldsymbol{\sigma}_0(\hat{X}_{t_i}) + \sum_{j=1}^{k} W^j_{t_i,t}\,\boldsymbol{\sigma}_j(\hat{X}_{t_i}) \\
&\qquad + \frac{1}{2}\sum_{j,l=1}^{k} W^j_{t_i,t} W^l_{t_i,t}\,\boldsymbol{\sigma}_{(j,l)}(\hat{X}_{t_i}) + \sum_{1\le j<l\le k} A^{j,l}_{t_i,t}(\boldsymbol{\sigma}_{(j,l)} - \boldsymbol{\sigma}_{(l,j)})(\hat{X}_{t_i}),
\end{aligned}$$

where $A^{j,l}_{s,t} = \frac{1}{2}\int_s^t (W^j_{s,r} \circ dW^l_r - W^l_{s,r} \circ dW^j_r)$ is the so-called Lévy area.

Method 3: Stochastic Taylor expansion ($\gamma = 1$)

Theorem (HGS 24)

If $\sigma_i, [\sigma_i, \sigma_j]$ spans the tangent bundle everywhere, then $\hat{X}_t$ has a smooth positive density.

Idea of proof: We make a fiber bundle $F \to \mathbb{R}^d$ and we can lift the process there and show that this is sub-Riemannian. Then we average over each fiber for the result.

Method 3: Stochastic Taylor expansion ($\gamma = 1$)

For approximating the score

$$\begin{aligned}&-2t\log p_t(x_0, y) = d_g(x_0, y)^2\\ &\asymp |x_0^1 - y^1|^2 + \cdots + |x_0^k - y^k|^2 + |x_0^{k+1} - y^{k+1}| + \cdots + |x_0^d - y^d|\end{aligned}$$

whenever $\sigma_i, [\sigma_i, \sigma_j]$ spans the tangent bundle everywhere.
Approach

- Take steps according to approximations using the Taylor expansions and approximations of the Levi area.
- Using the above identity (perhaps scaled with some constants for each coordinate) to approximate the score.

Method 3 Example: Heisenberg group

Equation on $\mathbb{R}^3$ (done for $d = 2k + 1$ in the paper):

$$dX_t = \left(\partial_x - \frac{X_t^2}{2}\partial_z\right) \circ dW_t^1 + \left(\partial_y + \frac{X_t^1}{2}\partial_z\right) \circ dW_t^2$$

Stochastic Taylor expansion X_t equals X_t itself: Solution

$$X_t = X_s \cdot X_{s,t},$$

$$X_{s,t} = \left(W_{s,t}^1, W_{s,t}^2, A_{s,t}\right),$$

$$W_{s,t}^j = W_t^j - W_s^j, \qquad A_{s,t} = \frac{1}{2}\int_0^t (W_{s,r}^1 \circ dW_r^2 - W_{s,r}^2 \circ dW_r^1).$$

$$(x_1, y_1, z_1) \cdot (x_2, y_2, z_2) = (x_1 + x_2, y_1 + y_2, z_1 + z_2 + 1/2 \cdot (x_1 y_2 - x_2 y_1))$$

Method 3 Example: Heisenberg group

Equation on $\mathbb{R}^3$ (done for $d = 2k + 1$ in the paper):

$$dX_t = \left(\partial_x - \frac{X_t^2}{2}\partial_z\right) \circ dW_t^1 + \left(\partial_y + \frac{X_t^1}{2}\partial_z\right) \circ dW_t^2$$

Stochastic Taylor expansion X_t equals X_t itself. Approximation

$$\hat{X}_{t_{i+1}} = \hat{X}_{t_i} \cdot \Delta_i X,$$

$$\Delta_i X = \left(\Delta_i W^1, \Delta_i W^2_{s,t}, \Delta_i A_{s,t}\right),$$

$$\Delta_i A = \frac{1}{2}\left(\Delta_i c_{2,1}\Delta_i W^1 - \Delta_i c_{1,1}\Delta_i W^2\right) + \frac{1}{2}\sum_{m=1}^{K_2-1}\left(\Delta_i c_{1,m}\Delta_i c_{2,m+1} - \Delta_i c_{1,m+1}\Delta_i c_{2,m}\right)$$

$$\Delta_i W^j \sim N(0, \Delta t), \qquad \Delta_i c_{j,m} \sim N\left(0, \frac{\Delta t}{2m+1}\right).$$

Method 3 Example: Heisenberg group

$q_0, q \in \mathbb{R}^3$:

$$\sigma = \partial_x - y/2\partial_z, \qquad \tau = \partial_y + x/2\partial_z.$$

$dX_t = \sigma(X_t) \circ dW_t^1 + \tau(X_t) \circ dW_t^2$. Sub-Riemannian distance

$$d_g(q_0, q) = d_g(0, q_0^{-1} \cdot q) =: f(q_0^{-1} \cdot q).$$

$$f(x, y, 0)^2 = x^2 + y^2, \qquad f(0, 0, z)^2 = 4\pi|z|.$$

$$f(x, y, z)^2 \approx \hat{f}(x, y, z)^2 = x^2 + y^2 + 4\pi|z|.$$

$$\nabla^E \hat{f}^2 = 2(x - \pi y \,\mathrm{sgn}(|z|))\sigma + 2(y + \pi x \,\mathrm{sgn}(|z|))\tau$$

Method 3 Example: Heisenberg group

Score estimation comparison for the Heisenberg group.

$$\text{Minimizing } \mathscr{E}(\theta) = \sum_{i=0}^{n-1} \int_0^{\Delta t} e_{t_i, t_i+t}(\theta)\, dt + C \approx \sum_{i=0}^{n-1} \Delta_i \mathscr{E}(\theta) + C$$

$\hat{X}_{t_{i+1}} = \hat{X}_{t_i} \cdot \Delta_i \hat{X}_i$. Estimated score $S^\theta = S^{\theta,1}\sigma + S^{\theta,2}\tau$.

- Euler-Maruyama: $\Delta_i X = (\Delta_i W^1, \Delta_i W^2, 0)$,

$$\Delta_i \mathscr{E}_i(\theta) = \left\| (S^{\theta,1}, S^{\theta,1}) - \frac{1}{t}(W^1, W^2) \right\|^2.$$

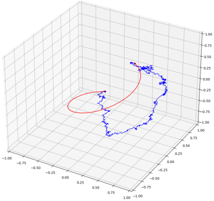
- Taylor: $\Delta_i X = (\Delta_i W^1, \Delta_i W^2, \Delta_i A)$,

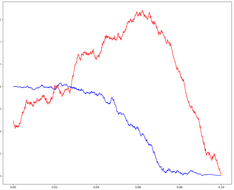
$$\Delta_i \mathscr{E}_i(\theta) = \left\| (S^{\theta,1}, S^{\theta,1}) - \frac{1}{t}(W^1, W^2) - \frac{\operatorname{sgn}(X_t^3)}{t}(-W^2, W^1) \right\|^2.$$

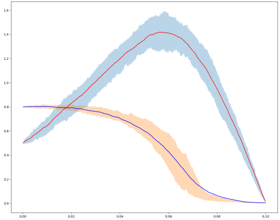
Method 3 Example: Heisenberg group

- Taylor: $\Delta_i X = (\Delta_i W^1, \Delta_i W^2, \Delta_i A)$,

$$\Delta_i \mathscr{E}_i(\theta) = \left\| (S^{\theta,1}, S^{\theta,1}) - \frac{1}{t}(W^1, W^2) - \frac{\operatorname{sgn}(X_t^3)}{t}(-W^2, W^1) \right\|^2 .$$







Thank you very much

Merci beaucoup